Processus de Fleming-Viot, distributions quasi-stationnaires et marches aléatoires en interaction de type champ moyen

Anh-Thi Marie Noémie Thai

THÈSE

Pour l'obtention du grade de

DOCTEUR DE L'UNIVERSITÉ PARIS-EST

ÉCOLE DOCTORALE MATHEMATIQUES ET SCIENCES ET TECHNOLOGIE DE L'INFORMATION ET DE LA COMMUNICATION

DISCIPLINE : MATHÉMATIQUES

Présentée par

Marie-Noémie Thai

Processus de Fleming-Viot, distributions quasi-stationnaires et marches aléatoires en interaction de type champ moyen

Directeurs de thèse : **Amine Asselah et Djalil Chafaï**

Soutenue le 27 novembre 2015
Devant le jury composé de

M. Amine ASSELAH	Université Paris-Est Créteil	Directeur de thèse
M. Vincent BANSAYE	École Polytechnique	Rapporteur
M. Djalil CHAFAÏ	Université Paris Dauphine	Directeur de thèse
M. Nicolas CHAMPAGNAT	INRIA, Université de Lorraine	Rapporteur
M. Jean-François DELMAS	École Nationale des Ponts et Chaussées	Examinateur
Mme. Ellen SAADA	CNRS, Université Paris Descartes	Présidente

Thèse préparée au
Laboratoire LAMA CNRS UMR 8050
Université de Paris-Est Marne-la-Vallée
5, boulevard Descartes, Champs-sur-Marne
77454 Marne-la-Vallée cedex 2, France

RDMath IdF
Domaine d'Intérêt Majeur (DIM)
en Mathématiques

îledeFrance

Remerciements

Mes travaux de thèse, présentés dans ce manuscrit, n'auraient pu aboutir sans l'aide et la présence de nombreuses personnes que je tiens à remercier sincèrement ici.

Je tiens tout d'abord à remercier mes directeurs Amine Asselah et Djalil Chafaï de m'avoir guidée et soutenue durant ces trois années. Je tiens à les remercier pour la confiance qu'ils m'ont accordée en acceptant d'encadrer ce travail doctoral. J'ai eu beaucoup de chance de pouvoir bénéficier de leur savoir mathématique et de leur disponibilité. Je tiens en particulier à remercier Djalil Chafaï pour sa qualité d'écoute, ses encouragements dans mes périodes de doute et de toute la liberté qu'il m'a laissée.

Je souhaiterais exprimer ma profonde gratitude à Robert Eymard pour avoir accepté d'encadrer mon mémoire de M2 et de m'avoir ainsi initié à la recherche. J'ai pris énormément de plaisir à travailler avec Robert Eymard, ses qualités humaines, sa joie de vivre, sa culture mathématique, tant d'aspects qui en font une personne remarquable. Merci pour votre générosité et surtout un Grand Merci d'avoir cru en moi.

Je remercie Vincent Bansaye et Nicolas Champagnat qui m'ont fait l'honneur de rapporter ma thèse et Jean-François Delmas et Ellen Saada d'avoir accepté de faire partie de mon jury.

Je tiens également à remercier la région île de France d'avoir financé mes années de thèse en m'accordant la bourse DIM RDM-IDF.

Je souhaite remercier les membres du LAMA pour leur accueil, leur écoute et leurs conseils et en particulier Christiane Lafargue et Audrey Patout pour s'être occupées de toutes les démarches administratives. Egalement un grand merci à Sylvie Cach pour s'être occupée de toutes les démarches pour mon voyage à Buenos Aires et pour avoir pris le temps de répondre à bon nombre de mes questions et cela toujours avec le sourire.

Les années de thèse ne sauraient être dissociées des enseignements que j'ai assurés en tant que monitrice. Je tiens à remercier les professeurs avec qui j'ai travaillé et en particulier Thierry Berkover pour tous ses conseils et ses renseignements concernant les postes PRAG.

Comment finir cette thèse sans parler des doctorants et des jeunes docteurs. Je remercie Ali, Arnaud, Bertrand, Harry, Marwa, Paolo, Pierre(Feron), Pierre(Youssef), Rania, Sébastien et Xavier pour tous les bons moments que l'on a passé ensemble. Sans vous je pense que cette thèse aurait été différente. En particulier, un Grand Merci à Marwa, Pierre (Feron) et Rania pour leur écoute, leur soutien permanent et pour tous les moments de rire, d'émotions et de complicité que l'on a partagés. Un Grand Grand Merci à Bertrand pour être arrivé à un moment de ma thèse où tout allait mal. Merci de m'avoir tant aidé tout le long de ma thèse. Bertrand mon troisième directeur ? Je souhaite également remercier Ludovic Goudenège pour les bons moments passés ensemble, pour son aide en simulation et surtout pour sa grande patience envers moi. En commençant cette thèse, je n'aurai jamais imaginé que j'en ressortirai avec de véritables amis.

Enfin, un Immense Merci à mes parents, Marie-Christine et Van Minh, ma soeur Elodie ainsi qu'à un de mes amis David. Merci pour votre présence, votre soutien et vos encouragements. Mon dernier remerciement s'adresse à toi, Tony. Merci de m'avoir supporté, rassuré, consolé, et tant donné. Vous m'avez toujours poussée à aller plus loin, vous m'avez aidée à surmonter tous les moments difficiles et surtout vous avez toujours cru en moi. Sans vous je ne serais pas là aujourd'hui, cette thèse je vous la dédie.

Résumé

Dans cette thèse nous étudions le comportement asymptotique de systèmes de particules en interaction de type champ moyen en espace discret, systèmes pour lesquels l'interaction a lieu par l'intermédiaire de la mesure empirique. Dans la première partie de ce mémoire, nous nous intéressons aux systèmes de particules de type Fleming-Viot : les particules se déplacent indépendamment suivant une dynamique markovienne jusqu'au moment où l'une d'entre elles touche un état absorbant. A cet instant, la particule absorbée choisit uniformément une autre particule et saute sur sa position. L'ergodicité du processus est établie dans le cadre de marches aléatoires sur $\mathbb{N}$ avec dérive vers l'origine et pour une dynamique proche de celle du graphe complet. Pour ce dernier, nous obtenons une estimation quantitative de la convergence en temps long à l'aide de la courbure de Wasserstein. Nous montrons de plus la convergence de la distribution empirique stationnaire vers une unique distribution quasi-stationnaire, quand le nombre de particules tend vers l'infini. Dans la deuxième partie de ce mémoire, nous nous intéressons au comportement en temps long et quand le nombre de particules devient grand, d'un système de processus de naissance et mort pour lequel les particules interagissent à chaque instant par le biais de la moyenne de leurs positions. Nous établissons l'existence d'une limite macroscopique, solution d'une équation non linéaire ainsi que le phénomène de propagation du chaos avec une estimation quantitative et uniforme en temps.

Mots clés : interaction de type champ moyen - processus de Fleming-Viot - distance de Wasserstein - couplage - propagation du chaos - marche aléatoire

Abstract

In this thesis we study the asymptotic behavior of particle systems in mean field type interaction in discrete space, where the system acts over one fixed particle through the empirical measure of the system. In the first part of this thesis, we are interested in Fleming-Viot particle systems : the particles move independently of each other until one of them reaches an absorbing state. At this time, the absorbed particle jumps instantly to the position of one of the other particles, chosen uniformly at random. The ergodicity of the process is established in the case of random walks on $\mathbb{N}$ with a dirft towards the origin and on complete graph dynamics. For the latter, we obtain a quantitative estimate of the convergence described by the Wasserstein curvature. Moreover, under the invariant measure, we show the convergence of the empirical measure towards the unique quasi-stationary distribution as the size of the system tends to infinity. In the second part of this thesis, we study the behavior in large time and when the number of particles is large of a system of birth and death processes where at each time a particle interacts with the others through the mean of theirs positions. We establish the existence of a macroscopic limit, solution of a non linear equation and the propagation of chaos phenomenon with quantitative and uniform in time estimate.

Keywords : mean field type interaction - Fleming-Viot process - Wasserstein distance - coupling - propagation of chaos - random walk

Table des matières

Introduction

Cette thèse porte sur l'étude des systèmes de particules en interaction de type champ moyen en espace discret et de leur approximation. Dans le modèle qui nous intéresse, les particules se déplacent indépendamment suivant une dynamique markovienne jusqu'au moment où l'une d'entre elles touche l'état absorbant. A cet instant, la particule absorbée choisit uniformément une autre particule et saute sur sa position. Ce système, appelé processus de Fleming-Viot, présente deux difficultés

1. le caractère localisé du site d'absorption.
2. la longue portée des sauts que les particules font.

Ces difficultés se retrouvent notamment dans l'étude des processus de Fleming-Viot ayant une dérive constante vers l'origine, où seule l'ergodicité a pu être établie. Il n'existe aucune preuve générale de convergence de la densité empirique d'équilibre (stationnaire), lorsque le nombre de particules tend vers l'infini. Afin de simplifier le problème, nous allons considérer dans un premier temps le modèle du graphe complet et son extension naturelle. Dans un second temps, nous considérons un modèle local où les sauts à longue portée sont omis. Il peut être vu comme une version discrète de celui introduit pour l'étude des équations de McKean-Vlasov. Les chapitres 3 et 4 de ce manuscrit relaxent respectivement les difficultés 1 et 2. Afin d'étudier le comportement asymptotique de ces systèmes, nous introduisons dans un premier temps les outils liés à l'ergodicité (chapitre 1), puis étudions par la suite les différents modèles.

Le manuscrit se décompose en 4 chapitres :

– Dans le premier chapitre, nous introduisons différents outils permettant d'établir l'ergodicité des processus de Markov. En particulier, nous introduisons les notions de couplage et de fonctions de Lyapunov et mettons en évidence le lien entre ces notions et le comportement en temps long d'un processus de Markov.

– Le chapitre 2, constitué de l'article [10] en collaboration avec Amine Asselah, donne l'ergodicité d'un processus de Fleming-Viot conduit par une marche aléatoire sur $\mathbb{N}$ avec dérive vers l'origine.
– Le chapitre 3, basé sur l'article [34] en collaboration avec Bertrand Cloez, est consacré à l'étude d'un modèle sans géométrie où chaque particule peut passer d'un état à un autre avec probabilité strictement positive.
– Pour terminer, le chapitre 4 est consacré à l'étude des processus de naissance et mort en interaction de type champ moyen, pour lesquels les particules interagissent à chaque instant par l'intermédiaire de la moyenne de leurs positions. Dans les chapitres 3 et 4 nous établissons des théorèmes limites en grande taille et en temps long pour la mesure empirique décrivant le système de particules.

Pour comprendre l'intérêt porté au processus de Fleming-Viot, nous allons établir le lien existant entre ce processus et les distributions dites quasi-stationnaires.
En 1874, Galton et Watson [104] ont introduit les processus de branchement, appelés aussi processus de Galton-Watson, afin d'étudier le phénomène d'extinction de noms de familles aristocratiques. Ces processus permettent de modéliser la dynamique d'une population dont les individus (ou particules ou cellules ...) vivent, se reproduisent et meurent. Aujourd'hui, la théorie des processus de branchement présente un panorama d'une grande richesse : applications en biologie, généalogie ou chimie. De plus, l'étude mathématique de ces processus intègre des situations de plus en plus complexes : processus de branchement avec immigration [68] ou en environnement aléatoire [11, 15, 93]. Dans l'évolution d'une population, une question fondamentale est de déterminer sa probabilité d'extinction. Il semble qu'en 1938, Kolmogorov [73] a initié l'étude en estimant la probabilité qu'une population soit encore vivante après un grand nombre de générations. Dans le cas où il y a extinction presque sûre d'une population, le temps d'extinction peut être long comparé à l'échelle de temps de l'observation et il a été constaté que la taille des populations fluctue pendant un long moment avant que l'extinction ne se produise réellement. Il est alors intéressant d'étudier le comportement en temps long de la population conditionnée à ne pas s'éteindre et à la notion qui lui est lié : la notion de *quasi-stationnarité*.

0.1 Processus conditionnés

On considère un processus de Markov à temps continu $(X_t)_{t\geq 0}$ irréductible sur $\Lambda \cup \{0\}$ avec Λ un espace dénombrable ou fini et 0 un état absorbant et dont les taux de transition sont donnés par la matrice $Q = (Q_{x,y})$ avec pour convention $Q_{x,x} = -\sum_{y\neq x} Q_{x,y}$ pour tout $x \in \Lambda$. On suppose que le processus $(X_t)_{t\geq 0}$ n'explose pas. Pour toute distribution initiale μ, on note par μT_t sa loi au temps t conditionnée à la non-absorption jusqu'au temps t. Elle est définie pour toute fonction positive f sur Λ et pour tout $t \geq 0$ par

$$\mu T_t f = \frac{\sum_{z\in\Lambda} P_t f(z)\mu(z)}{\sum_{z\in\Lambda} P_t \mathbb{1}_{\{0\}^c}(z)\mu(z)},$$

où on pose $f(0) = 0$ et où $P_t = e^{tQ}$ est le semi-groupe associé à la matrice Q. μT_t est alors l'unique solution de l'équation de Kolmogorov non linéaire suivante :

$$\begin{cases} \partial_t \mu T_t(x) = \sum_{y\in\Lambda} Q_{y,x}\ \mu T_t(y) + \sum_{y\in\Lambda} Q_{y,0}\ \mu T_t(y)\mu T_t(x), \quad x \in \Lambda \\ \mu T_0 = \mu. \end{cases} \tag{1}$$

Une question intéressante est celle du comportement asymptotique de μT_t :

1. La limite de μT_t, pour t tendant vers l'infini existe-t-elle ?
2. Si oui, dépend-elle de la loi initiale ? Peut-on obtenir une vitesse de convergence ?

Les premiers mathématiciens à s'y être intéressé sont Kolmogorov et Yaglom dans le cas des processus de Galton-Watson. En 1938, Kolmogorov [73] a donné des estimations de la probabilité de survie d'une population et a montré que dans le cas critique et sous-critique cette probabilité décroît quand le nombre de générations tend vers l'infini, mais plus lentement dans le cas critique que dans le cas sous-critique. Suite à cette étude, Yaglom a montré en 1947 [106], que le processus de Galton-Watson sous-critique conditionné à survivre admet une distribution limite, indépendante de la loi initiale, appelée *limite de Yaglom* :

Définition 0.1 (Limite de Yaglom)**.** *Une limite de Yaglom est une mesure de probabilité ν sur Λ telle que, pour tout $x \in \Lambda$*

$$\lim_{t\to+\infty} \mathbb{P}_x(X_t \in \cdot \mid t < \tau_0) = \nu(\cdot),$$

où τ_0 est le temps d'absorption en 0 défini par $\tau_0 = \inf\{t > 0, X_t = 0\}$.

La preuve repose sur l'analyse de la fonction génératrice du processus conditionné à la non-absorption et se trouve dans le livre de Athreya et Ney [12, p. 16].

0.2 Distributions quasi-stationnaires

La limite de Yaglom, si elle existe, est une *distribution quasi-stationnaire* c'est-à-dire une distribution invariante pour la dynamique conditionnée à ne pas s'éteindre.

Définition 0.2 (Distributions quasi-stationnaires (QSD) [*])**.** *Une distribution quasi-stationnaire (QSD) pour la matrice Q est une mesure de probabilité ν sur Λ invariante pour $(T_t)_{t\geq 0}$, c'est-à-dire*

$$\forall t \geq 0, \quad \nu T_t = \nu.$$

Quand elle existe, la limite de Yaglom est unique. Mais cela n'implique pas l'unicité de la distribution quasi-stationnaire. C'est le cas par exemple des processus de naissance et mort sur $\mathbb{N}$, voir [84] pour plus de détails.

Par (1), on remarque que ν est une QSD si et seulement si ν vérifie l'équation non-linéaire

$$\sum_{y\in\Lambda} Q_{y,x}\nu(y) + \sum_{y\in\Lambda} Q_{y,0}\ \nu(y)\nu(x) = 0, \tag{2}$$

qui est équivalente à

$$\nu Q(x) = -\left(\sum_{y\in\Lambda} Q_{y,0}\ \nu(y)\right)\nu(x).$$

Autrement dit, une QSD est un vecteur propre à gauche pour la restriction à Λ de la matrice Q de valeur propre $-\sum_{y\in\Lambda} Q_{y,0}\nu(y)$.

L'équation (2) peut être interprétée de la manière suivante : ν est la mesure invariante du processus sur Λ de générateur Q^ν donné, pour tout $x, y \in \Lambda$, par

$$Q^\nu(x,y) = Q_{x,y} + Q_{x,0}\nu(y). \tag{3}$$

Autrement dit, ν vérifie $\nu Q^\nu = 0$.

Dans le cas des chaînes de Markov à espace fini ou dénombrable, l'étude des QSD a été initiée par Darroch, Seneta et Vere-Jones [36, 37, 91]. Par la suite, cet axe de recherche n'a

∗. En anglais distribution quasi-stationnaire se dit quasi-stationary distribution d'où la notation QSD.

cessé de se développer (bibliographie sur les QSD mise en place par Pollett [87]). Toutes les propriétés générales sur les QSD peuvent être trouvées dans le survey de Méléard et Villemonais [84].

Espace d'état fini. Dans le cadre d'un espace d'état fini, Darroch et Seneta (1967) ont montré l'existence d'une unique distribution quasi-stationnaire ν et la convergence exponentielle de la loi μT_t du processus conditionné à la non-extinction vers ν, indépendamment de la distribution initiale μ.

Théorème 0.3 (Darroch et Seneta [37])**.** *Supposons que Λ soit fini et que le processus sur Λ de taux $\{Q_{x,y}, x, y \in \Lambda\}$ soit irréductible. Alors il existe $\theta > 0$ et $c > 0$ tels que*

$$\sup_{\mu \in \mathcal{M}_1(\Lambda)} d_{VT}(\mu T_t, \nu) \leq c e^{-\theta t}, \tag{4}$$

où $\mathcal{M}_1(\Lambda)$ désigne l'ensemble des mesures de probabilité sur Λ et pour tout $\mu, \mu' \in \mathcal{M}_1(\Lambda)$, $d_{VT}(\mu, \mu') = \frac{1}{2} \sum_{x \in \Lambda} |\mu(x) - \mu'(x)|$ la distance en variation totale.

Remarque 0.4. *Le paramètre θ apparaissant dans (4) est la distance entre la première et seconde valeur propre de Q.*

La preuve repose sur le Théorème de Perron-Frobenius qui donne une condition suffisante pour qu'une matrice admette une valeur propre de module maximal, de multiplicité 1.

Théorème 0.5 (Théorème de Perron-Frobenius)**.** *Soit Q une matrice carrée positive, irréductible et apériodique sur un espace d'état fini. Alors*

- *Il existe une valeur propre r réelle, strictement positive, de multiplicité 1, telle que pour toute autre valeur propre λ*

  $$|\lambda| < r.$$

  De plus, le sous-espace propre correspondant à la valeur propre r est de dimension 1.
- *A cette valeur propre maximale r correspond des vecteurs propres à gauche et à droite dont les coordonnées sont strictement positives.*

Se plaçant sous une hypothèse d'irréductibilité du processus de Markov, Diaconis et Miclo [45] ont récemment donné des estimations quantitatives de la convergence de la loi du processus conditionné à la non-absorption vers l'unique QSD, et ce quelque soit la distribution initiale. Pour cela, les auteurs réduisent l'étude de la convergence vers la distribution quasi-stationnaire à celle de la convergence d'un processus de Markov vers son état d'équilibre [45, Théorème 1]. Par la transformée de Doob, cette réduction d'étude est basée sur la seule connaissance du ratio $\dfrac{\max \varphi}{\min \varphi}$, où φ est la fonction propre associée à la matrice restreinte aux sites non-absorbants (l'existence de φ étant garantie par le Théorème de Perron-Frobenius) et pour lequel des bornes sont données dans [46].

Espace d'état dénombrable. Quand l'espace est dénombrable et contrairement aux processus de Markov irréductibles pour lesquels il y a au plus une distribution invariante, l'existence et l'unicité de QSD ne sont pas garanties. C'est le cas de la marche aléatoire simple $p-q$, étudié par Cavender [28], qui admet une infinité de QSD quand le drift est négatif ($q > p$) et aucune sinon. Sur l'espace des entiers naturels $\mathbb{N}$, les processus les plus étudiés sont les processus de naissance et mort [28, 47, 56, 84, 91, 98, 103]. En temps discret, ces processus ont été étudiés par Seneta et Vere-Jones [91] et Ferrari, Martínez et Picco [53]. En temps continu, Van Doorn [47] en donne une caractérisation complète : un processus de naissance et mort a 0, 1 ou une infinité de distributions quasi-stationnaires. En cas d'existence de plusieurs QSD, il y en a une parmi toutes les autres, dont le temps moyen d'extinction est minimale : elle est appelée **QSD minimale**.

Définition 0.6 (Distribution quasi-stationnaire minimale)**.** *Soit τ_0^μ le temps d'absorption du processus $(X_t)_{t\geq 0}$ de distribution initiale μ. Une QSD est dite minimale et est notée ν_{qs}^* si*

$$\mathbb{E}(\tau_0^{\nu_{qs}^*}) = \inf\{\mathbb{E}(\tau_0^\nu), \nu \textit{ vérifiant } (2)\}.$$

Si ν est une distribution quasi-stationnaire alors par la propriété de Markov, il existe une constante $\lambda(\nu) > 0$ telle que

$$\mathbb{P}_\nu(\tau_0 > t) = e^{-\lambda(\nu)t} \quad \forall t \geq 0.$$

Ainsi, partant d'une distribution quasi-stationnaire ν le temps d'absorption τ_0 suit une loi exponentielle de paramètre $\lambda(\nu)$ (dépendant de la QSD ν mais indépendant du temps t). Notamment, pour tout $0 < \alpha < \lambda(\nu)$

$$\mathbb{E}_\nu(e^{\alpha\tau_0}) < +\infty.$$

L'existence de moment exponentiel du temps d'absorption τ_0 est donc une condition nécessaire à l'existence de QSD. En particulier, puisque le temps d'extinction d'un processus de Galton-Watson critique vérifie $E_x(\tau_0) = +\infty$ pour tout $x > 0$, nous en déduisons que ce processus n'admet pas de distribution quasi-stationnaire.

En 1995, Ferrari, Kesten, Martínez et Picco montrent que si $\Lambda = \mathbb{N}$ et si $\lim_{x\to\infty} \mathbb{P}(\tau_0 < t \mid X_0 = x) = 0$ alors l'existence de moment exponentiel de τ_0 est également une condition suffisante pour l'existence d'une QSD [54, Théorème 1.1].

Théorème 0.7 (Ferrari, Kesten, Martínez et Picco [54])**.** *Supposons que $\Lambda = \mathbb{N}$ et que pour tout $t \geq 0$ $\lim_{x\to\infty} \mathbb{P}(\tau_0 < t \mid X_0 = x) = 0$ et pour tout $x \in \Lambda$ $\mathbb{P}_x(\tau_0 < \infty) = 1$. Alors une condition nécessaire et suffisante pour l'existence d'une QSD est*

$$\mathbb{E}_x(e^{\alpha\tau_0}) < \infty, \text{ pour un certain } x \in \mathbb{N}^* \text{ et } \alpha > 0.$$

Pour montrer ce théorème, Ferrari, Kesten, Martínez et Picco introduisent un processus de renouvellement sur $\mathbb{N}$: Partant d'une distribution initiale $\mu \in \mathbb{N}$, on considère un processus de Markov X^μ évoluant selon la dynamique de X jusqu'à ce qu'il touche l'état absorbant 0. A cet instant, il saute avec probabilité μ sur une nouvelle position dans $\mathbb{N}$. La Q-matrice est alors la matrice Q^μ donnée par (3). Se plaçant sous la condition $\mathbb{E}_\mu(\tau_0) < \infty$ où τ_0 est le temps d'absorption en 0, les auteurs considèrent la fonction $\Phi : \mu \mapsto \Phi(\mu)$ avec $\Phi(\mu)$ la mesure invariante du processus X^μ. Ils montrent que l'ensemble des points fixes de Φ n'est pas vide en montrant que les hypothèses du théorème du point fixe de Schauder s'appliquent (toute fonction continue d'un compact sur lui-même a un point fixe). D'autre part, les auteurs prouvent que les QSD sont des points fixes de Φ.

Sous la dynamique de ce processus, Jacka et Roberts ont montré l'existence et l'unicité d'une QSD [70, Proposition 4.4] sous la condition de Doeblin $\inf_{x\neq z} Q_{x,z} > \sup_{x\in\mathbb{N}} Q_{x,0}$ avec $\inf_{x\neq z} Q_{x,z} > 0$. En espace d'état fini, cette caractérisation se retrouve dans le papier d'Aldous, Flannery et Palacios [4] dans lequel les auteurs proposent une méthode de simulation de la QSD d'une chaîne de Markov à temps discret. L'idée principale est de remplaçer la mesure de redistribution μ du processus X^μ par la mesure d'occupation du processus. Les résultats de convergence obtenus ont récemment été améliorés par Benaïm et Cloez [16].

La notion de QSD pour un processus X_t est liée à l'étude du comportement en temps long du processus conditionné à la non-extinction. En effet (voir par exemple [84]), une mesure de probabilité ν est une distribution quasi-stationnaire si et seulement s'il existe $\mu \in \mathcal{M}_1(\Lambda)$ telle que

$$\lim_{t\to+\infty} \mathbb{P}_\mu (X_t \in \cdot \mid t < \tau_0) = \nu(\cdot).$$

En cas d'existence de QSD, on aimerait obtenir des résultats de convergence et des estimations de la vitesse de convergence. En 2012, pour un processus descendant de l'infini, Martínez, San Martín et Villemonais ont donné sur $\mathbb{N}$, un critère d'existence et d'unicité de QSD et ont montré la convergence exponentielle du processus conditionné vers l'unique QSD sous la distance en variation totale, et ce quelque soit la distribution initiale [81, Théorème 1]. Dans un cadre plus général, Champagnat et Villemonais [30] donnent des conditions nécessaires et suffisantes pour obtenir, sous la distance en variation totale, la convergence exponentielle vers une unique QSD, uniformément en la condition initiale.

Les QSD vérifient une équation non linéaire (équation (2)), elles sont donc difficilement simulables. Pour pallier à ce problème, Burdzy, Holyst, Ingermann et March ont proposé en 1996, une méthode d'approximation des QSD dans le cas des mouvements browniens sur un domaine borné [25]. Cette méthode est basée sur l'étude d'un système de particules en interaction, appelé *système de Fleming-Viot*.

0.3 Processus de Fleming-Viot

Soit $(X_t)_{t\geq 0}$ un processus de Markov à temps continu irréductible sur $\Lambda \cup \{0\}$ avec Λ un espace dénombrable ou fini et 0 un état absorbant et dont la Q-matrice est donnée par $Q = (Q_{x,y})$.

On considère N particules $X^1, \ldots, X^N$ évoluant de manière indépendante suivant la loi de $(X_t)_{t\geq 0}$ jusqu'à ce que l'une d'entre elles touche l'état absorbant 0. A cet instant, la particule absorbée choisit uniformément une autre particule et saute sur sa position. Entre les absorptions, chaque particule évolue de manière indépendante les unes des autres. Dans ce modèle, le nombre de particules reste constant : aucune particule ne se crée et aucune n'est détruite. Un tel système est appelé système de Fleming-Viot (FV). Malgré la même appellation, ce processus diffère de celui introduit par Fleming et Viot [57], mais ressemble plus au système de particules de type Moran [41, 42]. La dynamique du processus de Fleming-Viot est similaire à celle du processus de renouvellement introduit par Ferrari, Kesten, Martínez et Picco. La différence étant que les particules sont redistribuées selon la distribution empirique (dépendante du temps) et non pas selon la distribution initiale.

On peut considérer le système de Fleming-Viot en regardant les particules dans leur individualité c'est-à-dire en leur donnant à chacune une étiquette ou au contraire penser aux particules comme étant indistinguables et ne considérer que le nombre de particules en chaque élément de Λ que l'on appelera *site*. Soit η représentant le vecteur d'occupations avec pour tout $k \in \Lambda$, $\eta(k) = \eta^{(N)}(k)$ le nombre de particules au site k. Alors le processus $(\eta_t)_{t\geq 0}$ est un

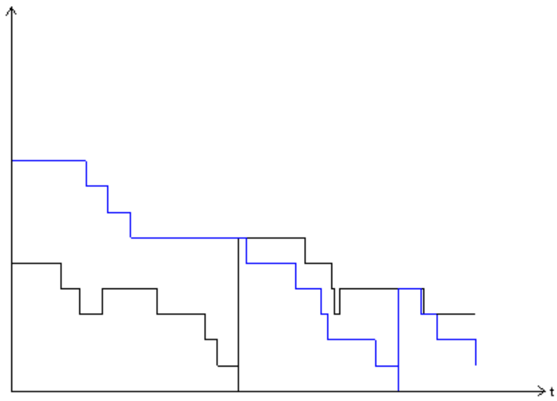

Figure 1 – Une trajectoire du système de Fleming-Viot avec deux particules.

processus de Markov d'espace d'état $E = E^{(N)}$ défini par

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

Le générateur † du processus de Fleming-Viot est donné, pour toute fonction bornée f, par

$$\mathcal{L}f(\eta) = \mathcal{L}^{(N)} f(\eta) = \sum_{i \in \Lambda} \eta(i) \left[\sum_{j \in \Lambda} \left(Q_{i,j} + Q_{i,0} \frac{\eta(j)}{N-1} \right) \left(f(T_{i \to j}\eta) - f(\eta) \right) \right], \tag{5}$$

pour tout $\eta \in E$, où, si $\eta(i) \neq 0$, $T_{i \to j}\eta$ est défini par $T_{i \to j}\eta = \eta$ si $i = j$ et pour $i \neq j$

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

Dans cette thèse, l'objectif est d'étudier le comportement asymptotique du processus. Plus précisément,

1. L'ergodicité du processus de Fleming-Viot à N fixé.
2. La convergence du processus de Fleming-Viot quand $N \to \infty$ et à temps fixé.

†. La définition d'un générateur est rappelée dans le chapitre 1.

3. La convergence de la densité empirique du processus de Fleming-Viot sous la mesure invariante.

Autrement dit, soit μ_t^N la distribution empirique du système de particules ‡ définie par

$$\mu_t^N = \frac{1}{N}\sum_{i=1}^{N}\delta_{X_t^i} = \frac{1}{N}\sum_{k\in\Lambda}\eta_t(k)\delta_{\{k\}}. \tag{6}$$

Sous quelles conditions les limites suivantes existent et peut-on (en cas d'existence) les quantifier ?

$$\begin{tikzcd}
& \mu_t^N \arrow[dl, "N\to\infty"'] \arrow[dr, "t\to\infty"] & \\
\mu T_t \arrow[dr, "t\to\infty"'] & & \mu_\infty^N \arrow[dl, "N\to\infty"] \\
& \nu &
\end{tikzcd}$$

Conjecture. *La mesure ν est l'unique distribution quasi-stationnaire minimale : $\nu = \nu_{qs}^*$.*

Le modèle introduit par Burdzy, Holyst, Ingermann et March (cas de mouvements browniens tués au bord d'un ouvert) pour répondre au problème d'approximation de la première fonction propre du Laplacien avec conditions de bord du type Dirichlet, a été étudié par Grigorescu et Kang [61] et Bieniek, Burdzy et Finch[17] dans le cadre d'un domaine lipschitzien. Cette étude a initialement été réalisée par Burdzy, Holyst et March [25, Théorèmes 1.3,1.4] mais les auteurs ont signalé que leur preuve était incomplète.

Expliquons pourquoi μT_t devrait être proche de μ_t^N. Supposons que μ soit proche de μ_0^N et posons pour tout $k \in \Lambda$, $v(k,t) = \mu T_t(k)$ et $u(k,t) = \mathbb{E}_\eta[\mu_t^N(k)]$ le semi-groupe de η. Alors,

$$\partial_t v(k,t) = \sum_{i\in\Lambda} Q_{i,k} v(i,t) + \sum_{i\in\Lambda} Q_{i,0} v(i,t) v(k,t), \tag{7}$$

et

$$\partial_t u(k,t) = \sum_{i\in\Lambda} Q_{i,k} u(i,t) + \sum_{i\in\Lambda} Q_{i,0} u(i,t) u(k,t) - \frac{Q_{k,0}}{N-1} u(k,t) + R_k(t), \tag{8}$$

‡. Dans le Chapitre 3 de cette thèse, la distribution empirique à l'instant t sera notée $m(\eta_t)$. Nous changeons la notation afin de pointer du doigt le fait qu'aucune étiquette n'est donnée aux particules, seul le vecteur d'occupation est pris en compte.

où

$$R_k(t) = \sum_{i\in\Lambda} Q_{i,0}\left(\frac{N}{N-1}\mathbb{E}_\eta(\mu_t^N(i)\mu_t^N(k)) - \mathbb{E}_\eta(\mu_t^N(i))\mathbb{E}_\eta(\mu_t^N(k))\right).$$

Les équations (7) et (8) sont alors similaires s'il existe une constance C telle que

$$\left|\mathbb{E}_\eta\left[\mu_t^N(i)\mu_t^N(k)\right] - \mathbb{E}_\eta\left[\mu_t^N(i)\right]\mathbb{E}_\eta\left[\mu_t^N(k)\right]\right| \le \frac{C}{N},$$

autrement dit, si les nombres d'occupations de deux sites distincts deviennent indépendants quand N tend vers l'infini. Cette asymptotique indépendance est appelée **propagation du chaos** et sera développée à la section 0.4.

Tout comme les QSD, de nombreuses études du processus de Fleming-Viot ont été menées, que ce soit en espace d'état fini [8] ou en espace d'état dénombrable [9, 55, 63]. Cependant, quand le nombre de QSD est infini, des problèmes restent encore ouverts. C'est le cas du processus de Fleming-Viot dont les particules suivent la dynamique de la marche aléatoire simple $p - q$ avec une dérive vers l'origine ($q > p$). Ce processus de Fleming-Viot s'interprète comme un système de N files d'attente M/M/1 en interaction : quand une file se vide, on duplique l'une des autres files [10, 39].

Espace d'état fini. Dans le cadre d'un espace d'état fini, nous avons vu qu'il y avait unicité de la distribution quasi-stationnaire. Pour $N \ge 2$, le processus de FV étant un processus de Markov irréductible, il est ergodique et admet donc une unique mesure invariante. Sous cette mesure invariante, Asselah, Ferrari et Groisman ont montré dans [8] la convergence de la distribution empirique du FV vers l'unique QSD. La démonstration repose sur le contrôle des corrélations entre deux particules et la borne obtenue dans ce cas est uniforme sur l'ensemble des distributions initiales.

Espace d'état dénombrable. Tout comme l'existence et l'unicité des QSD, l'ergodicité du processus de Fleming-Viot n'est plus garantie en espace dénombrable. En 2007, pour des processus satisfaisant la condition de Doeblin donnée par

$$\sum_{z\in\Lambda} \inf_{x\in\Lambda\setminus\{z\}} Q_{x,z} > \sup_{x\in\Lambda} Q_{x,0},$$

Ferrari et Marić ont montré l'ergodicité du processus de Fleming-Viot et la convergence sous la mesure invariante de sa densité empirique vers une unique QSD (quand N tend vers l'infini)[55]. Comme pour le cas fini, l'élément clé de la preuve est le contrôle des corrélations entre deux particules. Plus récemment et sous des hypothèses plus faibles, nous avons obtenu avec Bertrand Cloez [34] ces mêmes résultats de convergence mais avec des estimations quantitatives, généralisant ainsi ceux de Ferrari et Marić (Chapitre 3). Notre originalité réside dans l'utilisation du couplage pour obtenir un taux explicite et sous une certaine distance de Wasserstein, de la convergence exponentielle du processus de Fleming-Viot vers son état d'équilibre. Si l'on s'intéresse au processus conditionné, Martínez, San Martín et Villemonais donnent une condition suffisante pour l'existence et l'unicité d'une QSD

$$\inf_{x\in\mathbb{N}^*\setminus\{K\}} \left(Q(x,0) + \sum_{z\in K} Q_{x,z} \right) > \sup_{x\in\mathbb{N}^*} Q_{x,0},$$

où K est un sous-ensemble fini de $\mathbb{N}^*$. Ils montrent ainsi la convergence exponentielle de la loi du processus conditionné vers celle-ci avec un taux explicite de convergence [81, Théorème 3], généralisant le résultat de Ferrari et Marić. Si l'on compare maintenant cette condition avec celle obtenue avec Bertrand Cloez, elle est plus faible et plus générale. Cependant, nos résultats donnent des informations quand t est petit. Lorsque nos conditions sont satisfaites, celle de Martínez, San Martín et Villemonais également, mais notre taux de convergence est plus explicite que la leur.

Pour les processus admettant une infinité de QSD, la question est de savoir vers quelle QSD le processus de Fleming-Viot converge (si convergence il y a). A notre connaissance, il existe uniquement que deux processus pour lesquelles une preuve de la convergence vers la QSD minimale a été fournie. Ce phénomène est appelé *principe de sélection* dans le sens où le FV "sélectionne" la QSD minimale parmi toutes les QSD. Le premier processus est celui du Galton-Watson sous-critique pour lequel, Asselah, Ferrari, Groisman et Jonckheere [9] ont montré l'ergodicité du Fleming-Viot associé ainsi que la convergence de la distribution empirique à l'équilibre vers la QSD minimale. Un résultat similaire a récemment été démontré par Villemonais [103] pour quelques processus de naissance et mort et pour lequel l'élément clé de l'ergodicité est l'existence d'une **fonction de Lyapunov**. Cependant, les arguments utilisés dans ce papier ne s'appliquent pas dans le cas des marches aléatoires.

Les systèmes de Fleming-Viot présentent une forte similitude avec ceux introduits pour l'étude des équations de McKean-Vlasov. En effet, dans les deux cas, l'interaction d'une particule avec toutes les autres particules a lieu par le biais de la distribution empirique du système. On parle alors de système de particules en interaction de type champ moyen.

0.4 Systèmes en interaction de type champ moyen

Définition 0.8. *On dit qu'un système de particules est en interaction de type champ moyen lorsque l'interaction a lieu à travers la mesure empirique.*

L'étude de ces systèmes trouve ses racines en physique et plus particulièrement en mécanique statistique avec Kac [72] puis McKean [82] afin de modéliser les collisions entre particules dans un gaz. Cette approche de type champ moyen s'est par la suite développée dans divers domaines tels que la biologie [24, 41, 42] ou les réseaux [22, 39]. L'état du système à un instant t est donné par le N-uplet $(X_t^{1,N}, \dots, X_t^{N,N})$ où $X^{i,N}$ représente la position de la $i^{ème}$ particule pour tout $1 \leq i \leq N$. Quand N devient grand, il devient impossible de regarder le comportement de chacune des particules, seul le comportement moyen est observable. On en vient alors à considérer des quantités de la forme $\frac{1}{N}\sum_{i=1}^{N} \varphi(X_t^{i,N})$ pour φ une fonction test, obtenues à partir de la mesure empirique μ_t^N du système définie par (6). L'état du système peut donc être entièrement décrit par sa mesure empirique. Se pose alors la question de la convergence de cette mesure quand le nombre de particules N tend vers l'infini.
Dès que le système est en interaction, la dynamique d'une particule donnée n'est jamais indépendante de celle des autres particules. Une propriété d'indépendance ne peut donc être valable qu'à la limite, on parle alors de **chaos** : pour tout k fixé, la loi de k particules parmi N tend vers la loi de k particules indépendantes identiquement distribuées quand N tend vers l'infini. On dit qu'il y a **propagation du chaos** si le caractère chaotique d'un système de particules initiales est préservée au cours du temps : si les particules initiales $X_0^{i,N}$ sont indépendantes identiquement distribuées de loi u_0 alors pour tout k fixé, la loi $\mu_t^{(k)}$ du k-uplet $(X_t^{1.N}, \dots, X_t^{k,N})$ converge au sens de la topologie faible des mesures vers la loi $u_t^{\otimes k}$ de k particules indépendantes de même loi u_t quand N vers l'infini. Sznitman [95, Proposition 2.2] (voir aussi Méléard [83, Proposition 4.2]) a montré que dans le cas de particules échangeables, la propagation du chaos est équivalente à la convergence en loi de la mesure empirique du système vers la mesure déterministe u_t. Autrement dit, pour toute fonction ϕ continue bornée sur l'espace des mesures de probabilité, muni de la topologie faible, on a

$$\mathbb{E}\phi(\mu_t^N) \underset{N\to+\infty}{\longrightarrow} \phi(u_t).$$

Cette limite fréquemment appelée *limite de champ moyen*, est caractérisée comme étant l'unique solution faible d'une équation aux dérivées partielles non linéaire de la forme

$$\frac{d}{dt}\langle u_t, f\rangle = \langle u_t, \mathcal{G}_{u_t} f\rangle, \tag{9}$$

où $\mathcal{G}_{(\cdot)}$ est un opérateur défini par le comportement du système de particules, f est une fonction continue bornée sur l'espace d'état Λ et où pour toute mesure de probabilité u

$$\langle u, f\rangle = \sum_{k\in\Lambda} f(k)u(\{k\}).$$

La notion de propagation du chaos, introduite par Kac [72] pour l'étude de l'équation de Boltzmann, jouit donc de la double propriété de donner le comportement asymptotique d'un système de particules en interaction et de donner une approximation des solutions d'une équation aux dérivées partielles non linéaire. Depuis, ce phénomène de propagation du chaos a été étudié pour divers modèles [39, 83, 95](et les références se trouvant à l'intérieur).

La limite $\lim\limits_{N\to+\infty} \mu_t^N$ étant une mesure dépendante du temps, on peut étudier son comportement quand t tend vers l'infini. Considérons maintenant le comportement en temps long du système i.e $\lim\limits_{t\to+\infty} \mu_t^N$. Quand N est grand, obtient-on un résultat proche de l'asymptotique de u_t ? Cela revient alors à se demander si les limites peuvent s'intervertir :

$$\lim_{t\to\infty}\lim_{N\to\infty} \mu_t^N = \lim_{N\to\infty}\lim_{t\to\infty} \mu_t^N ?$$

ou encore s'il existe une mesure ν tel que le diagramme suivant soit commutatif

$$\begin{tikzcd}
& \mu_t^N \arrow[dl, "N\to\infty"'] \arrow[dr, "t\to\infty"] & \\
u_t \arrow[dr, "t\to\infty"'] & & \mu_\infty^N \arrow[dl, "N\to\infty"] \\
& \nu &
\end{tikzcd}$$

Exemples d'équations non linéaires.

a) Dans le cas des diffusions ($\Lambda = \mathbb{R}^d$), les équations de McKean-Vlasov, équations issues de la théorie cinétique des gaz, font parties des équations non linéaires les plus étudiées [38, 59, 77, 78, 82, 83]. L'équation des milieux granulaires, notamment étudiée par Malrieu, définie par

$$\frac{\partial u}{\partial t} = \operatorname{div}\left(\nabla u + u\left(\nabla V + \nabla W * u\right)\right), \tag{10}$$

en est un exemple. Ici, $*$ désigne la convolution en la variable d'espace et V et W sont des fonctions convexes régulières.

b) En espace dénombrable, l'opérateur $\mathcal{G}_{(\cdot)}$ défini en (9) peut représenter la dynamique d'un processus de naissance et mort dont un exemple est donné pour $i \in \Lambda$ par

$$\mathcal{G}_{\|u_t\|} f(i) = b_i(f(i+1) - f(i)) + (d_i + \|u_t\|)(f(i-1) - f(i)),$$

où $\|u\|$ représente le premier moment de u.

D'après la formule d'Itô, on peut associer à l'équation non linéaire (9) un processus dont la loi marginale au temps t est solution de (9). Une étude naturelle est celle de son comportement en temps long. Intéressons-nous plus particulièrement au cas des diffusions (le cas discret étant traité par la suite) et revenons à l'exemple des équations de McKean-Vlasov. Dans ce cas, le processus associé à l'équation (10) évolue selon l'équation différentielle stochastique

$$\left\{\begin{array}{rcl} d\overline{X}_t & = & \sqrt{2}dB_t - \nabla V(\overline{X}_t)dt - \nabla W * u_t(\overline{X}_t)dt, \\ \mathrm{Law}(\overline{X}_t) & = & u_t(x)dx, \end{array}\right.$$

où $(B_t)_{t\geq 0}$ est un mouvement brownien à valeurs dans $\mathbb{R}^d$ et $\mathrm{Law}(X)$ représente la loi de X. Dans l'interprétation physique, les potentiels V et W représentent respectivement des potentiels extérieur et d'attraction. Ce processus est dit non linéaire dans la mesure où sa loi intervient dans les coefficients à travers le terme de convolution. Pour étudier ce processus, on introduit un système de particules en interaction de type champ moyen $(X_t^{1,N}, \ldots, X_t^{N,N})$ où pour tout $1 \leq i \leq N$, $X_t^{i,N}$ vérifie l'équation différentielle stochastique

$$\left\{\begin{array}{rcl} dX_t^{i,N} & = & \sqrt{2}dB_t^{i,N} - \nabla V(X_t^{i,N})dt - \dfrac{1}{N}\displaystyle\sum_{j=1}^{N} \nabla W(X_t^{i,N} - X_t^{j,N})dt, \\ X_0^{i,N} & = & X_0^i, \end{array}\right.$$

où les $(B_t^{i,N})$ sont N mouvements browniens indépendants, les variables aléatoires $(X_0^i)_{i\in\{1,\ldots,N\}}$ sont indépendantes identiquement distribuées et indépendantes des $B_\cdot^i$.

Remarque 0.9. *Les particules sont dirigées par des mouvements browniens indépendants mais* *en interaction par leur mesure empirique. En effet,*

$$\frac{1}{N}\sum_{j=1}^{N}\nabla W(X_t^{i,N}-X_t^{j,N})=\frac{1}{N}\sum_{j=1}^{N}\nabla W*\mu_t^N(X_t^{i,N}).$$

Afin d'établir le phénomène de propagation du chaos, une des méthodes utilisées est celle du couplage consistant à introduire une dynamique à N particules indépendantes et de montrer que les trajectoires du modèle de départ sont proches des trajectoires du modèle auxiliaire. Dans l'exemple considéré, cela revient à introduire des processus indépendants $(\overline{X}_t^{i,N})_{t\geq 0}$ de loi u_t en chaque instant t (où u_t est la loi du processus $(\overline{X}_t)_{t\geq 0}$ solution de $d\overline{X}_t=\sqrt{2}dB_t-\nabla V(\overline{X}_t)dt-\nabla W*u_t(\overline{X}_t)$ dt), tels que pour $i=1,\ldots,N$

$$\begin{cases} d\overline{X}_t^{i,N} &= \sqrt{2}dB_t^{i,N}-\nabla V(\overline{X}_t^{i,N})dt-\nabla W*u_t(\overline{X}_t^{i,N})dt,\\ \overline{X}_0^{i,N} &= X_0^{i,N},\end{cases}$$

où $(B_t^{i,N})_{t\geq 0}$ est le mouvement brownien dirigeant l'évolution de $X_t^{i,N}$ pour chaque i.

De ce couplage, Malrieu donne pour la première fois un résultat de propagation de chaos uniforme en temps [77, Théorème 3.3]

$$\sup_{t\geq 0}\mathbb{E}\left(|X_t^{i,N}-\overline{X}_t^{i,N}|^2\right)\leq\frac{C}{N}, \tag{11}$$

basée sur la formule d'Itô et des hypothèses de convexité des potentiels.

En notant $\mathcal{W}_1$ la distance de Wasserstein définie, pour μ et μ' deux mesures de probabilités, par

$$\mathcal{W}_1(\mu,\mu')=\inf_{\substack{X\sim\mu\\ Y\sim\mu'}}\mathbb{E}|d(X,Y)|,$$

l'infimum étant pris sur les couples de variables aléatoires de lois marginales μ et μ', et d étant la distance euclidienne sur $\mathbb{R}^d$, l'inégalité (11) assure, avec des taux explicites

1. La convergence faible de la loi d'une particule vers u_t : en notant $u_t^{1,N}$ la première marge de la loi du système

$$\mathcal{W}_1(u_t^{1,N},u_t)\leq\frac{C}{\sqrt{N}}.$$

2. La propagation du chaos du système de particules : pour k fixé

$$\mathcal{W}_1(\mu_t^{(k)},u_t^{\otimes k})\leq\frac{kC}{\sqrt{N}},$$

où $\mu_t^{(k)}$ représente la loi de k particules parmi N.

3. La convergence de la mesure empirique μ_t^N vers la loi u_t : sous des hypothèses d'existence de moments pour u_t

$$\sup_{\|\varphi\|_{Lip}\leq 1} \mathbb{E}\Big|\frac{1}{N}\sum_{i=1}^N \varphi(X_t^{i,N}) - \int \varphi du_t\Big| \leq \frac{K}{\sqrt{N}},$$

où $\|\varphi\|_{Lip}$ est donné par

$$\|\varphi\|_{Lip} = \sup_{x\neq y} \frac{|\varphi(x)-\varphi(y)|}{d(x,y)}.$$

Ce résultat de propagation du chaos est également un élément important pour obtenir des inégalités de déviations et plus particulièrement des estimations quantitatives de

$$\sup_{\|\varphi\|_{Lip}\leq 1} \mathbb{P}\left[\Big|\frac{1}{N}\sum_{i=1}^N \varphi(X_t^{i,N}) - \int \varphi du_t\Big| > \epsilon\right] \text{ en fonction de N.}$$

De tels résultats ont été obtenus par Malrieu [77] : Sous des hypothèses de convexité des potentiels (stricte convexité pour le potentiel V) et basée sur le critère de Bakry-Emery, il montre que la loi de X_t^N satisfait une inégalité de Sobolev logarithmique de constante γ indépendante de t et N. Via l'argument de Herbst, l'inégalité de Sobolev permet alors de montrer que la loi du système satisfait une inégalité de concentration gaussienne autour de sa moyenne. En particulier, pour tout $\epsilon > 0$

$$\sup_{\|\varphi\|_{Lip}\leq 1} \mathbb{P}\Big[\Big|\frac{1}{N}\sum_{i=1}^N \varphi(X_t^{i,N}) - \mathbb{E}\left(\frac{1}{N}\sum_{i=1}^N \varphi(X_t^{i,N})\right)\Big| > \epsilon\Big] \leq 2\exp\left(-\frac{\gamma}{2}N\epsilon^2\right).$$

Combinant ce résultat avec celui de la propagation du chaos, Malrieu obtient (cf. [77, Corollaire 3.9])

$$\sup_{\|\varphi\|_{Lip}\leq 1} \mathbb{P}\left[\Big|\frac{1}{N}\sum_{i=1}^N \varphi(X_t^{i,N}) - \int \varphi du_t\Big| > \epsilon + \frac{C}{\sqrt{N}}\right] \leq 2\exp\left(-\frac{\gamma}{2}N\epsilon^2\right).$$

Quelques années plus tard, Bolley, Guillin et Villani [21] donnent un résultat de déviation plus fort et estiment

$$\mathbb{P}(\sup_{0\leq t\leq T} \mathcal{W}_1(\mu_t^N, u_t) > \epsilon).$$

Sous certaines conditions, incluant une condition lipschitzienne sur la force d'interaction ∇W, d'intégrabilité sur la mesure initiale u_0 et des conditions sur la taille N du système de particules,

les auteurs montrent [21, Théorème 1.7] qu'il existe des constantes K, $C(T,\epsilon)$ telles que

$$\mathbb{P}(\sup_{0\le t\le T} \mathcal{W}_1(\mu_t^N, u_t) > \epsilon) \le C(T,\epsilon)\exp\left(-KN\epsilon^2\right).$$

Pour cela, Bolley, Guillin et Villani reprennent le couplage précédemment défini en introduisant un système de particules indépendants $(\overline{X}_t^{i,N})_{1\le i\le N}$ de même conditions initiales que X_t^N et dont chaque particule $\overline{X}_t^{i,N}$ est dirigée par le mouvement brownien de $X_t^{i,N}$, $1\le i\le N$. Ils réduisent ainsi le problème de convergence de μ_t^N vers u_t à un problème de convergence de ν_t^N vers u_t où $\nu_t^N = \dfrac{1}{N}\sum_{i=1}^N \delta_{\overline{X}_t^{i,N}}$ est la densité empirique du système $(\overline{X}_t^{i,N})_{1\le i\le N}$. Après avoir montré que u_t vérifie une inégalité de transport T_1 et en utilisant par la suite un argument de continuité, ils obtiennent la décroissance exponentielle de $\mathbb{P}(\sup_{0\le t\le T} \mathcal{W}_1(\nu_t^N, u_t) > \epsilon)$.

0.5 Modèles étudiés et principaux résultats

Processus de Fleming-Viot avec dérive constante

Considérons N marches aléatoires simples en temps continu sur $\mathbb{N}$ avec une dérive constante vers l'origine (modèle $p-q$ avec $q>p$). Le processus de Fleming-Viot qui lui est associé, que l'on notera FVRW, est le premier modèle auquel nous nous sommes intéressés et le premier résultat obtenu est celui de l'ergodicité (via le critère de Foster).

Théorème 0.10. *Il existe des constantes strictement positives $K, \alpha, \delta_0, A, c_1, c_2$ telles que pour tout $N\in\mathbb{N}$, pour $T = A\log(N)$ et pour tout $\delta<\delta_0$ et $x\in\mathbb{N}^N$, on ait*

$$E\left[\exp\left(\delta\max(X_T)\right)\Big|X_0 = x\right] - \exp\left(\delta\max(x)\right) < -\,c_1\,\mathbb{1}_{\max(x)>K\log(N)}e^{\delta\max(x)} + c_2e^{\delta\alpha\log(N)}. \tag{12}$$

Par conséquent, pour N assez grand il existe une unique mesure invariante λ_N pour le processus de Fleming-Viot. En intégrant par rapport à λ_N, il existe $C>0$ tel que pour tout N et $\delta<\delta_0$

$$\int \exp(\delta\max(x))d\lambda_N(x) \;\le\; C\exp\left(\delta\alpha\log(N)\right). \tag{13}$$

Chercher une fonction de Lyapunov peut-être hasardeux, donc pour établir le Théorème 0.10, nous nous sommes inspirés de la stratégie établie par Asselah, Ferrari, Groisman et Jonckheere [9, Proposition 1.2]. L'idée est de plonger le processus de Fleming-Viot dans un processus de branchement multitype dont le nombre d'individus croît avec le temps mais dont les branchements sont indépendants des positions. Pour une description détaillée du processus de branchement, le lecteur pourra aller voir la section 3 de [9]. Cette stratégie est d'autant plus intéressante qu'elle entraîne l'ergodicité du processus et un contrôle du maximum autrement dit de la marche la plus à droite.

Plus précisément, considérons N marches aléatoires évoluant de manière indépendante les unes des autres. Supposons qu'à l'instant initial le maximum se trouve à une hauteur supérieure à $3L$ (L fixée) et observons le processus sur une période de temps $[0, T]$. Pour visualiser l'évolution du processus, nous divisons la population des marches en 2 :

- les marches rouges qui sont loin de l'origine (supérieures à L),
- les marches noires proches de l'origine (inférieures à L).

La dynamique des marches est la suivante : Quand une marche noire touche l'origine, elle saute sur la position d'une des marches rouges ou d'une des marches noires. Si elle saute sur une marche rouge alors celle-ci donne naissance à deux nouvelles marches rouges. Autrement dit, quand une marche noire saute sur une rouge elle devient elle-même rouge. Mais si elle saute sur une marche noire alors on aura deux nouvelles marches noires. Le principe est similaire quand une marche rouge touche l'origine. Les marches rouges et noires ne sont donc pas indépendantes. Pour créer de l'indépendance on introduit une autre catégorie de marches, les marches vertes. Elles se comportent de manière identique aux rouges tant qu'elles ne touchent pas 0 c'est-à-dire que les deux populations se superposent. Cependant quand une rouge (et donc une verte) touche l'origine, la rouge saute, tandis que la verte associée se déplace sur $\mathbb{Z}$ sans saut en 0. Les marches vertes sont donc des marches aléatoires indépendantes dont les temps de branchement sont indépendants de leurs positions.

Pour que l'inégalité (12) soit vraie, il faudrait qu'au temps T, le maximum (qui se trouve parmi les marches rouges ou vertes) ait décru. Par exemple, on aimerait qu'à l'instant T ce maximum se trouve dans l'intervalle $[2L, \frac{5L}{2}]$. Trois évènements peuvent mettre à mal cette décroissance :

- Une marche rouge touche 0 avant l'instant T.
- Une marche noire dépasse $2L$ i.e dépasse le maximum des rouges.
- Une marche verte ne descend pas assez (i.e descend de beaucoup moins que sa distance usuelle en un temps T).

En utilisant une méthode de premier moment inspirée par Zeitouni [107] et des propriétés de grandes déviations de marche aléatoire simple, nous montrons que les mauvais évènements se produisent avec une probabilité négligeable, et qu'en dehors d'eux le maximum décroît.

La marche aléatoire simple sur $\mathbb{N}$ avec dérive vers l'origine admet une famille de QSD (Cavender [28]), dont les expressions sont données dans le chapitre 2 ou dans [28, 80]. Une conjecture a été établie dans [10, 63, 79] dans laquelle, quand N tend vers l'infini, le processus de Fleming-Viot à l'équilibre converge vers la QSD minimale. Aucune preuve générale n'a pu être fournie, mais les simulations établies par Marić [80] confirment cette conjecture.

Autour du graphe complet

Parmi toutes les situations traitées liées au processus de Fleming-Viot, l'existence de la mesure invariante a, en général, toujours pu être prouvée mais son expression n'est jamais explicite. De plus, quand il y a convergence vers l'équilibre, le taux de convergence n'est généralement pas explicite. Il était donc intéressant de se demander s'il existe un modèle pour lequel ce taux le serait : c'est le cas du graphe complet. On considère N particules évoluant uniformément sur $K+1$ sites $\{0, 1, \ldots, K\}$. Quand une particule touche le site 0, elle saute instantanément vers la position d'une autre particule. En notant $\eta(i)$ le nombre de particules au site i, pour tout $i \in \{1, \ldots, K\}$, la dynamique d'une particule est la suivante :

$$i \to j \text{ avec un taux } \frac{1}{K} + \frac{1}{K}\frac{\eta(j)}{N-1}.$$

Le graphe complet est un modèle très particulier car tous les sites qui le composent sont symétriques c'est-à-dire jouent tous le même rôle. Du fait de sa géométrie, beaucoup de calculs sont rendus accessibles tels que celui de la mesure invariante ou encore celui des corrélations. Pour étudier le comportement en temps long du processus de Fleming-Viot associé, on introduit deux distances :

- La distance en variation totale : Pour $\eta, \eta' \in E = \left\{\eta : \{1, \ldots K\} \to \mathbb{N} \mid \sum_{i=1}^{K} \eta(i) = N\right\}$, $d(\eta, \eta') = \frac{1}{2}\sum_{j=1}^{K} |\eta(j) - \eta'(j)|$.
- La distance de Wasserstein : $\mathcal{W}_d(\mu, \mu') = \inf \mathbb{E}\left[d(\eta, \eta')\right]$ où l'infimum est pris sur tous les couples de lois marginales μ et μ'.

Ces distances sont basées sur l'idée de couplage. Autrement dit, étant données deux mesures μ et ν, comment peut-on fabriquer deux variables aléatoires X et Y de lois respectives μ, ν afin de minimiser $\mathbb{E}|X-Y|$ ou $\mathbb{P}(X \neq Y)$? Dans le cas du graphe complet, une illustration du couplage est visible dans les figures 2,3 et détaillé dans le chapitre 3 : On considère deux systèmes de particules de type Fleming-Viot, disons η et η', de configurations initiales différentes (sur les figures, cela est représenté par les boules rouges et les carrés bleus). On sélectionne une particule de chaque configuration (boule et carré verts), on appellera alors paire de particules ou couple l'association d'une particule de la configuration η avec celle de η'. Un couplage est réussi lorsqu'au cours du temps le nombre de particules non couplées tend vers 0, autrement dit lorsqu'il y a de plus en plus de paires de particules qui sont sur les mêmes sites. Dans le système de Fleming-Viot le nombre de paires n'est pas croissant comme l'illustre la figure 2 : les particules vertes sont sur le même site et si elles sont envoyées (avec probabilité $\frac{1}{K}$) sur l'origine (site 0), elles peuvent aboutir sur deux sites l et m distincts. Au contraire, comme le montre la figure 3, l'absence d'interaction va en favoriser l'augmentation : avec probabilité $\frac{1}{K}$, les particules peuvent sauter sur un même site, distinct de leurs sites initiaux. Ou avec probabilité $\frac{1}{K}$, une particule peut sauter sur la position de l'autre.

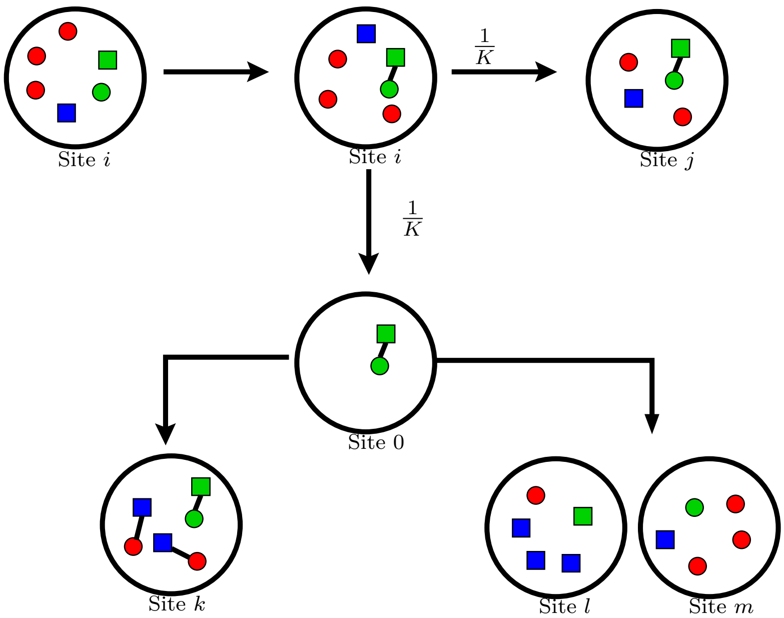

Figure 2 – Cas où les particules sélectionnées partent du même site.

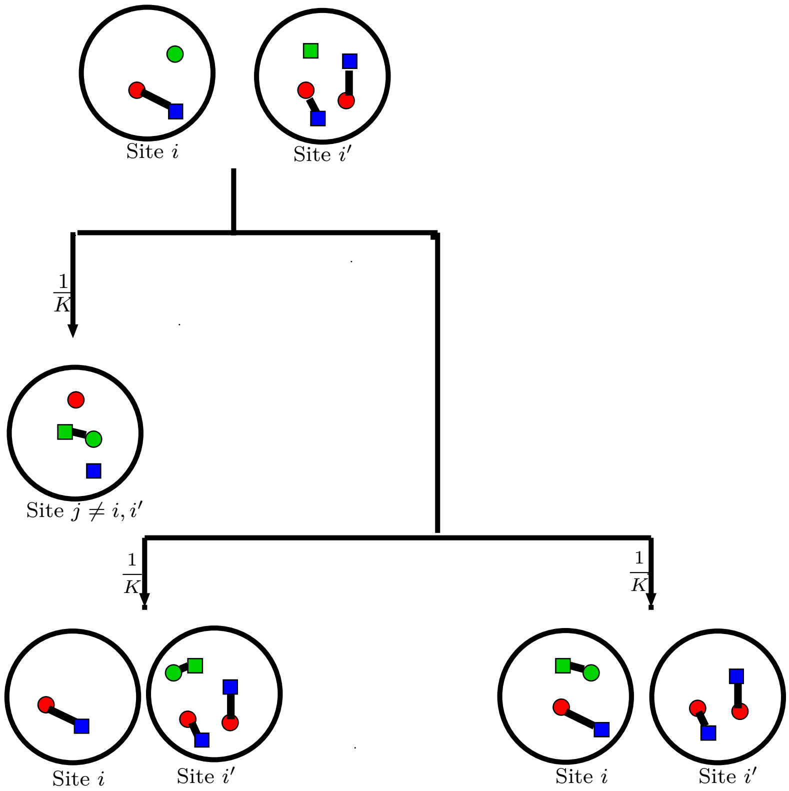

Figure 3 – Cas où les particules sélectionnées partent de site différents et ne meurent pas

L'approche proposée reste valable au delà du graphe complet pour une certaine classe de Fleming-Viot discrets. On se place sur un espace dénombrable Λ et on considère un processus de Markov dont les taux de sauts vérifient

$$\forall i \neq j \in \Lambda,\ Q_{i,j} > 0 \text{ et } \forall i \in \Lambda,\ Q_{i,0} > 0.$$

Théorème 0.11 (Convergence exponentielle sous la distance de Wasserstein)**.** *Soit*

$$\lambda = \inf_{i,i' \in \Lambda} \left(Q_{i,i'} + Q_{i',i} + \sum_{j \neq i,i'} Q_{i,j} \wedge Q_{i',j} \right) \text{ et } p_0 : i \mapsto Q_{i,0}, \forall i \in \Lambda.$$

Pour tous processus $(\eta_t)_{t>0}$ *et* $(\eta'_t)_{t>0}$ *générés par* (5)*, et pour tout* $t \geq 0$,

$$\mathcal{W}_d(\mathrm{Law}(\eta_t), \mathrm{Law}(\eta'_t)) \leq e^{-\rho t} \mathcal{W}_d(\mathrm{Law}(\eta_0), \mathrm{Law}(\eta'_0)),$$

avec $\rho = \lambda - (\sup(p_0) - \inf(p_0))$*. En particulier, si* $\rho > 0$ *alors il existe une unique distribution* *invariante ν_N satisfaisant pour tout $t \geq 0$,*

$$\mathcal{W}_d(\mathrm{Law}(\eta_t), \nu_N) \leq e^{-\rho t}\mathcal{W}_d(\mathrm{Law}(\eta_0), \nu_N).$$

Ce résultat, obtenu en collaboration avec Bertrand Cloez, montre qu'en moyenne, si l'on considère deux nuages de particules suivant la dynamique du Fleming-Viot alors ces nuages se rapprochent avec le temps. Le paramètre λ défini dans le Théorème 0.11 donne la "chance" de pouvoir coupler en un instant. Dans le cas du graphe complet $\rho = \lambda = 1$ et le résultat est optimal en terme de contraction. Par contre dans le cas d'un processus de naissance et mort $\lambda = 0$.

A notre connaissance c'est la première fois qu'un tel résultat de convergence du processus de Fleming-Viot avec un taux explicite a pu être établi. Ce résultat s'avèrera être un élément important pour montrer la décroissance des corrélations avec N. Mais la clé de sa démonstration est la relation de commutation entre l'opérateur du *carré du champ* et le *semi-groupe* de η

$$\mathrm{Var}_\eta(g(\eta_t)) = P_t(g^2)(\eta) - (P_t g)^2(\eta) = \int_0^t P_s \Gamma P_{t-s} g(\eta) ds. \tag{14}$$

Contrairement aux articles [55] et [8], la borne obtenue est uniforme en temps quand ρ est strictement positif. Sous cette hypothèse de strict positivité de ρ, découle deux conséquences importantes du Théorème 0.11 :

1. L'existence et l'unicité d'une QSD et la convergence exponentielle de la loi du processus conditionné vers celle-ci.
2. La convergence de la mesure invariante du système de particules vers la QSD.

Dans le cas particulier de l'espace à deux points, que nous développons dans le chapitre 3, l'utilisation de ce couplage nous permet d'obtenir le trou spectral comme taux de convergence. Expliciter ce trou spectral est difficile mais en utilisant les inégalités de Hardy et en mimant les arguments de Miclo [86], nous montrons que quelque soit la valeur des paramètres, le trou spectral est toujours minoré par une constante strictement positive ne dépendant pas de N.

Processus de naissance et mort en interaction de type champ moyen

Considérons un système de N particules $X^{1,N}, \ldots, X^{N,N}$ sur $\mathbb{N} = \{0, 1, 2 \ldots\}$, chaque particule évoluant selon un processus de naissance et mort en interaction de type champ moyen. Autrement dit, une particule évolue avec un taux dépendant de sa position et de la moyenne des positions. En notant pour tout $i \in \mathbb{N}$, b_i et d_i les taux de naissance et mort, $q^+, q^- : \mathbb{N}\times\mathbb{R}_+ \to \mathbb{R}_+$

les taux d'interaction de vie et de mort et M^N la moyenne des positions, la dynamique d'une particule à un instant t est la suivante :

$$\begin{array}{lllll} i & \to & i+1 & \text{avec taux } b_i + q^+(i, M_t^N), & \\ i & \to & i-1 & \text{avec taux } d_i + q^-(i, M_t^N), & \text{pour } i \geq 1 \\ i & \to & j & \text{avec taux } 0, & \text{si } j \notin \{i-1, i+1\}. \end{array}$$

Ce modèle, pouvant être vu comme une version discrète de celui introduit pour l'étude des équations de McKean-Vlasov, s'avère être une bonne approximation de l'équation non linéaire (9) avec pour $i \in \mathbb{N}$

$$\mathcal{G}_{\|u_t\|} f(i) = \left(b_i + q^+(i, \|u_t\|)\right)(f(i+1) - f(i)) + \left(d_i + q^-(i, \|u_t\|)\right)(f(i-1) - f(i))\mathbb{1}_{i>0}.$$

Autrement dit, la mesure empirique aléatoire μ_t^N des N particules converge, au sens faible des distributions, vers la mesure déterministe u_t solution de l'équation (9) au temps t. Une question importante est le comportement à l'infini de u_t : existence d'une mesure stationnaire et vitesse de convergence vers cette mesure ou encore distance entre deux solutions de l'équation (9) partant de conditions différentes.
Sous des hypothèses de convexité sur les taux de naissance et mort

- Il existe $\lambda > 0$ tel que
$$\nabla^+(d - b) \geq \lambda,$$
où pour tout $n \geq 0$ et $f : \mathbb{N} \to \mathbb{R}_+$
$$\nabla^+(f)(n) = f(n+1) - f(n),$$

et des conditions lipschitziennes sur les taux d'interaction,

- Pour tout $(k_1, l_1), (k_2, l_2) \in \mathbb{N} \times \mathbb{R}_+$
$$\left| \left(q^+ - q^-\right)(k_1, l_1) - \left(q^+ - q^-\right)(k_2, l_2) \right| \leq \alpha \left(|k_1 - k_2| + |l_1 - l_2|\right),$$

nous obtenons la convergence exponentielle en temps de u_t vers une mesure limite u_∞ sous la distance de Wasserstein $\mathcal{W}_1$ (distance sur $\mathbb{R}$ associée à $|\cdot|$)

Théorème 0.12. *Soient $(u_t)_{t\geq 0}$ et $(v_t)_{t\geq 0}$ les solutions de (9) de conditions initiales respectives u_0 et v_0. Alors*
$$\mathcal{W}_1(u_t, v_t) \leq e^{-(\lambda - 2\alpha)t} \mathcal{W}_1(u_0, v_0).$$

En particulier, si $\lambda - 2\alpha > 0$ *le processus non linéaire associé à l'équation* (9) *admet une unique mesure invariante* u_∞ *et*

$$\mathcal{W}_1(u_t, u_\infty) \leq e^{-(\lambda-2\alpha)t}\mathcal{W}_1(u_0, u_\infty).$$

Ce théorème est basé sur deux faits :

1. La propagation du chaos est uniforme en temps : il existe une constante $K > 0$ telle que

$$\sup_{t\geq 0} \mathbb{E}|X_t^{1,N} - \overline{X}_t^1| \leq \frac{K}{\sqrt{N}},$$

   où $\overline{X}_t^1$ est une copie du processus non linéaire.

2. La convergence à l'équilibre du système de particules quand t tend vers l'infini : pour $(X_t^N)_{t\geq 0}$ et $(Y_t^N)_{t\geq 0}$ deux systèmes de particules en interaction de type champ moyen et pour tout $t \geq 0$

$$\mathcal{W}_{\|\cdot\|_{\ell^1}}(\mathrm{Law}(X_t^N), \mathrm{Law}(Y_t^N)) \leq e^{-(\lambda-2\alpha)t}\mathcal{W}_{\|\cdot\|_{\ell^1}}(\mathrm{Law}(X_0^N), \mathrm{Law}(Y_0^N)),$$

   où $\mathcal{W}_{\|\cdot\|_{\ell^1}}$ est la distance de Wasserstein sur $\mathbb{R}^N$ associée à la norme ℓ^1. Autrement dit pour μ et ν deux lois sur $\mathbb{N}^N$

$$\mathcal{W}_{\|\cdot\|_{\ell^1}}(\mu, \nu) = \inf_{\substack{X\sim\mu \\ Y\sim\nu}} \mathbb{E}\left[\sum_{i=1}^N |X^i - Y^i|\right].$$

Les points 1. et 2. se montrent par une méthode de couplage (détaillée dans le Chapitre 4). L'idée associée au phénomène de propagation du chaos est de coupler, à N fixé, N particules du système avec N processus non linéaires indépendants de loi u_t à chaque instant t et de même positions initiales que celles du système de particules. Les hypothèses précédentes permettent alors l'obtention d'une telle estimation uniforme en temps.

Chapitre 1

Outils mathématiques

Dans ce chapitre, nous allons rappeler les définitions de semi-groupe, de générateur ou de fonctions de Lyapunov et décrire brièvement les inégalités de Poincaré et de Sobolev logarithmique. Ces inégalités fournissent toutes les deux des vitesses de convergence du semi-groupe de $(X_t)_{t\geq 0}$ vers sa mesure invariante. Tous les résultats et démonstrations peuvent être trouvés dans [5, 13].

1.1 Semi-groupe en temps continu

Soit $(X_t)_{t\geq 0}$ un processus de Markov sur un espace Λ au plus dénombrable, non explosif, admettant une mesure invariante π. Le semi-groupe $(P_t)_{t\geq 0}$ associé à X_t défini pour toute fonction $f : \Lambda \to \mathbb{R}$ mesurable et bornée et pour tout $x \in \Lambda$ par

$$P_t f(x) = \mathbb{E}[f(X_t)|X_0 = x] = \int_\Lambda f(y) P_t(x, dy),$$

joue en temps continu le rôle des probabilités de transitions pour les chaînes de Markov. Pour tout $t \geq 0$, l'opérateur P_t vérifie les propriétés suivantes :

- $P_0 = Id$ est la fonction identité.
- $\forall s, t \geq 0$

$$P_{t+s} = P_{s+t} = P_t \circ P_s = P_s \circ P_t.$$

Partant d'une distribution initiale μ, la loi de X_t est décrite pour toute fonction continue bornée f par

$$E[f(X_t)] = \mu P_t f = \int_\Lambda \int_\Lambda f(y) P_t(x, dy) \mu(dx).$$

Considérons l'espace $L^2(\pi)$. L'ensemble des fonctions f dans $L^2(\pi)$ pour lesquelles la convergence

$$\mathcal{L} f(x) = \lim_{t \to 0} \frac{P_t f(x) - f(x)}{t} = \partial_t P_t f|_{t=0}(x)$$

a lieu est appelé le domaine du générateur $\mathcal{L}$ et est noté $D(\mathcal{L})$. L'opérateur $\mathcal{L}$ ainsi défini est appelé générateur infinitésimal associé à $(P_t)_{t \geq 0}$. En général, il y a équivalence entre la donnée du semi-groupe et celle du générateur puisque

$$\forall t \geq 0,\ \partial_t P_t(f) = P_t(\mathcal{L} f) = \mathcal{L}(P_t f).$$

Ces équations sont appelées équations de Chapman-Kolmogorov forward et backward. De ces équations découlent deux propriétés :

- Une caractérisation de la mesure invariante : une mesure de probabilité π est dite invariante si

  $$\forall t \geq 0, \quad \pi P_t = \pi.$$

  Les équations de Chapman-Kolmogorov donnent alors

  $$\forall f \in D(\mathcal{L}), \quad \int_\Lambda \mathcal{L} f(x) \pi(dx) = 0.$$

- Une relation de commutation entre l'opérateur du *carré du champ* Γ et le *semi-groupe* de X_t : si $X_0 = x$ alors pour toute fonction $g \in D(\mathcal{L})$ suffisamment régulière [5, Ch.2],

  $$\mathrm{Var}_x(g(X_t)) = P_t(g^2)(x) - (P_t g)^2(x) = 2 \int_0^t P_s \Gamma P_{t-s} g(x) ds$$

  où

  $$\Gamma(f, g) = \frac{1}{2} \left[\mathcal{L}(fg) - f \mathcal{L}(g) - g \mathcal{L}(f) \right] \text{ et } \Gamma(f) = \Gamma(f, f).$$

En effet, posons pour tout $s \in [0, t]$ et $x \in \Lambda$, $\Psi(s) = P_s \left[(P_{t-s} g)^2 \right](x)$ et $\psi(s) = P_{t-s} g$. Alors

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

Donc

$$\mathrm{Var}_x(g(X_t)) = \Psi(t) - \Psi(0) = 2\int_0^t P_s \Gamma P_{t-s} g(x) ds.$$

1.2 Critère de Foster-Lyapunov

Les fonctions de Lyapunov ont été développées pour l'étude de la stabilité des systèmes dynamiques, systèmes décrits par des équations différentielles ordinaires. Depuis, les méthodes basées sur les fonctions de Lyapunov ont été étendues aux processus de Markov (en temps discret ou continu) afin d'y établir des conditions de non-explosion et des propriétés de stabilité.

Définition 1.1 (Fonction de Lyapunov). *On dit que $V : \Lambda \to [1, +\infty[$ est une fonction de Lyapunov pour $(X_t)_{t\geq 0}$ s'il existe $t_0 \geq 0$ et deux constantes $\gamma \in (0,1)$ et $d \geq 0$ telles que pour tout $x \in \Lambda$ et $t > t_0$*

$$P_t V(x) \leq \gamma V(x) + d.$$

Une condition suffisante pour trouver une fonction de Lyapunov est l'existence de deux constantes $c > 0$ et $0 < d < +\infty$ telles que

$$\mathcal{L}V(x) \leq -cV(x) + d, \quad (\textit{Condition de dérive}) \tag{1.1}$$

où $\mathcal{L}$ est le générateur associé à $(P_t)_{t\geq 0}$.

Le concept de stabilité des processus de Markov à temps continu a principalement été développé par Meyn et Tweedie [85] pour lequel ils fournissent un critère garantissant l'existence d'une probabilité invariante et l'exponentielle ergodicité des probabilités de transition vers celle-ci [85, Théorème 6.1], [48]. Ce critère est basé sur la recherche d'une fonction de Lyapunov vérifiant la condition de dérive (1.1). Cependant, les taux de convergence ne sont pas explicites et en pratique, trouver une fonction de Lyapunov peut être difficile.

Théorème 1.2 (Théorème 6.1 Meyn et Tweedie [85]). *Soit V une fonction positive telle que $\lim_{|x|\to+\infty} V(x) = +\infty$ et satisfaisant (1.1). Alors si $(X_t)_{t\geq 0}$ est irréductible récurrent positif, $(X_t)_{t\geq 0}$ est exponentiellement ergodique i.e il existe $\beta < 1$ et $B < \infty$ tels que*

$$d_{VT}(P_t(x,.), \pi) \leq B(V(x)+1)\beta^t, \quad t \geq 0, x \in \Lambda.$$

Le théorème qui suit, connu sous le nom de critère de Foster-Lyapunov [88, Théorème 8.13], garantit l'ergodicité d'un processus de Markov c'est-à-dire la convergence du processus vers une unique mesure invariante. Ce critère a notamment permis de montrer l'ergodicité du processus de Fleming-Viot associé au processus de Galton-Watson sous critique [9] et du processus de Fleming-Viot associé à la marche aléatoire sur $\mathbb{N}$ [10].

Théorème 1.3 (Critère de Foster-Lyapunov)**.** *S'il existe une fonction* $V : \Lambda \to \mathbb{R}_+$*, des constantes* $K, \gamma > 0$ *telles que :*

1. $\mathbb{E}_x[V(X_1)] - V(x) \leq -\gamma$*, pour* $V(x) > K$*,*
2. *l'ensemble* $F = \{x : V(x) \leq K\}$ *est fini,*
3. $\mathbb{E}_x[V(X_1)] < \infty$ *pour tout* $x \in \Lambda$*.*

Alors le processus de Markov $(X_t)_{t\geq 0}$ *est ergodique.*

1.3 Vitesse de convergence

En cas d'ergodicité d'un processus, on aimerait estimer la vitesse de convergence vers l'équilibre. Pour cela, il existe plusieurs méthodes comme celles liées aux inégalités fonctionnelles ou au couplage.

1.3.1 Couplage

L'idée du couplage est de trouver des valeurs quantitatives de la convergence par l'introduction de distances entre deux mesures.

Définition 1.4 (Couplage)**.** *Soit* μ *et* ν *deux mesures de probabilité sur* Λ*. Un couplage de* μ *et* ν *est un couple* (X, Y) *de variables aléatoires tel que* X *suit la loi* μ *et* Y *la loi* ν *.*

Définition 1.5 (Distance en variation totale)**.** *La distance en variation totale entre deux mesures de probabilités* μ *et* μ' *est donnée par*

$$d_{VT}(\mu, \mu') = \frac{1}{2} \sup_{\|f\|_\infty \leq 1} \left(\int f d\mu - \int f d\mu' \right)$$

ou de manière équivalente par

$$d_{VT}(\mu, \mu') = \inf_{\substack{X \sim \mu \\ X' \sim \mu'}} \mathbb{P}\left(X \neq X'\right),$$

où l'infimum porte sur les couplages de μ *et* μ'*.*

Soit d une distance sur Λ et $p \geq 1$. On définit $\mathcal{P}_p(\Lambda)$ l'espace des mesures de probabilités μ sur Λ tel que le moment $\int_\Lambda d(x,x_0)^p d\mu(x)$ soit fini pour un (et donc tout) $x_0 \in \Lambda$.

Définition 1.6 (Distance de Wasserstein)**.** *La distance de Wasserstein d'ordre p entre deux mesures μ et μ' dans $\mathcal{P}_p(\Lambda)$ associée à la distance d est définie par*

$$\mathcal{W}_p(\mu,\mu') = \inf_{\substack{X\sim\mu \\ Y\sim\mu'}} \left(\mathbb{E}\left[d(X,Y)\right]^p\right)^{\frac{1}{p}},$$

où l'infimum porte sur les couples de variables aléatoires de lois marginales μ et μ'.

Remarque 1.7. *Si $d(x,y) = \mathbb{1}_{x\neq y}$ alors $\mathcal{W}_1 = d_{VT}$.*

Théorème 1.8 (Dualité de Kantorovich-Rubinstein [99])**.** *Pour toutes mesures de probabilités μ et ν on a*

$$\mathcal{W}_1(\mu,\nu) = \sup_{f\in Lip_1} \left(\int f d\mu - \int f d\mu'\right),$$

où Lip est l'ensemble des fonctions lipschitziennes f par rapport à la distance d i.e

$$\|f\|_{Lip} =: \sup_{x\neq y} \frac{|f(x)-f(y)|}{d(x,y)} < \infty$$

et Lip_1 est l'ensemble des fonctions 1 Lipschitziennes i.e l'ensemble des fonctions f tel que $\|f\|_{Lip} \leq 1$.

Pour étudier le comportement en temps long des processus, on est amené à construire des couplages : on construit un couple de processus (X,Y) de positions initiales différentes et dont les marginales suivent la dynamique donnée par le générateur, de telle manière que les variables aléatoires soient de plus en plus proches. Ceci donne un contrôle de la distance de Wasserstein. Pour en obtenir une de la distance en variation totale, il faut réussir à rendre les deux variables aléatoires égales le plus souvent possible.

Ces distances sont particulièrement adaptées dans le cadre des problèmes de couplage car n'importe quel couplage en donne une majoration. Il faut alors se demander si les majorations obtenues sont optimales.

Théorème 1.9 (Existence et unicité d'une mesure invariante)**.** *La courbure de Wasserstein de $(P_t)_{t\geq 0}$ est la plus grande constante $\rho \in \mathbb{R}$ telle que pour tout $x,y \in \Lambda$ et $t \geq 0$*

$$\mathcal{W}_1(\delta_x P_t, \delta_y P_t) \leq e^{-\rho t}\mathcal{W}_1(\delta_x,\delta_y).$$

Si $\rho > 0$ alors il existe une unique mesure invariante π et pour tout $x \in \Lambda$ et $t \geq 0$

$$\mathcal{W}_1(\delta_x P_t, \pi) \leq e^{-\rho t} \mathcal{W}_1(\delta_x, \pi).$$

Une preuve de ce théorème se trouve dans [31].

Corollaire 1 (Trou spectral et courbure de Wasserstein)**.** *On suppose que la courbure de Wasserstein ρ est strictement positive. On suppose de plus que la mesure invariante π est réversible. Alors ρ minore le trou spectral c'est-à-dire le premier élément non nul du spectre [13].*

Remarque 1.10 (Cas où $\rho \leq 0$)**.** *Le cas $\rho \leq 0$ ne nous donne aucune information sur le trou spectral. Cependant, les inégalités de Hardy [5, Chapitre 6] permettent d'en obtenir des estimations. Un exemple d'utilisation de ces inégalités est donné à la section 3.4.4 du chapitre 3 dans le cas d'un processus de naissance et mort et les arguments utilisés ont été inspirés sur ceux de Miclo [86].*

1.3.2 Inégalités fonctionnelles

Les inégalités fonctionnelles dite de Poincaré ou de Sobolev Logarithmique fournissent des estimations de la convergence en temps long du processus de Markov $(X_t)_{t\geq 0}$. Elles peuvent être vues comme des inégalités portant sur le générateur [5, 13]. Soit π la mesure invariante associée au semi-groupe de générateur $\mathcal{L}$. On suppose que la mesure π est une probabilité.

Définition 1.11 (Inégalités de Poincaré et Sobolev Logarithmique)**.** *On dit que π satisfait une inégalité de Poincaré, resp. Sobolev Logarithmique, s'il existe une constante $c_1 > 0$, resp. $c_2 > 0$ telle que, pour toute fonction f suffisamment régulière*

$$\mathrm{Var}_\pi(f) \leq -c_1 \int_\Lambda f(x)\mathcal{L}f(x)\pi(dx),$$

$$\textit{resp. } \mathrm{Ent}_\pi(f^2) \leq -c_2 \int_\Lambda f(x)\mathcal{L}f(x)\pi(dx),$$

où pour $f \in L^2(\mu)$

$$\mathrm{Var}_\pi(f) = \int_\Lambda \left(f(x) - \int_\Lambda f(y)\pi(dy)\right)^2 \pi(dx)$$

$$\textit{et } \mathrm{Ent}_\pi(f) = \int_\Lambda f\log(f)d\pi - \left(\int_\Lambda f d\pi\right)\log\left(\int_\Lambda f d\pi\right).$$

L'inégalité de Poincaré est équivalente à la convergence exponentielle du semi-groupe dans $L^2(\pi)$ et dans le cas des diffusions, l'inégalité de Sobolev Logarithmique équivaut à une convergence du semi-groupe au sens de l'entropie.

Chapitre 2

Fleming-Viot process with constant drift *

2.1 Introduction

There is recent interest in approximating the limiting law of irreducible Markov processes conditioned not to hit some (forbidden) state [74, 102]. This limiting law is not guaranteed to exist, but when it does it is called a quasi-stationary distribution (QSD). For QSD in the context of Birth and Death chains, we refer to [54] (see also [47]) : the situation treated there is one in which there is a one parameter family of QSD.

QSD are neither well understood, nor easily amenable to simulation. One proposal made by Burdzy, Holyst, Ingerman, and March [24] (in a particular setting) is to consider N independent Markov processes except that when one reaches the forbidden state, it jumps to the state of one of the other processes, chosen uniformly at random. The natural conjecture is that the empirical measure, under the stationary measure, converges to the QSD as the number N of processes goes to infinity. It is also natural to conjecture that the selected QSD is *the minimal*, in terms of average time needed to reach the forbidden state.

In this note, we consider N random walks on $\mathbb{N}$, in continuous time, with a drift towards the origin. When one random walk reaches the origin, it jumps instantly to the position of one of the other $N-1$ walks, chosen uniformly at random. We call Fleming-Viot the interacting random walks just described. Indeed, this dynamics has a genetic interpretation when the positions of the walks are thought of as evolving genetic traits of N individuals with selection and branching :

$*$. In collaboration with Amine Asselah [10]

– *Selection* : the forbidden state (here 0) is a lethal trait.

– *Branching* : when an individual dies, another one chosen uniformly at random, branches. Here the branching is linked with selection so as to keep the population size constant. This model is related to N-Brownian Branching Motions proposed by Brunet and Derrida in [23].

In this note, we establish a Foster's criteria, which gives ergodicity, as well as a control of small exponential moments of the rightmost walk. To state our main result, let ξ_T denote the position of the N interacting walks at time T, and let $\mathbb{E}[\cdot|\xi_0 = \xi]$ denote average with respect of the law of the process $\{\xi_t,\ t \geq 0\}$ with initial condition ξ.

Theorem 2.1. *There are positive constants $K, \alpha, \kappa, \delta_0, A, c_1, c_2, c_3$ such that for $N \in \mathbb{N}$, time $T = A\log(N)$, for any $\delta < \delta_0$, and $\xi \in \mathbb{N}^N$, we have*

$$\begin{aligned}\mathbb{E}\left[\exp\Big(\delta\max(\xi_T)\Big)\Big|\xi_0 = \xi\right] - \exp\Big(\delta\max(\xi)\Big) < &- c_1\,\mathbb{1}_{\max(\xi)>K\log(N)}e^{\delta\max(\xi)} \\ &+ c_2\mathbb{1}_{\max(\xi)>K\log(N)}e^{-\kappa T}e^{\delta\max(\xi)} + c_3e^{\delta\alpha\log(N)}.\end{aligned} \tag{2.1}$$

As a consequence, for N large enough there is a unique invariant measure λ_N for Fleming-Viot. Integrating over λ_N, there are $\beta, C > 0$ such that for any N, and $\delta < \delta_0$

$$\int \exp(\delta\max(\xi))d\lambda_N(\xi) \ \leq\ C\exp\Big(\delta\beta\log(N)\Big). \tag{2.2}$$

This first elementary step is an important ingredient in the proof of the conjecture we alluded to above. Also, it might be of independent interest in view of recent deep and comprehensive studies on the rightmost position in branching random walks [1–3, 6, 7, 23, 69, 76]. This selection of recent works is far from being exhaustive, but already shows the vitality of this issue.

The rest of this note is organized as follows. In Section 2.2 we define the model, and recall well-known large deviations estimates. In Section 2.3, we explain how to divide walks into groups with little correlations over a well chosen time period. In Section 2.4, we estimate the probability that the maximum displacement does not decrease. Finally, in Section 2.5, we establish Foster's criteria.

2.2 Model and Preliminaries

Here, we deal with continuous-time nearest neighbor random walks on $\mathbb{N}$, with rate p to jump right, and rate $q = 1 - p > p$ to jump left. The drift is $-v$ with $v = q - p > 0$. A single walk makes N_t jumps in the time period $[0, t]$, and its increments are denoted $X_1, \ldots, X_n$, with

$\mathbb{E}[X_i] = -v$, and $\bar{X}_i = X_i + v$ denotes the centered variable. Note that

$$\mathbb{P}\Big(\sum_{i=1}^{T}(X_i + v) \geq xT\Big) \leq \exp(-TI(x)), \tag{2.3}$$

with

$$I(x) = \sup_{\lambda > 0}\{\lambda x - \Lambda(\lambda)\} \quad \text{with} \quad \Lambda(\lambda) = \log(pe^{\lambda} + qe^{-\lambda}) + \lambda v. \tag{2.4}$$

Due to the nearest neighbor jumps of our walk, we have

$$I(v+1) = \log\Big(\frac{1}{1-q}\Big) \quad \text{and for} \quad x > v+1, \quad I(x) = \infty. \tag{2.5}$$

We define also $x \mapsto \tilde{I}(x) = 1 - \exp(-I(x))$, which is discontinuous at $v+1$ with

$$\tilde{I}(v) < \tilde{I}(v+1) = q \quad \text{and for} \quad x > v+1, \quad \tilde{I}(x) = 1. \tag{2.6}$$

Note that if N_T is Poisson of mean T, then

$$\mathbb{P}\Big(\sum_{i \leq N_T}(X_i + v) \geq xN_T\Big) \leq \exp(-T\tilde{I}(x)). \tag{2.7}$$

On Poisson tails. We need two rough tail estimates on the Poisson clocks. Both are obvious and well-known. For any χ and T positive, we have

$$\mathbb{P}(N_T \geq eT + \chi) \leq \exp(-T - \chi), \tag{2.8}$$

and

$$\mathbb{P}(N_T \leq \frac{1}{e}T - \chi) \leq \exp(-(1 - 2/e)T - \chi), \tag{2.9}$$

Both are obtained readily by Chebychev's inequality. Indeed, we obtain (2.8) from

$$\mathbb{P}(N_T \geq eT + \chi) \leq e^{-eT-\chi}\mathbb{E}[e^{N_T}] = \exp(-T - \chi), \tag{2.10}$$

and we obtain (2.9) from

$$\mathbb{P}(N_T \leq \frac{1}{e}T - \chi) \leq e^{T/e-\chi}\mathbb{E}[e^{-N_T}] = \exp(-(1 - 2/e)T - \chi). \tag{2.11}$$

2.3 Independence

On the multitype branching of [9]. A key idea introduced in [9] is to embed the Fleming-Viot process into a multitype branching process whose space displacements and branching mechanism are independent, and which is *attractive*. We refer to Section 3 of [9] for a description of the multitype branching process, and recall here its main features. Assume that we start with N interacting random walks. This defines N types with which we associate N independent exponential clocks of intensity q with marks. The time realizations of the clock of type i have marks in the set of labels $\{1, \ldots, N\} \backslash \{i\}$, and each mark is chosen uniformly at random from the $N-1$ symbols. When clock i rings, and when its mark is j, each walk of type j branches into two children : one of type i and one of type j. The two children behave as independent random walks starting at the position of their parent. If $\mathbb{D}_T$ denotes the population of individuals alive at time T, and $|\mathbb{D}_T|$ denotes its cardinal, it is easy to see the equality $\mathbb{E}[|\mathbb{D}_T|] = |\mathbb{D}_0| \exp(qT)$. For an individual v alive at time T, we denote by $t \mapsto S_v(t)$ its trajectory for $t \in [0, T]$.

Independent groups of walks. A drawback of the multitype branching process is an exponentially growing population. Since, we use a time of order $\log(N)$, we cannot use here such an embedding. Even though in the Fleming-Viot process, all particles interact which each other, a simple observation is that as long as a particle has not touched the origin its trajectory is a sum of independent increments even though this trajectory might influence others.

To create some independence between walks, we decompose the interacting walks in two sets at time 0. We first fix a time T and a length L to be chosen later.

- The *blacks*, whose initial position is below L.
- The *reds*, whose initial position is above L.

Then, color changes as follows : if a black walk jumps on a red walk, it becomes instantly red. We interpret this jump as a *red binary branching*. Now, red walks are not independent from black walks because they might touch the origin before time T, and jump onto a black position. However, if $vT \ll L$, we expect this to be rare. To obtain independence, we add another color to our description : each red walk is coupled with a green walk which behaves identically in terms of move or branching but with green children, except that when a green walk reaches the origin it continues its drifted motion on $\mathbb{Z}$ (without selection mechanism). Thus, green walks behaves like independent random walks with branching at the times a black particle hit zero and chooses the label of a green walk. If R_0 is the first time one of the red walks touches the origin, we have that at time T, on the event $\{R_0 > T\}$, red and green positions are identical. The point of introducing green walks is that their branching times is independent of their positions. We denote with D_T^r, D_T^g, D_T^b the respective labels of red, green and black walks at time T. Also $D_T = D_T^r \cup D_T^g \cup D_T^b = \{1, \dots, N\}$, and we still denote by $t \mapsto S_v(t)$, the trajectory of $v \in D_T$.

When embedding a group of walks into a branching process, we denote with $\mathbb{D}_T^r, \mathbb{D}_T^g, \mathbb{D}_T^b$ the respective labels of red, green and black individuals in the mutlitype branching processes.

The key idea here is to work on a time of order $\log(N)$, to control the black walks by a multitype branching process, but to let the red walks (or rather the green walks) grow as in Fleming-Viot with a population bounded by N, and with branching due to independent black walks.

On the choice of time T and length L. We choose T large enough so that $q+\log(N)/T < 1$. We actually need a little more. We need that κ as defined below be positive :

$$\kappa = \min\Big(1 - 2/e, \tilde{I}(\frac{v}{2}) - \frac{\log(N)}{T}, 1 - q - \frac{\log(N)}{T}\Big) > 0. \tag{2.12}$$

Once $T = A\log(N)$ satisfies (2.12), we set $L = eT$.

2.4 When things go wrong

We wish to estimate the probability of the event where the maximum displacement does not decrease. We assume in this section that $\max\limits_{i \le N} \xi_i > 3L$. We define $B(T, L)$ (the *bad set*) as containing the following events :

– One red walk reaches the origin before time T (i.e. $\{R_0 \le T\}$).
– One black walk travels a distance L upwards in a period $[0, T]$.
– The maximum displacement of a green walk in a time T is above $-\frac{v}{2e}T$.

Thus, on the complement on $B(T, L)$, green and red are identical, and

$$\max_{v \in D_T} S_v(T) - \max_{v \in D_0} S_v(0) \le \max_{v \in D_T}\Big(S_v(T) - S_v(0)\Big) = \max_{v \in D_T^g}\Big(S_v(T) - S_v(0)\Big) < -\frac{v}{2e}T,$$

which implies, under the assumption $\max\limits_{i \le N} \xi_i > 3L$, that if $M(T) = \max\limits_{v \in D_T} S_v(T)$

$$\mathbb{E}\Big[\mathbb{1}_{B^c(T,L)} \exp\Big(\delta(M(T) - M(0))\Big)\Big|\xi(0) = \xi\Big] \le \exp\Big(-\frac{v\delta}{2e}T\Big). \tag{2.13}$$

We next estimate the probability of each event making up $B(T, L)$, with the following outcome.

Lemma 2.2. *For any $\xi \in \mathbb{N}^N$, we have, with κ as in (2.12),*

$$\mathbb{P}(B(T,L)\Big|\xi(0)=\xi) \leq 4\exp(-\kappa T). \tag{2.14}$$

2.4.1 A red walk does reach 0

Recall that $L = eT$. We embed the Fleming-Viot into a branching multitype, while keeping the red coloring. We need to estimate the probability that one red displacement gets below L units in a time period $[0,T]$. Note that to realize $\{R_0 < T\}$, there is $v \in \mathbb{D}_T^r$ such that the number of its time jumps N_T must be larger than L, and this is what we use (we recall that (2.12) implies that $\kappa \leq 1 - q - \frac{\log(N)}{T}$)

$$\begin{aligned} \mathbb{P}(R_0 < T\Big|\xi(0)=\xi) \leq & \mathbb{E}[|\mathbb{D}_T^r|] \times \mathbb{P}\Big(\exists t \leq T,\ \sum_{i \leq N_t} X_i < -eT\Big) \\ \leq & \mathbb{E}[|\mathbb{D}_T^r|] \times \mathbb{P}\Big(N_T > eT\Big) \leq N e^{qT} e^{-T} \leq e^{-\kappa T}. \end{aligned} \tag{2.15}$$

2.4.2 A large black displacement

Recall that at the time a black reaches 0, and jumps on a red walk, it ceases to be black to become red. We bound here the black walks with a multitype branching, assuming in this section that blacks remain blacks even if jumping on a red walk, with the effect that we are overestimating the black population. The estimates are similar to these of Section 2.4.1. We use that to make L steps right, a black walk must make L time-marks ($N_T > L$), and this event is estimated in (2.15).

2.4.3 Green's maximum too high

The key point is that the green branching times are independent of positions of the green. They depend only on the history of black walks. Also, the population of green walks is bounded by N. Thus, it is crucial here not to use the multitype branching of [9] : we estimate the probability that $\{\max_{v \in D_T^g}(S_v(T) - S_v(0)) > -vT/(2e)\}$. Define $\mathcal{N}_T(\gamma)$ as the number of green walks whose displacement during time period $[0,T]$ is larger than γ. Then,

$$\mathbb{E}[\mathcal{N}_T(\gamma)\Big|\xi(0)=\xi] = \mathbb{E}[|D_T^g|\Big|\xi(0)=\xi] \times \mathbb{P}\Big(\sum_{i \leq N_T} X_i > \gamma\Big). \tag{2.16}$$

The reason is the independence of the branching times and displacements of the green walks. For $v \in D_T^g$, assume there are ν branchings before time T, say at times $T_1, \ldots, T_\nu$, and we have (for $X_i^{(k)}$ i.i.d. independent from $\{T_i, i \in \mathbb{N}\}$)

$$S_v(T) - S_v(0) = \sum_{i \in N[0,T_1]} X_i^{(1)} + \cdots + \sum_{i \in N[T_\nu, T]} X_i^{(\nu)}. \tag{2.17}$$

As one conditions first on the black history up to time T, one fixes the times $T_1, \ldots, T_\nu$, and obtain that $N[0,T_1] + \cdots + N[T_\nu, T]$ sum up to a Poisson variable $N[0,T]$ of intensity T, and most importantly

$$\sum_{i \in N[0,T_1]} X_i^{(1)} + \cdots + \sum_{i \in N[T_\nu, T]} X_i^{(\nu)} = \sum_{i \in N[0,T]} X_i, \tag{2.18}$$

where the $\{X_i,\ i \in \mathbb{N}\}$ are i.i.d. increments independent of $N[0,T]$. We obtain, with κ defined in (2.12),

$$\begin{aligned}
\mathbb{P}\Big(\max_{v \in D_T^g}(S_v(T) - S_v(0)) > -\frac{v}{2e}T \Big| \xi(0) = \xi\Big) \le \mathbb{E}\left[\mathcal{N}_T\Big(-\frac{vT}{2e}\Big)\Big|\xi(0) = \xi\right]& \\
\le N\mathbb{P}\Big(\sum_{i \le N_T} \bar{X}_i > vN_T - \frac{v}{2e}T\Big)& \\
\le N\Big(\mathbb{P}\Big(\sum_{i \le N_T} \bar{X}_i > \frac{v}{2}N_T\Big) + \mathbb{P}\Big(N_T < \frac{1}{e}T\Big)\Big)& \\
\le N\Big(\exp\Big(-T\tilde{I}(\frac{v}{2})\Big) + \exp\Big(-(1-2/e)T\Big)\Big) \le 2\exp(-\kappa T).&
\end{aligned} \tag{2.19}$$

2.5 Foster's criteria

We start with an estimate on the tail, and of the exponential moments.

2.5.1 On exponential moments

We deal here with the multitype branching process. Recall that $\mathbb{D}_0 = \{1, \ldots, N\}$, and let $S(0) = \{S_v(0), v \in \mathbb{D}_0\}$.

Lemma 2.3. *For any T satisfying (2.12), and any $\chi > 0$*

$$\mathbb{P}\Big(\max_{v \in \mathbb{D}_T}\Big(S_v(T) - S_v(0)\Big) > eT + \chi \Big| S(0) = \xi\Big) \le \exp(-\chi). \tag{2.20}$$

Proof. Since the branching mechanism is independent of positions

$$\mathbb{P}\Big(\max_{v \in \mathbb{D}_T}\Big(S_v(T) - S_v(0)\Big) > eT + \chi \Big| S(0) = \xi\Big) \le \mathbb{E}[|\mathbb{D}_T|] \times \mathbb{P}\Big(\sum_{i=1}^{N_T} X_i > eT + \chi\Big). \tag{2.21}$$

Now, if $\bar{X}_i$ denotes the centered variable, note that since the walk is nearest neighbor

$$\mathbb{P}\Big(\sum_{i=1}^{N_T} \bar{X}_i > (v+1)N_T\Big) = 0.$$

Now,

$$\begin{aligned}
\mathbb{P}\Big(\sum_{i=1}^{N_T} X_i > eT + \chi\Big) =& \mathbb{P}\Big(\sum_{i=1}^{N_T} \bar{X}_i > vN_T + eT + \chi\Big) \\
\leq& \mathbb{P}\Big(\sum_{i=1}^{N_T} \bar{X}_i > (v+1)N_T\Big) + \mathbb{P}(N_T > eT + \chi) \\
\leq& 0 + \mathbb{P}(N_T > eT + \chi)
\end{aligned} \tag{2.22}$$

Now, N_T is a Poisson variable of mean T, the standard estimate (2.8) leads to

$$\mathbb{P}\Big(N_T > eT + \chi\Big) \leq \exp\Big(-T - \chi\Big). \tag{2.23}$$

Also, we have $\mathbb{E}[|\mathbb{D}_T|] \leq N\exp(qT)$, and (2.21) and the choice of T in (2.12) yield

$$\mathbb{P}\Big(\max_{v\in\mathbb{D}_T}\Big(S_v(T) - S_v(0)\Big) > eT + \chi \Big| S(0) = \xi\Big) \leq Ne^{qT}e^{-T-\chi} \leq e^{-\chi}. \tag{2.24}$$

□

We can state our main estimate.

Lemma 2.4. *Assume that* $(1-q)T > log(N)$, *and* $\delta < 1$. *Then, we have*

$$\mathbb{E}\left[\exp\Big(\delta\Big(\max_{i\leq N}\xi_i(T) - \max_{i\leq N}\xi_i(0)\Big)\Big)\Big|\xi(0) = \xi\right] \leq \frac{1}{1-\delta}e^{\delta eT}. \tag{2.25}$$

Proof. For any random variable X, we have

$$\begin{aligned}
\mathbb{E}\left[e^{\delta X}\right] =& 1 + \int_0^\infty \delta e^{\delta u}\mathbb{P}(X > u)du \\
\leq& 1 + \int_0^{eT} \delta e^{\delta u}du + \int_{eT}^\infty \delta e^{\delta u}\mathbb{P}(X > u)du \\
\leq& e^{\delta eT}\Big(1 + \int_0^\infty \delta e^{\delta u}\mathbb{P}(X > u + eT)du\Big).
\end{aligned} \tag{2.26}$$

Now, using the tail estimate (2.20), we have

$$\mathbb{E}\left[\exp\left(\delta \max_{v\in\mathbb{D}_T}\left(S_v(T)-S_v(0)\right)\right)\right] \leq e^{\delta e T}\left(1+\int_0^\infty \delta e^{\delta u}e^{-u}\right) \leq \frac{e^{\delta e T}}{1-\delta}. \tag{2.27}$$

We now use the following bound to conclude

$$\begin{aligned}\mathbb{E}\left[\exp\left(\delta\Big(\max_{i\leq N}\xi_i(T)-\max_{i\leq N}\xi_i(0)\Big)\right)\Big|\xi(0)=\xi\right] \leq & \mathbb{E}\left[\exp\left(\delta\Big(\max_{v\in\mathbb{D}_T}S_v(T)-\max_{v\in\mathbb{D}_0}S_v(0)\Big)\right)\Big|S(0)=\xi\right]\\ \leq & \mathbb{E}\left[\exp\left(\delta\max_{v\in\mathbb{D}_T}\left(S_v(T)-S_v(0)\right)\right)\Big|S(0)=\xi\right].\end{aligned} \tag{2.28}$$

□

2.5.2 Proof of Theorem 2.1

We recall the general strategy of the proof of Proposition 1.2 of [9] (The Foster criteria). We have a bad set $B(T,L)$ (which depends on T and L) which contains the cases where the maximum increases over a period $[0,T]$, or when black walks win over or influence red ones. First, there is a set K on which we do not expect the maximum to decrease, with

$$K=\left\{\max_v(S_v(0))<3L\right\}. \tag{2.29}$$

Then, there is a good set where the maximum decreases :

$$K^c\cap B^c(T,L)\subset G=\left\{\max_{v\in D_T}(S_v(T))-\max_{v\in D_T}(S_v(0))\leq -\frac{v}{2e}T\right\}. \tag{2.30}$$

Now, set $M_t=\max S_v(t)$. When ξ is the initial configuration, and when we work with $\xi\in K^c$, we have using Cauchy-Schwarz

$$\begin{aligned}\mathbb{1}_{\xi\in K^c}\left(\mathbb{E}\left[e^{\delta M_T}\Big|\xi(0)=\xi\right]-e^{\delta M_0}\right) =& \mathbb{1}_{\xi\in K^c}e^{\delta M_0}\left(\mathbb{E}\left[e^{\delta(M_T-M_0)}\left(\mathbb{1}_{B(T,L)}+\mathbb{1}_{B^c(T,L)}\right)\Big|\xi(0)=\xi\right]-1\right)\\ \leq& \mathbb{1}_{\xi\in K^c}e^{\delta M_0}\left(\mathbb{P}\Big(B(T,L)\Big|\xi(0)=\xi\Big)\mathbb{E}\left[\exp(2\delta(M_T-M_0))\Big|\xi(0)=\xi\right]\right)^{1/2}\\ &-\mathbb{1}_{\xi\in K^c}e^{\delta M_0}\left(1-e^{-\delta vT/2}\right).\end{aligned} \tag{2.31}$$

We know from Lemma 2.4 that for $\delta < 1/2$

$$\mathbb{E}\left[\exp(2\delta(M_T - M_0))\middle|\xi(0) = \xi\right] \leq \frac{e^{2\delta eT}}{1-2\delta}.$$

Note also that on the set K, we obtain by using Lemma 2.4

$$\mathbb{1}_K\Big(\mathbb{E}\left[e^{\delta M_T}\middle|\xi(0) = \xi\right] - e^{\delta M_0}\Big) \leq \mathbb{1}_K\frac{\exp(3\delta L + \delta eT)}{1-\delta}. \tag{2.32}$$

Thus, adding (2.31) and (2.32), we obtain

$$\mathbb{E}\left[e^{\delta M_T}\right] - e^{\delta M_0} \leq \mathbb{1}_K\frac{e^{3\delta L+\delta eT}}{1-\delta} - \mathbb{1}_{K^c}\Big(1-e^{-\delta vT/2}\Big)e^{\delta M_0} + \mathbb{1}_{K^c}\Big(\mathbb{P}(B(T,L)\big|\xi(0)=\xi)\frac{e^{2\delta eT}}{1-2\delta}\Big)^{1/2}e^{\delta M_0}. \tag{2.33}$$

We use now Lemma 2.2, and choose δ small enough so that $\kappa > 4\delta e$ with the result

$$\mathbb{E}\left[e^{\delta M_T}\right] - e^{\delta M_0} \leq \frac{e^{3\delta L+\delta eT}}{1-\delta} - \mathbb{1}_{K^c}\Big(1-e^{-\delta vT/2}\Big)e^{\delta M_0} + \mathbb{1}_{K^c}\frac{e^{-\kappa T/4}}{\sqrt{1-2\delta}}e^{\delta M_0}. \tag{2.34}$$

Inequality (2.34) is a Foster's criteria (see [88, Theorems 8.6 and 8.13]). This implies the first part of Theorem 2.1.

Now, as we integrate (2.34) with respect to the invariant measure, the left hand side of (2.34) vanishes, and we obtain

$$\Big(1-e^{-\delta vT/2}\Big)\int_{K^c} e^{\delta M(\xi)}d\lambda^N(\xi) \leq \frac{\exp(3\delta L+\delta eT)}{1-\delta} + \exp(-\frac{\kappa}{4}T)\int_{K^c} e^{\delta M(\xi)}d\lambda^N(\xi). \tag{2.35}$$

With A large enough so that (2.12) holds with $T = A\log(N)$ and $L = eT$, the second part of Theorem 2.1 follows at once.

2.6 Conjecture and simulations

In this note, we consider random walks on $\mathbb{N}$ with a drift towards the origin. Cavender has shown in [28] that, in this case, there is a one parameter family of QSD. To describe this family, we give in a first time, a necessary and sufficient condition for a probability measure $\nu = (\nu_j)_{j\in\mathbb{N}}$ to be a QSD.

Theorem 2.5 (QSD). *The sequence $(\nu_j)_{j\in\mathbb{N}}$ is a QSD if and only if*

– $\forall j \in \mathbb{N}\ \nu_j \geq 0$ *and* $\sum_{j\in\mathbb{N}} \nu_j = 1$

- $\forall j \geq 2$,

$$\begin{cases} -(q+p)\nu_1 + q\nu_2 = -q\nu_1^2 \\ p\nu_{j-1} - (q+p)\nu_j + q\nu_{j+1} = -q\nu_1\nu_j \end{cases}$$

Proof. It comes from the fact that ν is a QSD if and only if ν is a left eigenvector for the restriction of the transition rates matrix to $\mathbb{N}^*$ with eigenvalue $-q\nu(1)$.

□

Theorem 2.6 (Expression of a QSD). *Let $\nu = (\nu_j)_{j\in\mathbb{N}}$ be a QSD. We assume that $0 < \nu_1 \leq \left(\sqrt{\frac{p}{q}} - 1\right)^2$. Then, setting $c = ((\nu_1 - 1 - \lambda)^2 - 4\lambda)^{\frac{1}{2}}$ with $\lambda = \frac{p}{q}$, we have*

- *if $c > 0$*

$$\nu_j = \frac{\nu_1}{c}\left[\left(\frac{\lambda + 1 - \nu_1 + c}{2}\right)^j - \left(\frac{\lambda + 1 - \nu_1 - c}{2}\right)^j\right],$$

- *if $c = 0$*

$$\nu_j = j\nu_1\left(\frac{\lambda + 1 - \nu_1}{2}\right)^{j-1}.$$

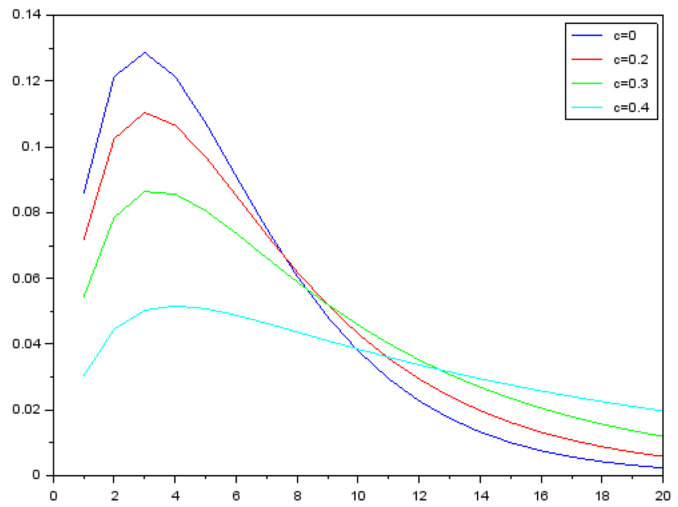

Figure 2.1 – Quasi-stationary distribution for Random Walk on $\mathbb{N}$ with $q = 2/3$, $p = 1/3$.

Remark 2.7.

- *QSDs are parametrized by c and so we should write $\nu = \nu_c = (\nu_{c,j})_{j\in\mathbb{N}}$.*

- *As $\lambda < 1$ we have*
$$0 \leq c < \sqrt{(1-\lambda)^2} = |1-\lambda| = 1-\lambda.$$
- *The minimal value $c = 0$ would correspond to the minimal QSD.*

Proof. Let us fix $\nu_1 > 0$. As ν is a QSD, Theorem 2.5 gives, for all $j \in \mathbb{N}^*$

$$p\nu_j - (q+p)\nu_{j+1} + q\nu_{j+2} = -q\nu_1\nu_{j+1}.$$

So putting $\lambda = \dfrac{p}{q} < 1$, we have the linear equation with constant coefficients

$$\nu_{j+2} - (\lambda + 1 - \nu_1)\nu_{j+1} + \lambda\nu_j = 0.$$

Solving the characteristic equation $y^2 - (\lambda + 1 - \nu_1)y + \lambda = 0$ ends the proof. □

Since the random walk has infinitely many QSDs, the question is about knowing toward which QSD converges the associated Fleming-Viot process, under the stationary distribution. To our knowledge, only two results are known. The first one is given by Asselah, Ferrari, Groisman and Jonckheere in [9] in the subcritical Galton-Watson case.

Theorem 2.8 (Asselah, Ferrari, Groisman, Jonckheere [9])**.** *Consider a subcritical Galton-Watson process whose offspring law has some finite positive exponential moment. Let ν_0 be the minimal quasi-stationary distribution for the process conditioned on non-extinction. Then,*

- *For each $N \geq 1$, the associated N-particle Fleming-Viot system is ergodic with invariant distribution λ^N.*
- *The empirical measure of FV, $m(\cdot, \xi)$ converges to the minimal QSD :*

$$\forall x \in \mathbb{N}^*, \quad \lim_{N\to+\infty} \int |m(x,\xi) - \nu_0(x)| d\lambda^N(\xi) = 0.$$

The second one is given by Villemonais in [103] where, owing to a Lyapunov-type criterion, proves the ergodicity and the convergence of a Fleming-Viot type particle system to the minimal quasi-stationary distribution of some birth and death processes. But the arguments used do not apply for the pure drift birth and death process (i.e the random walk).

However, it has been conjectured in [79] that as N goes to infinity, the empirical equilibrium measure approaches the minimal QSD.

Conjecture (Marić[79])**.** *The empirical measure of FVRW, under the stationary measure, converges to the minimal QSD as N goes to infinity.*

No general proof has been provided, but the simulations established by Marić confirm this conjecture.

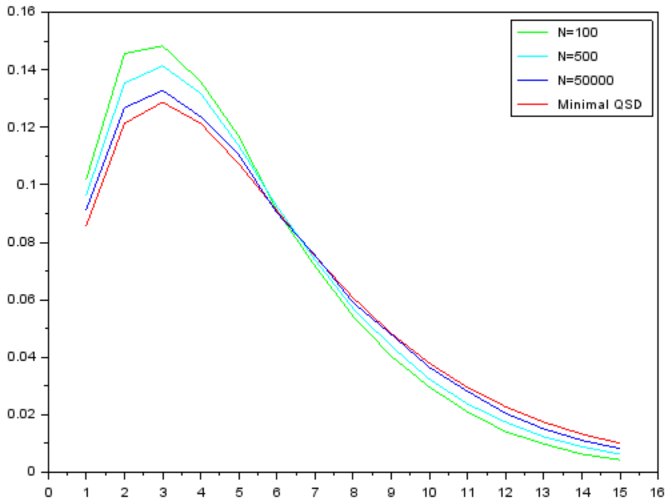

Figure 2.2 – Approximation of the FVRW by the minimal QSD.

Chapitre 3

Around the complete graph *

3.1 Introduction

We consider a (time-continuous) Moran type model, referred to as the Fleming-Viot process in the literature [25, 55], 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 [35, 84]. 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[42, 55, 102], the present chapter is concerned with deriving bounds for the rate of convergence. Our first main result, namely Theorem 3.1, 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 Theorem 3.3 and Theorem 3.11. This is also a new result even if [55] 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 3.2) which can be compared to the results of [42, 63, 102], a quantitative convergence of the conditioned semigroup (Corollary 3.4) comparable to the results of [43, 81] and uniform bound (in time) as N tends to infinity (see Corollary 3.5), which seems to be new in discrete space but already proven

$*$. In collaboration with Bertrand Cloez [34], to appear in SPA for diffusion processes in [89] 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 $f : F \mapsto \mathbb{R}$, gives

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

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 function f on F^*, by

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

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

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

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 [25], and Del Moral, Guionnet, Miclo [41, 42] 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

$$\mathcal{L}f(\eta) = \mathcal{L}^{(N)}f(\eta) = \sum_{i\in F^*} \eta(i) \left[\sum_{j\in F^*} (f(T_{i\to j}\eta) - f(\eta)) \left(Q_{i,j} + p_0(i)\frac{\eta(j)}{N-1} \right) \right], \tag{3.1}$$

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

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

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

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

For $\varphi : F^* \to \mathbb{R}$ 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 :

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

where all limits are in distribution and the limits $(b), (c)$ are taken as N tends to infinity. More precisely, Theorem 3.1 gives a bound for the limit (a), Theorem 3.2 for the limit (b), Corollary 3.5 for the limit (c) and finally Corollary 3.4 for the limit (d).

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

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

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. 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. 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. We are not able to find, even in the literature, a closed formula for its spectral gap. However, we give a lower bound not depending on the number of particles. The proofs are based on our main general theorem (coupling type argument) and a generalisation of [86] (Hardy's inequalities type argument). 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 [55] by :

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

In particular, if there exists $j\in F^*$ 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 η, $\eta'\in E$, 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

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

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'. The law of a random variable X is denoted by $\mathrm{Law}(X)$ and, along the chapter, we assume that

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

Our first main result is :

Theorem 3.1 (Wasserstein exponential ergodicity)**.** *If $\rho = \lambda - (\sup(p_0) - \inf(p_0))$ then for any processes $(\eta_t)_{t>0}$ and $(\eta'_t)_{t>0}$ generated by (3.1), and for any $t \geq 0$, we have*

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

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

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

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 [101], 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 [67, 85] and, contrary to us, the dependence on N of the rates and bounds are unknown. So, this gives less informations.

When the death rate p_0 is constant, our bound is optimal in terms of contraction. See for instance section 3.3, 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. See for instance Section 3.4, where we study the case where F^* contains only two elements. Even though in this case the study seems to be easy, we are not able to give a closed formula for the spectral gap (even if we give a lower bound in the general case). 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 3.8.

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 [95]. 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) $\mathbf{Q_1} = \sup_{i\in F^*} \sum_{\substack{j\in F^* \\ j\neq i}} Q_{i,j} < +\infty$ *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 3.2 (Convergence to the conditioned process)**.** *Under Assumption (A) and for $t \geq 0$, there exists $B, C > 0$ such that, for all $\eta \in E$, and any probability measure μ, we have*

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

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 3.3 (Covariance estimates)**.** *Let ρ be defined in Theorem 3.1. 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 [8, 55].

The previous theorem is a consequence of Theorem 3.11 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 [8, 55]. Indeed, as our proof differs completely to [8, 55] (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 $t \to +\infty$ 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 3.2 is a generalization of [8, Theorem 1.3], [55, Theorem 1.2] and of [63, Theorem 2.2]. Our assumptions are weaker and our convergence estimate is in a stronger form. We can also cite [42, Theorem 1.1] and [102, 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 3.2 differs from all these theorems ; it seems simpler and is only based on a Gronwall argument and the correlation estimates.

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 3.3 and 3.4. As all constants are explicit, the previous theorem allows us to consider parameters depending on N and to understand how the particle system evolves when Q varies with the size of the population ; see for instance Remark 3.25.

Two main consequences

We summarize two important consequences of our main theorems. Firstly, as ρ, defined in Theorem 3.1, does not depend on N, we can take the limit as $N \to +\infty$ in Theorem 3.1. 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 3.4 (Convergence to the QSD)**.** *Suppose that ρ is positive and that Assumption (A)*

holds. For any probability measure μ, ν, we have

$$\forall t \geq 0,\ d_{TV}(\mu T_t, \nu T_t) \leq e^{-\rho t} d_{TV}(\mu, \nu)\,. \tag{3.2}$$

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

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

This corollary is closely related to several previous work [37], [43, Theorem 1.1], [81, Theorem 3] and [55, Theorem 1.1]. When F is finite, the oldest result dates from 1967 [37] 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 [55, Theorem 1.1] in a stronger convergence and with an estimation of the rate of convergence. As in [43, 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 [43, Theorem 1.1] they only need that $\sup(p_0) < +\infty$ but, their mixing condition is stronger than ours. Finally [81, 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 3.13.

Without limiting results, several works establish existence and/or uniqueness of a QSD ; see [35, 84] for references. Let us mention the main result of [54]. In this article, the authors prove that, under the condition that $F = \mathbb{N}$, X is irreducible over $\mathbb{N}^*$ and

$$\lim_{x\to\infty} \mathbb{P}\left(T_0 < t \mid X_0 = x\right) = 0, \tag{3.3}$$

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 (3.3) can be restrictive. In particular, if the death rate is constant over $\mathbb{N}$, condition (3.3) 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 3.5 (Uniform bounds)**.** *If $\rho > 0$, then under the assumptions of Theorem 3.2, there* *exist $K_0, \gamma > 0$ such that , for every $\eta \in E$,*

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

All constants are explicit and given in (3.20)*.*

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

$$\mathbb{E}\left[|m(\eta)(\varphi)-\nu_{\mathrm{qs}}(\varphi)|\right]\leq\frac{K_0}{N^\gamma}, \tag{3.4}$$

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 [8, 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 3.14. Note also that, closely related, article [89] gives a similar result when the underlying dynamics is diffusive instead of discrete. Its approach is completly different and based on martingale properties and on spectral properties associated to Schrödinger equation.

The remainder of the chapter is as follows. Section 3.2 gives the proofs of our main theorems ; Subsection 3.2.1 contains the proof of Theorem 3.1, Subsection 3.2.2 the proof of Theorem 3.2 and the last subsection the proof of the corollaries. We conclude the chapter with Sections 3.3 and 3.4, 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.

3.2 Proof of the main theorems

In this section, we prove Theorems 3.1 and 3.2 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

$$\mathcal{L}f(\eta)=\sum_{i\in F^*}\eta(i)\sum_{j\in F^*}\left(Q_{i,j}+p_0(i)\frac{\eta(j)}{N-1}\right)\left(f(T_{i\to j}\eta)-f(\eta)\right). \tag{3.5}$$

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

Remark 3.6 (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*

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

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 $\inf(p_0)$ and why our propagation of chaos result depends on it.

Remark 3.7 (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 [18]. However, the assumption that p_0 is bounded is trivially sufficient to guarantee this non-explosion.*

3.2.1 Proof of Theorem 3.1

Proof of Theorem 3.1. We build a coupling between two Fleming-Viot particle systems, $(\eta_t)_{t\geq 0}$ and $(\eta'_t)_{t\geq 0}$, generated by (3.1), 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

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

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

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

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 can not do this but we can maximize the probability to coalesce. Indeed there is $N - d_1(\eta, \eta')$ 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 is $\eta(i) = \operatorname{card}\{k \mid \xi_k = i\}$ 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

$$\sum_{i \in F^*} (\eta(i) - \eta'(i))_+ = \sum_{i \in F^*} (\eta'(i) - \eta(i))_+ = d_1(\eta, \eta') = N - \sum_{i \in F^*} \eta(i) \wedge \eta'(i), \tag{3.6}$$

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

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

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. A Markov process having only

A_Q as jumps rate corresponds to a coupling of two systems of N particles evolving independently from each others. 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

$$\left(1 - \frac{\sum_{k\in F^*} \eta(k) \wedge \eta'(k) - 1}{N-1}\right) \frac{(\eta(j) - \eta'(j))_+}{\sum_{k\in F^*}(\eta(k) - \eta'(k))_+} \frac{(\eta'(j') - \eta(j'))_+}{\sum_{k\in F^*}(\eta'(k) - \eta(k))_+}, \tag{3.7}$$

    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 (3.7) 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 theses 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,

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

and

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

Our coupling is totally defined. Lemma 3.10 below shows that if a measurable function f on $E \times E$ does not depend on its second (resp. first) variable ; that is with a slight abuse of notation :

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

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

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

Now,

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

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 [31, 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 3.8 (Coalescent time estimate)**.** *For all $t\geq 0$, we have*

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

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

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

The proof is simple and given for sake of completeness.

Proof. Using Theorem 3.1, we find

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

□

Remark 3.9 (Generalization). *As we can see at the end of the chapter, 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 [29, 33, 49]. Another way is to find a contraction after a certain time : it is the Foster-Lyapunov-type techniques [14, 67, 85]. This type 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 chapter, 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 3.10 (Marginals of process generated by $\mathbb{L}$ are generated by $\mathcal{L}$). *With the notation of the proof of Theorem 3.1, 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 :*

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

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

$$\mathbb{L}_Q f(\eta,\eta') = \sum_{i\in F^*} (\eta(i)\wedge\eta'(i)) \sum_{j\in F^*} Q_{i,j}(f(T_{i\to j}\eta) - f(\eta)) \tag{3.8}$$

$$+ \sum_{i,i'\in F^*} \frac{(\eta(i)-\eta'(i))_+ \times (\eta'(i')-\eta(i'))_+}{d_1(\eta,\eta')} \tag{3.9}$$

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

$$+ \sum_{i,i'\in F^*} \frac{(\eta(i)-\eta'(i))_+ \times (\eta'(i')-\eta(i'))_+}{d_1(\eta,\eta')} \tag{3.10}$$

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

$$+ \sum_{i,i'\in F^*} \frac{(\eta(i)-\eta'(i))_+ \times (\eta'(i')-\eta(i'))_+}{d_1(\eta,\eta')} Q_{i,i'}(f(T_{i\to i'}\eta) - f(\eta)) \tag{3.11}$$

Using (3.6), we find

$$\begin{aligned}(3.9) + (3.10) + (3.11) &= \sum_{i,i'\in F^*} \frac{(\eta(i)-\eta'(i))_+ \times (\eta'(i')-\eta(i'))_+}{d_1(\eta,\eta')} \sum_{\substack{j\in F^*\\ j\neq i}} Q_{i,j}(f(T_{i\to j}\eta) - f(\eta)) \\ &= \sum_{i\in F^*} (\eta(i)-\eta'(i))_+ \sum_{\substack{j\in F^*\\ j\neq i}} Q_{i,j}(f(T_{i\to j}\eta) - f(\eta)) \\ &= \sum_{i\in F^*} (\eta(i)-\eta'(i))_+ \sum_{j\in F^*} Q_{i,j}(f(T_{i\to j}\eta) - f(\eta)).\end{aligned}$$

We deduce that

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

On the other hand,

$$\begin{aligned}
\mathbb{L}_p f(\eta,\eta') = &\sum_{i\in F^*} p_0(i)\,(\eta(i)\wedge\eta'(i)) \qquad (3.12)\\
&\times \sum_{j\in F^*}\left[\frac{\eta(j)\wedge\eta'(j)}{N-1} + \sum_{j'\in F^*}\frac{(\eta(j)-\eta'(j))_+\times(\eta'(j')-\eta(j'))_+}{(N-1)d_1(\eta,\eta')}\right](f(T_{i\to j}\eta)-f(\eta))\\
&+\sum_{i,i'\in F^*}(p_0(i)\wedge p_0(i'))\,\frac{(\eta(i)-\eta'(i))_+\times(\eta'(i')-\eta(i'))_+}{d_1(\eta,\eta')} \qquad (3.13)\\
&\times \sum_{j\in F^*}\left[\frac{\eta(j)\wedge\eta'(j)}{N-1} + \sum_{j'\in F^*}\frac{(\eta(j)-\eta'(j))_+\times(\eta'(j')-\eta(j'))_+}{(N-1)d_1(\eta,\eta')}\right](f(T_{i\to j}\eta)-f(\eta))\\
&+\sum_{i,i'\in F^*}(p_0(i)-p_0(i'))_+\,\frac{(\eta(i)-\eta'(i))_+\times(\eta'(i')-\eta(i'))_+}{d_1(\eta,\eta')} \qquad (3.14)\\
&\times \sum_{j\in F^*}\frac{\eta(j)}{N-1}(f(T_{i\to j}\eta)-f(\eta))
\end{aligned}$$

We have,

$$\begin{aligned}
(3.12)+(3.13) = &\sum_{i\in F^*}\left[p_0(i)\,(\eta(i)\wedge\eta'(i)) + \sum_{i'\in F^*}(p_0(i)\wedge p_0(i'))\,\frac{(\eta(i)-\eta'(i))_+\times(\eta'(i')-\eta(i'))_+}{d_1(\eta,\eta')}\right]\\
&\times\sum_{j\in F^*}\frac{\eta(j)}{N-1}(f(T_{i\to j}\eta)-f(\eta))
\end{aligned}$$

and

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

We deduce that

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

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. □

3.2.2 Proofs of Theorems 3.2 and 3.3

The proof of Theorem 3.2 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

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

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

$$\begin{aligned}\Gamma f(\eta) &= \mathcal{L}(f^2)(\eta) - 2f(\eta)\mathcal{L}f(\eta) \\ &= \sum_{i,j\in F^*} \eta(i)\left(Q_{i,j} + p_0(i)\frac{\eta(j)}{N-1}\right)\left(f(T_{i\to j}\eta) - f(\eta)\right)^2.\end{aligned} \tag{3.15}$$

We present now an improvement of Theorem 3.3.

Theorem 3.11 (Correlations for Lipschitz functional)**.** *Let g, h be two 1-Lipschitz mappings on (E, d_1) ; namely*

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

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

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

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

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

where η_{t-s}, η'_{t-s} evolve as Fleming-Viot particle systems with initial conditions η and $T_{i\to j}\eta$. Thus, using Theorem 3.1, we obtain

$$\begin{aligned} |\ S_{t-s}g(T_{i\to j}\eta) - S_{t-s}g(\eta)\ | &\leq \mathcal{W}_{d_1}(\mathrm{Law}(\eta_{t-s}), \mathrm{Law}(\eta'_{t-s})) \\ &\leq e^{-\rho(t-s)} d_1(T_{i\to j}\eta, \eta) \\ &\leq e^{-\rho(t-s)} \mathbb{1}_{i\neq j}. \end{aligned} \tag{3.16}$$

Hence,

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

Indeed, using (3.15) and (3.16) 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

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

□

Proof of Theorem 3.3. 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 3.12 (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 (3.1), and for any $t > 0$, we have*

$$\mathcal{W}_{d_1}(\mathrm{Law}(\eta_t), \mathrm{Law}(\eta'_t)) \leq Ce^{-\lambda t}\mathcal{W}_{d_1}(\mathrm{Law}(\eta_0), \mathrm{Law}(\eta'_0)), \tag{3.17}$$

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

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

The previous theorem is an instance of this implication with $C = 1$. For instance, Equation (3.17) is obtained when the state space F^* contains only two points.

Proof of Theorem 3.2. The proof is based on a bias-variance type decomposition. The variance is bounded through Theorem 3.3 and the bias through Gronwall-type argument. More precisely, for $t \geq 0$, we have

$$\sup_{\|\varphi\|_\infty\leq 1}\mathbb{E}_\eta\left[|m(\eta_t)(\varphi) - \mu T_t\varphi|\right] \leq \sup_{\|\varphi\|_\infty\leq 1}\mathbb{E}_\eta\left[|m(\eta_t)(\varphi) - \overline{m}(\eta_t)(\varphi)|\right] + 2d_{\mathrm{TV}}(\overline{m}(\eta_t), \mu T_t), \tag{3.18}$$

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

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

where $g_\varphi : \eta \mapsto \dfrac{1}{2}\displaystyle\sum_{k\in F^*} \eta(k)\varphi(k) = \dfrac{N}{2}m(\eta)(\varphi)$ is a 1-Lipschitz function. So by Theorem 3.11 we have

$$\sup_{\|\varphi\|_\infty \leq 1} \mathbb{E}_\eta\left[|m(\eta_t)(\varphi) - \overline{m}(\eta_t)(\varphi)|\right] \leq \sqrt{2\rho^{-1}(1-e^{-2\rho t})(\mathbf{Q_1} + \mathbf{p})N^{-1}}.$$

Now, to study the bias term in (3.18), 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

$$\forall t \geq 0,\ \partial_t \mu T_t(j) = \sum_{i\in F^*} \left(Q_{i,j}\ \mu T_t(i) +\ p_0(i)\ \mu T_t(i)\ \mu T_t(j)\right). \tag{3.19}$$

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 (3.1), 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

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

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

$$\begin{aligned}
\varepsilon(t) &= \sum_{k\in F^*} \left| u_k(0) - v_k(0) + \int_0^t \partial_s (u_k(s) - v_k(s)) ds \right| \\
&\leq \varepsilon(0) + \sum_{k\in F^*} \int_0^t \left| \sum_{i\in F^*} Q_{i,k}(u_i(s) - v_i(s)) \right| + \sum_{k\in F^*} \int_0^t \left(\frac{p_0(k)}{N-1} u_k(s) + |R_k(s)| \right) ds \\
&\quad + \sum_{k\in F^*} \int_0^t \left| \sum_{i\in F^*} p_0(i) \left[v_i(s)(u_k(s) - v_k(s)) + u_k(s)(u_i(s) - v_i(s)) \right] \right| ds \\
&\leq \varepsilon(0) + \int_0^t (\mathbf{Q_1} + 2\mathbf{p})\, \epsilon(s) ds + \frac{\mathbf{p}t}{N-1} + \int_0^t \sum_{k\in F^*} |R_k(s)| ds.
\end{aligned}$$

However, by Cauchy-Schwarz inequality and Theorem 3.11 with the 1-Lipschitz function $g : \eta \mapsto \frac{1}{2\mathbf{p}} \sum_{i\in F^*} p_0(i)\eta(i)$, we have

$$\begin{aligned}
\sum_{k\in F^*} |R_k(t)| &\leq \sum_{k\in F^*} \mathbb{E}_\eta \left(m(\eta_t)(k) \left| \sum_{i\in F^*} p_0(i) m(\eta_t)(i) - \mathbb{E}_\eta \left(\sum_{i\in F^*} p_0(i) m(\eta_t)(i) \right) \right| \right) + \mathbf{p}(N-1)^{-1} \\
&\leq 2\mathbf{p} N^{-1} \mathrm{Var}_\eta (g(\eta_t))^{\frac{1}{2}} + \mathbf{p}(N-1)^{-1} \\
&\leq \mathbf{p} \sqrt{2\rho^{-1}(1 - e^{-2\rho t})(\mathbf{Q_1} + \mathbf{p}) N^{-1}} + \mathbf{p}(N-1)^{-1}.
\end{aligned}$$

If $c_t = \rho^{-1}(1 - e^{-2\rho t})$, $B = \mathbf{Q_1} + 2\mathbf{p}$ then Gronwall's lemma gives

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

for some $A > 0$.

□

3.2.3 Proof of the corollaries

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

Proof of Corollary 3.4. The proof is based on an approximation of the conditioned semigroups by two particle systems. Theorem 3.1 gives a contraction for these particle systems. We then use Theorem 3.2 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 3.1 establishes the existence of a coupling between $\eta_t^{(N)}$ and $\widetilde{\eta}_t^{(N)}$, where each of its components is generated by (3.5), 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

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

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 3.2 and taking the limit $N \to +\infty$, we find

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

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

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

Furthermore all sequences in the expectations are increasing. Thus, taking the limit $n \to +\infty$, we obtain (3.2). Finally, the existence of a QSD can be proved as in the proof of [81, 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 [84, Proposition 1]. □

Remark 3.13 (Weaker assumptions). *Conclusion of Corollary 3.4 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 [102, 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 3.5. The proof is based on an "interpolation" between the bounds obtained in Corollary 3.4 and Theorem 3.2.

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

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

where $(\widetilde{\eta}_t)_{t\geq 0}$ is a Markov process generated by (3.1) 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, by Theorem 3.2, which is a uniform estimate on the initial condition, there exist $B, C > 0$ such that

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

On the other hand, from Corollary 3.4, 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

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

this gives

$$\sup_{\|\varphi\|_\infty \leq 1} \mathbb{E}_\eta\left[|m(\eta_t)(\varphi) - m(\eta)T_t\varphi|\right] \leq \frac{B+\rho}{B}\left(\frac{BC}{\rho\sqrt{N}}\right)^{\frac{\rho}{B+\rho}}. \tag{3.20}$$

□

Remark 3.14 (Weaker assumptions). *In the previous corollary, it is enough to assume that there exist $C > 0$ and $\lambda > 0$ such that*

$$\forall t \geq 0,\ d_{TV}(\mu T_t, \nu T_t) \leq Ce^{-\lambda t}, \tag{3.21}$$

to obtain a uniform bound. Some sufficient conditions to obtain (3.21) are given in [37, 43, 81]. We can also use a bound of convergence for the Fleming-Viot particle system as in Theorem 3.1.

As an application, a bound such as (3.4) 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.

3.3 Complete graph dynamics

In all this section, we study the example of a random walk on the complete graph. Let us fix $K \in \mathbb{N}^*$, $p > 0$ and $N \in \mathbb{N}^*$, the dynamics of this example is as follows : we consider a model with N particles and $K+1$ vertices $0, 1, \ldots, K$. The N particles move on the K vertices $1, \ldots, K$ uniformly at random and jump to 0 with rate p. 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

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

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

$$\mathcal{L}f(\eta) = \sum_{i=1}^{K} \eta(i) \left[\sum_{j=1}^{K} (f(T_{i \to j}\eta) - f(\eta)) \left(\frac{1}{K} + p\frac{\eta(j)}{N-1}\right)\right], \tag{3.22}$$

for every function f and $\eta \in E$.

A process generated by (3.22) is an instance of inclusion processes studied in [60, 64, 65]. It is then related to models of heat conduction. One main point of [60, 64] 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 (3.22) and we give this expression in Subsection 3.3.3. They also study different scaling limits which seem to be irrelevant for our problems.

Another application of this example comes from population genetics. Indeed,this model can also be referred as *neutral evolution*, see for instance [50, 105]. 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.

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 $\mathbb{P}_\mu = \mathbb{E}_\mu[\mathbb{1}.]$; similarly Cov_μ and Var_μ are defined with respect to $\mathbb{E}_\mu$.

3.3.1 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 $1/p$ 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 $\{1,\ldots,K\}$. 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.

3.3.2 Correlations at fixed time

The special form of $\mathcal{L}$, defined at (3.22), makes the calculation of the two-particle correlations at fixed time easy.

Theorem 3.15 (Two-particle correlations). *For all $k,l\in\{1,\ldots,K\}$, $k\neq l$ and any probability measure μ on E, we have for all $t\geq 0$*

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

Remark 3.16 (Limit $t\to+\infty$). *By the previous theorem, we find for any probability measure μ*

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

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

Remark 3.17 (Limit $N\to+\infty$). *If $Cov_\mu\left(\eta_0(k),\eta_0(l)\right)\neq 0$ then for all $k,l\in\{1,\ldots,K\}$, $k\neq l$ and any probability measure μ, we have*

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

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

Proof of Theorem 3.15. For $k,l \in \{1,..,K\}$, let $\psi_{k,l}$ be the function $\eta \mapsto \eta(k)\eta(l)$. Applying the generator (3.22) to $\psi_{k,l}$ we obtain

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

So, for all $t \geq 0$,

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

Using Kolmogorov's equation, we have

$$\partial_t \mathbb{E}_\mu(\eta_t(k)\eta_t(l)) = -\frac{2K(N-1+p)}{K(N-1)}\mathbb{E}_\mu(\eta_t(k)\eta_t(l)) + \frac{N-1}{K}(\mathbb{E}_\mu(\eta_t(k)) + \mathbb{E}_\mu(\eta_t(l))). \tag{3.23}$$

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 (3.23) ends the proof. □

3.3.3 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.

Theorem 3.18 (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*

$$\nu_N(\{\eta\}) = Z^{-1} \prod_{i=1}^{K} \prod_{j=0}^{\eta(i)-1} \frac{N-1+Kpj}{j+1},$$

where Z is a normalizing constant.

This result was already proved in [60, Section 4] and [64, Theorem 2.1] but we give it for sake of completeness.

Proof. A measure ν is reversible if and only if it satisfies the following balance equation

$$\nu(\{\eta\})C(\eta,\xi) = \nu(\{\xi\})C(\xi,\eta) \tag{3.24}$$

where $\xi = T_{i\to j}\eta$ and $C(\eta,\xi) = \mathcal{L}\mathbb{1}_\xi(\eta) = \eta(i)(K^{-1} + p\eta(j)(N-1)^{-1})$.

Due to the geometry of the complete graph, it is natural to consider that ν has the following form

$$\nu(\{\eta\}) = \frac{1}{Z}\prod_{i=1}^{K} l(\eta(i)),$$

where $l : \{0,\dots,N\} \to [0,1]$ is a function and Z is a normalizing constant. From (3.24), we have

$$l(\eta(i))l(\eta(j))\eta(i)(N-1+Kp\eta(j)) = l(\eta(i)-1)l(\eta(j)+1)(\eta(j)+1)(N-1+Kp(\eta(i)-1)),$$

for all $\eta \in E$ and $i,j \in \{1,\dots K\}$. Hence,

$$\frac{l(n)}{l(n-1)}\frac{n}{N-1+Kp(n-1)} = \frac{l(m)}{l(m-1)}\frac{m}{N-1+Kp(m-1)} = u,$$

for every $m,n \in \{1,\dots,N\}$ and some $u \in \mathbb{R}$. Finally,

$$\nu(\{\eta\}) = \prod_{i=1}^{K}\left(u^{\eta(i)}\prod_{j=0}^{\eta(i)-1}\frac{N-1+Kpi}{i+1}l(0)\right) = l(0)^K u^N \prod_{i=1}^{K}\prod_{j=0}^{\eta(i)-1}\frac{N-1+Kpj}{j+1},$$

and $Z = 1/(l(0)^K u^N)$. □

In particular, we have directly

Corollary 3.19 (Invariant distribution when $p = 1/K$)**.** *If $p = 1/K$ then 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

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

Corollary 3.20 (Marginal laws when $p = 1/K$). *If $p = 1/K$ then 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 satisfying $\sum_{i=1}^{p} n_i = n$, we have

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

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

$$|Cov_{\nu_N}(\eta(i)/N, \eta(j)/N)| \sim_N \frac{p+1}{K^2 N},$$

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 3.3.4, we have

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

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

$$\int \eta(i)^2 d\nu_N = \frac{N\left[(2N+K-2)(N-1) + 2KNp + K(N-1)\right]}{2K^2(N-1+p)}.$$

Finally,

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

and thus, for $i \neq j$,

$$|\mathrm{Cov}_{\nu_N}(\eta(i)/N, \eta(j)/N)| \sim_N \frac{p+1}{K^2N}.$$

□

Remark 3.22 (Proof through coalescence methods)**.** *Maybe we can use properties of Kingman's coalescent type process (which is a dual process) to recover some of our results (as for instance the previous correlation estimates). Indeed, after an interacting event, all individuals evolve independtly and it is enough to look when the first mutation happens (backwards in time) on one of the genealogical tree branches. Nevertheless, we prefer to use another approach based on Markovian techniques.*

Remark 3.23 (Number of sites)**.** *Theorem 3.21 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. For instance when $p = 1/K$ and if η is distributed under the invariant measure, then*

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

The previous theorem shows that the occupation numbers of two distinct sites become non-correlated when the number of particles increases. In fact, Theorem 3.21 leads to a propagation of chaos :

Corollary 3.24 (Convergence to the QSD)**.** *We have*

$$\mathbb{E}_{\nu_N}\left[d_{TV}(m(\eta),\pi_K)\right] \leq \sqrt{\frac{K(p+1)}{N}},$$

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

Proof. By the Cauchy-Schwarz inequality, we have

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

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

The previous bound is better than the bound obtained in Theorem 3.2 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 3.2 comes from, with the notations of its proof, the estimation of $|u_k(t)-v_k(t)|$ and Gronwall Lemma.

Remark 3.25 (Parameters depending on N)**.** *A nice application of explicit rates of convergence is to consider parameters depending on N. For instance, we can now consider that $p=p_N$ depends on N, this does not change neither the conditioned semi-goup nor the QSD but this changes the dynamics of our interacting-particle system. The last corollary gives that if $\lim_{N\to\infty} p_N/N = 0$ then the empirical measure converges to the uniform measure.*

3.3.4 Long time behavior and spectral analysis of the generator

In this subsection, we point out the optimality of Theorem 3.1 in this special case. Conditions in Theorem 3.1, 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.

Corollary 3.26 (Wasserstein contraction)**.** *For any processes $(\eta_t)_{t>0}$ and $(\eta'_t)_{t>0}$ generated by (3.22), and for any $t\geq 0$, we have*

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

In particular, when (η'_0) follows the invariant distribution ν_N associated to (3.22), we get for every $t \geq 0$

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

In particular, if λ_1 is the smallest positive eigenvalue of $-\mathcal{L}$, defined at (3.22), 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

$$\lim_{t \to +\infty} e^{2\lambda t}\|R_t f - \nu_N(f)\|^2_{L^2(\nu_N)} = 0, \tag{3.25}$$

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 [13, 90]. On the other hand, if $\lambda < 1$ then, by Theorem 3.1, we have

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

and then (3.25) 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

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

Remark 3.27 (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 $\sum_{i=1}^K a_i\varphi_i^{(1)}$ 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 an eigenvector of L with the same eigenvalue. Thus, it's enough to have an eigenvector of L to obtain an eigenvector of $\mathcal{L}$.*

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

Lemma 3.28 (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

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

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, satisfying $\mathcal{L}\psi_k^{(l)} = -\lambda_l \psi_k^{(l)}$ (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. $\square$

Remark 3.29 (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 3.30 (Marginals)**.** *For each k, the random process $(\eta_t(k))_{t \geq 0}$, which is a marginal of* *a process generated by* (3.22)*, is a Markov process on* $\mathbb{N}_N = \{0, \dots, N\}$ *generated by*

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

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* $\lambda_0, \lambda_1, \dots, \lambda_N$*.*

3.4 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

$$\begin{aligned}\mathcal{L}f(\eta) &= \eta(1)\left(a + p_0(1)\frac{\eta(2)}{N-1}\right)(f(T_{1\to 2}\eta) - f(\eta)) \\ &+ \eta(2)\left(b + p_0(2)\frac{\eta(1)}{N-1}\right)(f(T_{2\to 1}\eta) - f(\eta)).\end{aligned} \tag{3.26}$$

3.4.1 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 [84, 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 $A = a - b + p_0(1) - p_0(2)$. Also set $\nu = v_+/(v_+(1)+v_+(2))$. From these properties, we deduce that

Lemma 3.31 (Convergence to the QSD)**.** *There exists a constant $C > 0$ such that for every initial distribution μ, we have*

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

Proof. See [84, Theorem 7] and [84, Remark 3]. □

Note that

$$\lambda_+ - \lambda_- = \sqrt{(a+b)^2 + 2(a-b)(p_0(1)-p_0(2)) + (p_0(1)-p_0(2))^2} > a+b-(\sup(p_0)-\inf(p_0))$$

when $\sup(p_0) > \inf(p_0)$.

3.4.2 Explicit formula of the invariant distribution

Firstly note that, as

$$\forall \eta \in E, \eta(1)+\eta(2) = N,$$

each marginal of $(\eta_t)_{t\geq 0}$ is a Markov process :

Lemma 3.32 (Markovian marginals)**.** *The random process $(\eta_t(1))_{t\geq 0}$, which is a marginal of a process generated by (3.26), is a Markov process generated by $\mathcal{G}$ defined by*

$$\mathcal{G} f(n) = b_n(f(n+1)-f(n)) + d_n(f(n-1)-f(n)), \tag{3.27}$$

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 $(\eta_t)_{t\geq 0}$. $\square$

From this result and the already known results on birth and death processes [29, 31], we deduce that $(\eta_t(1))_{t\geq 0}$ admits an invariant and reversible distribution π given by

$$\pi(n) = u_0 \prod_{k=1}^{n} \frac{b_{k-1}}{d_k} \quad \text{and} \quad u_0^{-1} = 1 + \sum_{k=1}^{N} \frac{b_0 \cdots b_{k-1}}{d_1 \cdots d_k},$$

for every $n \in \mathbb{N}_N$. 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, as $\eta_t(2) = N - \eta_t(1)$, the process $(\eta_t(2))_{t\geq 0}$ is a Markov process whose invariant distribution is also easily calculable. The invariant law of $(\eta_t)_{t\geq 0}$, is then given by

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

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.

3.4.3 Rate of convergence

Applying Theorem 3.1, in this special case, we find :

Corollary 3.33 (Wasserstein contraction)**.** *For any processes $(\eta_t)_{t>0}$ and $(\eta'_t)_{t>0}$ generated by (3.26), and for any $t \geq 0$, we have*

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

where $\rho = a + b - (\sup(p_0) - \inf(p_0))$. In particular, when (η'_0) follows the invariant distribution ν_N of (3.26), we get for every $t > 0$

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

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 induces a coupling between two processes generated by $\mathcal{G}$ defined by (3.27). More precisely, let $\mathbb{L} = \mathbb{L}_Q + \mathbb{L}_p$ be the generator of our coupling introduced in the proof of Theorem 3.1 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 satisfies, 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

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

Now, for any sequence of positive numbers $(u_k)_{k\in\{0,\ldots,N-1\}}$, 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

$$\lambda_u = \min_{k\in\{0,\ldots,N-1\}} \left[d_{k+1} - d_k\frac{u_{k-1}}{u_k} + b_k - b_{k+1}\frac{u_{k+1}}{u_k} \right],$$

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 $t \ge 0$,

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

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 (3.26), and for any $t \ge 0$, we have

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

Finally, using [31, 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

$$\mathcal{W}_{d_1}(\mathrm{Law}(\eta_t), \mathrm{Law}(\eta'_t)) \le C(N)e^{-\lambda_N t}\mathcal{W}_{d_1}(\mathrm{Law}(\eta_0), \mathrm{Law}(\eta'_0)), \tag{3.28}$$

where $C(N)$ and λ_N are two constants depending on N. In conclusion, the coupling introduced in Theorem 3.1 gives the optimal rate of convergence but we are not able to express a precise expression of λ_N and $C(N)$. Nevertheless, in the section that follows, we will prove that, whatever the value of the parameters, the spectral gap is always bounded from below by a positive constant not depending on N.

3.4.4 A lower bound for the spectral gap

In this subsection, we study the evolution of $(\lambda_N)_{N\geq 0}$. Calculating λ_N for small value of N (it is the eigenvalue of a small matrix) and some different parameters show that, in general, this sequence is not monotone and seems to converge to $\lambda_+ - \lambda_-$. We are not able to prove this, but as it is trivial that for all $N \geq 0$, $\lambda_N > 0$, we can hope that it is bounded from below. The aim of this section is to prove this fact.

Firstly, using similar arguments of subsection 3.3.4, we have $\lambda_N \geq \rho$, for every $N \geq 0$. This result does not give us information in the case $\rho \leq 0$. However, we can use Hardy's inequalities [5, Chapter 6] and mimic some arguments of [86] to obtain :

Theorem 3.34 (A lower bound for the spectral gap)**.** *If $\rho \leq 0$ then there exists $c > 0$ such that*

$$\forall N \geq 0, \;\; \lambda_N > c.$$

The rest of this subsection aims to prove this result. Hardy's inequalities are mainly based on the estimation of the quantities $B_{N,+}$ and $B_{N,-}$ defined for every $i \in \mathbb{N}$ by

$$B_{N,+}(i) = \max_{x>i} \left(\sum_{y=i+1}^{x} \frac{1}{\pi(y) d_y} \right) \pi([x, N]), \tag{3.29}$$

and

$$B_{N,-}(i) = \max_{x<i} \left(\sum_{y=x}^{i-1} \frac{1}{\pi(y) b_y} \right) \pi([1, x]).$$

We recall that $\pi = \pi_N$ is the invariant distribution defined in Subsection 3.4.2 and jumps rates b and d also depend on N.

More precisely, [86, Proposition 3] shows that if one wants to get a "good" lower bound of the spectral gap, one only needs to guess an "adequate choice" of i and to apply the estimate

$$\lambda_N \geq \frac{1}{4 \max\{B_{N,+}(i), B_{N,-}(i)\}}.$$

So, we have to find an upper bound for these two quantities. Before to give it, let us prove that the invariant distribution π is unimodal. Indeed, it will help us to choose an appropriate i.

Lemma 3.35 (Unimodality of π)**.** *The sequence $(\pi(i+1)/\pi(i))_{i\geq 0}$ is decreasing.*

Proof of Lemma 3.35. For all $i \in \{1, \dots, N\}$, we set

$$g(i) = \frac{\pi(i+1)}{\pi(i)} = \frac{(N-i)(b(N-1) + i p_0(2))}{(i+1)((a + p_0(1))(N-1) - i p_0(1))}.$$

It follows that

$$g(i+1)-g(i)=\frac{\Lambda_N(i)}{(i+1)((a+p_0(1))(N-1)-ip_0(1))(i+2)((a+p_0(1))(N-1)-(i+1)p_0(1))}$$

where

$$\begin{aligned}\Lambda_N(i) &= (N-i-1)(b(N-1)+(i+1)p_0(2))(i+1)((a+p_0(1))(N-1)-ip_0(1))\\ &\quad -(N-i)(b(N-1)+ip_0(2))(i+2)((a+p_0(1))(N-1)-(i+1)p_0(1))\\ &= -\left[b(N-1)-p_0(2)\right]\left[(N+1)\left(a(N-1)-p_0(1)\right)+p_0(1)(N-i)(N-i-1)\right]\\ &\quad -p_0(2)\left(i^2+3i+2\right)\left(a(N-1)-p_0(1)\right)\\ &\leq 0.\end{aligned}$$

We deduce the result. □

Proof of Theorem 3.34. Without less of generality, we assume that $p_0(1)\geq p_0(2)$ and we recall that $\rho\leq 0$. We would like to know where π reaches its maximum i^* since it will be a good candidate to estimate $B_{N,+}(i^*)$ and $B_{N,-}(i^*)$. From the previous lemma, to find it, we look when $\pi(i+1)/\pi(i)$ is close to one. We have, for all $i\in\{1,\dots,N\}$,

$$\frac{\pi(i+1)}{\pi(i)}=\frac{b_i}{d_{i+1}}=1+\frac{(p_0(1)-p_0(2))(i-i_1)(i-i_2)}{(i+1)\left((a+p_0(1))(N-1)-ip_0(1)\right)}, \tag{3.30}$$

where i_1 and i_2 are the two real numbers given by

$$i_1=\frac{N(a+b+p_0(1)-p_0(2))-(a+b+2p_0(1))-\sqrt{\Delta}}{2(p_0(1)-p_0(2))}$$

and

$$i_2=\frac{N(a+b+p_0(1)-p_0(2))-(a+b+2p_0(1))+\sqrt{\Delta}}{2(p_0(1)-p_0(2))},$$

where

$$\begin{aligned}\Delta &= \left[N(a+b+p_0(1)-p_0(2))-(a+b+2p_0(1))\right]^2\\ &\quad -4(N-1)(bN-a-p_0(1))(p_0(1)-p_0(2)).\end{aligned}$$

In particular, $1 \leq i_1 \leq N \leq i_2$. Furthermore, if $\lfloor . \rfloor$ denotes the integer part then

$$\frac{\pi(\lfloor i_1 \rfloor + 2)}{\pi(\lfloor i_1 \rfloor + 1)} \leq 1 \leq \frac{\pi(\lfloor i_1 \rfloor + 1)}{\pi(\lfloor i_1 \rfloor)}.$$

Let us define $m_N = \lfloor i_1 \rfloor + 1$ and $l_N = 2(\lfloor \sqrt{N} \rfloor + 1)$. Using a telescopic product, we have

$$\frac{\pi(m_N + l_N)}{\pi(m_N)} = \frac{\pi(m_N + l_N - \lfloor \sqrt{N} \rfloor - 1)}{\pi(m_N)} \prod_{j=1}^{\lfloor \sqrt{N} \rfloor + 1} \frac{\pi(m_N + l_N - j + 1)}{\pi(m_N + l_N - j)},$$

Using Lemma 3.35 and the previous calculus, we have that the sequences $(\pi(i))_{i \geq m_N}$ and $(\pi(i+1)/\pi(i))_{i \geq 0}$ are decreasing and then

$$\frac{\pi(m_N + l_N)}{\pi(m_N)} \leq \left(\frac{\pi(m_N + l_N - \lfloor \sqrt{N} \rfloor)}{\pi(m_N + l_N - \lfloor \sqrt{N} \rfloor - 1)} \right)^{\lfloor \sqrt{N} \rfloor + 1}.$$

Now using (3.30) and some equivalents, there exists a constant $\delta_1 > 0$ (not depending on N) such that

$$\frac{\pi(m_N + l_N - \lfloor \sqrt{N} \rfloor)}{\pi(m_N + l_N - \lfloor \sqrt{N} \rfloor - 1)} \leq 1 - \frac{\delta_1}{\sqrt{N}}.$$

Using the fact that $1 - x \leq e^{-x}$ for all $x \geq 0$, we finally obtain $\pi(m_N + l_N)/\pi(m_N) \leq e^{-\delta_1}$. Similar arguments entail the existence of $\delta_2 > 0$ (also not depending on N) such that $\pi(m_N - l_N)/\pi(m_N) \leq e^{-\delta_2}$. In conclusion, using Lemma 3.35, we have shown that for all $i \geq m_N$ and $j \leq m_N$, the following inequalities holds :

$$\pi(i + l_N) \leq e^{-\delta_1} \pi(i) \quad \text{and} \quad \pi(j - l_N) \leq e^{-\delta_2} \pi(j).$$

We are now armed to evaluate $B_{N,+}(m_N)$ defined in (3.29). Firstly, using the expressions of the death rate d and m_N, there exist $\gamma > 0$ (not depending on N) and $N_0 \geq 0$ such that for all $N \geq N_0$ and all $i \geq m_N + 1$, $d_i \geq \gamma N$. Let us fix $x \geq m_N + 1$, using that $(\pi(i))_{i \geq m_N}$ is decreasing, we have

$$\begin{aligned}\sum_{y=m_N+1}^{x} \frac{1}{\pi(y)} &= \sum_{\{i,k|m_N+1\le k-il_N\le x\}} \frac{1}{\pi(k-il_N)}\\ &\le \sum_{\{i,k|m_N+1\le k-il_N\le x\}} \frac{e^{-\delta_1 i}}{\pi(k)}\\ &\le \frac{1}{1-e^{-\delta_1}} \sum_{k=x-l_N+1}^{x} \frac{1}{\pi(k)}\\ &\le \frac{l_N}{\pi(x)} \frac{1}{1-e^{-\delta_1}}.\end{aligned}$$

Similarly, we have

$$\pi([x,N]) = \sum_{\{k,i|x\le k+il_N\le N\}} \mathbb{1}_{\{x+il_N\le N\}} \Pi_N(k+il_N) \le \frac{l_N \pi(x)}{1-e^{-\delta_1}}.$$

Using these three estimates, we deduce that, for every $N \ge N_0$,

$$B_{N,+}(m_N) \le \frac{1}{\gamma N}\left(\frac{l_N}{1-e^{-\delta_1}}\right)^2 \le \frac{1}{\gamma N}\left(\frac{2(\sqrt{N}+1)}{1-e^{-\delta_1}}\right)^2 \le \frac{16}{\gamma(1-e^{-\delta_1})}.$$

The study of $B_{N,-}(m_N)$ is similar. $\square$

3.4.5 Correlations

Using Theorem 3.11, we have

Corollary 3.36 (Correlations). *If $(\eta_t)_{t\ge 0}$ is a process generated by (3.26) then we have for all $t \ge 0$,*

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

If $\rho \le 0$, the right-hand side of the previous inequality explodes as t tends to infinity whereas these correlations are bounded by 1. Nevertheless, using Theorem 3.11, Remark 3.12 and Inequality (3.28), we can prove that there exists two constants $C'(N)$, depending on N, and K, which does not depend on N, such that

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

where $C(N)$ is defined in (3.28). Even if Theorem 3.34 gives an estimate of λ_N, $C(N)$ 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 $\sup(p_0) \neq \inf(p_0)$.

Chapitre 4

Birth and Death Process in Mean Field type Interaction *

4.1 Introduction

The concept of *mean field interaction* arised in statistical physics with Kac [72] and then McKean [82] in order to describe the collisions between particles in a gas, and has later been applied in other areas such as biology or communication networks. A particle system is in mean field interaction when the system acts over one fixed particle through the empirical measure of the system. For continuous interacting diffusions, this linear particle system has been introduced in order to approximate the solution of a non linear equation, the so-called McKean-Vlasov equation, and has been extensively studied by many authors. But, to our knowledge, there are few results in discrete space.

In this paper, we give a discrete version of the particle approximation of the McKean-Vlasov equations. We consider a system of N particles $X^{1,N}, \ldots, X^{N,N}$ evolving in $\mathbb{N} = \{0, 1, 2, \ldots\}$, each one according to a birth and death process in mean field type interaction. Namely, a single particle evolves with a rate which depends on both its own position and the mean position of all particles. At time 0, the particles are independent and identically distributed and at time $t > 0$, the interaction is given in terms of the mean of the particles at this time. Let $q^+ : \mathbb{N} \times \mathbb{R}_+ \to \mathbb{R}_+$ and $q^- : \mathbb{N} \times \mathbb{R}_+ \to \mathbb{R}_+$ be the interaction functions. For any particle

∗. Submitted to Bernoulli [97]

$k \in \{1, \dots, N\}$ and position $i \in \mathbb{N}$, the transition rates at time t, given $\left(X_t^{1,N}, \dots, X_t^{N,N}\right)$ and $X_t^{k,N} = i$ are

$$\begin{array}{lllll} i & \to & i+1 & \text{with rate } b_i + q^+(i, M_t^N), & \\ i & \to & i-1 & \text{with rate } d_i + q^-(i, M_t^N), & \text{for } i \geq 1 \\ i & \to & j & \text{with rate } 0, & \text{if } j \notin \{i-1, i+1\}, \end{array}$$

where $b_i > 0$ for $i \geq 0$, $d_i > 0$ for $i \geq 1$, $d_0 = q^-(0, m) = 0$ for all $m \in \mathbb{R}_+$ and M_t^N denotes the mean of the N particles at time t, defined by

$$M_t^N = \frac{1}{N} \sum_{i=1}^N X_t^{i,N}.$$

By the positivity of the rates b_i, d_i, the process is irreducible but not necessary reversible. The generator of the particle system acts on bounded functions $f : \mathbb{N}^N \to \mathbb{R}$ as follows

$$\begin{aligned} \mathcal{L}f(x) = \sum_{i=1}^N \Big[& \left(b_{x_i} + q^+(x_i, M^N)\right) \left(f(x + e_i) - f(x)\right) \\ & + \left(d_{x_i} + q^-(x_i, M^N)\right) \left(f(x - e_i) - f(x)\right) \mathbb{1}_{x_i > 0} \Big], \end{aligned} \tag{4.1}$$

where $e_i, i \in \mathbb{N}$ is the canonical vector and

$$M^N = M^N(x) = \frac{1}{N} \sum_{i=1}^N x_i.$$

For such systems, we consider the limiting behavior as time and the size of the system go to infinity. As N tends to infinity, this leads to the propagation of chaos phenomenon : the law of a fixed number of particles becomes asymptotically independent as the size of the system goes to infinity [95]. Sznitman [95, Proposition 2.2] (see also Méléard [83, Proposition 4.2]) showed that, in the case of exchangeable particles, this is equivalent to the convergence in law of the empirical measure to a deterministic measure. This limiting measure, denoted by u and often called the mean field limit, is characterized by being the unique weak solution of the nonlinear master equation

$$\begin{cases} \dfrac{d}{dt} \langle u_t, f \rangle = \langle u_t, \mathcal{G}_{u_t} f \rangle \\ \qquad u_0 \in \mathcal{M}_1(\mathbb{N}), \end{cases} \tag{4.2}$$

where $\mathcal{M}_1(\mathbb{N})$ is the set of probability measures on $\mathbb{N}$ and $\mathcal{G}_{(\cdot)}$ is the operator defined through the behavior of the particle system by

$$\mathcal{G}_{u_t} f(i) = \Big(b_i + q^+(i, \|u_t\|)\Big) \left(f(i+1) - f(i)\right) + \Big(d_i + q^-(i, \|u_t\|)\Big) \left(f(i-1) - f(i)\right) \mathbb{1}_{i>0}, \quad (4.3)$$

for every $i \in \mathbb{N}$. For any probability measure u and bounded function f, the linear form $\langle u, f\rangle$ is defined by

$$\langle u, f\rangle = \sum_{k\in\mathbb{N}} f(k)u(\{k\}),$$

and

$$\|u\| = \langle u(dx), x\rangle \text{ is the first moment of } u.$$

We use the notation $\|\cdot\|$ to be consistent with the previous works [39, 52]. The propagation of chaos phenomenon can be seen as giving both the asymptotic behavior of an interacting particle system and an approximation of solutions of nonlinear differential equations.

We introduce a stochastic process $(\overline{X}_t)_{t\geq 0}$ whose time marginals are solutions of the nonlinear equation (4.2). This process is defined as a solution of the following martingale problem : for a nice test function φ

$$\varphi(\overline{X}_t) - \varphi(\overline{X}_0) - \int_0^t \mathcal{G}_{u_s}\varphi(\overline{X}_s)ds \text{ is a martingale,}$$

where $u_s = u \circ \overline{X}_s^{-1}$ and $u_0 = u \circ \overline{X}_0^{-1}$.

Under some assumptions, the existence and uniqueness of the solution of the martingale problem are insured. If $q^- \equiv 0$, this has been proved by Dawson, Tang and Zhao in [39] and before by Feng and Zheng [52] in the case where $q^+(i, \|u_t\|) = \|u_t\|$ for all $i \in \mathbb{N}$. This unique solution is a solution of (4.2) (cf [39, Theorem 2], [52, Theorem 1.2]).

For any $t \geq 0$, let us denote by μ_t^N the empirical distribution of $(X^{1,N}, \dots, X^{N,N})$ at time t defined by

$$\mu_t^N = \frac{1}{N}\sum_{i=1}^N \delta_{X_t^{i,N}} \in \mathcal{M}_1(\mathbb{N}). \quad (4.4)$$

This measure is a random measure on $\mathbb{N}$ and we note that the first moment of μ_t^N is exactly the value of M_t^N. Our goal is to quantify the following limits (if they exist)

$$
\begin{tikzcd}
& \mu_t^N \arrow[dl, "N\to\infty"'] \arrow[dr, "t\to\infty"] & \\
u_t \arrow[dr, "t\to\infty"'] & & \mu_\infty^N \arrow[dl, "N\to\infty"] \\
& u_\infty &
\end{tikzcd}
$$

Similar problems for interacting diffusions of the form

$$dX_t^{i,N} = \sqrt{2}dB_t^{i,N} - \nabla V(X_t^{i,N})dt - \frac{1}{N}\sum_{j=1}^{N} \nabla W(X_t^{i,N} - X_t^{j,N})dt \quad 1 \leq i \leq N,$$

have been studied in several works [27, 77, 78]. Here, $(B_t^{i,N})_{1\leq i\leq N}$ are N independent Brownian motions, V and W are two potentials and the symbol ∇ stands for the gradient operator. The mean field limit associated to these dynamics satisfies the so-called McKean-Vlasov equation.

Our initial motivation was the study of the Fleming-Viot type particle systems. In this system, the particles evolve as independent copies of a Markov process until one of them reaches the absorbing state. At this moment, the absorbed particle goes instantaneously to a state chosen with the empirical distribution of the particles remaining in the state space. For example, if we consider N random walks on $\mathbb{N}$ with a drift towards the origin (cf [10, 80]), the Fleming-Viot system can be interpreted as a system of N $M/M/1$ queues in interaction : when a queue is empty, another duplicates. The Fleming-Viot process has been introduced in order to approximate the solutions of a nonlinear equation : the quasi-stationary distributions. For results and related methods, we refer to [8–10, 24, 34, 55] and references therein.

Long time behavior of the particle system

Our starting point is the long time behavior of the interacting particle system. We show that under some conditions, the particle system converges exponentially fast to equilibrium for a suitable Wasserstein coupling distance. Let us describe first the different distances that we use and the assumptions that we make. For $x, \overline{x} \in \mathbb{N}^N$, let d be the ℓ^1-distance defined by

$$d(x,\overline{x}) = \sum_{k=1}^{N} |x_k - \overline{x}_k|, \tag{4.5}$$

and for any two probability measures μ and μ' on $\mathbb{N}^N$, let $\mathcal{W}_d(\mu, \mu')$ be the Wasserstein coupling distance between these two laws defined by

$$\mathcal{W}_d(\mu, \mu') = \inf_{\substack{X \sim \mu \\ \overline{X} \sim \mu'}} \mathbb{E}\left[d(X, \overline{X})\right], \tag{4.6}$$

where the infimum runs over all the couples of random variables with marginal laws μ and μ'. Let us assume that :

Assumption.

(*A*) *(Convexity condition) There exists* $\lambda > 0$ *such that*

$$\nabla^+(d - b) \geq \lambda,$$

where for every $n \geq 0$ *and* $f : \mathbb{N} \to \mathbb{R}_+$

$$\nabla^+(f)(n) = f(n+1) - f(n).$$

(*B*) *(Lipschitz condition) The function* q^+ *(resp.* q^-*) is non-decreasing (resp. non-increasing) in its second component and there exists* $\alpha > 0$ *such that for any* $(k_1, l_1), (k_2, l_2) \in \mathbb{N} \times \mathbb{R}_+$

$$|\left(q^+ - q^-\right)(k_1, l_1) - \left(q^+ - q^-\right)(k_2, l_2)| \leq \alpha \left(|k_1 - k_2| + |l_1 - l_2|\right).$$

An example of rates satisfying the assumptions (A) and (B) is given in Example 4.7. Denoting by $\mathrm{Law}(Y)$ the law of a random variable Y, we have

Theorem 4.1 (Long time behavior)**.** *Assume that Assumptions* (A) *and* (B) *are satisfied. Then, for all* $t \geq 0$

$$\mathcal{W}_d(\mathrm{Law}(X_t^N), \mathrm{Law}(Y_t^N)) \leq e^{-(\lambda - 2\alpha)t} \mathcal{W}_d(\mathrm{Law}(X_0^N), \mathrm{Law}(Y_0^N)), \tag{4.7}$$

where $(X_t^N)_{t \geq 0}$ *and* $(Y_t^N)_{t \geq 0}$ *are two processes generated by* (4.1).
In particular, if $\lambda - 2\alpha > 0$ *there exists a unique invariant distribution* λ_N *satisfying for every* $t \geq 0$,

$$\mathcal{W}_d(\mathrm{Law}(X_t^N), \lambda_N) \leq e^{-(\lambda - 2\alpha)t} \mathcal{W}_d(\mathrm{Law}(X_0^N), \lambda_N).$$

To our knowledge, it is the first theorem which establishes an exponential convergence of this interacting particle system in discrete state space and with an explicit rate. For continuous interacting diffusions, Malrieu [77, 78] proved that under some assumptions of convexity of the potentials and based on the Bakry-Émery criterion, the system of particles associated with the McKean-Vlasov equations converges exponentially fast to equilibrium with a rate that does not depend on N in terms of relative entropy.

Propagation of chaos

When N is large, we would like to show that the random empirical distribution μ_t^N is close to the deterministic measure u_t, solution of (4.2) at time t. For this convergence, one of the interesting points is to obtain a quantitative bound.

This behavior has been studied by Dawson,Tang and Zhao [39] and Dawson and Zheng [40], the key ingredient of the proof being the tightness of $\{\mu^N : N \geq 1\}$. For interacting diffusions, this has been studied particularly by Sznitman [95], Méléard [83] or Malrieu [77, 78]. The central limit theorem and large deviation principles for such models have also been studied by many authors [38, 51, 82, 92, 94, 96].

In the following theorem, we prove the propagation of chaos property and show that this property is uniform in time. For this purpose, we introduce a family of independent processes $(\overline{X}^i)_{i\in\{1,\ldots,N\}}$ of law u_t at each time t such that for $i \in \{1, \ldots, N\}$

- $\overline{X}_0^i = X_0^{i,N}$
- the transition rates of $\overline{X}^i$ at time t are given by

$$
\begin{array}{lllll}
i & \rightarrow & i+1 & \text{with rate } b_i + q^+(i, \|u_t\|), & \\
i & \rightarrow & i-1 & \text{with rate } d_i + q^-(i, \|u_t\|), & \text{for } i \geq 1 \\
i & \rightarrow & j & \text{with rate } 0, & \text{if } j \notin \{i-1, i+1\}.
\end{array}
$$

For $i \in \{1, \ldots, N\}$ and $t \geq 0$, the process $\overline{X}_t^i$ is said to be nonlinear in the sense that its dynamic depends on its law.

Theorem 4.2 (Uniform Propagation of chaos). *Let $\delta > 0$ and $\beta(\delta) := \inf_{x\in\mathbb{N}^*} \left(d_x e^{-\delta} - b_x\right)$ and let $K_1 = \alpha \left(\|u_0\| + \dfrac{b_0}{\lambda - 2\alpha}\right)$. Assume that Assumptions (A) and (B) are satisfied. Then, if $\lambda - 2\alpha > 0$ and $\beta(\delta) - K_1 > 0$, there exists a coupling and a constant $K > 0$ such that*

$$
\sup_{t\geq 0} \mathbb{E}|X_t^{1,N} - \overline{X}_t^1| \leq \frac{K}{\sqrt{N}}. \tag{4.8}
$$

For continuous interacting diffusions, Malrieu [77, 78] establishes for the first time a uniform (in time) propagation of chaos in the case of uniform convexity of the potentials. Cattiaux, Guillin, Malrieu [27] extend his results when the potentials are no more uniformly convex. The proof is based on a coupling argument and Itô's formula.

A consequence of the uniform propagation of chaos phenomenon is the convergence of the empirical measure μ_t^N to the solution u_t of the nonlinear equation. This convergence is well known but, here, we give an explicit rate. To express this convergence, we set for a function $\varphi : \mathbb{N} \to \mathbb{R}$

$$\mu_t^N(\varphi) = \sum_{k\in\mathbb{N}} \varphi(k)\mu_t^N(k), \text{ and } u_t(\varphi) = \sum_{k\in\mathbb{N}} \varphi(k)u_t(k).$$

Moreover, we denote by $\|\varphi\|_{Lip}$ the quantity defined by

$$\|\varphi\|_{Lip} = \sup_{\substack{x,y\in\mathbb{N}\\ x\neq y}} \frac{|\varphi(x)-\varphi(y)|}{|x-y|}.$$

Corollary 4.3 (Convergence of the empirical measure)**.** *Under the assumptions of Theorem 4.2, and if $\lambda - 2\alpha > 0$, there exists a constant $C > 0$ such that*

$$\sup_{t\geq 0} \sup_{\|\varphi\|_{Lip}\leq 1} \mathbb{E}|\mu_t^N(\varphi) - u_t(\varphi)| \leq \frac{C}{\sqrt{N}}. \tag{4.9}$$

In particular, from Markov's inequality, we have for $\epsilon > 0$

$$\sup_{t\geq 0} \sup_{\|\varphi\|_{Lip}\leq 1} \mathbb{P}\left(\left| \frac{1}{N}\sum_{i=1}^{N} \varphi(X_t^{i,N}) - \int \varphi du_t \right| > \epsilon \right) \leq \frac{C}{\epsilon\sqrt{N}}. \tag{4.10}$$

One can expect to improve this bound and obtain an exponential one. An exponential bound has been obtained by Malrieu [77] for interacting diffusions : as said previously, under some assumptions on the convexity of potentials, Malrieu showed that the law of $(X_t^N)_{t\geq 0}$ satisfies a Logarithmic Sobolev inequality with a constant independent of t and N. As a consequence, via the Herbst's argument, the law of the system satisfies a Gaussian concentration inequality around its mean. Our particle system does not verify a Logarithmic Sobolev inequality, this is the difference with interacting diffusions and that is the difficulty. Under an additional assumption on the number of particles N, Bolley, Guillin and Villani [21] improve the deviation inequality of Malrieu and obtain an exponential bound of $\mathbb{P}(\sup_{0\leq t\leq T} \mathcal{W}(\mu_t^N, u_t) > \epsilon)$ for every

$\epsilon > 0$ and $T \geq 0$. A Poisson type deviation bound has been established by Joulin [71] for the empirical measure of birth and death processes with unbounded generator. With the distance that we introduced (namely the ℓ_1-distance), we can not apply directly his results. Indeed, one of the hypotheses is the existence of a constant V such that

$$\| \sum_y d^2(\cdot, y) Q(\cdot, y) \|_\infty \leq V^2,$$

where Q is the transition rates matrix and d a metric on $\mathbb{N}$. However, if we consider the ℓ_1-distance, V is infinite. So, to apply Joulin's results, we need to choose another distance in such a way that the previous assumption is satisfied.

Although we are not able to provide an exponential bound of (4.10), we can measure how the empirical measure μ_t^N is close to the law u_t in a stronger way. For this purpose, let us consider the Wasserstein distance $\mathcal{W}_1$ associated with the $d(x, y) = |x - y|$ for $x, y \in \mathbb{R}$.

Corollary 4.4 (Deviation inequality). *Under the assumptions of Theorem 4.2, and if $\lambda - 2\alpha > 0$, there exists a constant $\overline{C} > 0$ such that*

$$\sup_{t \geq 0} \mathbb{P}\left(\mathcal{W}_1(\mu_t^N, u_t) > \epsilon\right) \leq \frac{\overline{C}}{\epsilon\sqrt{N}}.$$

Long time behavior of the nonlinear process

The long time behavior of u_t is a consequence of Theorems 4.1 and 4.2. We express the convergence of u_t to equilibrium with an explicit rate, under the Wasserstein distance $\mathcal{W}$.

Theorem 4.5 (Long time behavior). *Let us assume that the assumptions of Theorem 4.2 hold. Let $(u_t)_{t\geq 0}$ and $(v_t)_{t\geq 0}$ be the solutions of (4.2) with initial conditions u_0 and v_0 respectively. Then, under the assumption $\lambda - 2\alpha > 0$*

$$\mathcal{W}_1(u_t, v_t) \leq e^{-(\lambda - 2\alpha)t} \mathcal{W}_1(u_0, v_0). \tag{4.11}$$

In particular, the nonlinear process $(\overline{X}_t)_{t\geq 0}$ associated with the equation (4.2) has a unique invariant measure u_∞ and

$$\mathcal{W}_1(u_t, u_\infty) \leq e^{-(\lambda - 2\alpha)t} \mathcal{W}_1(u_0, u_\infty). \tag{4.12}$$

The exponential convergence to equilibrium of the nonlinear process has been obtained by Malrieu [77, 78] for interacting diffusions under the Wasserstein distance $\mathcal{W}_2$. To prove this convergence, he used the uniform propagation of chaos, the exponential convergence to equilibrium of the particle system and the Talagrand transport inequality T_2 (connecting the Wasserstein distance and the relative entropy). Later, Cattiaux, Guillin and Malrieu [27] complete his result by giving the distance between two solutions of the McKean-Vlasov equation (granular media equation) starting at different points.

Finally, from Corollary 4.3 and Theorem 4.5, we have the convergence of the empirical measure under the invariant distribution that we denote by μ_∞^N.

Corollary 4.6 (Convergence under the invariant distribution)**.** *Assume that the assumptions of Corollary 4.3 are satisfied. Assume moreover that $\lambda - 2\alpha > 0$. Then*

$$\sup_{\|\varphi\|_\infty \leq 1} \mathbb{E}[|\mu_\infty^N(\varphi) - u_\infty(\varphi)|] \leq \frac{C}{\sqrt{N}}.$$

The remainder of the paper is as follows. Section 4.2 gives the proof of Theorem 4.1, Section 4.3 the proofs of the propagation of chaos phenomenon (Theorem 4.2) and Corollaries 4.3 and 4.4. Finally, Section 4.4 is devoted to the proof of Theorem 4.5. We conclude the paper with Section 4.5, where we state the non-explosion and the positive recurrence of the interacting particle system.

4.2 Proof of Theorem 4.1

First of all, there exists a Lyapunov function for the process $(X_t^N)_{t\geq 0}$ which ensures its non-explosion and positive recurrence (see Appendix). Combining with the irreducibility of the process $(X_t^N)_{t\geq 0}$, the Foster-Lyapunov criteria ensures its ergodicity and even its exponential ergodicity, see [85]. Before giving the demonstration of Theorem 4.1 we give an example where Assumptions (A) and (B) are satistied.

Example 4.7 (Transition rates example)**.** *Let $p < q$ two positive constants and $a \geq 1$. For $k \in \mathbb{N}$ and $l \in \mathbb{R}_+$ let*

$$b_k = pk^a, \;\; d_k = qk^a, \;\; q^+(k,l) = (l-k)_+ \textit{ and } q^-(k,l) = (k-l)_+,$$

where for $x \in \mathbb{R}$, $x_+ = \max(x, 0)$. The interaction functions q^+ and q^- mean that the more the particles are far from their mean, the more they tend to come closer to it. Then, the transition *rates satisfy the assumptions* (A) *and* (B) *with* $\lambda = q - p$ *and* $\alpha = 2$*. We have equality in assumption* (A) *for the* $M/M/\infty$ *queue* $b_k = p$ *and* $d_k = qk$ *or for the linear case* $b_k = pk$ *and* $d_k = qk$ *for* $k \in \mathbb{N}$*.*

Remark 4.8 (Caputo-Dai Pra-Posta results)**.** *If we consider Assumption* (A) *and the following assumptions* $\nabla^+ b \leq 0$ *and* $\nabla^+ d \geq 0$*, Caputo, Dai Pra and Posta [26] obtain estimates for the rate of exponential convergence to equilibrium of the birth and death process without interaction* ($q^+ = q^- \equiv 0$)*, in the relative entropy sense. Moreover, this exponential decay of relative entropy is convex in time. To do this, the authors control the second derivative of the relative entropy and show that its convexity leads to a modified logarithmic Sobolev inequality. Nevertheless, one of the key points of their approach is a condition of reversibility. In our case, the reversibility is not assumed and their results do not hold.*

Proof of Theorem 4.1. We build a coupling between two particle systems generated by (4.1), $X = (X^1, \ldots, X^N) \in \mathbb{N}^N$ and $Y = (Y^1, \ldots, Y^N) \in \mathbb{N}^N$, starting respectively from some random configurations $X_0, Y_0 \in \mathbb{N}^N$. The coupling that we introduce is the same as [39] or [52]. Let $\mathbb{L} = \mathbb{L}_1 + \mathbb{L}_2$ be the generator of the coupling defined by

$$\begin{aligned}
\mathbb{L}_1 f(X,Y) = & \sum_{i=1}^N (b_{X^i} \wedge b_{Y^i}) \left(f(X + e_i, Y + e_i) - f(X,Y) \right) \\
& + \sum_{i=1}^N (d_{X^i} \wedge d_{Y^i}) \left(f(X - e_i, Y - e_i) - f(X,Y) \right) \\
& + \sum_{i=1}^N (b_{X^i} - b_{Y^i})_+ \left(f(X + e_i, Y) - f(X,Y) \right) \\
& + \sum_{i=1}^N (b_{Y^i} - b_{X^i})_+ \left(f(X, Y + e_i) - f(X,Y) \right) \\
& + \sum_{i=1}^N (d_{Y^i} - d_{X^i})_+ \left(f(X, Y - e_i) - f(X,Y) \right) \\
& + \sum_{i=1}^N (d_{X^i} - d_{Y^i})_+ \left(f(X - e_i, Y) - f(X,Y) \right)
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{L}_2 f(X,Y) &= \sum_{i=1}^{N} \left(q^+(X^i, M^{N,1}) \wedge q^+(Y^i, M^{N,2})\right) \left(f(X+e_i, Y+e_i) - f(X,Y)\right) \\
&+ \sum_{i=1}^{N} \left(q^+(Y^i, M^{N,2}) - q^+(X^i, M^{N,1})\right)_+ \left(f(X, Y+e_i) - f(X,Y)\right) \\
&+ \sum_{i=1}^{N} \left(q^+(X^i, M^{N,1}) - q^+(Y^i, M^{N,2})\right)_+ \left(f(X+e_i, Y) - f(X,Y)\right) \\
&+ \sum_{i=1}^{N} \left(q^-(X^i, M^{N,1}) \wedge q^-(Y^i, M^{N,2})\right) \left(f(X-e_i, Y-e_i) - f(X,Y)\right) \\
&+ \sum_{i=1}^{N} \left(q^-(Y^i, M^{N,2}) - q^-(X^i, M^{N,1})\right)_+ \left(f(X, Y-e_i) - f(X,Y)\right) \\
&+ \sum_{i=1}^{N} \left(q^-(X^i, M^{N,1}) - q^-(Y^i, M^{N,2})\right)_+ \left(f(X-e_i, Y) - f(X,Y)\right),
\end{aligned}$$

where $M^{N,1}$ (resp. $M^{N,2}$) represents the mean of the particle system X (resp. Y). We can easily verify that if a measurable function f on $\mathbb{N}^N \times \mathbb{N}^N$ does not depend on its second (resp. first) variable ; that is, with a slight abuse of notation :

$$\forall X, Y \in \mathbb{N}^N,\ f(X,Y) = f(X) \text{ (resp. } f(X,Y) = f(Y)),$$

then $\mathbb{L}f(X,Y) = \mathcal{L}f(X)$ (resp. $\mathbb{L}f(X,Y) = \mathcal{L}f(Y)$), where $\mathcal{L}$ is defined in (4.1). This property ensures that the couple $(X_t, Y_t)_{t\geq 0}$ generated by $\mathbb{L}$ is a well-defined coupling of processes generated by $\mathcal{L}$. Applying the generator $\mathbb{L}$ to the distance d defined in (4.5), we obtain, on the one hand

$$\mathbb{L}_1 d(X,Y) = \sum_{i=1}^{N} K_i,$$

where for all $i \in \{1, \dots, N\}$

$$\begin{aligned}
K_i &= (b_{X^i} - b_{Y^i})_+ \left(|X^i + 1 - Y^i| - |X^i - Y^i| \right) \\
&+ (b_{Y^i} - b_{X^i})_+ \left(|X^i - Y^i - 1| - |X^i - Y^i| \right) \\
&+ (d_{Y^i} - d_{X^i})_+ \left(|X^i - Y^i + 1| - |X^i - Y^i| \right) \\
&+ (d_{X^i} - d_{Y^i})_+ \left(|X^i - 1 - Y^i| - |X^i - Y^i| \right).
\end{aligned}$$

Under Assumption (A) and using the fact that for all $x, y \geq 0$, $(x-y)_+ - (y-x)_+ = x - y$, there exists $\lambda > 0$ such that for $i \in \{1, \dots, N\}$

$$\begin{aligned}
K_i &= (b_{X^i} - b_{Y^i} + d_{Y^i} - d_{X^i}) \mathbb{1}_{X^i > Y^i} + (b_{Y^i} - b_{X^i} + d_{X^i} - d_{Y^i}) \mathbb{1}_{X^i < Y^i} \\
&\leq -\lambda \left((X^i - Y^i) \mathbb{1}_{X^i > Y^i} + (Y^i - X^i) \mathbb{1}_{X^i < Y^i} \right) \\
&= -\lambda |X^i - Y^i|.
\end{aligned}$$

Thus,

$$\mathbb{L}_1 d(X, Y) \leq -\lambda d(X, Y).$$

On the other hand,

$$\mathbb{L}_2 d(X, Y) = \sum_{i=1}^{N} H_i,$$

where for all $i \in \{1, \dots, N\}$

$$\begin{aligned}
H_i &= \left(q^+(X^i, M^{N,1}) - q^+(Y^i, M^{N,2}) \right)_+ \left(|X^i + 1 - Y^i| - |X^i - Y^i| \right) \\
&+ \left(q^+(Y^i, M^{N,2}) - q^+(X^i, M^{N,1}) \right)_+ \left(|X^i - Y^i - 1| - |X^i - Y^i| \right) \\
&+ \left(q^-(X^i, M^{N,1}) - q^-(Y^i, M^{N,2}) \right)_+ \left(|X^i - Y^i - 1| - |X^i - Y^i| \right) \\
&+ \left(q^-(Y^i, M^{N,2}) - q^-(X^i, M^{N,1}) \right)_+ \left(|X^i - Y^i + 1| - |X^i - Y^i| \right).
\end{aligned}$$

Using the fact that for all $x, y \geq 0$, $(x-y)_+ + (y-x)_+ = |x-y|$, we have,

$$\begin{aligned} H_i &= \Big((q^+ - q^-)(X^i, M^{N,1}) - (q^+ - q^-)(Y^i, M^{N,2})\Big)\, \mathbb{1}_{X^i > Y^i} \\ &+ \Big((q^+ - q^-)(Y^i, M^{N,2}) - (q^+ - q^-)(X^i, M^{N,1})\Big)\, \mathbb{1}_{X^i < Y^i} \\ &+ \Big(|q^+(X^i, M^{N,1}) - q^+(Y^i, M^{N,2})| + |q^-(X^i, M^{N,1}) - q^-(Y^i, M^{N,2})|\Big)\mathbb{1}_{X^i = Y^i}. \end{aligned}$$

Under Assumption (B), the growth of q^+ and the decrease of q^- on the second component imply, for $i \in \{1, \ldots, N\}$ such that $X^i = Y^i$

$$\begin{aligned} J_i &:= |q^+(X^i, M^{N,1}) - q^+(X^i, M^{N,2})| + |q^-(X^i, M^{N,1}) - q^-(X^i, M^{N,2})| \\ &= \Big((q^+ - q^-)(X^i, M^{N,2}) - (q^+ - q^-)(X^i, M^{N,1})\Big)\, \mathbb{1}_{X^i = Y^i} \mathbb{1}_{M^{N,1} < M^{N,2}} \\ &+ \Big((q^+ - q^-)(X^i, M^{N,1}) - (q^+ - q^-)(X^i, M^{N,2})\Big)\, \mathbb{1}_{X^i = Y^i} \mathbb{1}_{M^{N,1} > M^{N,2}}. \end{aligned}$$

We deduce that, under Assumption (B), there exists $\alpha > 0$ such that for $i \in \{1, \ldots, N\}$

$$H_i \leq \alpha \left(|X^i - Y^i| + |M^{N,1} - M^{N,2}|\right),$$

and thus the definition of $M^{N,l}, l = 1, 2$ implies that

$$\mathbb{L}_2 d(X, Y) \leq 2\alpha d(X, Y).$$

We deduce that $\mathbb{L}d(X,Y) \leq -(\lambda - 2\alpha) d(X,Y)$. 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's Lemma, we have, for every $t \geq 0$, $\mathbb{P}_t d \leq e^{-(\lambda - 2\alpha)t} d$; namely

$$\mathbb{E}[d(X_t, Y_t)] \leq e^{-(\lambda - 2\alpha)t} \mathbb{E}[d(X_0, Y_0)].$$

Taking the infimum over all couples (X_0, Y_0), the claim follows. □

A new condition on the interaction rates. If we replace the Lipschitz condition in Assumption (B) by the following one : there exist $\alpha, \zeta > 0$ such that for any $(k_1, l_1), (k_2, l_2) \in \mathbb{N} \times \mathbb{R}_+$, $k_1 \geq k_2$

$$(q^+ - q^-)(k_1, l_1) - (q^+ - q^-)(k_2, l_2) \leq -\alpha\,(k_1 - k_2) + \zeta\,(l_1 - l_2)\,. \tag{4.13}$$

Then, under this condition and Assumption (A), we have for any processes $(X_t^N)_{t\geq 0}$ and $(Y_t^N)_{t\geq 0}$ generated by (4.1) and for any $t \geq 0$,

$$\mathcal{W}_d(\mathrm{Law}(X_t^N), \mathrm{Law}(Y_t^N)) \leq e^{-((\lambda+\alpha)-\zeta)t}\mathcal{W}_d(\mathrm{Law}(X_0^N), \mathrm{Law}(Y_0^N)).$$

If $l_1 = l_2$, the condition (4.13) is a convexity condition on the first variable. For $l_1 \neq l_2$, the term $\zeta(l_1 - l_2)$ represents the fluctuations of the barycenters. The resulting rate is slightly better than the one before, but we can find an example of interaction rates for which we obtain an optimal rate of convergence in Theorem 4.1. This example is inspired by Malrieu's model when the interaction potential is $W(x, y) = a(x - y)^2$, $a > 0$. We assume that for a particle system $X = (X^1, \dots, X^N)$, the interaction birth and death rates are given respectively by

$$q_X^+(X^i) = a\frac{1}{N}\sum_{j=1}^{N}\left(X^i - X^j\right)_- \text{ and } q_X^-(X^i) = a\frac{1}{N}\sum_{j=1}^{N}\left(X^i - X^j\right)_+,$$

where $a > 0$, and the generator is given by

$$\mathcal{L}f(x) = \sum_{i=1}^{N}\left[\left(b_{x_i} + q_{\cdot}^+(x_i)\right)(f(x+e_i) - f(x)) + \left(d_{x_i} + q_{\cdot}^-(x_i)\right)(f(x-e_i) - f(x))\,\mathbb{1}_{x_i>0}\right]. \tag{4.14}$$

Theorem 4.9. *Assume that Assumption (A) is satisfied. Then, for the processes $(X_t^N)_{t\geq 0}$ and $(Y_t^N)_{t\geq 0}$ generated by (4.14), we have for all $t \geq 0$*

$$\mathcal{W}_d(\mathrm{Law}(X_t^N), \mathrm{Law}(Y_t^N)) \leq e^{-\lambda t}\mathcal{W}_d(\mathrm{Law}(X_0^N), \mathrm{Law}(Y_0^N)).$$

Démonstration. Using the same coupling and the same notations as in the proof of Theorem 4.1, we have

$$\mathbb{L}_2 d(X, Y) = \sum_{i=1}^{N} H_i,$$

where for all $i \in \{1, \dots, N\}$

$$\begin{aligned}
H_i &= \left((q_X^+ - q_X^-)(X^i) - (q_Y^+ - q_Y^-)(Y^i) \right) \mathbb{1}_{X^i > Y^i} \\
&+ \left((q_Y^+ - q_Y^-)(Y^i) - (q_X^+ - q_X^-)(X^i, M^{N,1}) \right) \mathbb{1}_{X^i < Y^i} \\
&+ \left(|q_X^+(X^i) - q_Y^+(X^i)| + |q_X^-(X^i) - q_Y^-(X^i)| \right) \mathbb{1}_{X^i = Y^i}. \\
&\leq a \left[-\left(X^i - Y^i \right) + \frac{1}{N} \sum_{i=1}^N \left(X^j - Y^j \right) \right] \mathbb{1}_{X^i > Y^i} \\
&+ a \left[-\left(Y^i - X^i \right) + \frac{1}{N} \sum_{i=1}^N \left(Y^j - X^j \right) \right] \mathbb{1}_{Y^i > X^i} \\
&+ a \frac{1}{N} \sum_{i=1}^N |X^j - Y^j| \mathbb{1}_{X^i = Y^i}.
\end{aligned}$$

Thus,

$$H_i \leq -a|X^i - Y^i| + a\frac{1}{N} \sum_{i=1}^N |X^j - Y^j|,$$

and

$$\sum_{i=1}^N H_i \leq 0.$$

We deduce that $\mathbb{L}d(X, Y) \leq -\lambda d(X, Y)$. □

4.3 Proof of Theorem 4.2

Let us give an important consequence of Theorem 4.2 : with explicit rate, we have the propagation of chaos for the system of interacting particles.

Corollary 4.10 (Strong Propagation of chaos). *Let $\mu_t^{(k,N)}$ be the law of k particles among N at time t and u_t be the law of the nonlinear process. Then,*

$$\sup_{t \geq 0} \mathcal{W}_d(\mu_t^{(k,N)}, u_t^{\otimes k}) \leq \frac{kK}{\sqrt{N}}.$$

Proof. By exchangeability, the k-marginals of $\mathrm{Law}(X_t^N)$ do not depend on the choice of coordinates. Thus,

$$\mathcal{W}_d(\mu_t^{(k,N)}, u_t^{\otimes k}) \leq \sum_{i=1}^k \mathbb{E}[|X_t^{i,N} - \overline{X}_t^i|].$$

□

Proof of Theorem 4.2. To prove the propagation of chaos phenomenon, we construct a coupling between the particle system $(X^{1,N}, \dots, X^{N,N})$ and N independent nonlinear processes $(\overline{X}^i, \dots, \overline{X}^N)$. For $i \in \{1, \dots, N\}$ and $t \geq 0$

- $\overline{X}^i_0 = X^{i,N}_0$
- $\mathrm{Law}(\overline{X}^i_t) = u_t$
- the transition rates of $\overline{X}^i$ at time t are given by

$$\begin{array}{ccccc} i & \to & i+1 & \text{with rate } b_i + q^+(i, \|u_t\|), & \\ i & \to & i-1 & \text{with rate } d_i + q^-(i, \|u_t\|), & \text{for } i \geq 1 \\ i & \to & j & \text{with rate } 0, & \text{if } j \notin \{i-1, i+1\}. \end{array}$$

Using the coupling and the notations introduced in the proof of Theorem 4.1, we have under Assumption (A)

$$\mathbb{L}_1 d(X, \overline{X}) \leq -\lambda d(X, \overline{X}),$$

and under Assumption (B)

$$\mathbb{L}_2 d(X, \overline{X}) \leq \alpha d(X, \overline{X}) + \alpha N |M^N - \|u\||,$$

where we recall that the distance d is the l^1-distance. But, for $t \geq 0$

$$\mathbb{E}|M^N_t - \|u_t\|| \leq \mathbb{E}\Big|M^N_t - \frac{1}{N}\sum_{i=1}^N \overline{X}^i_t\Big| + \mathbb{E}\Big|\frac{1}{N}\sum_{i=1}^N \overline{X}^i_t - \|u_t\|\Big|.$$

On the one hand

$$\mathbb{E}\Big|M^N_t - \frac{1}{N}\sum_{i=1}^N \overline{X}^i_t\Big| \leq \frac{1}{N}\mathbb{E}\left(\sum_{i=1}^N |X^{i,N}_t - \overline{X}^i_t|\right) = \frac{1}{N}\mathbb{E}(d(X_t, \overline{X}_t)).$$

On the other hand, by independence and Cauchy Schwarz inequality

$$\begin{aligned}
\mathbb{E}\Big|\frac{1}{N}\sum_{i=1}^{N}\overline{X}_t^i - \|u_t\|\Big| &= \mathbb{E}\Big|\frac{1}{N}\sum_{i=1}^{N}\overline{X}_t^i - \mathbb{E}(\overline{X}_t^1)\Big| \\
&= \frac{1}{N}\mathbb{E}\Big|\sum_{i=1}^{N}\overline{X}_t^i - \mathbb{E}\left(\sum_{i=1}^{N}\overline{X}_t^i\right)\Big| \\
&\leq \frac{1}{N}\left(\mathrm{Var}\left(\sum_{i=1}^{N}\overline{X}_t^i\right)\right)^{\frac{1}{2}} \\
&\leq \frac{1}{N}\left(\sum_{i=1}^{N}\mathrm{Var}(\overline{X}_t^i)\right)^{\frac{1}{2}} \\
&\leq \frac{1}{\sqrt{N}}\mathrm{Var}(\overline{X}_t^1)^{\frac{1}{2}}.
\end{aligned}$$

Now, the process $(\overline{X}_t^1)_{t\geq 0}$ has finite exponential moments, uniform in time, as soon as it is finite at time 0.

Lemma 4.11 (Exponential moment of $(\overline{X}_t^1)_{t\geq 0}$)**.** *Let $\delta > 0$ and $\beta(\delta) := \inf_{x\in\mathbb{N}^*}\left(d_x e^{-\delta} - b_x\right)$ and let $K_1 = \alpha\left(\|u_0\| + \dfrac{b_0}{\lambda - 2\alpha}\right)$. Then, under Assumptions (A) and (B) and if $\beta(\delta) - K_1 > 0$, $\sum_{i\geq 0} e^{\delta i}u_t(i)$ is finite for every $t \geq 0$ as soon as $\sum_{i\geq 0} e^{\delta i}u_0(i)$ is finite and*

$$\sum_{i\geq 0} e^{\delta i}u_t(i) \leq \sum_{i\geq 0} e^{\delta i}u_0(i) + \frac{b_0}{\beta(\delta) - K_1}.$$

Proof of Lemma 4.11. Let us first remark that, if $\lambda - 2\alpha > 0$

$$\|u_t\| \leq \|u_0\| + \frac{b_0}{\lambda - 2\alpha}.$$

Indeed, applying the operator $\mathcal{G}_{(\cdot)}$ defined in (4.3) to the function $f(i) = i$ we have

$$\begin{aligned}
\mathcal{G}_{u_t}f(i) &= b_i - d_i + \left(q^+ - q^-\right)(i, \|u_t\|) \\
&\leq -\lambda i + b_0 + \alpha\left(i + \|u_t\|\right).
\end{aligned}$$

By equation (4.2) we obtain

$$\frac{d}{dt}\|u_t\| \leq -\left(\lambda - 2\alpha\right)\|u_t\| + b_0$$

and Gronwall's lemma gives the result.
Now, let us take $f(i) = e^{\delta i}$, $\delta > 0$. Then, under Assumption (B)

$$\begin{aligned}
\mathcal{G}_{u_t} f(i) &\leq -(e^{\delta} - 1)e^{\delta i}\left[e^{-\delta} d_i - b_i\right] \mathbb{1}_{i>0} + (e^{\delta} - 1)e^{\delta i} q^{+}(i, \|u_t\|) + b_0(e^{\delta} - 1) \\
&\leq -\beta(\delta)(e^{\delta} - 1) f(i) + (e^{\delta} - 1) f(i) q^{+}(0, \|u_t\|) + b_0(e^{\delta} - 1) \\
&\leq -\beta(\delta)(e^{\delta} - 1) f(i) + \alpha(e^{\delta} - 1) f(i)\|u_t\| + b_0(e^{\delta} - 1) \\
&\leq -\left(\beta(\delta) - K_1\right)(e^{\delta} - 1) f(i) + b_0(e^{\delta} - 1).
\end{aligned}$$

Thus,
$$\frac{d}{dt}\sum_{i\geq 0} e^{\delta i} u_t(i) \leq -\left(\beta(\delta) - K_1\right)(e^{\delta} - 1)\sum_{i\geq 0} e^{\delta i} u_t(i) + b_0(e^{\delta} - 1).$$

□

We are now able to conclude the proof. For $t \geq 0$ and $i \in \{1, \ldots, N\}$, let

$$\gamma(t) = \mathbb{E}|X_t^{i,N} - \overline{X}_t^i|.$$

Then, by the equality $\partial_t \mathbb{P}_t f = \mathbb{P}_t \mathbb{L} f$ (where we recall that $\mathbb{P}$ is the semi-group associated with $\mathbb{L}$), Lemma 4.11 and the exchangeability of the marginals of the particle system, there exists $K > 0$ such that
$$\partial_t \gamma(t) \leq -(\lambda - 2\alpha)\gamma(t) + \frac{K}{\sqrt{N}}.$$
Gronwall's lemma gives for every $t \geq 0$

$$\gamma(t) \leq e^{-(\lambda-2\alpha)t}\gamma(0) + \frac{K}{\sqrt{N}}\left(1 - e^{-(\lambda-2\alpha)t}\right).$$

As the initial conditions are the same, we obtain (4.8). □

Proof of Corollary 4.3. Let φ be a function such that $\|\varphi\|_{Lip} \leq 1$. Then, using the same coupling as Theorem 4.2, we have

$$\begin{aligned}
\mathbb{E}|\mu_t^N(\varphi) - u_t(\varphi)| &= \mathbb{E}\Big|\frac{1}{N}\sum_{i=1}^N \varphi(X_t^{i,N}) - \mathbb{E}[\varphi(\overline{X}_t^1)]\Big| \\
&\leq \mathbb{E}\Big|\frac{1}{N}\sum_{i=1}^N \varphi(X_t^{i,N}) - \frac{1}{N}\sum_{i=1}^N \varphi(\overline{X}_t^i)\Big| + \mathbb{E}\Big|\frac{1}{N}\sum_{i=1}^N \varphi(\overline{X}_t^i) - \mathbb{E}[\varphi(\overline{X}_t^1)]\Big|.
\end{aligned}$$

By Theorem 4.2, there exists $K > 0$ such that

$$\mathbb{E}\Big|\frac{1}{N}\sum_{i=1}^{N}\varphi(X_t^{i,N}) - \frac{1}{N}\sum_{i=1}^{N}\varphi(\overline{X}_t^i)\Big| \leq \frac{K}{\sqrt{N}}.$$

Now, the Cauchy-Schwarz inequality and Lemma 4.11 imply the existence of a constant $C > 0$ such that

$$\mathbb{E}\Big|\frac{1}{N}\sum_{i=1}^{N}\varphi(\overline{X}_t^i) - \mathbb{E}[\varphi(\overline{X}_t^1)]\Big| \leq \frac{C}{\sqrt{N}}.$$

□

Proof of Corollary 4.4. Let ν^N be the empirical measure of the independent processes $(\overline{X}^i)_{i\in\{1,\ldots,N\}}$. Namely, for any $t \geq 0$

$$\nu_t^N = \frac{1}{N}\sum_{i=1}^{N}\delta_{\overline{X}_t^i}.$$

Then, the triangular inequality gives

$$\mathcal{W}_1(\mu_t^N, u_t) \leq \mathcal{W}_1(\mu_t^N, \nu_t^N) + \mathcal{W}_1(\nu_t^N, u_t).$$

As for any $t \geq 0$,

$$\mathcal{W}_1(\mu_t^N, \nu_t^N) \leq \frac{1}{N}\sum_{i=1}^{N}|X_t^{i,N} - \overline{X}_t^i|,$$

then, by Theorem 4.2, there exists a constant $K > 0$ such that

$$\begin{aligned}\mathbb{E}[\mathcal{W}_1(\mu_t^N, \nu_t^N)] &\leq \frac{1}{N}\sum_{i=1}^{N}\sup_{t\geq 0}\mathbb{E}|X_t^{i,N} - \overline{X}_t^i| \\ &\leq \frac{K}{\sqrt{N}}.\end{aligned}$$

On the other hand, by Lemma 4.11, the process $\overline{X}^i$, for every $i \in \{1,\ldots,N\}$, has finite exponential moments. So, applying Theorem 1 of [58] with $p = d = 1$ and $q > 2$, there exists a constant $\overline{K}$ such that

$$\mathbb{E}[\mathcal{W}_1(\nu_t^N, u_t)] \leq \frac{\overline{K}}{\sqrt{N}}.$$

We deduce that for every $t \geq 0$,

$$\mathbb{E}[\mathcal{W}_1(\mu_t^N, u_t)] \leq \frac{K + \overline{K}}{\sqrt{N}},$$

which by Markov's inequality ends the proof. $\square$

4.4 Proof of Theorem 4.5

Proof of Theorem 4.5. Let $(X_t^N)_{t\geq 0}$ and $(Y_t^N)_{t\geq 0}$ be two particle systems generated by (4.1) with initial laws $u_0^{\otimes N}$ and $v_0^{\otimes N}$ respectively. Let $u_t^{1,N}$ (resp. $v_t^{1,N}$) be the first marginal of $\mathrm{Law}(X_t^N)$ (resp. $\mathrm{Law}(Y_t^N)$). Then, the triangular inequality yields

$$\mathcal{W}_1(u_t, v_t) \leq \mathcal{W}_1(u_t, u_t^{1,N}) + \mathcal{W}_1(u_t^{1,N}, v_t^{1,N}) + \mathcal{W}_1(v_t^{1,N}, v_t).$$

The uniform propagation of chaos (Theorem 4.2) gives for every $t \geq 0$

$$\mathcal{W}_1(u_t, u_t^{1,N}) \leq \mathbb{E}|\overline{X}_t^1 - X_t^{1,N}| \leq \frac{K}{\sqrt{N}},$$

and

$$\mathcal{W}_1(v_t, v_t^{1,N}) \leq \frac{K}{\sqrt{N}}.$$

Now, by exchangeability of the marginals of the particles and by Theorem 4.1 we have

$$\begin{aligned}
\mathcal{W}_1(u_t^{1,N}, v_t^{1,N}) &\leq \frac{1}{N}\mathcal{W}_d(\mathrm{Law}(X_t^N), \mathrm{Law}(Y_t^N)) \\
&\leq \frac{1}{N}e^{-(\lambda-2\alpha)t}\mathcal{W}_d(\mathrm{Law}(X_0^N), \mathrm{Law}(Y_0^N)) \\
&\leq \frac{1}{N}e^{-(\lambda-2\alpha)t}N\mathcal{W}_1(u_0, v_0) \\
&\leq e^{-(\lambda-2\alpha)t}\mathcal{W}_1(u_0, v_0).
\end{aligned}$$

We deduce that

$$\mathcal{W}_1(u_t, v_t) \leq \frac{2K}{\sqrt{N}} + e^{-(\lambda-2\alpha)t}\mathcal{W}_1(u_0, v_0).$$

Taking the limit as N tends to infinity we obtain (4.11). Then, the sequence $(u_t)_{t\geq 0}$ is a Cauchy sequence for the $\mathcal{W}_1$-Wasserstein distance and thus admits a limit u_∞. $\square$

4.5 Appendix

In the following theorem, we prove the existence of a Lyapunov function. This function ensures the non-explosion and the positive recurrence of the particle system. Along this section, we note $\kappa = \lambda - 2\alpha$.

Theorem 4.12 (Lyapunov function)**.** *Let V be the function $x \in \mathbb{N}^N \mapsto \sum_{k=1}^{N} x_k$ and let us assume that Assumptions (A) and (B) are satisfied. Then, for $x \in \mathbb{N}^N$*

$$\mathcal{L}V(x) \leq -\kappa V(x) + b_0 N. \tag{4.15}$$

Thus, the function V is a Lyapunov function for $\mathcal{L}$. In particular, the process $(X_t^N)_{t\geq 0}$ is non-explosive and positive recurrent.

Under this assumption and without rate of convergence, the existence of a Lyapunov function combined with the irreducibility of the process (which implies that every compact is a small set) provides a sufficient criterion ensuring that the interacting particle system is ergodic [85, Theorem 6.1]. And the inequality below implies that under the invariant distribution, the mean of the particle system M^N is upper bounded.

Proof. **Lyapunov function.**
Let V be the function $x \mapsto \sum_{k=1}^{N} x_k$. Then,

$$\begin{aligned}\mathcal{L}V(x) &= \sum_{i=1}^{N} \left[b_{x_i} + q^+(x_i, M^N) - \left(d_{x_i} + q^-(x_i, M^N)\right) \mathbb{1}_{x_i>0}\right] \\ &= \sum_{i=1}^{N} \left[(b_{x_i} - d_{x_i}) + (q^+ - q^-)(x_i, M^N)\right] \mathbb{1}_{x_i>0} + \left(b_0 + q^+(0, M^N)\right) \sum_{i=1}^{N} \mathbb{1}_{x_i=0}.\end{aligned}$$

By Assumption (A)

$$\begin{aligned} d_{x_i} - b_{x_i} &= \sum_{n=0}^{x_i-1} (d_{n+1} - d_n - b_{n+1} + b_n) - b_0 \\ &\geq \lambda x_i - b_0, \end{aligned}$$

and Assumption (B) gives

$$(q^+ - q^-)(x_i, M^N) \leq \alpha \left(x_i + M^N\right) \text{ and } q^+(0, M^N) \leq \alpha M^N.$$

Thus,

$$\begin{aligned}\mathcal{L}V(x) &\leq -\lambda\sum_{i=1}^{N} x_i + b_0\sum_{i=1}^{N}\mathbb{1}_{x_i\geq 0} + \alpha\sum_{i=1}^{N} x_i + \alpha M^N\sum_{i=1}^{N}\mathbb{1}_{x_i\geq 0}\\ &\leq -\kappa V(x) + b_0 N.\end{aligned}$$

Non-explosion.

Let V be the Lyapunov function defined in Theorem 4.12, and for $x \in \mathbb{N}^N$ let W be the function defined by

$$W(x) = \frac{1}{N}\left(V(x) - \frac{b_0 N}{\kappa}\right).$$

Then W satisfies a simpler inequality than (4.15) given by

$$\mathcal{L}W(x) \leq -\kappa W(x).$$

Consider now the function $f : \mathbb{N}^N \times \mathbb{R}_+ \to \mathbb{R}$ defined by $f(x,t) = e^{\kappa t}W(x)$ and for A, let τ_A be the stopping time defined by $\tau_A = \inf\{t \geq 0, W(X_t) > A\}$. Then, as

$$M_t = f(X_t,t) - f(X_0,0) - \int_0^t \left(\frac{\partial}{\partial_t}f + \mathcal{L}f\right)(X_s,s)ds \quad \text{is a martingale,}$$

we have for $x \in \mathbb{N}^N$

$$\mathbb{E}_x[f(X_{t\wedge\tau_A}, t\wedge\tau_A)] = W(x) + \mathbb{E}_x\left[\int_0^{t\wedge\tau_A}\left(\frac{\partial}{\partial_t}f + \mathcal{L}f\right)(X_s,s)ds\right].$$

But,

$$\frac{\partial}{\partial_t}f + \mathcal{L}f \leq 0,$$

thus,

$$\mathbb{E}_x[e^{\kappa(t\wedge\tau_A)}W(X_{t\wedge\tau_A})] \leq W(x).$$

And finally by the definition of τ_A and W

$$\begin{aligned}\mathbb{P}_x(\tau_A < t) &\leq \frac{1}{A}\mathbb{E}_x\left[e^{\kappa\tau_A}W(X_{\tau_A})\mathbb{1}_{\tau_A\leq t}\right]\\ &\leq \frac{1}{A}\left(W(x) - e^{\kappa t}\mathbb{E}_x[W(X_t)\mathbb{1}_{\tau_A\geq t}]\right)\\ &\leq \frac{1}{A}\left(W(x) + b_0 e^{\kappa t}\right).\end{aligned}$$

Positive recurrence.

Let $A = \{x \in \mathbb{N}^N : \kappa V(x) < 2b_0 N\}$, H_A the hitting time of A and $H_A^n = H_A \wedge n$. As $\lim_{\|x\|\to+\infty} V(x) = +\infty$, then A is a finite set. Applying the same argument for $f(x,t) = e^{\frac{\kappa}{2}t}V(x)$ gives

$$\mathbb{E}_x[e^{\frac{\kappa}{2}H_A^n}V(X_{H_A^n})] \leq V(x) \quad \forall x \notin A.$$

By definition of H_A, we deduce that

$$\mathbb{E}_x[e^{\frac{\kappa}{2}H_A^n}] \leq \frac{V(x)\kappa}{2b_0 N} \quad \forall x \notin A,$$

and the monotonicity convergence theorem gives

$$\mathbb{E}_x[e^{\frac{\kappa}{2}H_A}] \leq \frac{V(x)\kappa}{2b_0 N} \quad \forall x \notin A. \tag{4.16}$$

Now, let $\sigma_A = \inf\{t > H_{A^c}, X_t \in A\}$ and $x \in A$. Using (4.16), we have

$$\mathbb{E}_x[e^{\frac{\kappa}{2}\sigma_A}] = \mathbb{E}_x[\mathbb{E}_{X_{H_{A^c}}}[e^{\frac{\kappa}{2}\sigma_A}]] \leq \frac{\kappa\mathbb{E}_x[V(X_{H_{A^c}})]}{2b_0 N}$$

Using the fact that $\mathcal{L}V(x) < +\infty$ for $x \in A$ ends the proof. □

Bibliographie

[1] L. Addario-Berry and B. Reed. Minima in branching random walks. *Ann. Probab.*, 37(3) :1044–1079, 05 2009. 34

[2] E. Aïdékon. Convergence in law of the minimum of a branching random walk. *Ann. Probab.*, 41(3A) :1362–1426, 05 2013.

[3] E. Aïdékon, J. Berestycki, É. Brunet, and Z. Shi. Branching Brownian motion seen from its tip. *Probability Theory and Related Fields*, 157(1-2) :405–451, 2013. 34

[4] D. Aldous, B. Flannery, and J. L. Palacios. Two applications of urn processes the fringe analysis of search trees and the simulation of quasi-stationary distributions of markov chains. *Probability in the Engineering and Informational Sciences*, 2 :293–307, 7 1988. 7

[5] C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto, and G. Scheffer. *Sur les inégalités de Sobolev logarithmiques*, volume 10 of *Panoramas et Synthèses [Panoramas and Syntheses]*. Société Mathématique de France, Paris, 2000. With a preface by Dominique Bakry and Michel Ledoux. 27, 28, 32, 87

[6] L.-P. Arguin, A. Bovier, and N. Kistler. Poissonian statistics in the extremal process of branching brownian motion. *Ann. Appl. Probab.*, 22(4) :1693–1711, 08 2012. 34

[7] L.-P. Arguin, A. Bovier, and N. Kistler. The extremal process of branching brownian motion. *Probability Theory and Related Fields*, 157(3-4) :535–574, 2013. 34

[8] 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. 11, 23, 53, 55, 96

[9] A. Asselah, P. A. Ferrari, P. Groisman, and M. Jonckheere. Fleming-Viot selects the minimal quasi-stationary distribution : The Galton-Watson case. *ArXiv e-prints*, June 2012. 11, 12, 18, 19, 30, 36, 38, 41, 44

[10] A. Asselah and M.-N. Thai. A note on the rightmost particle in a Fleming-Viot process. *ArXiv e-prints*, Dec. 2012. 2, 11, 20, 30, 33, 96

[11] K. B. Athreya and S. Karlin. Branching processes with Random Environments, ii : Limit Theorems. *Ann. Math. Statist.*, 42(6) :1843–1858, 12 1971. 2

[12] K. B. Athreya and P. E. Ney. *Branching processes*. Springer-Verlag, New York, 1972. Die Grundlehren der mathematischen Wissenschaften, Band 196. 4

[13] 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. 27, 32, 80

[14] 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. 62

[15] V. Bansaye. Surviving particles for subcritical branching processes in random environment. *Stochastic Process. Appl.*, 119(8) :2436–2464, 2009. 2

[16] M. Benaïm and B. Cloez. A stochastic approximation approach to quasi-stationary distributions on finite spaces. *Electron. Commun. Probab.*, 20 :no. 37, 1–13, 2015. 7

[17] M. Bieniek, K. Burdzy, and S. Finch. Non-extinction of a Fleming-Viot particle model. *Probability Theory and Related Fields*, 153(1-2) :293–332, 2012. 10

[18] 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. 56

[19] S. Bobkov and F. Götze. Exponential integrability and transportation cost related to logarithmic sobolev inequalities. *Journal of Functional Analysis*, 163(1) :1 – 28, 1999.

[20] S. Bobkov and M. Ledoux. *One-dimensional empirical measures, order statistics and Kantorovich transport distances.* 2014. Book in preparation. Available at http ://perso.math.univ-toulouse.fr/ledoux/files/2013/11/Order.statistics.10.pdf.

[21] F. Bolley, A. Guillin, and C. Villani. Quantitative concentration inequalities for empirical measures on non-compact spaces. *Probability Theory and Related Fields*, 137(3-4) :541–593, 2007. 17, 18, 99

[22] C. Bordenave, D. Mcdonald, and R. Proutiere. A particle system in interaction with a rapidly varying environment : Mean field limits and applications. *Networks and Heterogeneous Media (NHM)*, 2010. 13

[23] E. Brunet and B. Derrida. Effect of microscopic noise on front propagation. *Journal of Statistical Physics*, 103(1-2) :269–282, 2001. 34

[24] K. Burdzy, R. Holyst, D. Ingerman, and P. March. Configurational transition in a Fleming - Viot-type model and probabilistic interpretation of Laplacian eigenfunctions. *Journal of Physics A : Mathematical and General*, 29(11) :2633, 1996. 13, 33, 96

[25] K. Burdzy, R. Hołyst, and P. March. A Fleming-Viot particle representation of the Dirichlet Laplacian. *Comm. Math. Phys.*, 214(3) :679–703, 2000. 8, 10, 47, 48

[26] P. Caputo, P. Dai Pra, and G. Posta. Convex entropy decay via the Bochner–Bakry–Emery approach. *Ann. Inst. H. Poincaré Probab. Statist.*, 45(3) :734–753, 08 2009. 102

[27] P. Cattiaux, A. Guillin, and F. Malrieu. Probabilistic approach for granular media equations in the non uniformly convex case. *Probability Theory and Related Fields*, 140(1-2) :19–40, 2008. 96, 99, 101

[28] J. A. Cavender. Quasi-stationary distributions of birth-and-death processes. *Advances in Applied Probability*, pages 570–586, 1978. 6, 20, 42

[29] D. Chafaï and A. Joulin. Intertwining and commutation relations for birth–death processes. *Bernoulli*, 19(5A) :1855–1879, 2013. 62, 84

[30] N. Champagnat and D. Villemonais. Exponential convergence to quasi-stationary distribution and Q-process. *Probability Theory and Related Fields*, pages 1–41, 2015. 8

[31] M.-F. Chen. *From Markov chains to non-equilibrium particle systems.* World Scientific Publishing Co. Inc., River Edge, NJ, second edition, 2004. 32, 61, 84, 86

[32] B. Cloez. *Comportement asymptotique de processus avec sauts et applications pour des modèles avec branchement.* PhD thesis, Université Paris-Est, 2013.

[33] B. Cloez and M. Hairer. Exponential ergodicity for markov processes with random switching. *Bernoulli*, 21(1) :505–536, 02 2015. 62

[34] B. Cloez and M.-N. Thai. Quantitative results for the Fleming-Viot particle system and quasi-stationary distributions in discrete space. *Stochastic Processes and their Applications*, 2015. 2, 12, 47, 96

[35] P. Collet, S. Martínez, and J. San Martín. *Quasi-stationary distributions.* Probability and its Applications (New York). Springer, Heidelberg, 2013. Markov chains, diffusions and dynamical systems. 47, 54

[36] J. N. Darroch and E. Seneta. On quasi-stationary distributions in absorbing discrete-time finite Markov chains. *J. Appl. Probability*, 2 :88–100, 1965. 4

[37] J. N. Darroch and E. Seneta. On quasi-stationary distributions in absorbing continuous-time finite Markov chains. *J. Appl. Probability*, 4 :192–196, 1967. 4, 5, 54, 72

[38] D. Dawson and J. Gärtner. Long time behaviour of interacting diffusions. Stochastic calculus in application, Proc. Symp., Cambridge/UK 1987, Pitman Res. Notes Math. Ser. 197, 29-54 (1988)., 1988. 14, 98

[39] D. Dawson, J. Tang, and Y. Zhao. Balancing Queues by Mean Field Interaction. *Queueing Systems*, 49(3-4) :335–361, 2005. 11, 13, 14, 95, 98, 102

[40] D. Dawson and X. Zheng. Law of large numbers and central limit theorem for unbounded jump mean-field models. *Advances in Applied Mathematics*, 12(3) :293 – 326, 1991. 98

[41] 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. 8, 13, 48

[42] P. Del Moral and L. Miclo. A Moran particle system approximation of Feynman-Kac formulae. *Stochastic Process. Appl.*, 86(2) :193–216, 2000. 8, 13, 47, 48, 53

[43] 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. 47, 54, 72

[44] A. Dembo and O. Zeitouni. *Large deviations techniques and applications.* Applications of mathematics. Springer, New York, Berlin, Heidelberg, 1998.

[45] P. Diaconis and L. Miclo. On quantitative convergence to quasi-stationarity. *ArXiv e-prints*, June 2014. 5, 6

[46] P. Diaconis and L. Miclo. Estimates on the amplitude of the first Dirichlet eigenvector in discrete frameworks. *ArXiv e-prints*, May 2015. 6

[47] E. A. V. Doorn. Quasi-stationary distributions and convergence to quasi-stationarity of birth-death processes. *Advances in Applied Probability*, 23(4) :pp. 683–700, 1991. 6, 33

[48] D. Down, S. P. Meyn, and R. L. Tweedie. Exponential and Uniform Ergodicity of Markov Processes. *Ann. Probab.*, 23(4) :1671–1691, 10 1995. 29

[49] A. Eberle. Couplings, distances and contractivity for diffusion processes revisited. *ArXiv e-prints*, May 2013. 62

[50] A. Etheridge. *Some mathematical models from population genetics*, volume 2012 of *Lecture Notes in Mathematics*. Springer, Heidelberg, 2011. Lectures from the 39th Probability Summer School held in Saint-Flour, 2009. 73

[51] S. Feng. Large deviations for empirical process of mean-field interacting particle system with unbounded jumps. *Ann. Probab.*, 22(4) :2122–2151, 10 1994. 98

[52] S. Feng and X. Zheng. Solutions of a class of nonlinear master equations. *Stochastic Processes and their Applications*, 43(1) :65 – 84, 1992. 95, 102

[53] P. Ferrari, S. Martinez, and P. Picco. Some properties of quasi stationary distributions in the birth and death chains : A dynamical approach. In E. Tirapegui and W. Zeller, editors, *Instabilities and Nonequilibrium Structures III*, volume 64 of *Mathematics and Its Applications*, pages 177–187. Springer Netherlands, 1991. 6

[54] P. A. Ferrari, H. Kesten, S. Martinez, and P. Picco. Existence of quasi-stationary distributions. A renewal dynamical approach. *Ann. Probab.*, 23(2) :501–521, 1995. 7, 33, 54

[55] 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. 11, 12, 23, 47, 50, 53, 54, 96

[56] P. A. Ferrari, S. Martínez, and P. Picco. Existence of non trivial quasi stationary distributions in the birth and death chain. *Advances Applied Probability*, 24 :79581–3, 1992. 6

[57] W. H. Fleming and M. Viot. Some measure-valued Markov processes in population genetics theory. 1979. 8

[58] N. Fournier and A. Guillin. On the rate of convergence in wasserstein distance of the empirical measure. *Probability Theory and Related Fields*, 162(3-4) :707–738, 2015. 111

[59] T. Funaki. A certain class of diffusion processes associated with nonlinear parabolic equations. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 67(3) :331–348, 1984. 14

[60] C. Giardinà, F. Redig, and K. Vafayi. Correlation inequalities for interacting particle systems with duality. *J. Stat. Phys.*, 141(2) :242–263, 2010. 73, 75

[61] I. Grigorescu and M. Kang. Hydrodynamic limit for a Fleming-Viot type system. *Stochastic Process. Appl.*, 110(1) :111–143, 2004. 10

[62] P. Groisman and M. Jonckheere. Front propagation and quasi-stationary distributions : the same selection principle ? *ArXiv e-prints*, Apr. 2013.

[63] P. Groisman and M. Jonckheere. Simulation of quasi-stationary distributions on countable spaces. *Markov Process. Related Fields*, 19(3) :521–542, 2013. 11, 20, 47, 53

[64] S. Grosskinsky, F. Redig, and K. Vafayi. Condensation in the inclusion process and related models. *J. Stat. Phys.*, 142(5) :952–974, 2011. 73, 75

[65] S. Grosskinsky, F. Redig, and K. Vafayi. Dynamics of condensation in the symmetric inclusion process. *Electron. J. Probab.*, 18 :no. 66, 23, 2013. 73

[66] M. Hairer. Convergence of Markov Processes. http ://www.hairer.org/notes/Convergence.pdf.

[67] 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. 51, 62

[68] C. Heathcote. A branching process allowing immigration. *J. R. Stat. Soc., Ser. B*, 27 :138–143, 1965. 2

[69] Y. Hu and Z. Shi. Minimal position and critical martingale convergence in branching random walks, and directed polymers on disordered trees. *Ann. Probab.*, 37(2) :742–789, 03 2009. 34

[70] S. D. Jacka and G. O. Roberts. Weak convergence of conditioned processes on a countable state space. *Journal of Applied Probability*, 32(4) :pp. 902–916, 1995. 7

[71] A. Joulin. A new poisson-type deviation inequality for markov jump processes with positive wasserstein curvature. *Bernoulli*, 15(2) :pp. 532–549, 2009. 100

[72] M. Kac. Foundations of kinetic theory. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Volume 3 : Contributions to Astronomy and Physics*, pages 171–197, Berkeley, Calif., 1956. University of California Press. 13, 14, 93

[73] A. Kolmogorov. Zur lösung einer biologischen aufgabe. *Comm. Math. Mech. Chebyshev Univ. Tomsk*, 2 :1–6, 1938. 2, 3

[74] T. Leliévre. Two mathematical tools to analyze metastable stochastic processes. In A. Cangiani, R. L. Davidchack, E. Georgoulis, A. N. Gorban, J. Levesley, and M. V. Tretyakov, editors, *Numerical Mathematics and Advanced Applications 2011*, pages 791–810. Springer Berlin Heidelberg, 2013. 33

[75] Q. Liu and F. Watbled. Exponential inequalities for martingales and asymptotic properties of the free energy of directed polymers in a random environment. *Stochastic Processes and their Applications*, 119(10) :3101 – 3132, 2009.

[76] P. Maillard. Branching Brownian motion with selection of the N right-most particles : An approximate model. *ArXiv e-prints*, Dec. 2011. 34

[77] F. Malrieu. Logarithmic sobolev inequalities for some nonlinear pde's. *Stochastic Processes and their Applications*, 95(1) :109 – 132, 2001. 14, 16, 17, 96, 98, 99, 101

[78] F. Malrieu. Convergence to equilibrium for granular media equations and their euler schemes. *Ann. Appl. Probab.*, 13(2) :540–560, 05 2003. 14, 96, 98, 99, 101

[79] N. Marić. *Quasi-stationary distributions and Fleming-Viot processes*. PhD thesis, Universidade de Sao Paulo, 2006. 20, 44

[80] N. Marić. Fleming-Viot particle system driven by a random walk on \$\mathbb{N}\$. *Journal of Statistical Physics*, 160(3) :548–560, 2015. 20, 96

[81] S. Martínez, J. San Martín, and D. Villemonais. Existence and uniqueness of a quasistationary distribution for Markov processes with fast return from infinity. *J. Appl. Probab.*, 51(3) :756–768, 09 2014. 8, 12, 47, 54, 70, 72

[82] H. P. McKean. Propagation of chaos for a class of non-linear parabolic equations. In *Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ., 1967)*, pages 41–57. Air Force Office Sci. Res., Arlington, Va., 1967. 13, 14, 93, 98

[83] S. Méléard. Asymptotic behaviour of some interacting particle systems ; McKean-Vlasov and Boltzmann models. In D. Talay and L. Tubaro, editors, *Probabilistic Models for Nonlinear Partial Differential Equations*, volume 1627 of *Lecture Notes in Mathematics*, pages 42–95. Springer Berlin Heidelberg, 1996. 13, 14, 94, 98

[84] S. Méléard and D. Villemonais. Quasi-stationary distributions and population processes. *Probab. Surv.*, 9 :340–410, 2012. 4, 5, 6, 7, 47, 54, 71, 83

[85] 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. 29, 51, 62, 101, 113

[86] L. Miclo. An example of application of discrete Hardy's inequalities. *Markov Process. Related Fields*, 5(3) :319–330, 1999. 23, 32, 50, 87

[87] P. Pollett. Quasi-stationary distributions : a bibliography. http ://www.maths.uq .edu.au/∼pkp/papers/qsds/qsds.pdf. 5

[88] P. Robert. *Stochastic Networks and Queues*. Stochastic Modelling and Applied Probability Series. Springer-Verlag, New York, 2003. 29, 42

[89] M. Rousset. On the control of an interacting particle estimation of Schrödinger ground states. *SIAM J. Math. Analysis*, 38(3) :824–844, 2006. 48, 55

[90] 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. 80

[91] E. Seneta and D. Vere-Jones. On quasi-stationary distributions in discrete-time Markov chains with a denumerable infinity of states. *J. Appl. Probability*, 3 :403–434, 1966. 4, 6

[92] T. Shiga and H. Tanaka. Central limit theorem for a system of Markovian particles with mean field interactions. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 69(3) :439–459, 1985. 98

[93] W. L. Smith and W. E. Wilkinson. On branching processes in random environments. *Ann. Math. Statist.*, 40(3) :814–827, 06 1969. 2

[94] A.-S. Sznitman. Nonlinear reflecting diffusion process, and the propagation of chaos and fluctuations associated. *Journal of Functional Analysis*, 56(3) :311 – 336, 1984. 98

[95] 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. 13, 14, 52, 94, 98

[96] H. Tanaka and M. Hitsuda. Central limit theorem for a simple diffusion model of interacting particles. *Hiroshima Math. J.*, 11(2) :415–423, 1981. 98

[97] M.-N. Thai. Birth and Death Process in Mean Field Interaction. *Preprint*. 93

[98] E. A. van Doorn. Conditions for the existence of quasi-stationary distributions for birth-death processes with killing. *Memorandum No. 1949, Department of Applied Mathematics, University of Twente*, 2011. 6

[99] C. Villani. *Optimal transport : old and new*. Grundlehren der mathematischen Wissenschaften. Springer, Berlin, 2009. 31

[100] D. Villemonais. *Distributions quasi-stationnaires et méthodes particulaires pour l'approximation de processus conditionnés*. PhD thesis, École Polytechnique, CMAP, 2011.

[101] D. Villemonais. Interacting particle systems and Yaglom limit approximation of diffusions with unbounded drift. *Electron. J. Probab.*, 16 :no. 61, 1663–1692, 2011. 51

[102] D. Villemonais. General approximation method for the distribution of Markov processes conditioned not to be killed. *ESAIM : Probability and Statistics*, 18 :441–467, 1 2014. 33, 47, 53, 71

[103] D. Villemonais. Minimal quasi-stationary distribution approximation for a birth and death process. *Electron. J. Probab.*, 20 :no. 30, 1–18, 2015. 6, 12, 44

[104] H. W. Watson and F. Galton. On the probability of the extinction of families. *The Journal of the Anthropological Institute of Great Britain and Ireland*, 4 :138–144, 1875. 2

[105] G. A. Watterson. Reversibility and the age of an allele. I. Moran's infinitely many neutral alleles model. *Theoret. Population Biology*, 10(3) :239–253, 1976. 73

[106] A. M. Yaglom. Certain limit theorems of the theory of branching random processes. *Doklady Akad. Nauk SSSR (N.S.)*, 56 :795–798, 1947. 3

[107] O. Zeitouni. Branching random walks and gaussian fields notes for lectures, 2012. 19