QUANTITATIVE RESULTS FOR THE FLEMING-VIOT PARTICLE SYSTEM IN DISCRETE SPACE

BERTRAND CLOEZ AND MARIE-NOÉMIE THAI

ABSTRACT. We show, for a class of discrete Fleming-Viot type particle systems, that the convergence to the equilibrium is exponential for a suitable Wassertein 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 the two point space.

AMS 2000 Mathematical Subject Classification: 60K35, 60B10, 37A25.

Keywords: Fleming-Viot process - quasi-stationary distributions - coupling - Wasserstein distance - chaos propagation - commutation relation.

. CONTENTS

1. Introduction 1
Long time behavior 3
Propagation of chaos 4
Two main consequences 5
2. Proof of the main theorems 6
2.1. Proof of Theorem 1.1 6
2.2. Proof of Theorem 1.2 9
2.3. Proof of the corollaries 12
3. Complete graph dynamics 13
3.1. The associated killed process 14
3.2. Correlations at fixed time 14
3.3. Properties of the invariant measure 15
3.4. Long time behavior and spectral analysis of the generator 17
4. The two point space 18
4.1. The associated killed process 19
4.2. Explicit formula of the invariant distribution 19
4.3. Rate of convergence 20
4.4. Correlations 21
References 22

1. INTRODUCTION

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 $F = F^* \cup \{0\}$ 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 $f : F \mapsto \mathbb{R}$, gives

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

Date: December 9, 2013.
Partially supported by ANR-11-LABX-0040-CIMI within the program ANR-11-IDEX-0002-02.

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 function 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

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

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 not well understood, nor easily amenable to simulation. To avoid these difficulties, Burdzy, Holyst, Ingerman, March [5], and Del Moral, Guionnet, Miclo [10, 11] 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

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

Applying its generator to a bounded function f gives

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

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

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

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

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

We also set $m(\eta)(k) = m(\eta)(\{k\})$. The aim of this work is to quantify (if they hold) the following limits:

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

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 develop, in detail, the study of two examples. The first one concerns a random walk on the complete graph with sites $\{0, 1, \ldots, K\}$, killed when reaching 0. Namely

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

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. One interesting point of the complete graph approach is that it permits to reduce the difficulties of the Fleming-Viot to the interaction. Due to its simple geometry, several explicit formulas are obtained such as the invariant distribution, the correlations and the spectral gap. It seems to be new in the context of Fleming-Viot particle systems. The second example is the case where F^* contains only two elements. Its study is reduced to the study of a birth-death process with quadratic rates. For this example, the only trivial limit to quantify is the limit (d). 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 [14] by:

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

In particular, if there exists $j \in F^*$ such that for every $i \neq j$, $Q_{i,j} > c > 0$ then $\lambda \geq c$, for some 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: $d_1(\eta, \eta') = N d_{\mathrm{TV}}(m(\eta), m(\eta'))$. Indeed, recall that, for every two probability measures μ and μ', the total variation distance is given by

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

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

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

where the infimum runs again over all the couples of random variables with marginal laws μ and μ'. Along this paper, $\mathcal{L}(X)$ design the law of the random variable X. Along the paper, we assume that

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

Our first main result is:

Theorem 1.1 (Wasserstein exponential ergodicity)**.** *If $\rho = \lambda - (\max(p_0) - \min(p_0))$ then for any processes $(\eta_t)_{t>0}$ and $(\eta'_t)_{t>0}$ 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 verifying 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. When the death rate p_0 is constant, this 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 complex problem. See for instance Section 4, where we study the case where F^* contains only two elements. Even though in this case, the study seems to be easy we were not able to give a closed formula for the spectral gap. Also, note that the previous inequality is a contraction, 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. Two tagged particles in a large population of interacting ones behave in an almost independent way under some assumptions; see [24]. 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:

- (A_1) $\mathbf{Q_1} = \sup_{i\in F^*} \sum_{j\in F^*} Q_{i,j} < +\infty$ and $\mathbf{p} = \sup_{i\in F^*} p_0(i) < +\infty$;
- (A_2) $\mathbf{Q_2} = \sup_{j\in F^*} \sum_{i\in F^*} Q_{i,j} < +\infty$.

Under these (boundedness) assumptions, the particle system converges to the conditioned semigroup. 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 Assumptions (A_1) and (A_2), there exists an explicit constant $C > 0$ such that, for all $\eta \in E$, $k \in F^*$, $T \geq 0$ and any probability measure μ, we have*

$$\sup_{t\in[0,T]} \mathbb{E}_\eta \left[|m(\eta_t)(k) - \mu T_t(k)|\right] \leq e^{CT} \left(\frac{1}{\sqrt{N}} + d_{TV}(m(\eta), \mu) \right),$$

and when F is finite, we have

$$\sup_{t\in[0,T]} \mathbb{E}_\eta[d_{TV}(m(\eta_t), \mu T_t)] \leq e^{CT} |F^*| \left(\frac{1}{\sqrt{N}} + d_{TV}(m(\eta), \mu) \right).$$

The proof is based on an estimation of the correlation and on an argument of supersolution inspired by [16]. More precisely our correlation estimate is given by:

Theorem 1.3 (Covariance estimates)**.** *Under Assumption (A_1), there exists an explicit constant D such that, for all $k, l \in F^*$, $\eta \in E$ and $t \geq 0$ we have*

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

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.
The previous theorem is a consequence of Theorem 2.5 which gives the correlation of more general functional of η. The proof of these results 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, 14].

Theorem 1.2 is a generalization of [1, Theorem 1.3], [14, Theorem 1.2] and a slight generalization of [16, Theorem 2.2]. Indeed, $\sum_{k\in F^*} p_0(k) < +\infty$ is not necessary. We can also cite [11, Theorem 1.1] and [25, 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. Finally, we can improve the previous bound in the special case of the complete graph random walk but, in general, we do not know how improve it even when $\mathrm{card}(F^*) = 2$; see Sections 3 and 4.

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 $N \to +\infty$ 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 Assumptions (A_1) and (A_2) hold. For any probability measure μ, ν, we have*

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

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

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

This corollary is closely related to several previous work [9], [12, Theorem 1.1], [20, Theorem 3] and [14, Theorem 1.1]. When F is finite, the oldest result dates from 1967 [9] 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 [14, Theorem 1.1] in a stronger convergence and with an estimation of the rate of convergence. As in [12, 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 $(\max(p_0) - \min(p_0))$ is small enough. In [12, Theorem 1.1] they only need that $\max(p_0) < +\infty$ but, their mixing condition is stronger than ours. Finally [20, 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 Assumptions (A_1) and (A_2) are not necessary; see Remark 2.7.
Our second corollary gives a uniform bound for the limit (d):

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 $\eta \in E$,*

$$\sup_{t\geq 0} \mathbb{E}_\eta\left[|m(\eta_t)(k) - m(\eta)T_t(k)|\right] \leq \frac{K_0}{N^\gamma},$$

for every $k \in F^*$. Furthermore, if F^* is finite then there exists $K_1 > 0$ such that

$$\sup_{t\geq 0} \mathbb{E}_\eta[d_{TV}(m(\eta_t), m(\eta)T_t)] \leq \frac{K_1}{N^\gamma}.$$

All constants are explicit.

In particular, if η is distributed according to the measure ν_N, then under the assumptions of the previous corollary, there exist $K_0, K_1 > 0$ and $\gamma > 0$ such that

$$\mathbb{E}\left[|m(\eta)(k) - \nu_{\mathrm{qs}}(k)|\right] \leq \frac{K_0}{N^\gamma},$$

for every $k \in F^*$. Moreover, if F^* is finite then

$$\mathbb{E}[d_{\mathrm{TV}}(m(\eta), \nu_{\mathrm{qs}})] \leq \frac{K_1}{N^\gamma}.$$

Without rate of convergence, this limiting result was proved in [1, Theorem 2] when F is finite. Whereas, here, a rate of convergence, which is not of the right order (since $\gamma \ll 1/2$) 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 2.8.

The remainder of the paper is as follows. Section 2 gives the proofs of our main theorems; Subsection 2.1 contains the proof of Theorem 1.1, Subsection 2.2 the proof of Theorem 1.2 and the last subsection the proof of the corollaries. We conclude the paper with Sections 3 and 4, 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 $f : E \to \mathbb{R}$ and $\eta \in E$, is given by

$$\mathcal{L}f(\eta) = \sum_{i \in F^*} \eta(i) \sum_{j \in F^*} \left(Q_{i,j} + p_0(i) \frac{\eta(j)}{N-1} \right) (f(T_{i \to j}\eta) - f(\eta)) \,. \tag{3}$$

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 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 $\min(p_0)$ 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. An example of explosive Fleming-Viot particle system can be found in* [4]*. However, the assumption that p_0 is bounded is 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. This couple is Markovian and we describe it by expressing its generator $\mathbb{L}$; for every bounded function f and $\eta, \eta' \in E$, it is given by

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

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. The jumps rate A_p, corresponds to the redistribution dynamics and depends only on p_0. We will give the expression of A_Q and A_p by describing the jump dynamics of two selected particles of each configuration.

- If these two particles are in the same site ($i = i'$), we can couple them as follows

– either they jump in a same site ($j = j'$) with rates

$$A_Q(i,i,j,j) = \big(\eta(i) \wedge \eta'(i)\big)\, Q_{i,j}, \quad \text{and} \quad A_p(i,i,j,j) = p_0(i)\,\big(\eta(i) \wedge \eta'(i)\big)\,\frac{\eta(j) \wedge \eta'(j)}{N-1};$$

– or they jump in different sites ($j \neq j'$) with rate

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

- If these two particles are in different sites ($i \neq i'$)
  – they jump to the same site ($j = j'$) with rates

$$A_Q(i,i',j,j) = \frac{(\eta(i) - \eta'(i))_+ \times (\eta'(i') - \eta(i'))_+}{d_1(\eta,\eta')}\,\big(Q_{i,j} \wedge Q_{i',j}\big)\,\mathbf{1}_{j\neq i,i'},$$

and

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

  – they jump to different sites ($j \neq j'$) with rate

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

  – a particle jumps to another site while the other one does not jump; namely
    * if $i' = j'$ then the rates are given by

$$A_Q(i,j,i',i') = \frac{(\eta(i) - \eta'(i))_+ \times (\eta'(i') - \eta(i'))_+}{d_1(\eta,\eta')}(Q_{i,j} - Q_{i',j})_+\mathbf{1}_{j\neq i,i'}$$

and

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

    * similarly, if $i = j$ then the rates are given by

$$A_Q(i,i,i',j') = \frac{(\eta(i) - \eta'(i))_+ \times (\eta'(i') - \eta(i'))_+}{d_1(\eta,\eta')}(Q_{i',j'} - Q_{i,j'})_+\mathbf{1}_{j'\neq i,i'}$$

and

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

  – finally, a particle jumps to the site of the second one,
    * if $j = i' = j'$, with rate

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

    * if $j = i'$ and $i = j'$ with rate

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

We also set, for every measurable function f,

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

and

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

We have

$$\begin{aligned}\mathbb{L}_p d_1(\eta,\eta') &\le \sum_{i\in F^*} p_0(i)\left(\eta(i)\wedge\eta'(i)\right)\frac{d_1(\eta,\eta')}{N-1}\\ &\quad - \sum_{i,i'\in F^*}\left(p_0(i)\wedge p_0(i')\right)\frac{(\eta(i)-\eta'(i))_+(\eta'(i')-\eta(i'))_+}{d_1(\eta,\eta')}\sum_{j\in F^*}\frac{\eta(j)\wedge\eta'(j)}{N-1}\\ &\le (\max(p_0)-\min(p_0))d_1(\eta,\eta'),\end{aligned}$$

and

$$\mathbb{L}_Q d_1(\eta,\eta') \le -\lambda d_1(\eta,\eta').$$

We deduce that $\mathbb{L}d_1(\eta,\eta') \le -\rho d_1(\eta,\eta')$. Now let $(\mathbb{P}_t)_{t\ge 0}$ 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 $t\ge 0$, $\mathbb{P}_t d_1 \le e^{-\rho t}d_1$; namely

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

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]. □

Corollary 2.3 (Coalescent time estimate)**.** *For all $t\ge 0$, we have*

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

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

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

Proof. Using Theorem 1.1, we find

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

□

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 founded. The first one is to use another more appropriate distance. This technique is in general useful in other (Markovian) contexts* [6, 8, 13]*. Another way is to find a contraction after a certain time: it's the Lyapunov-type techniques* [3, 19, 22]*. Theses 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*

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

and all conclusions hold.

2.2. **Proof of Theorem 1.2.** 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

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

and

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

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

$$S_t f(\eta) = \mathbb{E}\left[f(\eta_t) \mid \eta_0 = \eta\right],$$

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

$$\begin{aligned}\Gamma f(\eta) &= \mathcal{L}(f^2)(\eta) - 2f(\eta)\mathcal{L}f(\eta) \\ &= \sum_{i,j\in F^*} \eta(i)\left(Q_{i,j} + p_0(i)\frac{\eta(j)}{N-1}\right)\left(f(T_{i\to j}\eta) - f(\eta)\right)^2.\end{aligned} \tag{4}$$

We present now an improvement of Theorem 1.3.

Theorem 2.5 (Correlation for Lipschitz functional)**.** *Let g, h be two $1-$Lipschitz mappings on (E, d_1); namely*

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

for every $\eta, \eta' \in E$. Under Assumption (A_1) we have for all $t \geq 0$ and $\eta \in E$,

$$|Cov_\eta(g(\eta_t), h(\eta_t))| \leq \frac{1 - e^{-2\rho t}}{2\rho}\left(N\mathbf{Q_1} + \mathbf{p}\frac{N^2}{N-1}\right),$$

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

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

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

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

and so,

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

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

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

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

$$\begin{aligned} \mid S_{t-s}g(T_{i\to j}\eta) - S_{t-s}g(\eta) \mid &\leq \mathcal{W}_{d_1}(\mathcal{L}(\eta_{t-s}), \mathcal{L}(\eta'_{t-s})) \\ &\leq e^{-\rho(t-s)} d_1(T_{i\to j}\eta, \eta) \\ &\leq e^{-\rho(t-s)} \mathbf{1}_{i\neq j}. \end{aligned} \tag{5}$$

Hence,

$$\|\Gamma S_{t-s}g\|_\infty = \sup_{\eta\in E} \mid \Gamma S_{t-s}g(\eta) \mid \leq e^{-2\rho(t-s)} \left(N\mathbf{Q_1} + \mathbf{p}\frac{N^2}{N-1} \right).$$

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

$$\begin{aligned} \mathrm{Cov}_\eta(g(\eta_t), h(\eta_t)) &\leq \mathrm{Var}_\eta(g(\eta_t))^{1/2} \mathrm{Var}_\eta(h(\eta_t))^{1/2} \\ &\leq \frac{1-e^{-2\rho t}}{2\rho} \left(N\mathbf{Q_1} + \mathbf{p}\frac{N^2}{N-1} \right). \end{aligned}$$

□

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 . □

Remark 2.6 (Generalization)**.** *Assume that 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*

$$\mathcal{W}_{d_1}(\mathcal{L}(\eta_t), \mathcal{L}(\eta'_t)) \leq Ce^{-\lambda t} \mathcal{W}_{d_1}(\mathcal{L}(\eta_0), \mathcal{L}(\eta'_0)), \tag{6}$$

then under the previous assumptions we have, for all $t \geq 0$,

$$Cov_\eta(\eta_t(k)/N, \eta_t(l)/N) \leq \frac{2C}{N^2} \frac{1-e^{-2\lambda t}}{\lambda} \left(N\mathbf{Q_1} + \mathbf{p}\frac{N^2}{N-1} \right).$$

A bound like (6) *is proved in the two point space case.*

Proof of Theorem 1.2. The proof is based on a bias-variance type decomposition. The variance is bounded via Theorem 1.3 and the bias via a Gronwall-type argument. More precisely, for $t \in [0, T]$ and $k \in F^*$, we have

$$\mathbb{E}_\eta \left[|m(\eta_t)(k) - \mu T_t(k)| \right] \leq \mathbb{E}_\eta \left[|m(\eta_t)(k) - \mathbb{E}_\eta[m(\eta_t)(k)]| \right] + |\mathbb{E}_\eta \left[m(\eta_t)(k) \right] - \mu T_t(k)|. \tag{7}$$

Theorem 1.3 and the Cauchy-Schwarz inequality give

$$\mathbb{E}_\eta \left[|m(\eta_t)(k) - \mathbb{E}_\eta(m(\eta_t)(k))| \right] \leq \sqrt{\mathrm{Var}_\eta(m(\eta_t)(k))} \leq \sqrt{c_t(N-1)^{-1}},$$

where $c_t = 2\rho^{-1}(1-e^{-2\rho t})(\mathbf{Q_1} + \mathbf{p})$. Now, to study the bias term in (7), we follow the proof of [16, Proposition 2.6]. So let us introduce the following notations

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

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

$$\forall t \geq 0,\ \partial_t \mu T_t(j) = \sum_{i\in F^*} \left(Q_{i,j}\ \mu T_t(i) + \ p_0(i)\ \mu T_t(i)\ \mu T_t(j) \right). \tag{8}$$

Thus

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

and using (1) and $\partial_t u_t = \mathcal{L}^* u_t$, 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

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

Let us prove that

$$|R_k(t)| \leq \mathbf{p}(c_t+1)(N-1)^{-1}. \tag{9}$$

We have

$$\begin{aligned}|R_k(t)| &= \left|\mathrm{Cov}_\eta\left(\sum_{i\in F^*} p_0(i)m(\eta_t)(i), m(\eta_t)(k)\right) + (N-1)^{-1}\sum_{i\in F^*} p_0(i)\mathbb{E}_\eta[m(\eta_t)(i)m(\eta_t)(k)]\right| \\ &\leq 4\mathbf{p}N^{-2}|\mathrm{Cov}_\eta(g(\eta_t),h(\eta_t))| + \mathbf{p}(N-1)^{-1},\end{aligned}$$

where g and h are the $1-$Lipschitz functions defined by

$$g:\eta\mapsto\frac{1}{2\mathbf{p}}\sum_{i\in F^*}p_0(i)\eta(i) \text{ and } h:\eta\mapsto\frac{\eta(k)}{2}.$$

Hence the inequality (9) comes from Theorem 2.5.
Now, for $k\in F^*$ and $t\geq 0$, we set $\delta_k(t)=u_k(t)-v_k(t)$, $\delta(t)=(\delta_k(t))_{k\in F^*}$, $\overline{\delta_k}(t)=Ae^{Bt}$ and $\overline{\delta}(t)=(\overline{\delta_k}(t))_{k\in F^*}$. At this stage, we do not fix the value of A and B, but we allow ourselves the freedom to tune it at the end of the proof. We have $\partial_t\delta_k(t)=\mathcal{T}_+(\delta(t))$, where

$$\mathcal{T}_+(a)=\sum_{i\in F^*}Q_{i,k}a_i+\sum_{i\in F^*}p_0(i)v_i(t)a_k+\sum_{i\in F^*}p_0(i)u_k(t)a_i-\frac{p_0(k)}{N-1}u_k(t)+R_k(t), \tag{10}$$

for any sequence $a=(a_k)_{k\in F^*}$. Now, from (9), we have

$$\begin{aligned}\mathcal{T}_+(\overline{\delta}(t)) &= Ae^{Bt}\left[\sum_{i\in F^*}Q_{i,k}+\sum_{i\in F^*}p_0(i)v_i(t)+\sum_{i\in F^*}p_0(i)u_k(t)\right]-\frac{p_0(k)}{N-1}u_k(t)+R_k(t) \\ &\leq Ae^{Bt}\left[\sum_{i\in F^*}Q_{i,k}+\sum_{i\in F^*}p_0(i)v_i(t)+\sum_{i\in F^*}p_0(i)u_k(t)\right]+\frac{p_0(k)}{N-1}u_k(t)+R_k(t) \\ &\leq Ae^{Bt}\left[\mathbf{Q_2}+2\mathbf{p}\right]+\frac{\mathbf{p}(c_t+2)}{N-1}.\end{aligned}$$

We find that

$$\forall k\in F^*,\ \mathcal{T}_+(\overline{\delta}(t))\leq\partial_t\overline{\delta_k}(t)=ABe^{Bt},$$

if $A\geq\dfrac{1}{N-1}$ and

$$B=\begin{cases}\mathbf{Q_2}+4\mathbf{p}+2(-\rho^{-1})(\mathbf{Q_1}+\mathbf{p}) & \text{if } \rho<0\\ \mathbf{Q_2}+4\mathbf{p}+2\rho^{-1}(\mathbf{Q_1}+\mathbf{p}) & \text{if } \rho>0\\ \mathbf{Q_2}+4\mathbf{p}+4(\mathbf{Q_1}+\mathbf{p}) & \text{if } \rho=0.\end{cases} \tag{11}$$

Now, if $\overline{\delta}_k(0)\geq\delta_k(0)$ (namely $A\geq\sup_{k\in F^*}|m(\eta_0)(k)-\mu(k)|$), a classical argument of super-solution gives $\overline{\delta}_k(t)\geq\delta_k(t)$, for every $t\geq 0$ and $k\in F^*$; see [16, Lemma 2.7] for details. In the same way, setting for any configuration $a=(a_k)_{k\in F^*}$

$$\mathcal{T}_-(a)=\sum_{i\in F^*}Q_{i,k}a_i+\sum_{i\in F^*}p_0(i)v_i(t)a_k+\sum_{i\in F^*}p_0(i)u_k(t)a_i+\frac{p_0(k)}{N-1}u_k(t)-R_k(t), \tag{12}$$

the same argument applied to $\overline{\delta}+\delta_k$ gives

$$\delta_k(t)\leq Ae^{Bt} \text{ and } \delta_k(t)\geq -Ae^{Bt},\ \forall t\geq 0,\forall k\in F^*,$$

which ends the proof. □

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 $(\widetilde{m}_0^{(N)})_{N\geq 0}$ 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 $\widetilde{\eta}_0^{(N)} = (N\widetilde{m}_0^{(N)}(k))_{k\in F^*} \in E^{(N)}$, 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 a coupling between $\eta_t^{(N)}$ and $\widetilde{\eta}_t^{(N)}$, where each of its components is generated by (3), with initial condition $(\eta_0^{(N)}, \widetilde{\eta}_0^{(N)})$) which satisfies

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

Now let us prove that we can take the limit $N \to +\infty$. 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

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

The previous bound gives

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

Using Theorem 1.2 and taking the limit $N \to +\infty$, we find

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

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

$$\lim_{N\to+\infty} d_{\mathrm{TV}}(m_0^{(N)}, \mu) = \lim_{N\to+\infty} d_{\mathrm{TV}}(\widetilde{m}_0^{(N)}, \nu) = 0.$$

Thus, taking the limit $n \to +\infty$, we obtain (2). Finally, the existence of a QSD can be proved as in the proof of [20, Theorem 1]. More precisely, let μ be any probability measure on F^*. We have, for all $s, t \geq 0$ such that $s \geq t$,

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

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 [21, Proposition 1]. □

Remark 2.7 (Weaker assumptions)**.** *Assumptions* (A_1) *and* (A_2) *are not necessary (and even useless) in the previous corollary. Indeed, we can use* [25, Theorem 1] *and a similar argument of approximation. However, we used this proof for 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.

We only give the proof for the first inequality; the second one follows by the same argument. Let us fix $t > 0$ and $u \in [0, 1]$. By the Markov property, we have

$$\begin{aligned}\mathbb{E}_\eta[|m(\eta_t)(k) - m(\eta)T_t(k)|] &\leq \mathbb{E}_\eta\left[|m(\eta_t)(k) - m(\eta_{tu})T_{t(1-u)}(k)|\right] \\ &\quad + \mathbb{E}_\eta\left[|m(\eta_{tu})T_{t(1-u)}(k) - m(\eta)T_t(k)|\right] \\ &\leq \mathbb{E}_\eta\left[\widetilde{\mathbb{E}}_{\eta_{tu}}\left[|m(\widetilde{\eta}_{t(1-u)})(k) - m(\eta_{tu})T_{t(1-u)}(k)|\right]\right] \\ &\quad + \mathbb{E}_\eta\left[d_{\mathrm{TV}}(m(\eta_{tu})T_{t(1-u)}, m(\eta)T_{ut}T_{t(1-u)})\right],\end{aligned}$$

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

$$\tilde{\mathbb{E}}_{\eta_{tu}}\left[|m(\widetilde{\eta}_{t(1-u)})(k) - m(\eta_{tu})T_{t(1-u)}(k)|\right] \leq \frac{e^{Bt(1-u)}}{\sqrt{N-1}},$$

where B is defined in (11). On the other hand, from Corollary 1.4, we have

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

Choosing

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

this gives

$$\mathbb{E}_\eta\left[|m(\eta_t)(k) - m(\eta)T_t(k)|\right] \leq \frac{B+\rho}{B}\left(\frac{B}{\rho\sqrt{N-1}}\right)^{\frac{\rho}{B+\rho}}.$$

When F is finite, the same arguments give

$$\mathbb{E}_\eta[d_{TV}(m(\eta_t)(k) - m(\eta)T_t(k))] \leq \frac{B+\rho}{B}\left(\frac{B|F^*|}{2\rho\sqrt{N-1}}\right)^{\frac{\rho}{B+\rho}}.$$

□

Remark 2.8 (Weaker assumptions)**.** *We can weaken the assumption $\rho > 0$ in the previous corollary. Indeed, it is enough to assume that there exist $C > 0$ and $\lambda > 0$ such that*

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

Some sufficient conditions are given in [9, 12, 20]*. We can also use a bound of convergence for the Fleming-Viot particle system as in Theorem 1.1. In particular, when F^* is finite, the particle system converges, uniformly in time, to the conditioned process; hence, if η is distributed by the invariant distribution of the particle system (it exists since E is finite) then it converges in law towards the quasi-stationary distribution.*

3. Complete graph dynamics

In all this section, we study the example of a random walk on the complete graph. Let us fix $K \in \mathbb{N}^*$ and $N \in \mathbb{N}^*$, the dynamics of this example is as follows: we consider a model with N particles which move on the $K+1$ vertices $0, 1, \ldots, K$, of a complete graph uniformly at random. When a particle reaches the node 0, it jumps instantaneously over another particle chosen uniformly at random. This particle system corresponds to the model previously cited with parameters

$$Q_{i,j} = p_0(i) = \frac{1}{K}, \quad \forall i, j \in F^* = \{1, \ldots, K\}.$$

The generator of the associated Fleming-Viot process is then given by

$$\mathcal{L}f(\eta) = \sum_{i=1}^{K}\eta(i)\left[\sum_{j=1}^{K}(f(T_{i\to j}\eta) - f(\eta))\left(\frac{1}{K} + \frac{1}{K}\frac{\eta(j)}{N-1}\right)\right], \tag{13}$$

for every function f and $\eta \in E$. A process generated by (13) is an instance of inclusion processes studied in [15, 17, 18]. It is then related to models of heat conduction and has applications in population genetics. One main point of [15, 17] is a criterion ensuring the existence and reversibility of an invariant distribution for the inclusion processes. In particular, they give an explicit formula of the invariant distribution of a process generated by (13) and we introduce its expression in Subsection 3.3. They also study different scaling limits which seem to be irrelevant for our problems.
In all this section, for any probability measure μ on E, we set in a classical manner $\mathbb{E}_\mu[\cdot] = \int_{F^*} \mathbb{E}_x[\cdot]\mu(dx)$ and $\mathbb{P}_\mu = \mathbb{E}_\mu[\mathbf{1}_\cdot]$; similarly Cov_μ and Var_μ are defined with respect to $\mathbb{E}_\mu$.

3.1. **The associated killed process.** We define the process $(X_t)_{t\geq 0}$ by setting

$$X_t = \begin{cases} Z_t & \text{if } t < \tau \\ 0 & \text{if } t \geq \tau, \end{cases}$$

where τ is an exponential variable with mean K and $(Z_t)_{t\geq 0}$ is the classical complete graph random walk (i.e. without extinction) on $\{1, \dots, K\}$. We have, for any bounded function f,

$$T_t f(x) = \mathbb{E}\left[f(X_t) \mid X_0 = x, X_t \neq 0\right], \quad t \geq 0, x \in F^*.$$

The conditional distribution of X_t is simply given by the distribution of Z_t :

$$\mathbb{P}(X_t = i \mid X_t \neq 0) = \mathbb{P}(Z_t = i).$$

The study of $(Z_t)_{t\geq 0}$ is trivial. Indeed, it converges exponentially fast to the uniform distribution π_K on $\{1, \dots, K\}$. We deduce that for all $t \geq 0$ and all initial distribution μ,

$$d_{\mathrm{TV}}(\mu T_t, \pi_K) = \sum_{i=1}^{K} |\mathbb{P}_\mu(X_t = i \mid \tau > t) - \pi_K(i)| \leq e^{-t}.$$

Thus in this case, the conditional distribution of X converges exponentially fast to the Yaglom limit π_K.

3.2. **Correlations at fixed time.** The special form of $\mathcal{L}$, defined at (13), makes the calculation of the two-particle correlations at fixed time easy.

Theorem 3.1 (Two-particle correlations)**.** *For all $k, l \in \{1, \dots, K\}$, $k \neq l$ and any probability measure μ on E, we have for all $t \geq 0$*

$$\begin{aligned} \mathit{Cov}_\mu(\eta_t(k), \eta_t(l)) = {} & \mathbb{E}_\mu\left[\eta_0(k)\eta_0(l)\right] e^{-\frac{2(KN-K+1)}{K(N-1)}t} \\ & + \frac{-NK - 2N + K}{K(KN - K + 2)}(\mathbb{E}_\mu\left[\eta_0(k)\right] + \mathbb{E}_\mu\left[\eta_0(l)\right])e^{-t} \\ & - \mathbb{E}_\mu\left[\eta_0(k)\right]\mathbb{E}_\mu\left[\eta_0(l)\right] e^{-2t} + \frac{-N^2K + NK - N^2}{K^2(KN - K + 1)}. \end{aligned}$$

Remark 3.2 (Limit $t \to +\infty$)**.** *By the previous theorem, we find for any probability measure μ*

$$\lim_{t\to+\infty} \mathit{Cov}_\mu(\eta_t(k), \eta_t(l)) = \frac{-N^2K + NK - N^2}{K^2(KN - K + 1)} = \mathit{Cov}(\eta(k), \eta(l)),$$

where η is distributed according to the invariant distribution; it exists since the state space is finite, see the next section.

Remark 3.3 (Limit $N \to +\infty$)**.** *For all $k, l \in \{1, \dots, K\}$, $k \neq l$ and any probability measure μ, if $\mathit{Cov}_\mu\left(\eta_0(k), \eta_0(l)\right) \neq 0$ then we have*

$$\mathit{Cov}_\mu\left(\frac{\eta_t(k)}{N}, \frac{\eta_t(l)}{N}\right) \sim_N e^{-2t}\mathit{Cov}_\mu\left(\frac{\eta_0(k)}{N}, \frac{\eta_0(l)}{N}\right),$$

where $u_N \sim_N v_N$ iff $\lim_{N\to+\infty} \dfrac{u_N}{v_N} = 1$.

Proof of Theorem 3.1. For $k,l \in \{1,..,K\}$, let $\psi_{k,l}$ be the function $\eta \mapsto \eta(k)\eta(l)$. Applying the generator (13) to $\psi_{k,l}$ we obtain

$$\mathcal{L}\psi_{k,l}(\eta) = -\frac{2(KN-K+1)}{K(N-1)}\eta(k)\eta(l) + \frac{N-1}{K}(\eta(k)+\eta(l)).$$

So, for all $t \geq 0$,

$$\mathcal{L}\psi_{k,l}(\eta_t) = -\frac{2(KN-K+1)}{K(N-1)}\eta_t(k)\eta_t(l) + \frac{N-1}{K}(\eta_t(k)+\eta_t(l)).$$

Using Kolmogorov's equation, we have

$$\partial_t \mathbb{E}_\mu(\eta_t(k)\eta_t(l)) = -\frac{2(KN-K+1)}{K(N-1)}\mathbb{E}_\mu(\eta_t(k)\eta_t(l)) + \frac{N-1}{K}(\mathbb{E}_\mu(\eta_t(k)) + \mathbb{E}_\mu(\eta_t(l))). \quad (14)$$

Now if $\varphi_k(\eta) = \eta(k)$ then $\mathcal{L}\varphi_k(\eta) = \dfrac{N}{K} - \eta(k)$. We deduce that, for every $t \geq 0$,

$$\partial_t \mathbb{E}_\mu(\eta_t(k)) = \frac{N}{K} - \mathbb{E}_\mu(\eta_t(k)) \text{ and } \mathbb{E}_\mu(\eta_t(k)) = \mathbb{E}_\mu(\eta_0(k))e^{-t} + \frac{N}{K}.$$

Solving equation (14) ends the proof. □

3.3. **Properties of the invariant measure.** As $(\eta_t)_{t\geq 0}$ is an irreducible Markov chain on a finite state space, it is straightforward that it admits a unique invariant measure. In fact, this invariant distribution is reversible and we know its expression. In what follows, we shall give an explicit expression of this measure and some properties.

Theorem 3.4 (Invariant distribution)**.** *The process $(\eta_t)_{t\geq 0}$ admits a unique invariant and reversible measure ν_N, which is defined, for every $\eta \in E$, by*

$$\nu_N(\{\eta\}) = Z^{-1}\prod_{i=1}^{K}\binom{N-2+\eta(i)}{N-2},$$

where Z is a normalizing constant given by

$$Z = \binom{(K+1)N-K-1}{KN-K-1}.$$

Proof. Due to the simple geometry of the complete graph, it is easy to see that the invariant measure is a product measure. The proof of this result is already proved in a more general setting in [15, Section 4] and [17, Theorem 2.1]. □

Corollary 3.5 (Marginal laws)**.** *For all $i \in \{1,\dots,K\}$ we have*

$$\mathbb{P}_{\nu_N}(\eta(i) = x) = \frac{1}{Z}\binom{N-2+x}{N-2}\binom{KN-K-x}{(K-1)N-K},$$

Proof. Firstly let us recall the Vandermonde binomial convolution type formula: let $n, n_1, \dots, n_p$ be some non-negative integers verifying $\sum_{i=1}^{p} n_i = n$, we have

$$\binom{r-1}{n-1} = \sum_{r_1+\cdots+r_p=r}\prod_{j=1}^{p}\binom{r_j-1}{n_j-1}.$$

The proof is based on the power series decomposition of $z \mapsto (z/(1-z))^n = \prod_{i=1}^p (z/(1-z))^{n_i}$. Using this formula, we find

$$\begin{aligned}\mathbb{P}_{\nu_N}(\eta(i)=x) &= \sum_{\overline{x}\in E_1} \mathbb{P}_{\nu_N}(\eta=(x_1,\dots,x_{i-1},x,x_{i+1}\dots,x_K)) \\ &= \frac{1}{Z}\binom{N-2+x}{N-2}\sum_{\overline{x}\in E_1}\prod_{l=1}^{i-1}\prod_{l=i+1}^{K}\binom{N-2+x_l}{N-2} \\ &= \frac{1}{Z}\binom{N-2+x}{N-2}\binom{(K-1)(N-1)+N-x-1}{(K-1)(N-1)-1},\end{aligned}$$

where $E_1 = \{\overline{x}=(x_1,\dots,x_{i-1},x_{i+1}\dots,x_K) | x_1+\dots+x_{i-1}+x_{i+1}\dots+x_K = N-x\}$. □

We are now able to express the particle correlations under this invariant measure.

Theorem 3.6 (Correlation estimates)**.** *For all $i \neq j \in \{1,\dots,K\}$, we have*

$$|Cov_{\nu_N}(\eta(i)/N,\eta(j)/N)| \sim_N \frac{K+1}{K^3N},$$

Proof. Let η be a random variable with law ν_N. As $\eta(1),\dots,\eta(K)$ are identically distributed and $\sum_{i=1}^K \eta(i) = N$ we have

$$\mathrm{Cov}_{\nu_N}(\eta(i)/N,\eta(j)/N) = -\frac{\mathrm{Var}_{\nu_N}(\eta(i)/N)}{K-1}.$$

Using the results of Section 3.4, we have

$$\mathcal{L}(\eta(i)^2) = \eta(i)^2\left[-2-\frac{2}{K(N-1)}\right]+\eta(i)\left[\frac{2N}{K}+\frac{2N}{K(N-1)}+\frac{K-2}{K}\right]+\frac{N}{K}.$$

Using the fact that $\int \mathcal{L}(\eta(i)^2)d\nu_N = 0$ and $\int \eta(i)d\nu_N = \frac{N}{K}$, we deduce that

$$\int \eta(i)^2 d\nu_N = \frac{(2N^2+(K-2)(N-1))N+NK(N-1)}{2K(KN-(K-1))}.$$

Finally,

$$\mathrm{Var}_{\nu_N}(\eta(i)) = \int \eta(i)^2 d\nu_N - \left(\int \eta(i)d\nu_N\right)^2 = \frac{N^2(K^2-1)-NK(K-1)}{K^2(KN-(K-1))},$$

and thus, for $i \neq j$,

$$|\mathrm{Cov}_{\nu_N}(\eta(i)/N,\eta(j)/N)| \sim_N \frac{K+1}{K^3N}.$$

□

Remark 3.7 (Number of sites)**.** *Theorem 3.6 gives the rate of the decay of correlations with respect to the number of particles, but we also have a rate with respect to the number of sites K: if η is distributed under the invariant measure,*

$$|Cov_{\nu_N}(\eta(i)/N,\eta(j)/N)| \sim_K \frac{1}{K(K-1)N}.$$

The previous theorem shows that the occupation numbers of two distinct sites become non-correlated when the number of particles increases. In fact, Theorem 3.6 leads to a propagation of chaos:

Corollary 3.8 (Convergence to the QSD)**.** *We have*

$$\mathbb{E}_{\nu_N}\left[d_{TV}(m(\eta),\pi_K)\right] \leq \sqrt{\frac{K}{N}},$$

where π_K is the uniform measure on $\{1,\dots,K\}$.

Proof. By the Cauchy-Schwarz inequality, we have

$$\mathbb{E}_{\nu_N}\left[\left|\frac{\eta(k)}{N}-\frac{1}{K}\right|\right] \leq \left(\mathbb{E}_{\nu_N}\left[\left|\frac{\eta(k)}{N}-\frac{1}{K}\right|^2\right]\right)^{\frac{1}{2}} = \mathrm{Var}_{\nu_N}\left(\frac{\eta(k)}{N}\right)^{1/2} \leq \sqrt{\frac{K^2-1}{K^3N}}.$$

Summing over $\{1,\dots,K\}$ ends the proof. □

The previous bound is better than the bound obtained in Theorem 1.2 and its corollaries. This comes from the absence of bias term. Indeed,

$$\forall k \in F^*,\ \mathbb{E}_{\nu_N}[m(\eta)(k)] = \frac{1}{K} = \pi_K(k).$$

The bad term in Theorem 1.2 comes from, with the notations of its proof, the estimation of $|u_k(t)-v_k(t)|$ and Gronwall Lemma.

3.4. **Long time behavior and spectral analysis of the generator.** In this subsection, we point out the optimality of Theorem 1.1 in this special case. It gives

Corollary 3.9 (Wasserstein contraction)**.** *For any processes $(\eta_t)_{t>0}$ and $(\eta'_t)_{t>0}$ generated by* (13)*, and for any $t \geq 0$, we have*

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

In particular, when (η'_0) follows the invariant distribution ν_N associated to (13)*, we get for every $t \geq 0$*

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

In particular, if λ_1 is the smallest positive eigenvalue of $-\mathcal{L}$, defined at (13), then we have

$$1 = \rho \leq \lambda_1.$$

Indeed, on the one hand, let us recall that, as the invariant measure is reversible, λ_1 is the largest constant such that

$$\lim_{t\to+\infty} e^{2\lambda t}\|R_t f - \nu_N(f)\|^2_{L^2(\nu_N)} = 0, \tag{15}$$

for every $\lambda < \lambda_1$ and $f \in L^2(\nu_N)$, where $(R_t)_{t\geq 0}$ is the semi-group generated by $\mathcal{L}$. See for instance [2, 23]. On the other hand, if $\lambda < 1$ then, by Theorem 1.1, we have

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

and then (15) 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 $\varphi_k^{(1)}$ verifies

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

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

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

Remark 3.10 (Complete graph random walk)**.** *If $(a_i)_{1\leq i\leq K}$ is a sequence such that $\sum_{i=1}^K a_i = 0$ then the function $\sum_{i=1}^K \varphi_i^{(1)}$ is an eigenvector of $\mathcal{L}$. However, if L is the generator of the classical complete graph random walk, $La = -a$ and then a is also an eigenvector of L with the same eigenvalue.*

Let us finally give the following result on the spectrum of $\mathcal{L}$:

Lemma 3.11 (Spectrum of $-\mathcal{L}$)**.** *The spectrum of $-\mathcal{L}$ is included in*

$$\left\{ \sum_{i=1}^{K} \lambda_{l_i} \mid l_1, \dots, l_K \in \{0, \dots, N\} \right\},$$

where

$$\forall l \in \{0, \dots, N\},\ \lambda_l = l + \frac{l(l-1)}{K(N-1)}.$$

Proof. For every $k \in \{1, \dots, K\}$ and $l \in \{0, \dots, N\}$, we have

$$\mathcal{L}\varphi_k^{(l)}(\eta) = -\lambda_l \varphi_k^{(l)}(\eta) + Q_{l-1}(\eta),$$

where Q_{l-1} is a polynomial whose degree is less than $l-1$. A straightforward recurrence shows that whether there exists or not a polynomial function $\psi_k^{(l)}$, whose degree is l, verifying $\mathcal{L}\psi_k^{(l)} = -\lambda_l \psi_k^{(l)}$ (namely $\psi_k^{(l)}$ is an eigenvector of $\mathcal{L}$). Indeed, it is possible to have $\psi_k^{(l)} = 0$ since the polynomial functions are not linearly independent (F is finite). More generally, for all $l_1, \dots, l_K \in \{1, \dots, N\}$, there exists a polynomial Q with K variables, whose degree with respect to the i^{th} variable is strictly less than l_i, such that the function $\phi : \eta \mapsto \prod_{i=1}^{K} \eta(k_i)^{l_i} + Q(\eta)$ satisfies

$$\mathcal{L}\phi = -\lambda\phi \text{ where } \lambda = \sum_{i=1}^{K} \lambda_{l_i}.$$

Again, provided that $\phi \neq 0$, ϕ is an eigenvector and λ an eigenvalue of $-\mathcal{L}$. Finally, as the state space is finite, using multivariate Lagrange polynomial, we can prove that every function is polynomial and thus we capture all the eigenvalues. □

Remark 3.12 (Cardinal of E)**.** *As card$(F^*) = K$, we have*

$$\textit{card}(E) = \binom{N+K-1}{K-1} = \frac{(N+K-1)!}{N!(K-1)!}.$$

In particular, the number of eigenvalues is finite and less than card(E).

Remark 3.13 (Marginals)**.** *For each k, the random process $(\eta_t(k))_{t\geq 0}$, which is a marginal of a process generated by* (13)*, is a Markov process on $\mathbb{N}_N = \{0, \dots, N\}$ generated by*

$$\begin{aligned}\mathcal{G}f(x) = \frac{(N-x)}{K}\left(1 + \frac{x}{N-1}\right)(f(x+1) - f(x)) \\ + \frac{x}{K}\left(K - 1 + \frac{N-x}{N-1}\right)(f(x-1) - f(x)),\end{aligned}$$

for every function f on $\mathbb{N}_N$ and $x \in \mathbb{N}_N$. We can express the spectrum of this generator. Indeed, let $\varphi_l : x \mapsto x^l$, for every $l \geq 0$. The family $(\varphi_l)_{0\leq l\leq N}$ is linearly independent as can be checked with a Vandermonde determinant. This family generates the L^2-space associated to the invariant measure since this space has a dimension equal to $N+1$. Now, similarly to the proof of the previous lemma, we can prove the existence of $N+1$ polynomials, which are eigenvectors and linearly independent, whose eigenvalues are $\lambda_0, \lambda_1, \dots, \lambda_N$.

4. The two point space

We consider a Markov chain defined on the states $\{0, 1, 2\}$ where 0 is the absorbing state. Its infinitesimal generator G is defined by

$$G = \begin{bmatrix} 0 & 0 & 0 \\ p_0(1) & -a - p_0(1) & a \\ p_0(2) & b & -b - p_0(b), \end{bmatrix}$$

where $a, b > 0$, $p_0(1), p_0(2) \geq 0$ and $p_0(1) + p_0(2) > 0$. The generator of the Fleming-Viot process with N particles applied to bounded functions $f : E \to \mathbb{R}$ reads

$$\begin{aligned}\mathcal{L}f(\eta) = \eta(1)\left(a + p_0(1)\frac{\eta(2)}{N-1}\right)(f(T_{1\to 2}\eta) - f(\eta)) \\ + \eta(2)\left(b + p_0(2)\frac{\eta(1)}{N-1}\right)(f(T_{2\to 1}\eta) - f(\eta)).\end{aligned} \tag{16}$$

4.1. **The associated killed process.** The long time behavior of the conditionned process is related to the eigenvalues and eigenvectors of the matrix:

$$M = \begin{bmatrix} -a - p_0(1) & a \\ b & -b - p_0(2) \end{bmatrix}.$$

Indeed see [21, section 3.1]. Its eigenvalues are given by

$$\lambda_+ = \frac{-(a + b + p_0(1) + p_0(2)) + \sqrt{(a - b + p_0(1) - p_0(2))^2 + 4ab}}{2},$$

$$\lambda_- = \frac{-(a + b + p_0(1) + p_0(2)) - \sqrt{(a - b + p_0(1) - p_0(2))^2 + 4ab}}{2},$$

and the corresponding eigenvectors are respectively given by

$$v_+ = \begin{pmatrix} a \\ -A + \sqrt{A^2 + 4ab} \end{pmatrix} \text{ and } v_- = \begin{pmatrix} a \\ -A - \sqrt{A^2 + 4ab} \end{pmatrix},$$

where $A = a - b + p_0(1) - p_0(2)$. From these properties, we deduce that

Lemma 4.1 (Convergence to the QSD)**.** *There exists a constant $C > 0$ such that for every initial distribution μ, we have*

$$\forall t \geq 0,\ d_{TV}(\mu T_t, v_+) \leq Ce^{-(\lambda_+ - \lambda_-)t}.$$

Proof. See [21, Theorem 7] and [21, Remark 3]. □

Note that

$\lambda_+ - \lambda_- = \sqrt{(a+b)^2 + 2(a-b)(p_0(1) - p_0(2)) + (p_0(1) - p_0(2))^2} > a + b - (\max(p_0) - \min(p_0))$ when $\max(p_0) > \min(p_0)$.

4.2. **Explicit formula of the invariant distribution.** Firstly note that each marginal is a Markov process:

Lemma 4.2 (Markovian marginals)**.** *The random process $(\eta_t(1))_{t\geq 0}$, which is a marginal of a process generated by* (16)*, is a Markov process generated by $\mathcal{G}$ defined by*

$$\mathcal{G}f(n) = b_n(f(n+1) - f(n)) + d_n(f(n-1) - f(n)), \tag{17}$$

for any function f and $n \in \mathbb{N}_N = \{0, \ldots, N\}$, where

$$b_n = (N - n)\left(b + p_0(2)\frac{n}{N-1}\right) \text{ and } d_n = n\left(a + p_0(1)\frac{N-n}{N-1}\right).$$

Proof. For every $\eta \in E$, we have $\eta = (\eta(1), N - \eta(1))$ thus the Markov property and the generator are easily deducible from the properties of $(\eta_t)_{t\geq 0}$. □

From this result and the already known results on birth and death processes [6, 7], we deduce that $(\eta_t(1))_{t\geq 0}$ admits an invariant and reversible distribution π given by

$$\pi(n) = u_0 \prod_{k=1}^{n} \frac{b_{k-1}}{d_k} \text{ and } u_0^{-1} = 1 + \sum_{k=1}^{N} \frac{b(0)\cdots b(k-1)}{d(1)\cdots d(k)},$$

for every $n \in \mathbb{N}_{\mathrm{N}}$. This gives

$$\pi(n) = u_0 \binom{N}{n} \prod_{k=1}^{n} \frac{b(N-1) + (k-1)p_0(2)}{a(N-1) + (N-k)p_0(1)},$$

and

$$u_0^{-1} = 1 + \prod_{k=1}^{N} \frac{b(N-1) + kp_0(2)}{a(N-1) + kp_0(1)}.$$

Similarly, the process $(\eta_t(2))_{t\geq 0}$ is also a Markov process whose invariant distribution is also easily calculable. The invariant law of $(\eta_t)_{t\geq 0}$, is then given by

$$\nu_N((r_1, r_2)) = \pi\left(\{r_1\}\right), \quad \forall (r_1, r_2) \in E.$$

Note that if p_0 is not constant then we can not find a basis of orthogonal polynomials in the L^2 space associated to ν_N. It is then very difficult to express the spectral gap or the decay rate of the correlations without using our main results.

4.3. **Rate of convergence.** Applying Theorem 1.1, in this special case, we find:

Corollary 4.3 (Wasserstein contraction)**.** *For any processes $(\eta_t)_{t>0}$ and $(\eta'_t)_{t>0}$ generated by* (16)*, and for any $t \geq 0$, we have*

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

where $\rho = a+b-(\max(p_0)-\min(p_0))$. In particular, when (η'_0) follows the invariant distribution ν_N of (16)*, we get for every $t > 0$*

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

This result is not optimal. Nevertheless, the error does not come from our coupling choice but it comes from how we estimate the distance. Indeed, this coupling induce a coupling between two processes generated by $\mathcal{G}$ defined by (17). More precisely, let $\mathbb{L} = \mathbb{L}_Q + \mathbb{L}_p$ be the generator of our coupling introduced in the proof of Theorem 1.1 in this special case. We set $\mathbb{G} = \mathbb{G}_Q + \mathbb{G}_p$, where for any $n, n' \in \mathbb{N}_N$ and f on $E \times E$,

$$\mathbb{L}_Q f((n, N-n), (n', N-n')) = \mathbb{G}_Q \varphi_f(n, n'),$$

$$\mathbb{L}_p f((n, N-n), (n', N-n')) = \mathbb{G}_p \varphi_f(n, n'),$$

and $\varphi_f(n, n') = f((n, N-n), (n', N-n'))$. It verifies, for any function f on $\mathbb{N}_{\mathrm{N}}$ and $n' > n$ two elements of $\mathbb{N}_{\mathrm{N}}$,

$$\begin{aligned}
\mathbb{G}_Q f(n, n') = {} & na\left(f(n-1, n'-1) - f(n, n')\right) \\
& + (N-n')b\left(f(n+1, n'+1) - f(n, n')\right) \\
& + (n'-n)b\left(f(n+1, n') - f(n, n')\right) \\
& + (n'-n)a\left(f(n, n'-1) - f(n, n')\right),
\end{aligned}$$

and

$$\begin{aligned}\mathbb{G}_p f(n,n') &= p_0(1)\frac{n(N-n')}{N-1}\big(f(n-1,n'-1)-f(n,n')\big)\\ &+ p_0(2)\frac{n(N-n')}{N-1}\big(f(n+1,n'+1)-f(n,n')\big)\\ &+ p_0(1)\frac{n(n'-n)}{N-1}\big(f(n-1,n')-f(n,n')\big)\\ &+ p_0(2)\frac{(N-n')(n'-n)}{N-1}\big(f(n,n'+1)-f(n,n')\big)\\ &+ p_0(2)\frac{n(n'-n)}{N-1}\big(f(n+1,n')-f(n,n')\big)\\ &+ p_0(1)\frac{(N-n')(n'-n)}{N-1}\big(f(n,n'-1)-f(n,n')\big).\end{aligned}$$

Now, for any sequence of positive numbers $(u_k)_{k\in\{0,\dots,N-1\}}$, we introduce the distance δ_u defined by

$$\delta_u(n,n') = \sum_{k=n}^{n'-1} u_k,$$

for every $n, n' \in \mathbb{N}_{\mathrm{N}}$ such that $n' > n$. For all $n \in \mathbb{N}_{\mathrm{N}}\backslash\{N\}$, we have $\mathbb{G}\delta_u(n,n+1) \le -\lambda_u\delta_u(n,n+1)$ where

$$\lambda_u = \min_{k\in\{0,\dots,N-1\}}\left[d(k+1) - d(k)\frac{u_{k-1}}{u_k} + b(k) - b(k+1)\frac{u_{k+1}}{u_k}\right],$$

and thus, by linearity, $\mathbb{G}\delta_u(n,n') \le -\lambda_u\delta_u(n,n')$, for every $n, n' \in \mathbb{N}_{\mathrm{N}}$. This implies that for any processes $(X_t)_{t\ge0}$ and $(X'_t)_{t\ge0}$ generated by $\mathcal{G}$, and for any $t \ge 0$,

$$\mathcal{W}_{\delta_u}(\mathcal{L}(X_t),\mathcal{L}(X'_t)) \le e^{-\lambda_u t}\mathcal{W}_{\delta_u}(\mathcal{L}(X_0),\mathcal{L}(X'_0)).$$

Note that, for every $n, n' \in \mathbb{N}_{\mathrm{N}}$, we have

$$\min(u)d_1((n,N-n),(n',N-n')) \le \delta_u(n,n') \le \max(u)d_1((n,N-n),(n',N-n')),$$

and then for any processes $(\eta_t)_{t\ge0}$ and $(\eta'_t)_{t\ge0}$ generated by (16), and for any $t \ge 0$, we have

$$\mathcal{W}_{d_1}(\mathcal{L}(\eta_t),\mathcal{L}(\eta'_t)) \le \frac{\max(u)}{\min(u)}e^{-\lambda_u t}\mathcal{W}_{d_1}(\mathcal{L}(\eta_0),\mathcal{L}(\eta'_0)).$$

Finally, using [7, Theorem 9.25], there exists a positive sequence v such that $\lambda_v = \max_u \lambda_u > 0$ is the spectral gap of the birth and death process $(\eta_t(1))_{t\ge0}$. These parameters depend on N and so we should write the previous inequality as

$$\mathcal{W}_{d_1}(\mathcal{L}(\eta_t),\mathcal{L}(\eta'_t)) \le C(N)e^{-\lambda_N t}\mathcal{W}_{d_1}(\mathcal{L}(\eta_0),\mathcal{L}(\eta'_0)), \tag{18}$$

where $C(N)$ and λ_N are two constants depending on N. In conclusion, the coupling introduced in Theorem 1.1 gives the optimal rate of convergence but we are not able to express this rate and its dependence on N.

4.4. **Correlations.** Using Theorem 2.5, we have

Corollary 4.4 (Correlations)**.** *If $(\eta_t)_{t\ge0}$ is a process generated by* (16) *then we have for all $t \ge 0$,*

$$\mathit{Cov}(\eta_t(k)/N,\eta_t(l)/N) \le \frac{2}{N^2}\frac{1-e^{-2\rho t}}{\rho}\left(N(a\vee b) + \max(p_0)\frac{N^2}{N-1}\right).$$

If $\rho \leq 0$, the right-hand side of the two previous inequalities explodes as t tends to infinity whereas these correlations are bounded by 1. Nevertheless, using Theorem 2.5, Remark 2.6 and Inequality (18), we can prove that there exists two constants $C'(N)$, depending on N, and K, which does not depend on N, such that

$$\sup_{t\geq 0} \mathrm{Cov}(\eta_t(k)/N, \eta_t(l)/N) \leq C'(N) = \frac{KC(N)}{N\lambda_N},$$

where $C(N)$ is defined in (18). Unfortunately, $C(N)$ is not (completely) explicit and we do not know if the right-hand side of the previous expression tends to 0 as N tends to infinity. This example shows the difficulty of finding explicit and optimal rates of the convergence towards equilibrium and the decay of correlations; it also illustrates that our main results are extremely useful when $\max(p_0) \neq \min(p_0)$.

Acknowledgement: We would like to 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 while a postdoctoral scholar 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, and P. Groisman. Quasistationary distributions and Fleming-Viot processes in finite spaces. *J. Appl. Probab.*, 48(2):322–332, 2011.

[2] D. Bakry. L'hypercontractivité et son utilisation en théorie des semigroupes. In *Lectures on probability theory (Saint-Flour, 1992)*, volume 1581 of *Lecture Notes in Math.*, pages 1–114. Springer, Berlin, 1994.

[3] D. Bakry, P. Cattiaux, and A. Guillin. Rate of convergence for ergodic continuous Markov processes: Lyapunov versus Poincaré. *J. Funct. Anal.*, 254(3):727–759, 2008.

[4] M. Bieniek, K. Burdzy, and S. Pal. Extinction of Fleming-Viot-type particle systems with strong drift. *Electron. J. Probab.*, 17:no. 11, 15, 2012.

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

[6] D. Chafaï and A. Joulin. Intertwining and commutation relations for birth–death processes. *Bernoulli*, 19(5A):1855–1879, 2013.

[7] M.-F. Chen. *From Markov chains to non-equilibrium particle systems*. World Scientific Publishing Co. Inc., River Edge, NJ, second edition, 2004.

[8] B. Cloez and M. Hairer. Exponential ergodicity for Markov processes with random switching. *ArXiv e-prints*, Mar. 2013.

[9] J. N. Darroch and E. Seneta. On quasi-stationary distributions in absorbing continuous-time finite Markov chains. *J. Appl. Probability*, 4:192–196, 1967.

[10] P. Del Moral and A. Guionnet. On the stability of measure valued processes with applications to filtering. *C. R. Acad. Sci. Paris Sér. I Math.*, 329(5):429–434, 1999.

[11] P. Del Moral and L. Miclo. A Moran particle system approximation of Feynman-Kac formulae. *Stochastic Process. Appl.*, 86(2):193–216, 2000.

[12] P. Del Moral and L. Miclo. On the stability of nonlinear Feynman-Kac semigroups. *Ann. Fac. Sci. Toulouse Math. (6)*, 11(2):135–175, 2002.

[13] A. Eberle. Couplings, distances and contractivity for diffusion processes revisited. *ArXiv e-prints*, May 2013.

[14] P. A. Ferrari and N. Marić. Quasi stationary distributions and Fleming-Viot processes in countable spaces. *Electron. J. Probab.*, 12:no. 24, 684–702, 2007.

[15] C. Giardinà, F. Redig, and K. Vafayi. Correlation inequalities for interacting particle systems with duality. *J. Stat. Phys.*, 141(2):242–263, 2010.

[16] P. Groisman and M. Jonckheere. Simulation of quasi-stationary distributions on countable spaces. *ArXiv e-prints*, June 2012.

[17] S. Grosskinsky, F. Redig, and K. Vafayi. Condensation in the inclusion process and related models. *J. Stat. Phys.*, 142(5):952–974, 2011.

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

[19] M. Hairer and J. C. Mattingly. Yet another look at Harris' ergodic theorem for Markov chains. In *Seminar on Stochastic Analysis, Random Fields and Applications VI*, volume 63 of *Progr. Probab.*, pages 109–117. Birkhäuser/Springer Basel AG, Basel, 2011.

[20] S. Martinez, J. San Martin, and D. Villemonais. Existence and uniqueness of a quasi-stationary distribution for Markov processes with fast return from infinity. *ArXiv e-prints*, Feb. 2012.
[21] S. Méléard and D. Villemonais. Quasi-stationary distributions and population processes. *Probab. Surv.*, 9:340–410, 2012.
[22] S. P. Meyn and R. L. Tweedie. Stability of Markovian processes. III. Foster-Lyapunov criteria for continuous-time processes. *Adv. in Appl. Probab.*, 25(3):518–548, 1993.
[23] L. Saloff-Coste. Lectures on finite Markov chains. In *Lectures on probability theory and statistics (Saint-Flour, 1996)*, volume 1665 of *Lecture Notes in Math.*, pages 301–413. Springer, Berlin, 1997.
[24] A.-S. Sznitman. Topics in propagation of chaos. In *École d'Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Math.*, pages 165–251. Springer, Berlin, 1991.
[25] D. Villemonais. General approximation method for the distribution of Markov processes conditioned not to be killed. *ArXiv e-prints*, June 2011.

(Bertrand CLOEZ) UNIVERSITÉ DE TOULOUSE, INSTITUT DE MATHÉMATIQUES DE TOULOUSE, CNRS UMR5219, FRANCE
E-mail address: bertrand.cloez@univ-toulouse.fr

(Marie-Noémie THAI) LABORATOIRE D'ANALYSE ET DE MATHÉMATIQUES APPLIQUÉES, CNRS UMR8050, UNIVERSITÉ PARIS-EST MARNE-LA-VALLÉE, FRANCE
E-mail address: marie-noemie.thai@univ-paris-est.fr