

Quantitative results for the Fleming–Viot particle system and quasi-stationary distributions in discrete space

Bertrand Cloez^{a,b,*}, Marie-Noémie Thai^c

^a UMR INRA-SupAgro MISTEA, Montpellier, France

^b EPI INRA-INRIA MODEMIC, Sophia-Antipolis, France

^c Laboratoire d'Analyse et de Mathématiques Appliquées, CNRS UMR8050,
Université Paris-Est Marne-la-Vallée, France

Received 1 August 2014; received in revised form 24 September 2015; accepted 25 September 2015

Available online 9 October 2015

Abstract

We show, for a class of discrete Fleming–Viot (or Moran) type particle systems, that the convergence to the equilibrium is exponential for a suitable Wasserstein coupling distance. The approach provides an explicit quantitative estimate on the rate of convergence. As a consequence, we show that the conditioned process converges exponentially fast to a unique quasi-stationary distribution. Moreover, by estimating the two-particle correlations, we prove that the Fleming–Viot process converges, uniformly in time, to the conditioned process with an explicit rate of convergence. We illustrate our results on the examples of the complete graph and of N particles jumping on two points.

© 2015 Elsevier B.V. All rights reserved.

MSC: 60K35; 60B10; 37A25

Keywords: Fleming–Viot process; quasi-stationary distributions; Coupling; Wasserstein distance; Chaos propagation; Commutation relation

* Corresponding author at: UMR INRA-SupAgro MISTEA, Montpellier, France.

E-mail addresses: bertrand.cloez@supagro.inra.fr (B. Cloez), marie-noemie.thai@univ-paris-est.fr (M.-N. Thai).

1. Introduction

This paper deals with a (time-continuous) Moran type model, referred to as the Fleming–Viot process in the literature [5,18], which approximates Markov semigroup conditioned on non-absorption. Briefly, when considering a time-continuous Markov chain, an interesting question is about the quasi-stationary distribution of the process which is killed at some rate, see for instance [10,25]. Instead of conditioning on non-killing, it is possible to start N copies of the Markov chain and, instead of being killed, one chain jumps randomly on the state of another one. The resulting process is a version of the Moran model that we will call Fleming–Viot. While the convergence of the large-population limit of the Moran model to the quasi-stationary distribution was already shown under some assumptions [13,18,31], the present paper is concerned with deriving bounds for the rate of convergence. Our first main result, namely [Theorem 1.1](#), establishes the exponential ergodicity of the particle system with an explicit rate. This seems to be a novelty. As a consequence, we prove that the correlations between particles vanish uniformly in time, see [Theorems 1.3](#) and [2.6](#). This is also a new result even if [18] gives a similar bound heavily depending on time. As application, we also give new proofs for some more classical but important results as a rate of convergence as N tends to infinity ([Theorem 1.2](#)) which can be compared to the results of [13,20,31], a quantitative convergence of the conditioned semi-group ([Corollary 1.4](#)) comparable to the results of [14,24] and uniform bound (in time) as N tends to infinity (see [Corollary 1.5](#)), which seems to be new in discrete space but already proven for diffusion processes in [28] with an approach based on martingale inequality and spectral theory associated to Schrödinger equation.

Let us now be more precise and introduce our model. Let $(Q_{i,j})_{i,j \in F^*}$ be the transition rate matrix of an irreducible and positive recurrent continuous time Markov process on a discrete and countable state space F^* . Set $F = F^* \cup \{0\}$ where $0 \notin F^*$ and let $p_0 : F^* \mapsto \mathbb{R}_+$ be a non-null function. The generator of the Markov process $(X_t)_{t \geq 0}$, with transition rate Q and death rate p_0 , when applied to bounded functions $f : F \mapsto \mathbb{R}$, reads

$$Gf(i) = p_0(i)(f(0) - f(i)) + \sum_{j \in F^*} Q_{i,j}(f(j) - f(i)),$$

for every $i \in F^*$ and $Gf(0) = 0$. If this process does not start from 0, then it moves according to the transition rate Q until it jumps to 0 with rate p_0 ; the state 0 is absorbing. Consider the process $(X_t)_{t \geq 0}$ generated by G with initial law μ and denote by μT_t its law at time t conditioned on non absorption up to time t . That is defined, for all non-negative functions f on F^* , by

$$\mu T_t f = \frac{\mu P_t f}{\mu P_t \mathbf{1}_{\{0\}^c}} = \frac{\sum_{y \in F^*} P_t f(y) \mu(y)}{\sum_{y \in F^*} P_t \mathbf{1}_{\{0\}^c}(y) \mu(y)},$$

where $(P_t)_{t \geq 0}$ is the semigroup generated by G and we use the convention $f(0) = 0$. For every $x \in F^*$, $k \in F^*$ and non-negative function f on F^* , we also set

$$T_t f(x) = \delta_x T_t f \quad \text{and} \quad \mu T_t(k) = \mu T_t \mathbf{1}_{\{k\}}, \quad \forall t \geq 0.$$

A quasi-stationary distribution (QSD) for G is a probability measure ν_{qs} on F^* satisfying, for every $t \geq 0$, $\nu_{\text{qs}} T_t = \nu_{\text{qs}}$. The QSD are neither well understood, nor easily amenable to simulation. To avoid these difficulties, Burdzy, Holyst, Ingeman, March [5], and Del Moral, Guionnet, Miclo [12,13] introduced, independently from each other, a Fleming–Viot or Moran

type particle system. This model consists of finitely many particles, say N , moving in the finite set F^* . Particles are neither created nor destroyed. It is convenient to think of particles as being indistinguishable, and to consider the occupation number η with, for $k \in F^*$, $\eta(k) = \eta^{(N)}(k)$ representing the number of particles at site k . Each particle follows independent dynamics with the same law as $(X_t)_{t \geq 0}$ except when one of them hits state 0; at this moment, this individual jumps to another particle chosen uniformly at random. The configuration $(\eta_t)_{t \geq 0}$ is a Markov process with state space $E = E^{(N)}$ defined by

$$E = \left\{ \eta : F^* \rightarrow \mathbb{N} \mid \sum_{i \in F^*} \eta(i) = N \right\}.$$

Applying its generator to a bounded function f gives

$$\mathcal{L}f(\eta) = \mathcal{L}^{(N)}f(\eta) = \sum_{i \in F^*} \eta(i) \left[\sum_{j \in F^*} (f(T_{i \rightarrow j}\eta) - f(\eta)) \left(Q_{i,j} + p_0(i) \frac{\eta(j)}{N-1} \right) \right], \quad (1)$$

for every $\eta \in E$, where, if $\eta(i) \neq 0$, the configuration $T_{i \rightarrow j}\eta$ is defined by

$$T_{i \rightarrow j}\eta(i) = \eta(i) - 1, \quad T_{i \rightarrow j}\eta(j) = \eta(j) + 1, \quad \text{and} \quad T_{i \rightarrow j}\eta(k) = \eta(k) \quad k \notin \{i, j\}.$$

For $\eta \in E$, the associated empirical distribution $m(\eta)$ of the particle system is given by

$$m(\eta) = \frac{1}{N} \sum_{k \in F^*} \eta(k) \delta_{\{k\}}.$$

For $\varphi : F^* \rightarrow \mathbb{R}$ and $k \in F^*$, we also set $m(\eta)(\varphi) = \sum_{j \in F^*} \varphi(j) m(\eta)(\{j\})$ and $m(\eta)(k) = m(\eta)(\{k\})$. The aim of this work is to quantify (if they hold) the following limits:

$$\begin{array}{ccc} m(\eta_t^{(N)}) & \xrightarrow[t \rightarrow +\infty]{(a)} & m(\eta_\infty^{(N)}) \\ (b) \downarrow & & \downarrow (c) \\ m(\eta_0)T_t & \xrightarrow[t \rightarrow +\infty]{(d)} & \nu_{\text{qs}} \end{array}$$

where all limits are in distribution and the limits (b), (c) are taken as N tends to infinity. More precisely, [Theorem 1.1](#) gives a bound for the limit (a), [Theorem 1.2](#) for the limit (b), [Corollary 1.5](#) for the limit (c) and finally [Corollary 1.4](#) for the limit (d).

To illustrate our main results, we briefly describe two examples. Those examples are very simple when you are interested by the study of $(T_t)_{t \geq 0}$ (QSD, rate of convergence, etc.) but there are important problems (and even some open questions) on the particle system (invariant distribution, rate of convergence, etc.). The first example concerns a random walk on the complete graph with sites $\{1, \dots, K\}$ and constant killing rate. Namely

$$\forall i, j \in \{1, \dots, K\}, \quad i \neq j, \quad Q_{i,j} = \frac{1}{K}, \quad p_0(i) = p > 0.$$

The quasi-stationary distribution is trivially the uniform distribution. However, the associated particle system does not behave as independent identically distributed copies of uniformly distributed particles and its behavior is less trivial. The second example is the case where F^* contains only two elements. The study of $(T_t)_{t \geq 0}$ is classically reduced to the study of a 2×2 matrix. The study of the particle system, for its part, is reduced to the study of a

birth–death process with quadratic rates. The analysis of these two examples shows the subtlety of Fleming–Viot processes.

Long time behavior. To bound the limit (a), we introduce the parameter λ defined by

$$\lambda = \inf_{i,i' \in F^*} \left(Q_{i,i'} + Q_{i',i} + \sum_{j \neq i,i'} Q_{i,j} \wedge Q_{i',j} \right).$$

This parameter controls the ergodicity of a Markov chain with transition rate Q without killing. Note that λ is slightly larger than the ergodic coefficient α defined in [18] by:

$$\alpha = \sum_{j \in F^*} \inf_{i \neq j} Q_{i,j}.$$

In particular, if there exist $j \in F^*$ and $c > 0$ such that for every $i \neq j$, $Q_{i,j} > c$, then $\lambda \geq c$. Before expressing our results, let us describe the different distances that we use. We endow E with the distance d_1 defined, for all $\eta, \eta' \in E$, by

$$d_1(\eta, \eta') = \frac{1}{2} \sum_{j \in F} |\eta(j) - \eta'(j)|,$$

which is the total variation distance between $m(\eta)$ and $m(\eta')$ up to a factor N : $d_1(\eta, \eta') = Nd_{TV}(m(\eta), m(\eta'))$. Indeed, recall that, for every two probability measures μ and μ' , the total variation distance is given by

$$d_{TV}(\mu, \mu') = \frac{1}{2} \sup_{\|f\|_\infty \leq 1} \left(\int f d\mu - \int f d\mu' \right) = \inf_{\substack{X \sim \mu \\ X' \sim \mu'}} \mathbb{P}(X \neq X'),$$

where the infimum runs over all the couples of random variables with marginal laws μ and μ' . Now, if μ and μ' are two probability measures on E , the d_1 —Wasserstein distance between these two laws is defined by

$$\mathcal{W}_{d_1}(\mu, \mu') = \inf_{\substack{\eta \sim \mu \\ \eta' \sim \mu'}} \mathbb{E}[d_1(\eta, \eta')],$$

where the infimum runs again over all the couples of random variables with marginal laws μ and μ' . The law of a random variable X is denoted by $\mathcal{L}(X)$ and, along the paper, we assume that

$$\sup(p_0) < \infty.$$

Our first main result is:

Theorem 1.1 (Wasserstein Exponential Ergodicity). *If $\rho = \lambda - (\sup(p_0) - \inf(p_0))$, then for any processes $(\eta_t)_{t \geq 0}$ and $(\eta'_t)_{t \geq 0}$ generated by (1), and for any $t \geq 0$, we have*

$$\mathcal{W}_{d_1}(\mathcal{L}(\eta_t), \mathcal{L}(\eta'_t)) \leq e^{-\rho t} \mathcal{W}_{d_1}(\mathcal{L}(\eta_0), \mathcal{L}(\eta'_0)).$$

In particular, if $\rho > 0$, then there exists a unique invariant distribution ν_N satisfying for every $t \geq 0$,

$$\mathcal{W}_{d_1}(\mathcal{L}(\eta_t), \nu_N) \leq e^{-\rho t} \mathcal{W}_{d_1}(\mathcal{L}(\eta_0), \nu_N).$$

To our knowledge, it is the first theorem which establishes an exponential convergence for the Fleming–Viot particle system with an explicit rate. Note anyway that in [32], it is shown that the particle system is exponentially ergodic, when the underlying dynamics follows a certain stochastic differential equation. Its proof is based on Foster–Lyapunov techniques [26,23] and, contrary to us, the dependence on N of the rates and bounds are unknown. So, this gives less information.

When the death rate p_0 is constant, our bound is optimal in terms of contraction. See for instance Section 3, where the example of a random walk on the complete graph is developed. When the death rate is not constant, this bound is not optimal, for instance if the state space is finite, we can have $\rho < 0$ even if the process can converge exponentially fast. Indeed, it can be an irreducible Markov process on a finite state space. Nevertheless, finding a general optimal bound is a difficult problem. Also, note that the previous inequality is a contraction, this gives some information for small times and is more than a convergence result. Finally the previous convergence is stronger than a convergence in total variation distance as can be checked with Corollary 2.3.

Propagation of chaos. In general, two tagged particles in a large population of interacting ones behave in an almost independent way under some assumptions; see [30]. In our case, two particles are almost independent when N is large and this gives the convergence of $(m(\eta_t))_{t \geq 0}$ to $(T_t)_{t \geq 0}$.

To prove this result, we will assume that:

Assumption (Boundedness Assumption).

(A) $\mathbf{Q}_1 = \sup_{i \in F^*} \sum_{j \in F^*, j \neq i} Q_{i,j} < +\infty$ and $\mathbf{p} = \sup_{i \in F^*} p_0(i) < +\infty$.

Under this assumption, the particle system converges to the conditioned semi-group. Moreover, when the state space is finite, this convergence is quantified in terms of total variation distance. To express this convergence, we set

$$\mathbb{E}_\eta[f(X)] = \mathbb{E}[f(X) \mid \eta_0 = \eta],$$

for every bounded function f , every $\eta \in E$ and every random variable X .

Theorem 1.2 (Convergence to the Conditioned Process). Under Assumption (A) and for $t \geq 0$, there exist $B, C > 0$ such that, for all $\eta \in E$, and any probability measure μ , we have

$$\sup_{\|\varphi\|_\infty \leq 1} \mathbb{E}_\eta[|m(\eta_t)(\varphi) - \mu T_t \varphi|] \leq C e^{Bt} \left(\frac{1}{\sqrt{N}} + d_{TV}(m(\eta), \mu) \right).$$

All constants are explicit and detailed in the proof (in particular, they do not depend on N and t).

The proof is based on an estimation of correlations and on a Gronwall-type argument. More precisely our correlation estimate is given by:

Theorem 1.3 (Covariance Estimates). Let ρ be defined in Theorem 1.1. Under Assumption (A), we have for all $k, l \in F^*$, $\eta \in E$ and $t \geq 0$

$$\left| \mathbb{E}_\eta \left[\frac{\eta_t(k)}{N} \frac{\eta_t(l)}{N} \right] - \mathbb{E}_\eta \left[\frac{\eta_t(k)}{N} \right] \mathbb{E}_\eta \left[\frac{\eta_t(l)}{N} \right] \right| \leq \frac{2(\mathbf{Q}_1 + \mathbf{p})}{N} \frac{1 - e^{-2\rho t}}{\rho},$$

with the convention $(1 - e^{-2\rho t})\rho^{-1} = 2t$ when $\rho = 0$.

This theorem gives a decay of the variances and the covariances of the marginals of η . Actually, it does not give any information on the correlation but this slight abuse of language is used to be consistent with other previous works [1,18].

The previous theorem is a consequence of Theorem 2.6 which gives some bounds on the correlations of more general functional of η . The proof of this result comes from a commutation relation between the *carré du champs* operator and the semigroup of η . This commutation-type relation gives a decay of the variance and thus, by the Cauchy–Schwarz inequality, of the correlations. The previous bound is uniform in time when $\rho > 0$ and it generalizes several previous work [1,18]. Indeed, as our proof differs completely to [1,18] (proof based on a comparison with the voter model), we are able to use more complex functional of η and our bounds are uniform in time. In particular, taking the limit $t \rightarrow +\infty$ when $\rho > 0$, we have the decay of the correlations under the invariant distribution of $(\eta_t)_{t \geq 0}$. This seems to be new (in discrete or continuous state space).

Theorem 1.2 is a generalization of [1, Theorem 1.3], [18, Theorem 1.2] and of [20, Theorem 2.2]. Our assumptions are weaker and our convergence estimate is in a stronger form. We can also cite [13, Theorem 1.1] and [31, Theorem 1] which give the same kind of bound with a less explicit constant. However, these two theorems cover a more general setting. This theorem permits to extend the properties of the particle system to the conditioned process; see the next subsection. The proof of Theorem 1.2 differs from all these theorems; it seems simpler and is only based on a Gronwall argument and the correlation estimates.

Two main consequences. We summarize two important consequences of our main theorems. Firstly, as ρ , defined in Theorem 1.1, does not depend on N , we can take the limit as $N \rightarrow +\infty$ in Theorem 1.1. This gives an “easy-to-verify” criterion to prove the existence, uniqueness of a quasi-stationary distribution and the exponential convergence of the conditioned process to it.

Corollary 1.4 (Convergence to the QSD). *Suppose that ρ is positive and that Assumption (A) holds. For any probability measure μ, ν , we have*

$$\forall t \geq 0, \quad d_{TV}(\mu T_t, \nu T_t) \leq e^{-\rho t} d_{TV}(\mu, \nu). \quad (2)$$

In particular, there exists a unique quasi-stationary distribution ν_{qs} for $(T_t)_{t \geq 0}$ and for any probability measure μ , we have

$$\forall t \geq 0, \quad d_{TV}(\mu T_t, \nu_{qs}) \leq e^{-\rho t}.$$

This corollary is closely related to several previous works [11], [14, Theorem 1.1], [24, Theorem 3] and [18, Theorem 1.1]. When F is finite, the oldest result dates from 1967 [11] where Darroch and Seneta give a similar bound without additional assumption. Nevertheless, the constants are less explicit because the proof is based on Perron–Frobenius Theorem. The other results are more recent. Under a slightly weaker condition, we recover [18, Theorem 1.1] in a stronger convergence and with an estimation of the rate of convergence. As in [14, Theorem 1.1], a mixing condition for Q and a regularity one for p_0 are assumed to obtain an exponential convergence to a QSD; namely, we assume that λ is large enough and $(\sup(p_0) - \inf(p_0))$ is small enough. In [14, Theorem 1.1] they only need that $\sup(p_0) < +\infty$ but, their mixing condition is stronger than ours. Finally [24, Theorem 3] gives a weaker condition to obtain an exponential convergence with (generally) a lower and less explicit rate of convergence when our result applies. Also note that Assumption (A) is not necessary; see Remark 2.8.

Without limiting results, several works establish existence and/or uniqueness of a QSD; see [10,25] for references. Let us mention the main result of [17]. In this article, the authors

prove that, under the condition that $F = \mathbb{N}$, X is irreducible over \mathbb{N}^* and

$$\lim_{x \rightarrow \infty} \mathbb{P}(T_0 < t \mid X_0 = x) = 0, \quad (3)$$

where T_0 is the hitting time of 0, a sufficient and necessary condition for the existence of a QSD is that T_0 has an exponential moment. Even if this result is sharp (it is an equivalence), condition (3) can be restrictive. In particular, if the death rate is constant over \mathbb{N} , condition (3) never holds. In contrast, our results are better in this case, because, it is enough that $\lambda > 0$ to have existence of a QSD.

Our second corollary gives a uniform bound for the limits (b) and (c), namely the convergences as N tends to infinity:

Corollary 1.5 (Uniform Bounds). *If $\rho > 0$, then under the assumptions of Theorem 1.2, there exist $K_0, \gamma > 0$ such that, for every $\eta \in E$,*

$$\sup_{t \geq 0} \sup_{\|\varphi\|_\infty \leq 1} \mathbb{E}_\eta [|m(\eta_t)(\varphi) - m(\eta)T_t\varphi|] \leq \frac{K_0}{N^\gamma}.$$

All constants are explicit and given in (20).

In particular, if η is distributed according to the invariant measure ν_N , then under the assumptions of the previous corollary, there exist $K_0 > 0$ and $\gamma > 0$ such that

$$\mathbb{E} [|m(\eta)(\varphi) - \nu_{\text{qs}}(\varphi)|] \leq \frac{K_0}{N^\gamma}, \quad (4)$$

for every φ satisfying $\|\varphi\|_\infty \leq 1$. Namely, under its invariant distribution, the particle system converges to the QSD. Without rate of convergence, this limiting result was proved in [1, Theorem 2] when F is finite. Whereas, here, a rate of convergence is given. To our knowledge, it is the first bound of convergence for this limit. Whenever F^* is finite, the conclusion of the previous corollary holds with a less explicit γ even when $\rho \leq 0$; see Remark 2.9. Note also that, closely related, article [28] gives a similar result when the underlying dynamics is diffusive instead of discrete. Its approach is completely different and based on martingale properties and on spectral properties associated to Schrödinger equation.

The remainder of the paper is as follows. Section 2 gives the proofs of our main theorems; Section 2.1 contains the proof of Theorem 1.1, Section 2.2 the proof of Theorem 1.2 and the last subsection the proof of the corollaries. We conclude the paper with Section 3, where we give the two examples mentioned above. The first one illustrates the sharpness of our results. The study of the second one is reduced to a very simple process for which few properties are known. It illustrates the need of general theorems as those previously introduced.

2. Proof of the main theorems

In this section, we prove Theorems 1.1 and 1.2 and the corollaries stated before. Let us recall that the generator of the Fleming–Viot process with N particles applied to bounded functions $f : E \rightarrow \mathbb{R}$ and $\eta \in E$, is given by

$$\mathcal{L}f(\eta) = \sum_{i \in F^*} \eta(i) \sum_{j \in F^*} \left(Q_{i,j} + p_0(i) \frac{\eta(j)}{N-1} \right) (f(T_{i \rightarrow j}\eta) - f(\eta)). \quad (5)$$

Now let us give two remarks about the dynamics of the Fleming–Viot particle system.

Remark 2.1 (*Translation of the Death Rate*). Let $(P_t)_{t \geq 0}$ and $(P'_t)_{t \geq 0}$ be two semi-groups with the same transition rate Q but different death rates p_0, p'_0 and let $(T_t)_{t \geq 0}, (T'_t)_{t \geq 0}$ be their corresponding conditioned semi-groups respectively. Using the fact that

$$P_t \mathbf{1}_{\{0\}^c} = \mathbb{E} \left[e^{-\int_0^t p_0(X_s) ds} \right] \quad \text{and} \quad P'_t \mathbf{1}_{\{0\}^c} = \mathbb{E} \left[e^{-\int_0^t p'_0(X'_s) ds} \right],$$

for every $t \geq 0$, it is easy to see that $(T_t)_{t \geq 0} = (T'_t)_{t \geq 0}$ as soon as $p_0 - p'_0$ is constant. This invariance by translation is not conserved by the Fleming–Viot processes. The larger the p_0 is, the more jumps are obtained and the larger the variance becomes. This is why our criterion about the existence of QSD does not depend on $\inf(p_0)$ and why our propagation of chaos result depends on it.

Remark 2.2 (*Non-explosion*). The particle dynamics guarantees the existence of the process $(\eta_t)_{t \geq 0}$ under the condition that there is no explosion. In other words, our construction is global as long as the particles only jump finitely many times in any finite time interval. We naturally assume that the Markov process with transition Q is not explosive but it is not enough for the existence of the particle system. Indeed, an example of explosive Fleming–Viot particle system can be found in [4]. However, the assumption that p_0 is bounded is trivially sufficient to guarantee this non-explosion.

2.1. Proof of Theorem 1.1

Proof of Theorem 1.1. We build a coupling between two Fleming–Viot particle systems, $(\eta_t)_{t \geq 0}$ and $(\eta'_t)_{t \geq 0}$, generated by (1), starting respectively from some random configurations η_0, η'_0 in E . We will prove that they will be closer and closer.

Let us begin by roughly describing our coupling and then be more precise. For every $t \geq 0$, we set $\xi(t) = \xi = (\xi_1, \dots, \xi_N) \in (F^*)^N$ and $\xi'(t) = \xi' = (\xi'_1, \dots, \xi'_N)$ the respective positions of the N particles of the two configurations η_t and η'_t . Then

$$\forall i \in F^*, \quad \eta_t(i) = \text{card}\{1 \leq k \leq N \mid \xi_k = i\} \quad \text{and} \quad \eta'_t(i) = \text{card}\{1 \leq k \leq N \mid \xi'_k = i\}.$$

Distance $d_1(\eta, \eta')$ represents the number of particles which are not in the same site; namely, changing the indexation,

$$d_1(\eta, \eta') = \text{card}\{1 \leq k \leq N \mid \xi_k \neq \xi'_k\}.$$

We then couple our two processes in order to maximize the chance that two particles coalesce. In a first time, we forget the interaction; we have two systems of N particles evolving independently from each others. If two particles are in the same site, $\xi_k = \xi'_k$, then the Markov property entails that we can make them jump together. When two particles are not in the same site, we can choose our jumps time in such a way that one goes to the second one, with positive probability. These steps are represented by the jumps rate A_Q below.

Nevertheless, the situation is trickier when we consider the interaction. Indeed, let us now disregard the underlying dynamics and only regard the interaction. If two particles are in the same site, $\xi_k = \xi'_k$, then they have to be killed and jump over the other particles. If the empirical measures are the same $\eta = \eta'$, then we can couple the two particles in such a way they die at the same time (because they are in the same site) and jump in the same site (because the empirical measures are equal). If $\eta \neq \eta'$, then we cannot do this but we can maximize the probability to coalesce. Indeed there are $N - d_1(\eta, \eta')$ particles which are in the same site and then a probability

$(N - d_1(\eta, \eta'))/(N - 1)$ to coalesce. If two particles are not in the same site, $\xi_k \neq \xi'_k$, we can try to kill one before the other and put it in the same site. This is also not always possible.

Before expressing precisely the jumps rates, let us give some explanations. We call first configuration the particles represented by $\{\xi_k\}$ and the second configuration the particles represented by $\{\xi'_k\}$. We speak about couple of particles when there are two particles coming from different configurations. There are $\eta(i) = \text{card}\{k \mid \xi_k = i\}$ particles on the site i and we can write

$$\eta(i) = (\eta(i) - \eta'(i))_+ + \eta(i) \wedge \eta'(i),$$

where $(\cdot)_+ = \max(0, \cdot)$. The part $\eta(i) \wedge \eta'(i)$ represents the number of couples of particles on i and $(\eta(i) - \eta'(i))_+$ the rest of particles coming from the first configuration. Note that

$$\sum_{i \in F^*} (\eta(i) - \eta'(i))_+ = \sum_{i \in F^*} (\eta'(i) - \eta(i))_+ = d_1(\eta, \eta') = N - \sum_{i \in F^*} \eta(i) \wedge \eta'(i). \quad (6)$$

Now, we describe in detail our coupling. It is Markovian and we describe it by expressing its generator and its jumps rate; for every bounded function f and $\eta, \eta' \in E$, its generator \mathbb{L} is given by

$$\mathbb{L}f(\eta, \eta') = \sum_{i, i', j, j' \in F^*} A(i, i', j, j')(f(T_{i \rightarrow j}\eta, T_{i' \rightarrow j'}\eta') - f(\eta, \eta')),$$

where we decompose the jump rate A into two parts $A = A_Q + A_p$. The jumps rate A_Q , that depends only on the transition rate Q , corresponds to the jumps related to the underlying dynamics, namely it is the dynamics when a particle does not die. A Markov process having only A_Q as jumps rate corresponds to a coupling of two systems of N particles evolving independently from each other. The jumps rate A_p , corresponds to the redistribution dynamics and depends only on p_0 ; it does not depend on the underlying dynamics but only on the interaction. The construction of A_Q is then more classic and the construction of A_p is new and specific to this interaction. In what follows, we give the expressions of A_p and A_Q ; the points i, i', j, j' are always different in twos.

- There are $\eta(i) \wedge \eta'(i)$ couples of particles on site $i \in F^*$.
 - For each couple, both particles can jump to the same site $j \in F^*$, at the same time and through the underlying dynamics. This gives the following jumps rate:

$$A_Q(i, i, j, j) = (\eta(i) \wedge \eta'(i)) Q_{i, j}.$$

- Both of them can die at the same time. With probability $\frac{\eta(j) \wedge \eta'(j)}{N-1}$, they can jump to the same site j ; this gives

$$A_p(i, i, j, j) = p_0(i) (\eta(i) \wedge \eta'(i)) \frac{\eta(j) \wedge \eta'(j)}{N-1}.$$

With probability $\frac{\eta(i) \wedge \eta'(i) - 1}{N-1}$, both particles jump where they come from and, so, this changes anything. With probability

$$\left(1 - \frac{\sum_{k \in F^*} \eta(k) \wedge \eta'(k) - 1}{N-1}\right) \frac{(\eta(j) - \eta'(j))_+}{\sum_{k \in F^*} (\eta(k) - \eta'(k))_+} \frac{(\eta'(j') - \eta(j'))_+}{\sum_{k \in F^*} (\eta'(k) - \eta(k))_+}, \quad (7)$$

they can jump to two different sites j, j' . Indeed, with probability $1 - \frac{\sum_{k \in F^*} \eta(k) \wedge \eta'(k) - 1}{N-1}$, they can jump in different sites, and conditionally on this event, with probability $\frac{(\eta(j) - \eta'(j))_+}{\sum_{k \in F^*} (\eta'(k) - \eta(k))_+}$, the first particle jumps in site j and, with probability $\frac{(\eta'(j') - \eta(j'))_+}{\sum_{k \in F^*} (\eta'(k) - \eta(k))_+}$, the second one jumps in site j' . Probability (7) is equal to

$$\frac{(\eta(j) - \eta'(j))_+ \cdot (\eta'(j') - \eta(j'))_+}{(N-1)d_1(\eta, \eta')}.$$

In short, this gives the following jump rates:

$$A_p(i, i, j, j') = p_0(i) (\eta(i) \wedge \eta'(i)) \frac{(\eta(j) - \eta'(j))_+ \cdot (\eta'(j') - \eta(j'))_+}{(N-1)d_1(\eta, \eta')}.$$

- For every site $i \in F^*$ there are $(\eta(i) - \eta'(i))_+$ particles from the first configuration which are not in a couple. For each of these particles, we choose, uniformly at random, a particle of the second configuration (which is not coupled with another particle as in the first point). This particle, chosen at random, is on the site $i' \in F^*$ with probability

$$\frac{(\eta'(i') - \eta(i'))_+}{\sum_k (\eta'(k) - \eta(k))_+} = \frac{(\eta'(i') - \eta(i'))_+}{d_1(\eta, \eta')}.$$

- For one of these new couple of particles coming from sites $i \neq i'$, both particles can jump at the same time to the same site j (different from i, i'), through the underlying dynamics; this gives

$$A_Q(i, i', j, j) = (\eta(i) - \eta'(i))_+ \cdot \frac{(\eta'(i') - \eta(i'))_+}{d_1(\eta, \eta')} \cdot (Q_{i,j} \wedge Q_{i',j}).$$

Nevertheless, these two particles do not have the same jump rates (because they do not come from the same site), so it is possible that one jumps to another site while the other one does not jump (also through the underlying dynamics); this gives

$$A_Q(i, i', j, i') = (\eta(i) - \eta'(i))_+ \cdot \frac{(\eta'(i') - \eta(i'))_+}{d_1(\eta, \eta')} \cdot (Q_{i,j} - Q_{i',j})_+,$$

and

$$A_Q(i, i', i, j') = (\eta(i) - \eta'(i))_+ \cdot \frac{(\eta'(i') - \eta(i'))_+}{d_1(\eta, \eta')} \cdot (Q_{i',j'} - Q_{i,j'})_+.$$

Also, one of them can jump to the site of the second one:

$$A_Q(i, i', i', i') = \frac{(\eta(i) - \eta'(i))_+ \cdot (\eta'(i') - \eta(i'))_+}{d_1(\eta, \eta')} Q_{i,i'},$$

and

$$A_Q(i, i', i, i) = \frac{(\eta(i) - \eta'(i))_+ \cdot (\eta'(i') - \eta(i'))_+}{d_1(\eta, \eta')} Q_{i',i}.$$

- We focus now our attention on the redistribution dynamics. We would like that both particles of a couple die at the same time and jump to the same site j (where a couple of particles exists; that is with probability $\frac{\eta(j) \wedge \eta'(j)}{N-1}$). This gives:

$$A_p(i, i', j, j) = (\eta(i) - \eta'(i))_+ \cdot \frac{(\eta'(i') - \eta(i'))_+}{d_1(\eta, \eta')} \cdot (p_0(i) \wedge p_0(i')) \cdot \frac{\eta(j) \wedge \eta'(j)}{N-1}.$$

But, even if they die at same time, they can jump to different sites with rate

$$A_p(i, i', j, j') = (p_0(i) \wedge p_0(i')) \frac{(\eta(i) - \eta'(i))_+ (\eta'(i') - \eta(i'))_+}{d_1(\eta, \eta')} \\ \cdot \frac{(\eta(j) - \eta'(j))_+ (\eta'(j') - \eta(j'))_+}{(N-1)d_1(\eta, \eta')}.$$

However, this is not always possible to kill them at the same time. If they do not, then the dying particle jumps uniformly to a particle of its configuration; this gives

$$A_p(i, i', j, i') = (\eta(i) - \eta'(i))_+ \cdot \frac{(\eta'(i') - \eta(i'))_+}{d_1(\eta, \eta')} \cdot (p_0(i) - p_0(i'))_+ \cdot \frac{\eta(j)}{N-1},$$

and

$$A_p(i, i', i, j') = (\eta(i) - \eta'(i))_+ \cdot \frac{(\eta'(i') - \eta(i'))_+}{d_1(\eta, \eta')} \cdot (p_0(i') - p_0(i))_+ \cdot \frac{\eta(j')}{N-1}.$$

We set, for every measurable function f ,

$$\mathbb{L}_Q f(\eta, \eta') = \sum_{i, i', j, j' \in F^*} A_Q(i, i', j, j') (f(T_{i \rightarrow j} \eta, T_{i' \rightarrow j'} \eta') - f(\eta, \eta')),$$

and

$$\mathbb{L}_p f(\eta, \eta') = \sum_{i, i', j, j' \in F^*} A_p(i, i', j, j') (f(T_{i \rightarrow j} \eta, T_{i' \rightarrow j'} \eta') - f(\eta, \eta')).$$

Our coupling is totally defined.

Lemma 2.5 shows that if a measurable function f on $E \times E$ does not depend on its second (resp. first); that is with a slight abuse of notation:

$$\forall \eta, \eta' \in E, \quad f(\eta, \eta') = f(\eta) \text{ (resp. } f(\eta, \eta') = f(\eta')),$$

then $\mathbb{L}f(\eta, \eta') = \mathcal{L}f(\eta)$ (resp. $\mathbb{L}f(\eta, \eta') = \mathcal{L}f(\eta')$). This property ensures that the couple $(\eta_t, \eta'_t)_{t \geq 0}$ generated by \mathbb{L} is well a coupling of processes generated by \mathcal{L} (that is of Fleming–Viot processes)

Now, let us prove that the distance between η_t and η'_t decreases exponentially. We have

$$\mathbb{L}_p d_1(\eta, \eta') \leq \sum_{i \in F^*} p_0(i) (\eta(i) \wedge \eta'(i)) \frac{d_1(\eta, \eta')}{N-1} - \sum_{i, i' \in F^*} (p_0(i) \wedge p_0(i')) \\ \times \frac{(\eta(i) - \eta'(i))_+ (\eta'(i') - \eta(i'))_+}{d_1(\eta, \eta')} \sum_{j \in F^*} \frac{\eta(j) \wedge \eta'(j)}{N-1} \\ \leq (\sup(p_0) - \inf(p_0)) \frac{d_1(\eta, \eta')}{N-1} (N - d_1(\eta, \eta')) \\ \leq (\sup(p_0) - \inf(p_0)) d_1(\eta, \eta').$$

Now,

$$\begin{aligned}\mathbb{L}_{\mathcal{Q}} d_1(\eta, \eta') &\leq - \sum_{i, i' \in F^*} \left(\mathcal{Q}_{i, i'} + \mathcal{Q}_{i', i} + \sum_{j \neq i, i'} \mathcal{Q}_{i, j} \wedge \mathcal{Q}_{i', j} \right) \\ &\quad \times \frac{(\eta(i) - \eta'(i))_+ (\eta'(i') - \eta(i'))_+}{d_1(\eta, \eta')} \\ &\leq -\lambda d_1(\eta, \eta').\end{aligned}$$

We deduce that $\mathbb{L} d_1(\eta, \eta') \leq -\rho d_1(\eta, \eta')$. Now let $(\mathbb{P}_t)_{t \geq 0}$ be the semi-group associated with the generator \mathbb{L} . Using the equality $\partial_t \mathbb{P}_t f = \mathbb{P}_t \mathbb{L} f$ and Gronwall Lemma, we have, for every $t \geq 0$, $\mathbb{P}_t d_1 \leq e^{-\rho t} d_1$; namely

$$\mathbb{E}[d_1(\eta_t, \eta'_t)] \leq e^{-\rho t} \mathbb{E}[d_1(\eta_0, \eta'_0)].$$

Taking the infimum over all couples (η_0, η'_0) , the claim follows. The existence and the uniqueness of an invariant distribution come from classical arguments; see for instance [7, Theorem 5.23]. \square

As it is easy to see that the distance \mathcal{W}_{d_1} is larger than the total variation distance, we have the following consequence:

Corollary 2.3 (Coalescent Time Estimate). *For all $t \geq 0$, we have*

$$d_{TV}(\mathcal{L}(\eta_t), \mathcal{L}(\eta'_t)) \leq e^{-\rho t} \mathcal{W}_{d_1}(\mathcal{L}(\eta_0), \mathcal{L}(\eta'_0)).$$

In particular, if $\rho > 0$ the invariant distribution ν_N satisfies

$$d_{TV}(\mathcal{L}(\eta_t), \nu_N) \leq e^{-\rho t} \mathcal{W}_{d_1}(\mathcal{L}(\eta_0), \nu_N).$$

The proof is simple and given for sake of completeness.

Proof. Using Theorem 1.1, we find

$$\begin{aligned}d_{TV}(\mathcal{L}(\eta_t), \mathcal{L}(\eta'_t)) &= \inf_{\substack{\eta_t \sim \mathcal{L}(\eta_t) \\ \eta'_t \sim \mathcal{L}(\eta'_t)}} \mathbb{E} \left[\mathbf{1}_{\eta_t \neq \eta'_t} \right] \\ &\leq \inf_{\substack{\eta_t \sim \mathcal{L}(\eta_t) \\ \eta'_t \sim \mathcal{L}(\eta'_t)}} \mathbb{E} [d_1(\eta_t, \eta'_t)] = \mathcal{W}_{d_1}(\mathcal{L}(\eta_t), \mathcal{L}(\eta'_t)) \\ &\leq e^{-\rho t} \mathcal{W}_{d_1}(\mathcal{L}(\eta_0), \mathcal{L}(\eta'_0)). \quad \square\end{aligned}$$

Remark 2.4 (Generalization). As we can see at the end of the paper, in the case where F^* contains only two elements, the coupling that we use is pretty good but our estimation of the distance is (in general) too rough. There is some natural way to change the bound/criterion that we found. The first one is to use another more appropriate distance. This technique is in general useful in other (Markovian) contexts [6,8,15]. Another way is to find a contraction after a certain time: it is the Foster–Lyapunov-type techniques [3,23,26]. These types of techniques give more general criteria but are useless for small times and the formulas we get are less explicit. All of these techniques will give different criteria that are not necessarily better. Finally note that, in all

the paper, we can replace ρ by

$$\rho' = \inf_{i, i' \in F^*} \left\{ p_0(i) \wedge p_0(i') + Q_{i, i'} + Q_{i', i} + \sum_{j \neq i, i'} Q_{i, j} \wedge Q_{i', j} \right\} - \sup(p_0),$$

and all conclusions hold. Indeed, we have to bound directly $\mathbb{L}d_1$ instead of bounding separately $\mathbb{L}_Q d_1$ and $\mathbb{L}_p d_1$.

Lemma 2.5 (*Marginals of Process Generated by \mathbb{L} are Generated by \mathcal{L}*). *With the notation of the proof of Theorem 1.1, let f be a measurable function on $E \times E$ not depending on its second (resp. first) variable; that is with a slight abuse of notation:*

$$\forall \eta, \eta' \in E, \quad f(\eta, \eta') = f(\eta) \text{ (resp. } f(\eta, \eta') = f(\eta')).$$

We have $\mathbb{L}f(\eta, \eta') = \mathcal{L}f(\eta)$ (resp. $\mathbb{L}f(\eta, \eta') = \mathcal{L}f(\eta')$). In particular, the couple $(\eta_t, \eta'_t)_{t \geq 0}$ generated by \mathbb{L} is well a coupling of processes generated by \mathcal{L} .

Proof. Let f be such a function. On the one hand

$$\mathbb{L}_Q f(\eta, \eta') = \sum_{i \in F^*} (\eta(i) \wedge \eta'(i)) \sum_{j \in F^*} Q_{i, j} (f(T_{i \rightarrow j} \eta) - f(\eta)) \quad (8)$$

$$+ \sum_{i, i' \in F^*} \frac{(\eta(i) - \eta'(i))_+ \times (\eta'(i') - \eta(i'))_+}{d_1(\eta, \eta')} \quad (9)$$

$$\times \sum_{j \in F^*, j \neq i, i' \in F^*} Q_{i, j} \wedge Q_{i', j} (f(T_{i \rightarrow j} \eta) - f(\eta))$$

$$+ \sum_{i, i' \in F^*} \frac{(\eta(i) - \eta'(i))_+ \times (\eta'(i') - \eta(i'))_+}{d_1(\eta, \eta')} \quad (10)$$

$$\times \sum_{j \in F^*, j \neq i, i'} (Q_{i, j} - Q_{i', j})_+ (f(T_{i \rightarrow j} \eta) - f(\eta))$$

$$+ \sum_{i, i' \in F^*} \frac{(\eta(i) - \eta'(i))_+ \times (\eta'(i') - \eta(i'))_+}{d_1(\eta, \eta')}$$

$$\times Q_{i, i'} (f(T_{i \rightarrow i'} \eta) - f(\eta)). \quad (11)$$

Using (6), we find

$$\begin{aligned} (9) + (10) + (11) &= \sum_{i, i' \in F^*} \frac{(\eta(i) - \eta'(i))_+ \times (\eta'(i') - \eta(i'))_+}{d_1(\eta, \eta')} \\ &\quad \times \sum_{j \in F^*, j \neq i, i' \in F^*} Q_{i, j} (f(T_{i \rightarrow j} \eta) - f(\eta)) \\ &= \sum_{i \in F^*} (\eta(i) - \eta'(i))_+ \sum_{j \in F^*, j \neq i \in F^*} Q_{i, j} (f(T_{i \rightarrow j} \eta) - f(\eta)) \\ &= \sum_{i \in F^*} (\eta(i) - \eta'(i))_+ \sum_{j \in F^*} Q_{i, j} (f(T_{i \rightarrow j} \eta) - f(\eta)). \end{aligned}$$

We deduce that

$$\mathbb{L}_Q f(\eta, \eta') = \sum_{i \in F^*} \eta(i) \sum_{j \in F^*} Q_{i,j} (f(T_{i \rightarrow j} \eta) - f(\eta)).$$

On the other hand,

$$\mathbb{L}_p f(\eta, \eta') = \sum_{i \in F^*} p_0(i) (\eta(i) \wedge \eta'(i)) \quad (12)$$

$$\begin{aligned} & \times \sum_{j \in F^*} \left[\frac{\eta(j) \wedge \eta'(j)}{N-1} + \sum_{j' \in F^*} \frac{(\eta(j) - \eta'(j))_+ \times (\eta'(j') - \eta(j'))_+}{(N-1)d_1(\eta, \eta')} \right] (f(T_{i \rightarrow j} \eta) - f(\eta)) \\ & + \sum_{i, i' \in F^*} (p_0(i) \wedge p_0(i')) \frac{(\eta(i) - \eta'(i))_+ \times (\eta'(i') - \eta(i'))_+}{d_1(\eta, \eta')} \quad (13) \end{aligned}$$

$$\begin{aligned} & \times \sum_{j \in F^*} \left[\frac{\eta(j) \wedge \eta'(j)}{N-1} + \sum_{j' \in F^*} \frac{(\eta(j) - \eta'(j))_+ \times (\eta'(j') - \eta(j'))_+}{(N-1)d_1(\eta, \eta')} \right] (f(T_{i \rightarrow j} \eta) - f(\eta)) \\ & + \sum_{i, i' \in F^*} (p_0(i) - p_0(i'))_+ \frac{(\eta(i) - \eta'(i))_+ \times (\eta'(i') - \eta(i'))_+}{d_1(\eta, \eta')} \quad (14) \\ & \times \sum_{j \in F^*} \frac{\eta(j)}{N-1} (f(T_{i \rightarrow j} \eta) - f(\eta)). \end{aligned}$$

We have,

$$\begin{aligned} (12) + (13) &= \sum_{i \in F^*} \left[p_0(i) (\eta(i) \wedge \eta'(i)) + \sum_{i' \in F^*} (p_0(i) \wedge p_0(i')) \right. \\ & \quad \times \left. \frac{(\eta(i) - \eta'(i))_+ \times (\eta'(i') - \eta(i'))_+}{d_1(\eta, \eta')} \right] \\ & \quad \times \sum_{j \in F^*} \frac{\eta(j)}{N-1} (f(T_{i \rightarrow j} \eta) - f(\eta)) \end{aligned}$$

and

$$\begin{aligned} (14) &= \sum_{i \in F^*} \left[p_0(i) (\eta(i) - \eta'(i))_+ - \sum_{i' \in F^*} (p_0(i) \wedge p_0(i')) \right. \\ & \quad \times \left. \frac{(\eta(i) - \eta'(i))_+ \times (\eta'(i') - \eta(i'))_+}{d_1(\eta, \eta')} \right] \\ & \quad \times \sum_{j \in F^*} \frac{\eta(j)}{N-1} (f(T_{i \rightarrow j} \eta) - f(\eta)). \end{aligned}$$

We deduce that

$$\mathbb{L}_p f(\eta, \eta') = \sum_{i \in F^*} p_0(i) \eta(i) \sum_{j \in F^*} \frac{\eta(j)}{N-1} (f(T_{i \rightarrow j} \eta) - f(\eta)).$$

Finally,

$$\mathbb{L}f(\eta, \eta') = \mathbb{L}_Q f(\eta, \eta') + \mathbb{L}_p f(\eta, \eta') = \mathcal{L}f(\eta).$$

By a symmetry argument, the result also holds when f only depends on its second component. \square

2.2. Proofs of Theorems 1.2 and 1.3

The proof of Theorem 1.2 is done in two steps. Firstly, we estimate the correlations between the number of particles over the sites and then we estimate the distance in total variation via the Kolmogorov equation. Let us introduce some notations. For every bounded functions f, g , every $\eta \in E$ and every random variable X , we set

$$\text{Cov}_\eta[f(X), g(X)] = \mathbb{E}_\eta[f(X)g(X)] - \mathbb{E}_\eta[f(X)]\mathbb{E}_\eta[g(X)],$$

and

$$\text{Var}_\eta[f(X)] = \text{Cov}_\eta[f(X), f(X)].$$

Let $(S_t)_{t \geq 0}$ be the semigroup of $(\eta_t)_{t \geq 0}$ defined by

$$S_t f(\eta) = \mathbb{E}_\eta[f(\eta_t)],$$

for every $t \geq 0$, $\eta \in E$ and bounded function f . If μ is a probability measure on E and $t \geq 0$, then μS_t is the measure defined by

$$\mu S_t f = \int_E S_t f(y) \mu(dy).$$

It represents the law of η_t when η_0 is distributed according to μ . We also introduce the *carré du champ* operator Γ defined, for any bounded function f and $\eta \in E$, by

$$\begin{aligned} \Gamma f(\eta) &= \mathcal{L}(f^2)(\eta) - 2f(\eta)\mathcal{L}f(\eta) \\ &= \sum_{i,j \in F^*} \eta(i) \left(Q_{i,j} + p_0(i) \frac{\eta(j)}{N-1} \right) (f(T_{i \rightarrow j} \eta) - f(\eta))^2. \end{aligned} \quad (15)$$

We present now an improvement of Theorem 1.3.

Theorem 2.6 (Correlations for Lipschitz Functional). *Let g, h be two 1-Lipschitz mappings on (E, d_1) ; namely*

$$|g(\eta) - g(\eta')| \leq d_1(\eta, \eta') \quad \text{and} \quad |h(\eta) - h(\eta')| \leq d_1(\eta, \eta'),$$

for every $\eta, \eta' \in E$. Under Assumption (A) we have for all $t \geq 0$ and $\eta \in E$,

$$|\text{Cov}_\eta(g(\eta_t), h(\eta_t))| \leq \frac{1 - e^{-2\rho t}}{2\rho} N (\mathbf{Q}_1 + \mathbf{p}),$$

with the convention $(1 - e^{-2\rho t})\rho^{-1} = 2t$ when $\rho = 0$.

In particular, if $\rho > 0$, then the previous bound is uniform.

Proof. For any function g on E and $t \geq 0$, we have

$$\text{Var}_\eta(g(\eta_t)) = S_t(g^2)(\eta) - (S_t g)^2(\eta) = \int_0^t S_s \Gamma S_{t-s} g(\eta) ds.$$

Indeed, setting, for any $s \in [0, t]$ and $\eta \in E$, $\Psi_\eta(s) = S_s [(S_{t-s} g)^2](\eta)$ and $\psi(s) = S_{t-s} g$, we get

$$\forall s \geq 0, \quad \Psi'_\eta(s) = S_s [\mathcal{L}\psi^2 - 2\psi \mathcal{L}\psi](\eta) = S_s \Gamma \psi(s)(\eta),$$

and so,

$$\text{Var}_\eta(g(\eta_t)) = \Psi_\eta(t) - \Psi_\eta(0) = \int_0^t S_s \Gamma S_{t-s} g(\eta) ds.$$

Now, if g is a 1-Lipschitz mapping with respect to d_1 , then

$$|S_{t-s} g(T_{i \rightarrow j} \eta) - S_{t-s} g(\eta)| \leq \mathbb{E} [|g(\eta'_{t-s}) - g(\eta_{t-s})|] \leq \mathbb{E} [d_1(\eta_{t-s}, \eta'_{t-s})],$$

where η_{t-s}, η'_{t-s} evolve as Fleming–Viot particle systems with initial conditions η and $T_{i \rightarrow j} \eta$. Thus, using Theorem 1.1, we obtain

$$\begin{aligned} |S_{t-s} g(T_{i \rightarrow j} \eta) - S_{t-s} g(\eta)| &\leq \mathcal{W}_{d_1}(\mathcal{L}(\eta_{t-s}), \mathcal{L}(\eta'_{t-s})) \\ &\leq e^{-\rho(t-s)} d_1(T_{i \rightarrow j} \eta, \eta) \\ &\leq e^{-\rho(t-s)} \mathbf{1}_{i \neq j}. \end{aligned} \tag{16}$$

Hence,

$$\|\Gamma S_{t-s} g\|_\infty = \sup_{\eta \in E} |\Gamma S_{t-s} g(\eta)| \leq N e^{-2\rho(t-s)} (\mathbf{Q}_1 + \mathbf{p}).$$

Indeed, using (15) and (16) we have

$$\begin{aligned} |\Gamma S_{t-s} g(\eta)| &\leq e^{-2\rho(t-s)} \left(N \mathbf{Q}_1 + \mathbf{p} \sum_{i \in F^*} \eta(i) \sum_{j \neq i} \frac{\eta(j)}{N-1} \right) \\ &\leq e^{-2\rho(t-s)} \left(N \mathbf{Q}_1 + \frac{\mathbf{p}}{N-1} \sum_{i \in F^*} \eta(i) (N - \eta(i)) \right) \\ &\leq e^{-2\rho(t-s)} \left(N \mathbf{Q}_1 + \frac{\mathbf{p}}{N-1} N(N-1) \right). \end{aligned}$$

Finally, the Cauchy–Schwarz inequality and the first part of the proof give

$$\begin{aligned} |\text{Cov}_\eta(g(\eta_t), h(\eta_t))| &\leq \text{Var}_\eta(g(\eta_t))^{1/2} \text{Var}_\eta(h(\eta_t))^{1/2} \\ &\leq N \frac{1 - e^{-2\rho t}}{2\rho} (\mathbf{Q}_1 + \mathbf{p}). \quad \square \end{aligned}$$

Proof of Theorem 1.3. Fix $l \in F^*$ and set $\varphi_l : \eta \mapsto \eta(l)$. The function $\varphi_l/2$ is a 1-Lipschitz mapping with respect to d_1 , so we apply the previous theorem. \square

Remark 2.7 (Generalization). A slight modification of the proof shows that if Assumption (A) holds and there exist $C > 0$ and $\lambda > 0$ such that for any processes $(\eta_t)_{t>0}$ and $(\eta'_t)_{t>0}$ generated

by (1), and for any $t > 0$, we have

$$\mathcal{W}_{d_1}(\mathcal{L}(\eta_t), \mathcal{L}(\eta'_t)) \leq C e^{-\lambda t} \mathcal{W}_{d_1}(\mathcal{L}(\eta_0), \mathcal{L}(\eta'_0)), \quad (17)$$

then we have, for all $t \geq 0$,

$$\text{Cov}_\eta(\eta_t(k)/N, \eta_t(l)/N) \leq \frac{2C}{N} \frac{1 - e^{-2\lambda t}}{\lambda} (\mathbf{Q}_1 + \mathbf{p}).$$

The previous theorem is an instance of this implication with $C = 1$. For instance, Eq. (17) is obtained when the state space F^* contains only two points in [9].

Proof of Theorem 1.2. The proof is based on a bias–variance type decomposition. The variance is bounded through Theorem 1.3 and the bias through Gronwall-type argument. More precisely, for $t \geq 0$, we have

$$\begin{aligned} \sup_{\|\varphi\|_\infty \leq 1} \mathbb{E}_\eta [|m(\eta_t)(\varphi) - \mu T_t \varphi|] &\leq \sup_{\|\varphi\|_\infty \leq 1} \mathbb{E}_\eta [|m(\eta_t)(\varphi) - \bar{m}(\eta_t)(\varphi)|] \\ &\quad + 2d_{TV}(\bar{m}(\eta_t), \mu T_t), \end{aligned} \quad (18)$$

where $\bar{m}(\eta_t)$ is the empirical mean measure; namely $\bar{m}(\eta_t)(k) = \mathbb{E}[m(\eta_t)(k)]$, for every $k \in F^*$. Let φ be a function such that $\|\varphi\|_\infty \leq 1$. Cauchy–Schwarz inequality gives

$$\mathbb{E}_\eta [|m(\eta_t)(\varphi) - \bar{m}(\eta_t)(\varphi)|] \leq 2N^{-1} \text{Var}(g_\varphi(\eta_t))^{1/2},$$

where $g_\varphi : \eta \mapsto \frac{1}{2} \sum_{k \in F^*} \eta(k) \varphi(k) = \frac{N}{2} m(\eta)(\varphi)$ is a 1-Lipschitz function. So by Theorem 2.6 we have

$$\sup_{\|\varphi\|_\infty \leq 1} \mathbb{E}_\eta [|m(\eta_t)(\varphi) - \bar{m}(\eta_t)(\varphi)|] \leq \sqrt{2\rho^{-1}(1 - e^{-2\rho t})(\mathbf{Q}_1 + \mathbf{p})N^{-1}}.$$

Now, to study the bias term in (18), let us introduce the following notations

$$u_k(t) = \mathbb{E}_\eta[m(\eta_t)(k)] \quad \text{and} \quad v_k(t) = \mu T_t(k).$$

It is well known that $(\mu T_t)_{t \geq 0}$ is the unique measure solution to the (non-linear) Kolmogorov forward type equations: $\mu T_0 = \mu$, and

$$\forall t \geq 0, \quad \partial_t \mu T_t(j) = \sum_{i \in F^*} (Q_{i,j} \mu T_t(i) + p_0(i) \mu T_t(i) \mu T_t(j)). \quad (19)$$

Thus

$$\partial_t v_k(t) = \sum_{i \in F^*} Q_{i,k} v_i(t) + \sum_{i \in F^*} p_0(i) v_i(t) v_k(t).$$

Also, $u_k(t) = \mathbb{E}_\eta[m(\eta_t)(k)] = S_t f(\eta)$, where $f : \eta \mapsto m(\eta)(k)$ and $(S_t)_{t \geq 0}$ is the semi-group of $(\eta_t)_{t \geq 0}$, thus, using (1), the equality $\partial_t S_t f = \mathcal{L} S_t f$ and the convention that $p_0(i) + \sum_{j \in F^*} Q_{i,j} = 0$ for every $i \in F^*$, we find

$$\partial_t u_k(t) = \sum_{i \in F^*} Q_{i,k} u_i(t) + \sum_{i \in F^*} p_0(i) u_i(t) u_k(t) - \frac{p_0(k)}{N-1} u_k(t) + R_k(t),$$

where

$$\begin{aligned} R_k(t) &= \sum_{i \in F^*} p_0(i) \left(\frac{N}{N-1} \mathbb{E}_\eta(m(\eta_t)(i)m(\eta_t)(k)) - \mathbb{E}_\eta(m(\eta_t)(i)) \mathbb{E}_\eta(m(\eta_t)(k)) \right) \\ &= \mathbb{E}_\eta \left(\left(\sum_{i \in F^*} p_0(i)m(\eta_t)(i) \right) m(\eta_t)(k) \right) - \mathbb{E}_\eta \left(\sum_{i \in F^*} p_0(i)m(\eta_t)(i) \right) \mathbb{E}_\eta(m(\eta_t)(k)) \\ &\quad + (N-1)^{-1} \mathbb{E}_\eta \left(\left(\sum_{i \in F^*} p_0(i)m(\eta_t)(i) \right) m(\eta_t)(k) \right). \end{aligned}$$

For $t \geq 0$, let us define $\epsilon(t) = \sum_{k \in F^*} |u_k(t) - v_k(t)| = 2d_{\text{TV}}(\overline{m}(\eta_t), \mu T_t)$. Using triangular inequality, Fubini–Tonelli Theorem and Assumption (A), we have

$$\begin{aligned} \epsilon(t) &= \sum_{k \in F^*} \left| u_k(0) - v_k(0) + \int_0^t \partial_s(u_k(s) - v_k(s)) ds \right| \\ &\leq \epsilon(0) + \sum_{k \in F^*} \int_0^t \left| \sum_{i \in F^*} Q_{i,k}(u_i(s) - v_i(s)) \right| + \sum_{k \in F^*} \int_0^t \left(\frac{p_0(k)}{N-1} u_k(s) + |R_k(s)| \right) ds \\ &\quad + \sum_{k \in F^*} \int_0^t \left| \sum_{i \in F^*} p_0(i) [v_i(s)(u_k(s) - v_k(s)) + u_k(s)(u_i(s) - v_i(s))] \right| ds \\ &\leq \epsilon(0) + \int_0^t (\mathbf{Q}_1 + 2\mathbf{p}) \epsilon(s) ds + \frac{\mathbf{p}t}{N-1} + \int_0^t \sum_{k \in F^*} |R_k(s)| ds. \end{aligned}$$

However, by Cauchy–Schwarz inequality and [Theorem 2.6](#) with the 1-Lipschitz function $g : \eta \mapsto \frac{1}{2\mathbf{p}} \sum_{i \in F^*} p_0(i)\eta(i)$, we have

$$\begin{aligned} \sum_{k \in F^*} |R_k(t)| &\leq \sum_{k \in F^*} \mathbb{E}_\eta \left(m(\eta_t)(k) \left| \sum_{i \in F^*} p_0(i)m(\eta_t)(i) - \mathbb{E}_\eta \left(\sum_{i \in F^*} p_0(i)m(\eta_t)(i) \right) \right| \right) \\ &\quad + \mathbf{p}(N-1)^{-1} \\ &\leq 2\mathbf{p}N^{-1} \text{Var}_\eta(g(\eta_t))^{\frac{1}{2}} + \mathbf{p}(N-1)^{-1} \\ &\leq \mathbf{p} \sqrt{2\rho^{-1}(1 - e^{-2\rho t})(\mathbf{Q}_1 + \mathbf{p})N^{-1}} + \mathbf{p}(N-1)^{-1}. \end{aligned}$$

If $c_t = \rho^{-1}(1 - e^{-2\rho t})$, $B = \mathbf{Q}_1 + 2\mathbf{p}$, then Gronwall's lemma gives

$$\begin{aligned} \epsilon(t) &\leq \epsilon(0)e^{Bt} + \int_0^t e^{B(t-s)} \left(\frac{\mathbf{p}\sqrt{2B}}{\sqrt{N}} \sqrt{c_s} + \frac{2\mathbf{p}}{N-1} \right) ds \\ &\leq \left(\epsilon(0) + \frac{2\mathbf{p}}{(N-1)B} + \frac{\mathbf{p}\sqrt{2B}}{\sqrt{N}} \int_0^t e^{-Bs} \sqrt{c_s} ds \right) e^{Bt} \\ &\leq \left(\epsilon(0) + \frac{A}{\sqrt{N}} \right) e^{Bt}, \end{aligned}$$

for some $A > 0$. \square

2.3. Proof of the corollaries

In this subsection, we give the proofs of corollaries given in the introduction.

Proof of Corollary 1.4. The proof is based on an approximation of the conditioned semigroups by two particle systems. [Theorem 1.1](#) gives a contraction for these particle systems. We then use [Theorem 1.2](#) and a discretization argument to prove that it implies a contraction for the conditioned semigroups.

Let $(m_0^{(N)})_{N \geq 0}$ and $(\tilde{m}_0^{(N)})_{N \geq 0}$ be two sequences of probability measures that converge to μ and ν respectively, as N tends to infinity, and such that $\eta_0^{(N)} = (Nm_0^{(N)}(k))_{k \in F^*} \in E^{(N)}$ and $\tilde{\eta}_0^{(N)} = (N\tilde{m}_0^{(N)}(k))_{k \in F^*} \in E^{(N)}$, for every $N \geq 0$. The existence of these two sequences can be proved via the law of large numbers. Now, for each $N \geq 0$ and $t \geq 0$, [Theorem 1.1](#) establishes the existence of a coupling between $\eta_t^{(N)}$ and $\tilde{\eta}_t^{(N)}$, where each of its components is generated by (5), with initial condition $(\eta_0^{(N)}, \tilde{\eta}_0^{(N)})$ which satisfies

$$N^{-1} \mathbb{E} \left[d_1(\eta_t^{(N)}, \tilde{\eta}_t^{(N)}) \right] \leq e^{-\rho t} d_{\text{TV}} \left(m_0^{(N)}, \tilde{m}_0^{(N)} \right).$$

Now let us prove that we can take the limit $N \rightarrow +\infty$. Since F is countable and discrete, there exists an increasing sequence of finite sets $(F_n^*)_{n \geq 0}$ such that $F^* = \cup_{n \geq 0} F_n^*$ and

$$d_{\text{TV}}(\mu T_t, \nu T_t) = \frac{1}{2} \sum_{k \in F^*} |\mu T_t \mathbf{1}_{\{k\}} - \nu T_t \mathbf{1}_{\{k\}}| = \lim_{n \rightarrow +\infty} \frac{1}{2} \sum_{k \in F_n^*} |\mu T_t \mathbf{1}_{\{k\}} - \nu T_t \mathbf{1}_{\{k\}}|.$$

The previous bound gives

$$\mathbb{E} \left[\frac{1}{2} \sum_{k \in F_n^*} \left| \frac{\eta_t^{(N)}(k)}{N} - \frac{\tilde{\eta}_t^{(N)}(k)}{N} \right| \right] \leq N^{-1} \mathbb{E} \left[d_1(\eta_t^{(N)}, \tilde{\eta}_t^{(N)}) \right] \leq e^{-\rho t} d_{\text{TV}} \left(m_0^{(N)}, \tilde{m}_0^{(N)} \right).$$

Using [Theorem 1.2](#) and taking the limit $N \rightarrow +\infty$, we find

$$\frac{1}{2} \sum_{k \in F_n^*} |\mu T_t \mathbf{1}_{\{k\}} - \nu T_t \mathbf{1}_{\{k\}}| \leq e^{-\rho t} d_{\text{TV}}(\mu, \nu).$$

Indeed, as we work in discrete space, the convergence in distribution is equivalent to that in total variation distance:

$$\lim_{N \rightarrow +\infty} d_{\text{TV}}(m_0^{(N)}, \mu) = \lim_{N \rightarrow +\infty} d_{\text{TV}}(\tilde{m}_0^{(N)}, \nu) = 0.$$

Furthermore all sequences in the expectations are increasing. Thus, taking the limit $n \rightarrow +\infty$, we obtain (2). Finally, the existence of a QSD can be proved as in the proof of [24, Theorem 1]. More precisely, let μ be any probability measure on F^* . We have, for all $s, t \geq 0$ such that $s \geq t$,

$$d_{\text{TV}}(\mu T_t, \mu T_s) = d_{\text{TV}}(\mu T_t, \mu T_{s-t+t}) = d_{\text{TV}}(\mu T_t, (\mu T_{s-t}) T_t) \leq e^{-\rho t}.$$

Thus $(\mu T_t)_{t \geq 0}$ is a Cauchy sequence for the total variation distance and thus admits a limit ν_{qs} . This measure is then proved to be a QSD by standard arguments; see for instance [25, Proposition 1]. \square

Remark 2.8 (Weaker Assumptions). Conclusion of [Corollary 1.4](#) is also right if $\rho > 0$ and Assumption (A) does not hold. Indeed Assumption (A) is necessary to have the convergence

of the particle system to the conditioned semi-group. But, as the particle system does not explode, from [31, Theorem 1], this convergence is true whatever Assumption (A) holds or not. Nevertheless, we used this proof (and this additional and not so strong assumption) for the sake of completeness.

We can now proceed to the proof of the second corollary.

Proof of Corollary 1.5. The proof is based on an “interpolation” between the bounds obtained in Corollary 1.4 and Theorem 1.2.

Let us fix $t > 0$, $u \in [0, 1]$ and φ a function such that $\|\varphi\|_\infty \leq 1$. By the Markov property, we have

$$\begin{aligned} \mathbb{E}_\eta [|m(\eta_t)(\varphi) - m(\eta)T_t\varphi|] &\leq \mathbb{E}_\eta [|m(\eta_t)(\varphi) - m(\eta_{tu})T_{t(1-u)}\varphi|] \\ &\quad + \mathbb{E}_\eta [|m(\eta_{tu})T_{t(1-u)}\varphi - m(\eta)T_t\varphi|] \\ &\leq \sup_{\|\varphi\|_\infty \leq 1} \mathbb{E}_\eta [\tilde{\mathbb{E}}_{\eta_{tu}} [|m(\tilde{\eta}_{t(1-u)})(\varphi) - m(\eta_{tu})T_{t(1-u)}\varphi|]] \\ &\quad + \mathbb{E}_\eta [d_{\text{TV}}(m(\eta_{tu})T_{t(1-u)}, m(\eta)T_{ut}T_{t(1-u)})], \end{aligned}$$

where $(\tilde{\eta}_t)_{t \geq 0}$ is a Markov process generated by (1) and where, for all $\eta \in E$, we denote by $\tilde{\mathbb{E}}_\eta$ the conditional expectation of $(\tilde{\eta}_t)_{t \geq 0}$ given the event $\{\tilde{\eta}_0 = \eta\}$. On the one hand, by Theorem 1.2, which is a uniform estimate on the initial condition, there exist $B, C > 0$ such that

$$\sup_{\|\varphi\|_\infty \leq 1} \tilde{\mathbb{E}}_{\eta_{tu}} [|m(\tilde{\eta}_{t(1-u)})(\varphi) - m(\eta_{tu})T_{t(1-u)}\varphi|] \leq \frac{Ce^{Bt(1-u)}}{\sqrt{N}}.$$

On the other hand, from Corollary 1.4, we have

$$\mathbb{E}_\eta [d_{\text{TV}}(m(\eta_{tu})T_{t(1-u)}, m(\eta)T_{ut}T_{t(1-u)})] \leq e^{-\rho t(1-u)}.$$

Choosing

$$u = 1 + \frac{1}{t(B + \rho)} \log \left(\frac{BC}{\rho\sqrt{N}} \right),$$

this gives

$$\sup_{\|\varphi\|_\infty \leq 1} \mathbb{E}_\eta [|m(\eta_t)(\varphi) - m(\eta)T_t\varphi|] \leq \frac{B + \rho}{B} \left(\frac{BC}{\rho\sqrt{N}} \right)^{\frac{\rho}{B+\rho}}. \quad \square \quad (20)$$

Remark 2.9 (Weaker Assumptions). In the previous corollary, it is enough to assume that there exist $C > 0$ and $\lambda > 0$ such that

$$\forall t \geq 0, \quad d_{\text{TV}}(\mu T_t, \nu T_t) \leq Ce^{-\lambda t}, \quad (21)$$

to obtain a uniform bound. Some sufficient conditions to obtain (21) are given in [11,14,24]. We can also use a bound of convergence for the Fleming–Viot particle system as in Theorem 1.1.

As an application, a bound such as (4) always holds when F^* is finite. More precisely, the particle system converges, uniformly in time, to the conditioned process; hence, if the initial distribution is the invariant distribution of the particle system (which exists since E is finite), then it converges in law towards the quasi-stationary distribution.

3. Examples

The ideas developed previously are applied to two specific situations: complete graph dynamics and the two point case. A more detailed study of these two examples is available in [9, Section 3] and [9, Section 4].

3.1. Complete graph dynamics

We consider a random walk on the complete graph with sites $\{1, \dots, K\}$, $K \in \mathbb{N}^*$ and constant killing rate $p > 0$. Namely

$$\forall i, j \in \{1, \dots, K\}, \quad Q_{i,j} = \frac{1}{K}, \quad p_0(i) = p.$$

This setting translates to a setting called *neutral evolution* in the population genetics literature [16,33]. More precisely, consider N individuals possessing one type in $F^* = \{1, \dots, K\}$ at time t . Each pair of individuals interacts at rate p . Upon an interacting event, one individual dies and the other one reproduces. In addition, every individual changes its type (mutates) at rate 1 and chooses uniformly at random a new type in F^* . The measure $m(\eta_t)$ gives the proportions of types. The kind of mutation we consider here is often referred to as parent-independent or the house-of-cards model.

Due to the geometry of the complete graph, several explicit formulas can be obtained such as the invariant distribution, the two-particle correlations, ... For instance the invariant distribution is reversible and computed in [19,21,22].

The most interesting fact is that conditions in Theorem 1.1, which seems to be a bit strong, are tight in the complete graph dynamics. In that case, $\lambda = \rho = 1$ and the bound obtained is optimal in terms of contraction. Moreover, the rate that we obtain is exactly the spectral gap. Indeed, on the one hand, let us recall that, as the invariant measure is reversible, the spectral gap λ_1 is the largest constant such that

$$\lim_{t \rightarrow +\infty} e^{2\theta t} \|R_t f - v_N(f)\|_{L^2(v_N)}^2 = 0, \quad (22)$$

for every $\theta < \lambda_1$ and $f \in L^2(v_N)$, where $(R_t)_{t \geq 0}$ is the semi-group generated by \mathcal{L} . See for instance [2,29]. On the other hand, if $\theta < 1$, then by Theorem 1.1, we have

$$\begin{aligned} e^{2\theta t} \|R_t f - v_N(f)\|_{L^2(v_N)}^2 &= e^{2\theta t} \int_E ((\delta_\eta R_t) f - (v_N R_t) f)^2 v_N(d\eta) \\ &\leq 2e^{2\theta t} \|f\|_\infty^2 \int_E \mathcal{W}_{d_1}(\delta_\eta R_t, v_N R_t)^2 v_N(d\eta) \\ &\leq 2e^{2(\theta-1)t} \|f\|_\infty^2 \int_E \mathcal{W}_{d_1}(\delta_\eta, v_N)^2 v_N(d\eta), \end{aligned}$$

and then (22) holds. Now, the constant functions are trivially eigenvectors of \mathcal{L} associated with the eigenvalue 0, and if, for $k \in \{1, \dots, K\}$, $l \geq 1$ we set $\varphi_k^{(l)} : \eta \mapsto \eta(k)^l$, then the function $\varphi_k^{(1)}$ satisfies

$$\mathcal{L}\varphi_k^{(1)} = N/K - \varphi_k^{(1)}.$$

In particular $\varphi_k^{(1)} - N/K$ is an eigenvector and 1 is an eigenvalue of $-\mathcal{L}$. This gives $\lambda_1 \leq 1$ and finally $\lambda_1 = 1$ is the smallest eigenvalue of $-\mathcal{L}$. By the reversibility, we have a Poincaré (or

spectral gap) inequality

$$\forall t \geq 0, \quad \|R_t f - \nu_N(f)\|_{L^2(\nu_N)}^2 \leq e^{-2t} \|f - \nu_N(f)\|_{L^2(\nu_N)}^2.$$

3.2. The two point space example

The second interesting model is the two point space case. We consider a Markov chain defined on the states $\{0, 1, 2\}$ where 0 is the absorbing state. The invariant distribution is explicit since here, the study of the particle system is reduced to the one of $(\eta_t(1))_{t \geq 0}$ which is a birth and death process with quadratic rates. Applying [Theorem 1.1](#), the result is not optimal. Nevertheless, the error does not come from our coupling choice but it comes from how we estimate the distance. Indeed, as can be seen with the proof of [\[7, Theorem 9.25\]](#), one can use our coupling to obtain the spectral gap as rate of convergence. Moreover, even if our main theorem does not give an exponential decay in any case, one can prove that the spectral gap of the Fleming–Viot process is always bounded from below by a positive constant not depending on the number of particles. The proof relies on Hardy’s inequalities type arguments [\[27\]](#) and is detailed in [\[9\]](#).

Acknowledgments

We would like to thank the referees for their careful reading of the manuscript. We would like to also thank Amine Asselah and Djalil Chafaï for valuable discussions. The work of B. Cloez was partially supported by the CIMI (Centre International de Mathématiques et d’Informatique) Excellence program (11-LABX-0040-CIMI) while a postdoctoral scholar, by ANR PIECE (12-JS01-0006) and by ANR MANEGE (09-BLAN-0215). M.-N. Thai was supported by grants from Région Ile-de-France.

References

- [1] A. Asselah, P.A. Ferrari, P. Groisman, Quasistationary distributions and Fleming–Viot processes in finite spaces, *J. Appl. Probab.* 48 (2) (2011) 322–332.
- [2] D. Bakry, L’hypercontractivité et son utilisation en théorie des semigroupes, in: *Lectures on Probability Theory* (Saint-Flour, 1992), in: *Lecture Notes in Math.*, vol. 1581, Springer, Berlin, 1994, pp. 1–114.
- [3] D. Bakry, P. Cattiaux, A. Guillin, Rate of convergence for ergodic continuous Markov processes: Lyapunov versus Poincaré, *J. Funct. Anal.* 254 (3) (2008) 727–759.
- [4] M. Bieniek, K. Burdzy, S. Pal, Extinction of Fleming–Viot-type particle systems with strong drift, *Electron. J. Probab.* 17 (11) (2012) 15.
- [5] K. Burdzy, R. Holyst, P. March, A Fleming–Viot particle representation of the Dirichlet Laplacian, *Comm. Math. Phys.* 214 (3) (2000) 679–703.
- [6] D. Chafaï, A. Joulin, Intertwining and commutation relations for birth–death processes, *Bernoulli* 19 (5A) (2013) 1855–1879.
- [7] M.-F. Chen, *From Markov Chains to Non-equilibrium Particle Systems*, second ed., World Scientific Publishing Co. Inc., River Edge, NJ, 2004.
- [8] B. Cloez, M. Hairer, Exponential ergodicity for Markov processes with random switching. *ArXiv e-prints*, Mar. 2013.
- [9] B. Cloez, M.-N. Thai, Quantitative results for the Fleming–Viot particle system and quasi-stationary distributions in discrete space. Preliminary and unpublished version available on <http://arxiv.org/abs/1312.2444v2>.
- [10] P. Collet, S. Martínez, J. San Martín, *Quasi-stationary Distributions*, in: *Probability and its Applications* (New York), Springer, Heidelberg, 2013, Markov chains, diffusions and dynamical systems.
- [11] J.N. Darroch, E. Seneta, On quasi-stationary distributions in absorbing continuous-time finite Markov chains, *J. Appl. Probab.* 4 (1967) 192–196.
- [12] P. Del Moral, A. Guionnet, On the stability of measure valued processes with applications to filtering, *C. R. Acad. Sci. Paris I* 329 (5) (1999) 429–434.

- [13] P. Del Moral, L. Miclo, A Moran particle system approximation of Feynman–Kac formulae, *Stochastic Process. Appl.* 86 (2) (2000) 193–216.
- [14] P. Del Moral, L. Miclo, On the stability of nonlinear Feynman–Kac semigroups, *Ann. Fac. Sci. Toulouse Math.* (6) 11 (2) (2002) 135–175.
- [15] A. Eberle, Couplings, distances and contractivity for diffusion processes revisited. ArXiv e-prints, May 2013.
- [16] A. Etheridge, Some Mathematical Models from Population Genetics, in: *Lecture Notes in Mathematics*, vol. 2012, Springer, Heidelberg, 2011, Lectures from the 39th probability summer school held in Saint-Flour, 2009.
- [17] P.A. Ferrari, H. Kesten, S. Martinez, P. Picco, Existence of quasi-stationary distributions. A renewal dynamical approach, *Ann. Probab.* 23 (2) (1995) 501–521.
- [18] P.A. Ferrari, N. Marić, Quasi stationary distributions and Fleming–Viot processes in countable spaces, *Electron. J. Probab.* 12 (24) (2007) 684–702.
- [19] C. Giardinà, F. Redig, K. Vafayi, Correlation inequalities for interacting particle systems with duality, *J. Stat. Phys.* 141 (2) (2010) 242–263.
- [20] P. Groisman, M. Jonckheere, Simulation of quasi-stationary distributions on countable spaces. ArXiv e-prints, June 2012.
- [21] S. Grosskinsky, F. Redig, K. Vafayi, Condensation in the inclusion process and related models, *J. Stat. Phys.* 142 (5) (2011) 952–974.
- [22] S. Grosskinsky, F. Redig, K. Vafayi, Dynamics of condensation in the symmetric inclusion process, *Electron. J. Probab.* 18 (66) (2013) 23.
- [23] M. Hairer, J.C. Mattingly, Yet another look at Harris’ ergodic theorem for Markov chains, in: *Seminar on Stochastic Analysis, Random Fields and Applications VI*, in: *Progr. Probab.*, vol. 63, Birkhäuser/Springer Basel AG, Basel, 2011, pp. 109–117.
- [24] S. Martinez, J. San Martin, D. Villemonais, Existence and uniqueness of a quasi-stationary distribution for Markov processes with fast return from infinity. ArXiv e-prints, Feb. 2012.
- [25] S. Méléard, D. Villemonais, Quasi-stationary distributions and population processes, *Probab. Surv.* 9 (2012) 340–410.
- [26] S.P. Meyn, R.L. Tweedie, Stability of Markovian processes. III. Foster–Lyapunov criteria for continuous-time processes, *Adv. Appl. Probab.* 25 (3) (1993) 518–548.
- [27] L. Miclo, An example of application of discrete Hardy’s inequalities, *Markov Process. Related Fields* 5 (3) (1999) 319–330.
- [28] M. Rousset, On the control of an interacting particle estimation of Schrödinger ground states, *SIAM J. Math. Anal.* 38 (3) (2006) 824–844.
- [29] L. Saloff-Coste, Lectures on finite Markov chains, in: *Lectures on Probability Theory and Statistics (Saint-Flour, 1996)*, in: *Lecture Notes in Math.*, vol. 1665, Springer, Berlin, 1997, pp. 301–413.
- [30] A.-S. Sznitman, Topics in propagation of chaos, in: *École d’Été de Probabilités de Saint-Flour XIX—1989*, in: *Lecture Notes in Math.*, vol. 1464, Springer, Berlin, 1991, pp. 165–251.
- [31] D. Villemonais, General approximation method for the distribution of Markov processes conditioned not to be killed. ArXiv e-prints, June 2011.
- [32] D. Villemonais, Interacting particle systems and Yaglom limit approximation of diffusions with unbounded drift, *Electron. J. Probab.* 16 (61) (2011) 1663–1692.
- [33] G.A. Watterson, Reversibility and the age of an allele. I. Moran’s infinitely many neutral alleles model, *Theor. Popul. Biol.* 10 (3) (1976) 239–253.