THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PSL

Préparée à l'Université Paris Dauphine

Modèles de Moran multi-alléliques et distributions quasi-stationnaires

Soutenue par

Josué CORUJO RODRIGUEZ

Le 3 décembre 2021

École doctorale n°543

École Doctorale de Dauphine

Spécialité

Sciences

Composition du jury :

M. Djalil CHAFAÏ Université Paris-Dauphine	*Directeur de thèse*
M. Bertrand CLOEZ INRAE de Montpellier	*Examinateur*
M. Pierre DEL MORAL INRIA de Bordeaux	*Examinateur*
Mme. Simona GRUSEA INSA de Toulouse	*Co-encadrante*
M. Amaury LAMBERT École Normale Supérieure	*Examinateur*
Mme. Vlada LIMIC CNRS, Université de Strasbourg	*Rapporteure*
Mme. Amandine VÉBER CNRS, Université de Paris	*Présidente*
M. Denis VILLEMONAIS Université de Lorraine	*Rapporteur*

Liste des publications

Ce manuscrit est basé sur les publications et pré-publications suivantes[1]:

- B. Cloez and J. Corujo,
  Uniform in time propagation of chaos for a Moran model,
  arXiv e-prints 2107.10797v2 (2021) (soumis).

- J. Corujo,
  On the spectrum and ergodicity of a neutral multi-allelic Moran model,
  arXiv e-prints 2010.08809v2 (2021) (soumis).

- J. Corujo,
  Dynamics of a Fleming – Viot type particle system on the cycle graph,
  Stochastic Processes and their Applications **136** (2021), 57–91,
  DOI: 10.1016/j.spa.2021.02.001.

Pendant mes années de doctorat, j'ai également eu l'occasion de poursuivre deux autres projets de collaborations. Les trois publications suivantes sont des résultats de ces échanges. Celles-ci ne sont pas incluses dans ce manuscrit.

- J. Corujo and J. E. Valdés,
  Further results on stochastic orderings and aging classes in systems with age replacement,
  Probability in the Engineering and Informational Sciences (2021), 1–30,
  DOI: 10.1017/S0269964821000036.

- J. Corujo, D. Flores-Peñaloza, C. Huemer, P. Pérez-Lantero and C. Seara,
  Matching Random Colored Points with Rectangles,
  In: M. Rahman, K. Sadakane, WK. Sung (eds) WALCOM: Algorithms and Computation. WALCOM 2020. Lecture Notes in Computer Science, vol 12049. Springer, Cham. (2020)
  DOI: 10.1007/978-3-030-39881-1_22.

- W. Rodríguez, O. Mazet, S. Grusea, A. Arredondo, J. Corujo, S. Boitard and L. Chikhi,
  The IICR and the non-stationary structured coalescent: towards demographic inference with arbitrary changes in population structure,
  Heredity (2018), 663–678,
  DOI: 10.1038/s41437-018-0148-0.

[1] Une liste de mes publications est disponible sur https://sites.google.com/view/josuecorujo/research.

Contents

Résumé

L'objectif principal de cette thèse est d'étudier l'évolution, en temps long et pour une grande taille de population, des modèles de Moran multi-alléliques, qui sont des processus de Markov à temps continu et à espace discret, inspirés de modèles mathématiques pour la biologie. Nous nous intéressons à l'étude, entre autres aspects, de la relation entre le processus de Moran, compris comme un système de particules en interaction, et la théorie des distributions quasi-stationnaires. Plus précisément, nous exhibons des phénomènes de propagation du chaos lorsque la taille de la population est grande, et nous établissons des contrôles quantitatifs de la convergence en temps long vers l'équilibre. Les principaux résultats sont divisés en trois chapitres. Dans le premier chapitre, nous montrons que la mesure de probabilité empirique induite par le système de particules converge, lorsque la taille de la population est grande, vers la loi d'une chaîne de Markov absorbante conditionnée à ne pas être absorbée. De plus, nous établissons un contrôle de cette convergence, en prouvant une propagation du chaos uniforme en temps. Nous prouvons également la normalité asymptotique du biais et nous fournissons une expression explicite pour la variance asymptotique, utilisée ensuite pour définir un autre système de particules avec une erreur quadratique plus petite. Dans le deuxième chapitre, nous considérons un modèle plus simple où l'espace d'état est fini et le taux de mortalité est uniforme. Dans ce contexte, nous trouvons une expression explicite pour le spectre du générateur du système de particules en termes de spectre de la matrice des taux de mutation. De plus, nous étudions l'ergodicité du processus et, pour un schéma particulier de mutation, mutation indépendante des parents, nous sommes en mesure de prouver l'existence de phénomènes de cutoff pour les distances de variation totale et chi-deux. Le troisième chapitre est consacré à l'étude d'un cas particulier, où le processus de mutation correspond à une marche aléatoire asymétrique sur le graphe cyclique. Nous montrons que ce modèle possède une solvabilité remarquable, malgré le fait qu'il soit non-réversible avec une distribution invariante non-explicite.

Mots clés : Processus de Moran multi-allélique ; systèmes de particules de type Fleming-Viot ; distribution quasi-stationnaire ; propagation du chaos ; ergodicité ; phénomène de cutoff

Abstract

The main goal of this thesis is to study the evolution of a multi-allelic Moran model, which is a continuous-time discrete state Markov process, inspired by biological applications. We study, among many other aspects, the relation between the Moran process, understood as an interacting particle system, and the theory of quasi-stationary distributions. More precisely, we prove the existence of a propagation of chaos phenomenon when the population size is large, and we study the quantitative control for the long time convergence to stationarity by spectral arguments. The main results are divided in three chapters. In the first chapter we show that the empirical probability measure induced by the particle system converges, when the number of particles goes to infinity, to the law of an absorbing Markov process conditioned to non-absorption. Furthermore, we establish a control on this convergence, by proving a uniform in time propagation of chaos. We also prove the asymptotic normality of the bias and we provide an explicit expression for the asymptotic variance, which is later used to define another particle system with smaller quadratic error. In the second chapter, we consider a simpler model where the state space is finite and the killing rate is uniform. In this context we find an explicit expression for the spectrum of the particle system generator in terms of the spectrum of the mutation rate matrix. Moreover, we study the ergodicity of the process and, for a particular mutation scheme, which is the parent independent mutation, we are able to prove the existence of cutoff phenomena in the total variation and chi-square distances. The third chapter is devoted to the study of a particular case, where the mutation process is driven by an asymmetric random walk on the cycle graph. We show that this model has a remarkable exact solvability, despite the fact that it is non-reversible with non-explicit invariant distribution.

Keywords: multi-allelic Moran process; Fleming – Viot particle system; quasi-stationary distribution; propagation of chaos; ergodicity; cutoff phenomenon.

Introduction (version française)

Cette thèse est basée sur l'étude d'un système markovien de particules en interaction, à temps continu et à espace discret, modélisant l'évolution d'une population de N individus de différents types donnés par les éléments d'un ensemble discret E. En bref, chaque individu change de type selon un processus de mutation markovien, indépendamment des autres. L'interaction se produit lors des événements de reproduction où un individu meurt et un autre choisi au hasard, possiblement le même, est dupliqué. Ce type de reproduction est de type Moran et nous appelons ce processus *modèle de Moran multi-allélique* [Dur08; EG09].

Deux quantités asymptotiques d'intérêt apparaissent dans l'étude de ces processus : la distribution stationnaire du processus de particules (la distribution limite lorsque le temps tend vers l'infini) et la distribution empirique limite induite par le processus de particules dans E (la limite lorsque le nombre de particules tend vers l'infini). L'existence de ces limites et les vitesses de convergence sont d'un grand intérêt pour des raisons théoriques et pratiques. Cette thèse est centrée sur l'étude des vitesses de ces convergences décrites précédemment.

Nous présentons d'abord le modèle de Moran multi-allélique et décrivons plus explicitement les limites qui retiendront notre attention dans la suite de l'exposé. Nous décrivons ensuite les motivations et les problèmes auxquels cette thèse s'adresse. Enfin, nous résumons les principaux résultats de la thèse, qui sont présentés et démontrés en détail dans les chapitres suivants.

Le modèle de Moran multi-allélique

Considérons E un espace d'état discret. L'espace d'état du modèle de Moran multi-allélique est le N simplexe discret

$$\mathcal{E}_N := \left\{ \eta : E \to \mathbb{N} \;\middle|\; \sum_{x \in E} \eta(x) = N \right\},$$

La distribution empirique induite par $\eta \in \mathcal{E}_N$ est définie par

$$m(\eta) = \sum_{x \in E} \frac{\eta(x)}{N} \delta_x \in \mathcal{M}_1(E),$$

où $\mathcal{M}_1(E)$ est l'ensemble des mesures de probabilité sur E.

Soit Q le générateur d'une chaîne de Markov irréductible sur E, non-explosive à temps continu. Considérons les taux $V_\mu(x, y) \geq 0$, pour tout $x \neq y \in E$ et $\mu \in \mathcal{M}_1(E)$.

Le processus de Moran multi-allélique est une chaînes de Markov à temps continu sur $\mathcal{E}_N$. Le processus est dans l'état $\eta \in \mathcal{E}_N$ s'il existe $\eta(x) \in \{0, 1 \ldots, N\}$ individus avec le type allélique x, pour chaque $x \in E$. Entre deux événements de reproduction, les N individus évoluent comme des copies indépendantes du processus de mutation généré par $Q = (Q_{x,y})_{x,y \in E}$. Dans ce sens, nous appelons $Q_{x,y}$, pour $x, y \in E$, les *taux de mutation*.

Les événements de reproduction consistent en la mort d'un individu de type x, qui est alors retiré de la population, et en la reproduction d'un individu de type y, qui ajoute un individu y

à la population. Cela se produit avec le taux $\eta(y)/N \cdot V_{m(\eta)}(x,y)$. Par conséquent, le taux de transition de $\eta \in \mathcal{E}_N$, avec $\eta(x) > 0$, à $\eta - \mathbf{e}_x + \mathbf{e}_y$ est

$$\eta(x)\left(Q_{x,y} + \frac{\eta(y)}{N} V_{m(\eta)}(x,y)\right),$$

pour tout $x \neq y \in E$, où $\eta - \mathbf{e}_x + \mathbf{e}_y$ est l'élément de l'ensemble $\mathcal{E}_N$ définie comme suit

$$(\eta - \mathbf{e}_x + \mathbf{e}_y)(z) = \begin{cases} \eta(z) & \text{if} \quad z \notin \{x,y\}, \\ \eta(x) - 1 & \text{if} \quad z = x, \\ \eta(y) + 1 & \text{if} \quad z = y. \end{cases}$$

Par conséquent, le modèle de Moran décrit précédemment est une famille de chaînes de Markov à temps continu sur $\mathcal{E}_N$, avec générateurs $\mathcal{Q}_N$, qui agissent sur toute fonction f sur $\mathcal{E}_N$ comme suit

$$(\mathcal{Q}_N f)(\eta) = \sum_{x,y \in E} \eta(x)\left(Q_{x,y} + \frac{\eta(y)}{N} V_{m(\eta)}(x,y)\right)[f(\eta - \mathbf{e}_x + \mathbf{e}_y) - f(\eta)], \tag{1}$$

pour tout $\eta \in \mathcal{E}_N$.

Nous détaillerons plus loin des exemples particuliers mais, pour l'instant, voyons que lorsque $V_{m(\eta)}(x,y)$ est constant, chaque individu meurt au même taux et le parent est choisi uniformément au hasard parmi les individus présents dans la population (ceci explique le terme $\eta(y)/N$ dans le taux de transition). On peut aussi interpréter ce taux par le point de vue opposé : chaque individu se reproduit à un taux constant et l'individu mourant est choisi uniformément au hasard. Ceci est souvent appelé *sélection neutre* dans la littérature écologique, mais nos modèles permettent de choisir divers $V_{m(\eta)}(x,y)$ non constants. Dans ce sens, nous appelons les taux $V_\mu(x,y)$, pour tout $x,y \in E$ et $\mu \in \mathcal{M}_1(E)$, les *taux de sélection*.

Notez que la dynamique de reproduction dépend en général à la fois des types du parent et du descendant, et peut également dépendre de la distribution empirique induite par la configuration de la population à l'instant en cours, dans un sens que nous préciserons ultérieurement dans les Hypothèses (G1) et (C1).

La distribution empirique induite par le processus de particules au temps t est donnée par $m(\eta_t^{(N)})$. Le but de ce travail de thèse est de comprendre les limites et les vitesses de convergence de $m(\eta_t^{(N)})$ lorsque $t \to \infty$ et $N \to \infty$.

Dans toute la thèse, la condition suivante est supposée :

$$\|V\| := \sup_{\mu \in \mathcal{M}_1(E)} \sup_{x,y \in E} V_\mu(x,y) < \infty.$$

Notons que la condition de non-explosion sur Q et la condition de borne sur $\|V\|$ excluent la possibilité d'un nombre infini de sauts en temps fini. Ainsi, le processus est bien défini pour tout $t \geq 0$.

Le modèle de Moran multi-allélique généré par $\mathcal{Q}_N$ est une extension, pour $K > 2$, du modèle bi-allélique étudié par [Cor17]. En général, lors de la généralisation du modèle de Moran pour plus de deux types alléliques, les taux de sélection étudiés dépendent que du type du descendant, c'est-à-dire $V_\mu(x,y) = V^{\mathrm{b}}(y)$, pour tous les $x,y \in E$ et $\mu \in \mathcal{M}_1(E)$, taux appelé *sélection à la naissance* ou *sélection sur la fertilité*[2] [Dur08; MW09; Eth11]. En outre, dans les applications biologiques, le modèle avec *sélection à la mort* ou *sélection à la viabilité*[3] a également été considéré, lorsque les taux de sélection ne dépendent que du type

[2] *Selection at birth* ou *fecundity selection* en anglais.

[3] *Selection at death* ou *viability selection* en anglais.

du parent, c'est-à-dire, $V_\mu(x,y) = V^{\mathrm{d}}(x)$ [MW09], pour tous les $x, y \in E$ et $\mu \in \mathcal{M}_1(E)$. Cependant, l'importance de ce dernier modèle va au-delà de ses interprétations biologiques : ce processus est également appelé *processus de particules de Fleming – Viot*, processus de particules en interaction permettant d'approcher la *distribution quasi-stationnaire* d'une chaîne de Markov absorbante. Nous discuterons ultérieurement la relation entre le modèle de Moran considéré ici, les semigroupes de Feynman – Kac, la théorie des distributions quasi-stationnaires et les processus de particules de Fleming – Viot.

Nous avons que $\mathcal{Q}_N = \mathcal{Q}_N^{\mathrm{mut}} + \mathcal{Q}_N^{\mathrm{sel}}$, où ces générateurs agissent sur chaque $f \in \mathcal{B}_d(\mathcal{E}_N)$ comme suit :

$$(\mathcal{Q}_N^{\mathrm{mut}} f)(\eta) = \sum_{x,y \in E} \eta(x) Q_{x,y} [f(\eta - \mathbf{e}_x + \mathbf{e}_y) - f(\eta)],$$

$$(\mathcal{Q}_N^{\mathrm{sel}} f)(\eta) = \sum_{x,y \in E} \eta(x) \frac{\eta(y)}{N} V_{m(\eta)}(x,y) [f(\eta - \mathbf{e}_x + \mathbf{e}_y) - f(\eta)],$$

pour tout $\eta \in \mathcal{E}_N$, où $\mathcal{B}_b(\mathcal{E}_N)$ dénote l'ensemble des fonctions réelles bornées sur $\mathcal{E}_N$. Notons que l'irréductibilité de Q implique l'irréductibilité de $\mathcal{Q}_N$. Le générateur $\mathcal{Q}_N^{\mathrm{mut}}$ pilote le mouvement indépendant de N particules selon le taux de mutation Q. Par ailleurs, $\mathcal{Q}_N^{\mathrm{sel}}$ génère une chaîne de Markov absorbante sur $\mathcal{E}_N$. En effet, une fois qu'un site $y \in E$ est vide (le processus est dans un état η tel que $\eta(y) = 0$) aucune particule ne peut sauter en y. Ainsi, les états η_x tels que $\eta_x(x) = N$, pour $x \in E$, sont absorbants pour le processus généré par $\mathcal{Q}_N^{\mathrm{sel}}$.

Essayons d'avoir une idée plus claire de la limite de la mesure empirique induite par ce processus de particules lorsque le nombre de particules tend vers l'infini. Par l'équation de Kolmogorov, nous savons que $\partial_t \mathbb{E}_\eta[m_x(\eta_t)] = \mathbb{E}_\eta\left[(\mathcal{Q}_N\, m_x)(\eta_t)\right]$, où m_x représente la distribution empirique induite par η au point $x \in E$, c'est-à-dire, $m_x : \eta \mapsto \eta(x)/N$. Or,

$$(\mathcal{Q}_N^{\mathrm{mut}} m_x)(\eta) = \sum_{y \in E} Q_{y,x} m_y(\eta),$$

pour tous $x \in E$ and $\eta \in \mathcal{E}_N$. D'un autre côté,

$$(\mathcal{Q}_N^{\mathrm{sel}} m_x)(\eta) = -\, m_x(\eta) \sum_{y \in E} m_y(\eta) [V_{m(\eta)}(x,y) - V_{m(\eta)}(y,x)].$$

Enfin, nous obtenons

$$\partial_t \mathbb{E}_\eta[m_x(\eta_t)] = \sum_{y \in E} Q_{y,x} \mathbb{E}_\eta[m_y(\eta_t)] - \sum_{y \in E} [V_{m(\eta)}(x,y) - V_{m(\eta)}(y,x)] \mathbb{E}_\eta[m_x(\eta_t) m_y(\eta_t)].$$

Lorsque le nombre d'individus $N \to \infty$, nous nous attendons à ce que le processus présente un *phénomène de propagation du chaos* :

$$|\mathbb{E}_\eta[m_x(\eta_t) m_y(\eta_t)] - \mathbb{E}_\eta[m_x(\eta_t)] \mathbb{E}_\eta[m_y(\eta_t)]| \xrightarrow[N \to \infty]{} 0.$$

Autrement dit, les positions des particules deviennent asymptotiquement indépendantes. De cette manière, la distribution empirique induite par le processus approche la solution de l'équation différentielle ordinaire non linéaire suivante :

$$\partial_t \gamma_t(x) = \sum_{y \in E} Q_{y,x} \gamma_t(y) - \sum_{y \in E} [V_{\gamma_t}(x,y) - V_{\gamma_t}(y,x)] \gamma_t(x) \gamma_t(y), \tag{2}$$

pour tout $x \in E$.

Précisons maintenant la dépendance des taux de sélection par rapport à la distribution empirique induite par les particules. Supposons que les taux de sélection permettent la décomposition suivante

$$V_\mu(x,y) = V_\mu^{\mathrm{d}}(x) + V_\mu^{\mathrm{b}}(y) + V_\mu^{\mathrm{s}}(x,y), \tag{3}$$

où $\mu \mapsto V_\mu^{\mathrm{d}}$ et $\mu \mapsto V_\mu^{\mathrm{b}}$ sont des applications continues bornées de $(\mathcal{M}_1(E), \|\cdot\|_{\mathrm{TV}})$ sur $(\mathcal{B}_+(E), \|\cdot\|)$, et $\mu \mapsto V_\mu^{\mathrm{s}}$ est une autre application continue bornée de $(\mathcal{M}_1(E), \|\cdot\|_{\mathrm{TV}})$ sur $(\mathcal{B}_+(E \times E), \|\cdot\|)$. De plus, V_μ^{s} est symétrique dans $E \times E$, pour chaque $\mu \in \mathcal{M}_1(E)$. Nous supposons également qu'il existe une fonction $\Lambda \in \mathcal{B}_b(E)$, tel que

$$\Lambda = V_\mu^{\mathrm{b}} - V_\mu^{\mathrm{d}}.$$

Ainsi, (2) se réduit comme suit : pour toute fonction ϕ sur E nous avons

$$\partial_t \gamma_t(\phi) = \gamma_t\Big((Q + \Lambda)\phi - \gamma_t(\Lambda) \cdot \phi\Big). \tag{4}$$

En effet, notons que

$$\begin{aligned}\sum_{y \in E} \phi(x)[V_\gamma(x,y) - V_\gamma(y,x)]\gamma(x)\gamma(y) &= \gamma(Q\phi) + \gamma(V_\gamma^{\mathrm{d}})\gamma(\phi) - \gamma(V_\gamma^{\mathrm{d}}\phi) + \gamma(V_\gamma^{\mathrm{b}}\phi) - \gamma(V_\gamma^{\mathrm{b}})\gamma(\phi) \\ &= \gamma(Q\phi + \Lambda\phi - \gamma(\Lambda)\phi).\end{aligned}$$

Remarque 1 (Indépendance de la composante symétrique dans (3)). Notons que la fonction symétrique V_μ^{s} est absente dans (4). En d'autres termes, le système de particules se rapprocherait de la même mesure déterministe, qui est donnée par la solution de (4), indépendamment de l'élément symétrique dans la décomposition (3) des taux de sélection.

Remarque 2 (Feynman – Kac pour le modèle avec taux de sélection généraux). Voir [Del04, p. 25], et les références qui y sont données, pour une interprétation du type Feynman – Kac pour l'équation différentielle (2) avec des taux de sélection assez généraux.

Le schéma suivant résume les quatre convergences qui nous étudions dans la suite :

$$\begin{array}{ccc} m(\eta_t^{(N)}) & \xrightarrow[t\to\infty]{} & \nu_N \\ N\big\downarrow & & \big\downarrow N \\ \gamma_t & \xrightarrow[t\to\infty]{} & \nu_\infty \end{array} \tag{5}$$

où ν_N est la distribution stationnaire du processus généré par $\mathcal{Q}_N$ et $(\gamma_t)_{t\geq 0}$ est la solution de (4), quand ils existent.

Semigroupes de Feynman – Kac et distributions quasi-stationnaires

Considérons $(X_t)_{t\geq 0}$ la chaîne de Markov sur E générée par Q. Définissons le semigroupe de Feynman – Kac pour chaque $\phi \in \mathcal{B}_b(E)$ par :

$$P_t^\Lambda(\phi) : x \mapsto \mathbb{E}_x\left[\phi(X_t)\exp\left\{\int_0^t \Lambda(X_s)\mathrm{d}s\right\}\right],$$

dont le générateur est $Q + \Lambda$. La version normalisée du semigroupe, qui est définie par :

$$\mu_t(\phi) := \frac{\mu_0 P_t^\Lambda(\phi)}{\mu_0 P_t^\Lambda(\mathbf{1})}, \tag{6}$$

où $\mathbf{1}$ désigne la fonction $x \mapsto 1$ sur E, est la solution de l'équation différentielle non linéaire (4) avec valeur initiale $\mu_0(\phi)$ pour $t = 0$ [Del04, Eq. (1.17)]. Pour une étude exhaustive et compréhensible des formules de Feynman – Kac, nous renvoyons au livre de Del Moral [Del04].

Notons que $(\mu_t)_{t\geq 0}$ tel que défini ci-dessus est invariant par translation de la fonction Λ. En effet, pour tout réel β on obtient

$$\mu_t(\phi) = \frac{\mu_0 P_t^{\Lambda-\beta}(\phi)}{\mu_0 P_t^{\Lambda-\beta}(\mathbf{1})}.$$

En particulier, en prenant $\beta = \sup \Lambda$, on peut toujours interpréter $(\mu_t)_{t\geq 0}$ comme la distribution d'une chaîne de Markov absorbante conditionnée à la non absorption jusqu'au temps t et avec un taux de mortalité $\kappa = \sup \Lambda - \Lambda$. Ceci relie naturellement l'étude du comportement de $(\mu_t)_{t\geq 0}$ lorsque $t \to \infty$, à la théorie des *distributions quasi-stationnaires* (QSD), que nous présentons dans la section suivante.

Distributions quasi-stationaires

Notons par $(X_t)_{t\geq 0}$ une chaîne de Markov non-explosive irréductible à temps continu sur un espace discret E de générateur Q. Soit $\kappa : E \to \mathbb{R}_+$ une fonction bornée. Nous ajoutons un état absorbant $\partial \notin E$ et définissons la chaîne de Markov absorbante $(Y_t)_{t\geq 0}$ sur $E \cup \{\partial\}$ vérifiant

$$Y_t = \begin{cases} X_t & \text{si } \int_0^t \kappa(X_s)\mathrm{d}s < \xi \\ \partial & \text{sinon,} \end{cases}$$

où ξ est une variable aléatoire exponentielle de paramètre 1, indépendante de $(X_t)_{t\geq 0}$. En d'autres termes, la chaîne $(Y_t)_{t\geq 0}$ se comporte comme $(X_t)_{t\geq 0}$ sur E, et conditionnellement à être dans l'état $x \in E$, elle saute à ∂ (est absorbée) avec un taux $\kappa(x)$.

Alternativement, $(Y_t)_{t\geq 0}$ peut être définie comme la chaîne de générateur $\tilde{Q}$ agissant sur des fonctions bornées φ sur $E \cup \{\partial\}$ de la manière suivante

$$(\tilde{Q}\varphi) : x \mapsto \kappa(x)(\varphi(\partial) - \varphi(x)) + \sum_{y\in E} Q_{x,y}(\varphi(y) - \varphi(x)),$$

pour tout $x \in E$, et on fixe $(\tilde{Q}\varphi)(\partial) = 0$.

Notons τ_∂ le temps d'absorption. On a

$$\mathbb{E}_\mu[\phi(Y_t)] = \mathbb{E}_\mu[\phi(X_t) \mid t < \tau_\partial] = \frac{\mathbb{E}_\mu\left[\phi(X_t)\mathrm{e}^{-\int_0^t \kappa(X_s)\mathrm{d}s}\right]}{\mathbb{E}_\mu\left[\mathrm{e}^{-\int_0^t \kappa(X_s)\mathrm{d}s}\right]}. \tag{7}$$

Notons la ressemblance entre l'expression précédente et le semigroupe normalisé de Feynman – Kac défini par (6).

Remarque 3 (Lien entre la mesure γ_t dans (5) et une chaîne de Markov absorbante). Lorsque la décomposition (3) est vérifiée, la limite de $m(\eta_t^{(N)})$ lorsque $N \to \infty$, que nous avons notée γ_t dans (5), est la loi d'une chaîne de Markov absorbante conditionnée à la non absorption jusqu'au temps $t \geq 0$. De plus, le générateur d'une de ces chaînes de Markov absorbantes est $Q + \kappa$ où $\kappa := \sup \Lambda - \Lambda \in \mathcal{B}_b(E)$.

Le temps d'absorption τ_∂, qui est définie comme suit

$$\tau_\partial := \inf_{t\geq 0}\{X_t = \partial\},$$

est un temps d'arrêt. C'est-à-dire, pour tout $s \geq 0$, $X_s = \partial$ implique que $X_t = \partial$, pour tout $t \geq s$. Tout au long de ce manuscrit, nous supposons qu'en partant de n'importe quel $x \in E$, le processus sera finalement absorbé, presque sûrement. Concrètement, nous supposons que

$$\mathbb{P}_x(\tau_\partial < \infty) = 1, \ \ \forall x \in E.$$

Ainsi, la distribution stationnaire de $(X_t)_{t\geq 0}$, qui est la distribution concentrée en ∂, n'est pas intéressante. Cependant, il est intéressant de considérer la distribution stationnaire du processus conditionné à la non absorption, que l'on appelle la *distribution quasi-stationnaire*.

Définition 1 (Distribution quasi-stationnaire (QSD[4]))**.** Une distribution quasi-stationnaire pour la chaîne de Markov absorbante $(Y_t)_{t\geq 0}$ est toute $\mu \in \mathcal{M}_1(E)$ vérifiant

$$\mathbb{P}_\mu(Y_t \in \cdot \mid t < \tau_\partial) = \mu.$$

Définition 2 (Distribution quasi-limite (QLD[5]))**.** Une distribution quasi-limite pour la chaîne de Markov absorbante $(Y_t)_{t\geq 0}$ est toute $\mu \in \mathcal{M}_1(E)$ vérifiant

$$\lim_{t\to\infty} \mathbb{P}_\nu(Y_t \in \cdot \mid t < \tau_\partial) = \mu_{\mathrm{QLD}},$$

pour une certaine $\nu \in \mathcal{M}_1(E)$.

Lorsque la limite ci-dessus existe et est la même pour tout $\nu = \delta_x$, avec $x \in E$, la distribution quasi-limite est dite être la *limite de Yaglom*[6] ou la *distribution quasi-stationnaire minimale* de $(Y_t)_{t\geq 0}$.

De manière analogue à la théorie ergodique classique des chaînes de Markov, la QSD et la QLD sont des objets équivalents. De plus, une QSD minimale est toujours une QSD mais l'implication inverse n'est pas toujours vraie :

$$\boxed{\text{QSD}} \Leftrightarrow \boxed{\text{QLD}} \Leftarrow \boxed{\text{Limite de Yaglom (QSD minimale)}}.$$

Pour plus de détails sur l'étude des distributions quasi-stationnaires, nous renvoyons à l'article de Méléard et Villemonais [MV12], au livre de Collet, Martínez et San Martín [CMS13], et à l'article de van Doorn et Pollet [DP13], pour le contexte spécifique des espaces d'états discrets.

De plus, la définition de QLD et la Remarque 3 impliquent que la limite ν_∞ définie dans (5) peut être comprise comme une QSD.

Remarque 4 (La mesure ν_∞ dans (5) est une QSD)*.* Lorsque la décomposition (3) est vérifiée, la distribution de probabilité ν_∞ dans (5) est la QSD de la chaîne de Markov absorbante avec générateur $Q + \kappa$, où $\kappa := \sup \Lambda - \Lambda \in \mathcal{B}_b(E)$.

Au cours de ce travail, nous considérons une condition plus forte : la convergence exponentielle uniforme du semigroupe normalisé de Feynman – Kac. Nous supposons qu'il existe une distribution $\mu_\infty \in \mathcal{M}_1(E)$ et des constantes $C, \gamma > 0$, telles que

$$\|\mu_t - \mu_\infty\|_{\mathrm{TV}} \leq C\mathrm{e}^{-\gamma t}, \ \text{pour tous } \mu_0 \in \mathcal{M}_1(E) \text{ et } t \geq 0, \tag{8}$$

où $(\mu_t)_{t\geq 0}$ est définie comme dans (6).

La convergence exponentielle (8) est toujours vérifiée lorsque l'espace d'état est fini. Dans ce cas, la QSD est donnée par le vecteur propre normalisé à gauche du générateur Q, associé

[4]La notation QSD provient du nom en anglais : *quasi-stationary distribution*.

[5]La notation QLD provient du nom en anglais : *quasi-limit distribution*.

[6]Ce nom provient du mathématicien A. M. Yaglom, qui a commencé l'étude de ce type d'objets dans le cadre d'un processus de Galton – Watson sous-critique dans [Yag47].

à la deuxième plus grande valeur propre en module (SLEM[7]). Ceci a été prouvé par Darroch et Seneta [DS67] et le résultat est une conséquence du théorème de Perron – Frobenius [Per07; Fro12] (voir également [MV12, Thm. 8] pour le contexte spécifique des distributions quasi-stationnaires).

Le cas où E est dénombrable est plus délicat et a suscité beaucoup d'attention et plusieurs méthodes ont été appliquées. Grâce au travail exhaustif de Champagnat et Villemonais, notamment [CV16; CV17b], il est possible de décrire des hypothèses équivalentes à (8) et d'explorer les conséquences de la convergence exponentielle uniforme vers la QSD. Nous discutons plus en détail ce sujet dans la Section 1.2. Par la suite, énonçons simplement le résultat suivant qui clarifie les conséquences de (8) sur l'ergodicité exponentielle du semigroupe non normalisé $(P_t^\Lambda)_{t\geq 0}$.

Lemme 1 (Ergodicité exponentielle du semigroupe non normalisé)**.** *Supposons que* (8) *soit vérifié. Alors, il existe un unique triplet* $(\mu_\infty, h, \lambda) \in \mathcal{M}_1(E) \times \mathcal{B}_b(E) \times \mathbb{R}$ *d'éléments propres de* $Q + \Lambda$ *tels que* h *soit strictement positif,* $\mu_\infty(h) = 1$ *et tels que*

$$\mu_\infty P_t^\Lambda = \mathrm{e}^{\lambda t}\mu_\infty \ \text{et}\ P_t^\Lambda(h) = \mathrm{e}^{\lambda t} h.$$

De plus, il existe $C, \gamma > 0$ *tels que pour tout* $t \geq 0$ *:*

$$\sup_{\mu_0 \in \mathcal{M}(E)} \|\mathrm{e}^{-\lambda t}\mu_0 P_t^\Lambda - \mu_0(h)\mu_\infty\|_{\mathrm{TV}} \leq C\mathrm{e}^{-\gamma t}.$$

Par ailleurs, $\lambda \leq 0$ *lorsque* $\Lambda \leq 0$*.*

Ce résultat est essentiellement une conséquence du Théorème 2.1 de [CV17b] (Théorème 1.2.2 ci-dessous). Essentiellement, il dit que la convergence exponentielle uniforme du semigroupe normalisé implique celle du semigroupe non normalisé.

Systèmes de particules de Fleming – Viot

Le système de particules de Fleming – Viot consiste en N particules qui se déplacent dans E comme des copies indépendantes d'une chaîne de Markov absorbante, jusqu'à ce qu'une des particules soit absorbée. Lorsque cela se produit, cette particule saute instantanément et uniformément sur l'une des positions des autres particules. Les systèmes de particules de Fleming – Viot ont été introduits à l'origine et de manière indépendante par Del Moral, Guionnet et Miclo [DG99; DM00a][8] et Burdzy, Hołyst et March [BHM00] pour approcher la loi d'un processus de Markov conditionné à la non-absorption, et sa QSD dans le cadre d'un espace d'état continu. Ces processus de particules ont suscité beaucoup d'attention ces dernières années. Voir par exemple [DM03; Vil14; Cér+20; CV21] dans le cadre d'espaces d'états généraux, [FM07; GJ13; Ass+16; CT16b] pour le cadre d'espaces d'états dénombrables, et aussi [AFG11; LPR18] pour le cadre d'espaces d'états finis.

Le générateur du processus de Fleming – Viot est un cas particulier de processus de Moran multi-allélique avec *sélection à la naissance* ou *sélection sur la fertilité*, c'est-à-dire, les taux de sélection dépendent que du type de parent : $V_\mu(x, y) = \kappa(x)$, pour tous $x, y \in E$ et $\mu \in \mathcal{M}_1(E)$,

[7] La notation SLEM provient du nom en anglais : *second largest eigenvalue in modulus.*

[8] En fait, Del Moral et Miclo ont appelé ce processus *système de particules de type de Moran*, qui est peut-être un nom plus précis dans le cadre des espaces d'états discrets, afin d'éviter toute confusion avec l'existence d'un processus à valeur mesure lié au processus de Moran en génétique des populations introduit par Fleming et Viot [FV79] et nommé *processus de Fleming – Viot.* Voir [FV79, Annexe B] pour une discussion sur la relation entre le processus de Fleming – Viot (au sens de la génétique des populations) et le modèle de Moran multi-allélique, voir également [Fen10, §6.2].

avec $\kappa \in \mathcal{B}_b(E)$. Notons par $\mathcal{F}_N$ le générateur infinitésimal du processus de Fleming – Viot qui agit sur toute fonction bornée sur $\mathcal{E}_N$ de la manière suivante

$$(\mathcal{F}_N f)(\eta) = \sum_{x,y\in E} \eta(x) \left(Q_{x,y} + \frac{\eta(y)}{N}\kappa(x) \right) [f(\eta - \mathbf{e}_x + \mathbf{e}_y) - f(\eta)],$$

pour tout $\eta \in \mathcal{E}_N$.

La définition du processus de particules de Fleming – Viot est motivée par des objectifs de simulation. La simulation de la loi (7) et l'estimation de son QSD par les méthodes habituelles (comme la méthode de rejet) ne sont pas viables, puisque l'événement auquel on restreint l'observation a une probabilité tendant (généralement de façon exponentielle) vers zéro. Le système de particules de Fleming – Viot offre une méthode alternative. Sous certaines hypothèses, on peut prouver que lorsque $N \to \infty$, la mesure empirique induite par le système de particules s'approche de la distribution conditionnelle (7). Ainsi, lorsque $N, t \to \infty$, le système de particules peut être utilisé pour estimer la QSD associée à (7). Voir par exemple [GJ13] pour plus de détails sur les simulations de distributions quasi-stationnaires sur des espaces dénombrables.

Examinons, sans être exhaustif, l'état de l'art relatif à la convergence de la mesure empirique liée au système de particules de Fleming –Viot en tant qu'estimateur d'une distribution quasi-stationnaire. Ferrari et Marić [FM07] et Asselah et al. [AFG11] étudient la convergence de la distribution empirique induite par le processus de Fleming – Viot vers l'unique QSD dans des cadres d'espaces discrets dénombrables et finis, respectivement. Afin d'étudier la convergence du processus de particules sous la distribution stationnaire vers la QSD, Lelievre et al. [LPR18] prouvent un théorème central limite pour le cas des états finis. De plus, Villemonais [Vil15] et Asselah et al. [Ass+16] étudient la convergence vers la QSD minimale pour un processus de naissance et de mort et pour un processus de type Galton – Watson, respectivement. De même, Asselah et Thai [AT12] et Marič [Mar15] abordent l'étude du système de particules associé à une marche aléatoire sur $\mathbb{N}_0$ avec une dérive vers l'origine, qui est un état absorbant. Dans le scénario où l'espace d'état E est infini dénombrable (comme les processus de naissance et de mort et les processus de Galton – Watson), il peut exister une infinité de QSD. Il est donc important de garantir l'ergodicité du système de particules et de déterminer vers quelle QSD converge sa distribution empirique. Dans cette direction , Champagnat et Villemonais [CV21] étudient la convergence du processus de Fleming – Viot vers la QSD minimale sous des conditions générales, en fournissant également quelques exemples spécifiques. Voir aussi l'article de Cloez et Thai [CT16b] assurant, sous de fortes conditions de mélange, les convergences dans (5) avec des estimations quantitatives explicites pour la vitesse de convergence.

Principaux résultats

Le travail que nous présentons ici peut être vu comme une continuation de la recherche liée à l'ergodicité et à la propagation du chaos pour les processus de particules de Fleming – Viot. Dans le Chapitre 1, nous allons plus loin et étudions une classe plus générale de modèles : les modèles de Moran multi-alléliques avec des taux de sélection vérifiant (3), qui sont des systèmes de particules dont la mesure empirique s'approche également d'une QSD lorsque $t, N \to \infty$. Le Théorème 2 établie la convergence uniforme en temps d'ordre $1/\sqrt{N}$ de $m(\eta_t^{(N)})$ vers γ_t quand $N \to \infty$, comme dans le diagramme (5). En particulier, ce résultat assure une propagation du chaos uniforme dans le temps avec une vitesse d'ordre $1/\sqrt{N}$ pour le processus de Fleming – Viot. Ce problème a suscité une attention considérable ces dernières années. Voir, par exemple, [Vil14, Thm. 2.2], [CT16b, Thm. 1.5], [LPR18, Remarque 4.3] et [CV21, Thm. 2.3]. Dans le Théorème 3, nous allons plus loin et prouvons la normalité asymptotique de $\sqrt{N}\Big(m(\eta_T^{(N)})(\phi) - \gamma_T(\phi)\Big)$, lorsque $N \to \infty$, avec une expression explicite pour la variance, pour tout $T \geq 0$ and $\phi \in \mathcal{B}_b(E)$. La mesure empirique induite par le modèle de Moran avec fonction associé Λ, approche la même QSD du processus de particules de Fleming – Viot avec un taux de mortalité $\kappa = \sup \Lambda - \Lambda$ (voir les Remarques 1, 3 et 4). Cela pose naturellement la question suivante : quel système de particules est le meilleur estimateur d'une QSD d'intérêt ? Nous répondons partiellement à cette question dans le Corollaire 4 ci-dessous, en montrant un moyen de minimiser l'erreur quadratique asymptotique lorsque $t, N \to \infty$, de $(m(\eta_t))_{t\geq 0}$ comme estimateur de $(\mu_t)_{t\geq 0}$. Cependant, l'étude du processus de particules de Moran avec une erreur quadratique minimale mérite d'être approfondie.

Dans le Chapitre 2, nous considérons le cas où l'espace d'état des types alléliques est fini et le taux de sélection est uniforme. Dans le contexte du processus de Fleming – Viot, cela revient à considérer un taux de mortalité constant κ. Bien qu'il s'agisse d'un modèle assez simplifié, nos résultats sont particulièrement robustes : dans le Théorème 5, nous obtenons une expression explicite pour les valeurs propres du générateur du système de particules $\mathcal{Q}_N$ en fonction des valeurs propres de la matrice du taux de mutation Q, qui est seulement supposée irréductible, mais pas nécessairement diagonalisable. De plus, nous étudions le processus où le taux de mutation satisfait une condition, qui est appelée *indépendance du parent*. Dans ce cas, le processus de Moran est réversible et on peut calculer explicitement les vecteurs propres de son générateur et sa distribution stationnaire. Ensuite, nous allons plus loin et prouvons l'existence du *phénomène de cutoff* chi-deux (Théorème 8) et en variation totale (Théorème 9).

Dans le Chapitre 3 nous considérons également le processus de Moran où l'espace d'état des types alléliques est fini et le taux de sélection est uniforme, et de plus la matrice du taux de mutation est le générateur d'une marche aléatoire asymétrique sur le graphe de cycles $\mathbb{Z}/K\mathbb{Z}$, avec $K \in \mathbb{N}$. Nous montrons que ce modèle possède une remarquable résolubilité, malgré le fait qu'il soit non réversible de distribution invariante non explicite. Nos principaux résultats incluent la propagation quantitative du chaos et l'ergodicité exponentielle avec des constantes explicites, ainsi que des formules pour les covariances à l'équilibre en termes de polynômes de Tchebychev. Nous obtenons également une limite explicite uniforme en temps pour la convergence de la proportion de particules dans chaque état lorsque le nombre de particules tend vers l'infini.

Nous résumons maintenant les principaux résultats de la thèse, qui sont présentés plus en détail et prouvés ensuite dans les Chapitres 1, 2 and 3.

Propagation du chaos dans le modèle de Moran

Nous présentons ici les principaux résultats obtenus dans le Chapitre 1, qui sont basés sur la prépublication [CC21].

Dans ce chapitre, nous considérons le problème de l'étude de la convergence de la distribution empirique induite par le modèle de Moran lorsque le nombre d'individus tend vers l'infini. Naturellement, nous supposons que la mesure empirique induite par le processus de particules en $t = 0$ converge vers la distribution initiale $\mu_0 \in \mathcal{M}_1(E)$ dans $\mathbb{L}^p$, pour tout $p \geq 1$. Plus précisément, il existe une constante $C > 0$ telle que

$$\sup_{\phi \in \mathcal{B}_1(E)} \mathbb{E}[|m(\eta_0)(\phi) - \mu_0(\phi)|^p] \leq \frac{C}{N^{p/2}}, \tag{9}$$

où $\mathcal{B}_1(E)$ est l'ensemble des fonctions réelles sur E telles que $\|\phi\| \leq 1$. En particulier, ceci est vérifié lorsqu'initialement toutes les particules sont échantillonnées indépendamment avec la même distribution $\mu_0 \in \mathcal{M}_1(E)$. Ensuite, nous nous concentrons sur le contrôle de la quantité suivante

$$\sup_{\phi\in\mathcal{B}_1(E)} \mathbb{E}\left[\sup_{t\in A} |m(\eta_t^{(N)})(\phi) - \mu_t(\phi)|^p\right],$$

pour tout $p \geq 1$, où $(\mu_t)_{t\geq 0}$ est la solution de (4) avec la condition initiale $\mu_0 \in \mathcal{M}_1(E)$ et $A \subset \mathbb{R}_+$. Il s'agit d'un contrôle de la convergence de $m(\eta_t^{(N)})(\phi)$ ver $\gamma_t(\phi)$, quand $N \to \infty$, comme dans (5). Nous considérons deux cas : lorsque $A = [0, T]$, avec $T \geq 0$, et donc pouvant être étendu à tout compact sur $\mathbb{R}_+$; ou lorsque $A = \mathbb{R}_+$. Dans ce dernier cas, nos résultats fournissent un contrôle uniforme en temps de la distance $\mathbb{L}^p$ entre $m(\eta_t^{(N)})(\phi)$ et $\mu_t(\phi)$, pour tout $\phi \in \mathcal{B}_1(E)$.

Propagation du chaos avec un taux de sélection général

Tout d'abord, nous considérons le modèle où les taux de sélection peuvent être écrits de la manière suivante

$$V_\mu(x, y) = \sum_{i\geq 1} V_i^{\mathrm{d}}(x)V_i^{\mathrm{b}}(y) + V_\mu^{\mathrm{s}}(x, y), \tag{10}$$

où V_i^{d} et V_i^{b} sont uniformément bornés dans E et $\mu \mapsto V_\mu^{\mathrm{s}}$ est une fonction continue et bornée de $(\mathcal{M}_1(E), \|\cdot\|_{\mathrm{TV}})$ dans $(\mathcal{B}_b(E\times E), \|\cdot\|)$. De plus, V_μ^{s} est symétrique pour tout $\mu \in \mathcal{M}_1(E)$. Voir les Hypothèses (G1) et (G2) dans la Section 1.1.1 pour une version détaillée des hypothèses que nous supposons dans ce cas. Sous ces hypothèses, nous montrons l'existence d'une convergence uniforme sur les compacts au sens de l'inégalité suivante : pour tout $T \geq 0$ et $p \geq 1$, il existe une constante $C_{p,T} > 0$, telle que

$$\sup_{\phi\in\mathcal{B}_1(E)} \mathbb{E}\left[\sup_{t\in[0,T]} |m(\eta_t^{(N)})(\phi) - \mu_t(\phi)|^p\right]^{1/p} \leq \frac{C_{p,T}}{\sqrt{N}}, \tag{11}$$

où $(\mu_t)_{t\geq 0}$ est la solution de (2) avec la condition initiale $\mu_0 \in \mathcal{M}_1(E)$. Voir le Théorème 1.1.2 pour l'énoncé complet de ce résultat.

Ce résultat est une généralisation pour les modèles de Moran multi-alléliques de la Proposition 3.1 dans [Cor17], où la convergence en probabilité uniforme sur les compacts est montrée pour un modèle de Moran bi-allélique. Cette vitesse de convergence peut également être reliée aux résultats existants qui assurent la convergence de la mesure empirique induite par un processus de particules de type Moran (ou Fleming−Viot) vers la loi d'un processus absorbant conditionné à la non-absorption. Voir par exemple : [DM00b, Prop. 3.5] [DPR11, Lemma 3.1], [Vil14, Thm. 2.2] et [CT16a, Thm. 1.3]. Voir aussi [BC15, Thm. 3.1 et Remarque 3.2], où la convergence presque sûre (et aussi la convergence complète[9] [Gut13, Déf. 1.6]) est prouvée lorsque l'espace d'état est fini. À notre connaissance, le contrôle que nous établissons dans (11) est le premier résultat assurant la convergence uniforme sur les compacts dans $\mathbb{L}^p$, pour $p \geq 1$, avec une vitesse de convergence d'ordre $1/\sqrt{N}$ pour les modèles de Moran multi-alléliques, et aussi pour les processus de particules de Fleming−Viot dans des espaces d'états discrets dénombrables, non nécessairement finis. L'idée qui soutient la preuve est proche des méthodes de [Rou06] : elle consiste à trouver une martingale indexée par l'intervalle $[0, T]$, dont la valeur terminale au temps T est précisément $m(\eta_T)(\phi) - \mu_T(\phi)$ plus un terme dont sa norme $\mathbb{L}^p$ peut être contrôlée, pour tout $\phi \in \mathcal{B}_d(E)$. Par la suite, le résultat final vient par l'utilisation astucieuse d'un argument de type Grönwall, de manière similaire à la preuve de la Proposition 1 dans [MS19].

[9] Voir *complete convergence* ou *universal convergence* en anglais.

Propagation du chaos uniforme dans le temps avec des taux de sélection additifs

Pour une expression plus spécifique des taux de sélection, nous pouvons montrer une limite uniforme en temps pour la convergence de $(m(\eta_t))_{t\geq 0}$ vers $(\mu_t)_{t\geq 0}$, lorsque $N \to \infty$. En particulier, nous supposons que (3) est vérifié, c'est-à-dire, nous supposons que les taux de sélection peuvent être écrits comme suit

$$V_\mu(x,y) = V_\mu^{\mathrm{d}}(x) + V_\mu^{\mathrm{b}}(y) + V_\mu^{\mathrm{s}}(x,y),$$

et qu'il existe $\Lambda \in \mathcal{B}_b(E)$, indépendant de μ, et il existe une fonction Λ et une constante C_μ, dépendant de μ tels que

$$\Lambda = V_\mu^{\mathrm{b}} - V_\mu^{\mathrm{d}} + C_\mu.$$

Voir l'Hypothèse (C1) dans le Chapitre 1 pour plus de détails.

Exemple 1 (Taux de sélection indépendants de μ)**.** Lorsque les taux de sélection ne dépendent pas de μ, l'expression des taux de sélection se réduit à l'existence de $V^{\mathrm{d}}, V^{\mathrm{b}} \in \mathcal{B}_d(E)$ et d'un V^{s} symétrique dans $\mathcal{B}_d(E \times E)$ tels que

$$V(x,y) = V^{\mathrm{d}}(x) + V^{\mathrm{b}}(y) + V^{\mathrm{s}}(x,y).$$

Soit $\Lambda \in \mathcal{B}_d(E)$ une fonction fixée. Des exemples typiques de fonctions V^{b} et V^{d} vérifiant cette condition sont les suivants

$$V^{\mathrm{b}} = (\Lambda - c)^+ \quad \text{et} \quad V^{\mathrm{d}} = (\Lambda - c)^-,$$

pour une constante fixe $c \in \mathbb{R}$, où nous utilisons la notation standard

$$(x)^+ := \max\{x, 0\} \quad \text{et} \quad (x)^- := -\min\{x, 0\}.$$

Ce sont en fait les taux de sélection considérés par Angeli et al. [AGJ21, §3.3] dans le contexte des algorithmes de clonage[10]. De plus, le cas $c = 0$ est considéré dans l'Exemple 3.1-(2) dans [Rou06]. Notons que dans ce cas, l'expression dans (10) est également vérifiée.

D'un point de vue biologique, le paramètre $c \in \mathbb{R}$ peut être considéré comme un paramètre de adaptabilité. Pour simplifier, supposons que V^{s} est nul, et notons par $\xi_t^{(i)}$ le type du i-ème individu, pour $1 \leq i \leq N$, au temps $t \geq 0$. Donc, si $\Lambda(\xi_t^{(i)}) \leq c$, le i-ème individu meurt et un autre choisi au hasard se reproduit avec le taux $(\Lambda(\xi_t^{(i)}) - c)^-$. Sinon, un individu choisi au hasard meurt et le i-ème individu se reproduit avec le taux $(\Lambda(\xi_t^{(i)}) - c)^+$.

Un autre exemple particulièrement intéressant est celui où $V^{\mathrm{b}} = 0$. Notons que le processus de Moran avec ces taux de sélection est en fait un système de particules de Fleming – Viot.

Exemple 2 (Taux de sélection dépendant de μ)**.** Considérons une fonction fixée $\Lambda \in \mathcal{B}_b(E)$. Des exemples typiques de fonctions V_μ^{b} et V_μ^{d} sont :

$$V_\mu^{\mathrm{b}} = (\Lambda - \mu(\Lambda))^+ \quad \text{et} \quad V_\mu^{\mathrm{d}} = (\Lambda - \mu(\Lambda))^-.$$

Ce sont les taux de sélection considérés dans [Del04, §1.5.2, p. 35], voir aussi Exemple 3.1-(3) in [Rou06].

Dans ce cas, l'interprétation biologique de $\mu(\Lambda)$ est similaire à celle du paramètre c dans l'Exemple 1. En effet, le paramètre de adaptabilité varie dans le temps en fonction de la configuration de la population.

[10] *Cloning algorithms* en anglais.

Nous sommes maintenant en mesure d'énoncer nos principaux résultats pour le modèle de Moran multi-allélique avec sélection additive.

Théorème 2 (Propagation du chaos uniforme dans le temps)**.** *Supposons que*

- *l'expression* (3) *pour le taux de sélection,*
- *l'hypothèse* (9) *sour la condition initiale, et*
- *la convergence uniforme établie par* (8) *pour* $(\mu_t)_{t\geq 0}$

sont verifiées. Alors, pour tout $p \geq 1$*, il existe une constante positive* C_p*, telle que*

$$\sup_{\phi\in\mathcal{B}_1(E)} \sup_{t\geq 0} \mathbb{E}\left[|m(\eta_t^{(N)})(\phi) - \mu_t(\phi)|^p\right]^{1/p} \leq \frac{C_p}{\sqrt{N}}.$$

Voir le Théorème 1.1.4 dans le Chapitre 1 pour une version détaillée de ce résultat et sa preuve.

Pour un N donné, si le processus $\left(\eta_t^{(N)}\right)_{t\geq 0}$ généré par $\mathcal{Q}_N$ admet une distribution stationnaire ν_N, alors, sous l'hypothèse du Théorème 2 nous obtenons

$$\sup_{\phi\in\mathcal{B}_1(E)} \mathbb{E}_{\nu_N}\left[|m(\eta_\infty)(\phi) - \mu_\infty(\phi)|^p\right]^{1/p} \leq \frac{C_p}{\sqrt{N}},$$

pour tout $p \geq 1$.

Obtenir une borne uniforme en temps comme celle fournie par le Théorème 2 est un problème difficile et ce type de résultats est peu fréquent dans la littérature. Del Moral et Guionnet dans [DG01, Thm. 3.1] ont prouvé un résultat similaire pour un modèle analogue mais en temps discret, où le potentiel est supposé uniformément borné, et avec une borne inférieure strictement positive. De plus, leur borne supérieure pour la vitesse de convergence en [DG01, Thm. 3.1] est d'ordre $1/N^\alpha$, avec $\alpha < 1/2$. Rousset [Rou06, Thm. 4.1] a prouvé une borne uniforme en temps dans $\mathbb{L}^p$ avec la même vitesse de convergence que dans notre résultat. Cependant, le modèle étudié par Rousset est dans le cadre d'un espace d'état continu et le processus de diffusion dirigeant le processus de mutation est supposé réversible. De même, Angeli et al. [AGJ21, Thm. 3.2] ont obtenu un résultat équivalent pour les processus de saut sur des espaces localement compacts dans le contexte des algorithmes de clonage, pour $p \geq 2$. Notre modèle est différent, puisque nous considérons le cas où l'espace d'état est discret, et pas nécessairement compact. De plus, dans (3) nous permettons aux taux de sélection de dépendre de la mesure de probabilité empirique induite par le système de particules, dans le même esprit que [Rou06]. Néanmoins, nos méthodes sont similaires à celles de Rousset [Rou06] et d'Angeli et al. [AGJ21] (voir aussi [DM00b, §3.3.1]) : elles consistent à trouver une martingale indexée par l'intervalle $[0, T]$, dont la valeur terminale au temps T est précisément $m(\eta_T)(\phi) - \mu_T(\phi)$ plus un terme dont sa norme $\mathbb{L}^p$ peut être contrôlée, pour tout $\phi \in \mathcal{B}_b(E)$. Par la suite, le résultat final vient d'un contrôle de la variation quadratique de la martingale et d'un principe de recurrence.

Nous obtenons également le résultat suivant : sous les hypothèses du Théorème 2, pour chaque $p \geq 1$, il existe une constante $C_p > 0$ telle que

$$\sup_{t\geq 0} \mathbb{E}\left[\left(\|m(\eta_t^{(N)}) - \mu_t\|_{\mathrm{w}}\right)^p\right]^{1/p} \leq \frac{C}{\sqrt{N}}, \quad \text{où} \quad \|\mu_1 - \mu_2\|_{\mathrm{w}} := \sum_{n\in M} 2^{-n}|\mu_1(x_n) - \mu_2(x_n)|,$$

pour une énumération (arbitraire mais fixé) $(x_n)_{n\in M}$ des éléments de E, avec $M \subset \mathbb{N}$.

Remarque 5 (Convergence presque sûre). L'inégalité précédente, pour $p = 4$, assure le convergence :

$$\sum_{N \geq 2} \mathbb{P}\left[\|m(\eta_t^{(N)}) - \mu_t\|_{\mathrm{w}} > \epsilon\right] < \infty,$$

pour tout $\epsilon > 0$. En conséquence, avec un argument du type Borel – Cantelli cela implique la convergence $m(\eta_T^{(N)}) \xrightarrow{\text{p.s.}} \mu_T$ quand $N \to \infty$ au sens faible, où p.s. note la convergence presque sûre. De plus, cette convergence ne dépend pas de l'espace où les variables aléatoires sont couplées.

Désignons par $\bar{m}(\eta_t)$ la mesure de probabilité empirique moyenne induite par η_t définie par

$$\bar{m}(\eta_t) := \sum_{x \in E} \mathbb{E}\left[\frac{\eta_t(x)}{N}\right] \delta_x \in \mathcal{M}_1(E).$$

De plus, désignons par $\xi_t^{(i)}$ la variable aléatoire qui represente le type du i-ème individu au temps $t \geq 0$, avec $i \in \{1, 2, \ldots, N\}$ et par $\mathrm{Law}(\xi_t^{(i)})$ la loi de cette variable aléatoire. Notons par $(\mathrm{e}^{tQ})_{t \geq 0}$ le semigroupe associé au processus de mutation généré par Q. Alors, sous les hypothèses de l'énoncé du Théorème 2, nous pouvons garantir l'existence d'une constante positive C telle que

$$\sup_{t \geq 0} \left\|\bar{m}(\eta_t^{(N)}) - \mu_t\right\|_{\mathrm{TV}} \leq \frac{C}{N}.$$

De plus, si la distribution initiale des N particules est échangeable, on a

$$\sup_{t \geq 0} \left\|\mathrm{Law}(\xi_t^i) - \mu_t\right\|_{\mathrm{TV}} \leq \frac{C}{N}.$$

On s'attend à ce que, lorsque les taux de sélection sont constants, la mesure générée par le système de particules soit un estimateur sans biais de la loi de la chaîne de Markov générée par Q, au sens où

$$\bar{m}(\eta_t^{(N)}) = \bar{m}(\eta_0)\mathrm{e}^{tQ}, \text{ pour tout } t \geq 0.$$

Voir par exemple [CT16b] et les résultats du Chapitre 3. Nous prouvons dans le Corollaire 1.3.6 que ce résultat est plus généralement valable lorsque $V = V^{\mathrm{s}}$ est symétrique.

Nos deux derniers résultats du Chapitre 1 sont destinés à l'étude de l'erreur quadratique asymptotique de l'approximation de μ_∞ par $m(\eta_T^{(N)})$ lorsque $T, N \to \infty$. Ces résultats sont très importants lorsque le processus de Moran est utilisé pour approcher une distribution quasi-stationnaire. Tout d'abord, nous fournissons une expression explicite pour l'expression asymptotique de $N\mathbb{E}\left[\left(m(\eta_T^{(N)}) - \mu_\infty(\phi)\right)^2\right]$. Définissons les erreurs quadratiques asymptotiques comme suit

$$\sigma_T^2(\phi) := \lim_{N \to \infty} N\mathbb{E}\left[\left(m(\eta_T)(\phi) - \mu_T(\phi)\right)^2\right] \quad \text{and} \quad \sigma_\infty^2(\phi) := \lim_{T \to \infty} \sigma_T^2(\phi),$$

pour toute fonction uniformément bornée ϕ dans E.

Tout de suite, nous prouvons la normalité asymptotique du biais et nous fournissons des expressions explicites pour $\sigma_T^2(\phi)$ et $\sigma_\infty^2(\phi)$. Ensuite, nous utilisons cette expression pour montrer comment définir un autre processus de Moran approchant la même distribution μ_∞, et avec une erreur quadratique asymptotique inférieure ou égale.

Définissons

$$S_\mu(\phi) := \sum_{x,y \in E} (\phi(x) - \phi(y))^2 V_\mu^{\mathrm{s}}(x,y)\mu(x)\mu(y), \tag{12}$$

pour tous $\phi \in \mathcal{B}_b(E)$ et $\mu \in \mathcal{M}_1(E)$. Nous obtenons ainsi le résultat suivant :

Théorème 3 (Normalité asymptotique)**.** *Supposons qu'initialement les N particules sont échantillonnées indépendamment selon $\mu_0 \in \mathcal{M}_1(E)$, et que les hypothèses du Théorème 2 sont vérifiées. Alors, pour toute fonction $\phi \in \mathcal{B}_b(E)$ et $T \geq 0$, on a que $\sqrt{N}(m(\eta_T)(\phi) - \mu_T(\phi))$ converge en loi vers une variable aléatoire centrée gaussienne de variance $\sigma_T^2(\phi)$, lorsque N tend vers l'infini. De plus,*

$$\sigma_\infty^2(\phi) = \mathrm{Var}_{\mu_\infty}(\phi) + \int_0^\infty \mathrm{e}^{-2\lambda s} S_{\mu_\infty}(P_s^\Lambda(\bar{\phi}_\infty))\mathrm{d}s + 2\int_0^\infty \mathrm{e}^{-2\lambda s}\mu_\infty\Big(P_s^\Lambda(\bar{\phi}_\infty)^2\Big(V_{\mu_\infty}^{\mathrm{b}} + \mu_\infty(V_{\mu_\infty}^{\mathrm{d}})\Big)\Big)\mathrm{d}s,$$

où $\mathrm{Var}_{\mu_\infty}$ représente la variance par rapport à μ_∞, $\bar{\phi}_\infty := \phi - \mu_\infty T(\phi)$, λ est la valeur propre dans l'énoncé du Lemme 1 et S_μ est défini dans (12).

Voir le Théorème 1.1.8 pour un énoncé détaillé de ce résultat, qui inclut une expression explicite pour $\sigma_T^2(\phi)$, pour tout $T \geq 0$. Notez que les deux intégrales dans l'expression de $\sigma_\infty^2(\phi)$ dans le Théorème 3 convergent comme conséquence du Lemme 1.

Lorsque V_μ^{b} et V_μ^{s} sont nuls et donc $\Lambda = -V^{\mathrm{d}} \leq 0$, on retrouve

$$\sigma_\infty^2(\phi) = \mathrm{Var}_{\mu_\infty}(\phi) - 2\lambda\int_0^\infty \mathrm{e}^{-2\lambda s}\,\mathrm{Var}_{\mu_\infty}\Big(P_s^\Lambda(\phi)\Big)\,\mathrm{d}s,$$

Lorsque le processus $(\eta_t)_{t\geq 0}$ est ergodique et converge en loi vers une certaine variable aléatoire η_∞, lorsque $t \to \infty$, le Théorème 3 établit que $\sqrt{N}(m(\eta_\infty)(\phi) - \mu_\infty(\phi))$ converge vers une loi gaussienne centrée de variance $\sigma_\infty^2(\phi)$, lorsque $N \to \infty$. En effet, rappelons qu'une suite gaussienne converge en loi si ses deux premiers moments convergent. En particulier, nous retrouvons (et étendons) le résultat récent de Lelièvre et al. [LPR18, Thm. 2.4] dans la cadre des espaces d'états finis. Remarquons que la constante négative λ dans l'expression précédente est l'opposée de celle dans [LPR18, Thm. 2.4].

Notons que les trois termes dans l'expression de $\sigma_\infty^2(\phi)$ sont positifs. De plus, la limite $(\mu_t)_{t\geq 0}$ est invariante par le choix de la composante symétrique V_μ^{s} dans (3), comme commenté dans la Remarque 1. En conséquence, pour un taux de sélection donné V_μ, nous pouvons obtenir un autre processus de Moran approchant la même distribution limite en prenant le taux de sélection $V_\mu - \Sigma_\mu \geq 0$, où Σ_μ est une fonction symétrique dans $\mathcal{B}_b(E \times E)$.

Corollaire 4 (Processus de Moran avec une réduction de l'erreur quadratique asymptotique)**.** *Supposons qu'initialement les N particules sont échantillonnées indépendamment selon $\mu_0 \in \mathcal{M}_1(E)$, et que les hypothèses du Théorème 2 sont vérifiées. Soit $(\eta_t)_{t\geq 0}$ et $(\eta_t^\star)_{t\geq 0}$ les processus de Moran avec les mêmes taux de mutation, et taux de sélection donnés par V_μ et $V_\mu - \Sigma_\mu$, respectivement, où*

$$\Sigma_\mu(x,y) := \min\Big\{V_\mu^{\mathrm{d}}(x), V_\mu^{\mathrm{b}}(x)\Big\}\mathbf{1}_{\{x\}} + \min\Big\{V_\mu^{\mathrm{d}}(y), V_\mu^{\mathrm{b}}(y)\Big\}\mathbf{1}_{\{y\}} + V_\mu^{\mathrm{s}}(x,y),$$

où $\mathbf{1}_A$ représente la fonction indicatrice de $A \subset E$. Alors,

$$\lim_{T\to\infty}\lim_{N\to\infty} N\mathbb{E}\left[\left(m(\eta_T^\star)(\phi) - \mu_\infty(\phi)\right)^2\right] \leq \lim_{T\to\infty}\lim_{N\to\infty} N\mathbb{E}\left[\left(m(\eta_T)(\phi) - \mu_\infty(\phi)\right)^2\right].$$

Voir Corollaire 1.1.9 pour une version plus détaillé de ce résultat. Notons que le taux de sélection $V_\mu - \Sigma_\mu$ dans l'énoncé du Corollaire 4 vérifie (3).

Spectre et ergodicité du modèle de Moran neutre

Nous présentons ici les principaux résultats obtenus dans la prépublication [Cor21b], qui sont présentés en détail dans le Chapitre 2.

Dans cette section, nous considérons que $E = [K]$, où $[K] := \{1, 2, \ldots, K\}$. La définition de l'espace d'état du Moran multi-allélique peut être réduite au N simplexe discret de dimension K :

$$\mathcal{E}_{K,N} := \left\{\eta \in [N]_0^K : \ |\eta| = N\right\},$$

où $[N]_0 := \{0, 1, \ldots, N\}$ et $|\cdot|$ représente la somme des éléments d'un vecteur. L'ensemble $\mathcal{E}_{K,N}$ a pour cardinal $\binom{K-1+N}{N}$.

Selon la notation introduite précédemment, la matrice des taux de mutation dans ce cas est la matrice : $Q = (Q_{i,j})_{i,j=1}^K$. De plus, le processus est neutre, c'est-à-dire que les taux de sélection sont constants et égaux à $p \geq 0$. Notons par $\mathcal{Q}_{N,p}$ le générateur infinitésimal du processus de Moran multi-allélique, neutre, qui agit sur une fonction réelle f sur $\mathcal{E}_{K,N}$ comme suit :

$$(\mathcal{Q}_{N,p} f)(\eta) := \sum_{i,j \in [K]} \eta(i) \left(Q_{i,j} + \frac{p}{N}\eta(j)\right)[f(\eta - \mathbf{e}_i + \mathbf{e}_j) - f(\eta)], \tag{13}$$

pour tout $\eta \in \mathcal{E}_{K,N}$, où $\mathbf{e}_k$ est le k-ième vecteur canonique de $\mathbb{R}^K$.

En d'autres termes, $\mathcal{Q}_{N,p}$ génére un processus de N individus, où chaque individu possède un des K types d'allèles possibles et où le type de l'individu change suite à deux processus : un processus de mutation où les individus mutent indépendamment les uns des autres et un processus de reproduction de type Moran, où les individus interagissent. Les N individus mutent indépendamment du type $i \in [K]$ au type $j \in [K] \setminus \{i\}$ avec un taux $Q_{i,j}$. De plus, avec un taux uniforme $p \geq 0$, un des N individus est uniformément choisi pour être retiré de la population et un autre, également choisi au hasard, est dupliqué. Notons que les transitions d'un individu dues à une reproduction ne sont pas indépendantes de la position des autres individus. Comme dans le modèle original, introduit par Moran [Mor58], le même individu retiré de la population peut être dupliqué, dans ce cas l'état du système ne change pas. Dans le cas où l'individu retiré ne peut pas être dupliqué, le facteur $\frac{p}{N}$ dans (13) doit être remplacé par $\frac{p}{N-1}$. Le processus généré par $\mathcal{Q}_{N,p}$ peut également être considéré comme un système de particules de Fleming – Viot, mais avec un taux de mortalité constant, c'est-à-dire avec κ constant.

Notre résultat principal dans cette section est une description complète du spectre de $\mathcal{Q}_{N,p}$, comme exprimé dans le théorème suivant.

Théorème 5 (Spectre de $\mathcal{Q}_{N,p}$)**.** *Supposons $K \geq 2$, $N \geq 1$ et $p \in [0, \infty)$. Notons par λ_k, $k \in [K-1]$, les $K-1$ racines non nulles, en comptant les multiplicités algébriques, du polynôme caractéristique de Q. Pour tout $\eta \in \bigcup_{L=1}^{N} \mathcal{E}_{K-1,L}$, définissons*

$$\lambda_{\eta,p} := \sum_{k=1}^{K-1} \eta(k)\lambda_k - \frac{p}{N}|\eta|(|\eta| - 1).$$

Alors, les valeurs propres de $\mathcal{Q}_{N,p}$, en comptant les multiplicités algébriques, sont 0 *et $\lambda_{\eta,p}$, pour $\eta \in \bigcup_{L=1}^{N} \mathcal{E}_{K-1,L}$.*

La preuve du Théorème 5 est donnée dans la Section 2.3.3. Notons que le spectre de $Q_{N,p}$ est donné en fonction du spectre de Q, et que le résultat est vrai même si Q n'est pas diagonalisable. Nos résultats incluent également une étude des vecteurs propres de $\mathcal{Q}_{N,p}$. En particulier, nous sommes en mesure de fournir des expressions explicites pour les vecteurs propres de $\mathcal{Q}_{N,p}$ associés aux valeurs propres de plus petit module. Les expressions que nous fournissons sont basées sur les vecteurs propres de Q. Voir le Théorème 2.1.3 dans le Chapitre 2 pour plus de détails.

Remarque 6 (Monotonie en N du spectre de $\mathcal{Q}_N$ $(p = 0)$). Le Théorème 5 implique que le spectre de $\mathcal{Q}_N$, lorsque $p = 0$, pour une valeur fixé de K, est une fonction croissante de N au sens de l'inclusion des ensembles.

Ergodicité du processus de Moran multi-allélique neutre

La relation entre les propriétés spectrales de $\mathcal{Q}_{N,p}$ et Q peut être utilisée pour estimer la vitesse de convergence vers la stationnarité du processus de Moran.

Nous nous intéressons à la relation entre le spectre d'une matrice de taux et la convergence vers la stationnarité du processus de Markov qu'elle génère. Définissons la *distance en variation totale maximale* vers la stationnarité du processus piloté par un générateur infinitésimal L d'une chaîne de Markov sur un espace discret Ω, noté $\mathrm{D}^L_{\mathrm{TV}}$, par

$$\mathrm{D}^L_{\mathrm{TV}}(t) := \max_{\mu \in \mathcal{M}_1(\Omega)} \left\| \mu \, \mathrm{e}^{tL} - \pi \right\|_{\mathrm{TV}}, \tag{14}$$

où le maximum est pris sur toutes les distributions initiales possibles sur Ω. En utilisant la convexité de la norme en variation totale, nous pouvons prouver que $\mathrm{D}^L_{\mathrm{TV}}(t) = \frac{1}{2}\|\mathrm{e}^{tL} - \Pi\|_\infty$, où Π représente la matrice dont chaque ligne est égale à π, et on note par $\|\cdot\|_\infty$ la norme infinie des matrices (cf. [LP17, Ch. 4]).

En conséquence du Théorème 5 et d'autres résultats du Chapitre 2, la deuxième plus grande valeur propre en module (SLEM) de $\mathrm{e}^{t\mathcal{Q}_{N,p}}$ est égale à celle de e^{tQ}, pour tout $t \geq 0$. La SLEM du générateur d'un chaîne de Markov finie est utile pour étudier la convergence asymptotique en variation totale. Par conséquent, comme application de ce Théorème 5, nous étudions l'ergodicité du processus piloté par $\mathcal{Q}_{N,p}$ en variation totale en utilisant les propriétés spectrales de Q.

Pour une fonction positive réelle f, on note par $\mathcal{O}(f)$ une autre fonction positive réelle telle que $C_1 f(t) \leq \mathcal{O}(f)(t) \leq C_2 f(t)$, pour deux constantes $0 < C_1 \leq C_2 < \infty$ et pour tout $t \geq T$, pour $T > 0$ suffisamment grand.

Corollaire 6 (Ergodicité exponentielle asymptotique en variation totale)**.** *Supposons que la* SLEM *de* e^{tQ} *est égale à* $\mathrm{e}^{-\rho t}$ *et que* $s \in \mathbb{N}$ *est la plus grande multiplicité dans le polynôme minimal de* e^{tQ} *de toutes les valeurs propres de module* $\mathrm{e}^{-\rho t}$*. Alors,*

$$\mathrm{D}^{\mathcal{Q}_{N,p}}_{\mathrm{TV}}(t) = \mathcal{O}\Big(\mathrm{D}^{Q}_{\mathrm{TV}}(t)\Big) = \mathcal{O}(t^{s-1}\mathrm{e}^{-\rho t}).$$

Le Corollaire 6 est prouvé dans la Section 2.4.

L'expression asymptotique dans le Corollaire 6 cache la relation entre le temps de mélange de la chaîne de Markov et le nombre d'individus dans la population et le paramètre de sélection p. Cependant, si Q a une valeur propre réelle $-\lambda < 0$, nous pouvons aller plus loin et trouver la borne inférieure suivante pour la convergence en variation totale vers la stationnarité au temps $\frac{\ln N - c}{2\lambda}$, pour tout $c \geq 0$.

Théorème 7 (Borne inférieure pour la convergence en variation totale)**.** *Supposons que* $K \geq 2$*,* $N \geq 2$ *et* $p \in [0, \infty)$ *et que* $-\lambda < 0$ *soit une valeur propre de* Q*. Alors*

$$\mathrm{D}^{\mathcal{Q}_{N,p}}_{\mathrm{TV}}\left(\frac{\ln N - c}{2\lambda}\right) \geq 1 - \kappa \mathrm{e}^{-c}.$$

Dans le Théorème 2.1.5, nous énonçons une version plus détaillée du Théorème 7, qui est prouvée dans la Section 2.4.1.

La borne inférieure fournie par le Théorème 7 assure que le temps de mélange du modèle de Moran multi-allélique neutre est au moins de l'ordre de $\ln N/2\lambda$. Nos résultats ne nous permettent pas de trouver une borne supérieure assurant l'existence d'un *phénomène de cutoff*. Une étude plus détaillée doit être faite dans cette direction. Cependant, pour le schéma de mutation indépendant des parents, une analyse plus détaillée peut être effectuée pour prouver l'existence d'un phénomène de cutoff dans les distances chi-deux et en variation totale, comme nous le verrons ensuite.

Le modèle de Moran multi-allélique neutre avec mutation indépendante du parent

Considérons la matrice de taux de mutation suivante :

$$Q_{\boldsymbol{\mu}} := \begin{pmatrix} -|\boldsymbol{\mu}|+\mu_1 & \mu_2 & \mu_3 & \dots & \mu_K \\ \mu_1 & -|\boldsymbol{\mu}|+\mu_2 & \mu_3 & \dots & \mu_K \\ \mu_1 & \mu_2 & -|\boldsymbol{\mu}|+\mu_3 & \dots & \mu_K \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu_1 & \mu_2 & \mu_3 & \dots & -|\boldsymbol{\mu}|+\mu_K \end{pmatrix}, \tag{15}$$

où $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_K) \in (0,\infty)^K$ et $|\boldsymbol{\mu}|$ représente la somme des éléments de $\boldsymbol{\mu}$. Définissons

$$(\mathcal{L}_{N,p} f)(\eta) := \sum_{i,j=1}^{K} \eta(i) \left(\mu_j + \frac{p}{N}\eta(j)\right) [f(\eta - \mathbf{e}_i + \mathbf{e}_j) - f(\eta)],$$

pour tout f sur $\mathcal{E}_{K,N}$ et tout $\eta \in \mathcal{E}_{K,N}$, le générateur infinitésimal du processus de Moran multi-allélique neutre avec matrice de taux de mutation $Q_{\boldsymbol{\mu}}$. Le processus généré par $\mathcal{L}_{N,p}$ est un cas particulier du processus de Moran multi-allélique neutre considéré précédemment, mais à la différence que le taux de mutation ne dépend que du type du nouvel individu, c'est-à-dire que le processus de mutation change chaque individu de type i en type j au taux μ_j, pour tous les $i, j \in [K]$. Il s'agit du processus de Moran multi-allélique neutre avec mutation *indépendante du parent* (cf. [Eth11]). Notez que le graphe associé à $Q_{\boldsymbol{\mu}}$ est le graphe complet, comme le montre la Figure 1

En utilisant le Théorème 5 nous pouvons exprimer le spectre de $\mathcal{L}_{N,p}$. En effet, pour $K \geq 2$, $N \geq 2$ et $p \geq 0$, le générateur infinitésimal $\mathcal{L}_{N,p}$ est diagonalisable avec des valeurs propres $\lambda_{n,p}$ de multiplicité $\binom{K+n-2}{n}$, où

$$\lambda_{n,p} := -|\boldsymbol{\mu}|n - \frac{p}{N}n(n-1),$$

pour $n \in [N]_0$. En particulier, le trou spectral de $\mathcal{L}_{N,p}$ est $\rho = |\boldsymbol{\mu}|$.

Remarque 7 (Modèle de graphe complet). Le *modèle de graphe complet* étudié par Cloez et Thai [CT16a] dans le contexte des processus de particules de Fleming−Viot est un cas particulier du processus réversible généré par $Q_{\boldsymbol{\mu}}$ ci-dessus, lorsque $\mu_j = \frac{1}{K}$, pour tout $j \in [K]$. Dans ce cas, les valeurs propres de la matrice de mutation sont $\beta_0 = 0$ et $\beta_1 = -1$, cette dernière avec multiplicité $K-1$. En particulier, l'expression précédente pour le spectre de $\mathcal{L}_{N,p}$ améliore le Lemme 2.14 en [CT16a].

Pour $K \geq 2$, $N \geq 2$ et $p > 0$, nous prouvons dans le Chapitre 2 que le processus généré par $\mathcal{Q}_{N,p}$ est réversible si et seulement si la matrice de mutation a la forme $Q_{\boldsymbol{\mu}}$ comme dans (15), pour un vecteur quelconque $\boldsymbol{\mu}$, et par conséquent $\mathcal{Q}_{N,p}$ peut s'écrire comme $\mathcal{L}_{N,p}$. La Section 2.5 est consacrée à l'étude des propriétés spectrales de $\mathcal{L}_{N,p}$, pour $p \geq 0$, et à ses implications dans l'étude de la convergence vers la stationnarité. Nos résultats dans cette section incluent une description complète de l'ensemble des valeurs propres et des fonctions propres de $\mathcal{L}_{N,p}$, ainsi que de sa fonction de transition. Les fonctions propres de $\mathcal{L}_{N,p}$, $p > 0$, sont explicitement données en termes de *polynômes de Hahn multivariés*, qui sont orthogonaux par rapport à la distribution

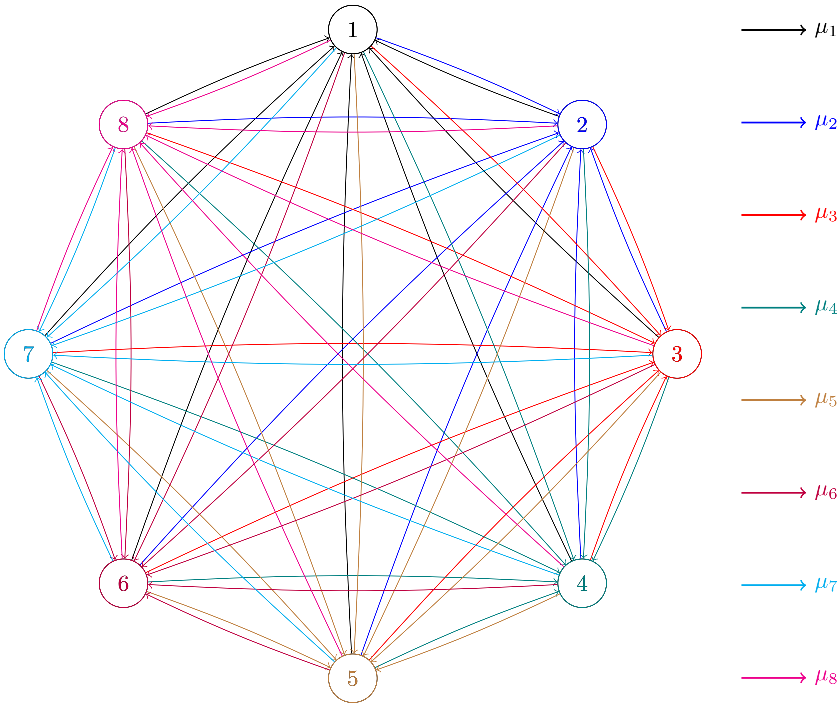


Figure 1: Graphe complet induit par le générateur infinitésimal $Q_{\boldsymbol{\mu}}$ (15) dans $[K]$. Dans cet exemple $K = 8$, la couleur de chaque flèche représente les taux μ_i, pour $1 \leq i \leq K$, qui sont en principe différents. Notez que la couleur des flèches ne dépend que du site d'arrivée.

stationnaire, qui est un mélange Dirichlet de distributions multinomiales (cf.[KM75; KZ09]). D'autre part, les fonctions propres de $\mathcal{L}_N$, pour $p = 0$, sont explicitement données en termes de *polynômes de Krawtchouk multivariés*, qui sont orthogonaux par rapport à la distribution stationnaire, qui est une distribution multinomiale (cf. [KM65; ZL09; DG14]).

Phénomène de cutoff

Le *phénomène de cutoff* constitue un riche sujet de recherche sur les chaînes de Markov depuis son introduction par les travaux de Aldous, Diaconis et Shahshahani dans les années 1980 [DS81; Ald83; AD86]). Une chaîne de Markov présente un cutoff si elle présente une transition brusque dans sa convergence vers la stationnarité. Parmi les notions de convergence les plus utilisées sont, comme nous le considérons ici, la distance en variation totale et la distance (ou divergence) chi-deux. Une bonne introduction à ce sujet peut être trouvée dans le livre de Levin et Peres [LP17, Ch. 18] et dans les recherches de Chen, Saloff-Coste et al. [Sal97; Che06; CS08; CS10; CHS17]. Voir aussi le cours de Salez [Sal21].

Un scénario typique pour l'existence d'un cutoff est une chaîne de Markov avec un haut degré de symétrie. Par conséquent, le phénomène de cutoff a été étudié en profondeur pour le mouvement de N particules indépendantes sur K sites. Ycart [Yca99] a étudié le cutoff en variation totale pour N particules indépendantes pilotés par une matrice de taux diagonalisable. Plus tard, Barrera et al. [BLY06] et Connor [Con10] ont étudié le cutoff de ce modèle pour d'autres notions de distance. Voir aussi [Lac15] [LP17, Ch. 20], [CHS17] et [CK18] pour des études plus récentes sur le cutoff sur ces modèles. Le modèle de Moran que nous considérons ici préserve le haut niveau de symétrie de ces chaînes de Markov, mais les mouvements des particules ne sont pas indépendants. En effet, les particules interagissent selon un processus de reproduction qui favorise les sauts vers les sites ayant de plus grandes proportions d'individus.

Avant de définir formellement le phénomène de cutoff, rappelons la *divergence chi-deux* (parfois appelée “distance”), qui apparaît naturellement dans le contexte des chaînes de Markov réversibles (cf. [LP17, Ch. 12 et 20], [DP17, Ch. 8], [Bré20, Ch. 9.1]). Soit Ω un ensemble discret. La divergence chi-deux de $\mu_2 \in \mathcal{M}_1(\Omega)$ par rapport à une distribution cible $\mu_1 \in \mathcal{M}_1(\Omega)$ est définie par

$$\chi^2(\mu_2 \mid \mu_1) := \sum_{\omega\in\Omega} \frac{[\mu_2(\omega) - \mu_1(\omega)]^2}{\mu_1(\omega)} = \|\mu_2 - \mu_1\|^2_{\frac{1}{\mu_1}},$$

où $\|\cdot\|_{\frac{1}{\mu_1}}$ représente la norme dans $l^2(\mathbb{R}^\Omega, \frac{1}{\mu_1})$, et $\frac{1}{\mu_1}$ est la mesure $\omega \mapsto 1/\mu_1(\omega)$.

La divergence du chi-deux n’est pas une métrique, mais une mesure de la différence entre deux distributions de probabilité. Notons que la divergence chi-deux, ainsi que la distance en variation totale, sont des cas particuliers des fonctions appelées $f-$divergences, qui mesurent la “différence” entre deux distributions de probabilité. Dans ce contexte, $\chi^2(\mu_2 \mid \mu_1)$ est également connue sous le nom de *divergence chi-deux de Pearson*.

Définissons les fonctions $\mathrm{d_{TV}}(\cdot, \eta)$ et $\chi^2(\cdot, \eta)$, comme suit

$$\mathrm{d_{TV}}(t, \eta) := \|\delta_\eta \mathrm{e}^{t\mathcal{L}_{N,p}} - \nu_{N,p}\|_{\mathrm{TV}} = \frac{1}{2} \sum_{\xi\in\mathcal{E}_{K,N}} \left| \left(\mathrm{e}^{t\mathcal{L}_{N,p}}\delta_\xi\right)(\eta) - \nu_{N,p}(\xi)\right|,$$

$$\chi^2(t, \eta) := \chi^2(\delta_\eta \mathrm{e}^{t\mathcal{L}_{N,p}} \mid \nu_{N,p}) = \sum_{\xi\in\mathcal{E}_{K,N}} \frac{\left[\left(\mathrm{e}^{t\mathcal{L}_{N,p}}\delta_\xi\right)(\eta) - \nu_{N,p}(\xi)\right]^2}{\nu_{N,p}(\xi)},$$

pour tout $t \geq 0$.

Les fonctions $\mathrm{d_{TV}}(\cdot, \eta)$ et $\chi^2(\cdot, \eta)$ sont donc des mesures de la convergence vers la loi stationnaire du processus généré par $\mathcal{L}_{N,p}$ au temps t et avec une configuration initiale $\eta \in \mathcal{E}_{K,N}$. Conformément à [Zho08; KZ09], nous les appelons respectivement *distance en variation totale* et *distance chi-deux* à la distribution stationnaire.

Lorsque le nombre d’individus varie, nous obtenons une famille infinie de chaînes de Markov finies à temps continu $\{(\mathcal{E}_{K,N}, \mathcal{L}_{N,p}, \nu_{N,p}), N \geq 2\}$. Pour chaque $N \geq 2$, notons par $\chi^2(t, N\mathbf{e}_k)$ (resp. $\mathrm{d_{TV}}(t, N\mathbf{e}_k)$) la distance chi-deux (resp. distance en variation totale) à la loi stationnaire du processus généré par $\mathcal{L}_{N,p}$ au temps t, lorsque la distribution initiale est concentrée en $N\mathbf{e}_k \in \mathcal{E}_{K,N}$. Notons que $\chi^2(0, N\mathbf{e}_k) \to \infty$ et $\mathrm{d_{TV}}(0, N\mathbf{e}_k) \to 1$, lorsque $N \to \infty$.

Définition 3 (Cutoff en chi-deux et en variation totale)**.** On dit que $\{\chi^2(\cdot, N\mathbf{e}_k), N \geq 2\}$ vérifie un (t_N, ω_N) cutoff en chi-deux si $t_N \geq 0$, $\omega_N \geq 0$, $\omega_N = o(t_N)$ et

$$\lim_{c\to\infty} \limsup_{N\to\infty} \chi^2(t_N + c\,\omega_N, N\mathbf{e}_k) = 0, \quad \lim_{c\to-\infty} \liminf_{N\to\infty} \chi^2(t_N + c\,\omega_N, N\mathbf{e}_k) = \infty.$$

De façon analogue, on dit que $\{\mathrm{d_{TV}}(\cdot, N\mathbf{e}_k), N \geq 2\}$ vérifie un (t_N, ω_N) cutoff en variation totale si $t_N \geq 0$, $\omega_N \geq 0$, $\omega_N = o(t_N)$ et

$$\lim_{c\to\infty} \limsup_{N\to\infty} \mathrm{d_{TV}}(t_N + c\,\omega_N, N\mathbf{e}_k) = 0, \quad \lim_{c\to-\infty} \liminf_{N\to\infty} \mathrm{d_{TV}}(t_N + c\,\omega_N, N\mathbf{e}_k) = 1.$$

Les suites $(t_N)_{N\geq 2}$ et $(\omega_N)_{N\geq 2}$ sont appelées suites de *cutoff* et de *fenêtres*, respectivement.

Voir la Définition 2.1 et la Remarque 2.1 dans [CS08].

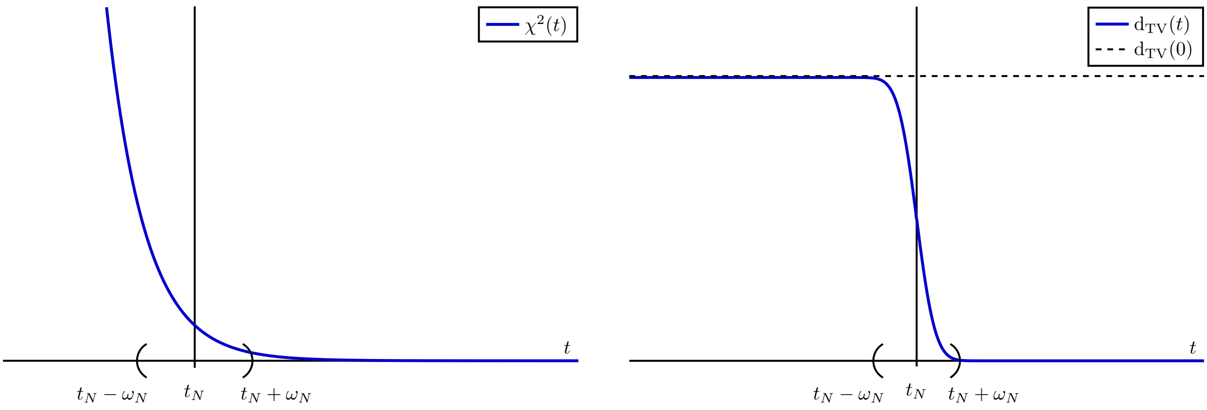
Le phénomène de cutoff décrit une transition abrupte dans la convergence vers l’équilibre : sur une période négligeable donnée par la suite fenêtre $(\omega_N)_{N>2}$, la distance de l’équilibre passe de près de sa valeur initiale à près de zéro à un moment donné par la suite cutoff $(t_N)_{N\geq 2}$.

Une condition plus forte pour l'existence d'un phénomène de cutoff en chi-deux (t_N, ω_N) (resp. en variation totale) est l'existence de la limite

$$G_k(c) := \lim_{N\to\infty} \chi^2(t_N + c\,\omega_N, N\mathbf{e}_k) \quad \left(\text{resp. } H_k(c) := \lim_{N\to\infty} \mathrm{d_{TV}}(t_N + c\,\omega_N, N\mathbf{e}_k)\right), \tag{16}$$

pour une fonction G_k (resp. H_k), avec $k \in [K]$, qui vérifie

$$\lim_{c\to-\infty} G_k(c) = \infty \text{ et } \lim_{c\to\infty} G_k(c) = 0, \left(\text{resp. } \lim_{c\to-\infty} H_k(c) = 1 \text{ et } \lim_{c\to\infty} H_k(c) = 0)\right).$$



(a) Profil d'un cutoff en chi-deux

(b) Profil d'un cutoff en variation totale

Figure 2: Graphiques des exemples des fonctions G_k (partie gauche) et H_k (partie droite) définies dans (16). Les deux graphiques montrent une transition brusque dans la convergence vers l'équilibre : sur une période négligeable donnée par la suite des fenêtres $(\omega_N)_{N>2}$, la distance à l'équilibre passe de près de sa valeur initiale à près de zéro à un moment donné par la suite des cutoffs $(t_N)_{N\geq 2}$.

En fait, dans ce cas, le cutoff (t_N, ω_N) est dit *fortement optimal*, voir la Définition 2.2 et la Proposition 2.2 dans [CS08]. Voir aussi les Sections 2.1 et 2.2 de [CS08] et le Chapitre 2 de [Che06] pour plus de détails sur la définition du cutoff (t_N, ω_N) et l'optimalité des suites des fenêtres. La Figure 2 montre des profils classiques de fonctions G_k et H_k tels que définis dans (16) pour les cutoffs en variation totale et chi-deux.

Les deux résultats suivants montrent l'existence de phénomènes de cutoff en chi-deux et en variation totale pour le processus de Moran multi-allélique généré par $\mathcal{L}_{N,p}$, pour $p \geq 0$, lorsque la distribution initiale est concentrée en $N\mathbf{e}_k$, pour $k \in [K]$. Dans le cas de la distance chi-deux, nous pouvons fournir explicitement le profil limite de la distance. De plus, nous prouvons que la distance en variation totale à la stationnarité du processus de mutation piloté par $\mathcal{L}_N$, c'est-à-dire pour $p = 0$, a un profil gaussien, lorsque tous les individus sont initialement du même type.

Théorème 8 (Profile limite du cutoff chi-deux lorsque $N \to \infty$)**.** *Pour $k \in [K]$, avec $K \geq 2$, $p \geq 0$ et pour tout $c \in \mathbb{R}$, on a*

$$\lim_{N\to\infty} \chi^2\left(t_{N,c}, N\mathbf{e}_k\right) = \exp\{K_{k,p}\mathrm{e}^{-c}\} - 1,$$

où $t_{N,c} = \dfrac{\ln N + c}{2|\boldsymbol{\mu}|}$ et $K_{k,p} = \dfrac{|\boldsymbol{\mu}|(|\boldsymbol{\mu}| - \mu_k)}{\mu_k(|\boldsymbol{\mu}| + p)}$. Ainsi, le processus de Markov généré par $\mathcal{L}_{N,p}$ vérifie un $\left(\frac{\ln N}{2|\boldsymbol{\mu}|}, 1\right)$ cutoff en chi-deux fortement optimal lorsque $N \to \infty$.

Figure 3 montre la convergence de $\chi^2(t_{N,c}, N\mathbf{e}_k)$, pour $t_{N,c} = \frac{\ln(N)+c}{2|\boldsymbol{\mu}|}$, vers $G_k(c) = \exp\{K_{k,p}\mathrm{e}^{-c}\} - 1$, lorsque $N \to \infty$.

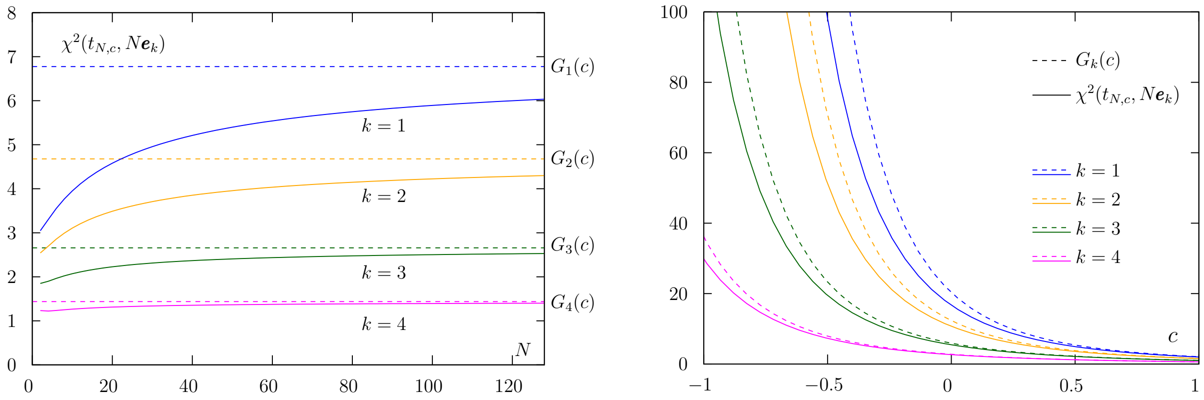


Figure 3: Pour les paramètres $K = 4$, $k \in [K]$, $\mu = (0.7, 0.8, 1.0, 1.3)$, $p = 1.7$: la partie gauche montre $\chi^2(t_{N,c}, N\mathbf{e}_k)$ comme fonction de N, $2 \leq N \leq 128$ pour $c = 0.4$, et la partie droite montre $G_k(c)$ et $\chi^2(t_{N,c}, N\mathbf{e}_k)$ comme fonctions de c, pour $-1 \leq c \leq 1$, for $N = 150$.

Théorème 9 (Cutoff en variation totale lorsque $N \to \infty$)**.** *Pour $k \in [K]$, avec $K \geq 2$, $p \geq 0$ et tout $c > 0$, on a*

$$\begin{aligned} \mathrm{d_{TV}}\left(t_{N,c}, N\mathbf{e}_k\right) &\geq 1 - 32|\mu|\kappa_k \mathrm{e}^{-c}, \\ \lim_{N\to\infty} \mathrm{d_{TV}}\left(t_{N,c}, N\mathbf{e}_k\right) &\leq \sqrt{\exp\{K_{k,p}\mathrm{e}^{-c}\} - 1}, \end{aligned}$$

où $t_{N,c} = \dfrac{\ln N + c}{2|\boldsymbol{\mu}|}$, $\kappa_k = \max\limits_{r:r\neq k} \dfrac{\mu_r \wedge \mu_k}{\mu_k}$ et $K_{k,p} = \dfrac{|\boldsymbol{\mu}|(|\boldsymbol{\mu}| - \mu_k)}{\mu_k(|\boldsymbol{\mu}| + p)}$. Ainsi, le processus de Markov généré par $\mathcal{L}_{N,p}$ vérifie un $\left(\frac{\ln N}{2|\boldsymbol{\mu}|}, 1\right)$ cutoff en variation totale lorsque $N \to \infty$.

De plus, lorsque $p = 0$, le profil limite de la distance en variation totale vérifie

$$\lim_{N\to\infty} \mathrm{d_{TV}}(t_{N,c}, N\mathbf{e}_k) = 2\Phi\left(\frac{1}{2}\sqrt{K_{k,0}\mathrm{e}^{-c}}\right) - 1,$$

où Φ est la fonction de répartition de la distribution normale standard. Alors, le processus généré par $\mathcal{L}_N$ vérifie un $\left(\frac{\ln N}{2|\boldsymbol{\mu}|}, 1\right)$ cutoff en variation totale lorsque $N \to \infty$.

Les preuves des Théorèmes 8 et 9 sont données dans la Section 2.5.1.

La Figure 4 montre la convergence de $\mathrm{d_{TV}}(t_{N,c}, N\mathbf{e}_k)$, avec $t_{N,c} = \frac{\ln(N)+c}{2|\boldsymbol{\mu}|}$, vers $H_k(c) = 2\Phi\left(\frac{1}{2}\sqrt{K_{k,0}\mathrm{e}^{-c}}\right) - 1$, lorsque $N \to \infty$.

Certains auteurs ont étudié l'existence d'un cutoff dans des modèles du type Moran. Par exemple, Donelly et Rodrigues [DR00] ont prouvé l'existence d'un cutoff pour le modèle de Moran neutre bi-allélique dans la distance en séparation[11]. Pour ce faire, ils ont utilisé une propriété de dualité du processus de Moran et ont trouvé une expression asymptotique pour la convergence en la distance en séparation pour un temps approprié mis à l'échelle, lorsque le nombre d'individus tend vers l'infini. Khare et Zhou [KZ09] ont trouvé des bornes pour la distance chi-deux pour un processus de Moran multi-allélique en temps discret qui implique l'existence d'un phénomène de cutoff. Diaconis et Griffiths [DG19] ont étudié l'existence d'un

[11] *Separation distance* en anglais.

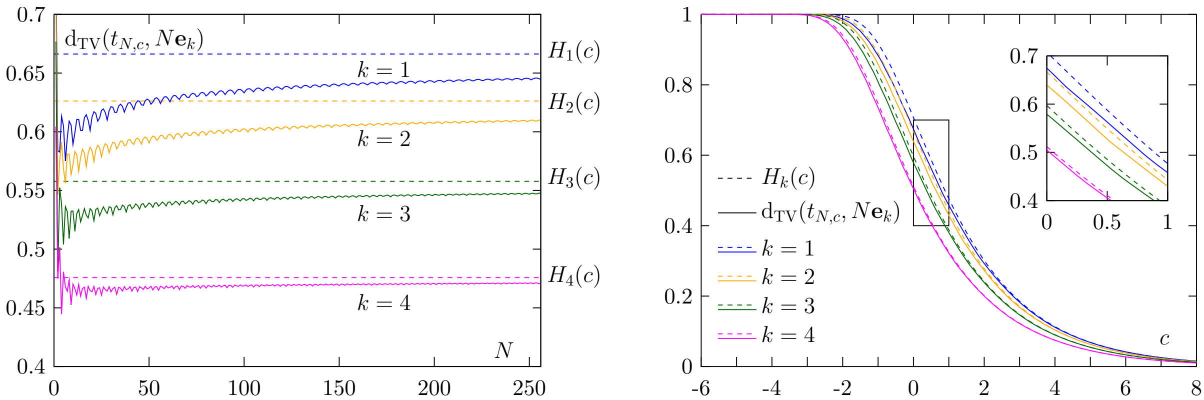


Figure 4: Pour les paramètres $K = 4$, $k \in [K]$, $\mu = (0.7, 0.8, 1.0, 1.3)$, $p = 0$: la partie gauche montre $\mathrm{d_{TV}}(t_{N,c}, N\mathbf{e}_k)$ comme fonction de N, $2 \leq N \leq 256$ avec $c = 0.17$, et la partie droite montre $H_k(c)$ et $\mathrm{d_{TV}}(t_{N,c}, N\mathbf{e}_k)$ comme fonctions de c, pour $-6 \leq c \leq 8$, avec $N = 100$.

cutoff en chi-deux et en variation totale pour un analogue en temps discret du processus de mutation généré par $\mathcal{L}_N$. Les Théorèmes 8 et 9 renforcent les résultats obtenus dans [KZ09] et [DG19], puisqu'ils donnent les profils limites pour la distance à l'équilibre en chi-deux et en variation totale, pour $p \geq 0$ et $p = 0$, respectivement.

Le Théorème 9 est, à notre connaissance, le premier résultat assurant l'existence d'un phénomène de cutoff en variation totale pour le modèle de Moran neutre avec $p > 0$ et mutation indépendante des parents.

Bornes quantitatives dans un exemple non réversible

Nous présentons ici les principaux résultats obtenus dans l'article [Cor21a], publié dans *Stochastic Processes and Their Applications*, qui sont présentés en détail dans le Chapitre 3.

L'objectif principal du Chapitre 3 est de fournir des estimations quantitatives pour les convergences du schéma (5) pour le modèle de Moran multi-allélique neutre, où le processus de mutation est piloté par une marche aléatoire asymétrique dans le graphe cyclique. En particulier, nous fournissons des estimations explicites pour les bornes du Théorème 2 ci-dessus.

Comme nous l'avons commenté précédemment, la convergence des distributions empiriques induites par un processus de particules de Moran (ou Fleming−Viot) défini sur des espaces d'états discrets, lorsque la taille de la population et le temps tendent vers l'infini, ont été assurés sous certaines hypothèses. Lorsque l'espace d'état est fini, comme dans le modèle considéré dans le Chapitre 3, il est bien connu que le processus de Moran est exponentiellement ergodique, et le phénomène de propagation du chaos est vérifié. On peut citer à titre d'exemple les travaux d'Asselah et al. [AFG11], Benaim et Cloez [BC15], Cloez et Thai [CT16b] et Lelièvre et al.[LPR18]. Cependant, il n'y a pas beaucoup de travaux dédiés à l'étude d'exemples spécifiques et fournissant des bornes explicites pour les convergences, même en considérant la neutralité comme nous le faisons ici. Dans cette direction, nous pouvons citer les travaux de Cloez et Thai [CT16a] et Del Moral et Jasra [DJ18].

Considérons la constante λ définie dans [CT16b] comme suit

$$\lambda = \inf_{x,y} \Big(Q_{x,y} + Q_{y,x} + \sum_{s \neq x,y} Q_{x,s} \wedge Q_{y,s} \Big), \tag{17}$$

où $Q = (Q_{x,y})_{x,y}$ est la matrice génératrice du processus jusqu'à l'absorption. Lorsque $\lambda = 0$ certains des résultats de [CT16b] ne sont pas vérifiés et la plupart des bornes deviennent trop grossières. Notons que $\lambda > 0$ pour les deux exemples étudiés dans [CT16a], mais que λ est égal à zéro pour les modèles où il existe deux sommets tels que la distance entre eux est supérieure à deux. La constante λ est en quelque sorte liée à la géométrie du graphe associé au processus de mutation du modèle de Moran. Par conséquent, il devient intéressant de trouver des bornes explicites pour la vitesse de convergence des processus de Moran avec des géométries plus complexes associées à leurs processus de mutation.

Dans cette section, et plus tard dans le Chapitre 3, nous nous concentrons sur la marche aléatoire sur le graphe cyclique $\mathbb{Z}/K\mathbb{Z}$, pour $K \geq 3$. Pour simplifier, nous supposons que les N particules sautent vers l'état absorbant avec le même taux. Même si dans ce cas la distribution du processus conditionnel est triviale, l'étude du processus de Moran neutre devient plus compliquée en raison de sa non réversibilité et de la géométrie du graphe cyclique. Nous fournissons des limites explicites pour la vitesse de convergence de la distribution empirique induite par le système de particules vers l'unique QSD lorsque t et N tendent vers l'infini. Cet exemple peut être considéré comme un pas de plus dans l'étude de la vitesse de convergence des processus de Moran (ou Fleming−Viot) avec des géométries plus générales.

Considérons le processus de Markov $(Z_t)_{t\geq 0}$ avec l'espace d'état $\mathbb{Z}/K\mathbb{Z} \cup \{\partial\}$, où $K \geq 3$ et ∂ est un état absorbant, de générateur infinitésimal donné par

$$\mathcal{G}f(x) = f(x+1) - f(x) + \theta[f(x-1) - f(x)] + p[f(\partial) - f(x)],$$

où $x \in \mathbb{Z}/K\mathbb{Z}$, $\mathcal{G}f(\partial) = 0$, avec $\theta, p \in \mathbb{R}_+^*$ et f est une fonction réelle définie sur $\mathbb{Z}/K\mathbb{Z} \cup \{\partial\}$. Autrement dit, $(Z_t)_{t\geq 0}$ est une marche aléatoire asymétrique sur le graphe cyclique K, qui saute avec des taux 1 et θ dans le sens horaire et antihoraire, respectivement. De plus, avec un taux uniforme p, le processus saute à l'état absorbant ∂, c'est-à-dire qu'il est tué. Notons que $\mathbb{Z}/K\mathbb{Z}$ est une classe irréductible. Le processus généré par $\mathcal{G}$ est un cas particulier des processus avec *mortalité uniforme dans un espace d'état fini* considérés par Méléard et Villemonais [MV12, §2.3].

Soit $(X_t)_{t\geq 0}$ la marche aléatoire asymétrique analogue sur le graphe cyclique $\mathbb{Z}/K\mathbb{Z}$, sans mortalité. Le générateur de ce processus, désigné par $\mathcal{H}$, est donné par

$$\mathcal{H}f(x) = f(x+1) - f(x) + \theta[f(x-1) - f(x)], \quad \text{pour tout} \quad x \in \mathbb{Z}/K\mathbb{Z}.$$

La Figure 5 montre un schéma des taux associés à $\mathcal{G}$ et $\mathcal{H}$ dans le graphe cyclique.

Notez que, en raison de la mortalité uniforme, le processus $(Z_t)_{t\geq 0}$ pourrait également être défini de la manière suivante

$$Z_t = \begin{cases} X_t & \text{si} \quad t < \tau_p \\ \partial & \text{si} \quad t \geq \tau_p, \end{cases}$$

où τ_p est une variable aléatoire exponentielle de moyenne $1/p$ et indépendante de la marche aléatoire $(X_t)_{t\geq 0}$. Cela signifie que la loi du processus $(Z_t)_{t\geq 0}$ conditionné à ne pas être absorbé jusqu'au temps $t \geq 0$ est donnée par $\mathbb{P}_\mu(Z_t = k \mid t < \tau_p) = \mathbb{P}_\mu(X_t = k)$, pour $k \in \mathbb{Z}/K\mathbb{Z}$ et pour chaque distribution initiale μ sur $\mathbb{Z}/K\mathbb{Z}$. En conséquence, la QSD de $(Z_t)_{t\geq 0}$ est la distribution stationnaire de $(X_t)_{t\geq 0}$, qui est la distribution uniforme sur $\mathbb{Z}/K\mathbb{Z}$, comme nous le prouvons dans le Théorème 3.1.1.

Maintenant, supposons que nous ayons N particules avec un comportement indépendant générés par $\mathcal{G}$, jusqu'à ce que l'une d'entre elles saute à l'état absorbant. Lorsque cela se produit, la particule saute instantanément et uniformément à l'une des positions des autres $N-1$ particules. Nous notons par $(\eta_t^{(N)})_{t\geq 0}$ le processus de Markov qui représente les positions des N particules dans le graphe cyclique K au temps t. Par abus de notation, nous redéfinissons l'espace d'état $\mathcal{E}_{K,N}$ comme suit

$$\mathcal{E}_{K,N} = \left\{ \eta : \mathbb{Z}/K\mathbb{Z} \to \mathbb{N} \mid \sum_{k=0}^{K-1} \eta(k) = N \right\}.$$

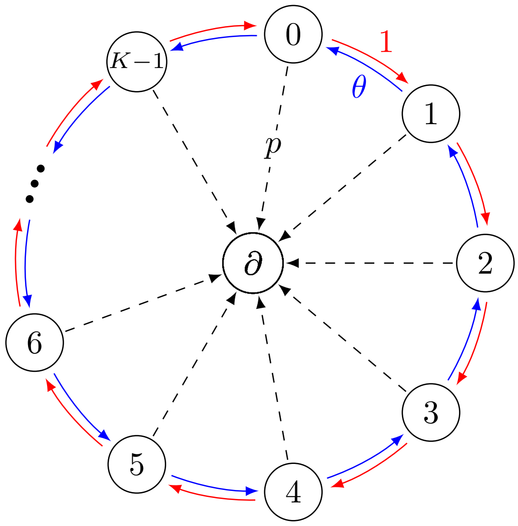


Figure 5: Schéma des taux associés aux marches aléatoires asymétriques sur le graphe cyclique avec et sans absorption, respectivement. Les flèches bleues et rouges représentent les interactions asymétriques de la marche aléatoire, tandis que les flèches en tirets représentent les taux d'absorption.

Le générateur du processus des N particules $(\eta_t^{(N)})_{t\geq 0}$, noté $\mathcal{Q}_{K,N}$, appliqué à une fonction f sur $\mathcal{E}_{K,N}$ donne

$$(\mathcal{Q}_{K,N}f)(\eta) = \sum_{i,j\in\mathbb{Z}/K\mathbb{Z}} \eta(i)\left(\mathbb{1}_{\{j=i+1\}} + \theta\mathbb{1}_{\{j=i-1\}} + p\frac{\eta(j)}{N-1}\right)[f(\eta - \mathbf{e}_i + \mathbf{e}_j) - f(\eta)],$$

où $\theta, p > 0$ et $\mathbf{e}_i$ est le i-ième vecteur canonique de $\mathbb{R}^K$. Dans le cadre de cette dynamique, chacune des N particules, quel que soit l'endroit où elle se trouve, peut sauter sur chaque site $j \in \mathbb{Z}/K\mathbb{Z}$ tel que $\eta(j) > 0$. Notons que le processus $(\eta_t^{(N)})_{t\geq 0}$ est donc irréductible. Par conséquent, il possède une distribution stationnaire unique notée ν_N.

La définition de $\mathcal{Q}_{K,N}$ diffère du générateur du processus de Moran multi-allélique défini dans (1) et aussi (13), car le dénominateur est $N-1$ à la place de N. Il s'agit d'une différence classique entre les définitions usuelles des processus de particules de Fleming−Viot et des modèles de Moran. Cependant, les résultats pour le générateur analogue à celui défini dans (13) peuvent être facilement retrouvés en prenant le taux de sélection p égal à $pN/(N-1)$. Nous conservons la définition de $\mathcal{Q}_{K,N}$ dans la suite de cette section et dans le Chapitre 3 en accord avec la version publiée [Cor21a].

La distribution empirique (aléatoire) $m(\eta_t^{(N)})$ approche la QSD du processus $(Z_t)_{t\geq 0}$ (cf.[AFG11]) qui, grâce au Théorème 3.1.1 dans le Chapter 3, est la distribution uniforme. Nous souhaitons étudier à quelle vitesse $m(\eta_t^{(N)})$ converge vers la distribution uniforme sur $\mathbb{Z}/K\mathbb{Z}$ lorsque t et N tendent vers l'infini. Considérons $\eta_\infty^{(N)}$ une variable aléatoire de distribution ν_N, la distribution stationnaire du processus $(\eta_t^{(N)})_{t\geq 0}$. Dans ce travail, nous développons une analyse similaire à celle de la dynamique du graphe complet dans [CT16b]. Les convergences dans (5) se réduisent dans ce cas à

$$\begin{tikzcd} m(\eta_t^{(N)}) \arrow[r, "t\to\infty"'] \arrow[d, "N"'] & m(\eta_\infty^{(N)}) \arrow[d, "N"] \\ \mathcal{L}(X_t) \arrow[r, "t\to\infty"'] & \mathcal{U}(\mathbb{Z}/K\mathbb{Z}) \end{tikzcd}$$

où les limites sont en loi, $\mathcal{L}(X_t)$ représente la loi de X_t au temps t et $\mathcal{U}(\mathbb{Z}/K\mathbb{Z})$ représente la distribution uniforme sur $\mathbb{Z}/K\mathbb{Z}$. Le Théorème 3.1.1 fournit des bornes exponentielles inférieure et supérieure pour la vitesse de convergence de $\mathcal{L}(Z_t \mid t < \tau_p)$ vers $\mathcal{U}(\mathbb{Z}/K\mathbb{Z})$ dans la norme L^2, lorsque $t \to \infty$. De même, le Corollaire 3.1.7 et le Théorème 3.1.9 donnent des bornes pour la vitesse de convergence de $m(\eta_t^{(N)})$ vers $\mathcal{L}(Z_t \mid t < \tau_p)$ et $m(\eta_\infty^{(N)})$ vers $\mathcal{U}(\mathbb{Z}/K\mathbb{Z})$, lorsque $N \to \infty$.

Le comportement quantitatif en temps long du système de N particules dans des espaces d'états dénombrables est étudié dans [CT16b]. En utilisant une technique de couplage et sous certaines conditions, une borne exponentielle est fournie pour la convergence de $\mathcal{L}(\eta_t^{(N)})$ vers ν_N au sens d'une *distance de Wasserstein* [CT16b, Thm. 1.1]. En particulier, le paramètre λ défini par (17) doit être positif. Ce n'est pas le cas de la marche aléatoire asymétrique sur le graphe cyclique K avec mortalité uniforme, lorsque $K \geq 6$. Une étude de cette convergence peut être effectuée en utilisant le spectre du générateur $\mathcal{Q}_{K,N}$, qui a été discuté dans la section précédente. En effet, en utilisant l'Exemple 2.4.1 dans le Chapitre 2, nous pouvons obtenir l'expression asymptotique suivante pour le profil de la convergence de la distance de variation totale vers la stationnarité :

$$\mathrm{D}_{\mathrm{TV}}^{\mathcal{Q}_{K,N}}(t) = \mathcal{O}\left(\mathrm{e}^{-\rho_K t}\right),$$

où $\rho_K = 2(1+\theta)\sin^2(\pi/K)$, et $\mathrm{D}_{\mathrm{TV}}^{\mathcal{Q}_{K,N}}$ est définie comme dans (14). Cependant, l'obtention d'une borne explicite pour cette convergence, ou la preuve de l'existence d'un phénomène de cutoff, sont des sujets de recherche à approfondir.

Introduction (English version)

This thesis is based on the study of a continuous-time discrete space Markovian interacting particle systems, modelling the evolution of a population of N individuals of different types given by the elements of a discrete set E. In words, each individual changes its type according to a Markovian mutation process independently of the others. The interaction occurs during the reproduction events where one individual dies and another randomly chosen, possibly the same, is duplicated. This type of reproduction is of Moran type and we call this process *multi-allelic Moran model* [Dur08; EG09].

Two asymptotic quantities of interest arise in the study of these processes: the stationary distribution of the particle process (the limit distribution when the time goes to infinity) and the limit empirical distribution induced by the particle process in E (the limit when the number of particles goes to infinity). The existence of these limits and the speed of the convergences are of high interest for theoretical and practical purposes. This thesis is centred on the study of the speed of these convergences previously described.

We first introduce the multi-allelic Moran model and describe more explicitly the limits that will attract our attention during the rest of the exposition. We then describe the motivations and the problem to which this thesis is addressed. Finally, we summarize the main results of the thesis, which are presented and demonstrated in detail in the subsequent chapters.

The multi-allelic Moran model

Consider E a discrete state space. The state space of the multi-allelic Moran model is the N discret simplex

$$\mathcal{E}_N := \left\{ \eta : E \to \mathbb{N} \;\Big|\; \sum_{x \in E} \eta(x) = N \right\}.$$

The empirical distribution induced by $\eta \in \mathcal{E}_N$ is defined by

$$m(\eta) = \sum_{x \in E} \frac{\eta(x)}{N} \delta_x \in \mathcal{M}_1(E),$$

where $\mathcal{M}_1(E)$ is the set of probability measures on E.

Let Q be the generator of a continuous-time, non-explosive, irreducible Markov chain, and consider the rates $V_\mu(x, y) \geq 0$, for all $x \neq y \in E$ and $\mu \in \mathcal{M}_1(E)$.

The multi-allelic Moran process is a continuous-time Markov chain evolving on $\mathcal{E}_N$. The process is at $\eta \in \mathcal{E}_N$ if there is $\eta(x)$ individuals of type x, for all $x \in E$. Between reproduction events, the N individuals evolve as independent copies of the mutation process generated by $Q = (Q_{x,y})_{x,y \in E}$. In this sense we call $Q_{x,y}$, for $x, y \in E$, the *mutation rates*.

Reproduction events consist of the death of an individual of type x, which is then removed from the population, and the reproduction of an individual of type y, which add an y individual to the population. This happens at rate $\eta(y)/N \cdot V_{m(\eta)}(x,y)$. Hence, the transition rate from $\eta \in \mathcal{E}_N$, with $\eta(x) > 0$, to $\eta - \mathbf{e}_x + \mathbf{e}_y$ is

$$\eta(x)\left(Q_{x,y} + \frac{\eta(y)}{N} V_{m(\eta)}(x,y)\right),$$

for every $x \neq y \in E$, where $\eta - \mathbf{e}_x + \mathbf{e}_y$ is the element in $\mathcal{E}_N$ satisfying

$$(\eta - \mathbf{e}_x + \mathbf{e}_y)(z) = \begin{cases} \eta(z) & \text{if} \quad z \notin \{x,y\}, \\ \eta(x) - 1 & \text{if} \quad z = x, \\ \eta(y) + 1 & \text{if} \quad z = y. \end{cases}$$

Hence, the Moran model considered previously described is a family of continuous-time Markov chains on $\mathcal{E}_N$, with generators $\mathcal{Q}_N$, which act on every function f on $\mathcal{E}_N$ as follows

$$(\mathcal{Q}_N f)(\eta) = \sum_{x,y\in E} \eta(x)\left(Q_{x,y} + \frac{\eta(y)}{N} V_{m(\eta)}(x,y)\right)[f(\eta - \mathbf{e}_x + \mathbf{e}_y) - f(\eta)], \tag{1}$$

for all $\eta \in \mathcal{E}_N$.

We will further detail particular examples but, for the moment, let us see that when $V_{m(\eta)}(x,y)$ is constant, each individual dies at the same rate and the parent is choose uniformly at random over the individuals that are present in the population (this explain the term $\eta(y)/N$ in the transition rate). We can also interpret this rate by the opposed point of view: each individual reproduces at a constant rate and the dying individual is chosen uniformly at random. This is often called *neutral selection* in ecology literature, but our models allow to choose various non constant $V_{m(\eta)}(x,y)$. In this sense, we call the rates $V_\mu(x,y)$, for $x,y \in E$ and $\mu \in \mathcal{M}_1(E)$, the *selection rates*.

Note that the reproduction dynamics depends in general on both the types of the parent and the offspring, and may also depend on the empirical distribution induced by the configuration of the population at the current time, in a sense that we will clarify further along in Assumptions (G1) and (C1).

The empirical distribution induced by the particle process at time t is given by $m(\eta_t^{(N)})$. The aim of this thesis is to understand the limits and convergence speeds of $m(\eta_t^{(N)})$ when $t \to \infty$ and $N \to \infty$.

Throughout this thesis the following boundedness condition holds:

$$\|V\| := \sup_{\mu\in\mathcal{M}_1(E)} \sup_{x,y\in E} V_\mu(x,y) < \infty.$$

Note that the non-explosion condition on Q and the bound condition on $\|V\|$ let out the possibility of an infinite number of jumps in finite time. Thus, the process is well-defined for all $t \geq 0$.

The multi-allelic Moran model generated by $\mathcal{Q}_N$ is an extension, for $K > 2$, of the bi-allelic model studied by [Cor17]. In general, when generalising the Moran model for more that two allelic types, the selection rates are taken depending only on the offspring type, i.e. $V_\mu(x,y) = V^{\mathrm{b}}(y)$, for all $x,y \in E$ and $\mu \in \mathcal{M}_1(E)$, which is called *selection at birth* or *fecundity selection* [Dur08; MW09; Eth11]. Besides, in biological applications it has been also considered the model with *selection at death* or *viability selection*, when the selection rates only depend on the parent type, i.e. $V_\mu(x,y) = V^{\mathrm{d}}(x)$ [MW09], for all $x,y \in E$ and $\mu \in \mathcal{M}_1(E)$. However, the importance of this last model goes outside its biological interpretations: this process is also called *Fleming – Viot particle process*, which is an interacting particle process intended for the approximation of the *quasi-stationary distribution* of an absorbing Markov chain. We will discuss later the relation among the Moran model considered here, Feynman – Kac semigroups, the theory of quasi-stationary distributions and Fleming – Viot particle processes.

We have that $\mathcal{Q}_N = \mathcal{Q}_N^{\mathrm{mut}} + \mathcal{Q}_N^{\mathrm{sel}}$, where these generators act on every $f \in \mathcal{B}_d(\mathcal{E}_N)$ as follows

$$(\mathcal{Q}_N^{\mathrm{mut}} f)(\eta) = \sum_{x,y \in E} \eta(x) Q_{x,y} [f(\eta - \mathbf{e}_x + \mathbf{e}_y) - f(\eta)],$$
$$(\mathcal{Q}_N^{\mathrm{sel}} f)(\eta) = \sum_{x,y \in E} \eta(x) \frac{\eta(y)}{N} V_{m(\eta)}(x,y) [f(\eta - \mathbf{e}_x + \mathbf{e}_y) - f(\eta)],$$

for all $\eta \in \mathcal{E}_N$, where $\mathcal{B}_d(\mathcal{E}_N)$ denotes the set of bounded real functions on $\mathcal{E}_N$. Note that the irreducibility of Q implies the irreducibility of $\mathcal{Q}_N$. The generator $\mathcal{Q}_N^{\mathrm{mut}}$ drives the independent movement of N particles according to the mutation rate Q. Besides, $\mathcal{Q}_N^{\mathrm{sel}}$ drives an absorbing Markov chain on $\mathcal{E}_N$. Indeed, once a site $y \in E$ is empty (the process is at η such that $\eta(y) = 0$) no particle can jump to y. Thus, the states η_x such that $\eta_x(x) = N$, for $x \in E$, are absorbing for the process generated by $\mathcal{Q}_N^{\mathrm{sel}}$.

Let us get some insight in the limit of the empirical mesure induced by this particle process when the number of particles tends towards infinity. By the Kolmogorov equation we know that $\partial_t \mathbb{E}_\eta[m_x(\eta_t)] = \mathbb{E}_\eta\left[(\mathcal{Q}_N\, m_x)(\eta_t)\right]$, where m_x stands for the empirical distribution induced by η on the point $x \in E$, i.e. $m_x : \eta \mapsto \eta(x)/N$. Let us thus compute $\mathcal{Q}_N^{\mathrm{mut}}\, m_x$:

$$(\mathcal{Q}_N^{\mathrm{mut}} m_x)(\eta) = \sum_{y \in E} Q_{y,x} m_y(\eta),$$

for every $x \in E$, for all $\eta \in \mathcal{E}_N$. On the other hand,

$$(\mathcal{Q}_N^{\mathrm{sel}} m_x)(\eta) = -\, m_x(\eta) \sum_{y \in E} m_y(\eta) [V_{m(\eta)}(x,y) - V_{m(\eta)}(y,x)].$$

Finally, we get

$$\partial_t \mathbb{E}_\eta[m_x(\eta_t)] = \sum_{y \in E} Q_{y,x} \mathbb{E}_\eta[m_y(\eta_t)] - \sum_{y \in E} [V_{m(\eta)}(x,y) - V_{m(\eta)}(y,x)] \mathbb{E}_\eta[m_x(\eta_t) m_y(\eta_t)].$$

When the number of individuals $N \to \infty$, we expect that the process exhibits a *propagation of chaos phenomenon*:

$$|\mathbb{E}_\eta[m_x(\eta_t) m_y(\eta_t)] - \mathbb{E}_\eta[m_x(\eta_t)] \mathbb{E}_\eta[m_y(\eta_t)]| \xrightarrow[N \to \infty]{} 0.$$

Namely, asymptotically the positions of the particles becomes independent. Thus, the empirical distribution induced by the process should approximate the solution of the following nonlinear differential equation:

$$\partial_t \gamma_t(x) = \sum_{y \in E} Q_{y,x} \gamma_t(y) - \sum_{y \in E} [V_{\gamma_t}(x,y) - V_{\gamma_t}(y,x)] \gamma_t(x) \gamma_t(y), \tag{2}$$

for all $x \in E$.

Let us now clarify the dependence of the selection rates on the empirical distribution induced by the particle system. Assume that the selection rates allow the following decomposition

$$V_\mu(x,y) = V_\mu^{\mathrm{d}}(x) + V_\mu^{\mathrm{b}}(y) + V_\mu^{\mathrm{s}}(x,y), \tag{3}$$

where $\mu \mapsto V_\mu^{\mathrm{d}}$ and $\mu \mapsto V_\mu^{\mathrm{b}}$ are continuous bounded maps from $(\mathcal{M}_1(E), \|\cdot\|_{\mathrm{TV}})$ to $(\mathcal{B}_+(E), \|\cdot\|)$, and $\mu \mapsto V_\mu^{\mathrm{s}}$ is a bounded continuous map from $(\mathcal{M}_1(E), \|\cdot\|_{\mathrm{TV}})$ to $(\mathcal{B}_+(E \times E), \|\cdot\|)$. In addition, V_μ^{s} is symmetric in $E \times E$, for every $\mu \in \mathcal{M}_1(E)$. We also assume that there exists a function $\Lambda \in \mathcal{B}_b(E)$ such that

$$\Lambda = V_\mu^{\mathrm{b}} - V_\mu^{\mathrm{d}}.$$

Thus, (2) reduces as follows: for every function ϕ on E we have

$$\partial_t \gamma_t(\phi) = \gamma_t\Big((Q + \Lambda)\phi - \gamma_t(\Lambda) \cdot \phi\Big). \tag{4}$$

Indeed, note that

$$\begin{aligned}\sum_{y \in E} \phi(x)[V_\gamma(x,y) - V_\gamma(y,x)]\gamma(x)\gamma(y) &= \gamma(Q\phi) + \gamma(V_\gamma^{\mathrm{d}})\gamma(\phi) - \gamma(V_\gamma^{\mathrm{d}}\phi) + \gamma(V_\gamma^{\mathrm{b}}\phi) - \gamma(V_\gamma^{\mathrm{b}})\gamma(\phi) \\ &= \gamma(Q\phi + \Lambda\phi - \gamma(\Lambda)\phi).\end{aligned}$$

Remark 1 (Independence on the symmetric component in (3)). Note that the symmetric function V_μ^{s} is absent in (4). In other words, the particle system would approximate the same deterministic measure, which is given by the solution of (4), independently of the symmetric element in the decomposition (3) of the selection rates.

Remark 2 (Feynman – Kac interpretation for the model with general selection rates). See [Del04, p. 25], and the references therein, for a Feynman – Kac interpretation for the differential equation (2) with rather general selection rates.

The following diagram summarises the four convergences that we study in the sequel:

$$\begin{array}{ccc} m(\eta_t^{(N)}) & \xrightarrow[t\to\infty]{} & \nu_N \\ N\big\downarrow & & \big\downarrow N \\ \gamma_t & \xrightarrow[t\to\infty]{} & \nu_\infty \end{array} \tag{5}$$

where ν_N denotes the stationary distribution of the process generated by $\mathcal{Q}_N$ and $(\gamma_t)_{t\geq 0}$ denotes the solution of (4), whenever they exist.

Feynman – Kac semigroups and quasi-stationary distributions

Consider $(X_t)_{t\geq 0}$ the Markov chain on E generated by Q. Let us define the Feynman – Kac semigroup for every $\phi \in \mathcal{B}_b(E)$ as follows

$$P_t^\Lambda(\phi) : x \mapsto \mathbb{E}_x\left[\phi(X_t)\exp\left\{\int_0^t \Lambda(X_s)\mathrm{d}s\right\}\right],$$

whose generator is $Q + \Lambda$. The normalised version of the semigroup, which is defined as follows

$$\mu_t(\phi) := \frac{\mu_0 P_t^\Lambda(\phi)}{\mu_0 P_t^\Lambda(\mathbf{1})}, \tag{6}$$

where $\mathbf{1}$ denotes the all-one function on E, is the solution of the nonlinear differential equation (4) with initial value $\mu_0(\phi)$ for $t = 0$ [Del04, Eq. (1.17)]. For an exhaustive and comprehensible study of the Feynman – Kac formulae, we refer to the book of Del Moral [Del04].

Note that $(\mu_t)_{t\geq 0}$ as defined above is invariant by translation of Λ. Indeed, for every real β we get

$$\mu_t(\phi) = \frac{\mu_0 P_t^{\Lambda-\beta}(\phi)}{\mu_0 P_t^{\Lambda-\beta}(\mathbf{1})}.$$

In particular, taking $\beta = \sup \Lambda$, we can always interpret $(\mu_t)_{t\geq 0}$ as the distribution of an absorbing Markov chain conditioned to non absorption up to time t and with killing rate $\kappa = \sup \Lambda - \Lambda$. This naturally relates the study of the behavior of $(\mu_t)_{t\geq 0}$ when $t \to \infty$, to the theory of *quasi-stationary distributions* (QSD), which is presented in the next section.

Quasi-stationary distributions

Let us denote by $(X_t)_{t\geq 0}$ an irreducible continuous-time non-explosive Markov chain on a discrete space E with generator Q. Let $\kappa : E \to \mathbb{R}_+$ be a bounded function. We add an absorbing state $\partial \notin E$, and we define the absorbing Markov chain $(Y_t)_{t\geq 0}$ on $E \cup \{\partial\}$, such that

$$Y_t = \begin{cases} X_t & \text{if } \int_0^t \kappa(X_s)\mathrm{d}s < \xi \\ \partial & \text{otherwise,} \end{cases}$$

where ξ is an exponential random variable with parameter 1, independent from $(X_t)_{t\geq 0}$. In words, $(Y_t)_{t\geq 0}$ evolves as $(X_t)_{t\geq 0}$ on E, and conditioned to be at $x \in E$, it jumps to ∂ (get absorbed) with rate $\kappa(x)$.

Alternatively, $(Y_t)_{t\geq 0}$ can be defined as the process with generator $\tilde{Q}$ acting on bounded functions φ on $E \cup \{\partial\}$ as follows:

$$(\tilde{Q}\varphi) : x \mapsto \kappa(x)(\varphi(\partial) - \varphi(x)) + \sum_{y\in E} Q_{x,y}(\varphi(y) - \varphi(x)),$$

for every $x \in E$, and we set $(\tilde{Q}\varphi)(\partial) = 0$.

Let us denote by τ_∂ the absorption time. We have

$$\mathbb{E}_\mu[\phi(Y_t)] = \mathbb{E}_\mu[\phi(X_t) \mid t < \tau_\partial] = \frac{\mathbb{E}_\mu\left[\phi(X_t)\mathrm{e}^{-\int_0^t \kappa(X_s)\mathrm{d}s}\right]}{\mathbb{E}_\mu\left[\mathrm{e}^{-\int_0^t \kappa(X_s)\mathrm{d}s}\right]}. \tag{7}$$

Notice the similarity between the previous expression and the normalised Feynman−Kac semigroup defined by (6).

Remark 3 (Link between the measure γ_t in (5) and an absorbing Markov chain)*.* When the decomposition (3) is verified, the limit of $m(\eta_t^{(N)})$ when $N \to \infty$, which we denoted γ_t in (5), is the law of an absorbing Markov chain conditioned on non absorption up to time $t \geq 0$. Furthermore, the generator of one of this absorbing Markov chains is $Q+\kappa$ where $\kappa := \sup \Lambda - \Lambda$.

The absorption time τ_∂, defined as

$$\tau_\partial := \inf_{t\geq 0}\{X_t = \partial\},$$

is a stopping time, which means that for all $s \geq 0$, $X_s = \partial$ implies $X_t = \partial$, for all $t \geq s$. Throughout this manuscript we assume that starting from any $x \in E$, the process eventually get absorbed, almost surely. Namely, we assume that

$$\mathbb{P}_x(\tau_\partial < \infty) = 1, \ \ \forall x \in E.$$

Thus, the stationary distribution of $(X_t)_{t\geq 0}$, which is the distribution concentrated at ∂, is not interesting. However, it is interesting to consider the stationary distribution of the process conditioned to non absorption, which is called the *quasi-stationary distribution.*

Definition 1 (Quasi-stationary distribution (QSD))**.** A quasi-stationary distribution for the absorbing Markov chain $(Y_t)_{t\geq 0}$ is any $\mu \in \mathcal{M}_1(E)$ satisfying

$$\mathbb{P}_\mu(Y_t \in \cdot \mid t < \tau_\partial) = \mu.$$

Definition 2 (Quasi-limit distribution (QLD))**.** A quasi-limit distribution for the absorbing Markov chain $(Y_t)_{t\geq 0}$ is any $\mu \in \mathcal{M}_1(E)$ satisfying

$$\lim_{t\to\infty} \mathbb{P}_\nu(Y_t \in \cdot \mid t < \tau_\partial) = \mu,$$

for any $\nu \in \mathcal{M}_1(E)$.

When the limit above holds true and is the same for all $\nu = \delta_x$, with $x \in E$, the distribution quasi-limite is said to be the *Yaglom limit*[12] or the *minimal quasi-stationary distribution* of $(Y_t)_{t\geq 0}$.

Analogously to the classic ergodic theory of Markov chains, the QSD and the QLD are equivalent objects. Moreover, a minimal QSD is always a QSD but the reversed implication is not always true:

$$\boxed{\text{QSD}} \Leftrightarrow \boxed{\text{QLD}} \Leftarrow \boxed{\text{Yaglom limit (minimal QSD)}}.$$

For more the datils on the study of quasi-stationary distributions, we refer to the survey of Méléard and Villemonais [MV12], the book of Collet, Martínez and San Martín [CMS13], and the paper of van Doorn and Pollet [DP13], for the specific context of dicrete state spaces.

The definition of QLD and Remark 3 imply that ν_∞ in (5) can be understood as a QSD.

Remark 4 (Measure ν_∞ in (5) is a QSD)*.* When the decomposition (3) is verified, the probability distribution ν_∞ in (5) is the QSD of the absorbing Markov chain with generator $Q + \kappa$, where $\kappa := \sup \Lambda - \Lambda \in \mathcal{B}_b(E)$.

During this work we consider a stronger condition: the uniform exponential convergence of the normalised Feynman – Kac semigroup. We assume that there exist a distribution $\mu_\infty \in \mathcal{M}_1(E)$ and $C, \gamma > 0$, such that

$$\|\mu_t - \mu_\infty\|_{\mathrm{TV}} \leq C\mathrm{e}^{-\gamma t}, \text{ for all } \mu_0 \in \mathcal{M}_1(E) \text{ and } t \geq 0, \tag{8}$$

where $(\mu_t)_{t\geq 0}$ is defined as in (6).

The exponential convergence (8) always verified when the state space if finite. In this case the QSD is given by the normalised left-eigenvector of the generator Q, associated to the second largest eigenvalue in modulus (SLEM). This was proved by Darroch and Seneta [DS67] and the result comes as a consequence of the Perron – Frobenius Theorem [Per07; Fro12] (see also [MV12, Thm. 8] for the specific context of quasi-stationary distributions).

The case where E is countable is more delicate and has attracted lots of attention and several methods have been applied. Thanks to the exhaustive work of Champagnat and Villemonais, specifically [CV16; CV17b], it is possible to describe hypothesis equivalent to (8) and explore the consequences of the uniform exponential convergence to the QSD. We further discuss this topics in Section 1.2. Meanwhile, let us just state the following result which clarify the consequences of (8) to the exponential ergodicity of the unnormalised semigroup $(P_t^\Lambda)_{t\geq 0}$.

Lemma 1 (Exponential ergodicity of the non normalised semigroup)**.** *Assume that* (8) *is verified. Then, there exist a unique triplet* $(\mu_\infty, h, \lambda) \in \mathcal{M}_1(E) \times \mathcal{B}_b(E) \times \mathbb{R}$*, of eigenelements of* $Q + \Lambda$ *such that* h *is strictly positive,* $\mu_\infty(h) = 1$ *and satisfy*

$$\mu_\infty P_t^\Lambda = \mathrm{e}^{\lambda t}\mu_\infty \text{ and } P_t^\Lambda(h) = \mathrm{e}^{\lambda t} h.$$

[12]This name comes from the mathematician A. M. Yaglom, who started the study of this type of objects for sub-critical Galton – Gatson process in [Yag47].

Moreover, there exist $C, \gamma > 0$ such that for all $t \geq 0$:

$$\sup_{\mu_0 \in \mathcal{M}(E)} \|\mathrm{e}^{-\lambda t}\mu_0 P_t^\Lambda - \mu_0(h)\mu_\infty\|_{\mathrm{TV}} \leq C\mathrm{e}^{-\gamma t}.$$

Furthermore, $\lambda \leq 0$ whether $\Lambda \leq 0$.

This result is basically a consequence of Theorem 2.1 of [CV17b] (Theorem 1.2.2 below). Essentially, it says that the uniform exponential convergence of the normalised semigroup implies that of the unnormalised semigroup.

Fleming – Viot particle systems

The Fleming – Viot particle system consists in N particles moving in E as independent copies of an absorbing Markov chain, until one of the particles gets absorbed. When this happens, this particle jumps instantaneously and uniformly to the positions of one of the other particles. The Fleming – Viot particle systems were originally and independently introduced by Del Moral, Guionnet and Miclo [DG99; DM00a][13], and Burdzy, Hołyst and March [BHM00] to approximate the law of a Markov process conditioned to non absorption, and its QSD in the continuous state space setting. These particle processes have attracted lots of attention in recent years. See for instance [DM03; Vil14; Cér+20; CV21] for general state spaces, [FM07; GJ13; Ass+16; CT16b] for countable state spaces, and even [AFG11; LPR18] for finite state spaces.

The Fleming – Viot particles process is a particular case of the multi-allelic Moran process with *selection at death* or *viability selection*, when the selection rates only depend on the parent type, i.e. $V_\mu(x, y) = \kappa(x)$, for all $x, y \in E$ and $\mu \in \mathcal{M}_1(E)$, with $\kappa \in \mathcal{B}_b(E)$. We denote by $\mathcal{F}_N$ the infinitesimal generator of the Fleming – Viot process which acts on every bounded function on $\mathcal{E}_N$ as follows

$$(\mathcal{F}_N f)(\eta) = \sum_{x,y \in E} \eta(x) \left(Q_{x,y} + \frac{\eta(y)}{N}\kappa(x) \right) [f(\eta - \mathbf{e}_x + \mathbf{e}_y) - f(\eta)],$$

for all $\eta \in \mathcal{E}_N$.

The definition of the Fleming – Viot particle process is motivated by simulation purposes. The simulation of the law (7) and the the estimation of its QSD by usual methods (as rejection sampling) are unviable, since to event to which one restrict the observation has probability tending (usually exponentially) to zero. The Fleming – Viot particle system offer an alternative method. Under some assumptions, it can be proved that the when $N \to \infty$, the empirical measure induced by the particle system approaches the conditional distribution (7). Thus, when $N, t \to \infty$, the particle system can be used to estimate the QSD associated to (7). See for example [GJ13] for more details about the simulations of quasi-stationary distributions on countable spaces.

Let us review, without being exhaustive, the state of the art related to the convergence of the empirical measure induced by Fleming – Viot as an etimator of a quasi-stationary distribution. Ferrari and Marić [FM07] and Asselah et al. [AFG11] study the convergence of the empirical distribution induced by the Fleming – Viot process to the unique QSD in countable and finite discrete space settings, respectively. With the aim to study the convergence of the particle process under the stationary distribution to the QSD, Lelièvre et al. [LPR18] prove a

[13] Actually, Del Moral and Miclo called this process *Moran-type particle system*, which is maybe a moreaccurate name in the discrete state space setting, in order to avoid confusion with the existence of a measure-valued process related to the Moran process in population genetics introduced by Fleming and Viot [FV79] and named *Fleming – Viot process*. See [FV79, Appendix B] for a discussion on the relationship between Fleming–Viot process (in the sense of population genetics) and the multi-allelic Moran model, see also [Fen10, §6.2].

Central Limit Theorem for the finite state case. Additionally, Villemonais [Vil15] and Asselah et al. [Ass+16] study the convergence to the minimal QSD in a birth-and-death process and in a Galton – Watson type model, respectively. Similarly, Asselah and Thai [AT12] and Mariĉ [Mar15] address the study of the N particle system associated to a random walk on $\mathbb{N}_0$ with a drift towards the origin, which is an absorbing state. In the scenario where the state space E is infinite (as for the Galton – Watson and the birth-and-death processes), they could exist an infinite number of QSD. It is thus important to ensure the ergodicity of the particle system and determine to which QSD it converges. In this direction, Champagnat and Villemonais [CV21] study the convergence of the Fleming – Viot process to the minimal QSD under general conditions, providing also some specific examples. See also the paper of Cloez and Thai [CT16b], ensuring, under strong mixing conditions, the convergences in (5) with explicit quantitative estimates for the speed of convergence.

Main results

The work we present here can be seen as a continuation of the research related to ergodicity and propagation of chaos on Fleming – Viot processes. In Chapter 1, we go further and study a more general class of models: the multi-allelic Moran models with selection rates satisfying (3), which are particle systems whose empirical measure also approaches a QSD when $t, N \to \infty$. Theorem 2 establishes the uniform in time convergence of order $1/\sqrt{N}$ of $m(\eta_t^{(N)}) \to \gamma_t$ as in diagram (5). In particular, this result ensure a uniform in time propagation of chaos with rate $1/\sqrt{N}$ for the Fleming – Viot process. This problem has attracted considerable attention in recent years. See, for instance, [Vil14, Thm. 2.2], [CT16b, Thm. 1.5], [LPR18, Remark 4.3] and [CV21, Thm. 2.3]. In Theorem 3 we go beyond and prove the asymptotic normality of $\sqrt{N}\Big(m(\eta_T^{(N)})(\phi) - \gamma_T(\phi)\Big)$, when $N \to \infty$, with an explicit expression for the variance, for all $T \geq 0$ and $\phi \in \mathcal{B}_b(E)$. The empirical distribution induced by the Moran model with associated function Λ, approaches the same QSD that approaches the Fleming – Viot particle process with killing rate $\kappa = \sup \Lambda - \Lambda$ (see Remarks 1, 3 and 4). This naturally arise the question: what particle system is the best estimator of a QSD of interest? We partially answer this question in Corollary 4 below showing a way to minimise the asymptotic quadratic error when $t, N \to \infty$, of $(m(\eta_t))_{t\geq 0}$ as an estimator of $(\mu_t)_{t\geq 0}$. However, the study of the Moran particle process with minimal quadratic error requires further research.

In Chapter 2 we consider the case where the state space of the allelic types is finite and the selection rate is uniform. In the context of Fleming – Viot process, this is equivalent to consider a constant killing rate κ. Although this is a quite simplified model, our results are particularly strong: in Theorem 5 we obtain an explicit expression for the eigenvalues of the generator of the particle system $\mathcal{Q}_N$ as function of the eigenvalues of the mutation rate matrix Q, which is only assumed to be irreducible, but not necessarily diagonalisable. Furthermore, we study the process where the mutation rate matrix is *parent independent*. In this case the Moran process is reversible and one can explicitly compute the eigenvectors of its generator and its stationary distribution. Then, we go further and prove the existence of cutoff phenomenon in the chi-square distance (Theorem 8) and in total variation (Theorem 9).

In Chapter 3 we also consider the process where the state space of the allelic types is finite and the selection rate is uniform, and moreover the mutation rate matrix is the generator of an asymmetric random walk on the cycle graph $\mathbb{Z}/K\mathbb{Z}$, with $K \in \mathbb{N}$. We show that this model has a remarkable exact solvability, despite the fact that it is non-reversible with non-explicit invariant distribution. Our main results include quantitative propagation of chaos and exponential ergodicity with explicit constants, as well as formulas for covariances at equilibrium in terms of the Chebyshev polynomials. We also obtain an explicit uniform in time bound for the convergence of the proportion of particles in each state when the number of particles goes to infinity.

We next summarise the main results of the thesis, which are presented in more details and proved later in Chapters 1, 2 and 3.

Propagation du chaos dans le modèle de Moran

Here we present the main results obtained in Chapter 1, which are based on the preprint [CC21].

In this chapter we consider the problem of studying the convergence of the empirical distribution induced by the Moran model when the number of individual tends toward infinity. Naturally, we assume that the empirical measure induced by the particle process at $t = 0$ converges towards the initial distribution $\mu_0 \in \mathcal{M}_1(E)$ in $\mathbb{L}^p$, for every $p \geq 1$. More precisely, there exists a constant $C > 0$ such that

$$\sup_{\phi \in \mathcal{B}_1(E)} \mathbb{E}[|m(\eta_0)(\phi) - \mu_0(\phi)|^p] \leq \frac{C}{N^{p/2}}, \tag{9}$$

where $\mathcal{B}_1(E)$ is the set of real function on E such that $\|\phi\| \leq 1$. In particular, this verified when initially all the particles are sampled independently with the same distribution $\mu_0 \in \mathcal{M}_1(E)$. Then, we focus on the control of the following quantity

$$\sup_{\phi \in \mathcal{B}_1(E)} \mathbb{E}\left[\sup_{t \in A} |m(\eta_t^{(N)})(\phi) - \mu_t(\phi)|^p\right],$$

for every $p \geq 1$, where $(\mu_t)_{t \geq 0}$ is the solution of (4) with initial condition $\mu_0 \in \mathcal{M}_1(E)$ and $A \subset \mathbb{R}_+$. This is a control of the speed of convergence of $m(\eta_t^{(N)})(\phi)$ to $\gamma_t(\phi)$, when $N \to \infty$, as in (5). We consider to cases: where $A = [0, T]$, with $T \geq 0$, and thus it can be extended to any compact on $\mathbb{R}_+$; or when $A = \mathbb{R}_+$. In this last case, our results provide a uniform in time control of the $\mathbb{L}^p$ distance between $m(\eta_t^{(N)})(\phi)$ and $\mu_t(\phi)$, for any $\phi \in \mathcal{B}_1(E)$.

Propagation of chaos with general selection rate

First, we consider a model where the selection rates can be written as follows

$$V_\mu(x, y) = \sum_{i \geq 1} V_i^{\mathrm{d}}(x) V_i^{\mathrm{b}}(y) + V_\mu^{\mathrm{s}}(x, y), \tag{10}$$

where V_i^{d} and V_i^{b} are uniformly bounded in E and $\mu \mapsto V_\mu^{\mathrm{s}}$ is a continuous and bounded from $(\mathcal{M}_1(E), \|\cdot\|_{\mathrm{TV}})$ to $(\mathcal{B}_b(E \times E), \|\cdot\|)$. Furthermore, V_μ^{s} is symmetric for every $\mu \in \mathcal{M}_1(E)$. See Assumptions (G1) and (G2) in Section 1.1.1 for a thorough version the hypothesis we assume in this case. Hence, we prove a uniform on compacts convergence in the sense of the following inequality: for every $T \geq 0$ and $p \geq 1$, there exists a constant $C_{p,T} > 0$, such that

$$\sup_{\phi \in \mathcal{B}_1(E)} \mathbb{E}\left[\sup_{t \in [0,T]} |m(\eta_t^{(N)})(\phi) - \mu_t(\phi)|^p\right]^{1/p} \leq \frac{C_{p,T}}{\sqrt{N}}, \tag{11}$$

where $(\mu_t)_{t \geq 0}$ is the solution of (2) with initial condition $\mu_0 \in \mathcal{M}_1(E)$. See Theorem 1.1.2 for the complete description of this result.

This result is a generalisation for multi-allelic Moran models of Proposition 3.1 in [Cor17], where the uniform convergence on compacts in probability is proved for a bi-allelic Moran model.

This speed of convergence can also be related to existing results which ensure the convergence of the empirical mesure induced by a Moran type (or Fleming – Viot) particle process towards the law of an absorbing process conditioned to non absorption. See for instance [DM00b, Prop. 3.5] [DPR11, Lemma 3.1], [Vil14, Thm. 2.2] and [CT16a, Thm. 1.3]. See also [BC15, Thm. 3.1 and Rmk. 3.2] where the almost surely convergence (and also the complete convergence [Gut13, Def. 1.6]) is proved when the state space is finite. As far as we know, the control we establish in (11) is the first result ensuring the convergence uniformly in compacts, for $p \geq 1$, with speed of convergence of order $1/\sqrt{N}$ for multi-allelic Moran models, and also for Fleming – Viot particle processes in discrete countable state spaces, not necessarily finite. The idea behind the proof is closed to the methods in [Rou06]: it consists in finding a martingale indexed by the interval $[0, T]$, whose terminal value at time T is precisely $m(\eta_T)(\phi) - \mu_T(\phi)$ plus a term whose $\mathbb{L}^p$-norm can be controlled, for any $\phi \in \mathcal{B}_d(E)$. Thereafter, the final result comes by clever use of a Grönwall type argument similarly to the proof of Proposition 1 in [MS19].

Uniform in time propagation of chaos with additive selection rates

Under a more specific expression for the selection rate, we can prove a uniform in time bound for the convergence of $(m(\eta_t))_{t\geq 0}$ towards $(\mu_t)_{t\geq 0}$, when $N \to \infty$. Namely, we assume that (3) is verified, i.e. we assume that the selection rates can be written as follows

$$V_\mu(x, y) = V_\mu^{\mathrm{d}}(x) + V_\mu^{\mathrm{b}}(y) + V_\mu^{\mathrm{s}}(x, y),$$

and that there exist $\Lambda \in \mathcal{B}_b(E)$, independent on μ, and a constant C_μ, depending on μ, such that

$$\Lambda = V_\mu^{\mathrm{b}} - V_\mu^{\mathrm{d}} + C_\mu, \text{ for all } \mu.$$

See Assumption (C1) in Chapter 1 for more details.

Example 1 (Selection rates independent on μ)**.** When the selection rates do not depend on μ, the expression for the selection rate reduces to the existence of $V^{\mathrm{d}}, V^{\mathrm{b}} \in \mathcal{B}_d(E)$ and a symmetric $V^{\mathrm{s}} \in \mathcal{B}_d(E \times E)$ such that

$$V(x, y) = V^{\mathrm{d}}(x) + V^{\mathrm{b}}(y) + V^{\mathrm{s}}(x, y).$$

Let $\Lambda \in \mathcal{B}_b(E)$ be a fixed function. Typical examples of functions V^{b} and V^{d} satisfying this condition are

$$V^{\mathrm{b}} = (\Lambda - c)^+ \quad \text{and} \quad V^{\mathrm{d}} = (\Lambda - c)^-,$$

for a fixed constant $c \in \mathbb{R}$, where we use the standard notation

$$(x)^+ := \max\{x, 0\} \quad \text{and} \quad (x)^- := -\min\{x, 0\}.$$

These are in fact the selections rates considered by Angeli et al. [AGJ21, §3.3] in the context of cloning algorithms. Moreover, the case $c = 0$ is considered in Example 3.1-(2) in [Rou06]. Note that in this case the expression in (10) is also verified.

From a biological point of view, the parameter $c \in \mathbb{R}$ can be seen as a fitness parameter. Let us assume V^{s} is null for simplicity, and denote by $\xi_t^{(i)}$ the type of the i-th individual, for $1 \leq i \leq N$, at time $t \geq 0$. Then, if $\Lambda(\xi_t^{(i)}) \leq c$, the individual i-th dies and another randomly chosen reproduces with rate $(\Lambda(\xi_t^{(i)}) - c)^-$. Otherwise, a random chosen individual individual dies and individual i-th reproduces with rate $(\Lambda(\xi_t^{(i)}) - c)^+$.

Another example of particular interest is when $V^{\mathrm{b}} = 0$. Notice that the Moran process with this selection rates is in fact a Fleming – Viot particle system .

Example 2 (Selection rates depending on μ)**.** Consider a fixed function $\Lambda \in \mathcal{B}_b(E)$. Typical examples of functions V_μ^{b} and V_μ^{d} are:

$$V_\mu^{\mathrm{b}} = (\Lambda - \mu(\Lambda))^+ \quad \text{and} \quad V_\mu^{\mathrm{d}} = (\Lambda - \mu(\Lambda))^-.$$

These are the selection rates considered in [Del04, §1.5.2, p. 35], see also Example 3.1-(3) in [Rou06].

In this case the biological interpretation of $\mu(\Lambda)$ is similar to that of the parameter c in the Example 1. Indeed, the fitness parameter varies in time according to the configuration of the population.

We are now in conditions to state our main results for the multi-allelic Moran model with additive selection.

Theorem 2 (Uniform in time propagation of chaos)**.** *Assume that*

- *the expression* (3) *for the selection rates,*
- *the assumption* (9) *for the initial condition, and*
- *the uniform convergence established in* (8) *for* $(\mu_t)_{t\geq 0}$

are verified. Then, for every $p \geq 1$ *there exists a constant* C_p*, such that*

$$\sup_{\phi\in\mathcal{B}_1(E)} \sup_{t\geq 0} \mathbb{E}\left[|m(\eta_t^{(N)})(\phi) - \mu_t(\phi)|^p\right]^{1/p} \leq \frac{C_p}{\sqrt{N}}.$$

See Theorem 1.1.4 in Chapter 1 for a detailed version of this results and its proof.

For a fixed N, if the process $(\eta_t^{(N)})_{t\geq 0}$ generated by $\mathcal{Q}_N$ allows a stationary distribution ν_N, then under the hypothesis in Theorem 2 we get

$$\sup_{\phi\in\mathcal{B}_1(E)} \mathbb{E}_{\nu_N}\left[|m(\eta_\infty)(\phi) - \mu_\infty(\phi)|^p\right]^{1/p} \leq \frac{C_p}{\sqrt{N}},$$

for all $p \geq 1$.

Obtaining a uniform in time bound as the one provided by Theorem 2 is a hard problem and this kind of results are uncommon in the literature. Del Moral and Guionnet in [DG01, Thm. 3.1] have proved a similar results for a resembling but discrete-time model, where the potential is assumed uniformly bounded and also bounded away from zero. Moreover, their upper bound for the speed of convergence in [DG01, Thm. 3.1] is of order $1/N^\alpha$, with $\alpha < 1/2$. Rousset [Rou06, Thm. 4.1] has proved a uniform in time bound in $\mathbb{L}^p$ with the same speed of convergence as our result. However, the model studied by Rousset is in continuous state space and the diffusion process driving the mutation process is assumed reversible. Similarly, Angeli et al. [AGJ21, Thm. 3.2] obtained an equivalent result for jump processes on locally compact spaces in the context of cloning algorithms, for $p \geq 2$. Our model is different, since we consider the case where the state space is discrete, non necessarily compact. Furthermore, in (3) we allow the selection rates to depend on the empirical probability measure induced by the particle system, in the same spirit of [Rou06]. Nonetheless, our methods are similar to those of Rousset [Rou06] and Angeli et al. [AGJ21] (see also [DM00b, §3.3.1]): they consist in finding a martingale indexed by the interval $[0, T]$, wwhose terminal value at time T is precisely $m(\eta_T)(\phi) - \mu_T(\phi)$ plus a term whose $\mathbb{L}^p$ norm can be controlled, for any $\phi \in \mathcal{B}_b(E)$. Thereafter, the final result comes by a control of the quadratic variation of the martingale and an induction principle.

We also get the next result: under the hypothesis in Theorem 2, for every $p \geq 1$, there exists a constant $C_p > 0$ such that

$$\sup_{t\geq 0} \mathbb{E}\left[\left(\|m(\eta_t^{(N)}) - \mu_t\|_{\mathrm{w}}\right)^p\right]^{1/p} \leq \frac{C}{\sqrt{N}}, \quad \text{where} \quad \|\mu_1 - \mu_2\|_{\mathrm{w}} := \sum_{n\in M} 2^{-n}|\mu_1(x_n) - \mu_2(x_n)|,$$

for an (arbitrary but fixed) enumeration $(x_n)_{n\in M}$ of the elements in E, with $M \subset \mathbb{N}$.

Remark 5 (Almost sure convergence). The previous inequality, for $p = 4$, ensure the convergence:

$$\sum_{N\geq 2} \mathbb{P}\left[\|m(\eta_t^{(N)}) - \mu_t\|_{\mathrm{w}} > \epsilon\right] < \infty,$$

for every $\epsilon > 0$. Hence, using a Borel – Cantelli argument we get the convergence $m(\eta_T^{(N)}) \xrightarrow{\text{a.s.}} \mu_T$ in the weak sense, where a.s. denotes the almost sure convergence. Furthermore, this convergence does not depend on the space where the random variables are coupled.

Let us denote by $\bar{m}(\eta_t)$ the mean empirical probability measure induced by η_t which is defined as

$$\bar{m}(\eta_t) := \sum_{x\in E} \mathbb{E}\left[\frac{\eta_t(x)}{N}\right] \delta_x \in \mathcal{M}_1(E).$$

Moreover, let us denote by $\xi_t^{(i)}$ the random variable given by the type of the i-th individual at time $t \geq 0$, for $i \in \{1, 2, \dots, N\}$, and by $\mathrm{Law}(\xi_t^{(i)})$ the law of this random variable. Let us denote by $(\mathrm{e}^{tQ})_{t\geq 0}$ the semigroup associated to the mutation process generated by Q. Hence, under the hypothesis in the statement of Theorem 2 we can ensure the existence of a positive constant C such that

$$\sup_{t\geq 0} \left\|\bar{m}(\eta_t^{(N)}) - \mu_t\right\|_{\mathrm{TV}} \leq \frac{C}{N}.$$

Moreover, if the initial distribution of the N particles is exchangeable, then

$$\sup_{t\geq 0} \left\|\mathrm{Law}(\xi_t^i) - \mu_t\right\|_{\mathrm{TV}} \leq \frac{C}{N}.$$

It is expected that when the selection rates are constant the measure generated by the particle system is a unbiased estimator of the law of the Markov chain generated by Q, in the sense that

$$\bar{m}(\eta_t^{(N)}) = \bar{m}(\eta_0)\mathrm{e}^{tQ}, \text{ for all } t \geq 0.$$

See e.g. [CT16b] and the results in Chapter 3. In Corollary 1.3.6, we prove that this result also holds when $V = V^{\mathrm{s}}$ is symmetric.

Our last two results in Chapter 1 are addressed to the study of the asymptotic square error of the approximation of μ_∞ by $m(\eta_T^{(N)})$ when $T, N \to \infty$. These results are highly important when the Moran process is used for approximating a quasi-stationary distribution. First, we provide an explicit expression for the asymptotic expression of $N\mathbb{E}\left[\left(m(\eta_T^{(N)}) - \mu_\infty(\phi)\right)^2\right]$. Let us define the asymptotic quadratic errors:

$$\sigma_T^2(\phi) := \lim_{N\to\infty} N\mathbb{E}\left[\left(m(\eta_T)(\phi) - \mu_T(\phi)\right)^2\right] \quad \text{and} \quad \sigma_\infty^2(\phi) := \lim_{T\to\infty} \sigma_T^2(\phi),$$

for every uniformly bounded function ϕ in E.

First, we prove the asymptotic normality of the bias and we provide explicit expressions for $\sigma_T^2(\phi)$ and $\sigma_\infty^2(\phi)$. Then, we use this expression to show how to define another Moran process approaching the same distribution μ_∞, with smaller or equal asymptotic square error.

Let us define

$$S_\mu(\phi) := \sum_{x,y\in E} \mu(x)\mu(y)V^{\mathrm{s}}_\mu(x,y)(\phi(x)-\phi(y))^2, \tag{12}$$

for every $\phi \in \mathcal{B}_b(E)$ and $\mu \in \mathcal{M}_1(E)$. We thus obtain the following result:

Theorem 3 (Asymptotic normality)**.** *Suppose that initially the N particles are independently sampled according to $\mu_0 \in \mathcal{M}_1(E)$, and the assumptions in the statement of Theorem 2 are verified. Then, for every $\phi \in \mathcal{B}_b(E)$ and $T \geq 0$, we have that $\sqrt{N}(m(\eta_T)(\phi) - \mu_T(\phi))$ converges in law towards a Gaussian centered random variable of variance $\sigma^2_T(\phi)$, when N goes to infinity. Moreover,*

$$\sigma^2_\infty(\phi) = \mathrm{Var}_{\mu_\infty}(\phi) + \int_0^\infty \mathrm{e}^{-2\lambda s} S_{\mu_\infty}(P^\Lambda_s(\bar{\phi}_\infty))\mathrm{d}s + 2\int_0^\infty \mathrm{e}^{-2\lambda s}\mu_\infty\Big(P^\Lambda_s(\bar{\phi}_\infty)^2\Big(V^{\mathrm{b}}_{\mu_\infty} + \mu_\infty(V^{\mathrm{d}}_{\mu_\infty})\Big)\Big)\mathrm{d}s,$$

where $\mathrm{Var}_{\mu_\infty}$ stands for the variance with respect to μ_∞, $\bar{\phi}_\infty := \phi - \mu_\infty(\phi)$, λ is the eigenvalue in the statement of Lemma 1 and S_μ is as defined in (12).

See Theorem 1.1.8 for a detailed version of this result, including an explicit expression for $\sigma^2_T(\phi)$, for every $T > 0$. Note that the two integrals in the expression of $\sigma^2_\infty(\phi)$ in Theorem 3 converge as a consequence of Lemma 1.

When V^{b}_μ and V^{s}_μ are null and thus $\Lambda = -V^{\mathrm{d}} \leq 0$, we get

$$\sigma^2_\infty(\phi) = \mathrm{Var}_{\mu_\infty}(\phi) - 2\lambda\int_0^\infty \mathrm{e}^{-2\lambda s}\,\mathrm{Var}_{\mu_\infty}\left(P^\Lambda_s(\phi)\right)\mathrm{d}s.$$

When the process $(\eta_t)_{t\geq 0}$ is ergodic and converges in law to some random variable η_∞, when $t \to \infty$, Theorem 3 states that $\sqrt{N}(m(\eta_\infty)(\phi) - \mu_\infty(\phi))$ converges to a centered Gaussian law of variance $\sigma^2_\infty(\phi)$, when $N \to \infty$. Indeed, recall that a Gaussian sequence converges in law if their first two moments converge. In particular, we recover (and extend) the recent result of Lellièvre et al. [LPR18, Thm. 2.4] for finite state spaces. Notice that the negative constant λ in the previous expression is the opposite of that in [LPR18, Thm. 2.4].

Note that the three summands on the expression of $\sigma^2_\infty(\phi)$ are positive. Moreover, the limit $(\mu_t)_{t\geq 0}$ is invariant by the choice of the symmetric component V^{s}_μ in (3), as commented in Remark 1. As a consequence, for a given selection rate V_μ we can obtain another Moran process approaching the same limit distribution taking the selection rate $V_\mu - \Sigma_\mu \geq 0$, where Σ_μ is a symmetric function in $\mathcal{B}_b(E \times E)$.

Corollary 4 (Moran process with smaller asymptotic square error)**.** *Suppose that initially the N particles are independently sampled according to $\mu_0 \in \mathcal{M}_1(E)$, and the assumptions in the statement of Theorem 2 are verified. Let $(\eta_t)_{t\geq 0}$ and $(\eta^\star_t)_{t\geq 0}$ be the Moran processes with the same mutation rates and selection rates given by V_μ and $V_\mu - \Sigma_\mu$, respectively, where*

$$\Sigma_\mu(x,y) := \min\left\{V^{\mathrm{d}}_\mu(x), V^{\mathrm{b}}_\mu(x)\right\}\mathbf{1}_{\{x\}} + \min\left\{V^{\mathrm{d}}_\mu(y), V^{\mathrm{b}}_\mu(y)\right\}\mathbf{1}_{\{y\}} + V^{\mathrm{s}}_\mu(x,y),$$

where $\mathbf{1}_A$, stands for the indicator function on $A \subset E$. Then,

$$\lim_{T\to\infty}\lim_{N\to\infty} N\mathbb{E}\left[\left(m(\eta^\star_T)(\phi) - \mu_\infty(\phi)\right)^2\right] \leq \lim_{T\to\infty}\lim_{N\to\infty} N\mathbb{E}\left[\left(m(\eta_T)(\phi) - \mu_\infty(\phi)\right)^2\right].$$

See Corollary 1.1.9 for a detailed version of this result. Note that the selection rate $V_\mu - \Sigma_\mu$ in the statement of Corollary 4 satisfies (3).

Spectrum and ergodicity of the neutral Moran model

Here we present the main results obtained in the preprint [Cor21b], which are presented in detail in Chapter 2.

During this section we will consider that $E = [K]$, where $[K] := \{1, 2, \ldots, K\}$. The definition of the state space of the multi-allelic Moran reduces to the K-dimensional N-discrete simplex:

$$\mathcal{E}_{K,N} := \left\{ \eta \in [N]_0^K : \ |\eta| = N \right\},$$

where $[N]_0 := \{0, 1, \ldots, N\}$ and $|\cdot|$ stands for the sum of elements in a vector. The set $\mathcal{E}_{K,N}$ has cardinality of $\binom{K-1+N}{N}$.

According to the notation previously introduced, the mutation rate matrix in this case is the matrix $Q = (Q_{i,j})_{i,j=1}^K$. Moreover, the process is neutral, i.e. the selection rates are constant and equal to $p \geq 0$. Let us denote by $\mathcal{Q}_{N,p}$, the infinitesimal generator of the neutral multi-allelic Moran process in this case, which acts on a real function f on $\mathcal{E}_{K,N}$ as follows:

$$(\mathcal{Q}_{N,p} f)(\eta) := \sum_{i,j \in [K]} \eta(i) \left(Q_{i,j} + \frac{p}{N} \eta(j) \right) [f(\eta - \mathbf{e}_i + \mathbf{e}_j) - f(\eta)], \tag{13}$$

for all $\eta \in \mathcal{E}_{K,N}$, where $\mathbf{e}_k$ is the k-th canonical vector of $\mathbb{R}^K$.

In words, $\mathcal{Q}_{N,p}$ drives a process of N individuals, where each individual has one of K possible types of alleles and where the type of the individual changes following two processes: a mutation process where individuals mutate independently of each other and a Moran type reproduction process, where the individuals interact. The N individuals mutate independently from type $i \in [K]$ to type $j \in [K] \setminus \{i\}$ with rate $Q_{i,j}$. In addition, with uniform rate $p \geq 0$, one of the N individuals is uniformly chosen to be removed from the population and another one, also randomly chosen, is duplicated. Note that the transitions of an individual due to a reproduction is not independent of the position of the other individuals. As in the original model, introduced by Moran [Mor58], the same individual removed from the population can be duplicated, in this case the state of the system does not change. In the instance where the removed individual cannot be duplicated, the factor $\frac{p}{N}$ in (13) must be replaced by $\frac{p}{N-1}$. The process generated by $\mathcal{Q}_{N,p}$ can be also seen as a Fleming–Viot particle system, but with uniform killing, i.e., with constant killing rate κ.

Our main result in this section is a complete description of the spectrum of $\mathcal{Q}_{N,p}$, which is expressed in the following theorem.

Theorem 5 (Spectrum of $\mathcal{Q}_{N,p}$)**.** *Assume $K \geq 2$, $N \geq 1$ and $p \in [0, \infty)$. Let us denote by λ_k, $k \in [K-1]$, the nonzero $K-1$ roots, counting algebraic multiplicities, of the characteristic polynomial of Q. For any $\eta \in \bigcup_{L=1}^{N} \mathcal{E}_{K-1,L}$, let us define*

$$\lambda_{\eta,p} := \sum_{k=1}^{K-1} \eta(k) \lambda_k - \frac{p}{N} |\eta|(|\eta| - 1).$$

Then, the eigenvalues of $\mathcal{Q}_{N,p}$, counting algebraic multiplicities, are 0 *and $\lambda_{\eta,p}$, for $\eta \in \bigcup_{L=1}^{N} \mathcal{E}_{K-1,L}$.*

The proof of Corollary 5 is given in Section 2.3.3. Note that the spectrum of $Q_{N,p}$ is given in function of the spectrum of Q, and it holds even when Q is not diagonalisable. Our results also include a study of the eigenvectors of $\mathcal{Q}_{N,p}$. In particular, we are able to provide explicit expressions for the eigenvectors of $\mathcal{Q}_{N,p}$ associated to the eigenvalues of smaller modulus. The expression we provide are based on the eigenvectors of Q. See Theorem 2.1.3 in Chapter 2 for more details.

Remark 6 (Monotony in N of the spectrum of $\mathcal{Q}_N$ $(p=0)$)*.* Theorem 5 implies that the spectrum of $\mathcal{Q}_N$, when $p=0$, for a fixed value of K, is an increasing function of N in the sense of the inclusion of sets.

Ergodicity of the neutral multi-allelic Moran process

The relation between the spectral properties of $\mathcal{Q}_{N,p}$ and Q can be used to estimate the speed of convergence to stationarity of the Moran process.

We are interested in the relationship between the spectrum of an infinitesimal rate matrix and the convergence to stationarity of the Markov process it drives. Let us define the *maximum total variation distance* to stationarity of the process driven by a infinitesimal generator L of a Markov chain on a discrete space Ω, denoted $\mathrm{D}^L_{\mathrm{TV}}$, as follows:

$$\mathrm{D}^L_{\mathrm{TV}}(t) := \max_{\mu\in\mathcal{M}_1(\Omega)} \left\|\mu\, \mathrm{e}^{tL} - \pi\right\|_{\mathrm{TV}}, \tag{14}$$

where the maximum runs over all possible initial distributions on Ω. Using the convexity of the total variation norm, we can prove that $\mathrm{D}^L_{\mathrm{TV}}(t) = \frac{1}{2}\|\mathrm{e}^{tL} - \Pi\|_\infty$, where Π stands for the matrix with every row equal to π, and $\|\cdot\|_\infty$ denotes the infinity norm of matrices (cf. [LP17, Ch. 4]).

As a consequence of Theorem 5 and other results in Chapter 2, the second largest eigenvalue in modulus (SLEM) of $\mathrm{e}^{t\mathcal{Q}_{N,p}}$ is equal to that of e^{tQ}, for all $t\geq 0$. The SLEM of the generator of a finite Markov chain is useful to study the asymptotic convergence in total variation. Hence, as an application of this Theorem 5, we study the ergodicity of the process driven by $\mathcal{Q}_{N,p}$ in total variation using the spectral properties of Q.

For a real positive function f we denote by $\mathcal{O}(f)$ another real positive function such that $C_1 f(t) \leq \mathcal{O}(f)(t) \leq C_2 f(t)$, for two constants $0 < C_1 \leq C_2 < \infty$ and for all $t \geq T$, for $T > 0$ large enough.

Corollary 6 (Asymptotic exponential ergodicity in total variation)**.** *Let us assume the* SLEM *of* e^{tQ} *equals* $\mathrm{e}^{-\rho t}$ *and let* $s\in\mathbb{N}$ *be the largest multiplicity in the minimal polynomial of* e^{tQ} *of all the eigenvalues with modulus* $\mathrm{e}^{-\rho t}$*. Then,*

$$\mathrm{D}^{\mathcal{Q}_{N,p}}_{\mathrm{TV}}(t) = \mathcal{O}\Big(\mathrm{D}^{Q}_{\mathrm{TV}}(t)\Big) = \mathcal{O}(t^{s-1}\mathrm{e}^{-\rho t}).$$

Corollary 6 is proved in Section 2.4.

The asymptotic expression in Corollary 6 hides the relation among the mixing time of the Markov chain and the number of individuals in the population and the selection parameter p. However, if Q has a real eigenvalue $-\lambda < 0$, we can go further and prove the following lower bound for the convergence in total variation to stationarity at time $\frac{\ln N - c}{2\lambda}$, for every $c\geq 0$.

Theorem 7 (Lower bound for convergence in total variation)**.** *Assume* $K\geq 2$*,* $N\geq 2$ *and* $p\in[0,\infty)$ *and let* $-\lambda<0$ *be an eigenvalue of* Q*. Then,*

$$\mathrm{D}^{\mathcal{Q}_{N,p}}_{\mathrm{TV}}\left(\frac{\ln N - c}{2\lambda}\right) \geq 1 - \kappa\mathrm{e}^{-c}.$$

In Theorem 2.1.5 we state a more detailed version of Theorem 7, which will be proved in Section 2.4.1.

The lower bound provided by Theorem 7 ensures that the mixing time of the neutral multi-allelic Moran model is at least of order of $\ln N/2\lambda$. Our results do not allow us to prove an upper bound ensuring the existence of a *cutoff phenomenon*. A further study needs to be done in this direction. However, for the parent independent mutation scheme, a further analysis can be done to prove the existence of a cutoff phenomenon in the chi-square and total variation distances, as we next discuss.

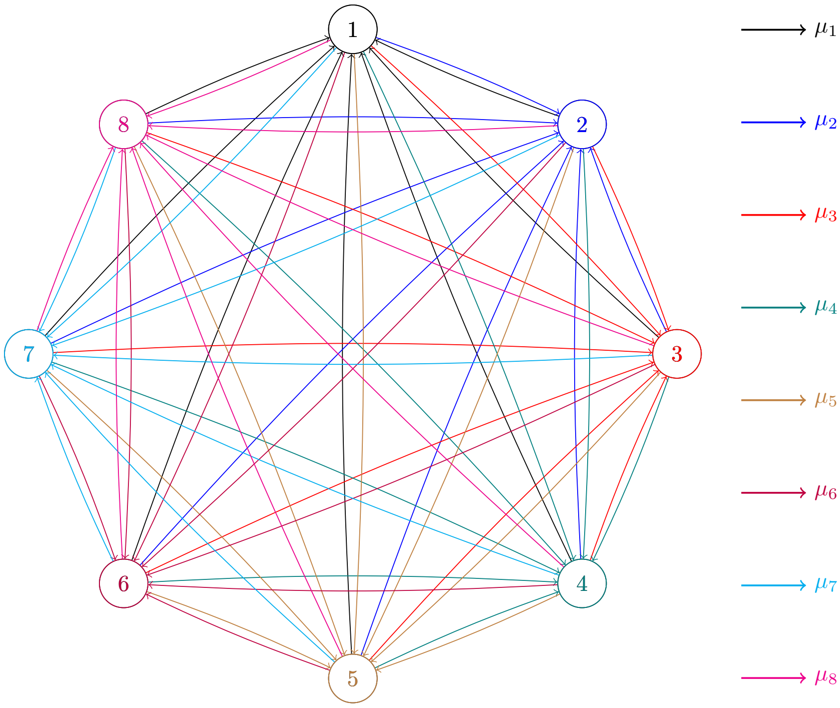


Figure 6: Complete graph induced by the infinitesimal generator $Q_{\boldsymbol{\mu}}$ (15) in $[K]$. In this example $K = 10$, the color of each arrow represents a different a rate μ_i, for $1 \leq i \leq K$, which in principle are different. Note that the color of the arrows only depends on the arrival site.

The neutral multi-allelic Moran model with parent independent mutation

Consider the following mutation rate matrix:

$$Q_{\boldsymbol{\mu}} := \begin{pmatrix} -|\boldsymbol{\mu}| + \mu_1 & \mu_2 & \mu_3 & \cdots & \mu_K \\ \mu_1 & -|\boldsymbol{\mu}| + \mu_2 & \mu_3 & \cdots & \mu_K \\ \mu_1 & \mu_2 & -|\boldsymbol{\mu}| + \mu_3 & \cdots & \mu_K \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu_1 & \mu_2 & \mu_3 & \cdots & -|\boldsymbol{\mu}| + \mu_K \end{pmatrix}, \tag{15}$$

where $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_K) \in (0, \infty)^K$ and $|\boldsymbol{\mu}|$ stands for the sum of the entries of $\boldsymbol{\mu}$. Let us define

$$(\mathcal{L}_{N,p} f)(\eta) := \sum_{i,j=1}^{K} \eta(i) \left(\mu_j + \frac{p}{N} \eta(j) \right) [f(\eta - \mathbf{e}_i + \mathbf{e}_j) - f(\eta)],$$

for every f on $\mathcal{E}_{K,N}$ and all $\eta \in \mathcal{E}_{K,N}$, the infinitesimal generator of the neutral multi-allelic Moran process with mutation rate matrix $Q_{\boldsymbol{\mu}}$. The process driven by $\mathcal{L}_{N,p}$ is a special case of the neutral multi-allelic Moran process considered before, but with the difference that the mutation rate only depends on the type of the new individual, i.e., the mutation process changes each type i individual to type j at rate μ_j, for all $i, j \in [K]$. This is the neutral multi-allelic Moran process with *parent independent* mutation (cf. [Eth11]). Note that the graph associated to $Q_{\boldsymbol{\mu}}$ is the complete graph, as show in Figure 6.

Using Theorem 5 we can provide the spectrum of $\mathcal{L}_{N,p}$. Indeed, for $K \geq 2$, $N \geq 2$ and $p \geq 0$, the infinitesimal generator $\mathcal{L}_{N,p}$ is diagonalisable with eigenvalues $\lambda_{n,p}$ with multiplicity $\binom{K+n-2}{n}$, where

$$\lambda_{n,p} := -|\boldsymbol{\mu}| n - \frac{p}{N} n(n-1),$$

for $n \in [N]_0$. In particular, the spectral gap of $\mathcal{L}_{N,p}$ is $\rho = |\boldsymbol{\mu}|$.

Remark 7 (Complete graph model). The *complete graph model* studied by Cloez and Thai [CT16a] in the context of Fleming – Viot particle processes is a particular case of the reversible process driven by $Q_{\boldsymbol{\mu}}$ above when $\mu_j = \frac{1}{K}$, for all $j \in [K]$. In this case, the eigenvalues of the mutation rate matrix are $\beta_0 = 0$ and $\beta_1 = -1$, this last one with multiplicity $K - 1$. In particular, the previous expression for the spectrum of $\mathcal{L}_{N,p}$ improves the Lemma 2.14 in [CT16a].

For $K \geq 2$, $N \geq 2$ and $p > 0$, in Chapter 2 we prove that the process driven by $\mathcal{Q}_{N,p}$ is reversible if and only if the mutation rate matrix has the form $Q_{\boldsymbol{\mu}}$ as in (15), for some vector $\boldsymbol{\mu}$, and consequently $\mathcal{Q}_{N,p}$ can be written as $\mathcal{L}_{N,p}$. Section 2.5 is devoted to the study of the spectral properties of $\mathcal{L}_{N,p}$, for $p \geq 0$, and its applications to the study of the convergence to stationarity. Our results in this section include a complete description of the set of eigenvalues and eigenfunctions of $\mathcal{L}_{N,p}$ and an explicit expression for its transition function. The eigenfunctions of $\mathcal{L}_{N,p}$, $p > 0$, are explicitly given in terms of *multivariate Hahn polynomials*, which are orthogonal with the stationary distribution, which follows a compound Dirichlet multinomial distribution (cf. [KM75; KZ09]). On the other hand, the eigenfunctions of $\mathcal{L}_N$, for $p = 0$, are explicitly given in terms of *multivariate Krawtchouk polynomials*, which are orthogonal with respect to stationary distribution, which follows a multinomial distribution (cf. [KM65; ZL09; DG14]).

Cutoff phenomenon

The *cutoff phenomenon* has been a rich topic of research on Markov chains since its introduction by the works of Aldous, Diaconis and Shahshahani in the 1980s (cf. [DS81; Ald83; AD86]). A Markov chain presents a cutoff if it exhibits an abrupt transition in its convergence to stationarity. Some of the most used notions of convergence are, as we consider here, the total variation distance and the chi-square distance (or divergence). A good introduction to this subject can be found in the book of Levin and Peres [LP17, Ch. 18] and in the work of Chen, Saloff-Coste et al. [Sal97; Che06; CS08; CS10; CHS17]. See the course of Salez [Sal21].

A typical scenario for the existence of a cutoff is a Markov chain with a high degree of symmetry. Hence, the cutoff phenomenon has been deeply studied for the movement on N independent particles on K sites, model which is usually known as *product chain*. Ycart [Yca99] studied the cutoff in total variation for N independent particles driven by a diagonalisable rate matrix. Later, Barrera et al. [BLY06] and Connor [Con10] studied the cutoff on this model according to other notions of distance. See also [Lac15] [LP17, Ch. 20], [CHS17] and [CK18] for more recent studies about the cutoff on product chains. The Moran model we consider here preserves the high level of symmetry of the product chain, but the movements of the particles are not independent. Indeed, the particles interact according to a reproduction process that favours the jumps to the sites with greater proportions of individuals.

Before formally defining the cutoff phenomenon, let us recall the *chi-square divergence* (sometimes called "distance"), which naturally arises in the context of reversible Markov chains (cf. [LP17, Ch. 12 and 20], [DP17, Ch. 8], [Bré20, Ch. 9.1]). Let Ω be a discrete set. The chi-square divergence of $\mu_2 \in \mathcal{M}_1(\Omega)$ with respect to the target distribution $\mu_1 \in \mathcal{M}_1(\Omega)$ is defined by

$$\chi^2(\mu_2 \mid \mu_1) := \sum_{\omega \in \Omega} \frac{[\mu_2(\omega) - \mu_1(\omega)]^2}{\mu_1(\omega)} = \|\mu_2 - \mu_1\|^2_{\frac{1}{\mu_1}},$$

where $\|\cdot\|_{\frac{1}{\mu_1}}$ stands for the norm in $l^2(\mathbb{R}^\Omega, \frac{1}{\mu_1})$, and $\frac{1}{\mu_1}$ is the measure $\omega \mapsto 1/\mu_1(\omega)$.

The chi-square divergence is not a metric, but a measure of the difference between two probability distributions. Note that the chi-square divergence, as well as the total variation distance, are special cases of the so called $f-$divergence functions, which measure the "difference" between two probability distributions. In this context, $\chi^2(\mu_2 \mid \mu_1)$ is also known as *Pearson chi-square divergence*.

Let us define the functions $\mathrm{d}_{\mathrm{TV}}(\cdot,\eta)$ and $\chi^2(\cdot,\eta)$, as follows

$$\mathrm{d}_{\mathrm{TV}}(t,\eta) := \|\delta_\eta \mathrm{e}^{t\mathcal{L}_{N,p}} - \nu_{N,p}\|_{\mathrm{TV}} = \frac{1}{2}\sum_{\xi\in\mathcal{E}_{K,N}} \left| \left(\mathrm{e}^{t\mathcal{L}_{N,p}}\delta_\xi\right)(\eta) - \nu_{N,p}(\xi)\right|,$$

$$\chi^2(t,\eta) := \chi^2(\delta_\eta \mathrm{e}^{t\mathcal{L}_{N,p}} \mid \nu_{N,p}) = \sum_{\xi\in\mathcal{E}_{K,N}} \frac{\left[\left(\mathrm{e}^{t\mathcal{L}_{N,p}}\delta_\xi\right)(\eta) - \nu_{N,p}(\xi)\right]^2}{\nu_{N,p}(\xi)},$$

for all $t \geq 0$.

The functions $\mathrm{d}_{\mathrm{TV}}(\cdot,\eta)$ and $\chi^2(\cdot,\eta)$ are thus measures of the convergence to stationary of the process driven by $\mathcal{L}_{N,p}$ at time t and with initial configuration $\eta \in \mathcal{E}_{K,N}$. In agreement with [Zho08; KZ09] we call them the total variation and the chi-square distances to stationarity, respectively.

As the number of individuals varies we obtain an infinite family of continuous-time finite Markov chains $\{(\mathcal{E}_{K,N}, \mathcal{L}_{N,p}, \nu_{N,p}), N \geq 2\}$. For each $N \geq 2$ let us denote by $\chi^2(t, N\mathbf{e}_k)$ (resp. $\mathrm{d}_{\mathrm{TV}}(t, N\mathbf{e}_k)$) the chi-square distance (resp. total variation distance) to stationarity of the process driven by $\mathcal{L}_{N,p}$ at time t, when the initial distribution is concentrated at $N\mathbf{e}_k \in \mathcal{E}_{K,N}$. Note that $\chi^2(0, N\mathbf{e}_k) \to \infty$ and $\mathrm{d}_{\mathrm{TV}}(0, N\mathbf{e}_k) \to 1$, when $N \to \infty$.

Definition 1 (Chi-square and total variation cutoff)**.** We say that $\{\chi^2(\cdot, N\mathbf{e}_k), N \geq 2\}$ exhibits a (t_N, ω_N) chi-square cutoff if $t_N \geq 0$, $\omega_N \geq 0$, $\omega_N = o(t_N)$ and

$$\lim_{c\to\infty} \limsup_{N\to\infty} \chi^2(t_N + c\,\omega_N, N\mathbf{e}_k) = 0, \quad \lim_{c\to-\infty} \liminf_{N\to\infty} \chi^2(t_N + c\,\omega_N, N\mathbf{e}_k) = \infty.$$

Analogously, we say that $\{\mathrm{d}_{\mathrm{TV}}(\cdot, N\mathbf{e}_k), N \geq 2\}$ exhibits a (t_N, ω_N) total variation cutoff if $t_N \geq 0$, $\omega_N \geq 0$, $\omega_N = o(t_N)$ and

$$\lim_{c\to\infty} \limsup_{N\to\infty} \mathrm{d}_{\mathrm{TV}}(t_N + c\,\omega_N, N\mathbf{e}_k) = 0, \quad \lim_{c\to-\infty} \liminf_{N\to\infty} \mathrm{d}_{\mathrm{TV}}(t_N + c\,\omega_N, N\mathbf{e}_k) = 1.$$

The sequences $(t_N)_{N\geq 2}$ and $(\omega_N)_{N\geq 2}$ are called *cutoff* and *window sequences*, respectively.

See Definition 2.1 and Remark 2.1 in [CS08].

The cutoff phenomenon describes an abrupt transition in the convergence to stationarity: over a negligible period given by the window sequence $(\omega_N)_{N>2}$, the distance from equilibrium drops from near its initial value to near zero at a time given by the cutoff sequence $(t_N)_{N\geq 2}$.

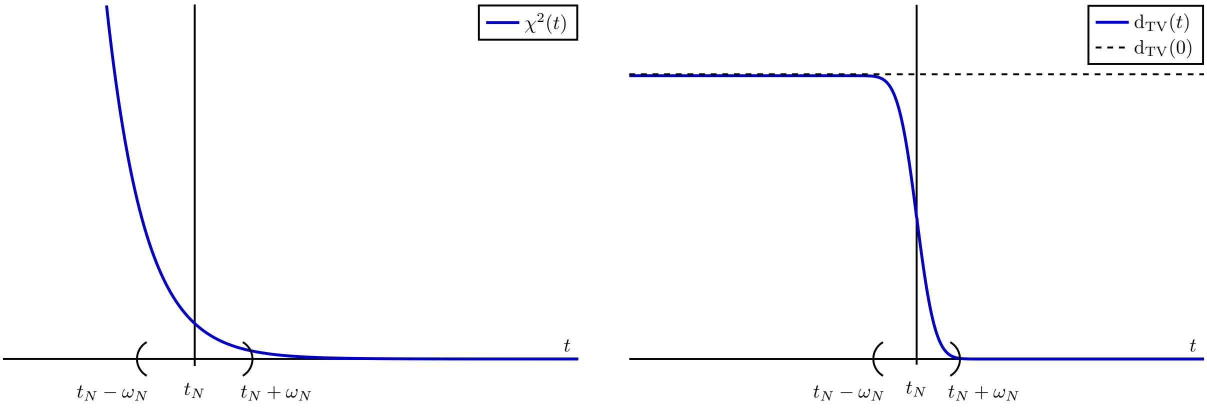
A stronger condition for the existence of a (t_N, ω_N) chi-square cutoff (resp. total variation cutoff) is the existence of the limit

$$G_k(c) := \lim_{N\to\infty} \chi^2(t_N + c\,\omega_N, N\mathbf{e}_k) \quad \left(\text{resp. } H_k(c) := \lim_{N\to\infty} \mathrm{d}_{\mathrm{TV}}(t_N + c\,\omega_N, N\mathbf{e}_k)\right), \tag{16}$$

for a function G_k (resp. H_k), for $k \in [K]$, satisfying:

$$\lim_{c\to-\infty} G_k(c) = \infty \text{ and } \lim_{c\to\infty} G_k(c) = 0, \left(\text{resp. } \lim_{c\to-\infty} H_k(c) = 1 \text{ and } \lim_{c\to\infty} H_k(c) = 0\right).$$

Actually, in this case the (t_N, ω_N) cutoff is said to be *strongly optimal*, see for instance Definition 2.2 and Proposition 2.2 in [CS08]. See Sections 2.1 and 2.2 of [CS08] and Chapter



(a) Profile of a chi-square cutoff

(b) Profile of a total variation cutoff

Figure 7: Schemes of functions G_k (left panel) and H_k (right panel) as defined in (16). Both plots show an abrupt transition in the convergence to stationarity: over a negligible period given by the window sequence $(\omega_N)_{N>2}$, the distance from equilibrium drops from near its initial value to near zero at a time given by the cutoff sequence $(t_N)_{N\geq 2}$.

2 in [Che06] for more details about the definition of (t_N, ω_N) cutoff and window optimality. Figure 7 shows classic profiles of functions G_k and H_k as defined in (16) for total variation and chi-square cutoffs.

The next two results establish the existence of cutoff phenomena in the chi-square and the total variation distances for the multi-allelic Moran process driven by $\mathcal{L}_{N,p}$, for $p \geq 0$, when the initial distribution is concentrated at $N\mathbf{e}_k$, for $k \in [K]$. In the chi-square case we are able to explicitly provide the limit profile of the distance. Moreover, we prove the total variation distance to stationarity of the mutation process driven by $\mathcal{L}_N$, i.e. for $p = 0$, has a Gaussian profile, when all the individuals are initially of the same type.

Theorem 8 (Limit profil of the chi-square cutoff when $N \to \infty$)**.** *For $k \in [K]$, with $K \geq 2$, $p \geq 0$ and every $c \in \mathbb{R}$, we have*

$$\lim_{N\to\infty} \chi^2\left(t_{N,c}, N\mathbf{e}_k\right) = \exp\{K_{k,p}\mathrm{e}^{-c}\} - 1,$$

where $t_{N,c} = \dfrac{\ln N + c}{2|\boldsymbol{\mu}|}$ *and* $K_{k,p} = \dfrac{|\boldsymbol{\mu}|(|\boldsymbol{\mu}| - \mu_k)}{\mu_k(|\boldsymbol{\mu}| + p)}$*. Consequently, the Markov process driven by $\mathcal{L}_{N,p}$ has a strongly optimal $\left(\frac{\ln N}{2|\boldsymbol{\mu}|}, 1\right)$ chi-square cutoff when $N \to \infty$.*

Figure 8 illustrates the convergence of $\chi^2(t_{N,c}, N\mathbf{e}_k)$, for $t_{N,c} = \frac{\ln(N)+c}{2|\boldsymbol{\mu}|}$, towards $G_k(c) = \exp\{K_{k,p}\mathrm{e}^{-c}\} - 1$, when $N \to \infty$.

Theorem 9 (Total variation cutoff when $N \to \infty$)**.** *For every $k \in [K]$, with $K \geq 2$, $p \geq 0$ and every $c > 0$, we have*

$$\mathrm{d}_{\mathrm{TV}}\left(t_{N,c}, N\mathbf{e}_k\right) \geq 1 - 32|\mu|\kappa_k\mathrm{e}^{-c},$$
$$\lim_{N\to\infty} \mathrm{d}_{\mathrm{TV}}\left(t_{N,c}, N\mathbf{e}_k\right) \leq \sqrt{\exp\{K_{k,p}\mathrm{e}^{-c}\} - 1},$$

where $t_{N,c} = \dfrac{\ln N + c}{2|\boldsymbol{\mu}|}$, $\kappa_k = \max\limits_{r:r\neq k} \dfrac{\mu_r \wedge \mu_k}{\mu_k}$ *and* $K_{k,p} = \dfrac{|\boldsymbol{\mu}|(|\boldsymbol{\mu}| - \mu_k)}{\mu_k(|\boldsymbol{\mu}| + p)}$*. Consequently, the Markov process driven by $\mathcal{L}_{N,p}$ exhibits a $\left(\frac{\ln N}{2|\boldsymbol{\mu}|}, 1\right)$ total variation cutoff when $N \to \infty$.*

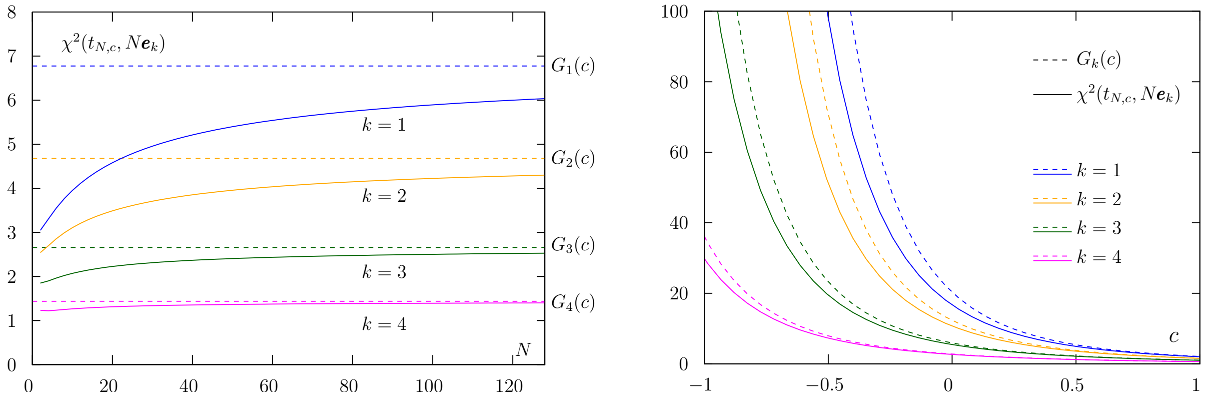


Figure 8: For parameters $K = 4$, $k \in [K]$, $\mu = (0.7, 0.8, 1.0, 1.3)$, $p = 1.7$: left panel shows $\chi^2(t_{N,c}, N\mathbf{e}_k)$ as a function of N, $2 \le N \le 128$ for $c = 0.4$, and right panel shows $G_k(c)$ and $\chi^2(t_{N,c}, N\mathbf{e}_k)$ as functions of c, with $-1 \le c \le 1$, for $N = 150$.

Moreover, when $p = 0$ the limit profile of the total variation distance satisfies

$$\lim_{N\to\infty} \mathrm{d_{TV}}(t_{N,c}, N\mathbf{e}_k) = 2\Phi\left(\frac{1}{2}\sqrt{K_{k,0}\mathrm{e}^{-c}}\right) - 1,$$

where Φ is the cumulative distribution function of the standard normal distribution. Thus, there exists a strongly optimal $\left(\frac{\ln N}{2|\boldsymbol{\mu}|}, 1\right)$ total variation cutoff for the process driven by $\mathcal{L}_N$ when $N \to \infty$.

Proof of Theorems 8 and 9 will be given in Section 2.5.1.

Figure 9 illustrates the convergence of $\mathrm{d_{TV}}(t_{N,c}, N\mathbf{e}_k)$, for $t_{N,c} = \frac{\ln(N)+c}{2|\boldsymbol{\mu}|}$, towards $H_k(c) = 2\Phi\left(\frac{1}{2}\sqrt{K_{k,0}\mathrm{e}^{-c}}\right) - 1$, when $N \to \infty$.

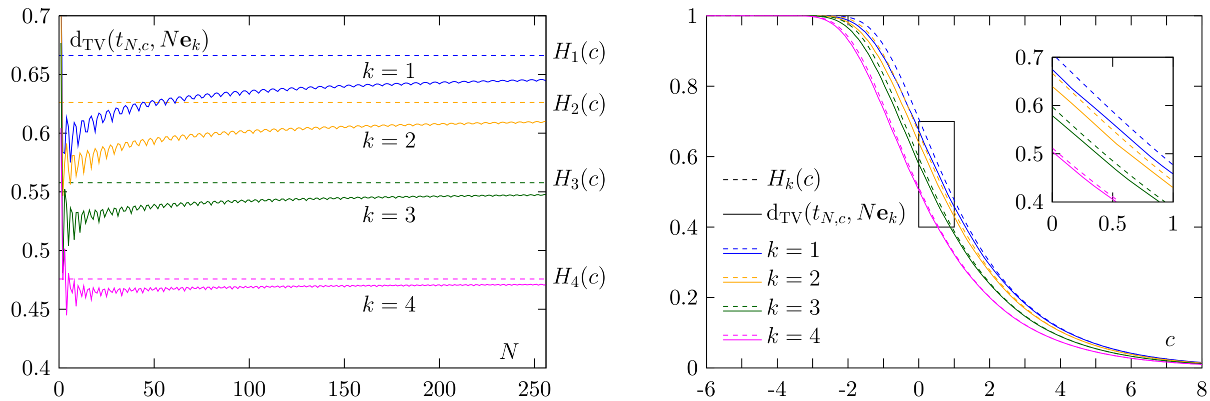


Figure 9: For parameters $K = 4$, $k \in [K]$, $\mu = (0.7, 0.8, 1.0, 1.3)$, $p = 0$: left panel shows $\mathrm{d_{TV}}(t_{N,c}, N\mathbf{e}_k)$ as a function of N, $2 \le N \le 256$ for $c = 0.17$ and right panel shows $H_k(c)$ and $\mathrm{d_{TV}}(t_{N,c}, N\mathbf{e}_k)$ as functions of c, with $-6 \le c \le 8$, for $N = 100$.

Some authors have studied the existence of a cutoff in Moran type models. For instance, Donelly and Rodrigues [DR00] proved the existence of a cutoff for the bi-allelic neutral Moran model in the separation distance. In order to do that, they used a duality property of the Moran process and found an asymptotic expression for the convergence in separation distance for a suitable scaled time, when the number of individuals tends to infinity. Khare and Zhou [KZ09] proved bounds for the chi-square distance in a discrete-time multi-allelic Moran process that implies the existence of a cutoff. Diaconis and Griffiths [DG19] studied the existence of a chi-square and total variation cutoffs for a discrete-time analogous of the mutation process generated by $\mathcal{L}_N$. Theorems 8 and 9 sharpen the results in [KZ09] and [DG19], since they provide the limit profiles for the chi-square and the total variation distances, for $p \geq 0$ and $p = 0$, respectively.

Theorem 9 is, as far as we know, the first result ensuring the existence of a total variation cutoff phenomenon for the neutral Moran model with $p > 0$ and parent independent mutation.

Quantitative bounds in a non-reversible example

Here we present the main results obtained in the article [Cor21a], published in *Stochastic Processes and Their Applications*, which are presented in detail in Chapter 3.

The main goal of Chapter 3 is to provide quantitative estimates for the convergences in scheme (5) for the neutral multi-allelic Moran model, where the mutation process is driven by an asymmetric random walk in the cycle graph. In particular, we provide explicit estimates for the bounds in Theorem 2 above.

As we previously commented, the convergence of the empirical distributions induced by a Moran (or Fleming – Viot) particle process defined on discrete state spaces, when the size of the population and the time increase have been assured under some assumptions. When the state space in finite, as in the model consider in Chapter 3, it is well-known that the Moran process is exponentially ergodic, and the propagation of chaos phenomenon is verified. See for examples the works of Asselah et al. [AFG11], Benaim and Cloez [BC15], Cloez and Thai [CT16b] and Lelièvre et al. [LPR18]. However, there are no many works dedicated to the study of specific examples and providing explicit bounds for the convergences, even considering neutrality as we do here. In this direction we can cite the work of Cloez and Thai [CT16a] and Del Moral and Jasra [DJ18].

Consider the quantity λ defined in [CT16b] as

$$\lambda = \inf_{x,y} \Big(Q_{x,y} + Q_{y,x} + \sum_{s \neq x,y} Q_{x,s} \wedge Q_{y,s} \Big), \tag{17}$$

where $Q = \left(Q_{x,y}\right)_{x,y}$ is the infinitesimal rate matrix of the process until absorption. When $\lambda = 0$ some of the results of [CT16b] do not hold and most of the bounds become too rough. Note that $\lambda > 0$ for the two examples studied in [CT16a], but λ is equal to zero for those models where there exist two vertices such that the distance between them is greater than two. The quantity λ is somehow related to the geometry of the graph associated to the mutation process of the Moran model. Hence, it becomes interesting to find explicit bounds for the speed of convergence of Moran processes with more complex geometries associated to their mutation processes.

In this section, and later in Chapter 3, we focus on the random walk on the cycle graph $\mathbb{Z}/K\mathbb{Z}$, for $K \geq 3$. Note that for this graph it holds that $\lambda = 0$ when $K \geq 6$. For simplicity, we assume that the N particles jump to the absorbing state with the same rate. Even if in this case the distribution of the conditional process is trivial, the study of the neutral Moran process becomes more complicated due to its non reversibility and the geometry of the cycle graph. We provide explicit bounds for the speed of the convergence of the empirical distribution induced by the particle system to the unique QSD when t and N tend to infinity. This example can be seen as a further step towards the study of the speed of convergence of Moran (or Fleming – Viot) processes with more general geometry.

Consider a Markov process $(Z_t)_{t\geq 0}$ with state space $\mathbb{Z}/K\mathbb{Z} \cup \{\partial\}$, where $K \geq 3$ and ∂ is an absorbing state. Specifically, the infinitesimal generator of the process is given by

$$\mathcal{G}f(x) = f(x+1) - f(x) + \theta[f(x-1) - f(x)] + p[f(\partial) - f(x)],$$

where $x \in \mathbb{Z}/K\mathbb{Z}$, $\mathcal{G}f(\partial) = 0$, with $\theta, p \in \mathbb{R}_+^*$ and f is a real function defined on $\mathbb{Z}/K\mathbb{Z} \cup \{\partial\}$. In words, $(Z_t)_{t\geq 0}$ is an asymmetric random walk on the K-cycle graph, which jumps with rates 1 and θ in the clockwise and the anti-clockwise directions, respectively. Also, with uniform rate p the process jumps to the absorbing state ∂, i.e., it is killed. Note that $\mathbb{Z}/K\mathbb{Z}$ is an irreducible class. The process generated by $\mathcal{G}$ is a particular case of the processes with *uniform killing in a finite state space* considered by Méléard and Villemonais [MV12, § 2.3].

Let $(X_t)_{t\geq 0}$ be the analogous asymmetric random walk on the cycle graph $\mathbb{Z}/K\mathbb{Z}$ without killing. The generator of this process, denoted by $\mathcal{H}$, is given by

$$\mathcal{H}f(x) = f(x+1) - f(x) + \theta[f(x-1) - f(x)], \quad \text{for all} \;\; x \in \mathbb{Z}/K\mathbb{Z}.$$

Figure 10 shows an scheme of the rates associated to $\mathcal{G}$ and $\mathcal{H}$ in the cycle graph.

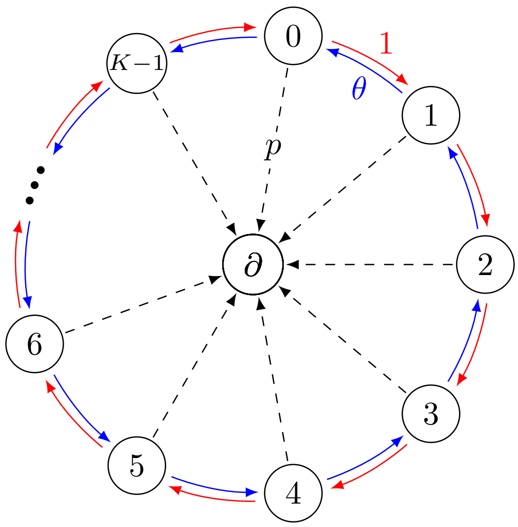


Figure 10: Scheme of the generators $\mathcal{G}$ and $\mathcal{H}$ of the asymmetric random walks on the cycle graph with and without absorption, respectively. The blue and red arrows represent the asymmetric interactions of the random walk, whereas the dashed arrows represents the absorption rates.

Note that, because of the uniform killing, the process $(Z_t)_{t\geq 0}$ could also be defined in the following way

$$Z_t = \begin{cases} X_t & \text{if} \quad t < \tau_p \\ \partial & \text{if} \quad t \geq \tau_p, \end{cases}$$

where τ_p is an exponential random variable with mean $1/p$ and independent of the random walk $(X_t)_{t\geq 0}$. This means that the law of the process $(Z_t)_{t\geq 0}$ conditioned to non-absorption is given by $\mathbb{P}_\mu(Z_t = k \mid t < \tau_p) = \mathbb{P}_\mu(X_t = k)$, for $k \in \mathbb{Z}/K\mathbb{Z}$ and for every initial distribution μ on $\mathbb{Z}/K\mathbb{Z}$. As a consequence, the QSD of $(Z_t)_{t\geq 0}$ is the stationary distribution of $(X_t)_{t\geq 0}$, which is the uniform distribution on $\mathbb{Z}/K\mathbb{Z}$, as we prove in Theorem 3.1.1.

Now, assume we have N particles with independent behavior driven by the generator $\mathcal{G}$, until one of them jumps to the absorbing state. When this happens, the particle instantaneously and uniformly jumps to one of the positions of the other $N-1$ particles. We denote by $(\eta_t^{(N)})_{t\geq 0}$

the Markov process, which accounts the positions of the N particles in the K-cycle graph at time t. Abusing notation, we redefine the state space $\mathcal{E}_{K,N}$ as follows

$$\mathcal{E}_{K,N} = \left\{ \eta : \mathbb{Z}/K\mathbb{Z} \to \mathbb{N}, \sum_{k=0}^{K-1} \eta(k) = N \right\}.$$

The generator of the N-particle process $(\eta_t^{(N)})_{t\geq 0}$, denoted by $\mathcal{Q}_{K,N}$, applied to a function f on $\mathcal{E}_{K,N}$ reads

$$(\mathcal{Q}_{K,N}f)(\eta) = \sum_{i,j\in\mathbb{Z}/K\mathbb{Z}} \eta(i) \left(\mathbb{1}_{\{j=i+1\}} + \theta \mathbb{1}_{\{j=i-1\}} + p\frac{\eta(j)}{N-1} \right) [f(\eta - \mathbf{e}_i + \mathbf{e}_j) - f(\eta)],$$

where $\theta, p > 0$ and $\mathbf{e}_i$ is the i-th canonical vector of $\mathbb{R}^K$. Under these dynamics, each of the N particles, no matter where it is, can jump to every site $j \in \mathbb{Z}/K\mathbb{Z}$ such that $\eta(j) > 0$. Note that the process $(\eta_t^{(N)})_{t\geq 0}$ is irreducible. Consequently, it has a unique stationary distribution denoted ν_N.

The definition of $\mathcal{Q}_{K,N}$ differs from the generator of the multi-allelic Moran process defined in (1) and also (13), because the denominator $N-1$ is in place of N. This is a classical difference between the usual definitions of Fleming−Viot particle process and Moran models. However, the results for the analogous generator of that defined in (13) can be easily recovered taking the selection rate p equal to $pN/(N-1)$. We keep the definition of $\mathcal{Q}_{K,N}$ in the rest of this section and in Chapter 3 in agreement with the published version [Cor21a].

The (random) empirical distribution $m(\eta_t^{(N)})$ approximates the QSD of the process $(Z_t)_{t\geq 0}$ (cf. [AFG11]) which due to Theorem 3.1.1 below is the uniform distribution. We are interested in studying how fast $m(\eta_t^{(N)})$ converges to the uniform distribution on $\mathbb{Z}/K\mathbb{Z}$ when both t and N tend to infinity. Consider $\eta_\infty^{(N)}$ a random variable with distribution ν_N, the stationary distribution of the process $(\eta_t^{(N)})_{t\geq 0}$. In this work we develop a similar analysis to that of the complete graph dynamics in [CT16b]. The convergences in (5) reduces in this case to

$$\begin{array}{ccc} m(\eta_t^{(N)}) & \xrightarrow[t\to\infty]{} & m(\eta_\infty^{(N)}) \\ N\big\downarrow & & \big\downarrow N \\ \mathcal{L}(X_t) & \xrightarrow[t\to\infty]{} & \mathcal{U}(\mathbb{Z}/K\mathbb{Z}) \end{array}$$

where the limits are in distribution, $\mathcal{L}(X_t)$ stands for the law of X_t at time t and $\mathcal{U}(\mathbb{Z}/K\mathbb{Z})$ denotes the uniform distribution on $\mathbb{Z}/K\mathbb{Z}$. Theorem 3.1.1 provides lower and upper exponential bounds for the speed of convergence of $\mathcal{L}(Z_t \mid t < \tau_p)$ to ν_{qs} in the 2-norm, when $t \to \infty$. Likewise, Corollary 3.1.7 and Theorem 3.1.9 give bounds for the speed of convergence of $m(\eta_t^{(N)})$ to $\mathcal{L}(Z_t \mid t < \tau_p)$ and $m(\eta_\infty^{(N)})$ to ν_{qs}, when $N \to \infty$.

The quantitative long time behavior of the N-particle system in countable state spaces is studied in [CT16b]. Using a coupling technique and under certain conditions, an exponential bound is provided for the convergence of $\mathcal{L}(\eta_t^{(N)})$ to ν_N in the sense of a *Wasserstein distance* [CT16b, Thm. 1.1]. In particular, the parameter λ defined by (17) needs to be positive. As we said, this is not the case of the asymmetric random walk on the K-cycle graph with uniform killing, when $K \geq 6$. A study of this convergence can be carried out using the spectrum of the generator $\mathcal{Q}_{K,N}$, which was discussed in the previous section. Indeed, using Example 2.4.1 in Chapter 2 we can get the following asymptotic expression for the profile of the convergence in total variation distance to stationarity:

$$\mathrm{D}_{\mathrm{TV}}^{\mathcal{Q}_{K,N}} = \mathcal{O}\left(\mathrm{e}^{-\rho_K t}\right),$$

where $\rho_K = 2(1+\theta)\sin^2(\pi/K)$, and $\mathrm{D}_{\mathrm{TV}}^{\mathcal{Q}_{K,N}}$ is defined as in (14). However, obtaining an explicit bound for this convergence, or proving the existence of a cutoff phenomenon, are topics of further research.

Chapter 1

Propagation of chaos of the Moran model *

This chapter is based on the preprint [CC21], submitted on September, 2021.

Abstract: The goal of this article is to study the limit of the empirical distribution induced by a mutation-selection multi-allelic Moran model, whose dynamic is given by a continuous-time irreducible Markov chain. The rate matrix driving the mutation is assumed irreducible and the selection rates are assumed uniformly bounded. The paper is divided into two parts. The first one deals with processes with general selection rates. For this case we are able to prove the propagation of chaos in $\mathbb{L}^p$ over the compacts, with speed of convergence of order $1/\sqrt{N}$. Further on, we consider a specific type of selection that we call additive selection. Essentially, we assume that the selection rate can be decomposed as the sum of three terms: a term depending on the allelic type of the parent (which can be understood as selection at death), another term depending on the allelic type of the descendant (which can be understood as selection at birth) and a third term which is symmetric. Under this setting, our results include a uniform in time bound for the propagation of chaos in $\mathbb{L}^p$ of order $1/\sqrt{N}$, and the proof of the asymptotic normality with zero mean and explicit variance, for the approximation error between the empirical distribution and its limit, when the number of individuals tend towards infinity. Additionally, we explore the interpretation of the Moran model with additive selection as a particle process whose empirical distribution approximates a quasi-stationary distribution, in the same spirit as the Fleming – Viot particle systems. We then address the problem of minimising the asymptotic quadratic error, when the time and the number of particles go to infinity.

1.1 Introduction and main results

This paper is devoted to the study of a mutation-selection multi-allelic Moran model consisting on $N \in \mathbb{N}$ individuals, which can be of different allelic types belonging to a discrete set E. The state space of the Moran model is the N discrete simplex

$$\mathcal{E}_N := \left\{ \eta : E \to \mathbb{N} \,\Big|\, \sum_{x \in E} \eta(x) = N \right\}.$$

*In collaboration with Bertrand Cloez.

The empirical distribution induced by $\eta \in \mathcal{E}_N$ is defined by

$$m(\eta) = \sum_{x\in E} \frac{\eta(x)}{N}\delta_x \in \mathcal{M}_1(E),$$

where $\mathcal{M}_1(E)$ is the set of probability measures on E. Let Q be the generator of a continuous-time, non-explosive, irreducible Markov chain, and consider some rates $V_\mu(x,y) \geq 0$, for all $x \neq y \in E$ and $\mu \in \mathcal{M}_1(E)$.

The multi-allelic Moran model is a continuous-time Markov chain evolving on $\mathcal{E}_N$. The process is at $\eta \in \mathcal{E}_N$ if there is $\eta(x)$ individuals of type x, for all $x \in E$. Between reproduction events, the N individuals evolve as independent copies of the mutation process generated by $Q = (Q_{x,y})_{x,y\in E}$. In this sense we call $Q_{x,y}$, for $x, y \in E$, the *mutation rates*. Reproduction events consist of the death of an individual of type x, which is then removed from the population, and the reproduction of an individual of type y, which add an y individual to the population. This happens at rate $\eta(y)/N \cdot V_{m(\eta)}(x,y)$. Hence, the transition rate from $\eta \in \mathcal{E}_N$, with $\eta(x) > 0$, to $\eta - \mathbf{e}_x + \mathbf{e}_y$ is

$$\eta(x)\left(Q_{x,y} + \frac{\eta(y)}{N} V_{m(\eta)}(x,y)\right),$$

for every $x \neq y \in E$, where $\eta - \mathbf{e}_x + \mathbf{e}_y$ is the element in $\mathcal{E}_N$ satisfying

$$(\eta - \mathbf{e}_x + \mathbf{e}_y)(z) = \begin{cases} \eta(z) & \text{if} \quad z \notin \{x, y\}, \\ \eta(x) - 1 & \text{if} \quad z = x, \\ \eta(y) + 1 & \text{if} \quad z = y. \end{cases}$$

We will further detail particular examples but, for the moment, let us see that when $V_{m(\eta)}(x,y)$ is constant, each individual dies at the same rate and the parent is choose uniformly at random over the individuals that are present in the population (this explain the term $\eta(y)/N$ in the transition rate). We can also interpret this rate by the opposed point of view: each individual reproduces at a constant rate and the dying individual is chosen uniformly at random. This is often called neutral selection in ecology literature, but our models allow to choose various non constant $V_{m(\eta)}(x,y)$. In this sense, we call the rates $V_\mu(x,y)$, for $x, y \in E$ and $\mu \in \mathcal{M}_1(E)$, the *selection rates.*

Note that the reproduction dynamics depends in general on both the types of the parent and the offspring, and may also depend on the empirical distribution induced by the configuration of the population at the current time, in a sense that we will clarify further along in Assumptions (G1) and (C1). The generator of the Moran model is denoted $\mathcal{Q} := \mathcal{Q}^{\text{mut}} + \mathcal{Q}^{\text{sel}}$, where $\mathcal{Q}^{\text{mut}}$ and $\mathcal{Q}^{\text{sel}}$ act on every function $f \in \mathcal{B}_b(\mathcal{E}_N)$ as follows

$$(\mathcal{Q}^{\text{mut}} f)(\eta) = \sum_{x,y\in E} \eta(x) Q_{x,y}[f(\eta - \mathbf{e}_x + \mathbf{e}_y) - f(\eta)],$$

$$(\mathcal{Q}^{\text{sel}} f)(\eta) = \frac{1}{N}\sum_{x,y\in E} \eta(x)\eta(y) V_{m(\eta)}(x,y)[f(\eta - \mathbf{e}_x + \mathbf{e}_y) - f(\eta)],$$

for every $\eta \in \mathcal{E}_N$. Throughout the paper the following boundedness condition holds:

$$\|V\| := \sup_{\mu\in\mathcal{M}_1(E)} \sup_{x,y\in E} V_\mu(x,y) < \infty. \tag{1.1}$$

Note that the non-explosion of the process generated by Q and the bound condition (1.1) let out the possibility of an infinite number of jumps in finite time. Thus, the process generated by $\mathcal{Q}$ is well-defined for all $t \geq 0$.

This Moran model is an extension, for $K > 2$, of the model studied by [Cor17]. In general, when generalising the Moran model for more than two allelic types, the selection rates are taken depending only on the children type, i.e. $V_\mu(x, y) = V^{\mathrm{b}}(y)$, for all $x, y \in E$ and $\mu \in \mathcal{M}_1(E)$, which is called *selection at birth* or *fecundity selection* [Dur08; MW09; Eth11]. Moreover, in biological applications it has been also considered models with *selection at death* or *viability selection*, when the selection rates only depend on the parent type, i.e. $V_\mu(x, y) = V^{\mathrm{d}}(x)$ [MW09], for all $x, y \in E$ and $\mu \in \mathcal{M}_1(E)$. However, the importance of this last model is beyond its biological interpretations: this process is also called *Fleming – Viot particle process*, which is an interacting particle process intended for the approximation of a *quasi-stationary distribution* (QSD) of an absorbing Markov chain conditioned on non-absorption. These particle processes have attracted lots of attention in recent years. See for instance [DM03; Vil14; Cér+20; CV21] for general state spaces, [FM07; GJ13; Ass+16; CT16b] for countable state spaces, and even [AFG11; LPR18] for finite state spaces. We will discuss later, in Section 1.2, the relation among the Moran model considered here, our results and the theory of QSD and Fleming – Viot particle processes.

1.1.1 Main results

Let us get some insight into the limit of the empirical measure induced by this particle process when the number of particles tends towards infinity. Let us denote by $(\eta_t)_{t\geq 0}$ the continuous-time Markov chain on $\mathcal{E}_N$, generated by $\mathcal{Q}$. Although the process generated by $\mathcal{Q}$ clearly depends on N and a better notation would be $(\eta_t^{(N)})_{t\geq 0}$, we keep this dependence implicit for the sake of simplicity. By the Kolmogorov equation we know that $\partial_t \mathbb{E}_\eta[m_x(\eta_t)] = \mathbb{E}_\eta\left[(\mathcal{Q}\, m_x)(\eta_t)\right]$, where m_x stands for the empirical distribution induced by η on the point $x \in E$, i.e. $m_x : \eta \mapsto \eta(x)/N$. Let us thus compute $\mathcal{Q}\, m_x$. It is easy to get

$$(\mathcal{Q}^{\mathrm{mut}} m_x)(\eta) = \sum_{y\in E} Q_{y,x} m_y(\eta), \tag{1.2}$$

for every $x \in E$, for all $\eta \in \mathcal{E}_N$.

On the other hand,

$$(\mathcal{Q}^{\mathrm{sel}} m_x)(\eta) = -m_x(\eta) \sum_{y\in E} m_y(\eta)[V_{m(\eta)}(x, y) - V_{m(\eta)}(y, x)]. \tag{1.3}$$

Finally, we get

$$\partial_t \mathbb{E}_\eta[m_x(\eta_t)] = \sum_{y\in E} Q_{y,x} \mathbb{E}_\eta[m_y(\eta_t)] - \sum_{y\in E} [V_{m(\eta_t)}(x, y) - V_{m(\eta_t)}(y, x)]\mathbb{E}_\eta[m_x(\eta_t)m_y(\eta_t)].$$

When the number of individuals N is large, we expect the Moran process to exhibit a *propagation of chaos phenomenon* and thus the empirical distribution induced by the process approximates the solution of the following nonlinear system of ordinary differential equations:

$$\partial_t \gamma_t(x) = \sum_{y\in E} Q_{y,x} \gamma_t(y) - \sum_{y\in E} [V_{\gamma_t}(x, y) - V_{\gamma_t}(y, x)]\gamma_t(x)\gamma_t(y),$$

for all $x \in E$. For every function ϕ on E we thus get the nonlinear differential equation

$$\partial_t \gamma_t(\phi) = \gamma_t(Q_{\gamma_t}\phi), \tag{1.4}$$

where $Q_\gamma := Q + \Pi_\gamma$ and

$$\Pi_\gamma \phi : x \mapsto \sum_{y\in E} \gamma(y) V_\gamma(x, y)[\phi(y) - \phi(x)],$$

for every probability distribution γ on E.

The main results we provide in this article are related to the speed of convergence of $(m(\eta_t))_{t\geq 0}$ towards $(\gamma_t)_{t\geq 0}$ when $N \to \infty$.

Propagation of chaos with general selection rate

We denote by $\|\cdot\|$ the uniform norm on the set of functions on E, defined by

$$\|\phi\| := \sup_{x\in E} |\phi(x)|.$$

Let $\mathcal{B}_b(E)$ be the set of bounded functions on E for the uniform norm and $\mathcal{B}_1(E) := \{\phi : E \to \mathbb{R} :\ \|\phi\| \leq 1\}$. For two probability distributions $\mu_1, \mu_2 \in \mathcal{M}_1(E)$ the total variation distance is defined as follows:

$$\|\mu_1 - \mu_2\|_{\mathrm{TV}} := \sup_{A\subset E} |\mu_1(A) - \mu_2(A)| = \frac{1}{2} \sup_{\phi\in\mathcal{B}_1(E)} |\mu_1(\phi) - \mu_2(\phi)| = \frac{1}{2} \sum_{x\in E} |\mu_1(x) - \mu_2(x)|,$$

where $\mu(\phi)$ stands for the mean of ϕ with respect to $\mu \in \mathcal{M}_1(E)$.

First, we consider the case where the selection rates satisfy the following hypothesis.

Assumption (G1) (General selection rate)**.** The selection rates are uniformly bounded as in (1.1) and there exists $V_i^{\mathrm{d}}, V_i^{\mathrm{b}}$, for $i \geq 1$, and a continuous, nonnegative function $\mu \mapsto V_\mu^{\mathrm{s}}$ from $(\mathcal{M}_1(E), \|\cdot\|_{\mathrm{TV}})$ to $(\mathcal{B}_b(E \times E), \|\cdot\|)$ such that V_μ^{s} is symmetric in $E \times E$ for every μ, and

$$V_\mu(x, y) = \sum_{i\geq 1} V_i^{\mathrm{d}}(x) V_i^{\mathrm{b}}(y) + V_\mu^{\mathrm{s}}(x, y),$$

for all $\mu \in \mathcal{M}_1(E)$ and $x, y \in E$. In addition, $V_\mu - V_\mu^{\mathrm{s}} \in \mathcal{B}_b(E \times E)$ and

$$\left\{ \sum_{i\geq 1} |V_i^{\mathrm{d}} - V_i^{\mathrm{b}}|,\ \sup_{i\geq 1} V_i^{\mathrm{d}},\ \sup_{i\geq 1} V_i^{\mathrm{d}} \right\} \subset \mathcal{B}_b(E).$$

The expression given by Assumption (G1) for V is rather general. It can be seen as a "discrete Taylor's expansion" for $V_\mu - V_\mu^{\mathrm{s}}$, with some boundedness conditions on the norm of the factors in the development. See also [Del04, p. 25], and the references therein, for a Feynman – Kac interpretation for the differential equation (1.4) satisfying (G1), and even a more general expression for the selection rates.

We also assume the existence of a solution of the differential equation (1.4).

Assumption (G2) (Existence of a solution)**.** For every $\mu_0 \in \mathcal{M}_1(E)$, there is a unique solution $(\mu_t)_{t\geq 0}$ of the differential equation (1.4). Namely, $(\mu_t)_{t\geq 0}$ is solution of the Cauchy problem

$$\partial_t \mu_t(\phi) = \mu_t(Q^\star_{\mu_t}\phi),$$

with initial condition $\mu_0(\phi)$, for every $\phi \in \mathcal{B}_d(E)$, where

$$Q^\star_\mu\phi : x \mapsto (Q\phi)(x) + \sum_{y\in E} \mu(y) V^\star(x, y)[\phi(y) - \phi(x)]$$

and $V^\star := V_\mu - V_\mu^{\mathrm{s}} \in \mathcal{B}_b(E)$.

Assumption (G2) is always verified when $\|Q\| < \infty$, since the differential equation (1.4) satisfies a Lipschitz condition. This includes in particular the case where E is finite. In general, when the generator is not bounded, the analysis is more delicate.

Consider also the following assumption.

Assumption (I) (Initial condition)**.** The empirical measure induced by the particle process at $t = 0$ converges towards the initial distribution $\mu_0 \in \mathcal{M}_1(E)$ in $\mathbb{L}^p$, for every $p \geq 1$. More precisely, for every $p \geq 1$, there exists a constant $C_p > 0$ such that

$$\sup_{\phi \in \mathcal{B}_1(E)} \mathbb{E}[|m(\eta_0)(\phi) - \mu_0(\phi)|^p] \leq \frac{C_p}{N^{p/2}}.$$

Note that Assumption (I) is verified when initially all the particles are sampled independently with distribution μ_0, as the next lemma shows.

Lemma 1.1.1 (Control of the initial error)**.** *Assume that initially the N particles are sampled independently according to $\mu_0 \in \mathcal{M}_1(E)$. Then, Assumption (I) is verified.*

The proof of Lemma 1.1.1 is deferred to Appendix 1.A. We include Assumption (I) in order to be able to apply our results to a wider class of situations, than that described in Lemma 1.1.1.

Theorem 1.1.2 (Propagation of chaos on compacts)**.** *Suppose that Assumptions (G1), (G2) and (I) are verified. Then, for every $T \geq 0$ and $p \geq 1$, there exists a constant $C_{p,T} > 0$, such that*

$$\sup_{\phi \in \mathcal{B}_1(E)} \mathbb{E}\left[\sup_{t \in [0,T]} |m(\eta_t)(\phi) - \mu_t(\phi)|^p\right]^{1/p} \leq \frac{C_{p,T}}{\sqrt{N}},$$

where $(\mu_t)_{t \geq 0}$ is as in Assumption (G2), with initial condition $\mu_0 \in \mathcal{M}_1(E)$ as in Assumption (I).

The proof of Theorem 1.1.2 is deferred to Section 1.3.2.

Let $(x_n)_{n \geq 1}$ be an enumeration of the elements in E. We define the following distance in $\mathcal{M}_1(E)$:

$$\|\mu_1 - \mu_2\|_{\mathrm{w}} := \sum_{k \geq 1} 2^{-k} |\mu_1(x_k) - \mu_2(x_k)|.$$

Note that the space $\mathcal{M}_1(E)$ with the convergence in law (the weak topology) is metrizable with this distance. As a consequence of Theorem 1.1.2 we get the following result.

Corollary 1.1.3 (Convergence of the empirical measure)**.** *Suppose that Assumptions (G1), (G2) and (I) are verified. Then, for every $T \geq 0$ and $p \geq 1$, there exists a constant $C_{p,T} > 0$, such that*

$$\mathbb{E}\left[\left(\sup_{t \in [0,T]} \|m(\eta_t) - \mu_t\|_{\mathrm{w}}\right)^p\right]^{1/p} \leq \frac{C_{p,T}}{\sqrt{N}}.$$

Corollary 1.1.3 is proved in Section 1.3.2.

Note that this ensures a functional convergence in $\mathbb{L}^p(\mathcal{C}([0,T], \mathcal{M}_1(E)))$:

$$m(\eta_\cdot) \xrightarrow[N \to \infty]{\mathbb{L}^p} \mu_\cdot,$$

with an estimation of the speed of convergence. Furthermore, Theorem 1.1.2, for $p = 4$, and a Borel−Cantelli argument imply the convergence $m(\eta_\cdot) \xrightarrow{\text{c.c.}} \mu_\cdot$ in the weak sense, where c.c. denotes the complete (or universal) convergence (cf. [Gut13, Def. 1.6]). In particular, this implies $m(\eta_\cdot) \xrightarrow{\text{a.s.}} \mu_\cdot$, when $N \to \infty$, in the weak sense, no matter in which space the random variables are coupled.

Theorem 1.1.2 is a generalisation for multi-allelic Moran models with more than two allelic types, of Proposition 3.1 in [Cor17], where the uniform convergence on compacts in probability is proved. The speed of convergence in Theorem 1.1.2 can also be related to existing results that ensure the convergence of the empirical measure induced by a Moran type (or Fleming – Viot) particle process towards the law of an absorbing process conditioned to non-absorption. See for instance [DM00b, Prop. 3.5] [DPR11, Lemma 3.1], [Vil14, Thm. 2.2], [CT16a, Thm. 1.1], and [AD20, Thm. 5.10 and Cor. 5.12]. See also [BC15, Thm. 3.1 and Rmk. 3.2] where the almost sure convergence (and also the complete convergence) is proved when the state space is finite. As far as we know, Theorem 1.1.2 and Corollary 1.1.3 are the first results ensuring the convergence uniformly on compacts in $\mathbb{L}^p$, for all $p \geq 1$, with speed of convergence of order $1/\sqrt{N}$ for multi-allelic Moran models with general selection rates in the sense of Assumption (G1), in discrete countable state spaces, not necessarily finite. The idea behind the proof is closed to the methods in [Rou06]: it consists in finding a martingale indexed by the interval $[0, T]$, whose terminal value at time T is precisely $m(\eta_T)(\phi) - \mu_T(\phi)$ plus a term whose $\mathbb{L}^p$ norm can be controlled, for any $\phi \in \mathcal{B}_b(E)$. Thereafter, the final result comes by a Grönwall type argument similarly to the proof of Proposition 1 in [MS19]. Nevertheless, [Rou06] does not contain any uniform bound as in Theorem 1.1.2.

Uniform in time propagation of chaos for additive selection rates

Under a more specific expression for the selection rates, we prove a uniform in time bound for the convergence of $(m(\eta_t))_{t\geq 0}$ towards $(\mu_t)_{t\geq 0}$, when $N \to \infty$. Consider the following kind of selection rates that we call *additive selection*.

Assumption (C1) (Additive selection)**.** The selection rates are uniformly bounded as in (1.1). Moreover, there exist two continuous nonnegative functions $\mu \mapsto V_\mu^{\mathrm{d}}$ and $\mu \mapsto V_\mu^{\mathrm{b}}$, from $(\mathcal{M}_1(E), \|\cdot\|_{\mathrm{TV}})$ to $(\mathcal{B}_b(E), \|\cdot\|)$; and a continuous, nonnegative function $\mu \mapsto V_\mu^{\mathrm{s}}$ from $(\mathcal{M}_1(E), \|\cdot\|_{\mathrm{TV}})$ to $(\mathcal{B}_b(E \times E), \|\cdot\|)$ such that V_μ^{s} is symmetric on $E \times E$, for every $\mu \in \mathcal{M}_1(E)$ and

$$V_\mu(x, y) = V_\mu^{\mathrm{d}}(x) + V_\mu^{\mathrm{b}}(y) + V_\mu^{\mathrm{s}}(x, y),$$

for all $x, y \in E$ and $\mu \in \mathcal{M}_1(E)$. Furthermore, there exist a function $\Lambda \in \mathcal{B}_b(E)$ and a real function $\mu \in \mathcal{M}_1(R) \mapsto C_\mu$ such that

$$\Lambda(x) = V_\mu^{\mathrm{b}}(x) - V_\mu^{\mathrm{d}}(x) + C_\mu, \tag{1.5}$$

for every $x \in E$.

Remark 1.1.1 (Selection rates independent on μ). When the selection rates do not depend on μ, Assumption (C1) reduces to the existence of $V^{\mathrm{d}}, V^{\mathrm{b}} \in \mathcal{B}_b(E)$ and a symmetric $V^{\mathrm{s}} \in \mathcal{B}_b(E \times E)$ such that

$$V(x, y) = V^{\mathrm{d}}(x) + V^{\mathrm{b}}(y) + V^{\mathrm{s}}(x, y).$$

Let $\Lambda \in \mathcal{B}_b(E)$ be a fixed function. Typical examples of functions V^{b} and V^{d} satisfying this condition are

$$V^{\mathrm{b}} = (\Lambda - c)^+ \text{ and } V^{\mathrm{d}} = (\Lambda - c)^-,$$

for a fixed constant $c \in \mathbb{R}$, where we use the standard notation

$$(x)^+ := \max\{x, 0\} \text{ and } (x)^- := -\min\{x, 0\}.$$

These are in fact the selections rates considered by Angeli et al. [AGJ21, §3.3] in the context of cloning algorithms. Moreover, the case $c = 0$ is considered in Example 3.1-(2) in [Rou06]. Note that in this case Assumption (G1) is also verified.

From a biological point of view, the parameter $c \in \mathbb{R}$ can be seen as a fitness parameter. Let us assume that V^{s}_μ is null for simplicity, and denote by $\xi_t^{(i)}$ the type of the i-th individual, for $1 \le i \le N$, at time $t \ge 0$. Then, if $\Lambda(\xi_t^{(i)}) \le c$, the i-th individual dies and another randomly chosen individual reproduces with rate $(\Lambda(\xi_t^{(i)}) - c)^-$. Otherwise, if $\Lambda(\xi_t^{(i)}) \ge c$ a random chosen individual dies and the i-th individual reproduces with rate $(\Lambda(\xi_t^{(i)}) - c)^+$.

Another example of particular interest is when $V^{\mathrm{b}} = 0$. Notice that the Moran process with these selection rates is in fact a Fleming–Viot particle process (cf. [FM07]). Later, in Section 1.2 we will consider in detail the interpretation of the Moran processes as particle systems approaching a quasi-stationary distribution.

Remark 1.1.2 (Selection rates depending on μ). Consider a fixed function $\Lambda \in \mathcal{B}_b(E)$. Typical examples of functions V^{b}_μ and V^{d}_μ are:

$$V^{\mathrm{b}}_\mu = (\Lambda - \mu(\Lambda))^+ \text{ and } V^{\mathrm{d}}_\mu = (\Lambda - \mu(\Lambda))^-.$$

These are the selection rates considered in [Del04, §1.5.2, p. 35], see also Example 3.1-(3) in [Rou06].

In this case the biological interpretation of $\mu(\Lambda)$ is similar to that of the parameter c in the previous remark. Indeed, the fitness coefficient evolves in time according to the evolution of the population.

Remark 1.1.3 (Additive selection and Feynman–Kac semigroups). Consider that Assumption (C1) is satisfied. Then, from (1.4) we can recover the nonlinear differential equation

$$\partial_t \gamma_t(\phi) = \gamma_t((Q + \Lambda)\phi - \gamma_t(\Lambda)\phi), \tag{1.6}$$

where Λ is as defined in (1.5). Indeed, let $Q^\star_\gamma$ defined as in Assumption (G2), namely

$$Q^\star_\gamma \phi : x \mapsto (Q\phi)(x) + \sum_{y \in E} \gamma(y) \left(V^{\mathrm{d}}(x) + V^{\mathrm{b}}(y)\right) [\phi(y) - \phi(x)]$$

then

$$\begin{aligned} \gamma(Q^\star_\gamma \phi) &= \gamma(Q\phi) + \gamma(V^{\mathrm{d}}_\gamma)\gamma(\phi) - \gamma(V^{\mathrm{d}}_\gamma \phi) + \gamma(V^{\mathrm{b}}_\gamma \phi) - \gamma(V^{\mathrm{b}}_\gamma)\gamma(\phi) \\ &= \gamma(Q\phi + \Lambda\phi - \gamma(\Lambda)\phi). \end{aligned} \tag{1.7}$$

We emphasise that the symmetric component V^{s} is not present in (1.6).

Consider the Feynman–Kac semigroup $(P_t^\Lambda)_{t \ge 0}$, where

$$P_t^\Lambda(\phi) : x \mapsto \mathbb{E}_x \left[\phi(X_t) \exp\left\{\int_0^t \Lambda(X_s) \mathrm{d}s\right\}\right] \tag{1.8}$$

whose generator is $Q + \Lambda$. Let us define the normalised version of this semigroup as follows

$$\mu_t(\phi) := \frac{\mu_0 P_t^\Lambda(\phi)}{\mu_0 P_t^\Lambda(\mathbf{1})}, \tag{1.9}$$

where $\mathbf{1}$ denotes the all-one function on E. Then, $(\mu_t)_{t \ge 0}$ is the solution of the nonlinear differential equation (1.6) with initial value $\mu_0(\phi)$ for $t = 0$ [Del04, Eq. (1.17)].

Remark 1.1.4 (Translation of the selection rate and QSD). Note that $(\mu_t)_{t \ge 0}$ as defined above is invariant by translation of the function Λ. Namely, for every real β we have that

$$\mu_t(\varphi) = \frac{\mu_0 P_t^{\Lambda - \beta}(\varphi)}{\mu_0 P_t^{\Lambda - \beta}(\mathbf{1})}. \tag{1.10}$$

In particular, taking $\beta = \sup \Lambda$, we can always interpret $(\mu_t)_{t\geq 0}$ as the distribution of an absorbed Markov chain conditioned to non-absorption up to time t with killing rate $\kappa = \sup \Lambda - \Lambda$. This naturally relates the study of the behaviour of $(\mu_t)_{t\geq 0}$ when $t \to \infty$, to the theory of quasi-stationary distributions (QSD), as we will discuss later in Section 1.2.

Consider the following assumptions related to the control in the norm $\mathbb{L}^p$ of the initial error and the exponential convergence of $(\mu_t)_{t\geq 0}$, as defined by (1.10), towards a unique limit, for every initial distribution on $\mu_0 \in \mathcal{M}_1(E)$.

Assumption (C2) (Uniform exponential ergodicity of the normalised semigroup)**.** There exist a distribution $\mu_\infty \in \mathcal{M}_1(E)$ and $C, \gamma > 0$, such that for every initial distribution $\mu_0 \in \mathcal{M}_1(E)$ and for all $t \geq 0$:

$$\|\mu_t - \mu_\infty\|_{\mathrm{TV}} \leq C\mathrm{e}^{-\gamma t}, \tag{1.11}$$

where $(\mu_t)_{t\geq 0}$ is defined as in (1.9).

We are now in a position to state our main results for the multi-allelic Moran model with additive selection.

Theorem 1.1.4 (Uniform in time propagation of chaos)**.** *Under Assumptions (I), (C1) and (C2), for every $p \geq 1$, there exists a constant C_p, such that*

$$\sup_{\phi\in\mathcal{B}_1(E)} \sup_{t\geq 0} \mathbb{E}\big[|m(\eta_t)(\phi) - \mu_t(\phi)|^p\big]^{1/p} \leq \frac{C_p}{\sqrt{N}}.$$

Theorem 1.1.4 is proved in Section 1.3.3.

Corollary 1.1.5 (Convergence of the empirical measure)**.** *Suppose that Assumptions (I), (C1) and (C2) are verified. Then, for every $p \geq 1$, there exists a constant $C_p > 0$, such that*

$$\sup_{t\geq 0} \mathbb{E}\big[(\|m(\eta_t) - \mu_t\|_{\mathrm{w}})^p\big]^{1/p} \leq \frac{C_p}{\sqrt{N}}.$$

The proof of Corollary 1.1.5 is analogous to that of Corollary 1.1.3 and we skip it for the seek of brevity.

For a fixed N, if the process $(\eta_t)_{t\geq 0}$ generated by $\mathcal{Q}$ allows a stationary distribution ν_N, then under the hypothesis in Theorem 1.1.4 we get

$$\sup_{\phi\in\mathcal{B}_1(E)} \mathbb{E}_{\nu_N}\big[|m(\eta_\infty)(\phi) - \mu_\infty(\phi)|^p\big]^{1/p} \leq \frac{C_p}{\sqrt{N}}, \tag{1.12}$$

for all $p \geq 1$.

Obtaining a uniform in time bound as the one provided by Theorem 1.1.4 is a hard problem and this kind of results are uncommon in the literature. Del Moral and Guionnet in [DG01, Thm. 3.1] have proved a similar result for a resembling but discrete-time model, where the potential function Λ is assumed uniformly bounded and also bounded away from zero. Moreover, their upper bound for the speed of convergence in [DG01, Thm. 3.1] is of order $1/N^\alpha$, with $\alpha < 1/2$. Rousset [Rou06, Thm. 4.1] has proved a uniform in time bound in $\mathbb{L}^p$ with the same speed of convergence as our result. However, the model studied by Rousset is in continuous state space and the diffusion process driving the mutation process is assumed reversible. Similarly, Angeli et al. [AGJ21, Thm. 3.2] obtained an equivalent result for jump processes on locally compact spaces in the context of cloning algorithms, for $p \geq 2$. See also Theorem 5.10 and Corollary 5.12 in [AD20], for a related result when $p = 2$. Our model is different, since we consider the case where the state space is discrete, not necessarily finite and in Assumption (C1) we allow the selection rates to depend on the empirical probability measure induced by the particle system, in the same spirit of [Rou06]. Nonetheless, our methods are similar to those of Rousset [Rou06] and Angeli et al. [AGJ21] (see also [DM00b, §3.3.1]): it consists in finding a martingale indexed by the interval $[0,T]$, whose terminal value at time T is precisely $m(\eta_T)(\phi) - \mu_T(\phi)$ plus a term whose $\mathbb{L}^p$ norm can be controlled, for any $\phi \in \mathcal{B}_b(E)$. Thereafter, the final result comes by a control of the quadratic variation of the martingale and an induction principle.

Remark 1.1.5 (Almost sure convergence). Corollary 1.1.5, for $p = 4$, and a Borel−Cantelli argument imply the convergence $m(\eta_T) \xrightarrow{\text{c.c.}} \mu_T$ in $\mathcal{M}_1(E)$, when $N \to \infty$, for every $T \geq 0$, where c.c. denotes the complete (or universal) convergence. In particular, this implies $m(\eta_T) \xrightarrow{\text{a.s.}} \mu_T$, when $N \to \infty$, for every $T \geq 0$. Note that in contrast with Corollary 1.1.3, Corollary 1.1.5 does not ensure the convergence in $\mathcal{C}([0,T], \mathcal{M}_1(E))$.

Let us denote by $\bar{m}(\eta_t)$ the mean empirical probability measure induced by η_t, which is defined as

$$\bar{m}(\eta_t) := \sum_{x \in E} \mathbb{E}\left[\frac{\eta_t(x)}{N}\right] \delta_x \in \mathcal{M}_1(E).$$

We recall that $\xi_t^{(i)}$ stands for the type of the i-th individual, for $1 \leq i \leq N$, at time $t \geq 0$. Let us denote by $\mathrm{Law}(\xi_t^{(i)})$ the law of $\xi_t^{(i)}$.

Theorem 1.1.6 (Bias estimate one ergodicity of one particle)**.** *Under Assumptions* (I), (C1) *and* (C2), *there exists a constant $C > 0$ such that*

$$\sup_{t \geq 0} \|\bar{m}(\eta_t) - \mu_t\|_{\mathrm{TV}} \leq \frac{C}{N}.$$

Moreover, if the initial distribution of the N particles is exchangeable, then

$$\sup_{t \geq 0} \left\|\mathrm{Law}(\xi_t^{(i)}) - \mu_t\right\|_{\mathrm{TV}} \leq \frac{C}{N}.$$

Theorem 1.1.6 are proved in Section 1.3.3.

It is expected that when selection rates are constant the empirical probability measure generated by the particle system is a unbiased estimator of the law of the Markov chain generated by Q, in the sense that

$$\bar{m}(\eta_t) = \bar{m}(\eta_0)\mathrm{e}^{tQ}, \text{ for all } t \geq 0.$$

See e.g. [CT16b] and [Cor21a]. We prove in Corollary 1.3.6 below that this result also holds when the selection rates are symmetric.

The following result ensures the exponential ergodicity of the unnormalised semigroup.

Lemma 1.1.7 (Exponential ergodicity of the unnormalised semigroup)**.** *Suppose that Assumptions* (C1) *and* (C2) *are verified. Then, there exists a unique triplet $(\mu_\infty, h, \lambda) \in \mathcal{M}_1(E) \times \mathcal{B}_b(E) \times \mathbb{R}$, of eigenelements of $Q + \Lambda$ such that h is strictly positive, $\mu_\infty(h) = 1$,*

$$\mu_\infty P_t^\Lambda = \mathrm{e}^{\lambda t}\mu_\infty \text{ and } P_t^\Lambda(h) = \mathrm{e}^{\lambda t}h.$$

Moreover, there exist $C, \gamma > 0$ such that for all $t \geq 0$:

$$\sup_{\mu_0 \in \mathcal{M}(E)} \|\mathrm{e}^{-\lambda t}\mu_0 P_t^\Lambda - \mu_0(h)\mu_\infty\|_{\mathrm{TV}} \leq C\mathrm{e}^{-\gamma t}. \tag{1.13}$$

Furthermore, $\lambda \leq 0$ whether $\Lambda \leq 0$.

This result is basically a consequence of Theorem 2.1 of [CV17b] (Theorem 1.2.2 below). We review this theorem and others results on the theory of quasi-stationary distribution in Section 1.2. Lemma 1.1.7 establishes an exponential control on the speed of convergence of the unnormalised semigroup. A similar estimate is stated in [AGJ21, Assumption 2.2] as hypothesis. However, their assumption implies that the eigenfunction h is constant, which in practice makes their assumption only valid when Λ ($\mathcal{V}$ in their notation) is constant.

Let us define

$$S_\mu(\phi) := \sum_{x,y\in E} (\phi(x) - \phi(y))^2 V^{\mathrm{s}}_\mu(x,y)\mu(x)\mu(y), \tag{1.14}$$

for every $\phi \in \mathcal{B}_b(E)$, and the operator $W_{t,T}$ for $t \le T$ as follows

$$W_{t,T} : \phi \mapsto \frac{P^\Lambda_{T-t}(\phi)}{\mu_t\left(P^\Lambda_{T-t}(\mathbf{1})\right)}, \tag{1.15}$$

Note that

$$\mu_t\left(P^\Lambda_{T-t}(\mathbf{1})\right) = \exp\left\{\int_t^T \mu_s(\Lambda)\mathrm{d}s\right\}. \tag{1.16}$$

Indeed, it is a consequence of the next two identities

$$\frac{\mathrm{d}}{\mathrm{d}t} \ln\left(\mu_0 P^\Lambda_t(\mathbf{1})\right) = \mu_t(\Lambda) \quad \text{and} \quad \mu_t\left(P^\Lambda_{T-t}(\mathbf{1})\right) = \frac{\mu_0\left(P^\Lambda_T(\mathbf{1})\right)}{\mu_0\left(P^\Lambda_t(\mathbf{1})\right)}.$$

Our last two results are addressed to the study of the asymptotic square error of the approximation of μ_T by $m(\eta_T)$ when $T, N \to \infty$. These results are highly important when the Moran process is used for approximating a quasi-stationary distribution. Let us define the asymptotic quadratic errors:

$$\begin{aligned}\sigma^2_T(\phi) &:= \lim_{N\to\infty} N\mathbb{E}\left[\left(m(\eta_T)(\phi) - \mu_T(\phi)\right)^2\right],\\ \sigma^2_\infty(\phi) &:= \lim_{T\to\infty} \sigma^2_T(\phi),\end{aligned}$$

for every $\phi \in \mathcal{B}_b(E)$. First, we prove the asymptotic normality of the bias and we provide explicit expressions for $\sigma^2_T(\phi)$ and $\sigma^2_\infty(\phi)$. Then, we use this expression to show how to define another Moran process approaching the same distribution μ_∞, with smaller or equal asymptotic square error.

In order to prove the asymptotic normality of the statistic $\sqrt{N}(m(\eta_T)(\phi) - \mu_T(\phi))$, for every $T \ge 0$, we naturally need to ask, in addition to the law of large numbers established by Assumption (I), for the existence of a central limit theorem on the initial empirical distribution, as stated in the following hypothesis.

Assumption (I′) (Asymptotic normality for initial empirical distribution)**.** For every $\phi \in \mathcal{B}_b(E)$, the empirical measure induced by the particle process at $t = 0$ satisfy the following condition: $\sqrt{N}(m(\eta_0)(\phi) - \mu_0(\phi))$ converges in law towards a centered Gaussian distribution of variance $\mu_0(\phi^2)$, when $N \to \infty$.

Analogously to Lemma 1.1.1, we have that Assumption (I′) is verified when initially the N particles are sampled independently according to $\mu_0 \in \mathcal{M}_1(E)$. The proof of this result is a simple consequence of the classical central limit theorem.

Theorem 1.1.8 (Asymptotic normality)**.** *Suppose that Assumptions (I), (I′), (C1) and (C2) are verified. Then, for every $\phi \in \mathcal{B}_b(E)$ and $T \geq 0$, we have that $\sqrt{N}(m(\eta_T)(\phi) - \mu_T(\phi))$ converges in law, when N goes to infinity, towards a Gaussian centered random variable of variance*

$$\sigma_T^2(\phi) = \mathrm{Var}_{\mu_T}(\phi) + \int_0^T S_{\mu_s}(W_{s,T}(\bar{\phi}_T))\mathrm{d}s + 2\int_0^T \mu_s\Big(W_{s,T}(\bar{\phi}_T)^2\Big(V_{\mu_s}^{\mathrm{b}} + \mu_s(V_{\mu_s}^{\mathrm{d}})\Big)\Big)\mathrm{d}s,$$

where Var_{μ_T} stands for the variance with respect to μ_T, $\bar{\phi}_T := \phi - \mu_T(\phi)$ and S_μ and $W_{t,T}$ are as defined in (1.14) and (1.15), respectively. Moreover,

$$\sigma_\infty^2(\phi) = \mathrm{Var}_{\mu_\infty}(\phi) + \int_0^\infty \mathrm{e}^{-2\lambda s} S_{\mu_\infty}(P_s^\Lambda(\bar{\phi}_\infty))\mathrm{d}s + 2\int_0^\infty \mathrm{e}^{-2\lambda s}\mu_\infty\Big(P_s^\Lambda(\bar{\phi}_\infty)^2\Big(V_{\mu_\infty}^{\mathrm{b}} + \mu_\infty(V_{\mu_\infty}^{\mathrm{d}})\Big)\Big)\mathrm{d}s,$$

where $\mathrm{Var}_{\mu_\infty}$ stands for the variance with respect to μ_∞, $\bar{\phi}_\infty := \phi - \mu_\infty(\phi)$ and λ is the eigenvalue in the statement of Lemma 1.1.7.

The proof of Theorem 1.1.8 can be found in Section 1.3.4. Note that the two integrals in the expression of $\sigma_\infty^2(\phi)$ in Theorem 1.1.8 converge as a consequence of Lemma 1.1.7.

Let us mention the relation between Theorem 1.1.8 and some existing results in the literature. When V^{s} is null, and the selection rates do not depend on μ, our result is related to Proposition 3.7 in [DM03]. Indeed, when taking the parameters of the model in [DM03] as follows: $V = 2V^{\mathrm{b}}$, $V' = 2V^{\mathrm{d}}$ and $\rho = 1/2$, Theorem 1.1.8 can be obtained from Proposition 3.7 in [DM03].

Moreover, when V_μ^{b} and V_μ^{s} are null and thus $\Lambda = -V^{\mathrm{d}} \leq 0$, we get

$$\sigma_\infty^2(\phi) = \mathrm{Var}_{\mu_\infty}(\phi) - 2\lambda\int_0^\infty \mathrm{e}^{-2\lambda s}\,\mathrm{Var}_{\mu_\infty}\Big(P_s^\Lambda(\phi)\Big)\,\mathrm{d}s.$$

When the process $(\eta_t)_{t\geq 0}$ is ergodic and converges in law to some random variable η_∞, when $t \to \infty$, Theorem 1.1.8 states that $\sqrt{N}(m(\eta_\infty)(\phi) - \mu_\infty(\phi))$ converges to a centered Gaussian law of variance $\sigma_\infty^2(\phi)$, when $N \to \infty$. Indeed, recall that a Gaussian sequence converges in law if their first two moments converge. In particular, we recover (and extend) the recent result of Lelièvre et al. [LPR18, Thm. 2.4] for finite state spaces. Notice that the negative constant λ in the previous expression for $\sigma_\infty^2(\phi)$ is the opposite of that in [LPR18, Thm. 2.4].

The expression for $\sigma_\infty^2(\phi)$, when $V_\mu^{\mathrm{s}} = 0$, is also similar to the expression for the asymptotic square error in Theorem 4.4 in [Rou06]. However, the results in [Rou06] do not include the asymptotic normality we prove in Theorem 1.1.8. See also Corollary 2.7 and Remark 2.8 [Cér+20] for a central limit theorem for the empirical measure induced by Fleming – Viot particle systems.

Note that the three summands in the expression of $\sigma_T^2(\phi)$ in Theorem 1.1.8 are positive, for every $T \geq 0$. Moreover, the limit $(\mu_t)_{t\geq 0}$ is invariant by the choice of the symmetric component V_μ^{s} in Assumption (C1). As a consequence, for a given selection rate V_μ we can obtain another Moran process approaching the same limit distribution taking the selection rate $V_\mu - \Sigma_\mu \geq 0$, where Σ_μ is a symmetric function in $\mathcal{B}_b(E \times E)$. We thus get the following result.

Corollary 1.1.9 (Moran process with smaller asymptotic square error)**.** *Suppose that Assumptions (I), (I′), (C1) and (C2) are verified. Let $(\eta_t)_{t\geq 0}$ and $(\eta_t^\star)_{t\geq 0}$ be the Moran processes with the same mutation rates and selection rates given by V_μ and $V_\mu - \Sigma_\mu$, respectively, where*

$$\Sigma_\mu(x,y) := \min\Big\{V_\mu^{\mathrm{d}}(x), V_\mu^{\mathrm{b}}(x)\Big\}\mathbf{1}_{\{x\}} + \min\Big\{V_\mu^{\mathrm{d}}(y), V_\mu^{\mathrm{b}}(y)\Big\}\mathbf{1}_{\{y\}} + V_\mu^{\mathrm{s}}(x,y),$$

where $\mathbf{1}_A$ stands for the indicator function on $A \subset E$. Then,

$$\lim_{N\to\infty} N\mathbb{E}\left[\left(m(\eta_T^\star)(\phi) - \mu_T(\phi)\right)^2\right] \leq \lim_{N\to\infty} N\mathbb{E}\left[\left(m(\eta_T)(\phi) - \mu_T(\phi)\right)^2\right],$$

for all $T \geq 0$.

Note that the selection rate $V_\mu - \Sigma_\mu$ in the statement of Corollary 1.1.9 satisfies Assumption (C1). The proof of the previous result is thus a simple consequence of Theorem 1.1.8. In Example 1.2.1 below we discuss the application of this result to the simple case of the bi-allelic Moran model, that is, when the cardinality of E is 2.

Structure of the paper

The rest of the paper is organised as follows. In Section 1.2 we discuss the relation of our results to the theory of quasi-stationary distributions. We are particularly interested in the consequences of Assumption (C2), namely, the uniform exponential convergence of the conditioned process to the quasi-stationary distribution, the spectral properties of the Markov semi-group and the ergodicity of the Feynman – Kac semigroup defined by (1.8). Next, we consider several examples of mutation and selection rates where Assumptions (C1) and (C2) are verified. We end this section with a discussion about our main results and their possible extensions. Finally, in Section 1.3, we prove our main results.

1.2 Links to the theory of QSD

Let us denote by $(X_t)_{t\geq 0}$ an irreducible continuous-time non-explosive Markov chain on a discrete space E with generator Q. Let $\kappa : E \to \mathbb{R}_+$ be a uniformly bounded function. Consider the absorbing Markov chain $(Y_t)_{t\geq 0}$ on $E \cup \{\partial\}$, where $\partial \notin E$ is an absorbing state, satisfying

$$Y_t = \begin{cases} X_t & \text{if } \int_0^t \kappa(X_s)\mathrm{d}s < \xi \\ \partial & \text{otherwise,} \end{cases}$$

where ξ is an exponential random variable with parameter 1, independent from $(X_t)_{t\geq 0}$. In words, $(Y_t)_{t\geq 0}$ evolves as $(X_t)_{t\geq 0}$ on E and conditioned to be at $x \in E$, it jumps to the absorbing state ∂ with rate $\kappa(x)$. Let us denote by τ_∂ the absorption time. A quasi-stationary distribution (QSD) for $(Y_t)_{t\geq 0}$, is a probability measure μ_{QSD} on E such that

$$\lim_{t\to\infty} \mathbb{P}_\mu[Y_t \in \cdot \mid t < \tau_\partial] = \mu_{\mathrm{QSD}},$$

for some probability measure μ on E. When the limit above holds true for all $\mu = \delta_x$, with $x \in E$, the distribution μ_{QSD} is said to be the *minimal quasi-stationary distribution* of $(Y_t)_{t\geq 0}$. We refer the reader to [MV12; CMS13; DP13] and references therein for a review of the classical results concerning quasi-stationary distributions.

Note that the limit above can be written also as follows

$$\lim_{t\to\infty} \frac{\mathbb{E}_\mu\left[\phi(X_t)\mathrm{e}^{-\int_0^t \kappa(X_s)\mathrm{d}s}\right]}{\mathbb{E}_\mu\left[\mathrm{e}^{-\int_0^t \kappa(X_s)\mathrm{d}s}\right]} = \mu_{\mathrm{QSD}}(\phi),$$

for every function ϕ on E such that $\mu_{\mathrm{QSD}}(\phi)$ exists. This limit is analogous to (1.11), since the definition of μ_t is invariant by translation of κ, as we commented in Remark 1.1.4. Thus, Assumption (C2) is equivalent to the uniform exponential convergence of a conditioned Markov chain to its QSD, and the probability measure μ_∞ is indeed the QSD of the process driven by Q and with killing rate $\kappa := \sup_x \Lambda(x) - \Lambda \geq 0$.

Remark 1.2.1 (The Moran model with additive selection approaches a QSD). According to our previous discussion, when Assumptions (I), (C1) and (C2) are verified, Theorem 1.1.4 implies that the empirical probability measure induced by the multi-allelic Moran model approaches the law of the conditioned Markov chain. Moreover, when the process generated by $\mathcal{Q}$ allows a stationary distribution ν_N, Theorem 1.1.4 also implies that $m(\eta_\infty)$, where η_∞ is distributed according to ν_N, approaches the QSD of this absorbing Markov chain in the $\mathbb{L}^p$ distance with rate $1/\sqrt{N}$. In particular, the Moran model with selection at birth, when V^{d}_μ and V^{s}_μ in (C1) are null, also approximate the QSD of the absorbing Markov chain driven by Q and with killing rate $\kappa := \sup_{x\in E} V^{\mathrm{b}}(x) - V^{\mathrm{b}}$.

Relying on known results related to the exponential convergence to quasi-stationary distribution, we can discuss equivalent conditions to Assumption (C2), and provide explicit examples where this assumption holds.

The first example is precisely when E is finite. In this case inequality (1.11) was proved by Darroch and Seneta [DS67] and the result comes as a consequence of the Perron – Frobenius Theorem (see [MV12, Thm. 8] for the specific context of quasi-stationary distributions).

The case where E is countable is more delicate and has attracted lots of attention and several methods have been applied. Thanks to the exhaustive work of Champagnat and Villemonais, specifically [CV16; CV17b], it is possible to describe hypothesis equivalent to Assumption (C2) and explore the consequences of the uniform exponential convergence to the QSD.

Let us consider the following assumption, which is the translation of that in [CV16] under our notation:

Assumption (A)**.** There exists a $\nu \in \mathcal{M}_1(E)$ and three positive constants t_0, c_1, c_2 such that

(A1) for all $x \in E$,
$$\frac{\delta_x\left(P^\Lambda_{t_0}(\phi)\right)}{\delta_x\left(P^\Lambda_{t_0}(\mathbf{1})\right)} \geq c_1 \nu(\phi),$$
for every positive function $\phi \in \mathcal{B}_b(E)$,

(A2) for all $x \in E$ and $t \geq 0$,
$$\nu\left(P^\Lambda_t(\mathbf{1})\right) \geq c_2 \cdot \delta_x\left(P^\Lambda_t(\mathbf{1})\right).$$

The first condition in (A) is related to the fact that the process comes back fast in a finite subset of E with positive probability. This is associated to the idea of processes *coming down from infinity*. The second condition in (A) implies that the highest non-absorption probability among those starting from a singleton of E, has the same order of that starting from distribution ν. We refer to [CV16, §2] for a deeper analysis of the consequences of Assumption (A) and its equivalent formulations.

Then, we have the next result:

Theorem 1.2.1 (Theorem 2.1 in [CV16])**.** *Let us assume that* (C1) *if verified. Then, Assumptions* (A) *and* (C2) *are equivalent. In addition, if Assumption* (A) *is satisfied, then* (1.11) *holds with the explicit bound*
$$\left\| \frac{\mu_0 P^\Lambda_t}{\mu_0\left(P^\Lambda_t(\mathbf{1})\right)} - \mu_\infty \right\|_{\mathrm{TV}} \leq 2(1 - c_1 c_2)^{\lfloor t/t_0 \rfloor},$$
for every $\mu_0 \in \mathcal{M}_1(E)$.

See also the anterior work of Del Moral and Miclo [DM02, Prop. 2.3 and 3.1], which studies the large time behaviour and stability of Feynman – Kac semigroups in continuous time.

The distribution μ_∞ is the quasi-stationary distribution of the process driven by Q and with killing rates $\sup_x \Lambda - \Lambda(x) \geq 0$. It is well known, see e.g. [MV12], that there exists $\lambda \in \mathbb{R}$ such that, for all $t \geq 0$,

$$\mu_\infty \left(P_t^\Lambda(\mathbf{1}) \right) = \mathrm{e}^{\lambda t} \quad \text{and} \quad \frac{\mu_\infty P_t^\Lambda}{\mu_\infty \left(P_t^\Lambda(\mathbf{1}) \right)} = \mathrm{e}^{-\lambda t} \cdot \mu_\infty P_t^\Lambda = \mu_\infty.$$

In particular, $\lambda \leq 0$ whether $\Lambda \leq 0$.

The next two results explore the consequences of Assumption (A) to the spectrum of $Q + \Lambda$.

Theorem 1.2.2 (Theorem 2.1 in [CV17b])**.** *Assume that (A) is verified. There exist a positive function h on E and $C > 0$ such that*

$$|\mathrm{e}^{-\lambda t} \cdot \delta_x \left(P_t^\Lambda(\mathbf{1}) \right) - h(x)| \leq C \mathrm{e}^{-\lambda t} \cdot \delta_x \left(P_t^\Lambda(\mathbf{1}) \right) \mathrm{e}^{-\gamma t},$$

where $\gamma > 0$ is as in (1.11). Moreover, $\mu_\infty(h) = 1$ and

$$(Q + \Lambda)(h) = \lambda h.$$

Corollary 1.2.3 (Corollaire 2.4 in [CV16])**.** *Assume that (A) is verified. If $\varphi \in \mathcal{B}_b(E)$ is a right eigenfunction for $Q + \Lambda$ for an eigenvalue β, then either*

1. *$\beta = \lambda$ and $\varphi = \mu_\infty(\varphi) h$, or*
2. *$\beta \leq \lambda - \gamma$, $\mu_\infty(\varphi) = 0$,*

where γ is as in the statement of (C2).

Then, as we commented in Section 1.1.1, the proof of Lemma 1.1.7 is an immediate consequence of Theorem 1.2.2.

1.2.1 Examples

In this section we consider several examples where Assumption (C2) holds, for the process with additive selection satisfying (C1).

As we commented, this assumption is always verified when the state space is finite. The first example we consider is precisely when $E = \{1, 2\}$. This example offers us the opportunity to compare our result with the existing results on bi-allelic Moran models and the Fleming–Viot particle process approximating the QSD of an absorbing Markov chain with two transient states.

Example 1.2.1 (Two-allelic Moran model)**.** Consider the two-allelic Moran model on $E = \{1, 2\}$ with mutation rate matrix

$$Q = \begin{pmatrix} -a & a \\ b & -b \end{pmatrix}$$

and selection rates $V_{1,2} = p$ and $V_{2,1} = q$, with $a, b > 0$ and $p, q \geq 0$. Let us assume, without loss of generality, that $p \leq q$.

The empirical probability measure induced by this Moran process approaches the QSD of the absorbing Markov chain on $E \cup \{\partial\}$, where ∂ is an absorbing state, with infinitesimal generator

$$\begin{pmatrix} -(a+p) & a & p \\ b & -(b+q) & q \\ 0 & 0 & 0 \end{pmatrix}.$$

See [Cor17] and [CT16a, §3] for a deeper treatment of this model and the limit behaviour of the interacting particle process approaching its QSD.

Theorem 1.1.2 applied in this case improves Proposition 3.1 in [Cor17] and Theorem 3.1 (see also Remark 3.2) in [BC15]. Furthermore, Theorem 1.1.4, and also (1.12), improve the control of the speed of convergence to stationarity of the bounds obtained in [CT16b, Cor. 1.5] and [CV21, Thm. 2.4].

Likewise, as a consequence of Corollary 1.1.9, we have that the Moran model with the same mutation rate matrix Q and with selection rates $V_{1,2} = 0$ and $V_{2,1} = q - p$, approaches the same QSD but with smaller asymptotic square error.

Consider now the case when $p = q$. In this case the empirical distribution induced by the particle system approaches the stationary distribution of the process generated by Q. When E is finite, the results about the spectrum of the generator $\mathcal{Q}$ in [Cor21b] imply that the asymptotic ergodicity is independent of the value of p. Besides, Corollary 1.1.9 implies that a minimal asymptotic variance is obtained when there is no selection, that is, when the particle system is simply given by N independent particles, where each of them is driven by Q.

We now focus on the classical birth and death Markov chain. The existence and uniqueness of QSD for these models have been well understood. We rely on existing results to find explicit conditions on the parameters of the birth and death chain that are equivalent to the existence of a unique QSD and the uniform exponential convergence.

Example 1.2.2 (Birth and death chain)**.** Consider two positive sequences $(b_x)_{x\geq 1}$ and $(d_x)_{x\geq 1}$ and the Markov chain on $\mathbb{N}$ with rate matrix

$$Q_{x,y} := \begin{cases} b_x & \text{if } x \geq 1 \text{ and } y = x+1 \\ d_x & \text{if } x \geq 2 \text{ and } y = x-1 \\ 0 & \text{otherwise,} \end{cases}$$

and $\Lambda := d_1 \mathbf{1}_{\{x=1\}}$. Note that 0 is an absorbing state and $\mathbb{N}$ is a transient class. Van Doorn [Doo91, Thm. 3.2] has found explicit condition characterising the three possible cases: there is no QSD, there exists a unique QSD or there exists an infinite continuum of QSDs. See also [MV12, §4]. Furthermore, Martínez et al. [MMV14, Thm. 2] have proved that the existence of a unique QSD is in fact equivalent to the uniform exponential convergence of the law of the conditioned process to its QSD, which is in fact Assumption (C2). In addition, this occurs if and only if

$$\sum_{k\geq 2} \frac{1}{d_k \alpha_k} \sum_{r \geq k} \alpha_r < \infty, \tag{1.17}$$

where $\alpha_r := \prod_{i=1}^{r-1} b_i \Big/ \prod_{i=2}^{r} d_i$. We refer also to [CV16, §4.1], where the uniform exponential convergence is ensured for some generalisations of the classical birth and death chain.

We end this section presenting two quantitative criteria on the transition rates and on the spectral elements, respectively, ensuring the uniform exponential convergence in (C2).

Example 1.2.3 (A criterion on the mutation and selection rates)**.** We next describe a criterion on the transition rates, which is in fact Theorem 3 in [MMV14]. Assume that (C1) is verified and the following condition holds: there exists a finite subset $K \subset E$ such that

$$\inf_{y \in E \setminus K} \left(\Lambda(y) + \sum_{x \in K} Q_{y,x} \right) > \sup_{y \in E} \Lambda(y).$$

Then, (C2) holds. This provides an easy condition on the mutation rates and Λ to verify Assumption (C2), which is applicable to a wide range of Moran processes with discrete countable state space. See also [CT16b, Thm. 1.1], where a stronger condition is asked in order to provide, via a coupling technique, explicit constants for the upper bound in (C2).

Example 1.2.4 (A spectral criterion)**.** Assume there exists a triplet $(\mu_\infty, h, \lambda) \in \mathcal{M}_1(E) \times \mathcal{B}_b(E) \times \mathbb{R}$, of eigenelements of $Q + \Lambda$ such that λ is an eigenvalue of $Q + \Lambda$, h is strictly positive, $\mu_\infty(h) = 1$,

$$\mu_\infty P_t^\Lambda = \mathrm{e}^{\lambda t}\mu_\infty \text{ and } P_t^\Lambda(h) = \mathrm{e}^{\lambda t} h.$$

Note that these are the eigenelements in the statement of Lemma 1.1.7. Let us also assume $\|h^{-1}\| \leq \infty$, which is always true if E is finite, and furthermore, there exists $\epsilon > 0$ such that the set

$$K_\epsilon := \{x \in E : \Lambda(x) \geq \lambda - \epsilon\}$$

is finite. Then, (C2) is verified.

The proof is based on the methods in [DMT95], and is very similar to the proofs of Proposition 3.2 in [Rou06] and Proposition A.5 in [AGJ21]. Consider the Doob's h-transform

$$P_t^{\Lambda,h} := \frac{1}{h}\mathrm{e}^{-\lambda t} P_t^\Lambda(h\cdot),$$

which is the semigroup associated to an irreducible continuous-time Markov chain on E with generator Q^h acting on every $\phi \in \mathcal{M}_1(E)$ as follows

$$Q^h(\phi) = \frac{1}{h}(Q + \Lambda - \lambda)(h\phi).$$

Furthermore, the process driven by Q^h has stationary distribution $\mu_\infty^h \in \mathcal{M}_1(E)$, satisfying $\mu_\infty^h(\phi) = \mu_\infty(h\phi)$, for every $\phi \in \mathcal{B}_b(E)$. Now, note that h^{-1} is bounded on K_ϵ, and consequently there exists $\beta > 0$ such that

$$Q^h(h^{-1}) = \frac{\Lambda - \lambda}{h} \leq -\frac{\epsilon}{h} + \beta \mathbb{1}_{K_\epsilon}.$$

Thus, condition $(\tilde{D})$ in [DMT95] is verified and using their Theorem 5.2-(c) we get the h^{-1}-uniform exponential ergodicity of $P_t^{\Lambda,h}$ as follows

$$\sup_{|g| \leq h^{-1}} \left| \frac{1}{h(x)}\mathrm{e}^{-\lambda t}\delta_x P_t^\Lambda(hg) - \mu_\infty(hg) \right| \leq \frac{C}{h(x)}\rho^t.$$

Multiplying by $h(x)$ the previous inequality and taking $\phi = hg \in \mathcal{B}_1(E)$, we get the uniform exponential ergodicity (1.13). Finally, it is not difficult to verify that Assumption (C2) also holds, using the exponential ergodicity (1.13) and the fact that $\|h^{-1}\| < \infty$.

Remark 1.2.2. In [AGJ21, Appendix], the authors state a similar result but they do not include the fact that $\|h^{-1}\|$ is bounded in their hypothesis. We have not found or understood the arguments making the authors claim that $\|h^{-1}\|$ is bounded when the state space is locally compact [AGJ21, p. 150]. We next provide an example of a birth and death chain whose generator allows an unbounded eigenfunction associated to its greatest eigenvalue. Indeed, let us consider the following parameters for the birth and death chain in Example 1.2.2: $b_i = b$, and $d_i = d$, for all $i \geq 2$, with $b < d$. Moreover, take $b_1 > b$ and $d_1 = d(e-1)$. Hence, taking $h : n \in \mathbb{N} \mapsto \mathrm{e}^{-n}$ we get $(Q + \Lambda)(h) = \lambda h$, for $\lambda = b(\mathrm{e}^{-1} - 1) + d(e-1) > 0$. Moreover, $K_\epsilon = \{1\}$ is finite (compact), but $\|r^{-1}\| = \infty$. In fact, the infinite sum (1.17) diverges, thus this birth and death chain allows an infnite number of QSDs (cf. [Doo91, Thm. 3.2]).

1.2.2 Discussion

In this paper we study the speed of the convergence of the empirical distribution induced by a multi-allelic Moran model to a family of probability distributions on $\mathcal{M}_1(E)$, which is the solution of a second order nonlinear differential equation. In the case where the selection is additive in the sense of Assumption (C1), the limit is in fact the law of a absorbing Markov chain conditioned to non-absorption. The multi-allelic Moran model we study here contains as a special case the Fleming – Viot particle system, which is an interacting particle system intended for the approximation of a quasi-stationary distribution.

We also study the Moran model with additive selection for numerical approximation purposes. As we commented in Remark 1.2.1, the Moran model with additive selection always approaches a QSD, which depends on the additive selection expression in (C1) only through the function Λ. Actually, one of the main goals of this article is to strengthen the relationship between research on Moran models with additive selection and Fleming – Viot particle process. Theorem 1.1.8 ensures to asymptotic normality of the approximation error made by the empirical probability measure induced by the particle system, when N is large. Using this result, Corollary 1.1.9 can be used to define another particle process approaching the same QSD with smaller asymptotic quadratic error, given a Moran process satisfying (C1) and (C2). However, the problem of finding the optimal selection rates for a fixed function Λ remains open.

The fact that the state space E is discrete is not necessarily for our proofs. Therefore, we expect to be able to extend all our results to more general Markov processes following the same methods.

There are lots of possible directions to continue this research. Maybe, the more natural is to weaken the condition (C2) and consider the case where there exists a minimal QSD but the exponential convergence is not uniform on $\mathcal{M}_1(E)$. Lots of research have been done for controlling the domains of attraction of the minimal QSD. See for example the works of Champagnat and Villemonais [Vil15; CV17a; CV20a; CV20b; CV21] and also the related works of Bansaye et al. [Ban+21; BCG20], and the references therein. Another interesting research direction is to improve the upper bound constants $C_{p,T}$ in Theorem 1.1.2. In this sense, the results of Arnaudon and Del Moral [AD20, Thm. 5.10 and Cor. 5.12] suggest that a bound of type $C_{p,T} = C_p T$ could hold. Hence, a future research direction would be to combine the approach of [AD20] and this paper to improve the upper bound in Theorem 1.1.2. Moreover, the results in [AD20, §5] could also be useful to obtain exponential concentration inequalities, which is a natural continuation of the research on the long time behaviour of the empirical measure induced by Moran type particle processes.

1.3 Proof of the main results

1.3.1 The associated martingale problem

For a Markovian generator L, its associated “carré-du-champ” operator, denoted Γ_L, is defined by

$$\Gamma_L : \phi \mapsto L(\phi^2) - 2\phi L\phi.$$

See, for example [Ané+00, Def. 2.5.1] for more details on the theory related to this operator.

It is not difficult to prove that $\Gamma_{\mathcal{Q}}$ satisfies

$$\Gamma_{\mathcal{Q}}(\psi)(\eta) = \sum_{x\in E} \eta(x) \sum_{y\in E} \left(Q_{x,y} + V_{m(\eta)}(x,y)\frac{\eta(y)}{N} \right) [\psi(\eta - \mathbf{e}_x + \mathbf{e}_y) - \psi(\eta)]^2,$$

where for every $\eta \in \mathcal{E}_N$. We recall that $m(\eta)$ denotes the empirical distribution induced by $\eta \in \mathcal{E}_N$. Moreover, $m(\eta)(\phi)$ stands for the mean of ϕ with respect to $m(\eta)$, for every $\phi \in \mathcal{B}_b(E)$.

Suppose that one of Assumptions (G1) or (C1) is verified. In either case, let us denote $V_\mu^\star := V_\mu - V_\mu^{\mathrm{s}}$.

Lemma 1.3.1. *Suppose that one of Assumptions (G1) or (C1) is verified. We have*

$$\mathcal{Q}(m(\cdot)(\phi)) = m(\cdot)\left(Q^\star_{m(\cdot)}(\phi)\right),$$
$$\Gamma_{\mathcal{Q}}(m(\cdot)(\phi)) = \frac{1}{N} m(\cdot)\left(\Gamma_{Q_{m(\cdot)}}(\phi)\right),$$

where

$$Q^\star_\mu \phi : x \mapsto (Q\phi)(x) + \sum_{y\in E} \mu(y) V^\star_\mu(x,y)[\phi(y) - \phi(x)],$$
$$Q_\mu \phi : x \mapsto (Q\phi)(x) + \sum_{y\in E} \mu(y) V_\mu(x,y)[\phi(y) - \phi(x)],$$

for every $\phi \in \mathcal{B}_b(E)$ and all $x \in E$.

Proof. The first equality is simply a consequence of (1.2) and (1.3), and the fact that

$$\mu(Q_\mu \phi) = \mu(Q^\star_\mu \phi). \tag{1.18}$$

Now, to prove the second equality, note that

$$\begin{aligned}
\Gamma_{\mathcal{Q}}(m(\cdot)(\phi))(\eta) &= \sum_{x\in E} \eta(x) \sum_{y\in E} \left(Q_{x,y} + V_{m(\eta)}(x,y)\frac{\eta(y)}{N}\right)[m(\eta - \mathbf{e}_x + \mathbf{e}_y)(\phi) - m(\eta)(\phi)]^2 \\
&= \frac{1}{N}\sum_{x\in E} \frac{\eta(x)}{N} \sum_{y\in E} \left(Q_{x,y} + V_{m(\eta)}(x,y)\frac{\eta(y)}{N}\right)[\phi(y) - \phi(x)]^2 \\
&= \frac{1}{N}\sum_{x\in E} \left(\sum_{y\in E} (Q_{x,y} + V_{m(\eta)}(x,y) m_y(\eta))[\phi(y) - \phi(x)]^2\right) m_x(\eta) \\
&= \frac{1}{N} m(\eta)\left(\Gamma_{Q_{m(\eta)}}(\phi)\right). \qquad \square
\end{aligned}$$

Using the Lemma 1.3.1 we can study the martingale problem associated to the process $(m(\eta_t)(\psi_t))_{t\geq 0}$.

Proposition 1.3.2 (Martingale decomposition)**.** *Let ψ be a function on $E \times \mathbb{R}_+$ such that $\psi_\cdot(x)$ is continuously differentiable in $\mathbb{R}_+$, for every $x \in E$ and $\psi_t(\cdot) \in \mathcal{B}_b(E)$, for every $t \in \mathbb{R}_+$. Then, the process $(\mathcal{M}_t(\psi_\cdot))_{t\geq 0}$ such that*

$$\mathcal{M}_t(\psi_\cdot) := m(\eta_t)(\psi_t) - m(\eta_0)(\psi_0) - \int_0^t m(\eta_s)\left(\partial_s \psi_s + Q^\star_{m(\eta_s)}(\psi_s)\right) \mathrm{d}s,$$

where $Q^\star_\mu$ is defined as in Lemma 1.3.1, is a local martingale, with predictable quadratic variation *given by*

$$\langle \mathcal{M}(\psi_\cdot)\rangle_t = \frac{1}{N}\int_0^t m(\eta_s)\left(\Gamma_{Q_{m(\eta_s)}}(\psi_s)\right)\mathrm{d}s.$$

Moreover,

$$|\Delta \mathcal{M}_t(\psi_t)| \leq \frac{2\|\psi_t\|}{N}.$$

Proof. The usual martingale problem associated to $(\eta_t)_{t\geq 0}$ implies that for every function ϕ on E, the process

$$\begin{aligned} t \mapsto m(\eta_t)(\phi) - m(\eta_0)(\phi) - \int_0^t \mathcal{Q}(m(\eta_s)(\phi))\mathrm{d}s \\ = m(\eta_t)(\phi) - m(\eta_0)(\phi) - \int_0^t m(\eta_s)(Q^\star_{m(\eta_s)}(\phi))\mathrm{d}s \end{aligned}$$

is a local martingale. Note that the equality is due to the first identity in Lemma 1.3.1. Then, for a function ψ on $E \times \mathbb{R}_+$, continuously differentiable in $\mathbb{R}_+$, the Itô formula implies that $(\mathcal{M}_t(\psi_\cdot))_{t\geq 0}$ is a local martingale, as desired.

The predictable quadratic variation is obtained using that

$$\langle \mathcal{M}(\psi)\rangle_t = \int_0^t \Gamma_{\mathcal{Q}}(m(\eta_s)(\psi_s))\mathrm{d}s,$$

and the final result comes from the second identity in Lemma 1.3.1.

The bound for the jump is due to the fact that each jump only concerns one particle that jumps from one position to another. □

Now, for a function ψ on $E \times \mathbb{R}_+$, continuously differentiable in $\mathbb{R}_+$, we get

$$\mathrm{d}m(\eta_t)(\psi_t) = \mathrm{d}\mathcal{M}_t(\psi_\cdot) + m(\eta_t)(\partial_t \psi_t + Q^\star_{m(\eta_t)}(\psi_t))\mathrm{d}t.$$

Thus, the empirical measure induced by the particle process is a perturbation of the dynamic given by (1.4), by a martingale whose jumps and predictable quadratic variation are of order $\frac{1}{N}$.

1.3.2 Proof of Theorem 1.1.2

Throughout this section we will suppose that the expression for the selection rates in Assumption (G1) is verified. We will denote $Q^\star_\mu = Q + \Pi^\star_\mu$ as in Lemma 1.3.1, namely

$$\Pi^\star_\mu \phi : x \mapsto \sum_{y\in E} \mu(y) V^\star(x,y)[\phi(y) - \phi(x)],$$

where

$$V^\star(x,y) := V_\mu(x,y) - V^{\mathrm{s}}_\mu(x,y) = \sum_{i\geq 1} V^{\mathrm{d}}_i(x) V^{\mathrm{b}}_i(y),$$

which is independent of $\mu \in \mathcal{M}_1(E)$.

The family of generators $(Q^\star_{\mu_t})_{t\geq 0}$ defines an inhomogeneous-time Markov chain, which is associated to a map $(s,t) \mapsto P(s,t)$, for all $s \leq t$ such that $P(s,s) = I$, for all $s \geq 0$ and satisfies the forward and backward Kolmogorov equations:

$$\begin{aligned} \partial_t P(s,t) &= P(s,t) Q^\star_{\mu_t}, \text{ for } t \geq s, \\ \partial_s P(s,t) &= -Q^\star_{\mu_s} P(s,t), \text{ for } s \leq t. \end{aligned} \tag{1.19}$$

See [FMS14] and the references therein. Moreover, using the forward Kolmogorov equation (1.19), we get that $(\mu_t)_{t\geq 0}$ as in Assumption (G2), satisfies the propagation equation $\mu_T = \mu_t P(t,T)$. Note that since $P(t,T)$ is the propagator of an inhomogenous Markov chain, we get $\|P(t,T)\| \leq 1$ for all $t \in [0,T]$, which implies

$$\int_0^T \|P(s,T)(\phi)\|^p \mathrm{d}s \leq T - t. \tag{1.20}$$

Let us now study the control in $\mathbb{L}^p$ norm for the martingales that are obtained taking the functions $t \in [0,T] \mapsto P(\cdot,T)(\phi)$ and $t \in [0,T] \mapsto P(\cdot,T)(\phi)^2$ in Proposition 1.3.2. Note that,

$$\partial_t \left(P(t,T)(\phi)\right) = -Q^\star_{\mu_t}\left(P(t,T)(\phi)\right).$$

From Proposition 1.3.2, we get the following local martingale for $t \in [0,T]$:

$$\begin{aligned}\mathcal{M}_t\left(P(\cdot,T)(\phi)\right) := \; & m(\eta_t)\left(P(t,T)(\phi)\right) - m(\eta_0)\left(P(0,T)(\phi)\right) \\ & - \int_0^t m(\eta_s)\left(Q^\star_{m(\eta_s)}\Big(P(s,T)(\phi)\Big) - Q^\star_{\mu_s}\Big(P(s,T)(\phi)\Big)\right) \mathrm{d}s.\end{aligned}$$

Similarly, we get

$$\begin{aligned}\mathcal{M}_t\left(P(\cdot,T)(\phi)^2\right) :=& \, m(\eta_t)\left(P(t,T)(\phi)^2\right) - m(\eta_0)\left(P(0,T)(\phi)\right) \\ & - \int_0^t m(\eta_s)\left(Q^\star_{m(\eta_s)}\Big(P(s,T)(\phi)^2\Big) - 2P(s,T)(\phi)\; Q^\star_{\mu_s}\Big(P(s,T)(\phi)\Big)\right) \mathrm{d}s.\end{aligned} \tag{1.21}$$

Moreover, by definition, $P(T,T)(\phi) = \phi$.

Lemma 1.3.3 (Control of the predictable quadratic variation)**.** *Assume that Assumption (G1) is verified. For every test function $\phi \in \mathcal{B}_1(E)$, we have*

$$N\Big\langle \mathcal{M}\left(P(\cdot,T)(\phi)\right)\Big\rangle_t \leq C(t+1) - \mathcal{M}_t\Big(P(\cdot,T)(\phi)^2\Big), \quad \textit{for all} \quad t \in [0,T].$$

Proof. The predictable quadratic variation of the martingale $\left(\mathcal{M}_t\left(P(\cdot,T)(\phi)\right)\right)_{t\in[0,T]}$ satisfies

$$\begin{aligned} N\Big\langle \mathcal{M}\Big(P(\cdot,T)(\phi)\Big)\Big\rangle_t &= \int_0^t m(\eta_s)\Big(\Gamma_{Q_{m(\eta_s)}}\left(P(s,T)(\phi)\right)\Big)\,\mathrm{d}s \\ &= \int_0^t m(\eta_s)\left(Q^\star_{m(\eta_s)}\left(P(s,T)(\phi)^2\right) - 2P(s,T)(\phi)\cdot Q_{m(\eta_s)}\Big(P(s,T)(\phi)\Big)\right)\mathrm{d}s, \end{aligned}$$

where the second equality holds because of the definition of carré-du-champ operator and (1.18).

Thus, using (1.21) we get

$$\begin{aligned} N\Big\langle \mathcal{M}\Big(P(\cdot,T)(\phi)\Big)\Big\rangle_t =& -\mathcal{M}_t\Big(P(\cdot,T)(\phi)^2\Big) - m(\eta_t)\Big(P(t,T)(\phi)^2\Big) + m(\eta_0)\Big(P(0,T)(\phi)^2\Big) \\ & + 2\int_0^t m(\eta_s)\Big(P(s,T)(\phi)\cdot\Big[\Big(Q^\star_{\mu_s} - Q_{m(\eta_s)}\Big)\Big(P(s,T)(\phi)\Big)\Big]\Big)\mathrm{d}s. \end{aligned}$$

Now, because of (1.20) and the boundedness conditions on V_μ in Assumption (G1) we can ensure the existence of a constant $C > 0$ such that

$$N\Big\langle \mathcal{M}\Big(P(\cdot,T)(\phi)\Big)\Big\rangle_t \leq C(t+1) - \mathcal{M}_t\left(P(\cdot,T)(\phi)^2\right). \qquad \square$$

The following lemma is a generalisation of the classical Burkholder – Davis – Gundy (BDG) inequality [Kal21, Thm. 20.12]. The lower bound is obtained from the classical BDG inequality. The proof of the upper bound can be found in [Rou06, Lemma 6.2].

Lemma 1.3.4 (BDG inequalities)**.** *Let $(\mathcal{M}_t)_{t\geq 0}$ be a quasi-left-continuous (i.e. with continuous predictable increasing process) locally square-integrable martingale with $M_0 = 0$ and bounded jumps*

$$\sup_{0\leq t\leq T} |\Delta\mathcal{M}_t| \leq a < +\infty.$$

Then, there exists a constant C, possibly depending on q, such that

$$\mathbb{E}\left[\left(\sup_{t\in[0,T]}\mathcal{M}_t\right)^{2^{q+1}}\right]\leq C\mathbb{E}\left[\left([\mathcal{M}]_T\right)^{2^q}\right]\leq C\sum_{k=0}^{q}a^{2^{q+1}-2^{k+1}}\mathbb{E}\left[\left(\langle\mathcal{M}\rangle_T\right)^{2^k}\right].$$

We are now in position to establish a control on quadratic variation of the martingale $\Big(\mathcal{M}_t\big(P(\cdot,T)(\phi)\big)\Big)_{t\in[0,T]}$.

Theorem 1.3.5 (Control of the quadratic variation)**.** *Assume that Assumption (G1) is verified. For all $p>0$ and all test function $\phi\in\mathcal{B}_1(E)$, there exists a positive C (possibly depending on p) such that*

$$\mathbb{E}\left[\left(\Big[\mathcal{M}\Big(P(\cdot,T)(\phi)\Big)\Big]_t\right)^p\right]\leq\frac{C(t+1)^p}{N^p},\quad \textit{for all}\quad t\in[0,T]$$

The proof of this result we provide below is inspired by the proof of Theorem 5.4 in [Rou06].

Proof. First, by localisation, we can suppose that the martingales are bounded. Now, we will prove the inequalities for $p=2^q$, and then using the Jensen inequality we will extend the result for all $p\geq1$. The result for $p\in(0,1)$ is simply a consequence of the result for $p=1$ and the Jensen inequality for concave functions.

We want to prove the following inequalities:

$$\mathbb{E}\left[\left(\Big\langle\mathcal{M}\Big(P(\cdot,T)(\phi)\Big)\Big\rangle_t\right)^{2^q}\right]\leq\frac{C(t+1)^{2^q}}{N^{2^q}},\quad \mathbb{E}\left[\left(\Big[\mathcal{M}\Big(P(\cdot,T)(\phi)\Big)\Big]_t\right)^{2^q}\right]\leq\frac{C(t+1)^{2^q}}{N^{2^q}}.$$

For $q=0$, the first inequality is consequence of Lemma 1.3.3 and the second one is due to the fact that $\left(\Big[\mathcal{M}\Big(P(\cdot,T)(\phi)\Big)\Big]-\Big\langle\mathcal{M}\Big(P(\cdot,T)(\phi)\Big)\Big\rangle\right)_{t\in[0,T]}$ is a local martingale.

We will prove the previous inequalities by induction. Let us assume they are true for q and lower. Thus, by Lemma 1.3.3 and the Minkowski inequality, there exists a $K>0$ such that

$$\begin{aligned}
I_p&:=\mathbb{E}\left[\left(N\Big\langle\mathcal{M}(P(\cdot,T)(\phi))\Big\rangle_t\right)^p\right]\\
&\leq\mathbb{E}\left[\left(K(t+1)+\Big|\mathcal{M}_t\Big(P(\cdot,T)(\phi)^2\Big)\Big|\right)^p\right]\\
&\leq\left(K(t+1)+\left(\mathbb{E}\left[\Big|\mathcal{M}_t\left(P(\cdot,T)(\phi)^2\right)\Big|^p\right]\right)^{1/p}\right)^p,
\end{aligned}$$

for all $p\geq1$. Using now the BDG inequality we get

$$\begin{aligned}
I_{2^{q+1}}&\leq\left(K(t+1)+\kappa\left(\mathbb{E}\left[\left(\Big[\mathcal{M}\Big(P(\cdot,T)(\phi)^2\Big)\Big]_t\right)^{2^q}\right]\right)^{1/2^{q+1}}\right)^{2^{q+1}}\\
&\leq\left(K(t+1)+\kappa\sqrt{\frac{t+1}{N}}\right)^{2^{q+1}}\leq C'(t+1)^{2^{q+1}},
\end{aligned}$$

where the second inequality holds by the induction hypothesis and the last one due to $N\geq1$ and $t+1\geq1$.

Now, the martingale $(\mathcal{M}_t(P(\cdot,T)(\phi)))_{t\in[0,T]}$ has jumps verifying

$$a\leq2\frac{\|P(\cdot,T)(\phi)\|}{N}\leq\frac{2}{N}.$$

Thus, using Lemma 1.3.4 we get

$$\mathbb{E}\left[\left([\mathcal{M}(P(\cdot,T)(\phi))]_t\right)^{2^{q+1}}\right] \leq C'' \sum_{k=0}^{q+1} \frac{\mathbb{E}\left[\left(\langle \mathcal{M}(P(\cdot,T)(\phi))\rangle_t\right)^{2^k}\right]}{N^{2^{q+2}-2^{k+1}}} \leq C'' \sum_{k=0}^{q+1} \frac{(t+1)^{2^k}}{N^{2^{q+2}-2^{k+1}+2^k}}$$
$$= \frac{C''}{N^{2^{q+2}}} \sum_{k=0}^{q+1} [N(t+1)]^{2^k} = \frac{C''(q+1)}{N^{2^{q+2}}} N^{2^{q+1}} (t+1)^{2^{q+1}}$$
$$\leq C \frac{(t+1)^{2^{q+1}}}{N^{2^{q+1}}}.$$

This concludes the proof for $p = 2^q$.

Now, for arbitrary p, there exists q such that $p \leq 2^q$. Thus, using the Jensen inequality (for the concave function $x \mapsto x^{p/2^q}$) we get

$$\mathbb{E}\left[\left(\left\langle \mathcal{M}\left(P(\cdot,T)(\phi)\right)\right\rangle_t\right)^p\right] \leq \left(\mathbb{E}\left[\left(\left\langle \mathcal{M}\left(P(\cdot,T)(\phi)\right)\right\rangle_t\right)^{2^q}\right]\right)^{p/2^q}$$
$$\leq \left(\frac{C(t+1)^{2^q}}{N^{2^q}}\right)^{p/2^q} \leq C^{p/2^q} \frac{(t+1)^p}{N^p}.$$

The result for $\mathbb{E}\left[\left([\mathcal{M}(P(\cdot,T)(\phi))]_t\right)^p\right]$ is analogously obtained. $\square$

Proof of Theorem 1.1.2. Let us denote $\psi_{s,T} := P(s,T)(\bar{\phi}_T)$, which satisfies the backward Kolmogorov equation (1.19).

We have that $(\mathcal{M}_t(\psi_{\cdot,T}))_{t\in[0,T]}$, defined as in Proposition 1.3.2, is a local martingale. Moreover,

$$\begin{aligned}\mathcal{M}_T(\psi_{\cdot,T}) &= m(\eta_T)(\psi_{T,T}) - m(\eta_0)(\psi_{0,T}) - \int_0^T m(\eta_s)\left(-Q^\star_{\mu_s}(\psi_{s,T}) + Q^\star_{m(\eta_s)}(\psi_{s,T})\right) \mathrm{d}s \\ &= m(\eta_T)(\phi) - \mu_T(\phi) - m(\eta_0)(\psi_{0,T}) - \int_0^T m(\eta_s)\left(\Pi^\star_{m(\eta_s)}(\psi_{s,T}) - \Pi^\star_{\mu_s}(\psi_{s,T})\right) \mathrm{d}s.\end{aligned} \tag{1.22}$$

Note that for any two probability measures λ and μ on E and for every function ψ on E we have

$$\lambda(\Pi^\star_\mu(\psi)) = -\mu(\widetilde{\Pi}_\lambda(\psi)), \tag{1.23}$$

where $\widetilde{\Pi}_\lambda$ acts on a test function ψ as follows:

$$\widetilde{\Pi}_\lambda(\psi) : x \mapsto \sum_{y\in E} \lambda(y) V^\star(y,x)[\psi(y) - \psi(x)].$$

Indeed, note that

$$\lambda\left(\Pi^\star_\mu(\psi)\right) = \sum_{x,y\in E} \lambda(x)\mu(y)V^\star(x,y)[\psi(y)-\psi(x)] = -\sum_{y\in E}\left(\sum_{x\in E}\lambda(x)V^\star(y,x)[\psi(x)-\psi(y)]\right)\mu(y)$$
$$= -\mu\left(\widetilde{\Pi}_\lambda(\psi)\right).$$

Using (1.22) and (1.23) we get

$$\mathcal{M}_T(\psi_{\cdot,T}) = m(\eta_T)(\phi) - \mu_T(\phi) - m(\eta_0)(\psi_{0,T}) + \int_0^T (m(\eta_s) - \mu_s)\left(\widetilde{\Pi}_{m(\eta_s)}(\psi_{s,T})\right) \mathrm{d}s, \tag{1.24}$$

where $\psi_{s,T} = P(s,T)(\bar{\phi}_T)$.

Hence, using (1.24), we can ensure the existence of a positive constant $C > 0$ such that

$$\sup_{t\le T} |m(\eta_t)(\phi) - \mu_t(\phi)|^p \le C\Big(|m(\eta_0)(\psi_{0,T}) - \mu_0(\psi_{0,T})|^p + \sup_{t\le T} |\mathcal{M}_t(\psi_{\cdot,T})|^p + R_p(T)\Big),$$

where

$$R_p(T) = \int_0^T \left|(m(\eta_s) - \mu_s)\left(\widetilde{\Pi}_{m(\eta_s)}(\psi_{s,T})\right)\right|^p \mathrm{d}s.$$

The initial error can be controlled using Assumption (I). Indeed, there exists $C_1 > 0$ such that

$$\mathbb{E}[|m(\eta_0)(\psi_{0,T}) - \mu_0(\psi_{0,T})|^p] \le \frac{C_1}{N^{p/2}}.$$

Furthermore, using Theorem 1.3.5 and BDG inequality we get

$$\mathbb{E}\left[\sup_{t\le T}\left|\mathcal{M}_t\Big(P(\cdot,t)(\bar{\phi}_T)\Big)\right|^p\right] \le \frac{C_2(T+1)^{p/2}}{N^{p/2}},$$

for all $p \ge 1$. Let us denote by λ_s, the (random) signed measure

$$\lambda_s := m(\eta_s) - \mu_s.$$

We have

$$R_p(T) \le C\left(\int_0^T \left|\lambda_s\left(\widetilde{\Pi}_{\mu_s}(\psi_{s,T})\right)\right|^p \mathrm{d}s + \int_0^T \left|\lambda_s\left(\widetilde{\Pi}_{\lambda_s}(\psi_{s,T})\right)\right|^p \mathrm{d}s\right).$$

The first term in the last expression can be controlled, since $\widetilde{\Pi}_{\mu_s}(\psi_{s,T})$ is not random. Indeed,

$$\begin{aligned} I_1(T) &:= \int_0^T \left|(m(\eta_s) - \mu_s)\left(\widetilde{\Pi}_{\mu_s}(\psi_{s,T})\right)\right|^p \mathrm{d}s \\ &\le 2^p\|V^\star\|^p \int_0^T \left|(m(\eta_s) - \mu_s)\left(\frac{\widetilde{\Pi}_{\mu_s}(\psi_{s,T})}{2\|V^\star\|}\right)\right|^p \mathrm{d}s. \end{aligned}$$

For the second term, note that

$$\begin{aligned} I_2(T) &:= \int_0^T \left|\sum_{i\ge1}\sum_{x,y\in E} \lambda_s(x)\lambda_s(y)V_i^{\mathrm{d}}(x)V_i^{\mathrm{b}}(y)\Big[\psi_{s,T}(y) - \psi_{s,T}(x)\Big]\right|^p \mathrm{d}s \\ &= \int_0^T \left|\sum_{i\ge1} \lambda_s(V_i^{\mathrm{d}} - V_i^{\mathrm{b}})\lambda_s(V_i^{\mathrm{b}}\psi_{s,T}) + \lambda_s(V_i^{\mathrm{b}})\lambda_s\Big((V_i^{\mathrm{b}} - V_i^{\mathrm{d}})\psi_{s,T}\Big)\right|^p \mathrm{d}s. \end{aligned}$$

Hence, using the Assumption (G1), we conclude that there exists a positive constant C_p, depending on p, such that

$$I_2(T) \le C_p \int_0^T \left|(m(\eta_s) - \mu_s)\left(\frac{1}{\kappa}\sum_{i\ge1}|V_i^{\mathrm{d}} - V_i^{\mathrm{b}}|\right)\right|^p + \left|(m(\eta_s) - \mu_s)\left(\frac{|\psi_{s,T}|}{\kappa}\sum_{i\ge1}|V_i^{\mathrm{d}} - V_i^{\mathrm{b}}|\right)\right|^p \mathrm{d}s,$$

where $\kappa = \left\|\sum_{i\ge1}|V_i^{\mathrm{d}} - V_i^{\mathrm{b}}|\right\| < \infty$.

Let us define

$$\Phi_p(t) := \sup_{\phi\in\mathcal{B}_1(E)} \mathbb{E}\left[\sup_{s\le t}|m(\eta_s)(\phi) - \mu_s(\phi)|^p\right].$$

Thus, taking the expectations of $I_1(T)$ and $I_2(T)$, we can ensure the existence of a positive constant B_p such that

$$\mathbb{E}[R_p(T)] \leq B_p \int_0^T \Phi_p(s)\mathrm{d}s.$$

Hence, there exists $K_{p,T} > 0$ such that

$$\Phi_p(T) \leq \frac{K_{p,T}}{N^{p/2}} + B_p \int_0^T \Phi_p(s)\mathrm{d}s,$$

which, using Grönwall inequality, gives

$$\Phi_p(T) \leq \frac{K_{p,T}}{N^{p/2}} \mathrm{e}^{B_p T}.$$

□

Proof of Corollary 1.1.3. Let $(x_n)_{n\geq 1}$ be the enumeration of the elements in E, in the definition of the distance $\|\cdot\|_{\mathrm{w}}$. Note that

$$\begin{aligned}
\mathbb{E}\left[\sup_{t\in[0,T]} \left(\|m(\eta_t) - \mu_t\|_{\mathrm{w}}\right)^p\right]^{1/p} &= \mathbb{E}\left[\sup_{t\in[0,T]} \left(\sum_{k\geq 1} 2^{-k} |m(\eta_t)(x_k) - \mu_t(x_k)|\right)^p\right]^{1/p} \\
&\leq \mathbb{E}\left[\left(\sum_{k\geq 1} 2^{-k} \sup_{t\in[0,T]} |m(\eta_t)(x_k) - \mu_t(x_k)|\right)^p\right]^{1/p} \\
&\leq \sum_{k\geq 1} 2^{-k} \mathbb{E}\left[\sup_{t\in[0,T]} |m(\eta_t)(x_k) - \mu_t(x_k)|^p\right]^{1/p} \\
&\leq \frac{C_{p,T}}{\sqrt{N}},
\end{aligned}$$

where the first inequality comes from interchanging the supremum and the infinite sum, the second is a consequence of the Minkowski's inequality for infinite sums and the last inequality is a consequence of Theorem 1.1.2. □

1.3.3 Proof of Theorems 1.1.4 and 1.1.6

In the rest of the paper we will assume that Assumption (C1) is verified. Namely, the selection rates can be written as follows

$$V_\mu(x,y) = V_\mu^{\mathrm{d}}(x) + V_\mu^{\mathrm{b}}(y) + V_\mu^{\mathrm{s}}(x,y),$$

where V_μ^{s} is symmetric. Note that, as we commented in Remark 1.1.3, under this assumption equation (1.4) becomes equivalent to

$$\partial_t \gamma_t(\phi) = \gamma_t\big((Q+\Lambda)\phi - \gamma_t(\Lambda)\phi\big),$$

where $\Lambda = V_\mu^{\mathrm{b}} - V_\mu^{\mathrm{d}}$. Moreover, we recall that the solution of this ordinary differential equation is given by the normalised Feynman−Kac semigroup:

$$\mu_t(\phi) := \frac{\mu_0 P_t^\Lambda(\phi)}{\mu_0 P_t^\Lambda(\mathbf{1})},$$

where

$$P_t^\Lambda(\phi) : x \mapsto \mathbb{E}_x\left[\phi(X_t)\exp\left\{\int_0^t \Lambda(X_s)\mathrm{d}s\right\}\right]$$

is the Feynman – Kac semigroup with generator $Q + \Lambda$.

Using (1.7) we can simplify the expression of the martingale $\big(\mathcal{M}(\psi_\cdot)\big)_{t\geq 0}$ in Proposition 1.3.2 as follows

$$\mathcal{M}_t(\psi_\cdot) = m(\eta_\cdot)(\psi_\cdot) - m(\eta_0)(\psi_0) - \int_0^\cdot m(\eta_s)(\partial_s\psi_s + (Q+\Lambda)(\psi_s) - m(\eta_s)(\Lambda)\cdot\psi_s)\mathrm{d}s, \quad (1.25)$$

for every bounded function ψ on $E\times\mathbb{R}$, such that $\psi_\cdot(x)$ is continuously differentiable in $\mathbb{R}_+$, for every $x\in E$; and $\psi_t(\cdot)\in\mathcal{B}_b(E)$, for every $t\in\mathbb{R}$. The previous expression is essential: for a suitable choice of the function ψ, we can control the integral part in the expression of $\mathcal{M}(\psi_\cdot)$.

When V is symmetric, and thus Λ is null, we obtain the following result as an immediate consequence of (1.25).

Corollary 1.3.6 (V_μ is symmetric)**.** *Assume that Assumption (C1) is verified in such a way that $V_\mu = V_\mu^{\mathrm{s}}$, and Assumptions (I) and (C2) are also verified. Then the process*

$$\Big(m(\eta_t)(\mathrm{e}^{(T-t)Q}(\phi)) - m(\eta_0)(\mathrm{e}^{TQ}(\phi))\Big)_{t\in[0,T]},$$

is a local martingale, for every $\phi\in\mathcal{B}_b(E)$. In particular, $\bar{m}(\eta_t) = \bar{m}(\eta_0)\mathrm{e}^{tQ}$, for all $t\geq 0$.

Proof of Corollary 1.3.6. Note that since the selection rates are symmetric, Λ as defined in (1.5) is null. The proof simply follows as a consequence of (1.25), taking $\psi_t = \mathrm{e}^{(T-t)Q}(\phi)$, for all $t\in[0,T]$ and $\phi\in\mathcal{B}_b(E)$. □

Let us define the operator

$$W_{t,T} : \phi \mapsto \frac{P^\Lambda_{T-t}(\phi)}{\mu_t P^\Lambda_{T-t}(\mathbf{1})},$$

which verifies the propagation equation $\mu_T(\phi) = \mu_t(W_{t,T}(\phi))$. Recall that $\bar{\phi}_T := \phi - \mu_T(\phi)$. We get $m(\eta_T)(W_{T,T}(\bar{\phi}_T)) = m(\eta_T)(\phi) - \mu_T(\phi)$, which is the difference we intend to control.

The following results establish a control of the uniform norm of $W_{t,T}$.

Lemma 1.3.7. *The operator $(W_{t,T})_{t\in[0,T]}$ verifies the following properties:*

a) Given $p\geq 1$, for any test function $\phi\in\mathcal{B}_1(E)$, there exists $C>0$ such that

$$\|W_{t,T}(\phi)\| \leq C, \quad \textit{and} \quad \int_t^T \|W_{s,T}(\phi)\|^p\mathrm{d}s \leq C(T-t).$$

b) There exists a $\rho\in(0,1)$, such that

$$\|W_{t,T}(\bar{\phi}_T)\| \leq C\rho^{T-t}, \quad \textit{and} \quad \int_t^T \|W_{s,T}(\bar{\phi}_T)\|^p\mathrm{d}s \leq C.$$

Proof of Lemma 1.3.7. The proof of this result is inspired by the proof of Lemma 5.1 in [Rou06], but we do not make any direct assumption of the spectrum of $Q+\Lambda$.

Note that $\mu_t(P^\Lambda_{T-t}(\mathbf{1})) = \mu_0 P^\Lambda_T(\mathbf{1})/\mu_0 P^\Lambda_t(\mathbf{1})$. Moreover, using Corollary 1.1.7, the function $t\mapsto \mathrm{e}^{-\lambda t}\mu_0(P^\Lambda_t(\mathbf{1}))$ is continuous and positive, going from 1 to $\mu_0(h)>0$. This proves part a).

To prove part b) of the lemma, note that

$$\mu_T(\phi) = \frac{\mu_t P^\Lambda_{T-t}(\phi)}{\mu_t P^\Lambda_{T-t}(\mathbf{1})} \quad \text{and} \quad W_{t,T}(\mu_T(\phi)) = \mu_T(\phi)\frac{P^\Lambda_{T-t}(\mathbf{1})}{\mu_t P^\Lambda_{T-t}(\mathbf{1})},$$

since $\mu_T(\phi)$ is constant. Thus,

$$\begin{aligned}\|W_{t,T}(\bar{\phi}_T)\| &= \left\| \frac{P^\Lambda_{T-t}(\phi)}{\mu_t P^\Lambda_{T-t}(\mathbf{1})} - \mu_T(\phi)\frac{P^\Lambda_{T-t}(\mathbf{1})}{\mu_t P^\Lambda_{T-t}(\mathbf{1})} \right\| \\ &= \left\| \frac{\mu_t P^\Lambda_{T-t}(\mathbf{1}) \cdot P^\Lambda_{T-t}(\phi) - \mu_t P^\Lambda_{T-t}(\phi) \cdot P^\Lambda_{T-t}(\mathbf{1})}{\left(\mu_t P^\Lambda_{T-t}(\mathbf{1})\right)^2} \right\| \leq C\rho^{T-t},\end{aligned}$$

where the last inequality is a consequence of the fact that the function $t \mapsto \mathrm{e}^{-\lambda t}\mu_0 P^\Lambda_t(\mathbf{1})$ is bounded away from zero, and the uniform convergence of $\mathrm{e}^{-\lambda t}P^\Lambda_t(\phi)$ towards $h\mu_\infty(\phi)$, when $t \to \infty$, claimed in Corollary 1.1.7.

$\square$

Let us study the control of the $\mathbb{L}^p$ norm of the martingales obtained from Proposition 1.3.2 using as argument function $W_{t,T}(\phi)$, with $\phi \in \mathcal{B}_b(E)$ and $t \in [0,T]$. Note that

$$\partial_t\Big(\mu_t(P^\Lambda_{T-t}(\mathbf{1}))\Big) = \partial_t\left(\frac{\mu_0 P^\Lambda_T(1)}{\mu_0 P^\Lambda_t(\mathbf{1})}\right) = -\frac{\mu_0 P^\Lambda_T(\mathbf{1})}{\mu_0 P^\Lambda_t(\mathbf{1})^2}\mu_0 P^\Lambda_t(\Lambda) = -\mu_t(P^\Lambda_{T-t}(\mathbf{1}))\mu_t(\Lambda).$$

Thus,

$$\begin{aligned}\partial_t W_{t,T}(\phi) &= -(Q+\Lambda)W_{t,T}(\phi) - \frac{\partial_t(\mu_t(P^\Lambda_{T-t}(\mathbf{1})))P^\Lambda_{T-t}(\phi)}{\mu_t P^\Lambda_{T-t}(1)^2} \\ &= -(Q+\Lambda)W_{t,T}(\phi) + \mu_t(\Lambda)W_{t,T}(\phi).\end{aligned}$$

Hence, $W_{t,T}(\phi)$ is solution of the Cauchy problem

$$\begin{cases} \partial_s\psi_s &= -((Q+\Lambda) - \mu_t(\Lambda))\psi_s \\ \psi_T &= \phi. \end{cases}$$

Let us denote $\psi_{s,T} := W_{t,T}(\phi)$, for any $\phi \in \mathcal{B}_b(E)$. Note that,

$$\begin{aligned}\partial_t\left(\psi_{t,T}\right) &= -\Big(Q+\Lambda - \mu_t(\Lambda)\Big)\psi_{t,T}, \\ \partial_t\left(\psi^2_{t,T}\right) &= -2\psi_{t,T}\cdot\Big((Q+\Lambda-\mu_t(\Lambda))\psi_{t,T}\Big).\end{aligned}$$

We are in position to define the martingales $\big(\mathcal{M}_t(\psi_{\cdot,T})\big)_{t\in[0,T]}$ and $\big(\mathcal{M}_t(\psi^2_{\cdot,T})\big)_{t\in[0,T]}$, as stated in Proposition 1.3.2. We recall that

$$\mu(Q^\star_\mu\phi) = \mu(Q+\Lambda-\mu(\Lambda)\phi), \tag{1.26}$$

where $Q^\star_\mu$ is defined as in Remark 1.1.3. This identity is proved in (1.7). Hence, we obtain the following simplified expressions for the martingales $\big(\mathcal{M}_t(\psi_{\cdot,T})\big)_{t\in[0,T]}$ and $\big(\mathcal{M}_t(\psi^2_{\cdot,T})\big)_{t\in[0,T]}$:

$$\mathcal{M}_t\left(\psi_{\cdot,T}\right) = m(\eta_t)\left(\psi_{t,T}\right) - m(\eta_0)\left(\psi_{0,T}\right) - \int_0^t m(\eta_s)\left(\psi_{s,T}\right)\left[m(\eta_s)(\Lambda) - \mu_s(\Lambda)\right]\mathrm{d}s, \tag{1.27}$$

$$\mathcal{M}_t\left(\psi^2_{\cdot,T}\right) = m(\eta_t)\left(\psi^2_{t,T}\right) - m(\eta_0)\left(\psi^2_{0,T}\right) - 2\int_0^t m(\eta_s)\left(\psi^2_{s,T}\right)\Big[\mu_s(\Lambda) - m(\eta_s)(\Lambda)\Big]\mathrm{d}s - \Psi_t, \tag{1.28}$$

where

$$\Psi_t := \int_0^t m(\eta_s)\Big((Q+\Lambda-m(\eta_s)(\Lambda))(\psi_{s,T}^2) - 2\psi_{s,T}\cdot(Q+\Lambda-m(\eta_s)(\Lambda))(\psi_{s,T})\Big)\mathrm{d}s.$$

Furthermore, note that

$$\mu(\phi\cdot Q_\mu^\star\phi) = \mu(\phi(Q+\Lambda-\mu(\Lambda))\phi) + \mu(\phi)\mu(\mathcal{V}_\mu\phi) - \mu(\phi^2V_\mu^{\mathrm{b}}) - \mu(\phi^2)\mu(V_\mu^{\mathrm{d}}), \tag{1.29}$$

where $\mathcal{V}_\mu := V_\mu^{\mathrm{d}} + V_\mu^{\mathrm{b}} \in \mathcal{B}_b(E)$, for every $\mu\in\mathcal{M}_1(E)$.

Thus, the predictable quadratic variation of $(\mathcal{M}_t(\psi_{\cdot,T}))_{t\in[0,T]}$ satisfies

$$\begin{aligned} N\Big\langle\mathcal{M}\Big(\psi_{\cdot,T}\Big)\Big\rangle_t &:= \int_0^t m(\eta_s)(\Gamma_{Q_{m(\eta_s)}}(\psi_{s,T})\mathrm{d}s \\ &= \int_0^t m(\eta_s)\left(Q^\star_{m(\eta_s)}(\psi_{s,T}^2) - 2\psi_{s,T}\cdot Q^\star_{m(\eta_s)}\psi_{s,T}\right)\mathrm{d}s + \int_0^t S_{m(\eta_s)}\left(\psi_{s,T}\right)\mathrm{d}s, \end{aligned}$$

where S_μ is defined as in (1.14), for every $\mu\in\mathcal{M}_1(E)$. Now, using (1.26) and (1.29) we obtain

$$\begin{aligned} N\Big\langle\mathcal{M}\Big(\psi_{\cdot,T}\Big)\Big\rangle_t = \Psi_t &- 2\int_0^t m(\eta_s)(\psi_{s,T})m(\eta_s)(\mathcal{V}_{m(\eta_s)}\psi_{s,T})\mathrm{d}s + 2\int_0^t m(\eta_s)(\psi_{s,T}^2V^{\mathrm{b}}_{m(\eta_s)})\mathrm{d}s \\ &+ 2\int_0^t m(\eta_s)(\psi_{s,T}^2)m(\eta_s)\left(V^{\mathrm{d}}_{m(\eta_s)}\right)\mathrm{d}s + \int_0^t S_{m(\eta_s)}\left(\psi_{s,T}\right)\mathrm{d}s. \end{aligned}$$

Then, using (1.28) we can substitute the value of Ψ_t into this last expression and get

$$\begin{aligned} N\Big\langle\mathcal{M}(\psi_{\cdot,T})\Big\rangle_t =& -\mathcal{M}_t(\psi_{\cdot,T}^2) + m(\eta_t)\left(\psi_{t,T}^2\right) - m(\eta_0)\left(\psi_{0,T}^2\right) \\ &- 2\int_0^t m(\eta_s)\left(\psi_{s,T}^2\right)\Big[\mu_s(\Lambda) - m(\eta_s)(\Lambda)\Big]\mathrm{d}s \\ &- 2\int_0^t m(\eta_s)(\psi_{s,T})m(\eta_s)(\mathcal{V}_{m(\eta_s)}\psi_{s,T})\mathrm{d}s + 2\int_0^t m(\eta_s)(\psi_{s,T}^2V^{\mathrm{b}}_{m(\eta_s)})\mathrm{d}s \\ &+ 2\int_0^t m(\eta_s)(\psi_{s,T}^2)m(\eta_s)\left(V^{\mathrm{d}}_{m(\eta_s)}\right)\mathrm{d}s + \int_0^t S_{m(\eta_s)}\left(\psi_{s,T}\right)\mathrm{d}s. \end{aligned} \tag{1.30}$$

The expression in (1.30) for the predictable quadratic variation of the martingale $\mathcal{M}(\psi_{\cdot,T})$ is a key ingredient in the proof of Theorem 1.1.8. Using this expression and Theorem 1.1.4, it is possible to obtain an asymptotic expression for the quadratic error of the empirical distribution induced by the particle system.

Let ψ be a function on $E\times\mathbb{R}_+$ as in the statement of Proposition 1.3.2. We denote by $(\mathcal{M}_t^T(\psi_{\cdot,T}))_{t\in[0,T]}$ the local martingale defined as follows

$$\mathcal{M}_t^T\left(\psi_{\cdot,T}(\phi)\right) := \mathcal{M}_T\left(\psi_{\cdot,T}(\phi)\right) - \mathcal{M}_t\left(\psi_{\cdot,T}(\phi)\right),$$

for all $t\in[0,T]$. We denote by $(\langle\mathcal{M}(\psi_\cdot)\rangle_t^T)_{t\in[0,T]}$, the predictable quadratic variation of the local martingale $(\mathcal{M}_t^T(\psi_{\cdot,T}))_{t\in[0,T]}$.

Using (1.30) we can prove the next two results, analogously to Lemma 1.3.3 and Theorem 1.3.5, and establish a control on the predictable quadratic variation and the quadratic variation of the martingale $(\mathcal{M}_t^T(\psi_{\cdot,T}))_{t\in[0,T]}$, respectively.

Lemma 1.3.8 (Control of the predictable quadratic variation)**.** *For every test function $\phi\in\mathcal{B}_1(E)$ and every $t\in[0,T]$ we have*

$$N\Big\langle\mathcal{M}\left(W_{\cdot,T}(\phi)\right)\Big\rangle_t^T \le C(T-t+1) - \mathcal{M}_t^T\Big(W_{\cdot,T}(\phi)^2\Big), \quad \textit{for all} \quad t\in[0,T],$$

and for $\bar{\phi}_T = \phi - \mu_T(\phi)$ *we have*

$$N\left\langle \mathcal{M}\left(W_{\cdot,T}\left(\bar{\phi}_T\right)\right)\right\rangle_t^T \leq C - \mathcal{M}_t^T\left(W_{\cdot,T}\left(\bar{\phi}_T\right)^2\right), \quad \textit{for all} \quad t \in [0,T].$$

Theorem 1.3.9 (Control of the quadratic variation)**.** *For all* $p > 0$ *and every test function* $\phi \in \mathcal{B}_1(E)$ *there exists a positive* C *(possibly depending on* p*) such that*

$$\mathbb{E}\left[\left(\left[\mathcal{M}\left(W_{\cdot,T}(\phi)\right)\right]_t^T\right)^p\right] \leq \frac{C(T-t+1)^p}{N^p},$$

and for a centered test function $\bar{\phi}_T = \phi - \mu_T(\phi)$*:*

$$\mathbb{E}\left[\left(\left[\mathcal{M}\left(W_{\cdot,T}(\bar{\phi}_T)\right)\right]_t^T\right)^p\right] \leq \frac{C}{N^p}.$$

The proofs of Lemma 1.3.8 and Theorem 1.3.9 are analogous to those of Lemma 1.3.3 and Theorem 1.3.5, respectively. They are obtained using Lemma 1.3.7 instead of (1.20). In particular, the second inequalities in both results are consequences of part b) of Lemma 1.3.7. See also the proofs of Lemma 5.3 and Theorem 5.4 in [Rou06]. We skip the proofs of Lemma 1.3.8 and Theorem 1.3.9 for the sake of brevity.

Let us define the nonlinear propagator associated to $(\mu_t)_{t\geq 0}$ as follows

$$\Phi_{t,T}(\nu) := \frac{\nu P^{\Lambda}_{T-t}}{\nu P^{\Lambda}_{T-t}(\mathbf{1})} \in \mathcal{M}_1(E).$$

By the semigroup property, $\Phi_{t,T}$ satisfies the propagation equation $\mu_T = \Phi_{t,T}(\mu_t)$. Using Assumption (C2) we can ensure the existence of $\rho \in (0,1)$ such that

$$\sup_{\nu\in\mathcal{M}_1(E)} \|\Phi_{t,T}(\nu) - \mu_\infty\|_{\mathrm{TV}} \leq C\rho^{T-t}.$$

Let us define

$$I_p(N) = \sup_{T\geq 0}\sup_{\phi\in\mathcal{B}_1(E)} \mathbb{E}\left[\left(m(\eta_T)(\phi) - \mu_T(\phi)\right)^p\right].$$

Our goal is to prove that $I_p(N) \leq C/N^{p/2}$. The method we use is similar to the one used by Rousset [Rou06] and Angeli et al. [AGJ21]. Broadly speaking, it consists in an induction principle. First let us prove the following result providing the initial case of the induction.

Lemma 1.3.10 (Initial case)**.** *There exists* $\epsilon > 0$ *independent of* p*, such that*

$$I_p(N) \leq \frac{C}{N^{\epsilon p/2}}.$$

Proof. Fix $T > 0$ and consider

$$m(\eta_T)(\phi) - \mu_T(\phi) = \underbrace{m(\eta_T)(\phi) - \Phi_{t,T}(m(\eta_t))(\phi)}_{:=a(t)} + \underbrace{\Phi_{t,T}(m(\eta_t))(\phi) - \mu_T(\phi)}_{:=b(t)}. \tag{1.31}$$

The idea is to control $a(t)$ using the stochastic error between t and T, and $b(t)$ using the limiting stability. Moreover, $b(0)$ is controlled by the error made by the initial condition.

Let us first control the term $\mathbb{E}[|a(t)|^p]$. Consider the finite variation process

$$A_{t_1}^{t_2} := \exp\left\{\int_{t_1}^{t_2} m(\eta_s)(\Lambda) - \mu_s(\Lambda)\mathrm{d}s\right\}.$$

Then,

$$\partial_s \left(A_t^s m(\eta_s)(W_{s,t}(\phi))\right) = A_t^s \mathrm{d}\mathcal{M}_s(W_{\cdot,T}(\phi)). \tag{1.32}$$

Indeed, the martingale problem in Proposition 1.3.2, for the function $W_{t,T}(\phi)$ yields

$$\mathrm{d}(m(\eta_t)(W_{t,T}(\phi))) = \mathrm{d}\mathcal{M}_t(W_{\cdot,T}(\phi)) + (\mu_t(\Lambda) - m(\eta_t))m(\eta_t)(W_{t,T}(\phi))\mathrm{d}t.$$

Hence,

$$\begin{aligned}
\partial_s \left(A_t^s m(\eta_s)(W_{s,T}(\phi))\right) &= \partial_s \left(A_t^s\right) m(\eta_s)(W_{s,T}(\phi)) + A_t^s \mathrm{d}\left(m(\eta_s)(W_{s,T}(\phi))\right), \\
&= A_t^s\Big(m(\eta_s)(\Lambda) - \mu_s(\Lambda)\Big)m(\eta_s)(W_{s,T}(\phi)) + A_t^s \mathrm{d}\left(m(\eta_s)(W_{s,T}(\phi))\right) \\
&= A_t^s \mathrm{d}\mathcal{M}_s(W_{\cdot,T}(\phi)),
\end{aligned}$$

where the last expression is obtained using (1.27).

Now, integrating from t to T in (1.32) and dividing by A_t^T we get

$$m(\eta_T)(\phi) - (A_t^T)^{-1} m(\eta_t)(W_{t,T}(\phi)) = (A_t^T)^{-1} \int_t^T A_t^s \mathrm{d}\mathcal{M}_s(W_{\cdot,T}(\phi)).$$

Note that

$$\Phi_{t,T}(m(\eta_t))(\phi) = \frac{(A_t^T)^{-1} m(\eta_t)(W_{t,T}(\phi))}{(A_t^T)^{-1} m(\eta_t)(W_{t,T}(\mathbf{1}))},$$

for all $t \leq T$. Thus,

$$\begin{aligned}
a(t) &= m(\eta_t)(\phi) - (A_t^T)^{-1} m(\eta_t)(W_{t,T}(\phi)) - \Phi_{t,T}(m(\eta_t))(\phi)\left[1 - (A_t^T)^{-1} m(\eta_t)(W_{t,T}(\mathbf{1}))\right] \\
&= (A_t^T)^{-1} \int_t^T A_t^s \mathrm{d}\mathcal{M}_s(W_{\cdot,T}(\phi)) - \Phi_{t,T}(m(\eta_t))(\phi)(A_t^T)^{-1} \int_t^T A_t^s \mathrm{d}\mathcal{M}_s(W_{\cdot,T}(\mathbf{1})).
\end{aligned}$$

Thus, we obtain the upper bound

$$|a(t)| \leq (A_t^T)^{-1}\left(\left|\int_t^T A_t^s \mathrm{d}\mathcal{M}_s\Big(W_{\cdot,T}(\phi)\Big)\right| + \left|\int_t^T A_t^s \mathrm{d}\mathcal{M}_s\Big(W_{\cdot,T}(\mathbf{1})\Big)\right|\right).$$

There exists a $K > 0$ such that

$$\begin{aligned}
\mathbb{E}\left[|a(t)|^p\right] &\leq K\mathrm{e}^{2p\|\Lambda\|(T-t)} \sup_{\varphi\in\mathcal{B}_1(E)} \mathbb{E}\left[\left|\int_t^T A_t^s \mathrm{d}\mathcal{M}_s(W_{s,t}(\varphi))\right|^p\right] \\
&\leq K\mathrm{e}^{2p\|\Lambda\|(T-t)} \sup_{\varphi\in\mathcal{B}_1(E)} \mathbb{E}\left[\left|\int_t^T (A_t^s)^2 \mathrm{d}\left[\mathcal{M}(W_{\cdot,t}(\varphi))\right]_s\right|^{p/2}\right],
\end{aligned}$$

where the second inequality holds by the BDG inequality. Then, using Theorem 1.3.9 we get

$$\begin{aligned}
\mathbb{E}\left[|a(t)|^p\right] &\leq K\mathrm{e}^{4p\|\Lambda\|(T-t)} \sup_{\varphi\in\mathcal{B}_1(E)} \mathbb{E}\left[\left|\left[\mathcal{M}(W_{\cdot,t}(\varphi))\right]_t^T\right|^{p/2}\right] \\
&\leq K\mathrm{e}^{4p\|\Lambda\|(T-t)} \frac{(T-t+1)^{p/2}}{N^{p/2}} \\
&\leq K\frac{\kappa^{p(T-t)}}{N^{p/2}},
\end{aligned}$$

where $\kappa = \mathrm{e}^{4\|\Lambda\|+1/2} > 1$.

Let us now control $\mathbb{E}[|b(t)|^p]$. As a consequence of Assumption (C2) there exists a $\rho \in (0,1)$ and a $C \geq 0$ such that

$$\mathbb{E}[|b(t)|^p] = \mathbb{E}[|\Phi_{t,T}(m(\eta_t))(\phi) - \Phi_{t,T}(\mu_t)(\phi)|^p] \leq C\rho^{p(T-t)}.$$

Now, for controlling $b(0)$, note that

$$\begin{aligned}
b(0) &= \Phi_{0,T}(m(\eta_0))(\phi) - \mu_T(\phi) \\
&= m(\eta_0)(W_{0,T}(\phi)) - \mu_0(W_{0,T}(\phi)) + \Phi_{0,T}(m(\eta_0))(\phi) - m(\eta_0)(W_{0,T}(\phi)) \\
&= m(\eta_0)(W_{0,T}(\phi)) - \mu_0(W_{0,T}(\phi)) + \Phi_{0,T}(m(\eta_0))(\phi)[1 - m(\eta_0)(W_{0,T}(\mathbf{1}))].
\end{aligned}$$

Thus, using Assumption (I) and the fact that $\mu_0(W_{0,T}(\mathbf{1})) = 1$, we get

$$\mathbb{E}[|b(0)|^p] \leq \frac{C}{N^{p/2}}.$$

Let us now establish the global control optimising the choice of the argument t in (1.31). We have

$$\mathbb{E}\left[|a(0) + b(0)|^p\right] \leq C\frac{\kappa^{pT} + 1}{N^{p/2}}, \tag{1.33}$$

$$\mathbb{E}\left[|a(t) + b(t)|^p\right] \leq C\frac{\kappa^{p(T-t)} + 1}{N^{p/2}} + C\rho^{p(T-t)}, \tag{1.34}$$

for all $t \in [0,T]$.

The key idea now is to find a t_ϵ such that κ^{t_ϵ}/N and ρ^{t_ϵ} are both equal to $1/N^\epsilon$. Let us take $t_\epsilon = \frac{\ln N}{\ln\kappa - \ln\rho}$ and $\epsilon = \frac{-\ln\rho}{\ln\kappa - \ln\rho}$. Then, we have

$$\frac{\kappa^{t_\epsilon}}{N^{1/2}} = \exp\left\{-\epsilon \ln N\right\} = \frac{1}{N^\epsilon} \quad \text{and} \quad \rho^{t_\epsilon} = \exp\{-\epsilon\ln N\} = \frac{1}{N^\epsilon}.$$

We thus obtain the desired inequality:

$$\mathbb{E}\left[|m(\eta_T)(\phi) - \mu_T(\phi)|^p\right] \leq \frac{C}{N^{\epsilon p/2}}.$$

Indeed, this inequality is obtained either from (1.33) when $T \leq \frac{1}{2}\frac{\ln N}{\ln\kappa - \ln\rho}$, since the expression in the upper bound is increasing in T, or from (1.34) otherwise taking $T - t = \frac{1}{2}\frac{\ln N}{\ln\kappa - \ln\rho}$. □

We proceed now to prove the induction step (equation (1.37) below), which together with the initial case proved in Lemma 1.3.10, concludes the proof of Theorem 1.1.4.

Proof of Theorem 1.1.4. Note that (1.25) taking as argument function $W_{\cdot,T}(\bar{\phi}_T)$ reduces to

$$\begin{aligned}
\mathcal{M}_T(W_{\cdot,T}(\bar{\phi}_T)) &= m(\eta_T)(\bar{\phi}_T) - m(\eta_0)(W_{0,T}(\bar{\phi}_T)) \\
&\quad - \int_0^T (\mu_s(\Lambda) - m(\eta_s)(\Lambda))m(\eta_s)(W_{s,T}(\bar{\phi}_T))\mathrm{d}s.
\end{aligned} \tag{1.35}$$

Hence,

$$|m(\eta_T)(\phi) - \mu_T(\phi)|^p \leq C\left(|m(\eta_0)(W_{0,T}(\bar{\phi}_T))|^p + |\mathcal{M}_T(W_{\cdot,T}(\bar{\phi}_T))|^p + R_p\right),$$

where

$$R_p = \left|\int_0^T (\mu_s(\Lambda) - m(\eta_s)(\Lambda))m(\eta_s)(W_{s,T}(\bar{\phi}_T))\mathrm{d}s\right|^p.$$

The initial error can be controlled using Assumption (I). Indeed,

$$\begin{aligned}
\mathbb{E}\left[|m(\eta_0)(W_{0,T}(\bar{\phi}_T))|^p\right] &= \mathbb{E}\left[|m(\eta_0)(W_{0,T}(\phi)) - \mu_T(\phi) + \mu_T(\phi) - \mu_T(\phi)m(\eta_0)(W_{0,T}(\mathbf{1}))|^p\right] \\
&\leq C_1\mathbb{E}\left[|m(\eta_0)(W_{0,T}(\phi)) - \mu_0(W_{0,T}(\phi))|^p\right] + C_1\mathbb{E}\left[|\mu_0(W_{0,T}(\mathbf{1})) - m(\eta_0)(W_{0,T}(\mathbf{1}))|^p\right] \\
&\leq \frac{C}{N^{p/2}}.
\end{aligned}$$

Furthermore, using Theorem 1.3.9 and BDG inequality we get

$$\mathbb{E}\left[|\mathcal{M}_T(W_{\cdot,T})(\bar{\phi}_T)|^p\right] \leq \frac{C}{N^{p/2}}.$$

Note that $\mu_s(W_{s,T}(\bar{\phi}_T)) = 0$. Using Hölder inequality we obtain

$$\begin{aligned}
R_p &= \left|\int_0^T (\mu_s(\Lambda) - m(\eta_s)(\Lambda))m(\eta_s)\left(\frac{W_{s,T}(\bar{\phi}_T)}{\|W_{s,T}(\bar{\phi}_T)\|}\right)\|W_{s,T}(\bar{\phi}_T)\|\mathrm{d}s\right|^p \\
&\leq \int_0^T |\mu_s(\Lambda) - m(\eta_s)(\Lambda)|^p\left|m(\eta_s)\left(\frac{W_{s,T}(\bar{\phi}_T)}{\|W_{s,T}(\bar{\phi}_T)\|}\right)\right|^p\|W_{s,T}(\bar{\phi}_T)\|\mathrm{d}s\left(\int_0^T\|W_{s,T}(\bar{\phi}_T)\|\mathrm{d}s\right)^{p-1} \\
&\leq \kappa\int_0^T\left|\mu_s\left(\frac{\Lambda}{\|\Lambda\|}\right) - m(\eta_s)\left(\frac{\Lambda}{\|\Lambda\|}\right)\right|^p\left|m(\eta_s)\left(\frac{W_{s,T}(\bar{\phi}_T)}{\|W_{s,T}(\bar{\phi}_T)\|}\right) - \mu_s\left(\frac{W_{s,T}(\bar{\phi}_T)}{\|W_{s,T}(\bar{\phi}_T)\|}\right)\right|^p\|W_{s,T}(\bar{\phi}_T)\|\mathrm{d}s.
\end{aligned}$$

Taking expectation and using the Cauchy−Schwarz inequality yield

$$\mathbb{E}\left[\left|\int_0^T (\mu_s(\Lambda) - m(\eta_s)(\Lambda))m(\eta_s)(W_{s,T}(\bar{\phi}_T))\mathrm{d}s\right|^p\right] \leq \kappa\int_0^T I_{2p}(N)\|W_{s,T}(\bar{\phi}_T)\|\mathrm{d}s \leq KI_{2p}(N). \tag{1.36}$$

Thus, for every $p \geq 1$ we get the inequality

$$I_p(N) \leq C\left(\frac{1}{N^{p/2}} + I_{2p}(N)\right), \tag{1.37}$$

which using Lemma 1.3.10 reduces to

$$I_p(N) \leq \frac{C}{N^{\min\{2\epsilon,1\}p/2}}.$$

By induction we obtain the bound

$$I_p(N) \leq \frac{C}{N^{p/2}}.$$

□

Proof of Theorem 1.1.6. Taking expectation in (1.35) we get

$$\mathbb{E}\left[m(\eta_T)(\phi)\right] - \mu_T(\phi) = \int_0^T \mathbb{E}\left[(\mu_s(\Lambda) - m(\eta_s)(\Lambda))m(\eta_s)\Big(W_{s,T}(\bar{\phi}_T)\Big)\right]\mathrm{d}s.$$

Using Cauchy−Schwarz inequality we obtain

$$|\mathbb{E}\left[m(\eta_T)(\phi)\right] - \mu_T(\phi)| \leq C_1\int_0^T I_2(N)\|W_{s,T}(\bar{\phi}_T)\|\mathrm{d}s \leq \frac{C}{N}.$$

Now, assume that initially the N particles are sampled according to an exchangeable distribution. Note that

$$\mathbb{E}\left[\frac{\eta_t(x)}{N}\right] = \frac{1}{N}\sum_{i=1}^{N}\mathbb{P}[\xi_t^i = x] = \mathbb{P}[\xi_t^j = x], \quad \forall j \in \{1,\ldots,N\},$$

where ξ_t^i denotes the position of the i-th particle of the process $(\eta_t)_{t\geq 0}$ at time $t \geq 0$. Note that the second equality holds because of the exchangeability of the particles. Thus, as a consequence of the first part of the theorem and the previous equality we get

$$\|\operatorname{Law}(\xi_t^{(i)}) - \mu_t\|_{\mathrm{TV}} \leq \frac{C}{N}.$$

□

1.3.4 Proof of Theorem 1.1.8

Under Assumption (C1), it is possible to find a simplified expression for the predictable quadratic variation of the martingale $(\mathcal{M}_t(\psi_{\cdot,T}))_{t\in[0,T]}$, where $\psi_{t,T} = W_{t,T}(\bar{\phi}_T)$, for $t \in [0,T]$. Indeed, from (1.30) we have

$$\begin{aligned} N\Big\langle \mathcal{M}(\psi_{\cdot,T})\Big\rangle_t =& -\mathcal{M}_t(\psi^2_{\cdot,T}) + m(\eta_t)\left(\psi^2_{t,T}\right) - m(\eta_0)\left(\psi^2_{0,T}\right) + 2\int_0^t m(\eta_s)(\psi^2_{s,T})m(\eta_s)\left(V^{\mathrm{d}}_{m(\eta_s)}\right)\mathrm{d}s \\ &+ 2\int_0^t m(\eta_s)\left(\psi^2_{s,T}V^{\mathrm{b}}_{m(\eta_s)}\right)\mathrm{d}s + \int_0^t S_{m(\eta_s)}\left(\psi_{s,T}\right)\mathrm{d}s + R_t, \end{aligned}$$

where

$$R_t := -2\int_0^t m(\eta_s)\left(\psi^2_{s,T}\right)\Big[\mu_s(\Lambda) - m(\eta_s)(\Lambda)\Big]\mathrm{d}s - 2\int_0^t m(\eta_s)(\psi_{s,T})m(\eta_s)(\mathcal{V}_{m(\eta_s)}\psi_{s,T})\mathrm{d}s.$$

The key component in the proof of Theorem 1.1.8 is a central limit theorem for the martingale $(\mathcal{M}_t(\psi_{\cdot,T}))_{t\in[0,T]}$. Let us first introduce an auxiliary result.

Consider the process $(\widetilde{\mathcal{M}}_t(W_{\cdot,T}(\bar{\phi}_T)))_{t\in[0,T]}$ defined as

$$\widetilde{\mathcal{M}}_t(W_{\cdot,T}(\bar{\phi}_T)) := \sqrt{N}m(\eta_0)(W_{0,T}(\bar{\phi}_T)) + \sqrt{N}\mathcal{M}_t(W_{\cdot,T}(\bar{\phi}_T)),$$

for $t \in [0,T]$. Then, $(\widetilde{\mathcal{M}}_t(W_{\cdot,T}(\bar{\phi}_T)))_{t\in[0,T]}$ is a martingale, with initial value

$$\widetilde{\mathcal{M}}_0(W_{\cdot,T}(\bar{\phi}_T)) = \sqrt{N}m(\eta_0)(W_{0,T}(\bar{\phi}_N)).$$

Proposition 1.3.11 (Central limit theorem)**.** *The martingale $(\widetilde{\mathcal{M}}_t(W_{\cdot,T}(\bar{\phi}_T)))_{t\in[0,T]}$ converges in law when $N \to \infty$ towards a Gaussian martingale whose variance at time $t \in [0,T]$ is $\sigma_t^2(\phi)$, defined as*

$$\sigma_t^2(\phi) := \mu_t(\psi^2_{t,T}) + 2\int_0^t \mu_s(\psi^2_{s,T})\mu_s\left(V^{\mathrm{d}}_{\mu_s}\right)\mathrm{d}s + 2\int_0^t \mu_s(\psi^2_{s,T}V^{\mathrm{b}}_{\mu_s})\mathrm{d}s + \int_0^t S_{\mu_s}\left(\psi_{s,T}\right)\mathrm{d}s,$$

and $\psi_{t,T} = W_{t,T}(\bar{\phi}_T)$.

Proof. Using Theorem 3.11 in [JS87, §8], and arguing as in the proofs of Proposition 3.31 in [DM00b] and Proposition 3.7 in [DM03], we only need to check that the result holds for the initial value $\widetilde{\mathcal{M}}_0(W_{\cdot,T}(\bar{\phi}_T)) = \sqrt{N}m(\eta_0)(W_{0,T}(\bar{\phi}_T))$ and that $N\langle\mathcal{M}(\psi_{\cdot,T})\rangle$ converges in probability to a continuous function, when N goes to infinity. The first point is in fact Assumption (I′). Furthermore, Theorem 1.1.4 implies, by a Borel–Cantelli argument, the following convergence:

$$m(\eta_s) \xrightarrow{\text{a.s.}} \mu_s,$$

when $N \to \infty$, for all $s \geq 0$, as we commented in Remark 1.1.5. Now, using Theorem 1.1.4 and reasoning as in (1.36), we easily prove that R_t converges to 0 in probability and that $N\langle \mathcal{M}(\psi_{\cdot,T})\rangle$ converges to the continuous function $\sigma_\cdot^2(\phi) - \sigma_0^2(\phi)$ in probability, when $N \to \infty$, which concludes the proof. □

Proof of Theorem 1.1.8. As a consequence of Proposition 1.3.11 and (1.27) we have that

$$\widetilde{\mathcal{M}}_T(W_{\cdot,T}(\bar{\phi}_T)) = \sqrt{N} m(\eta_T)(\phi) - \mu_T(\phi) - \sqrt{N}\int_0^t m(\eta_s)\,(\psi_{s,T})\,[m(\eta_s)(\Lambda) - \mu_s(\Lambda)]\,\mathrm{d}s$$

converges to a Gaussian random variable of variance $\sigma_T^2(\phi)$, when $N \to \infty$. Thus, the first part of Theorem 1.1.8 comes from the fact that

$$\sqrt{N}\int_0^t m(\eta_s)\,(\psi_{s,T})\,[m(\eta_s)(\Lambda) - \mu_s(\Lambda)]\,\mathrm{d}s$$

converges to 0 almost surely when $N \to \infty$, as a consequence of (1.36). Thus, $m(\eta_T)(\phi) - \mu_T(\phi)$ converges in law to a centered Gaussian law with variance

$$\begin{aligned}\sigma_T^2(\phi) = \mu_T\Big((\phi - \mu_T(\phi))^2\Big) + \int_0^T S_{\mu_s}(W_{s,T}(\bar{\phi}_T))\mathrm{d}s \\ + 2\int_0^T \mu_s\left(W_{s,T}(\bar{\phi}_T)^2 V_{\mu_s}^{\mathrm{b}}\right) + \mu_s(W_{s,T}(\bar{\phi}_T)^2)\mu_s(V_{\mu_s}^{\mathrm{d}})\mathrm{d}s.\end{aligned}$$

Consider now the change of variables $u = T - s$ in the last integral of the previous expression, and then take limit when $T \to \infty$. The final result comes due the the following convergences:

$$\begin{aligned}\mu_{T-s} &\xrightarrow[T\to\infty]{} \mu_\infty, \\ \bar{\phi}_T = \phi - \mu_T(\phi) &\xrightarrow[T\to\infty]{} \phi - \mu_\infty(\phi), \\ W_{T-s,T}(\bar{\phi}_T) &\xrightarrow[T\to\infty]{} \frac{P_s^\Lambda(\bar{\phi}_\infty)}{\mu_\infty P_s^\Lambda(\mathbf{1})} = \mathrm{e}^{-\lambda s} P_s^\Lambda(\bar{\phi}_\infty),\end{aligned}$$

where the last inequality is a consequence of (1.16) and of the equality $\mu_\infty(\Lambda) = \lambda$.

□

1.A Proof of Lemma 1.1.1

Let us first prove the following result, which has an independent interest.

Lemma 1.A.1 ($\mathbb{L}^p$ norm bound for sum of i.i.d. centered r.v.)**.** *Let us consider $Y_1, Y_2, \ldots$ a sequence of independent identically distributed random variables with zero-mean and finite second moment, such that $\mathbb{E}[|Y_1|^p] < \infty$, for a given $p \geq 1$. Then, there exists a universal constant C_p such that*

$$\left(\mathbb{E}\left[\left|\frac{1}{N}\sum_{i=1}^N Y_i\right|^p\right]\right)^{1/p} \leq \frac{C_p}{\sqrt{N}}.$$

Proof. First note that for $p \leq 2$ we get the following result as a consequence of Jensen inequality for concave functions:

$$\mathbb{E}\left[\left|\frac{1}{N}\sum_{i=1}^{N} Y_i\right|^p\right] = \mathbb{E}\left[\left(\left(\frac{1}{N}\sum_{i=1}^{N} Y_i\right)^2\right)^{p/2}\right] \leq \left(\mathbb{E}\left[\left(\frac{1}{N}\sum_{i=1}^{N} Y_i\right)^2\right]\right)^{p/2} = \left(\frac{\mathbb{E}[Y^2]}{N}\right)^{p/2}.$$

For $p > 2$, the proof follows from Marcinkiewicz − Zygmund inequality, which is a consequence of the BDG inequality for discrete-time martingales. Indeed, the Marcinkiewicz–Zygmund inequality (cf. [RL01]) ensures us that

$$\mathbb{E}\left[\left|\sum_{i=1}^{N} Y_i\right|^p\right] \leq \frac{K_p}{N^{p/2}}\mathbb{E}\left[|Y_1|^p\right]. \tag{1.38}$$

Thus,

$$\mathbb{E}\left[\left|\frac{1}{N}\sum_{i=0}^{N} Y_i\right|^p\right] \leq \frac{C_p}{N^{p/2}}, \text{ where } C_p = \begin{cases} (\mathbb{E}[Y_1^2])^{p/2} & \text{if} \quad p \leq 2 \\ K_p\mathbb{E}[|Y_1|^p] & \text{if} \quad p > 2. \end{cases}$$

□

Remark 1.A.1 (Qualitative results for the Marcinkiewicz–Zygmund constant K_p). See the work of Ren and Liang [RL01] for a qualitative study of the constant K_p in inequality (1.38). They show that $(K_p)^{1/p}$ grows like $\sqrt{p}$, when $p \to \infty$, and give the estimate $K_p \leq (3\sqrt{2})^p p^{p/2}$.

Proof of Lemma 1.1.1. Note that $m(\eta_0)(\phi) = \frac{1}{N}\sum_{i=1}^{N}\phi(\xi_0^{(i)})$, where $\xi_0^{(i)}$, for $i = 1, \ldots, N$ are independent random variables. Moreover, $\phi(\xi_0^{(i)})$ has mean $\mu_0(\phi)$, for all $i = 1, \ldots, N$. Thus,

$$m(\eta_0)(\phi) - \mu_0(\phi) = \sum_{i=1}^{N}\frac{\phi(\xi_0^{(i)}) - \mu_0(\phi)}{N},$$

can be written as a sum of N zero-mean random variables. The result comes from Lemma 1.A.1. □

Chapter 2

Spectrum and ergodicity of the neutral Moran model

This chapter is based on the preprint [Cor21b] submitted on May, 2021.

Abstract: The purpose of this chapter is to provide a complete description of the eigenvalues of the generator of a neutral multi-type Moran model, and the applications to the study of the speed of convergence to stationarity. The Moran model we consider is a non-reversible in general, continuous-time Markov chain with unknown stationary distribution. Specifically, we consider N individuals such that each one of them is of one type among K possible allelic types. The individuals interact in two ways: by an independent irreducible mutation process and by a reproduction process, where a pair of individuals is randomly chosen, one of them dies and the other reproduces. Our main result provides explicit expressions for the eigenvalues of the infinitesimal generator matrix of the Moran process, in terms of the eigenvalues of the jump rate matrix. As a consequence of this result, we study the convergence in total variation of the process to stationarity. Our results include a lower bound for the mixing time of the Moran process when the mutation process allows a real eigenvalue. Furthermore, we study in detail the spectral decomposition of the neutral multi-allelic Moran model with parent independent mutation scheme, which turns to be the unique mutation scheme that makes the neutral Moran process reversible. Under the parent independent mutation, we also prove the existence of a cutoff phenomenon in the chi-square and the total variation distances when initially all the individuals are of the same type and the number of individuals tends to infinity. Additionally, in the absence of reproduction, we prove that the the total variation distance to stationarity of the parent independent mutation process, when initially all the individuals are of the same type, has a Gaussian profile.

2.1 Introduction and main results

This chapter is devoted to the study of a continuous-time Markov model of N particles on K sites with interaction, which is known as the neutral multi-allelic Moran model in population genetics literature [EG09]: the K sites correspond to K allelic types in a population of N individuals. The state space of the process is the K-dimensional N-discrete simplex:

$$\mathcal{E}_{K,N} := \left\{ \eta \in [N]_0^K : \ |\eta| = N \right\}, \tag{2.1}$$

where $[N]_0 := \{0, 1, \dots, N\}$ and $|\cdot|$ stands for the sum of elements in a vector. The set $\mathcal{E}_{K,N}$ is a finite set with cardinality $\mathrm{Card}(\mathcal{E}_{K,N}) = \binom{K-1+N}{N}$. The process is in state $\eta \in \mathcal{E}_{K,N}$ if there are

$\eta(k) \in [N]_0$ individuals with allelic type $k \in [K] := \{1, 2, \ldots, K\}$. Consider $Q = (\mu_{i,j})_{i,j=1}^K$ the infinitesimal rate matrix of an irreducible Markov chain on $[K]$, which is called the *mutation matrix* of the Moran process. The infinitesimal generator of the neutral multi-allelic Moran process, denoted $\mathcal{Q}_{N,p}$, acts on a real function f on $\mathcal{E}_{K,N}$ as follows:

$$(\mathcal{Q}_{N,p}f)(\eta) := \sum_{i,j\in[K]} \eta(i) \left(\mu_{i,j} + \frac{p}{N}\,\eta(j)\right) [f(\eta - \mathbf{e}_i + \mathbf{e}_j) - f(\eta)]\,, \tag{2.2}$$

for all $\eta \in \mathcal{E}_{K,N}$, where $\mathbf{e}_k$ is the k-th canonical vector of $\mathbb{R}^K$ (cf. [EG09]). In words, $\mathcal{Q}_{N,p}$ drives a process of N individuals, where each individual has one of K possible types of alleles and where the type of the individual changes following two processes: a mutation process where individuals mutate independently of each other and a Moran type reproduction process, where the individuals interact. The N individuals mutate independently from type $i \in [K]$ to type $j \in [K] \setminus \{i\}$ with rate $\mu_{i,j}$. In addition, with uniform rate $p \geq 0$, one of the N individuals is uniformly chosen to be removed from the population and another one, also randomly chosen, is duplicated. Note that the transitions of an individual due to a reproduction is not independent of the position of the other individuals.

As in the original model, introduced by Moran [Mor58], the same individual removed from the population can be duplicated, in this case the state of the system does not change. In the instance where the removed individual cannot be duplicated, the factor $\frac{p}{N}$ in (2.2) must be replaced by $\frac{p}{N-1}$.

Note that $\mathcal{Q}_{N,p}$ can be decomposed as $\mathcal{Q}_{N,p} = \mathcal{Q}_N + \frac{p}{N}\,\mathcal{A}_N$, where $\mathcal{Q}_N$ and $\mathcal{A}_N$ are also infinitesimal generators of Markov chains acting on every $f \in \mathbb{R}^{\mathcal{E}_{K,N}}$ as follows

$$(\mathcal{Q}_N f)(\eta) := \sum_{i,j\in[K]} \eta(i)\mu_{i,j}\,[f(\eta - \mathbf{e}_i + \mathbf{e}_j) - f(\eta)]\,, \tag{2.3}$$

$$(\mathcal{A}_N f)(\eta) := \sum_{i,j\in[K]} \eta(i)\eta(j)\,[f(\eta - \mathbf{e}_i + \mathbf{e}_j) - f(\eta)]\,, \tag{2.4}$$

for every $\eta \in \mathcal{E}_{K,N}$. The processes driven by $\mathcal{Q}_N$ and $\mathcal{A}_N$ are called *mutation process* and *reproduction process*, respectively. In words, $\mathcal{Q}_N$ models the dynamic of N indistinguishable particles, where each one moves among K sites according to the process generated by the mutation rate matrix Q. This process is usually called *compound chain* (cf. [ZL09]). On the other hand, $\mathcal{A}_N$ models the dynamic where at uniform rate two individuals are randomly chosen and one of them changes its type for the type of the other one. This chapter is devoted to the study of the spectrum of $\mathcal{Q}_N$, $\mathcal{A}_N$ and $\mathcal{Q}_{N,p}$, and of the convergence to stationarity of the generated Markov processes. Before stating our main results in this direction, let us establish some notation.

We recall that if $V_n \in \mathbb{R}^K$, $1 \leq n \leq N$ are N vectors in $\mathbb{R}^K$, their tensor product is the vector $V_1 \otimes V_2 \otimes \cdots \otimes V_N$ defined by $(V_1 \otimes V_2 \otimes \cdots \otimes V_N)(k_1, k_2, \ldots, k_N) := V_1(k_1)V_2(k_2)\ldots V_N(k_N)$, for all $1 \leq k_n \leq K$ and $1 \leq n \leq N$. The tensor $V_1 \otimes V_2 \otimes \cdots \otimes V_N$ can be considered as a function on $[K]^N$. Actually, throughout this chapter we completely identify a real function f on $[K]^N$ and the tensor vector V_f such that $V_f(k_1, k_2, \ldots, k_N) = f(k_1, k_2, \ldots, k_N)$, for all $(k_1, k_2, \ldots, k_N) \in [K]^N$.

Let us denote by σ a permutation on $[N]$, i.e. an element of the symmetric group $\mathcal{S}_N$. Then, the permutation of $f \in \mathbb{R}^{[K]^N}$ by σ, denoted by σf, is defined by

$$\sigma f : (k_1, k_2, \ldots, k_N) \mapsto f(k_{\sigma(1)}, k_{\sigma(2)}, \ldots k_{\sigma(N)}),$$

for all $(k_1, k_2, \ldots, k_N) \in [K]^N$. In particular, for $V_1, V_2, \ldots, V_N \in \mathbb{R}^N$ we have

$$\sigma(V_1 \otimes V_2 \otimes \cdots \otimes V_N) = V_{\sigma^{-1}(1)} \otimes V_{\sigma^{-1}(2)} \otimes \cdots \otimes V_{\sigma^{-1}(N)}.$$

A real function f on $[K]^N$ is symmetric if $f = \sigma f$, for all σ in $\mathcal{S}_N$. Moreover, every function f on $[K]^N$ can be symmetrised by the projector Sym, defined as follows:

$$\mathrm{Sym}: \; f \mapsto \overline{f} = \frac{1}{N!} \sum_{\sigma \in \mathcal{S}_N} \sigma f. \tag{2.5}$$

Symmetric functions on $[K]^N$ are highly important in the sequel because of their relation to the functions on $\mathcal{E}_{K,N}$. Consider the application $\psi_{K,N} : \mathcal{E}_{K,N} \to [K]^N$ defined by

$$\psi_{K,N} : \eta \mapsto (\underbrace{1,1,\ldots,1}_{\eta(1)}, \underbrace{2,2,\ldots,2}_{\eta(2)}, \ldots, \underbrace{K,K,\ldots,K}_{\eta(K)}), \tag{2.6}$$

when the number of k in $\overbrace{k,k,\ldots,k}^{\eta(k)}$ is 0 if $\eta(k) = 0$. Note that for every symmetric function f on $[K]^N$, the function $\tilde{f} := f \circ \psi_{K,N}$ on $\mathcal{E}_{K,N}$ is well defined. Let U_0 be the all-one vector in $\mathbb{R}^K$ and $U_1, U_2, \ldots, U_{K-1} \in \mathbb{R}^K$ such that $\mathcal{U} := \{U_0, U_1, \ldots, U_{K-1}\}$ is a basis of $\mathbb{R}^K$. Note that this is the type of basis given by the eigenvectors of the diagonalisable rate matrix of dimension K of a Markov chain on $[K]$. For every $\eta \in \mathcal{E}_{K-1,L}$, for $1 \leq L \leq N$, let us also denote by $U_\eta \in \mathbb{R}^{[K]^N}$, $V_\eta \in \mathrm{Sym}\left(\mathbb{R}^{[K]^N}\right)$ and $\tilde{V}_\eta \in \mathbb{R}^{\mathcal{E}_{K,N}}$ the vectors defined by

$$U_\eta := U_{k_1} \otimes U_{k_2} \otimes \cdots \otimes U_{k_L} \otimes \underbrace{U_0 \otimes \cdots \otimes U_0}_{N-L \text{ times}}, \tag{2.7}$$

$$V_\eta := \mathrm{Sym}(U_\eta), \tag{2.8}$$

$$\tilde{V}_\eta := V_\eta \circ \psi_{K,N}, \tag{2.9}$$

where $(k_1, k_2, \ldots, k_L) = \psi_{K-1,L}(\eta)$, $\eta \in \mathcal{E}_{K-1,L}$ and $L \in [N]$. In Section 2.2 we analyse the link between the spaces $\mathrm{Sym}\left(\mathbb{R}^{[K]^N}\right)$ and $\mathbb{R}^{\mathcal{E}_{K,N}}$, and we clarify the nature of the definitions previously introduced. Next theorem clarifies the connection between the eigenstructures of Q and $\mathcal{Q}_N$

Theorem 2.1.1 (Eigenstructure of $\mathcal{Q}_N$)**.** *Assume $K \geq 2$, $N \geq 1$. Let $\mathcal{U} = \{U_0, U_1, \ldots, U_{r-1}\}$ be a set of r independent right eigenvectors of Q such that U_0 is the all-one vector. Let $\lambda_0 = 0, \lambda_1, \ldots, \lambda_{K-1}$ be the K complex roots of the characteristic polynomial of Q, counting algebraic multiplicities, such that $QU_k = \lambda_k U_k$, for $k \in \{0, 1, \ldots, r-1\}$. Consider λ_η defined as follows*

$$\lambda_\eta := \sum_{k=1}^{K-1} \eta(k) \lambda_k. \tag{2.10}$$

Then,

(a) The eigenvalues of $\mathcal{Q}_N$ are given by λ_η, for all $\eta \in \bigcup_{L=1}^{N} \mathcal{E}_{K-1,L}$.

(b) Every function $\tilde{V}_\eta$, as defined in (2.9), for $\eta \in \bigcup_{L=1}^{N} \mathcal{E}_{K-1,L}$ satisfying $\eta(r) = \cdots = \eta(K-1) = 0$ is a right eigenfunction of $\mathcal{Q}_N$ such that $\mathcal{Q}_N \tilde{V}_\eta = \lambda_\eta \tilde{V}_\eta$.

(c) In particular, if Q is diagonalisable, then $\mathcal{Q}_N$ is diagonalisable.

The proof of Theorem 2.1.1 can be found in Section 2.3.1. Theorem 2.1.1 can be seen as a continuous-time generalisation of the results provided by Zhou and Lange [ZL09] for the discrete-time analogous of the mutation process driven by $\mathcal{Q}_N$. We emphasize that our hypotheses do not require the mutation rate matrix Q to be diagonalisable.

Next result deals with the spectrum of $\mathcal{A}_N$.

Theorem 2.1.2 (Spectrum of $\mathcal{A}_N$)**.** *Assume $K \geq 2$ and $N \geq 2$. The eigenvalues of $\mathcal{A}_N$ are*

$$\begin{array}{rl} 0 & \textit{with multiplicity } K \textit{ and} \\ -L(L-1) & \textit{with multiplicity } \binom{K+L-2}{L}, \textit{ for } 2 \leq L \leq N. \end{array}$$

Additionally, the infinitesimal rate matrix $\mathcal{A}_N$ is diagonalisable.

The proof of Theorem 2.1.2 is deferred to Section 2.3.2. Theorem 2.1.2 can be seen as a generalisation, for $K \geq 3$, of the results in [Zho08, §4.2.2] for the discrete analogous of the reproduction process driven by $\mathcal{A}_N$, for $K = 2$.

Unlike in the independent mutation process, the dynamics of the neutral multi-allelic Moran process driven by $\mathcal{Q}_{N,p}$, for $p > 0$, is that of an interacting particle system, which makes harder the study of its spectrum. Our main result is precisely a complete description of the eigenvalues of the generator $\mathcal{Q}_{N,p}$, which is expressed in the following theorem.

Theorem 2.1.3 (Spectrum of $\mathcal{Q}_{N,p}$)**.** *Assume $K \geq 2$, $N \geq 1$ and $p \in [0, \infty)$. Let us denote by λ_k, $k \in [K-1]$, the nonzero $K-1$ roots, counting algebraic multiplicities, of the characteristic polynomial of Q. For any $\eta \in \bigcup_{L=1}^{N} \mathcal{E}_{K-1,L}$, let us define*

$$\lambda_{\eta,p} := \sum_{k=1}^{K-1} \eta(k)\lambda_k - \frac{p}{N}|\eta|(|\eta| - 1).$$

Then, the eigenvalues of $\mathcal{Q}_{N,p}$, counting algebraic multiplicities, are 0 and $\lambda_{\eta,p}$, for $\eta \in \bigcup_{L=1}^{N} \mathcal{E}_{K-1,L}$.

The proof of Theorem 2.1.3 is given in Section 2.3.3.

Remark 2.1.1 (Monotony in N of the spectrum of $\mathcal{Q}_{N,p}$)*.* Theorem 2.1.3 implies that the spectrum of $\mathcal{Q}_{N,p}$, for fixed values of K and $p \geq 0$, is an increasing function of N in the sense of the inclusion of sets.

Remark 2.1.2 (Relation to the spectrum of the Wright – Fisher diffusion)*.* The eigenstructure of the Wright – Fisher diffusion is a special case of the eigenstructure in a Lambda – Fleming – Viot process studied in [Gri14]. Theorem 5 of [Gri14], taking $W = 0$, gives the spectrum of the neutral Wright-Fisher diffusion, which coincides with the spectrum provided by Theorem 2.1.3. This is not surprising since the Wright – Fisher diffusion is the limit process for the Moran model (cf. [Eth11, Lemma 2.39]).

Applications to the ergodicity of neutral multi-allelic Moran process

The relation between the spectral properties of $\mathcal{Q}_{N,p}$ and Q can be used to estimate the speed of convergence to stationarity of the Moran process.

Let us first recall the *total variation distance*. For two probability measures μ_1 and μ_2 defined on the same discrete space Ω, the total variation distance is defined as follows:

$$\mathrm{d}^{\mathrm{TV}}(\mu_1, \mu_2) := \sup_{A \subset \Omega} |\mu_1(A) - \mu_2(A)| = \sup_{f:\Omega \to [-1,1]} \left| \int f \mathrm{d}\mu_1 - \int f \mathrm{d}\mu_2 \right| = \frac{1}{2} \|\mu_1 - \mu_2\|_1,$$

where $\| \cdot \|_1$ denotes the 1-norm in $\mathbb{R}^\Omega$.

The total variation distance to stationarity at time t of an ergodic process driven by a generator L on Ω, with initial distribution μ, is given by $\mathrm{d}^{\mathrm{TV}}(\mu \, \mathrm{e}^{tL}, \pi)$, where μ is the initial distribution on Ω and π is the stationary distribution of the process driven by L. We are interested in the relationship between the spectrum of an infinitesimal rate matrix and the convergence to stationarity of the Markov process it drives. Let us define the *maximum total variation distance* to stationarity of the process driven by L, denoted $\mathrm{D}_L^{\mathrm{TV}}$, as follows:

$$\mathrm{D}_L^{\mathrm{TV}}(t) := \max_{\mu} \mathrm{d}^{\mathrm{TV}}(\mu\, \mathrm{e}^{tL}, \pi),$$

where the maximum runs over all possible initial distributions on Ω. Using the convexity of d^{TV}, we can prove that $\mathrm{D}_L^{\mathrm{TV}}(t) = \frac{1}{2}\|\mathrm{e}^{tL} - \Pi\|_\infty$, where Π stands for the matrix with every row equal to π, and $\|\cdot\|_\infty$ denotes the infinity norm of matrices (cf. [LP17, Ch. 4]).

As a consequence of Theorem 2.1.3, the *second largest eigenvalue in modulus* (SLEM) of $\mathcal{Q}_{N,p}$ is equal to that of Q. The SLEM of the generator of the process is useful to study the asymptotic convergence of the process in total variation. Hence, in Section 2.4 we study the ergodicity of the process driven by $\mathcal{Q}_{N,p}$ in total variation using the spectral properties of Q. We also analyse several examples of neutral multi-allelic Moran processes with diagonalisable and non-diagonalisable mutation rate matrices.

For a real positive function f we denote by $\mathcal{O}(f)$ another real positive function such that $C_1 f(t) \leq \mathcal{O}(f)(t) \leq C_2 f(t)$, for two constants $0 < C_1 \leq C_2 < \infty$ and for all $t \geq T$, for $T > 0$ large enough.

Corollary 2.1.4 (Asymptotic exponential ergodicity in total variation)**.** *Let us denote by ρ the* SLEM *of Q and by $s \in \mathbb{N}$ the largest multiplicity in the minimal polynomial of Q of all the eigenvalues with modulus ρ. Then,*

$$\mathrm{D}_{\mathcal{Q}_{N,p}}^{\mathrm{TV}}(t) = \mathcal{O}(\mathrm{D}_Q^{\mathrm{TV}}(t)) = \mathcal{O}(t^{s-1}\mathrm{e}^{-\rho t}).$$

Corollary 2.1.4 is proved in Section 2.4.

The asymptotic expression in Corollary 2.1.4 hides the relation between the mixing time of the Markov chain and the number of individuals in the population. However, if we know the right eigenvector associated to a real eigenvalue $-\lambda < 0$ of Q, we can further prove the following lower bound for the convergence in total variation to stationarity at time $\frac{\ln N - c}{2\lambda}$, for every $c \geq 0$.

Theorem 2.1.5 (Lower bound for convergence in total variation)**.** *Assume $K \geq 2$, $N \geq 2$ and $p \in [0, \infty)$ and let $-\lambda < 0$ be an eigenvalue of Q with associated right-eigenvector $V = [v_1, \ldots, v_K]$. Let $\nu_{N,p}$ be the stationary distribution of the process driven by $\mathcal{Q}_{N,p}$ and let us denote*

$$t_{N,c} := \frac{\ln N - c}{2\lambda} \quad \text{and} \quad \kappa := 8(2\lambda + \|Q\|_\infty).$$

Then,

$$\mathrm{d}^{\mathrm{TV}}(\delta_{N\mathbf{e}_k}\mathrm{e}^{t_{N,c}\mathcal{Q}_{N,p}}, \nu_{N,p}) \geq 1 - \kappa\frac{\|V\|_\infty}{|v_k|}\,\mathrm{e}^{-c},$$

for all $c \geq 0$ and for any $k \in [K]$ such that $v_k \neq 0$. In particular,

$$\mathrm{D}_{\mathcal{Q}_{N,p}}^{\mathrm{TV}}\left(\frac{\ln N - c}{2\lambda}\right) \geq 1 - \kappa\mathrm{e}^{-c}.$$

The proof of Theorem 2.1.5 is deferred to Section 2.4.1.

The lower bound provided by Theorem 2.1.5 ensures that the mixing time of the neutral multi-allelic Moran model is at least of order of $\ln N/2\lambda$. Our results do not allow us to prove an upper bound ensuring the existence of a *cutoff phenomenon*. A further study needs to be done in this direction. However, for the parent independent mutation scheme, a further analysis can be done to prove the existence of a cutoff phenomenon in the chi-square and total variation distances, as we next discuss.

Study of the neutral multi-allelic Moran model with parent independent mutation

Consider the following mutation rate matrix:

$$Q_{\boldsymbol{\mu}} := \begin{pmatrix} -|\boldsymbol{\mu}| + \mu_1 & \mu_2 & \mu_3 & \dots & \mu_K \\ \mu_1 & -|\boldsymbol{\mu}| + \mu_2 & \mu_3 & \dots & \mu_K \\ \mu_1 & \mu_2 & -|\boldsymbol{\mu}| + \mu_3 & \dots & \mu_K \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu_1 & \mu_2 & \mu_3 & \dots & -|\boldsymbol{\mu}| + \mu_K \end{pmatrix}, \tag{2.11}$$

where $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_K) \in (0,\infty)^K$ and $|\boldsymbol{\mu}|$ stands for the sum of the entries of $\boldsymbol{\mu}$. Let us define

$$(\mathcal{L}_{N,p} f)(\eta) := \sum_{i,j=1}^{K} \eta(i) \left(\mu_j + \frac{p}{N}\eta(j)\right) [f(\eta - \mathbf{e}_i + \mathbf{e}_j) - f(\eta)],$$

for every f on $\mathcal{E}_{K,N}$ and all $\eta \in \mathcal{E}_{K,N}$, the infinitesimal generator of the neutral multi-allelic Moran process with mutation rate matrix $Q_{\boldsymbol{\mu}}$. The process driven by $\mathcal{L}_{N,p}$ is a special case of the neutral multi-allelic Moran process considered before, but with the difference that the mutation rate only depends on the type of the new individual, i.e. mutation changes each type i individual to type j at rate μ_j, for all $i, j \in [K]$. This is the neutral multi-allelic Moran process with *parent independent* mutation (cf. [Eth11]). Note that $\mathcal{L}_{N,p} = \mathcal{L}_N + \frac{p}{N}\mathcal{A}_N$, where $\mathcal{L}_N := \mathcal{L}_{N,0}$, satisfies

$$(\mathcal{L}_N f)(\eta) := \sum_{i,j=1}^{K} \eta(i)\mu_j [f(\eta - \mathbf{e}_i + \mathbf{e}_j) - f(\eta)],$$

for every f on $\mathcal{E}_{K,N}$ and all $\eta \in \mathcal{E}_{K,N}$.

The next result explicitly describes the spectrum of $\mathcal{L}_{N,p}$ and it is a consequence of Theorem 2.1.3.

Corollary 2.1.6 (Spectrum of $\mathcal{L}_{N,p}$)**.** *For $K \geq 2$, $N \geq 2$ and $p \geq 0$, the infinitesimal generator $\mathcal{L}_{N,p}$ is diagonalisable with eigenvalues λ_L with multiplicity $\binom{K+L-2}{L}$, where*

$$\lambda_{L,p} := -|\boldsymbol{\mu}|L - \frac{p}{N}L(L-1), \tag{2.12}$$

for $L \in [N]_0$. In particular, the spectral gap of $\mathcal{L}_{N,p}$ is $\rho = |\boldsymbol{\mu}|$.

Corollary 2.1.6 is proved in Section 2.5.1.

Remark 2.1.3 (Complete graph model). The *complete graph model* studied by Cloez and Thai [CT16a] in the context of Fleming−Viot particle processes is a particular case of the reversible process driven by $Q_{\boldsymbol{\mu}}$ above when $\mu_j = \frac{1}{K}$, for all $j \in [K]$. In this case, the eigenvalues of the mutation rate are $\beta_0 = 0$ and $\beta_1 = -1$, this last one with multiplicity $K - 1$. In particular, Corollary 2.1.6 improves the Lemma 2.14 in [CT16a].

For a real x and $n \in \mathbb{N}_0$, we denote by $x_{(n)}$, $x_{[n]}$ and $\binom{N}{\eta}$ the *increasing factorial coefficient*, the *decreasing factorial coefficient* and the *multinomial coefficient*, defined by

$$x_{(n)} := \prod_{k=0}^{n-1}(x+k), \quad x_{[n]} := \prod_{k=0}^{n-1}(x-k) \quad \text{and} \quad \binom{N}{\eta} := \frac{N!}{\prod_{j=1}^{K}\eta(j)!},$$

for all $n > 0$ and $\eta \in \mathcal{E}_{K,N}$, respectively. We set by convention $x_{(0)} := 1$ and $x_{[0]} := 1$, even for $x = 0$.

The *multinomial distribution* distribution on $\mathcal{E}_{K,N}$ with parameters N and $\mathbf{q} = (q_1, \dots, q_K) \in (0,1)^K$ such that $|\mathbf{q}| = 1$, denoted $\mathcal{M}(\cdot \mid N, \mathbf{q})$, satisfies

$$\mathcal{M}(\eta \mid N, \mathbf{q}) = \binom{N}{\eta} \prod_{i=1}^{K} q_i^{\eta(i)},$$

for all $\eta \in \mathcal{E}_{K,N}$. Furthermore, the *Dirichlet multinomial distribution* on $\mathcal{E}_{K,N}$ with parameters N and $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_K) \in (0, \infty)^K$, denoted $\mathcal{DM}(\cdot \mid N, \boldsymbol{\alpha})$, satisfies

$$\mathcal{DM}(\eta \mid N, \boldsymbol{\alpha}) = \frac{1}{|\boldsymbol{\alpha}|_{(N)}} \binom{N}{\eta} \prod_{k=1}^{K} (\alpha_k)_{(\eta(k))},$$

for all $\eta \in \mathcal{E}_{K,N}$. $\mathcal{DM}(\cdot \mid N, \boldsymbol{\alpha})$ is a mixture, using a *Dirichlet distribution*, of $\mathcal{M}(\cdot \mid N, \mathbf{q})$. See Mosimann [Mos62] for the original reference to the Dirichlet multinomial distribution and Johnson et al. [JKK05, §13.1], a classical reference on multivariate discrete distributions, for more details.

It is known in population genetics literature that the process driven by $\mathcal{L}_{N,p}$, for $p > 0$, is reversible with stationary distribution $\mathcal{DM}(\cdot \mid N, N\boldsymbol{\mu}/p)$, see e.g. [EG09]. Moreover, the stationary distribution of the process driven by $\mathcal{L}_N$ is $\mathcal{M}(\cdot \mid N, \boldsymbol{\mu}/|\boldsymbol{\mu}|)$, see e.g. [ZL09]. Let us define the distribution $\nu_{N,p}$ on $\mathcal{E}_{K,N}$, for all $p \geq 0$, as follows

$$\nu_{N,p}(\eta) := \begin{cases} \mathcal{DM}(\eta \mid N, N\boldsymbol{\mu}/p) & \text{if} \quad p > 0 \\ \mathcal{M}(\eta \mid \boldsymbol{\mu}/|\boldsymbol{\mu}|) & \text{if} \quad p = 0, \end{cases} \tag{2.13}$$

for all $\eta \in \mathcal{E}_{K,N}$. Then, $\nu_{N,p}$ is the stationary distribution of $\mathcal{L}_{N,p}$, for all $p \geq 0$. Besides, the stationary distribution is continuous when $p \to 0$, in the sense that

$$\lim_{p \to 0} \nu_{N,p}(\eta) = \nu_{N,0}(\eta) =: \nu_N(\eta),$$

for every $\eta \in \mathcal{E}_{K,N}$.

In their study of the spectral properties of the discrete-time analogous of $\mathcal{Q}_N$, Zhou and Lange [ZL09] mainly focus on the case where the process driven by Q is reversible, which is proved to be a necessary and sufficient condition for the reversibility of $\mathcal{Q}_N$. However, the reversibility of Q is not sufficient to ensure the reversibility of the neutral multi-allelic Moran model driven by $\mathcal{Q}_{N,p}$, for $p > 0$, as we discuss in Section 2.5.1. Going further, the next result characterises the reversible neutral multi-allelic Moran processes as those with parent independent mutation.

Lemma 2.1.7 (Reversible neutral Moran process and parent independent mutation)**.** *Assume $K \geq 2$, $N \geq 2$ and $p > 0$. The process driven by $\mathcal{Q}_{N,p}$ is reversible if and only if the mutation rate matrix has the form $Q_{\boldsymbol{\mu}}$ as in (2.11), for some vector $\boldsymbol{\mu}$, and consequently $\mathcal{Q}_{N,p}$ can be written as $\mathcal{L}_{N,p}$. Furthermore, the stationary distribution of the process driven by $\mathcal{L}_{N,p}$ is $\nu_{N,p}$ as defined by (2.13).*

The previous result is expected because of its analogy with the theory on the measure-valued Fleming – Viot process studied in [EK93]. Indeed, the measure-valued Fleming – Viot process is reversible if and only if its mutation factor is parent independent (see e.g. [EK93, Thm. 8.2] and [LSY99, Thm. 1.1]. Although the *"if part"* in Lemma 2.1.7 is well known in the theory related to Moran process, we have not found a explicit statement, nor a proof, of this equivalence between parent independent mutation scheme and the reversibility of the neutral multi-allelic Moran model considered here. Thus, for the sake of completeness, we provide a proof of Lemma 2.1.7 in Appendix 2.C.

Section 2.5 is devoted to the study of the spectral properties of $\mathcal{L}_{N,p}$, for $p \geq 0$, and its applications to the study of the convergence to stationarity. Our results in this section include a complete description of the set of eigenvalues and eigenfunctions of $\mathcal{L}_{N,p}$ and an explicit expression for its transition function. The eigenfunctions of $\mathcal{L}_{N,p}$, $p > 0$, are explicitly given in terms of *multivariate Hahn polynomials*, which are orthogonal with respect to the compound Dirichlet multinomial distribution (cf. [KM75; KZ09]). The eigenfunctions of $\mathcal{L}_N$, i.e. for $p = 0$, are explicitly given in terms of *multivariate Krawtchouk polynomials*, which are orthogonal with respect to the multinomial distribution (cf. [KM65; ZL09; DG14]).

Cutoff phenomenon

The *cutoff phenomenon* has been a rich topic of research on Markov chains since its introduction by the works of Aldous, Diaconis and Shahshahani in the 1980s (cf. [DS81; Ald83; AD86]). A Markov chain presents a cutoff if it exhibits an abrupt transition in its convergence to stationarity. Some of the most used notions of convergence are, as we consider here, the total variation and the chi-square distances. A good introduction to this subject can be found in the classic book of Levin and Peres [LP17, Ch. 18] and in the exhaustive work of Chen, Saloff-Coste et al. [Sal97; Che06; CS08; CS10; CHS17].

A typical scenario for the existence of a cutoff is a Markov chain with a high degree of symmetry. Hence, the cutoff phenomenon has been deeply studied for the movement on N independent particles on K sites, model which is usually known as *product chain*. Ycart [Yca99] studied the cutoff in total variation for N independent particles driven by a diagonalisable rate matrix. Later, Barrera et al. [BLY06] and Connor [Con10] studied the cutoff on this model according to other notions of distance. See also [Lac15] [LP17, Ch. 20], [CHS17] and [CK18] for more recent studies about the cutoff on product chains. The Moran model we consider here preserves the high level of symmetry of the product chain, but the movements of the particles are not independent. Indeed, the particles interact according to a reproduction process that favours the jumps to the sites with greater proportions of individuals.

Before formally defining the cutoff phenomenon, let us recall the *chi-square divergence* (sometimes called "distance"), which naturally arises in the context of reversible Markov chains. The chi-square divergence of μ_2 with respect to the target distribution μ_1 is defined by

$$\chi^2(\mu_2 \mid \mu_1) := \sum_{\omega \in \Omega} \frac{[\mu_2(\omega) - \mu_1(\omega)]^2}{\mu_1(\omega)} = \|\mu_2 - \mu_1\|^2_{\frac{1}{\mu_1}},$$

where $\|\cdot\|_{\frac{1}{\mu_1}}$ stands for the norm in $l^2(\mathbb{R}^\Omega, \frac{1}{\mu_1})$, and $\frac{1}{\mu_1}$ is the measure $\omega \mapsto 1/\mu_1(\omega)$.

The chi-square divergence is not a metric, but a measure of the difference between two probability distributions. Note that the chi-square divergence, as well as the total variation distance, are special cases of the so called $f-$divergence functions, which measure the "difference" between two probability distributions [NN14]. In this context, $\chi^2(\mu_2 \mid \mu_1)$ is also known as *Pearson chi-square divergence*.

Abusing notation, let us define the functions χ^2_η and $\mathrm{d}^{\mathrm{TV}}_\eta$, as follows

$$\mathrm{d}^{\mathrm{TV}}_\eta(t) := \mathrm{d}^{\mathrm{TV}}(\delta_\eta \mathrm{e}^{t\mathcal{L}_{N,p}}, \nu_{N,p}) = \frac{1}{2}\sum_{\xi\in\mathcal{E}_{K,N}} \left| \left(\mathrm{e}^{t\mathcal{L}_{N,p}}\delta_\xi\right)(\eta) - \nu_{N,p}(\xi)\right|,$$

$$\chi^2_\eta(t) := \chi^2(\delta_\eta \mathrm{e}^{t\mathcal{L}_{N,p}} \mid \nu_{N,p}) = \sum_{\xi\in\mathcal{E}_{K,N}} \frac{\left[\left(\mathrm{e}^{t\mathcal{L}_{N,p}}\delta_\xi\right)(\eta) - \nu_{N,p}(\xi)\right]^2}{\nu_{N,p}(\xi)}.$$

The functions $\mathrm{d}^{\mathrm{TV}}_\eta$ and χ^2_η are thus measures of the convergence to stationary of the process driven by $\mathcal{L}_{N,p}$ at time t and with initial configuration $\eta \in \mathcal{E}_{K,N}$. In agreement with [Zho08; KZ09] we call χ^2_η and $\mathrm{d}^{\mathrm{TV}}_\eta$ the *total variation* and the *chi-square distances* to stationarity, respectively.

As the number of individuals varies we obtain an infinite family of continuous-time finite Markov chains $\{(\mathcal{E}_{K,N}, \mathcal{L}_{N,p}, \nu_{N,p}), N \geq 2\}$. For each $N \geq 2$ let us denote by $\chi^2_{N\mathbf{e}_k}(t)$ (resp. $\mathrm{d}^{\mathrm{TV}}_{N\mathbf{e}_k}(t)$) the chi-square distance (resp. total variation distance) to stationarity of the process driven by $\mathcal{L}_{N,p}$ at time t, when the initial distribution is concentrated at $N\mathbf{e}_k \in \mathcal{E}_{K,N}$. Note that $\chi^2_{N\mathbf{e}_k}(0) \to \infty$ and $\mathrm{d}^{\mathrm{TV}}_{N\mathbf{e}_k}(0) \to 1$, when $N \to \infty$.

Definition 2 (Chi-square and total variation cutoff)**.** We say that $\{\chi^2_{N\mathbf{e}_k}(t), N \geq 2\}$ exhibits a (t_N, b_N) chi-square cutoff if $t_N \geq 0$, $b_N \geq 0$, $b_N = o(t_N)$ and

$$\lim_{c\to\infty}\limsup_{N\to\infty} \chi^2_{N\mathbf{e}_k}(t_N + c\, b_N) = 0, \quad \lim_{c\to-\infty}\liminf_{N\to\infty} \chi^2_{N\mathbf{e}_k}(t_N + c\, b_N) = \infty.$$

Analogously, we say that $\{\mathrm{d}^{\mathrm{TV}}_{N\mathbf{e}_k}(t), N \geq 2\}$ exhibits a (t_N, b_N) total variation cutoff if $t_N \geq 0$, $b_N \geq 0$, $b_N = o(t_N)$ and

$$\lim_{c\to\infty}\limsup_{N\to\infty} \mathrm{d}^{\mathrm{TV}}_{N\mathbf{e}_k}(t_N + c\, b_N) = 0, \quad \lim_{c\to-\infty}\liminf_{N\to\infty} \mathrm{d}^{\mathrm{TV}}_{N\mathbf{e}_k}(t_N + c\, b_N) = 1.$$

The sequences $(t_N)_{N\geq 2}$ and $(b_N)_{N\geq 2}$ are called *cutoff* and *window sequences*, respectively.

See Definition 2.1 and Remark 2.1 in [CS08].

The cutoff phenomenon describes an abrupt transition in the convergence to stationarity: over a negligible period given by the window sequence $(b_N)_{N>2}$, the distance from equilibrium drops from near its initial value to near zero at a time given by the cutoff sequence $(t_N)_{N\geq 2}$.

A stronger condition for the existence of a (t_N, b_N) chi-square cutoff (resp. total variation cutoff) is the existence of the limit

$$G_k(c) := \lim_{N\to\infty} \chi^2_{N\mathbf{e}_k}(t_N + c\, b_N) \quad \left(\text{resp. } H_k(c) := \lim_{N\to\infty} \mathrm{d}^{\mathrm{TV}}_{N\mathbf{e}_k}(t_N + c\, b_N)\right),$$

for a function G_k (resp. H_k), for $k \in [K]$, satisfying:

$$\lim_{c\to-\infty} G_k(c) = \infty \text{ and } \lim_{c\to\infty} G_k(c) = 0, \quad \left(\text{resp. } \lim_{c\to-\infty} H_k(c) = 1 \text{ and } \lim_{c\to\infty} H_k(c) = 0)\right).$$

Actually, in this case the (t_N, b_N) cutoff is said to be *strongly optimal*, see e.g. Definition 2.2 and Proposition 2.2 in [CS08]. See Sections 2.1 and 2.2 of [CS08] and Chapter 2 in [Che06] for more details about the definition of (t_N, b_N) cutoff and window optimality. Figure 7, in Section shows classic profiles of functions G_k and H_k as defined in (16) for total variation and chi-square cutoffs.

The next two results establish the existence of cutoff phenomena in the chi-square and the total variation distances for the multi-allelic Moran process driven by $\mathcal{L}_{N,p}$, for $p \geq 0$, when the initial distribution is concentrated at $N\mathbf{e}_k$, for $k \in [K]$. In the chi-square case we are able to explicitly provide the limit profile of the distance. Moreover, we prove the total variation distance to stationarity of the mutation process driven by $\mathcal{L}_N$, i.e. for $p=0$, has a Gaussian profile, when all the individuals are initially of the same type.

Theorem 2.1.8 (Strongly optimal chi-square cutoff when $N \to \infty$). *For $k \in [K]$, with $K \geq 2$, $p \geq 0$ and every $c \in \mathbb{R}$, we have*

$$\lim_{N\to\infty} \chi^2_{N\mathbf{e}_k}(t_{N,c}) = \exp\{K_{k,p}\mathrm{e}^{-c}\} - 1, \tag{2.14}$$

where $t_{N,c} = \dfrac{\ln N + c}{2|\boldsymbol{\mu}|}$ *and* $K_{k,p} = \dfrac{|\boldsymbol{\mu}|(|\boldsymbol{\mu}| - \mu_k)}{\mu_k(|\boldsymbol{\mu}|+p)}$. *Consequently, the Markov process driven by $\mathcal{L}_{N,p}$ has a strongly optimal $\left(\frac{\ln N}{2|\boldsymbol{\mu}|}, 1\right)$ chi-square cutoff when $N \to \infty$.*

Figure 8 in the introduction illustrates the convergence of $\chi^2(t_{N,c}, N\mathbf{e}_k)$, for $t_{N,c} = \frac{\ln(N)+c}{2|\boldsymbol{\mu}|}$, towards $G_k(c) = \exp\{K_{k,p}\mathrm{e}^{-c}\} - 1$, when $N \to \infty$.

Theorem 2.1.9 (Total variation cutoff when $N \to \infty$). *For every $k \in [K]$, with $K \geq 2$, $p \geq 0$ and every $c > 0$, we have*

$$\mathrm{d}^{\mathrm{TV}}_{N\mathbf{e}_k}\left(\frac{\ln N - c}{2|\boldsymbol{\mu}|}\right) \geq 1 - 32|\mu|\kappa_k \mathrm{e}^{-c},$$
$$\lim_{N\to\infty} \mathrm{d}^{\mathrm{TV}}_{N\mathbf{e}_k}\left(\frac{\ln N + c}{2|\boldsymbol{\mu}|}\right) \leq \sqrt{\exp\{K_{k,p}\mathrm{e}^{-c}\} - 1},$$

where $\kappa_k = \max_{r:r\neq k} \dfrac{\mu_r \wedge \mu_k}{\mu_k}$ *and* $K_{k,p} = \dfrac{|\boldsymbol{\mu}|(|\boldsymbol{\mu}| - \mu_k)}{\mu_k(|\boldsymbol{\mu}|+p)}$. *Consequently, the Markov process driven by $\mathcal{L}_{N,p}$ exhibits a $\left(\frac{\ln N}{2|\boldsymbol{\mu}|}, 1\right)$ total variation cutoff when $N \to \infty$.*

Moreover, when $p = 0$ the limit profile of the total variation distance satisfies

$$\lim_{N\to\infty} \mathrm{d}^{\mathrm{TV}}_{N\mathbf{e}_k}(t_{N,c}) = 2\Phi\left(\frac{1}{2}\sqrt{K_{k,0}\mathrm{e}^{-c}}\right) - 1,$$

where Φ is the cumulative distribution function of the standard normal distribution. Thus, there exists a strongly optimal $\left(\frac{\ln N}{2|\boldsymbol{\mu}|}, 1\right)$ total variation cutoff for the process driven by $\mathcal{L}_N$ when $N \to \infty$.

Proof of Theorem 2.1.8 and 2.1.9 will be given in Section 2.5.1.

Figure 9 in the introduction illustrates the convergence of $\mathrm{d}_{\mathrm{TV}}(t_{N,c}, N\mathbf{e}_k)$, for $t_{N,c} = \frac{\ln(N)+c}{2|\boldsymbol{\mu}|}$, towards $H_k(c) = 2\Phi\left(\frac{1}{2}\sqrt{K_{k,0}\mathrm{e}^{-c}}\right) - 1$, when $N \to \infty$.

During the proof of Theorem 2.1.8, we prove the following result which is of independent interest.

Corollary 2.1.10 (Law of the process driven $\mathcal{L}_N$). *The law of the process driven by $\mathcal{L}_N$ at time t when initially all the individuals are of type $k \in [K]$ is multinomial*

$$\mathcal{M}\left(\cdot \mid N, \frac{\boldsymbol{\mu}}{|\boldsymbol{\mu}|}(1 - \mathrm{e}^{-|\boldsymbol{\mu}|t}) + \mathrm{e}^{-|\boldsymbol{\mu}|t}\mathbf{e}_k\right).$$

Some authors have studied the existence of a cutoff in Moran type models. For instance, Donelly and Rodrigues [DR00] proved the existence of a cutoff for the two-allelic neutral Moran model in the separation distance. In order to do that, they used a duality property of the

Moran process and found an asymptotic expression for the convergence in separation distance for a suitable scaled time, when the number of individuals tends to infinity. Khare and Zhou [KZ09] proved bounds for the chi-square distance in a discrete-time multi-allelic Moran process that implies the existence of a cutoff. Diaconis and Griffiths [DG19] studied the existence of a chi-square and total variation cutoffs for a discrete-time analogous of the mutation process generated by $\mathcal{L}_N$. Theorems 2.1.8 and 2.1.9 sharpen the results in [KZ09] and [DG19], since they provide the limit profiles for the chi-square and the total variation distances, for $p \geq 0$ and $p = 0$, respectively. Besides, Theorem 2.1.9 is, as far as we know, the first result ensuring the existence of a total variation cutoff phenomenon for the neutral Moran model with parent independent mutation with $p > 0$.

Links with other models

Moran type models are fundamental in population genetics and other branches of applied mathematics [Dur08], [Eth11]. Simpler than the Wright – Fisher model, the Moran model is more tractable mathematically and several quantities of interest can be explicitly computed. There is a rich literature on Moran models in population genetics and other fields, since the seminal work of Moran [Mor58]. In particular, the study of spectral properties of the generator of a Markov process is an interesting and active topic of research in population genetics. See e.g. [KZ09], [ZL09], [MP14], [Möh18], [Möh19] and the references therein.

We want to remark that the utility of Moran processes is behind population genetics. For instance, the mutation process driven by $\mathcal{Q}_N$ is a particular case of the *zero range process*, where the *kinetics*, i.e. the rate at which the particles are expelled from one state, is proportional to the number of particles occupying that state. Moreover, the mutation process driven by $\mathcal{L}_N$ corresponds to the *mean-field* version of the zero range process. The very recent paper of Hermon and Salez [HS19] shows that the *Dirichlet form* of a zero range process can be controlled in terms of the *Dirichlet form* of a single particle. We believe that the methods in [HS19] could be very useful for the further study of the ergodicity of the Moran process driven by $\mathcal{Q}_{N,p}$, for $p \geq 0$, by controlling its *Dirichlet form*.

Consider a Markov process in $\mathcal{E}_{K,N}$ with generator $\mathcal{F}$ acting on a real function f on $\mathcal{E}_{K,N}$ as follows

$$(\mathcal{F}f)(\eta) = \sum_{i,j\in[K]} \eta(i) \left[f(\eta - \mathbf{e}_i + \mathbf{e}_j) - f(\eta) \right] \left(\mu_{i,j} + \frac{p_i}{N-1}\, \eta(j) \right), \tag{2.15}$$

for every $\eta \in \mathcal{E}_{K,N}$, where $p_i \geq 0$, for all $i \in [K]$. The process driven by $\mathcal{F}$ is a particular case of the countable state space continuous-time Markov processes introduced by Ferrari and Marić [FM07] to approximate the *quasi stationary distribution* (QSD) of an absorbing Markov chain on a countable space. Ferrari and Marić called these Markov chains *Fleming – Viot particle processes*. The random empirical distribution associated to the process driven by $\mathcal{F}$ has been proved to approximate the QSD of an absorbing Markov process driven by an irreducible rate matrix Q on $[K]$ which jumps, with rate p_i, from i to a fictitious absorbing state [AFG11]. This kind of N particle interacting process was originally introduced independently and simultaneously by Burdzy et. al. [BHM00] and Del Moral and Miclo [DM00b][1] in the continuous state space settings. The study of the evolution of the proportion of particles in each state for a Moran-type particle system driven by $\mathcal{F}$ is an active topic of research. In particular, many papers have

[1] Actually, Del Moral and Miclo called this process *Moran-type particle system*, which is maybe a more accurate name in the discrete state space setting, in order to avoid confusion with the existence of a measure-valued process related to the Moran process in population genetics introduced by Fleming and Viot [FV79] and named *Fleming – Viot process*. See [FV79, Appendix B] for a discussion on the relationship between Fleming –Viot process (in the sense of population genetics) and the multi-allelic Moran model, see also [Fen10, Section 6.2].

been focused on the convergence and the speed of convergence of the proportion of particles in each state when the time and the number of particles tend toward infinity. See e.g. [FM07], [AFG11], [CT16a], [CT16b], [Vil20] and the references therein. Note that the Fleming – Viot particle process generated by (2.15) is different from the classical Fleming – Viot measure-valued diffusion process, which can be obtained as a limit of particle systems also including mutations and reproductions, but with a different parameter scaling (cf. [EK93]).

The generator $\mathcal{F}$ is also interesting in population genetics. From this point of view, it models the evolution of a population with an irreducible mutation process driven by $Q = (\mu_{i,j})_{i,j=1}^K$ and *selection at death* given by the coefficients $(p_i)_{i=1}^K$ (cf. [MW09]). Unlike the other type of selection that has been mostly considered in population genetics, which is the *selection at reproduction* (cf. [Dur08], [MW09] and [Eth11]), which assumes that the rates p_i in the definition (2.15) do not depend on i but on j, i.e. on the type of the individual that is going to reproduce.

Note that when $p_i = p$, for all $i \in [K]$, the generator $\mathcal{F}$ reduces to $\mathcal{Q}_{N,p}$. Theorem 2.1.3 thus provides an explicit description for the eigenvalues of the Fleming – Viot (or Moran type) particle process with irreducible mutation rate matrix Q and the transition rate to the absorbing state is uniform on $[K]$, which is known in the theory of QSD as *uniform killing* [MV12, §2.3]. This is, for example, the case of the complete graph process studied by Cloez and Thai [CT16a] and the neutral Moran model process with circulant mutation rate matrix considered in [Cor21a].

Structure of the chapter

The rest of the chapter is organised as follows. In Section 2.2 we study the state spaces of the neutral multi-allelic Moran models, when the individuals are assumed distinguishable or indistinguishable, respectively. We particularly focus on the study of the vector spaces of real functions defined on the state spaces of these two models. The notations and results in Section 2.2 are used to prove our main theorems in Section 2.3. Sections 2.3.1, 2.3.2 and 2.3.3 are devoted to the proofs of Theorems 2.1.1, 2.1.2 and 2.1.3, respectively. In Section 2.4 we focus on the applications of our main results to the asymptotic exponential ergodicity in total variation distance of the process driven by $\mathcal{Q}_{N,p}$ to its stationary distribution, using the eigenstructure of Q. In particular, we prove Corollary 2.1.4 and Theorem 2.1.5. We also consider several examples of neutral multi-allelic Moran processes with diagonalisable and non-diagonalisable mutation rate matrices, throughout this chapter. In Section 2.5 we consider the neutral multi-allelic Moran process with parent independent mutation and provide a complete description of its eigenvalues and eigenfunctions. We also prove Theorems 2.1.8 and 2.1.9 about the existence of a cutoff phenomena in the chi-square and the total variation distances, when initially all the individuals are of the same type.

2.2 State spaces for distinguishable and indistinguishable particle processes

The Moran model can be seen as a system of N interacting particles on K sites moving according to a continuous-time Markov chain. For the same model, we study two different situations. Although the sites themselves are supposed to be distinguishable, the N particles can be considered either *distinguishable* or *indistinguishable*. According to both interpretations we describe two state spaces for the two Markov chains modelling the N independent particle systems. We study how the vector spaces of the real functions defined on those state spaces are related.

For N *distinguishable* particles on K sites, the state space of the model describes the location of each particle, i.e. it is the set $[K]^N$. This is the state space considered in [FM07] and [Eth11]. The set of real functions on $[K]$, denoted $\mathbb{R}^{[K]}$, may be endowed with a vector space structure.

Thus, the set of real functions on $[K]^N$ may be considered as a tensor product of N vectors in $\mathbb{R}^K$ as we commented in the introduction.

When the N particles are considered *indistinguishable*, what matters is the number of particles present at each of the K sites. The state space for this second model, as in [CT16b] and [EG09], is the set $\mathcal{E}_{K,N}$ defined by (2.1) with cardinality equal to $\mathrm{Card}\,(\mathcal{E}_{K,N}) = \binom{K-1+N}{N}$.

For any k, $1 \le k \le K$, let us denote by x_k the k-th coordinate function defined by

$$x_k : \eta = (\eta(1), \eta(2), \dots, \eta(K)) \in \mathcal{E}_{K,N} \mapsto \eta(k) \in \mathbb{R}.$$

Let us also denote by $\mathbf{x}^\alpha$ the monomial on $\mathcal{E}_{K,N}$ defined by

$$\mathbf{x}^\alpha := x_1^{\alpha_1} x_2^{\alpha_2} \dots x_K^{\alpha_K}, \tag{2.16}$$

where $\alpha \in \mathcal{E}_{K,L}$, for $L \in [N]$.

For $0 \le L \le N$, let us denote by $H_{K,L}$ the vector space of homogeneous polynomial functions of degree L in variables x_k, $1 \le k \le K$ on $\mathcal{E}_{K,N}$ and the null function. From the definition of $\mathcal{E}_{K,N}$, it follows that the function $\sum_{k=1}^K x_k$ is equal to the constant function equal to N. $H_{K,L}$ may be considered as a subspace of $H_{K,L'}$ when $0 \le L < L' \le N$ by identifying $P(x_1, x_2, \dots, x_K) \in H_{K,L}$ with

$$\frac{1}{N^{L'-L}} \left(\sum_{k=1}^K x_k \right)^{L'-L} P(x_1, x_2, \dots, x_K) \in H_{K,L'}.$$

We will say that the *degree of homogeneity* of an homogeneous polynomial P is L, if P is the sum of monomials $\mathbf{x}^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \dots x_K^{\alpha_K}$ with the same value of $|\alpha| = L$, and the value L is the smallest as possible. This corresponds to the fact that there is no factor equal to $x_1 + x_2 + \cdots + x_K$ in the factorisation of P. The *total degree* of a polynomial P is the minimum value of L such that $P = P_L + R$ where P_L is homogeneous of degree L and all the monomials in R have a maximum degree strictly less than L. Such an expression for P, which is not unique, may be obtained by replacing x_K by $N - \sum_{k=1}^{K-1} x_k$ in P and adding the monomials in $P(x_1, \dots, x_{K-1}, N - \sum_{k=1}^{K-1} x_k)$ of maximum total degree to define P_L.

The following interpolation result shows that $\mathbb{R}^{\mathcal{E}_{K,N}}$ is actually equal to $H_{K,N}$, the space of homogeneous polynomials of degree N.

Lemma 2.2.1 (Interpolation on $\mathcal{E}_{K,N}$)**.** *Let $K \ge 2$ and $N \ge 1$. Then*

(a) For any real function f on $\mathcal{E}_{K,N}$ there exists a unique homogeneous polynomial $P \in H_{K,N}$ of degree N such that $f(\eta) = P(\eta)$, for all $\eta \in \mathcal{E}_{K,N}$.

(b) The set of monomials of degree N

$$\mathcal{B}_{H_{K,N}} := \{\mathbf{x}^\alpha, \alpha \in \mathcal{E}_{K,N}\}$$

where $\mathbf{x}^\alpha$ is defined by (2.16), is a basis of $\mathbb{R}^{\mathcal{E}_{K,N}}$.

The proof of Lemma 2.2.1 is mostly technical and is deferred to Appendix 2.A.

Remark 2.2.1 (Dimension of $H_{K,N}$). As a consequence of Lemma 2.2.1-(b) we have that the dimension of $H_{K,N}$ equals $\binom{K+N-1}{N}$.

A natural link between the two state spaces is $\phi_{K,N} : [K]^N \to \mathcal{E}_{K,N}$, defined by

$$\phi_{K,N} : (k_1, k_2, \dots, k_N) \mapsto (\eta(1), \eta(2), \dots, \eta(K)), \tag{2.17}$$

where $\eta(k) = \mathrm{Card}(\{n,\ 1 \leq n \leq N,\ k_n = k\})$, for all $k \in [K]$. The function $\phi_{K,N}$ is obtained by *forgetting the identity* of the N particles. Note that $\psi_{K,N}$, defined in (2.6), is a right inverse of $\phi_{K,N}$, i.e. $\phi_{K,N} \circ \psi_{K,N} = \mathrm{Id}_{\mathcal{E}_{K,N}}$, where $\mathrm{Id}_{\mathcal{E}_{K,N}}$ stands for the identity function on $\mathcal{E}_{K,N}$.

Let us denote by Sym the *symmetrisation* endomorphism, acting on function $f \in \mathbb{R}^{[K]^N}$ as defined by (2.5). In fact, Sym is the projector onto the subspace of symmetric functions, denoted $\mathrm{Sym}(\mathbb{R}^{[K]^N})$.

Note that $\phi_{K,N}$ is a symmetric function on $[K]^N$. Furthermore, the equality $\phi_{K,N}(\mathbf{x}) = \phi_{K,N}(\mathbf{y})$ holds if and only if $\mathbf{y}$ is obtained from $\mathbf{x}$ by a permutation of its components. Hence, if f is symmetric and $\mathbf{x}$ and $\mathbf{y}$ are elements in $[K]^N$ such that $\phi_{K,N}(\mathbf{x}) = \phi_{K,N}(\mathbf{y})$, then $f(\mathbf{x}) = f(\mathbf{y})$.

In general, for every function f on $[K]^N$ it is not always possible to define a function $\tilde{f}$ on $\mathcal{E}_{K,N}$ such that $f = \tilde{f} \circ \phi_{K,N}$ holds. We claim that such a function $\tilde{f}$ exists if and only if f is symmetric.

Lemma 2.2.2 (Link between $\mathbb{R}^{\mathcal{E}_{K,N}}$ and $\mathrm{Sym}(\mathbb{R}^{[K]^N})$)**.** *The linear operator*

$$\Phi_{K,N} : f \in \mathrm{Sym}\left(\mathbb{R}^{[K]^N}\right) \mapsto f \circ \psi_{K,N} \in \mathbb{R}^{\mathcal{E}_{K,N}}, \tag{2.18}$$

where $\psi_{K,N}$ is defined by (2.6), is an isomorphism. In particular, the dimension of the space of symmetric functions on $[K]^N$ is

$$\dim\left(\mathrm{Sym}\left(\mathbb{R}^{[K]^N}\right)\right) = \binom{K+N-1}{N}.$$

Proof. Note that $\Phi_{K,N}$ is linear and well defined. Moreover, for any function h on $\mathcal{E}_{K,N}$, the function $h \circ \phi_{K,N}$ is symmetric on $[K]^N$ and satisfies $\Phi_{K,N}\left(h \circ \phi_{K,N}\right) = h$, proving that $\Phi_{K,N}$ is an isomorphism. $\square$

Lemma 2.2.2 justifies the well definiteness of $\tilde{V}_\eta$, defined by (2.9), for $\eta \in \bigcup_{L=1}^{N} \mathcal{E}_{K-1,L}$. The relationship between f and $\tilde{f}$ is shown in the following diagram:

$$\begin{tikzcd}
{[K]^N} \arrow[d, "\phi_{K,N}"'] \arrow[dr, "f"] & \\
\mathcal{E}_{K,N} \arrow[r, "\tilde{f}"'] & \mathbb{R}.
\end{tikzcd}$$

We denote by U_0 the K-dimensional all-one vector, which is always a right eigenvector associated to zero of every K-dimensional rate matrix of a continuous-time Markov chain. Let $K \geq 2$, $N \geq 2$ and $1 \leq L \leq N$ and let us consider L vectors $V_1, V_2, \ldots, V_L$ in $\mathbb{R}^K$, non-proportional to U_0, and f the function equal to the following symmetrised tensor product

$$f := \mathrm{Sym}(V_1 \otimes V_2 \otimes \cdots \otimes V_L \otimes \underbrace{U_0 \otimes \cdots \otimes U_0}_{N-L}) \in \mathrm{Sym}\left(\mathbb{R}^{[K]^N}\right).$$

Note that,

$$f(k_1, k_2, \ldots, k_N) = \frac{1}{N!} \sum_{\sigma \in \mathcal{S}_N} V_1(k_{\sigma(1)}) V_2(k_{\sigma(2)}) \times \cdots \times V_L(k_{\sigma(L)}). \tag{2.19}$$

We denote by $\mathcal{I}_{L,N}$, for $1 \leq L \leq N$, the set of all injective applications from $[L]$ to $[N]$. For every $\sigma \in \mathcal{S}_N$, the map $s_\sigma : n \in [L] \mapsto \sigma(n) \in \{\sigma(1), \ldots, \sigma(L)\}$ is an injective map in $\mathcal{I}_{L,N}$ and

σ is completely determined by this function s_σ and a bijective application $\beta : (L+1,\dots,N) \to [N]\setminus s_\sigma([L])$. For each s_σ, there are $(N-L)!$ such applications β. Thus, using (2.19) we obtain

$$f(k_1,k_2,\dots,k_N) = \frac{(N-L)!}{N!} \sum_{s\in\mathcal{I}_{L,N}} V_1(k_{s(1)})V_2(k_{s(2)}) \times \cdots \times V_L(k_{s(L)}).$$

In order to simplify the calculations we denote by $\xi(V_1,V_2,\dots,V_L)$ the function on $[K]^N$ defined by

$$\xi(V_1,V_2,\dots,V_L) : (k_1,k_2,\dots,k_N) \mapsto \sum_{s\in\mathcal{I}_{L,N}} V_1(k_{s(1)})V_2(k_{s(2)})\dots V_L(k_{s(L)}). \tag{2.20}$$

Note that $\xi(V_1,V_2,\dots,V_L) = \frac{N!}{(N-L)!} f$. Since $\xi(V_1,V_2,\dots,V_L)$ is symmetric, Lemma 2.2.2 ensures the existence of a unique function $\tilde{\xi}(V_1,V_2,\dots,V_L)$ on $\mathcal{E}_{K,N}$ given by

$$\tilde{\xi}(V_1,V_2,\dots,V_L) = \Phi_{K,N}\,\xi(V_1,V_2,\dots,V_L). \tag{2.21}$$

The following two equalities are thus satisfied:

$$\xi(V_1,\dots,V_L) = \tilde{\xi}(V_1,\dots,V_L)\circ\phi_{K,N}, \quad \tilde{\xi}(V_1,\dots,V_L) = \xi(V_1,\dots,V_L)\circ\psi_{K,N}, \tag{2.22}$$

where $\phi_{K,N}$ and $\psi_{K,N}$ are defined by (2.17) and (2.6), respectively.

The next result provides recursive expressions for the functions $\xi(V_1,\dots,V_L)$, defined in (2.20), and $\tilde{\xi}(V_1,\dots,V_L)$, defined in (2.21), for $L\in[N]$. Furthermore, we prove that $\tilde{V}_\eta$, as defined by (2.9), is a polynomial of total degree $|\eta|$, for $\eta \in \bigcup_{L=1}^N \mathcal{E}_{K-1,L}$.

Lemma 2.2.3. *The following properties are verified:*

(a) For $L=1$: if $V_1 = [a_1,a_2,\dots,a_K]^T$ is non-proportional to U_0, then $\xi(V_1)$ and $\tilde{\xi}(V_1)$, defined by (2.20) and (2.21), satisfy

$$\xi(V_1) : (k_1,k_2,\dots,k_N) \mapsto \sum_{i=1}^N V_1(k_i),$$
$$\tilde{\xi}(V_1) : (\eta(1),\eta(2),\dots,\eta(K)) \mapsto \sum_{j=1}^K a_j\eta(j). \tag{2.23}$$

(b) For any L, $2\le L\le N-1$: if the L vectors $V_i = [a_{i,1},a_{i,2},\dots,a_{i,K}]^T$, $1\le i\le L$, are non-proportional to U_0, then $\xi(V_1,\dots,V_L)$ and $\tilde{\xi}(V_1,\dots,V_L)$ satisfy

$$\xi(V_1,\dots,V_L) = \xi(V_1,\dots,V_{L-1})\xi(V_L) - \sum_{i=1}^{L-1} \xi(V_1,\dots,V_{i-1},V_i\odot V_L,V_{i+1},\dots,V_{L-1}),$$
$$\tilde{\xi}(V_1,\dots,V_L) = \tilde{\xi}(V_1,\dots,V_{L-1})\tilde{\xi}(V_L) - \sum_{i=1}^{L-1} \tilde{\xi}(V_1,\dots,V_{i-1},V_i\odot V_L,V_{i+1},\dots,V_{L-1}),$$

where $V_i\odot V_L$ stands for the Hadamard (componentwise) product of the vectors V_i and V_L.

In particular, when $L=2$ and the two vectors $V_1 = [a_1,a_2,\dots,a_K]^T$ and $V_2 = [b_1,b_2,\dots,b_K]^T$ are non-proportional to U_0, then $\tilde{\xi}(V_1,V_2)$ is the quadratic polynomial given by

$$\tilde{\xi}(V_1,V_2) = \tilde{\xi}(V_1)\tilde{\xi}(V_2) - \tilde{\xi}(V_1\odot V_2). \tag{2.24}$$

(c) For any L, $1 \leq L \leq N$: if the L vectors $V_i = [a_{i,1}, a_{i,2}, \dots, a_{i,K}]^T$, $1 \leq i \leq L$, are non-proportional to U_0, then $\tilde{\xi}(V_1, V_2, \dots, V_L)$ is a polynomial of total degree L satisfying

$$\tilde{\xi}(V_1, V_2, \dots, V_L) = \prod_{i=1}^{L} \tilde{\xi}(V_i) + q, \tag{2.25}$$

where q is a polynomial of total degree strictly less than L. In particular, $\tilde{V}_\eta$, as defined by (2.9), is a polynomial of total degree $|\eta|$, for $\eta \in \bigcup_{L=0}^{N} \mathcal{E}_{K-1,L}$.

The proof of Lemma 2.2.3 can be found in Appendix 2.A.

The following result helps us to construct from a basis of $\mathbb{R}^K$, three bases for the vector spaces $\mathbb{R}^{[K]^N}$, $\mathrm{Sym}(\mathbb{R}^{[K]^N})$ and $\mathbb{R}^{\mathcal{E}_{K,N}}$, respectively.

Proposition 2.2.4. *Let U_0 be the all-one vector in $\mathbb{R}^K$ and $U_1, U_2, \dots, U_{K-1} \in \mathbb{R}^K$ such that*

$$\mathcal{U} = \{U_0, U_1, \dots, U_{K-1}\}$$

is a basis of $\mathbb{R}^K$. The following statements hold:

a) $\mathcal{U}^N$, defined as $\mathcal{U}^N := \{W_1 \otimes W_2 \otimes \cdots \otimes W_N,$ where $W_i \in \mathcal{U}$, for $i \in [N]\}$ is a basis of $\mathbb{R}^{[K]^N}$.

b) $\mathcal{S}^N$, defined as

$$\mathcal{S}^N := \{\underbrace{U_0 \otimes \cdots \otimes U_0}_{N \ times}\} \cup \bigcup_{L=1}^{N} \{V_\eta, \eta \in \mathcal{E}_{K-1,L},\}$$

where V_η is defined by (2.8), is a basis of $\mathrm{Sym}\left(\mathbb{R}^{[K]^N}\right)$.

c) $\tilde{\mathcal{S}}^N$, defined as

$$\tilde{\mathcal{S}}^N := \{\underbrace{U_0 \otimes \cdots \otimes U_0}_{K \ times}\} \cup \bigcup_{L=1}^{N} \{\tilde{V}_\eta, \eta \in \mathcal{E}_{K-1,L}\}, \tag{2.26}$$

where $\tilde{V}_\eta$ is defined by (2.9), is a basis of $\mathbb{R}^{\mathcal{E}_{K,N}}$.

The proof of Proposition 2.2.4 is deferred to Appendix 2.A.

2.3 Spectrum of the neutral multi-allelic Moran process

The main goal of this section is to prove Theorem 2.1.3. In Section 2.3.1 we prove Theorem 2.1.1 describing the set of eigenvalues of the composition chain $\mathcal{Q}_N$ in terms of the eigenvalues of Q. Moreover, we construct right eigenvectors of $\mathcal{Q}_N$ using the symmetrised tensor product of right eigenvectors of Q. Later, in Section 2.3.2 we prove Theorem 2.1.2. Using the results in these two sections we prove Theorem 2.1.3 in Section 2.3.3.

2.3.1 Proof of Theorem 2.1.1

As we commented in Section 2.2, the N particles in the neutral multi-allelic Moran type process can be considered distinguishable or indistinguishable. Throughout the manuscript we suppose that Q is irreducible. Thus, 0 is a simple eigenvalue of Q with eigenvector U_0. The generator for the distinguishable case, denoted by $\mathcal{D}_N$, acts on a real function f on $[K]^N$ as follows

$$(\mathcal{D}_N f)(k_1, k_2, \dots, k_N) := \sum_{i=1}^{N} \sum_{k=1}^{K} \mu_{k_i,k}[f(k_1, \dots k_{i-1}, k, k_{i+1}, \dots, k_N) - f(k_1, \dots, k_N)],$$

for all $(k_1, k_2, \ldots, k_N) \in [K]^N$. If the function is given in a tensor product form, we get

$$\mathcal{D}_N(V_1 \otimes V_2 \otimes \cdots \otimes V_N) = \sum_{n=1}^{N} V_1 \otimes V_2 \otimes \cdots \otimes Q\, V_n \otimes \cdots \otimes V_N, \tag{2.27}$$

where $QV_n(k) := \sum_{r=1}^{K} \mu_{k,r} V_n(r) = \sum_{r=1}^{K} \mu_{k,r}(V_n(r) - V_n(k))$, for all $k \in [K]$.

Remark 2.3.1 ($\mathcal{D}_N$ as a Kronecker sum). In fact, the infinitesimal generator satisfies $\mathcal{D}_N = Q \oplus Q \oplus \cdots \oplus Q$, where $\oplus$ denotes the Kronecker sum. The well-known relationship between the exponential of a Kronecker sum and the Kronecker product of exponential matrices, namely:

$$\exp\{Q \oplus Q \oplus \cdots \oplus Q\} = \exp\{Q\} \otimes \exp\{Q\} \otimes \cdots \otimes \exp\{Q\},$$

makes clearer the idea that $\mathcal{D}_N$ is the infinitesimal generator of the system of N particles moving independently according to the infinitesimal generator Q. See [Pea65, Ch. XIV] and [Dav79, §2.2] for further details on the Kronecker sum.

The Markov chain generated by $\mathcal{D}_N$ is usually called *product chain*. The infinitesimal generator $\mathcal{D}_N$ inherits its spectral properties from those of Q. Namely, if π is the stationary distribution of Q, then $\pi \otimes \pi \otimes \cdots \otimes \pi$ is the stationary distribution of $\mathcal{D}_N$. Moreover, if $V_1, V_2, \ldots, V_N$ are N (not necessarily distinct) eigenvectors of Q, then $V_1 \otimes V_2 \otimes \cdots \otimes V_N$ is an eigenvector of $\mathcal{D}_N$. Consequently, if Q is diagonalisable, then $\mathcal{D}_N$ is also diagonalisable and the tensors products of vectors in an eigenbasis of Q form an eigenbasis of $\mathcal{D}_N$, as in Proposition 2.2.4-(a). In particular, if $\lambda_0 = 0, \lambda_1, \ldots, \lambda_{K-1}$ are the K complex eigenvalues of Q, then the eigenvalues of $\mathcal{D}_N$ are given by the sums of eigenvalues of Q, i.e. the spectrum of $\mathcal{D}_N$ is

$$\{z_0 + z_1 + \cdots + z_{K-1} : z_i \in \{\lambda_0, \lambda_1, \ldots, \lambda_{K-1}\}\}.$$

See Sections 12.4 and 20.4 in [LP17] for the proofs of these results and more details on product chains.

When the N particles are considered indistinguishable, the infinitesimal generator of the Markov chain, denoted by $\mathcal{Q}_N$, is that defined by (2.3), i.e.

$$(\mathcal{Q}_N f)(\eta) = \sum_{i,j \in [K]} \eta(i)\mu_{i,j} \left[f(\eta - \mathbf{e}_i + \mathbf{e}_j) - f(\eta)\right],$$

for all $\eta \in \mathcal{E}_{K,N}$ and for every function f on $\mathcal{E}_{K,N}$. Zhou and Lange [ZL09] noticed that $\mathcal{Q}_N$ is a lumped chain of $\mathcal{D}_N$ and used this fact to study the relationship between the spectral properties of both chains. They studied the eigenvalues and the left eigenfunctions of $\mathcal{Q}_N$. In particular, they proved that the stationary distribution of $\mathcal{Q}_N$ is multinomial with probability vector π, denoted $\mathcal{M}(\cdot \mid N, \pi)$, where π is the unique stationary probability of Q. Our approach differs from that on [ZL09]: we study the right eigenfunctions of $\mathcal{Q}_N$ using the connections between the real functions on $\mathcal{E}_{K,N}$ and the symmetric real functions on $[K]^N$ studied in Section 2.2. In addition, our methods allow us to explicitly describe the spectrum of $\mathcal{Q}_N$, for every mutation matrix Q generating an irreducible process, even when Q is non-diagonalisable. We first study the relationship between the generators $\mathcal{Q}_N$ and $\mathcal{D}_N$ through the operator $\Phi_{K,N}$.

Lemma 2.3.1 (Link between the generators $\mathcal{Q}_N$ and $\mathcal{D}_N$)**.** *For any symmetric function ξ on $[K]^N$, the function $\mathcal{D}_N\,\xi$ is also symmetric. In addition,*

$$\mathcal{Q}_N\,(\Phi_{K,N}\,\xi) = \Phi_{K,N}\,(\mathcal{D}_N\,\xi),$$

where $\Phi_{K,N}$ is defined by (2.18).

Proof. The symmetry of $\mathcal{D}_N\,\xi$ is a consequence of the symmetry of ξ and the linearity of $\mathcal{D}_N$.

For $\eta \in \mathcal{E}_{K,N}$ let us define $(k_1, k_2, \ldots, k_N) = \psi_{K,N}(\eta)$, i.e. k_i is the position on $[K]$ of the i-th particle according to the definition of $\psi_{K,N}$. We have

$$
\begin{aligned}
(\mathcal{D}_N\,\xi \circ \psi_{K,N})(\eta) &= \sum_{i=1}^{N}\sum_{k=1}^{K} \mu_{k_i,k}[\xi(k_1,\ldots,k_{i-1},k,k_{i+1},\ldots,k_N) - \xi(\psi_{K,N}(\eta))] \\
&= \sum_{k=1}^{K}\sum_{r=1}^{K}\sum_{i:\,k_i=r} \mu_{k_i,k}[\xi(k_1,\ldots,k_{i-1},k,k_{i+1},\ldots,k_N) - \xi(\psi_{K,N}(\eta))].
\end{aligned}
$$

Using the symmetry of ξ, for all η such that $\psi_{K,N}(\eta)(i) = r$ we obtain

$$
\xi(k_1,\ldots,k_{i-1},k,k_{i+1},\ldots,k_N) - \xi(\psi_{K,N}(\eta)) = \xi(\psi_{K,N}(\eta - \mathbf{e}_r + \mathbf{e}_k)) - \xi(\psi_{K,N}(\eta)).
$$

Thus,

$$
\begin{aligned}
(\mathcal{D}_N\,\xi \circ \psi_{K,N})(\eta) &= \sum_{k=1}^{K}\sum_{r=1}^{K}\sum_{i:\,k_i=r} \mu_{k_i,k}[\xi(\psi_{K,N}(\eta - \mathbf{e}_r + \mathbf{e}_k)) - \xi(\psi_{K,N}(\eta))] \\
&= \sum_{k=1}^{K}\sum_{r=1}^{K} \eta(r)\mu_{r,k}[\xi(\psi_{K,N}(\eta - \mathbf{e}_r + \mathbf{e}_k)) - \xi(\psi_{K,N}(\eta))] \\
&= (\mathcal{Q}_N\,\xi \circ \psi_{K,N})(\eta),
\end{aligned}
$$

for every $\eta \in \mathcal{E}_{K,N}$. $\square$

The following lemma describes all the eigenvalues of $\mathcal{Q}_N$, defined by (2.3), in the case where the mutation matrix is diagonalisable.

Lemma 2.3.2 (Eigenvalues of $\mathcal{Q}_N$ for diagonalisable Q)**.** *Assume Q is diagonalisable and*

$$
\mathcal{U} = \{U_0, U_1, \ldots, U_{K-1}\}
$$

is the basis of $\mathbb{R}^K$ formed by right eigenvectors of Q, such that U_0 is the all-one vector. Consider $\tilde{V}_\eta$ and λ_η defined as in (2.9) and (2.10), respectively. Then

(a) λ_η is an eigenvalue of $\mathcal{Q}_N$ with right eigenvector $\tilde{V}_\eta$.

(b) The spectrum of $\mathcal{Q}_N$ is formed by 0 and all λ_η for $\eta \in \bigcup_{L=1}^{N} \mathcal{E}_{K-1,L}$.

(c) $\mathcal{Q}_N$ is diagonalisable.

Proof. (a) For $\eta \in \mathcal{E}_{K-1,L}$ let us denote U_η as in (2.7). Because $Q\,U_0 = 0$ and $Q\,U_k = \lambda_k\,U_k$, $1 \le k \le K-1$, from (2.27), we get $\mathcal{D}_N(U_\eta) = \lambda_\eta U_\eta$. More generally, for every permutation $\sigma \in \mathcal{S}_N$, $\mathcal{D}_N\,(\sigma U_\eta) = \lambda_\eta(\sigma U_\eta)$, and thus, using the linearity of $\mathcal{D}_N$ we get

$$
\mathcal{D}_N\,V_\eta = \lambda_\eta V_\eta,
$$

where V_η is defined as in (2.8). Applying $\psi_{K,N}$ to both members of the previous equality we obtain $(\mathcal{D}_N V_\eta) \circ \psi_{K,N} = \lambda_\eta V_\eta \circ \psi_{K,N}$. Now, using Lemma (2.3.1), and the expressions (2.8) and (2.9), definitions of V_η and $\tilde{V}_\eta$, respectively, we obtain $\mathcal{Q}_N \tilde{V}_\eta = \lambda_\eta \tilde{V}_\eta$, which proves (a).

(b)-(c) Because $\mathcal{U}$ is a basis of $\mathbb{R}^K$, the set $\tilde{\mathcal{S}}^N$ as defined in (2.26) is a basis of $\mathbb{R}^{\mathcal{E}_{K,N}}$, due to Proposition 2.2.4-(c). Therefore, all the eigenvalues of $\mathcal{Q}_N$ are those described in part (b) and $\mathcal{Q}_N$ is diagonalisable. $\square$

Remark 2.3.2. Note that the results in Lemma 2.3.2 remains valid for all operator $\mathcal{Q}_N$ defined using a diagonalisable matrix Q, not necessarily a rate matrix, with complex entries and such that $QU_0 = \mathbf{0}$ and $\lambda_0 = 0$ has algebraic multiplicity equal to one.

Lemma 2.3.2 provides all the eigenvalues and right eigenvectors of $\mathcal{Q}_N$ when Q is diagonalisable. However, an ergodic rate matrix is not necessarily diagonalisable, as next example shows.

Example 2.3.1 (Non-diagonalisable rate matrix of an ergodic Markov chain)**.** Consider the infinitesimal rate matrix Q given by

$$Q = \begin{pmatrix} -9 & 7 & 2 \\ 1 & -7 & 6 \\ 5 & 7 & -12 \end{pmatrix} = W \begin{pmatrix} 0 & 0 & 0 \\ 0 & -14 & 1 \\ 0 & 0 & -14 \end{pmatrix} W^{-1},$$

where

$$W = \begin{pmatrix} 3/14 & 2 & 11/14 \\ 3/14 & -2 & -3/14 \\ 3/14 & 2 & -3/14 \end{pmatrix}, \quad W^{-1} = \begin{pmatrix} 1 & 7/3 & 4/3 \\ 0 & -1/4 & 1/4 \\ 1 & 0 & -1 \end{pmatrix}.$$

Then, Q is a non-diagonalisable rate matrix generating an ergodic Markov chain. Note that the unique stationary distribution of the process driven by Q is $(3/14, 1/2, 2/7)$.

Now, we want to extend the results in Lemma 2.3.2 to the case where the matrix Q is non-diagonalisable, as stated in Theorem 2.1.1. Let us first recall two known facts in the theory of real matrices. We denote by $M_n(\mathbb{R})$ and $M_n(\mathbb{C})$ the vector space of n-dimensional real and complex matrices, respectively. For a matrix $M \in M_n(\mathbb{C})$ we denote by $\mathrm{Spec}(M) \in \mathbb{C}^n$ its spectrum counting the algebraic multiplicities of the eigenvalues. It is known that the set of diagonalisable complex matrices is dense in $M_n(\mathbb{C})$. Serre [Ser10, Cor. 5.1], for instance, proves this result as a consequence of the Schur's Theorem. Using the same reasoning we can prove the following:

Fact 1 The set of diagonalisable complex matrices with each row summing to zero is dense in the set of the irreducible rate matrices: for every rate matrix $Q \in M_n(\mathbb{R})$ and $\epsilon > 0$ there exists a diagonalisable matrix $\bar{Q} \in M_n(\mathbb{C})$ such that $\|Q - \bar{Q}\| < \epsilon$. Moreover, $\bar{Q}$ can be chosen such that $0 \in \mathrm{Spec}(\bar{Q})$, with 0 having geometric multiplicity 1 and $\bar{Q}\, U_0 = \mathbf{0}$, where $\mathbf{0}$ denotes the K dimensional null column vector, i.e. each row of $\bar{Q}$ sums to zero.

The idea of the proof of Fact 1 is to modify diagonal elements in the upper-triangular matrix obtained by the Schur's Theorem [Ser10, Thm. 5.1] to get a matrix with n different eigenvalues, and thus diagonalisable. Indeed, since Q is an irreducible rate matrix, the eigenspace associated to the eigenvalue $\lambda_0 = 0$ has dimension one and it is generated by U_0. Moreover, the other $n-1$ complex eigenvalues have strictly negative real parts. Thus, it is possible to modify the diagonal of the upper triangular matrix obtained by the Schur's Theorem in such a way that the eigenvalues of the modified matrix, denoted $\bar{Q}$, are zero and $n-1$ complex numbers with different and strictly negative real parts. Furthermore, because of the Schur's factorisation, U_0 is also an eigenvector of $\bar{Q}$ associated to the null eigenvalue, i.e. $\bar{Q}U_0 = \mathbf{0}$.

Note that, since $M_n(\mathbb{C})$ is a finite dimensional vector space, the result in Fact 1 holds for every norm defined on $M_n(\mathbb{C})$. In the sequel we will use the uniform norm, denoted $\|\cdot\|_{\mathrm{Unif}}$, and defined as follows

$$\|A\|_{\mathrm{Unif}} := \max_{i,j} |a_{i,j}|,$$

for every matrix $A = (a_{i,j})_{i,j} \in M_n(\mathbb{C})$.

The second fact is related to the continuity of the eigenvalues of a matrix with respect to its entries. Consider the following distance between two sets of n elements in $\mathbb{C}$:

$$\mathrm{D}\left(\{z_i\}_{i=1}^n, \{\omega_i\}_{i=1}^n\right) := \inf_{\sigma\in\mathcal{S}_n} \max_j |z_j - \omega_{\sigma(j)}|,$$

where $\mathcal{S}_n$ denotes de symmetric group on $[n]$, for every $n \in \mathbb{N}$.

Fact 2 The eigenvalues are continuous with respect to the entries of the matrix in the following sense: consider $M \in M_n(\mathbb{C})$, then for all $\epsilon > 0$ there exists a $\delta > 0$ such that for every matrix $N \in M_n(\mathbb{C})$ such that $\|M - N\| < \delta$, then $\mathrm{D}\left(\mathrm{Spec}(M), \mathrm{Spec}(N)\right) < \epsilon$.

See e.g. [HM87] and [Ser10, Thm. 5.2] for a proof of Fact 2.

Proof of Theorem 2.1.1. From Lemma 2.3.2 we know that the statement of Theorem 2.1.1 holds for a diagonalisable rate matrix Q. Let us prove it in the general case using the Facts 1 and 2 we previously discussed.

For a mutation rate matrix $Q \in M_K(\mathbb{R})$ with spectrum $\mathrm{Spec}(Q) = \{0, \lambda_1, \dots, \lambda_{K-1}\}$, let us define by $\sigma_N(Q)$ the set formed by 0 and λ_η, for $\eta \in \bigcup_{L=1}^{K-1} \mathcal{E}_{K-1,L}$, where the values λ_k in the definition (2.10) of λ_η are those in $\mathrm{Spec}(Q)$. Then, proving Theorem 2.1.1-(a) is equivalent to prove that $\sigma_N(Q)$ is the spectrum of $\mathcal{Q}_N$, i.e. $\mathrm{D}\left(\mathrm{Spec}(\mathcal{Q}_N), \sigma_N(Q)\right) = 0$.

For a matrix $\bar{Q} \in M_K(\mathbb{C})$ whose rows sum to zero (not necessarily a rate matrix), let us define $\bar{\mathcal{Q}}_N$ similarly to the definition of $\mathcal{Q}_N$ (2.3), but with $\bar{Q}$ as mutation matrix instead of Q. As we commented in Remark 2.3.2, Lemma 2.3.2 remains valid and it ensures us that $\mathrm{Spec}(\bar{\mathcal{Q}}_N) = \sigma_N(\bar{Q})$. Thus, using the triangular inequality we get

$$\mathrm{D}\left(\mathrm{Spec}(\mathcal{Q}_N), \sigma_N(Q)\right) \leq \mathrm{D}\left(\mathrm{Spec}(\mathcal{Q}_N), \mathrm{Spec}(\bar{\mathcal{Q}}_N)\right) + \mathrm{D}\left(\mathrm{Spec}(\bar{\mathcal{Q}}_N), \sigma_N(Q)\right).$$

Moreover,

$$\begin{aligned} \|\mathcal{Q}_N - \bar{\mathcal{Q}}_N\|_{\mathrm{Unif}} &\leq N\|Q - \bar{Q}\|_{\mathrm{Unif}}, \\ \mathrm{D}\left(\mathrm{Spec}(\bar{\mathcal{Q}}_N), \sigma_N(Q)\right) &\leq N\,\mathrm{D}\left(\mathrm{Spec}(\bar{Q}), \mathrm{Spec}(Q)\right). \end{aligned}$$

Fix $\epsilon > 0$. Using Fact 2, we know there exist $\delta_1, \delta_2 > 0$ such that

$$\begin{aligned} \mathrm{D}\left(\mathrm{Spec}(\mathcal{Q}_N), \mathrm{Spec}(\bar{\mathcal{Q}}_N)\right) &\leq \frac{\epsilon}{2} \quad \text{if} \quad \|\mathcal{Q}_N - \bar{\mathcal{Q}}_N\|_{\mathrm{Unif}} < \delta_1, \\ \mathrm{D}\left(\mathrm{Spec}(\bar{Q}), \mathrm{Spec}(Q)\right) &\leq \frac{\epsilon}{2N} \quad \text{if} \quad \|Q - \bar{Q}\|_{\mathrm{Unif}} < \delta_2. \end{aligned}$$

Thus,

$$\mathrm{D}\left(\mathrm{Spec}(\mathcal{Q}_N), \sigma_N(Q)\right) \leq \frac{\epsilon}{2} + N\,\mathrm{D}(\mathrm{Spec}(\bar{Q}), \mathrm{Spec}(Q)) < \epsilon,$$

whenever $\|Q - \bar{Q}\|_{\mathrm{Unif}} < \min\{\delta_1/N, \delta_2\}$. Since ϵ can be taken arbitrary small, by Fact 1, the proof of (a) is finished.

The proof of (*b*) is exactly the same as the proof of (a) in Lemma 2.3.2. Note that, since $\eta(r) = \cdots = \eta(K-1) = 0$, the definition of $\tilde{V}_\eta$ only depends on the r linearly independent vectors forming $\mathcal{U}$. Finally, the result in (c) trivially comes from Lemma 2.3.2. ☐

Remark 2.3.3 (Alternative proof for Theorem 2.1.1). The *Jordan-Chevalley decomposition* is an elegant tool to find the eigenvalues of $\mathcal{Q}_N$ and prove Theorem 2.1.1. The Jordan-Chevalley decomposition ensures the existence of two matrices Q_{Diag} and Q_{Nil} such that $Q = Q_{\mathrm{Diag}} + Q_{\mathrm{Nil}}$. Moreover, Q_{Diag} is diagonalisable, Q_{Nil} is nilpotent, they commute and such a decomposition is unique. See [Ser10, Prop. 3.20] and [CEZ11] for more details about the Jordan-Chevalley decomposition. Then, it can be proved that the Jordan-Chevalley decomposition of $\mathcal{Q}_N$ is $\mathcal{Q}_N = (\mathcal{Q}_{\text{Diag}})_N + (\mathcal{Q}_{\text{Nil}})_N$, where $(\mathcal{Q}_{\text{Diag}})_N$ and $(\mathcal{Q}_{\text{Nil}})_N$ are defined similarly to $\mathcal{Q}_N$ in (2.3), substituting Q by Q_{Diag} and Q_{Nil}, respectively. Now, since the spectrum of $\mathcal{Q}_N$ is that of $(\mathcal{Q}_{\text{Diag}})_N$, the proof of Theorem 2.1.1 follows from Lemma 2.3.2.

2.3.2 Proof of Theorem 2.1.2

In this section, given $K \geq 2$ and $N \geq 2$, we consider the continuous-time Markov chain of N indistinguishable particles on K sites, with state space $\mathcal{E}_{K,N}$, where, with rate 1, any particle jumps to one of the positions of another particle chosen at random. We denote by $\mathcal{A}_N$ the infinitesimal generator of this reproduction process, which is defined in (2.4) as

$$(\mathcal{A}_N f)(\eta) = \sum_{i,j\in[K]} \eta(i)\eta(j)\left[f(\eta - \mathbf{e}_i + \mathbf{e}_j) - f(\eta)\right]$$

for every real function f and all $\eta \in \mathcal{E}_{K,N}$.

Remark 2.3.4 (First degree eigenfunctions of $\mathcal{A}_N$). Note that the states $\{N\,\mathbf{e}_k\}_{k=1}^{K} \subset \mathcal{E}_{K,N}$ are the only absorbing states for the interaction process generated by $\mathcal{A}_N$. Thus, the distribution concentrated at $N\,\mathbf{e}_k$, denoted $\delta_{\{N\mathbf{e}_k\}}$, is stationary for $\mathcal{A}_N$, for $k \in [K]$. It is not difficult to check that the real functions on $\mathcal{E}_{K,N}$, $x_0 \equiv 1$ and $x_k : \eta \mapsto \eta_k$, for $k \in [K-1]$, are linearly independent vectors of $\mathbb{R}^{\mathcal{E}_{K,N}}$ and they satisfy $\mathcal{A}_N x_k = 0$, for all $k = 0, 1, \ldots, K-1$. Thus, the right eigenspace associated to 0 is the space of homogeneous polynomials of degree 1, which has dimension K.

Actually, it can be proved that the generator $\mathcal{A}_N$ preserves the total degree of a polynomial, in the sense that the image of a polynomial is another polynomial of the same total degree. To prove Theorem 2.1.2 we first formally describe the preserving degree polynomial property satisfied by $\mathcal{A}_N$.

Lemma 2.3.3 ($\mathcal{A}_N$ preserves polynomial total degree)**.** *Assume $K \geq 2$ and $N \geq 2$. Let P be a polynomial on $\mathcal{E}_{K,N}$ of total degree L with $1 \leq L \leq N$. Then,*

$$\mathcal{A}_N V_P = -L(L-1)V_P + V_R,$$

where R is a polynomial with a total degree strictly less than L.

The proof of Lemma 2.3.3 is technical and it is deferred to Appendix 2.B. We proceed to prove Theorem 2.1.2.

Proof of Theorem 2.1.2. (a) For $K \geq 2$ and $N \geq 2$, let us define the sets $\mathcal{B}_L$ of monomials in $\mathcal{E}_{K,N}$ as follows

$$\mathcal{B}_0 := \{1\}, \quad \mathcal{B}_1 := \{x_1, x_2, \ldots, x_{K-1}\}, \quad \mathcal{B}_L := \{\mathbf{x}^\alpha,\ \alpha \in \mathcal{E}_{K-1,L}\},$$

for $2 \leq L \leq N$, where $\mathbf{x}^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_{K-1}^{\alpha_{K-1}}$ for $\alpha := (\alpha_1, \alpha_2, \ldots, \alpha_{K-1})$. Then, consider the ordered set

$$\mathcal{B} = \mathcal{B}_0 \cup \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_N.$$

The set $\mathcal{B}$ is a basis of the space of real functions on $\mathcal{E}_{K,N}$, due to Lemma 2.2.1-(b). The matrix similar to $\mathcal{A}_N$ with respect to this basis is $\bar{\mathcal{A}}_N = W^{-1}\mathcal{A}_N W$, where W is the matrix with P, with $P \in \mathcal{B}$, as column vectors. Thanks to the result in Lemma 2.3.3-(a), $\bar{\mathcal{A}}_N$ is a block upper triangular matrix, where the first diagonal block has size K and is a null matrix. The other diagonal blocks have size $\mathrm{Card}(\mathcal{E}_{K-1,L}) = \binom{K-2+L}{L}$ and are diagonal matrices with constant diagonal elements equal to $-L(L-1)$, with $2 \leq L \leq N$. This analysis gives us the eigenvalues of $\mathcal{A}_N$ are 0 with algebraic multiplicity K and $-L(L-1)$ with algebraic multiplicity $\binom{K-2+L}{L}$ for $2 \leq L \leq N$.

Now, using the block multiplication of matrices, it is not difficult to see that $(\bar{\mathcal{A}}_N)^n$ is also a block diagonal matrix, where the L-th block is a diagonal matrix of dimension $\binom{K-2+L}{L}$ with all the entries on the diagonal equal to $(-L(L-1))^n$, for $2 \leq L \leq N$. Thus, for every real polynomial Υ the matrix $\Upsilon(\bar{\mathcal{A}}_N) = W^{-1}\Upsilon(\mathcal{A}_N)W$ is a block diagonal matrix with diagonal elements $\Upsilon(-L(L-1))$. Taking

$$\Upsilon : s \mapsto s \prod_{L=2}^{N} [s + L(L-1)],$$

we get $\Upsilon(\bar{\mathcal{A}}_N) = \mathbf{0}_{K,N}$, where $\mathbf{0}_{K,N}$ is the $\binom{K-1+N}{N}$ dimensional null matrix. Thus, $\Upsilon(\mathcal{A}_N) = \mathbf{0}_{K,N}$ and Υ is necessarily the *minimal polynomial* of $\mathcal{A}_N$, which factors into distinct linear factors. We thus conclude that $\mathcal{A}_N$ is diagonalisable. □

Remark 2.3.5 (On the right eigenfunctions of $\mathcal{A}_N$). Theorem 2.1.2 does not provide a characterisation of the eigenspace associated to the eigenvalue $-L(L-1)$, for $L \in [N]$. For the special case $K = 2$, Watterson [Wat61] does provide such a decomposition for the discrete analogue of $\mathcal{A}_N$ in terms of cumulative sums of discrete Chebyshev polynomials. In addition, Zhou [Zho08, §4.2.2] provides an equivalent but simpler expression for the eigenvectors of the equivalent analogous of $\mathcal{A}_N$, for $K = 2$, in terms of univariate Hahn polynomials.

In the general case, it is possible to describe the eigenspaces associated to the first three eigenvalues of $\mathcal{A}_N$. As we commented in Remark 2.3.4, the right eigenspace associated to 0 is the space of homogeneous polynomials of first degree. Moreover, the right eigenspace associated to -2 has dimension $K(K-1)/2$ and it is generated by the set of monomials $\{x_k x_r,\ 1 \leq k < r \leq K\}$. Additionally, for $L = 3$, it is possible to prove that a simple basis of the right eigenspace associated to -6 has dimension $K(K+1)(K-1)/6$ and is given by eigenvectors $\{x_k^2 x_r - x_k x_r^2,\ 1 \leq k < r \leq K\} \cup \{x_k x_r x_s,\ 1 \leq k < r < s \leq K\}$. The complete characterisation of the eigenvectors of $\mathcal{A}_N$, for $K \geq 3$, is a topic of further research.

2.3.3 Proof of Theorem 2.1.3

This section is devoted to the proof of Theorem 2.1.3 providing a description of the spectrum of the neutral multi-allelic Moran process with generator $\mathcal{Q}_{N,p}$, defined by (2.2) as

$$(\mathcal{Q}_{N,p} f)(\eta) = \sum_{i,j \in [K]} \eta(i) \left(\mu_{i,j} + \frac{p}{N} \eta(j) \right) [f(\eta - \mathbf{e}_i + \mathbf{e}_j) - f(\eta)],$$

for every real function f in $\mathcal{E}_{K,N}$ and every $\eta \in \mathcal{E}_{K,N}$.

Assume $K \geq 2$, $N \geq 2$ and $p \in [0, \infty)$ and suppose that Q is diagonalisable with eigenvalues 0 and λ_k, for $k \in [K-1]$. For any $\eta \in \mathcal{E}_{K-1,L}$, with $L \in [N]$, let us recall the definition of $\lambda_{\eta,p}$:

$$\lambda_{\eta,p} = -L(L-1)\frac{p}{N} + \sum_{k=1}^{K-1} \eta(k)\lambda_k. \tag{2.28}$$

Then, we will prove that the eigenvalues of $\mathcal{Q}_{N,p}$ are 0 and all $\lambda_{\eta,p}$ for $\eta \in \bigcup_{L=1}^{N} \mathcal{E}_{K-1,L}$.

Proof of Theorem 2.1.3. It is straightforward to remark that

$$\mathcal{Q}_{N,p} = \mathcal{Q}_N + \frac{p}{N}\mathcal{A}_N,$$

where $\mathcal{Q}_N$ and $\mathcal{A}_N$ are the generators of the mutation and the reproduction processes defined by (2.3) and (2.4), respectively.

Let us first prove the result when the mutation rate matrix Q is diagonalisable. As proved in Lemma 2.3.2, the vector $\tilde{V}_\eta$ is an eigenvector of $\mathcal{Q}_N$ with eigenvalue λ_η, for $\eta \in \bigcup_{L=1}^N \mathcal{E}_{K-1,L}$. Let us denote by $\tilde{V}_0$ the all-one vector in $\mathbb{R}^{\mathcal{E}_{K,N}}$. Then, the set $\mathcal{B} = \{V_0\} \cup \{\tilde{V}_\eta,\ \eta \in \bigcup_{L=1}^N \mathcal{E}_{K-1,L}\}$ is a basis of $\mathbb{R}^{\mathcal{E}_{K,N}}$, thanks to Proposition 2.2.4-(c). Let us denote by W the matrix with the elements of $\mathcal{B}$ as columns such that $W^{-1}\mathcal{Q}_N W$ is a diagonal matrix with diagonal entries equal to 0 and λ_η, for $\eta \in \bigcup_{L=1}^N \mathcal{E}_{K-1,L}$.

For $1 \leq L \leq N$ and $\eta \in \mathcal{E}_{K-1,L}$, the expression (2.9) and Lemma 2.2.3-(c) ensure that $\tilde{V}_\eta$ is a polynomial of total degree equal to L. Using now Theorem 2.1.2-(b), we get

$$\mathcal{A}_N \tilde{V}_\eta = -L(L-1)\tilde{V}_\eta + R,$$

where R is a polynomial of total degree strictly less than L. This fact means that, like in Theorem 2.1.2-(c), $W^{-1}\mathcal{A}_N W$ is a block upper triangular matrix, where the diagonal blocks of size $\mathrm{Card}(\mathcal{E}_{K-1,L}) = \binom{K+L-2}{L}$ are diagonal matrices with constant diagonal elements equal to $-L(L-1)$, for $2 \leq L \leq N$. The first diagonal block of size K is a null matrix. It follows that the matrix

$$W^{-1}\mathcal{Q}_{N,p}W = W^{-1}\mathcal{Q}_N W + \frac{p}{N}W^{-1}\mathcal{A}_N W$$

is a block upper triangular matrix, where the first diagonal block has dimension one and is null, i.e. the first column is null. Moreover, the L-th diagonal block has dimension $\binom{K+L-2}{L}$ and its diagonal elements are the eigenvalues $\lambda_{\eta,p}$ with $\eta \in \mathcal{E}_{K-1,L}$, for $L \in [N]$. Thus, these are the eigenvalues of $\mathcal{Q}_{N,p}$.

Now, consider a general mutation matrix $Q \in M_K(\mathbb{R})$, not necessarily diagonalisable, with spectrum $\mathrm{Spec}(Q) = \{0, \lambda_1, \dots, \lambda_{K-1}\}$. Let us define by $\sigma_{N,p}(Q)$ the set formed by 0 and $\lambda_{\eta,p}$, for $\eta \in \bigcup_{L=1}^{K-1} \mathcal{E}_{K-1,L}$, where the values λ_k in the definition (2.28) of $\lambda_{\eta,p}$, are those in $\mathrm{Spec}(Q)$. Define $\bar{\mathcal{Q}}_{N,p}$ similarly to (2.2) but with a diagonalisable matrix $\bar{Q} \in M_K(\mathbb{C})$, whose rows have null sum (not necessarily a rate matrix), instead of Q. Then,

$$\|\mathcal{Q}_{N,p} - \bar{\mathcal{Q}}_{N,p}\|_{\mathrm{Unif}} = \|\mathcal{Q}_N - \bar{\mathcal{Q}}_N\|_{\mathrm{Unif}},$$
$$\mathrm{D}\left(\mathrm{Spec}(\bar{\mathcal{Q}}_{N,p}), \sigma_{N,p}(Q)\right) = \mathrm{D}\left(\mathrm{Spec}(\bar{\mathcal{Q}}_{K,N}), \sigma_N(Q)\right).$$

Hence, $\sigma_{N,p}(Q)$ is proved to be the spectrum of $\mathcal{Q}_{N,p}$, analogously to the proof of Theorem 2.1.1-(a).

□

Remark 2.3.6 (Alternative proof of Theorem 2.1.3). Another proof of Theorem 2.1.3 can be carried out using the Jordan form of the mutation rate matrix Q. Indeed, the vectors $\tilde{V}_\eta \in \mathbb{R}^{\mathcal{E}_{K,N}}$ can be defined using the basis of $\mathbb{R}^K$ that transforms Q in its normal Jordan form. Then, defining a suitable order among the vectors $\tilde{V}_\eta$, for $\eta \in \bigcup_{L=1}^N \mathcal{E}_{K-1,L}$, it is possible to show that $\mathcal{Q}_{N,p}$ is similar to an upper triangular matrix with the values $\lambda_{\eta,p}$ on the diagonal.

2.4 Applications to the ergodicity of the Moran process

This section devoted to some applications of the results in Section 2.3 to the study of the ergodicity of the process driven by $\mathcal{Q}_{N,p}$ in total variation, using spectral properties of Q. In this section we prove Corollary 2.1.4 and Theorem 2.1.5.

Next result establishes that the Jordan form of Q is a diagonal block in the Jordan form of $\mathcal{Q}_{N,p}$.

Corollary 2.4.1 (Jordan forms of Q and $\mathcal{Q}_{N,p}$)**.** *Consider $K \geq 2$, $N \geq 2$ and $p \geq 0$. If J is the Jordan form of Q, then the Jordan normal form of $\mathcal{Q}_{N,p}$ is $J \oplus J'$, where J' is a Jordan matrix of dimension $\binom{K-1+N}{N} - K$.*

In particular, Q and $\mathcal{Q}_{N,p}$ have that same SLEM.

Proof. The image by $\mathcal{Q}_{N,p}$ of a first degree polynomial is also a first degree polynomial, i.e. the space of first degree polynomials is invariant by $\mathcal{Q}_{N,p}$. Moreover, as a consequence of Lemma 2.3.1 we obtain

$$\mathcal{Q}_{N,p}\,\tilde{\xi}(V) = \mathcal{Q}_N\,\tilde{\xi}(V) = \Phi_{K,N}\,\mathcal{D}_N\xi(V) = \Phi_{K,N}\,\xi(QV) = \tilde{\xi}(QV).$$

Let $\mathcal{U} = \{U_0, \ldots, U_{K-1}\}$ by a Jordan basis of Q formed by generalised eigenvectors of Q. Since $\mathcal{Q}_{N,p}\,\tilde{\xi}(U_k) = \tilde{\xi}(Q\,U_k)$, for every $k \in [K-1]_0$, we have that $\{\tilde{\xi}(U_0), \ldots, \tilde{\xi}(U_{K-1})\}$ is a system of linearly independent generalised eigenvectors of $Q_{N,p}$. They are precisely the generalised eigenvectors of $\mathcal{Q}_{N,p}$ associated to the eigenvalues in $\mathrm{Spec}(Q) \subset \mathrm{Spec}(\mathcal{Q}_{N,p})$. We can complete this system to a Jordan basis of $\mathbb{R}^{\mathcal{E}_{K,N}}$, adding the generalised eigenvectors of the other eigenvalues on $\mathcal{Q}_{N,p}$. With respect to this Jordan basis $\mathcal{Q}_{N,p}$ becomes similar to $J \oplus J'$, where J is the Jordan matrix of Q and J' is a Jordan matrix of dimension $\binom{K-1+N}{N} - K$.

Note that the eigenvalues $\{\lambda_0, \lambda_1, \ldots, \lambda_{K-1}\}$ are those eigenvalues of $\mathcal{Q}_{N,p}$ of smallest modulus. We thus get that Q and $\mathcal{Q}_{N,p}$ have the same SLEM. □

Every irreducible finite Markov chain convergences exponentially to stationarity, see e.g. [LP17, Thm. 4.9]. In addition, the asymptotic speed of convergence is associated to the SLEM and the size of the largest Jordan block corresponding to any eigenvalue with this modulus. We recall that the size of the largest Jordan block associated to an eigenvalue λ is equal to the multiplicity of λ in the minimal polynomial of the rate matrix of the Markov chain.

Proof of Corollary 2.1.4. Let ρ be the SLEM of Q and s the largest multiplicity in the minimal polynomial of Q of all the eigenvalues with modulus ρ, or equivalently, the size of the largest Jordan block associated to eigenvalues with modulus ρ. Then,

$$\mathrm{D}^{\mathrm{TV}}_{Q}(t) = \mathcal{O}(t^{s-1}\mathrm{e}^{-\rho t}), \tag{2.29}$$

see e.g. [SRW15, Thm. 3.2]. The following result is a consequence of Corollary (2.29) and 2.4.1. □

The following example uses Corollary 2.1.4 to provide the rates for the exponential convergence to stationarity of the neutral multi-allelic Moran (Fleming−Viot particle) process considered in [Cor21a].

Example 2.4.1 (Circulant mutation rate matrix)**.** Consider the following mutation rate matrix

$$Q_\theta = \begin{pmatrix} -(1+\theta) & 1 & 0 & \ldots & 0 & \theta \\ \theta & -(1+\theta) & 1 & \ldots & 0 & 0 \\ 0 & \theta & -(1+\theta) & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 0 & 0 & \ldots & \theta & -(1+\theta) \end{pmatrix},$$

where $\theta \geq 0$. Q_θ is the infinitesimal generator of a simple asymmetric random walk on the K-cycle graph. The neutral multi-allelic Moran type process with mutation rate Q_θ was considered in [Cor21a]. Since Q_θ is circulant, it is possible to explicitly diagonalise it using the Fourier matrix. The eigenvalues of Q_θ are

$$\lambda_k = -2(1+\theta)\sin^2\left(\frac{\pi k}{K}\right) + \mathrm{i}(1-\theta)\sin\left(\frac{2\pi k}{K}\right),$$

for $0 \le k \le K-1$. Thus, the SLEM of Q_θ is $2(1+\theta)\sin^2\left(\frac{\pi}{K}\right)$, which is attained for two eigenvalues, each one of them with algebraic multiplicity equals to 1, for $\theta \neq 1$. When $\theta = 1$, the SLEM of Q_θ is $\lambda_1 = 4\sin^2\left(\frac{\pi}{K}\right)$ and it is attained for a unique eigenvalue with algebraic and geometric multiplicities equal to 2. Let $\mathcal{Q}_\theta$ be the infinitesimal generator of the neutral multi-allelic Moran process with mutation rate Q_θ. Then,

$$\mathrm{D}^{\mathrm{TV}}_{\mathcal{Q}_\theta}(t) = \mathcal{O}\left(\mathrm{e}^{-2(1+\theta)\sin^2\left(\frac{\pi}{K}\right)t}\right).$$

Example 2.4.2 (Convergence rate for a process with non-diagonalisable mutation rate matrix)**.** Consider Q as in Example 2.3.1 and $\mathcal{Q}_{N,p}$ the infinitesimal generator of the associated neutral multi-allelic Moran process with mutation rate matrix Q. Then, $\lambda_0 = 0$ and $\lambda_1 = \lambda_2 = -14$, because -14 has algebraic multiplicity 2. Then, for N fixed, the eigenvalues of $\mathcal{Q}_{N,p}$ are

$$\lambda_{L,p} := \eta(1)\lambda_1 + \eta(2)\lambda_2 - L(L-1)\frac{p}{N} = -14L - L(L-1)\frac{p}{N},$$

for $L \in [N]_0$. In addition, λ_L has algebraic multiplicity $\mathrm{Card}(\mathcal{E}_{2,L}) = L+1$.

Note that the minimal polynomial of Q is $m_Q : s \mapsto s\,(s+14)^2$ and according to the notation in Corollary 2.1.4 we get $\rho = 14$ and $s = 2$. Then,

$$\mathrm{D}^{\mathrm{TV}}_{\mathcal{Q}_{N,p}}(t) = \mathcal{O}\left(t\,\mathrm{e}^{-14t}\right).$$

Furthermore, according to Theorem 2.1.5 we get

$$\mathrm{D}^{\mathrm{TV}}_{\mathcal{Q}_{N,p}}\left(\frac{\ln N - c}{28}\right) \ge 1 - 416\mathrm{e}^{-c},$$

for all $c \ge 0$.

2.4.1 Proof of Theorem 2.1.5

First, let us denote by $\Gamma_{\mathcal{L}}$ the *"carré-du-champ"* operator associated to the Markov generator $\mathcal{L}$ on a state space $\mathcal{E}$, i.e.

$$\Gamma_{\mathcal{L}} f : \eta \mapsto (\mathcal{L}f^2)(\eta) - 2f(\eta)(\mathcal{L}f)(\eta),$$

for all $\eta \in \mathcal{E}$.

The *"carré-du-champ"* operator is associated to the evolution in time of the variance of the test function. Indeed,

$$\mathrm{Var}_\eta(f(\eta_t)) = \int_0^t \mathrm{e}^{s\mathcal{L}}\left(\Gamma_{\mathcal{L}}(\mathrm{e}^{(t-s)\mathcal{L}} f)\right)(\eta)\mathrm{d}s,$$

where $(\mathrm{e}^{t\mathcal{L}})_{t\ge 0}$ denotes the semigroup generated by $\mathcal{L}$. See, for example, [CT16b, p. 695].

Proof of Theorem 2.1.5. Our method of proof is based on a Wilson's method (cf. [LP17, Thm. 13.28]). Let us denote $V = [v_1, v_2, \ldots, v_K]$ a real right-eigenvector satisfying $QV = -\lambda V$. Then, using Theorem 2.1.3 and Lemma 2.2.3 (specifically equations (2.23) and (2.24)) we get that $\tilde{\xi}(V)$ and $\tilde{\xi}(V,V)$ are right-eigenfunctions of $\mathcal{Q}_{N,p}$ satisfying

$$
\begin{aligned}
(\mathrm{e}^{t\mathcal{Q}_{N,p}}\tilde{\xi}(V))(\eta) &= \mathrm{e}^{-t\lambda}\tilde{\xi}(V)(\eta)\\
(\mathrm{e}^{t\mathcal{Q}_{N,p}}\tilde{\xi}(V,V))(\eta) &= \mathrm{e}^{-2(\lambda+p/N)t}\tilde{\xi}(V,V)(\eta),
\end{aligned}
$$

for every $\eta \in \mathcal{E}_{K,N}$. We recall that from (2.24) we have

$$
\tilde{\xi}(V,V) = \tilde{\xi}(V)^2 - \tilde{\xi}(V \odot V).
$$

where $V \odot V = [v_1^2, \dots, v_K^2]$ is the componentwise square vector of V.

Thereafter, using $\tilde{\xi}(V)$ as a test function we get

$$
\mathrm{d}^{\mathrm{TV}}(\delta_{N\mathbf{e}_k}\mathrm{e}^{t_{N,c}\mathcal{Q}_{N,p}}, \nu_{N,p}) \geq \mathbb{P}_{N\mathbf{e}_k}\left[\tilde{\xi}(V)(\eta_t) \geq \mu_t/2\right] - \mathbb{P}_{\nu_{N,p}}\left[\tilde{\xi}(V)(\eta_\infty) \geq \mu_t/2\right], \tag{2.30}
$$

where $\mu_t = \mathbb{E}_{N\mathbf{e}_k}\left[\tilde{\xi}(V)(\eta_t)\right] = \mathrm{e}^{-t\lambda}Nv_k$. By Markov's and Chebyshev's inequalities we have that

$$
\begin{aligned}
\mathbb{P}_{\nu_{N,p}}\left[\tilde{\xi}(V)(\eta_\infty) \geq \frac{\mu_t}{2}\right] &\leq 4\mathrm{e}^{2t\lambda}\frac{\mathrm{Var}_{\nu_{N,p}}\left[\tilde{\xi}(V)(\eta_\infty)\right]}{|v_k|^2N^2},\\
\mathbb{P}_{N\mathbf{e}_k}\left[\tilde{\xi}(V)(\eta_t) \geq \frac{\mu_t}{2}\right] &\geq 1 - 4\mathrm{e}^{2t\lambda}\frac{\mathrm{Var}_{N\mathbf{e}_k}\left[\tilde{\xi}(V)(\eta_t)\right]}{|v_k|^2N^2}.
\end{aligned}
$$

Thus, plugging these last expressions into (2.30) we get

$$
\mathrm{d}^{\mathrm{TV}}(\delta_{N\mathbf{e}_k}\mathrm{e}^{t_{N,c}\mathcal{Q}_{N,p}}, \nu_{N,p}) \geq 1 - 8\frac{\mathrm{e}^{2\lambda t}}{|v_k|^2N}\sup_{t\geq 0}\frac{\mathrm{Var}_{N\mathbf{e}_k}\left[\tilde{\xi}(V)(\eta_t)\right]}{N}.
$$

We are interested in finding a lower bound for $\mathrm{D}^{\mathrm{TV}}_{\mathcal{Q}_N}$ at time $(\ln(N) - c)/2\lambda$. It remains to prove a bound for the last factor in the previous expression.

Note that

$$
\begin{aligned}
\Gamma_{\mathcal{Q}_{N,p}}\tilde{\xi}(V) &= \mathcal{Q}_{N,p}(\tilde{\xi}(V)^2) - 2\tilde{\xi}(V)\mathcal{Q}_{N,p}(\tilde{\xi}(V))\\
&= \mathcal{Q}_{N,p}(\tilde{\xi}(V,V)) + \mathcal{Q}_{N,p}\tilde{\xi}(V \odot V) - 2\tilde{\xi}(V)\mathcal{Q}_{N,p}(\tilde{\xi}(V))\\
&= -2\left(\lambda + \frac{p}{N}\right)\tilde{\xi}(V,V) + 2\lambda\tilde{\xi}(V)^2 + \tilde{\xi}(Q(V \odot V))\\
&= -2\frac{p}{N}\tilde{\xi}(V,V) + 2\lambda\tilde{\xi}(V \odot V) + \tilde{\xi}(Q(V \odot V)).
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\frac{\mathrm{Var}_{N\mathbf{e}_k}\left(\tilde{\xi}(V)(\eta_t)\right)}{N} &= \frac{1}{N}\int_0^t \mathrm{e}^{s\mathcal{Q}_{N,p}}\left(\Gamma_{\mathcal{Q}_{N,p}}(\mathrm{e}^{(t-s)\mathcal{Q}_{N,p}}\tilde{\xi}(V))\right)(N\mathbf{e}_k)\mathrm{d}s\\
&= \frac{1}{N}\int_0^t \mathrm{e}^{-\lambda(t-s)}\mathrm{e}^{s\mathcal{Q}_{N,p}}\left(\Gamma_{\mathcal{Q}_{N,p}}\tilde{\xi}(V)\right)(N\mathbf{e}_k)\mathrm{d}s\\
&= \frac{1}{N}\int_0^t \mathrm{e}^{-\lambda(t-s)}\mathrm{e}^{s\mathcal{Q}_{N,p}}\left(-2\frac{p}{N}\tilde{\xi}(V,V) + 2\lambda\tilde{\xi}(V \odot V) + \tilde{\xi}(Q(V \odot V))\right)(N\mathbf{e}_k)\mathrm{d}s.
\end{aligned}
$$

Note that

$$
\frac{1}{N}\int_0^t \mathrm{e}^{-\lambda(t-s)}\mathrm{e}^{s\mathcal{Q}_{N,p}}\left(-2\frac{p}{N}\tilde{\xi}(V,V)\right)(N\mathbf{e}_k)\mathrm{d}s = -2\left(1 - \frac{1}{N}\right)p\,v_k^2\,\mathrm{e}^{-\lambda t}\int_0^t \mathrm{e}^{-(\lambda+2p/N)s}\mathrm{d}s \leq 0.
$$

Then,

$$\begin{aligned}\frac{\mathrm{Var}_{N\mathbf{e}_k}\left(\tilde{\xi}(V)(\eta_t)\right)}{N} &\leq \frac{1}{N}\int_0^t \mathrm{e}^{-\lambda(t-s)}\mathrm{e}^{s\mathcal{Q}_{N,p}}\left(2\lambda\tilde{\xi}(V\odot V)+\tilde{\xi}(Q(V\odot V))\right)(\eta^\star)\mathrm{d}s\\ &\leq 2\lambda\left\|\frac{\tilde{\xi}(V\odot V)}{N}\right\|_\infty+\left\|\frac{\tilde{\xi}(Q(V\odot V))}{N}\right\|_\infty\\ &\leq (2\lambda+\|Q\|_\infty)\|V\|_\infty.\end{aligned}$$

We obtain the desired inequality.

The lower bound for $\mathrm{D}^{\mathrm{TV}}_{\mathcal{Q}_{N,p}}$ is obtained considering the initial distribution concentrated at $N\mathbf{e}_{k^\star}$, where $k^\star$ satisfies $|v_{k^\star}|=\|V\|_\infty$.

□

2.5 Neutral multi-allelic Moran type process with parent independent mutation

In this section we discuss some applications of the Theorem 2.1.3 and its consequences to the neutral multi-allelic Moran model with parent independent mutation scheme. We will use some well-known results on finite state reversible Markov chains and their convergence to stationarity. We refer the interested reader to [Sal97], [Bré20], [DP17] and [LP17]), for further details. We will focus on the case where the Moran process has *parent independent mutation* [Eth11]. In this case, the Moran process is reversible. In fact, as we claimed in Lemma 2.1.7, the neutral Moran process with $p>0$ is reversible if and only if its mutation matrix satisfies the *parent independent* condition. We explicitly diagonalise the infinitesimal generator of the neutral multi-allelic Moran process with parent independent mutation rate using the multivariate Hahn and Krawtchouk polynomials, which allows us to provide an explicit expression for the transition function of this process. Using these results we prove Theorems 2.1.8 and 2.1.9.

2.5.1 Proof of Theorems 2.1.8 and 2.1.9

Let us recall that the generator of the neutral multi-allelic Moran process with parent independent mutation defined by (2.5.1), which acts on a real function f on $\mathcal{E}_{K,N}$ as follows

$$(\mathcal{L}_{N,p}f)(\eta) := \sum_{i,j=1}^{K}\eta(i)\left[f(\eta-\mathbf{e}_i+\mathbf{e}_j)-f(\eta)\right]\left(\mu_j+p\frac{\eta(j)}{N}\right),$$

for all $\eta\in\mathcal{E}_{K,N}$. We next prove Corollary 2.1.6, which provides the spectrum of $\mathcal{L}_{N,p}$, for all $p\geq 0$.

Proof of Corollary 2.1.6. Since $Q_{\boldsymbol{\mu}}$ is an infinitesimal matrix, zero is one of its eigenvalues with right eigenfunction f_1, the K-dimensional all-one vector. Note that $\boldsymbol{\pi}:=\boldsymbol{\mu}/|\boldsymbol{\mu}|=(\mu_1/|\boldsymbol{\mu}|,\ldots,\mu_K/|\boldsymbol{\mu}|)$ is the unique stationary distribution of $Q_{\boldsymbol{\mu}}$, which is also reversible. Moreover, note that $Q_{\boldsymbol{\mu}}f=-|\boldsymbol{\mu}|(f+\langle f,\boldsymbol{\pi}\rangle)$, for every $f\in\mathbb{R}^{\mathcal{E}_{K,N}}$. Thus, every function f satisfying $\langle f,\boldsymbol{\pi}\rangle=0$ is a right eigenfunction of $Q_{\boldsymbol{\mu}}$, i.e. the eigenspace associated to $-|\boldsymbol{\mu}|$ is the space of the orthogonal functions to $\boldsymbol{\pi}$, which has dimension $K-1$. The expression (2.12) for the eigenvalues comes from Theorem 2.1.3. Since the process is reversible we obtain that $\mathcal{L}_{N,p}$ is diagonalisable. The spectral gap is obtained for $L=1$.

□

Multivariate orthogonal Hahn and Krawtchouk polynomials

The rest of the section is devoted to the characterisation of the eigenfunctions of $\mathcal{L}_{N,p}$ and the proof of Theorem 2.1.8. Let us establish some notation that will be useful in the sequel to study the eigenfunctions of $\mathcal{L}_{N,p}$. For a K-dimensional real vector $\mathbf{x}$ we define the following quantities:

$$|\mathbf{x}_i| := \sum_{j=1}^{i} x_j, \quad |\mathbf{x}^i| := \sum_{j=i}^{K} x_j.$$

We set by convention $|\mathbf{x}^i| := 0$, for all $i > K$.

The orthogonal polynomials we define below are indexed by the set $\bigcup_{L=0}^{N} \mathcal{E}_{K-1,L}$, where $\mathcal{E}_{K-1,0} = \{\mathbf{0}\}$ is the set formed by the $K-1$ dimensional null vector. We define the *multivariate Hahn polynomials* on $\mathcal{E}_{K,N}$, indexed by $\eta \in \mathcal{E}_{K-1,L}$, for $L \in [N]_0$, and denoted $H_\eta(\mathbf{x}; N, \boldsymbol{\alpha})$, as follows

$$H_\eta(\mathbf{x}; N, \boldsymbol{\alpha}) := \frac{1}{(N)_{[|\eta|]}} \prod_{k=1}^{K-1} (-N + |\mathbf{x}_{k-1}| + |\eta^{k+1}|)_{(\eta(k))} H_{\eta(k)}(x_k; M_k, \alpha_k, \gamma_k) \tag{2.31}$$

where $M_k = N - |\mathbf{x}_{k-1}| - |\eta^{k+1}|$, $\gamma_k = |\boldsymbol{\alpha}^{k+1}| + 2|\eta^{k+1}|$ and $H_n(x; M, \beta, \gamma)$ is the *univariate Hahn polynomial* defined by

$$\begin{aligned} H_n(x; M, \beta, \gamma) &:= {}_3F_2\left(\begin{matrix} -n, & n+\beta+\gamma-1, -x \\ \beta, & -M \end{matrix} \middle| 1 \right) \\ &= \sum_{j=0}^{n} \frac{(-n)_{(j)}(n+\beta+\gamma-1)_{(j)}(-x)_{(j)}}{\beta_{(j)}(-M)_{(j)}} \frac{1}{j!}. \end{aligned} \tag{2.32}$$

Note that for $\mathbf{0} \in \mathcal{E}_{K-1,0}$ we obtain $H_{\mathbf{0}}(\cdot\,; N, \boldsymbol{\alpha}) \equiv 1$. In addition, it is no difficult to check that $H_\eta(N\mathbf{e}_K; N, \boldsymbol{\alpha}) \equiv 1$, for all $\eta \in \bigcup\limits_{L=0}^{N} \mathcal{E}_{K-1,L}$.

We also define the *multivariate Krawtchouk polynomials* on $\mathcal{E}_{K,N}$ denoted $K_\eta(\mathbf{x}; N, \mathbf{q})$, indexed by $\eta \in \bigcup\limits_{L=0}^{N} \mathcal{E}_{K-1,L}$, with $\mathbf{q} \in (0,1)^K$ such that $|\mathbf{q}| = 1$, as the multivariate polynomials satisfying:

$$K_\eta(\mathbf{x}; N, \mathbf{q}) := \frac{1}{(N)_{[|\eta|]}} \prod_{k=1}^{K-1} (-N + |\mathbf{x}_{k-1}| + |\eta^{k+1}|)_{(\eta(k))} K_{\eta(k)}\left(x_k; M_k, \frac{q_k}{|\boldsymbol{q}^k|}\right) \tag{2.33}$$

where $M_k = N - |\mathbf{x}_{k-1}| - |\eta^{k+1}|$, and $K_n(x; N, q)$ is the *univariate Krawtchouk polynomial* defined by

$$\begin{aligned} K_n(x; N, q) &:= {}_2F_1\left(\begin{matrix} -n, & -x \\ & -N \end{matrix} \middle| \frac{1}{q} \right) \\ &= \sum_{j=0}^{n} \frac{(-n)_{(j)}(-x)_{(j)}}{(-N)_{(j)}} \frac{1}{j! q^j}. \end{aligned} \tag{2.34}$$

In addition, $K_{\mathbf{0}}(\cdot\,; N, \mathbf{q}) \equiv 1$, for $\mathbf{0} \in \mathcal{E}_{K-1,0}$, and $K_\eta(N\mathbf{e}_K; N, \mathbf{q}) \equiv 1$, for all $\eta \in \bigcup\limits_{L=0}^{N} \mathcal{E}_{K-1,L}$.

See [Ism05, Ch. 6] and [KLS10, Ch. 9] for more details about the univariate Hahn and Krawtchouk polynomials. We define the univariate Hahn and Krawtchouk polynomials in (2.32) and (2.34), respectively, using the hypergeometric functions notation which could be very useful for algebraic manipulations (cf. [KLS10, Ch. 10]). For instance, consider $\boldsymbol{\alpha} = N\boldsymbol{\mu}/p$ in the definition of Hahn polynomials, then

$$\begin{aligned}
\lim_{p\to 0^+} H_{\eta(k)}(x_k; M_k, \alpha_k, |\boldsymbol{\alpha}^{k+1}| + 2|\eta^{k+1}|) &= \lim_{p\to 0^+} H_{\eta(k)}\left(x_k; M_k, \frac{N\mu_k}{p}, \frac{N|\boldsymbol{\mu}^{k+1}|}{p} + 2|\eta^{k+1}|\right) \\
&= \lim_{p\to 0^+} {}_3F_2\left(\begin{matrix} -\eta(k), & \eta(k) + N\mu_k/p + N|\boldsymbol{\mu}^{k+1}|/p + 2|\eta^{k+1}| - 1, -x_k \\ N\mu_k/p, & -M_k \end{matrix} \middle| 1 \right) \\
&= {}_2F_1\left(\begin{matrix} -\eta(k), -x_k \\ -M_k \end{matrix} \middle| \frac{\mu_k + |\boldsymbol{\mu}^{k+1}|}{\mu_k} \right) \\
&= K_{\eta(k)}\left(x_k; N, \frac{\mu_k}{|\boldsymbol{\mu}^k|}\right),
\end{aligned}$$

for every $k \in [K]$, where the calculation of the limit in the third equation follows from [KLS10, Eq. (1.4.5)] and the last inequality follows from the definition of univariate Krawtchouk polynomials in (2.34). Now, using the previous limit and the definitions (2.31) and (2.33) of the multivariate Hahn and Krawtchouk polynomials we get

$$\lim_{p\to 0^+} H_\eta\left(\mathbf{x}; N, N\frac{\boldsymbol{\mu}}{p}\right) = K_\eta\left(\mathbf{x}; N, \frac{\boldsymbol{\mu}}{|\boldsymbol{\mu}|}\right).$$

Thus, similarly to how we define $\nu_{N,p}$ in (2.13), we define the multivariate polynomial $Q_\eta(\cdot, N, \boldsymbol{\mu}, p)$ by

$$Q_\eta(\mathbf{x}; N, \boldsymbol{\mu}, p) := \begin{cases} H_\eta\left(\mathbf{x}; N, \frac{N\boldsymbol{\mu}}{p}\right) & \text{if} \quad p > 0 \\ K_\eta\left(\mathbf{x}; N, \frac{\boldsymbol{\mu}}{|\boldsymbol{\mu}|}\right) & \text{if} \quad p = 0, \end{cases} \tag{2.35}$$

for every $\eta \in \bigcup_{L=0}^N \mathcal{E}_{K-1,L}$, and for all $\mathbf{x} \in \mathcal{E}_{K,N}$. Note that the functions $Q_\eta(\mathbf{x}; N, \boldsymbol{\mu}, p)$ are continuous when p tends towards zero, in the sense that:

$$\lim_{p\to 0^+} Q_\eta\left(\mathbf{x}; N, \boldsymbol{\mu}, p\right) = Q_\eta\left(\mathbf{x}; N, \boldsymbol{\mu}, 0\right),$$

for every $\mathbf{x} \in \mathcal{E}_{K,N}$. The following result sets some important properties of the multivariate Hahn and Krawtchouk polynomials.

Proposition 2.5.1 (Orthogonality of the Hahn and Krawtchouk polynomials)**.** *The multivariate polynomials Q_η defined by (2.35) satisfy the following properties:*

a) $Q_\eta(\cdot; N, \boldsymbol{\mu}, p)$ is a polynomial on $\mathcal{E}_{K,N}$ of total degree $|\eta|$, for every $\eta \in \bigcup_{L=0}^{N} \mathcal{E}_{K-1,L}$.

b) The polynomials $Q_\eta(\cdot; N, \boldsymbol{\mu}, p)$ are orthogonal on $\mathcal{E}_{K,N}$ with respect to the probability distribution $\nu_{N,p}$, defined by (2.13), i.e.

$$\begin{aligned}
\mathbb{E}_{\nu_{N,p}}\left[Q_\eta(\cdot\,; N, \boldsymbol{\mu}, p) Q_{\eta'}(\cdot\,; N, \boldsymbol{\mu}, p)\right] &= \sum_{\xi \in \mathcal{E}_{K,N}} Q_\eta(\xi\,; N, \boldsymbol{\mu}, p) Q_{\eta'}(\xi\,; N, \boldsymbol{\mu}, p) \nu_{N,p}(\xi) \\
&= d^2_{\eta,p}\, \delta_{\eta,\eta'},
\end{aligned}$$

for every $\eta, \eta' \in \bigcup_{L=0}^{N} \mathcal{E}_{K-1,L}$, where $\delta_{\eta,\eta'}$ stands for the Kronecker delta function and

$$d^2_{\eta,p} = \begin{cases} \dfrac{(|\boldsymbol{\alpha}| + N)_{(|\eta|)}}{(N)_{[|\eta|]} |\boldsymbol{\alpha}|_{(2|\eta|)}} \displaystyle\prod_{j=1}^{K-1} \frac{(|\boldsymbol{\alpha}^j| + |\eta^j| + |\eta^{j+1}| - 1)_{(\eta(j))} (|\boldsymbol{\alpha}^{j+1}| + 2|\eta^{j+1}|)_{(\eta(j))} \eta(j)!}{(\alpha_j)_{(\eta(j))}}, & p > 0 \\ \dfrac{1}{(N)_{[|\eta|]}} \displaystyle\prod_{j=1}^{K-1} \frac{(|\boldsymbol{\pi}^j|)^{\eta(j)} (|\boldsymbol{\pi}^{j+1}|)^{\eta(j)}}{\pi_j^{\eta(j)}} \eta(j)!, & p = 0, \end{cases}$$

where $\boldsymbol{\alpha} = N\boldsymbol{\mu}/p$ *and* $\boldsymbol{\pi} = \boldsymbol{\mu}/|\boldsymbol{\mu}|$.

See Theorem 5.4 in [IX07] and Proposition 2.1, also Remark 2.2, in [KZ09] for the proofs of these results on multivariate Hahn polynomials. See Theorem 6.2 in [IX07] and Proposition 2.4 in [KZ09] for the proofs for the multivariate Krawtchouk polynomials. The system of orthogonal polynomials for a fixed multinomial distribution is not unique. A general construction of the multivariate Krawtchouk polynomials can be found in [DG14].

Kernel polynomials for Dirichlet multinomial and multinomial distributions

Consider ν a multivariate distribution on $\mathcal{E}_{K,N}$ and $\{Q^0_\eta\}$ an orthonormal system of polynomials in $l^2(\mathbb{R}^{\mathcal{E}_{K,N}}, \nu)$. Then, the *kernel polynomial* associated to ν is defined by

$$h_n(\mathbf{x}, \mathbf{y}) := \sum_{|\eta|=n} Q^0_\eta(\mathbf{x}) Q^0_\eta(\mathbf{y}),$$

for all $\mathbf{x}, \mathbf{y} \in \mathcal{E}_{K,N}$ and for every $n \in [N]_0$. The kernel polynomials are invariant under the choice of the orthonormal systems, i.e. they only depend on the distribution ν. Kernel polynomials are used for manipulating sums of products of orthogonal polynomials. They are especially useful to obtain explicit expressions for the transition function of a reversible Markov chain with polynomial eigenfunctions, as we do below in Proposition 2.5.2.

We next review the expressions for the kernel polynomials of the Dirichlet multinomial and the multinomial distributions. Let us denote by $h_n(\mathbf{x}, \mathbf{y}; p)$ the n-th kernel polynomial of $\nu_{N,p}$, for all $n \in [N]_0$. Then, it can be proven that

$$h_n(N\mathbf{e}_k, N\mathbf{e}_k; p) = \binom{N}{n} \frac{(|\boldsymbol{\alpha}| + 2n - 1)(|\boldsymbol{\alpha}|)_{(n-1)}(|\boldsymbol{\alpha}| - \alpha_k)_{(n)}}{(|\boldsymbol{\alpha}| + N)_{(n)}(\alpha_k)_{(n)}}, \tag{2.36}$$

for all $p > 0$, see [KZ09, Eq. (2.18)].

For $p = 0$, $\nu_{N,0}$ follows a $\mathcal{M}(\cdot \mid N, \boldsymbol{\mu}/|\boldsymbol{\mu}|)$ distribution and its n-th kernel polynomial satisfies

$$h_n(\mathbf{x}, N\mathbf{e}_k; 0) = \sum_{m=0}^{n} \binom{N}{m} \binom{N-m}{n-m} (-1)^{n-m} \frac{(x_k)_{[m]}}{N_{[m]}} \left(\frac{\mu_k}{|\boldsymbol{\mu}|}\right)^{-m}, \tag{2.37}$$

and

$$h_n(N\mathbf{e}_k, N\mathbf{e}_k; 0) = \binom{N}{n} \left(\frac{|\boldsymbol{\mu}|}{\mu_k} - 1\right)^n. \tag{2.38}$$

For more details on the kernel polynomials for the multinomial distribution see e.g. [KZ09, Prop. 2.8] and [DG14]. Also, for more details on the kernel polynomials for the Dirichlet multinomial distribution see e.g. [KZ09, Prop. 2.6] and [GS13].

The following proposition shows that the right eigenfunctions of $\mathcal{L}_{N,p}$ are given by multivariate orthogonal polynomials defined by (2.35).

Proposition 2.5.2 (Eigenfunctions of $\mathcal{L}_{N,p}$)**.** *The right eigenfunctions of $\mathcal{L}_{N,p}$ are the multivariate polynomials $Q_\eta(\cdot\,; N, \boldsymbol{\mu}, p)$ with associated eigenvalue $\lambda_{L,p}$, for $\eta \in \mathcal{E}_{K-1,L}$, for $L \in [N]_0$. Moreover, the set of right eigenfunctions*

$$\left\{ Q_\eta(\cdot\,; N, \boldsymbol{\mu}, p), \eta \in \bigcup_{L=0}^{N} \mathcal{E}_{K-1,L} \right\}$$

is orthogonal in $l^2(\nu_{N,p})$, for all $p \geq 0$. In addition, the functions $\phi_\eta(\cdot\,; N, \boldsymbol{\mu}, p)$ defined by

$$\phi_\eta(\eta'; N, \boldsymbol{\mu}, p) := \nu_{N,p}(\eta') Q_\eta(\eta'; N, \boldsymbol{\mu}, p)$$

are left eigenfunctions of $\mathcal{L}_{N,p}$ and the set of left eigenfunctions is orthogonal in $l^2(1/\nu_{N,p})$.

Furthermore, the transition kernel of the Markov chain driven by $\mathcal{L}_{N,p}$ can be decomposed as follows:

$$(\mathrm{e}^{t\mathcal{L}_{N,p}}\delta_\xi)(\eta) = \nu_{N,p}(\xi)\left(1 + \sum_{L=1}^{N} \mathrm{e}^{\lambda_{L,p}t} h_L(\eta, \xi; p)\right), \tag{2.39}$$

where $h_L(\eta, \xi; p)$ is the kernel polynomial associated to $\nu_{N,p}$.

Griffiths and Spanò [GS13] give the expression (2.39) for the transition kernel of the process driven by $\mathcal{L}_{N,p}$, for $p > 0$, as an example of the usefulness of the kernel polynomials for the Dirichlet multinomial distribution. For the sake of brevity, we skip the proof of Proposition 2.5.2 because it comes from a standard method on reversible Markov chains with polynomials eigenfunctions. The interested reader could see, for example, the proofs of Propositions 4.7 and 4.10 in [KZ09].

The following result provides an explicit expression for the chi-square distance between the distribution of the Markov process driven by $\mathcal{L}_{N,p}$ starting at $N\mathbf{e}_k$ and its stationary distribution at a given time t.

Corollary 2.5.3 (Explicit expression for the chi-square distance)**.** *For $K \geq 2$, $N \geq 2$ and $p \geq 0$, we obtain the following explicit expression for the chi-square distance between the distribution of the reversible process driven by $\mathcal{L}_{N,p}$ at time t when the initial distribution is concentrated at $N\mathbf{e}_k$, for $k \in [K]$:*

$$\chi^2_{N\mathbf{e}_k}(t) = \begin{cases} \left[1 + \mathrm{e}^{-2|\boldsymbol{\mu}|t}\left(\frac{|\boldsymbol{\mu}|}{\mu_k} - 1\right)\right]^N - 1 & \text{if } p = 0 \\ \displaystyle\sum_{L=1}^{N} \mathrm{e}^{2\lambda_{L,p}t}\binom{N}{L}\frac{(|\boldsymbol{\alpha}| + 2L - 1)(|\boldsymbol{\alpha}|)_{(L-1)}(|\boldsymbol{\alpha}| - \alpha_k)_{(L)}}{(|\boldsymbol{\alpha}| + N)_{(L)}(\alpha_k)_{(L)}} & \text{if } p > 0 \end{cases} \tag{2.40}$$

Proof. Using classical results on reversible Markov chains, see e.g. [KZ09, Eq. (2.1)], we obtain the following equality for the chi-square distance:

$$\chi^2_{N\mathbf{e}_k}(t) = \sum_{L=1}^{N} \mathrm{e}^{2\lambda_{L,p}t} h(N\mathbf{e}_k, N\mathbf{e}_k; p),$$

where $h(N\mathbf{e}_k, N\mathbf{e}_k; p)$ stands for the kernel polynomials associated to $\nu_{N,p}$, as defined in (2.36) and (2.38). Thus, the expression for $\chi^2_{N\mathbf{e}_k}(t)$ in (2.40) simply comes from (2.36), when $p > 0$.

To prove the case when $p = 0$, note that (2.38) implies

$$\begin{aligned} \chi^2_{N\mathbf{e}_k}(t) &= \sum_{L=1}^{N} \mathrm{e}^{-2L|\boldsymbol{\mu}|t}\binom{N}{L}\left(\frac{|\boldsymbol{\mu}|}{\mu_k} - 1\right)^L \\ &= \left[1 + \mathrm{e}^{-2|\boldsymbol{\mu}|t}\left(\frac{|\boldsymbol{\mu}|}{\mu_k} - 1\right)\right]^N - 1. \end{aligned}$$

□

We now take advantage of the explicit expression in (2.40) to prove the existence of a strongly optimal cutoff in the chi-square distance for the multi-allelic Moran process with parent independent mutation when $N \to \infty$.

Proof of Theorem 2.1.8. Let us first prove the existence of the chi-square cutoff. When $p = 0$, for $t_{N,c} = \dfrac{\ln N + c}{2|\boldsymbol{\mu}|}$ we obtain

$$\begin{aligned}
\lim_{N\to\infty} \chi^2_{N\mathbf{e}_k}(t_{N,c}) &= \lim_{N\to\infty} \left[1 + \frac{\mathrm{e}^{-c}}{N}\left(\frac{|\boldsymbol{\mu}|}{\mu_k} - 1\right)\right]^N - 1 \\
&= \exp\left\{-\left(\frac{|\boldsymbol{\mu}|}{\mu_k} - 1\right)\mathrm{e}^{-c}\right\} - 1.
\end{aligned}$$

Now, since $K_{k,0} = |\boldsymbol{\mu}|/\mu_k - 1$, we have proved the existence of the limit (2.14) for $p = 0$.

Now, for $p > 0$ let us focus on expression (2.40). For every $L \in \mathbb{N}$ and $k \in [K]$, let us denote

$$\phi_{L,k}(N) := \frac{(|\boldsymbol{\alpha}| + 2L - 1)(|\boldsymbol{\alpha}|)_{(L-1)}(|\boldsymbol{\alpha}| - \alpha_k)_{(L)}}{(|\boldsymbol{\alpha}| + N)_{(L)}(\alpha_k)_{(L)}}.$$

We thus have

$$\begin{aligned}
\phi_{L,k}(N) &:= \frac{|\boldsymbol{\alpha}| + 2L - 1}{|\boldsymbol{\alpha}| + L - 1} \frac{\prod_{r=0}^{L-1}(|\boldsymbol{\alpha}| + r)(|\boldsymbol{\alpha}| - \alpha_k + r)}{\prod_{r=0}^{L-1}(|\boldsymbol{\alpha}| + N + r)(\alpha_k + r)} \\
&= \frac{N|\boldsymbol{\mu}|/p + 2L - 1}{N|\boldsymbol{\mu}|/p + L - 1}\left[\frac{|\boldsymbol{\mu}|(|\boldsymbol{\mu}| - \mu_k)}{\mu_k(|\boldsymbol{\mu}| + p)}\right]^L \frac{\prod_{r=0}^{L-1}\left(1 + \frac{p}{N|\boldsymbol{\mu}|}r\right)\left(1 + \frac{p}{N(|\boldsymbol{\mu}| - \mu_k)}r\right)}{\prod_{r=0}^{L-1}\left(1 + \frac{p}{N(|\boldsymbol{\mu}| + p)}r\right)\left(1 + \frac{p}{N\mu_k}r\right)}.
\end{aligned}$$

Hence, for all $L \in \mathbb{N}$ we get

$$\lim_{N\to\infty} \phi_{L,k}(N) = \left[\frac{|\boldsymbol{\mu}|(|\boldsymbol{\mu}| - \mu_k)}{\mu_k(|\boldsymbol{\mu}| + p)}\right]^L = (K_{k,p})^L.$$

Moreover,

$$\binom{N}{L} \underset{N}{\sim} \frac{N^L}{L!} \quad \text{and} \quad \mathrm{e}^{2\lambda_L t_N} \underset{N}{\sim} \frac{(\mathrm{e}^{-c})^L}{N^L},$$

where for two sequences (f_N) and (g_N) the notation $f_N \underset{N}{\sim} g_N$ means $f_N - g_N = o(g_N)$. According to (2.36) we have

$$h_L(N\mathbf{e}_k, N\mathbf{e}_k; p) = \binom{N}{L}\frac{(|\boldsymbol{\alpha}| + 2L - 1)(|\boldsymbol{\alpha}|)_{(L-1)}(|\boldsymbol{\alpha}| - \alpha_k)_{(L)}}{(|\boldsymbol{\alpha}| + N)_{(L)}(\alpha_k)_{(L)}}.$$

Plugging these asymptotic expressions in the L-th summand of (2.40) yields

$$\lim_{N\to\infty} \mathrm{e}^{2\lambda_L t_N} h_L(N\mathbf{e}_k, N\mathbf{e}_k; p) = \frac{(K_{k,p}\,\mathrm{e}^c)^L}{L!}.$$

Moreover,

$$\begin{aligned}
\mathrm{e}^{2\lambda_L t_N} h_L(N\mathbf{e}_k, N\mathbf{e}_k; p) &\le \mathrm{e}^{-L(c + \ln(N))} h_L(N\mathbf{e}_k, N\mathbf{e}_k; p) \\
&= \frac{\mathrm{e}^{-cL}}{N^L}\binom{N}{L}\frac{|\boldsymbol{\alpha}| + 2L - 1}{\boldsymbol{\alpha}| + L - 1}\frac{|\boldsymbol{\alpha}|_{(L)}(|\boldsymbol{\alpha}| - \alpha_k)_{(L)}}{(|\alpha| + N)_{(L)}(\alpha_k)_{(L)}} \\
&= \frac{\mathrm{e}^{-cL}}{L!}\frac{|\boldsymbol{\alpha}| + 2L - 1}{\boldsymbol{\alpha}| + L - 1}\prod_{r=0}^{L-1}\left[\frac{N - r}{N}\frac{|\boldsymbol{\alpha}| + r}{|\boldsymbol{\alpha}| + N + r}\frac{|\boldsymbol{\alpha}| - \alpha_k + r}{\alpha_k + r}\right] \\
&\le 3\frac{(\gamma\mathrm{e}^{-c})^L}{L!},
\end{aligned}$$

where $\gamma = \max\{1, K_{k,0}\}$.

For an arbitrary small $\epsilon > 0$ let us consider $M \in \mathbb{N}$ such that $3 \sum_{L=M+1}^{\infty} \frac{(\gamma \mathrm{e}^{-c})^L}{L!} \leq \frac{\epsilon}{3}$, and let N_ϵ be a positive integer such that

$$\left| \sum_{L=1}^{M} \mathrm{e}^{2\lambda_L t_N} h_L(N\mathbf{e}_k, N\mathbf{e}_k; p) - \sum_{L=1}^{M} \frac{(K_{k,p})^L}{L!} \right| \leq \frac{\epsilon}{3},$$

for all $N \geq N_\epsilon$. Note that

$$\sum_{L=M+1}^{\infty} \frac{(K_{k,p}\mathrm{e}^{-c})^L}{L!} \leq \frac{\epsilon}{3}.$$

Then, for all $N \geq N_\epsilon$, using the triangular inequality we have

$$\left| \sum_{L=1}^{N} \mathrm{e}^{2\lambda_L t_N} h_L(N\mathbf{e}_k) - \left(\exp\{K_{k,p}\mathrm{e}^{-c} - 1\}\right) \right| \leq \epsilon,$$

which concludes the proof for the chi-square cutoff for the process driven by $\mathcal{L}_{N,p}$, for $p \geq 0$. $\square$

Let us establish a result that will be very useful during the proof of Theorem 2.1.9.

Lemma 2.5.4 (Lemma A.2 in [NO21]). *Let $\psi_N \in (0,1)$, for all $N \in \mathbb{N}$, such that $N\psi_N \to \infty$, when $N \to \infty$. Then, for all $y \in \mathbb{R}$ we have*

$$\lim_{N\to\infty} \mathrm{d}^{\mathrm{TV}} \left(\mathrm{Bin}(N, \psi_N), \mathrm{Bin}\left(N, \psi_N + \sqrt{\frac{\psi_N(1-\psi_N)}{N}} y \right) \right) = 2\Phi\left(\frac{1}{2}|y|\right) - 1,$$

where where $\mathrm{Bin}(N, \psi)$ stands for the binomial distribution with N trials and probability of success ψ, and Φ is the cumulative distribution function of the standard normal distribution, i.e.

$$\Phi : t \mapsto \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-s^2/2} \mathrm{d}s.$$

This lemma characterises the limit profile of the total variation distance between two random variables B_1 and B_2, following binomial distributions, when the difference between their means is of the same order of the standard deviation of B_1/N. The proof can be found in the Appendix A.2 of the very recent work of Nestoridi and Olesker-Taylor [NO21].

Proof of Theorem 2.1.9. First note that the lower and upper bounds for $\mathrm{d}^{\mathrm{TV}}_{N\mathbf{e}_k}(t_{N,c})$ are simply consequences of Theorems 2.1.5 and 2.1.8, respectively. Indeed, for $c < 0$ and using Theorems 2.1.5 we have

$$\lim_{N\to\infty} \mathrm{d}^{\mathrm{TV}}_{N\mathbf{e}_k}(t_{N,c}) \geq 1 - \kappa \frac{\|V\|_\infty}{v_k} \mathrm{e}^{-c},$$

where $\kappa = 8(2|\boldsymbol{\mu}| + \|Q_{\boldsymbol{\mu}}\|_\infty) = 32|\boldsymbol{\mu}|$, and V is any right eigenvector of $Q_{\boldsymbol{\mu}}$ with eigenvalue $|\boldsymbol{\mu}|$. Finally, the desired inequality is obtained considering the eigenvector $V = 1/\mu_k \mathbf{e}_k - 1/\mu_s \mathbf{e}_s$, where $s \in [K]$ satisfies $\mu_s \wedge \mu_k = \min_{r:r\neq k} \mu_r \wedge \mu_k$.

Moreover, using the classical inequality between the chi-square and the total variation distances and Theorem 2.1.8 we get

$$\lim_{N\to\infty} \mathrm{d}^{\mathrm{TV}}_{N\mathbf{e}_k}(t_{N,c}) \leq \lim_{N\to\infty} \frac{1}{2} \sqrt{\chi^2_{N\mathbf{e}_k}(t_{N,c})} = \frac{1}{2} \sqrt{\exp\{K_{k,p}\mathrm{e}^{-c}\} - 1}.$$

This concludes to proof of the existence of the $\left(\frac{\ln N}{2|\boldsymbol{\mu}|},1\right)$ total variation cutoff.

Let us now prove the limit profile for the total variation distance when $p=0$. Using (2.37) and (2.39) we get

$$
\begin{aligned}
(\mathrm{e}^{t\mathcal{L}_N}\delta_\xi)(N\mathbf{e}_k) &= \nu_N(\xi)\sum_{L=0}^{N}\mathrm{e}^{-|\boldsymbol{\mu}|Lt}\sum_{m=0}^{L}\binom{N}{m}\binom{N-m}{L-m}(-1)^{L-m}\frac{(\xi_k)_{[m]}}{N_{[m]}}\left(\frac{\mu_k}{|\boldsymbol{\mu}|}\right)^{-m} \\
&= \nu_N(\xi)\sum_{m=0}^{N}\binom{N}{m}\frac{(\xi_k)_{[m]}}{N_{[m]}}\left(\frac{\mu_k}{|\boldsymbol{\mu}|}\right)^{-m}\sum_{L=m}^{N}\mathrm{e}^{-|\boldsymbol{\mu}|Lt}\binom{N-m}{L-m}(-1)^{L-m} \\
&= \nu_N(\xi)\sum_{m=0}^{N}\binom{N}{m}\frac{(\xi_k)_{[m]}}{N_{[m]}}\left(\frac{\mu_k}{|\boldsymbol{\mu}|}\right)^{-m}\mathrm{e}^{-|\boldsymbol{\mu}|mt}(1-\mathrm{e}^{-|\boldsymbol{\mu}|t})^{N-m} \\
&= \nu_N(\xi)(1-\mathrm{e}^{-|\boldsymbol{\mu}|t})^{N}\sum_{m=0}^{\xi_k}\binom{\xi_k}{m}\left[\frac{\mu_k}{|\boldsymbol{\mu}|}\mathrm{e}^{|\boldsymbol{\mu}|t}(1-\mathrm{e}^{-|\boldsymbol{\mu}|t})\right]^{-m} \\
&= \nu_N(\xi)(1-\mathrm{e}^{-|\boldsymbol{\mu}|t})^{N-\xi_k}\left[(1-\mathrm{e}^{-|\boldsymbol{\mu}|t})+\frac{|\boldsymbol{\mu}|\mathrm{e}^{-|\boldsymbol{\mu}|t}}{\mu_k}\right]^{\xi_k}.
\end{aligned}
$$

Thus, the process driven by $\mathcal{L}_N$ starting at $N\mathbf{e}_k$ at time t follows a $\mathcal{M}\left(\cdot\mid N,(1-\mathrm{e}^{-|\boldsymbol{\mu}|t})\frac{\boldsymbol{\mu}}{|\boldsymbol{\mu}|}+\mathrm{e}^{-|\boldsymbol{\mu}|t}\mathbf{e}_k\right)$ distribution, which proves Corollary 2.1.10. Moreover,

$$
\begin{aligned}
\mathrm{d}^{\mathrm{TV}}_{N\mathbf{e}_k}(t) &= \frac{1}{2}\sum_{\xi\in\mathcal{E}_{K,N}}\left|(\mathrm{e}^{t\mathcal{L}_N}\delta_\xi)(N\mathbf{e}_k)-\nu_N(\xi)\right| \\
&= \frac{1}{2}\sum_{L=0}^{N}\sum_{\substack{\xi\in\mathcal{E}_{K,N}:\\ \xi_k=L}}\nu_N(\xi)\left|(1-\mathrm{e}^{-|\boldsymbol{\mu}|t})^{N-\xi_k}\left[1-\mathrm{e}^{-|\boldsymbol{\mu}|t}+\frac{|\boldsymbol{\mu}|\mathrm{e}^{-|\boldsymbol{\mu}|t}}{\mu_k}\right]^{\xi_k}-1\right| \\
&= \frac{1}{2}\sum_{L=0}^{N}\binom{N}{L}\left(\frac{\mu_k}{|\boldsymbol{\mu}|}\right)^{L}\left(1-\frac{\mu_k}{|\boldsymbol{\mu}|}\right)^{N-L}\left|(1-\mathrm{e}^{-|\boldsymbol{\mu}|t})^{N-L}\left[1-\mathrm{e}^{-|\boldsymbol{\mu}|t}+\frac{|\boldsymbol{\mu}|\mathrm{e}^{-|\boldsymbol{\mu}|t}}{\mu_k}\right]^{L}-1\right| \\
&= \mathrm{d}^{\mathrm{TV}}\left(\mathrm{Bin}\left(N,\frac{\mu_k}{|\boldsymbol{\mu}|}\right),\mathrm{Bin}\left(N,\frac{\mu_k}{|\boldsymbol{\mu}|}(1-\mathrm{e}^{-|\boldsymbol{\mu}|t})+\mathrm{e}^{-|\boldsymbol{\mu}|t}\right)\right).
\end{aligned}
$$

Then, we have proved that we can write $\mathrm{d}^{\mathrm{TV}}(t,N\mathbf{e}_k)$ as the total variation distance between two binomial distributions with parameters N both and probabilities of success $\pi_k=\mu_k/|\boldsymbol{\mu}|$ and $\tilde{\pi}_k=\pi_k(1-\mathrm{e}^{-|\boldsymbol{\mu}|t})+\mathrm{e}^{-|\boldsymbol{\mu}|t}$, respectively. For $t_{N,c}=\frac{\ln N+c}{2|\mu|}$ we get

$$
\tilde{\pi}_k=\pi_k+\frac{\sqrt{\pi_k(1-\pi_k)}}{\sqrt{N}}\sqrt{\frac{1-\pi_k}{\pi_k}}\mathrm{e}^{-c/2}.
$$

Therefore, using Lemma 2.5.4 we obtain

$$
\lim_{N\to\infty}\mathrm{d}^{\mathrm{TV}}_{N\mathbf{e}_k}(t)=2\Phi\left(\frac{1}{2}\sqrt{K_{k,0}\mathrm{e}^{-c}}\right)-1,
$$

where $K_{k,0}=\frac{1-\pi_k}{\pi_k}=\frac{|\boldsymbol{\mu}|}{\mu_k}-1$.

□

2.6 Discussion and open problems

There are several future directions to explore in order to better understand Moran models. Despite the fact that it is non-reversible in general, the neutral multi-allelic Moran model with reversible mutation process seems an interesting model for both theoretical and practical reasons (cf. [SH17]). One possible first step to study the eigenfunctions of $\mathcal{Q}_{N,p}$ when Q is reversible, could be the study of the eigenfunctions of the generator of the reproduction process $\mathcal{A}_N$, for $K \geq 3$, extending the results in [Zho08, §4.2.2].

There are several ways to continue the study of the existence of cutoff phenomena for Moran processes. For example, using the results of Zhou and Lange [ZL09], it could be possible to prove the existence of a (strongly optimal) chi-square cutoff for the composition chain, when the process driven by the mutation matrix is reversible. A possible generalisation of Theorems 2.1.8 and 2.1.9 would be to prove the existence of a cutoff phenomenon for the Moran process with parent independent mutation, when initially all the individuals are not of the same type.

Another interesting problem to address is the study of the spectrum of the multi-allelic Moran process with selection. Under selection at birth the infinitesimal rate matrix of the process is reversible, but an explicit expression for its spectral gap is unknown. The multi-allelic Moran process with selection at death seems more complicated from the spectral point of view because it is non-reversible. However, this process is very interesting in population genetics but also, more generally, because of its interpretation as a Fleming−Viot particle system, which approximates the quasi-stationary distribution of a continuous-time Markov chain.

2.A Proofs of Lemmas 2.2.1 and 2.2.3, and Proposition 2.2.4

This section is devoted to the proofs of Lemmas 2.2.1 and 2.2.3, and Proposition 2.2.4.

Proof of Lemma 2.2.1. (a) Let us first prove that for any $\alpha \in \mathcal{E}_{K,N}$, there exists a unique polynomial $P_\alpha \in H_{K,N}$, product of N linear functions on $H_{K,1}$, such that $P_\alpha(\eta) = 1$ if $\eta = \alpha$ and 0 otherwise. Indeed, let us define the polynomial P_α by

$$P_\alpha : \mathbf{x} \in \mathcal{E}_{K,N} \mapsto \prod_{k=1}^{K} \prod_{a=0}^{\alpha_k - 1} \frac{x_k - a}{\alpha_k - a},$$

where $\prod_{a=0}^{\alpha_k-1}(x_k - a) = 1$ when $\alpha_k = 0$. Note that $P_\alpha = \mathbf{1}_\alpha$, for every $\alpha \in \mathcal{E}_{K,N}$. There are $\sum_{k=1}^{K} \alpha_k = N$ linear factors in the numerator. Also, each term $x_k - a$ may be replaced by $x_k - \frac{a}{N}\sum_{k=1}^{K} x_k$ when $a \neq 0$, so $P_\alpha(x)$ may be considered as a product of N linear functions on $H_{K,1}$, and because the uniqueness of such a function P_α is straightforward, (a) is proved.

Now, for every real function f on $\mathcal{E}_{K,N}$, the result is immediately obtained from (a) by setting

$$P := \sum_{\alpha \in \mathcal{E}_{K,N}} f(\alpha) P_\alpha.$$

(b) From part (b) we have that $\mathcal{B}_{H_{K,N}}$ is a generator system of $\mathbb{R}^{\mathcal{E}_{K,N}}$. Moreover,

$$\mathrm{Card}(\mathcal{B}_{H_{K,N}}) = \mathrm{Card}(\mathcal{E}_{K,N}) = \dim(\mathbb{R}^{\mathcal{E}_{K,N}}) = \binom{K-1+N}{N},$$

thus $\mathcal{B}_{H_{K,N}}$ is necessarily a basis of $\mathbb{R}^{\mathcal{E}_{K,N}}$. $\square$

Proof of Lemma 2.2.3. (a) For $L=1$: An injection $s:\{1\}\to\{1,2,\dots,N\}$ is characterised by $s(1)=i$. It follows from (2.20) that

$$\xi(V_1)(k_1,k_2,\dots,k_N)=\sum_{i=1}^{N}V_1(k_i),$$

which is a symmetric function. For every $\eta=(\eta(1),\eta(2),\dots,\eta(K))\in\mathcal{E}_{K,N}$, we have

$$\tilde{\xi}(V_1)(\eta)=(\xi(V_1)\circ\psi_{K,N})(\eta)=\sum_{j=1}^{K}V_1(j)\eta(j),$$

which finishes the proof of part (a).

(b) From (2.20), we get

$$\begin{aligned}\xi(V_1,V_2,\dots,V_L)(k_1,\dots,k_N)&=\sum_{s\in\mathcal{I}_{L-1,N}}V_1(k_{s(1)})\dots V_{L-1}(k_{s(L-1)})\sum_{i\in[N]\setminus s([L-1])}V_L(k_i)\\&=\sum_{s\in\mathcal{I}_{L-1,N}}V_1(k_{s(1)})\dots V_{L-1}(k_{s(L-1)})\left(\sum_{i=1}^{N}V_L(k_i)-\sum_{i=1}^{L-1}V_L(k_{s(i)})\right)\\&=\xi(V_1,V_2,\dots,V_{L-1})(k_1,\dots,k_N)\xi(V_L)(k_1,\dots,k_N)\\&\quad-\sum_{i=1}^{L-1}\xi(V_1,\dots,V_i\odot V_L,\dots,V_{L-1})(k_1,\dots,k_N).\end{aligned}$$

Using (2.22) we obtain the result for $\tilde{\xi}(V_1,V_2,\dots,V_L)$. The particular case $L=2$ comes from part (a).

(c) We can prove equation (2.25) by induction on L. For $L=1$ the result easily comes by (a). If we suppose that (2.25) is satisfied for L, for $2\le L<N-1$, then, using (b) and (a), we can check that (2.25) holds for $L+1$. □

Proof of Proposition 2.2.4. Since $\mathcal{U}$ is a basis of $\mathbb{R}^K$ we trivially have that $\mathcal{U}^N$ is a basis of $\mathbb{R}^{[K]^N}$, proving (a) (cf. Lemma 12.12 in [LP17]). To prove (b) we prove that each element of $\mathcal{U}^N$ has image in $\mathcal{S}^N$ by Sym, defined as in (2.5). First, $\mathrm{Sym}(U_0\otimes\dots\otimes U_0)=U_0\otimes\dots\otimes U_0$, since the constant function equal to one is symmetric. Furthermore, for every $W=W_1\otimes W_2\otimes\dots\otimes W_N\in\mathcal{U}^N$ there is a permutation $\sigma\in\mathcal{S}_N$ such that $\sigma W=U_\eta$, with $\eta\in\mathcal{E}_{K-1,L}$, where $L\in[N]$ is the number of components in the expression of W different from U_0. Thus, $\mathrm{Sym}(W)=\mathrm{Sym}(\sigma W)=V_\eta$, for $\eta\in\mathcal{E}_{K-1,L}$. We have not proved that $V_\eta\neq V_\alpha$, for $\eta\neq\alpha$. However, $\mathcal{S}^N$ is a generator system of $\mathrm{Sym}(\mathbb{R}^{[K]^N})$ satisfying

$$\begin{aligned}\mathrm{Card}\left(\mathcal{S}^N\right)&\le 1+\sum_{L=1}^{N}\mathrm{Card}(\mathcal{E}_{K-1,L})\\&=\sum_{L=0}^{N}\binom{K-2+L}{L}=\binom{K-1+N}{N},\end{aligned}$$

where the last equality is the well-known *Hockey−Stick identity* in combinatorics, see e.g. [LPV03]. Now, since

$$\dim\left(\mathrm{Sym}\left(\mathbb{R}^{[K]^N}\right)\right)=\binom{K-1+N}{N},$$

we have that $\mathcal{S}^N$ is a generator system with a minimal number of vectors, therefore it is a basis of $\mathrm{Sym}(\mathbb{R}^{[K]^N})$. To prove (c) simply note that each element in $\tilde{\mathcal{S}}^N$ is the image by the isomorphism $\Phi_{K,N}$ of an element in $\mathcal{S}^N$. □

2.B Proof of Lemma 2.3.3

Proof of Lemma 2.3.3. Without lost of generality we can only prove the result for the monomials on $\mathcal{E}_{K,N}$. Consider m a monomial on $\mathcal{E}_{K,N}$ of total degree $|\alpha| = L$ with $0 \leq L \leq N$. Then, we want to prove that

$$\mathcal{A}_N V_m = -L(L-1)V_m + V_q,$$

where q is a polynomial with a total degree strictly less than L.

As we commented in Remark 2.3.4, the result is true for $L = 1$. Let us assume $L \geq 2$ and consider the monomial $m : \eta \mapsto \prod_{r=1}^{K} \eta(r)^{\alpha_r}$. Evaluating V_m in $\mathcal{A}_N$, defined by (2.4), we obtain

$$(\mathcal{A}_N V_m)(\eta) = \sum_{k,r:k\neq r} \left(\prod_{s\notin\{k,r\}} \eta(s)^{\alpha_s} \right) \left[(\eta(k)-1)^{\alpha_k}(\eta(r)+1)^{\alpha_r} - \eta(k)^{\alpha_k}\eta(r)^{\alpha_r} \right] \eta(k)\eta(r), \tag{2.41}$$

for all $\eta \in \mathcal{E}_{K,N}$. Then, from the Newton's binomial formula, we get

$$\eta(k)(\eta(k)-1)^{\alpha_k} = \eta(k)^{\alpha_k+1} - \alpha_k\eta(k)^{\alpha_k} + \frac{\alpha_k(\alpha_k-1)}{2}\eta(k)^{\alpha_k-1} + a(\eta(k)),$$

where $a(\eta(k))$ is a polynomial in $\eta(k)$ with degree strictly less than $\alpha_k - 1$ if $\alpha_k \geq 2$ and null otherwise. In the same way, we get

$$\eta(r)(\eta(r)+1)^{\alpha_r} = \eta(r)^{\alpha_r+1} + \alpha_r\eta(r)^{\alpha_r} + \frac{\alpha_r(\alpha_r-1)}{2}\eta(r)^{\alpha_r-1} + b(\eta(r)),$$

where $b(\eta(r))$ is a polynomial in $\eta(r)$ with degree strictly less than $\alpha_r - 1$ if $\alpha_r \geq 2$ and null otherwise.

Using this expansion in (2.41) and regrouping terms with total degree in $\eta(k)$ and $\eta(r)$ strictly less than $\alpha_k + \alpha_r$ give

$$\begin{aligned}
(\mathcal{A}_N V_m)(\eta) = &\sum_{k,r:k\neq r} \left(\prod_{s\notin\{k,r\}} \eta(s)^{\alpha_s} \right) \left(\alpha_r\eta(k)^{\alpha_k+1}\eta(r)^{\alpha_r} - \alpha_k\eta(k)^{\alpha_k}\eta(r)^{\alpha_r+1} \right) \\
&+ \sum_{k,r:k\neq r} \left(\prod_{s\notin\{k,r\}} \eta(s)^{\alpha_s} \right) \frac{\alpha_r(\alpha_r-1)}{2}\eta(k)^{\alpha_k+1}\eta(r)^{\alpha_r-1} \\
&- \sum_{k,r:k\neq r} \left(\prod_{s\notin\{k,r\}} \eta(s)^{\alpha_s} \right) \alpha_k\alpha_r\eta(k)^{\alpha_k}\eta(r)^{\alpha_r} \\
&+ \sum_{k,r:k\neq r} \left(\prod_{s\notin\{k,r\}} \eta(k)^{\alpha_s} \right) \frac{\alpha_k(\alpha_k-1)}{2}\eta(k)^{\alpha_k-1}\eta(r)^{\alpha_r+1} + w(\eta),
\end{aligned} \tag{2.42}$$

where w is a polynomial in η of total degree strictly less than $\sum_k \alpha_k = L$. The first sum in the right member of (2.42) is null because the antisymmetry in k, r of its summands. The third term is

$$-\sum_{k,r:k\neq r} \left(\prod_{s\notin\{k,r\}} \eta(s)^{\alpha_s} \right) \alpha_k\alpha_r\eta(k)^{\alpha_k}\eta(r)^{\alpha_r} = -c_1 p(\eta),$$

with

$$c_1 = \sum_{k,r:k\neq r} \alpha_k\alpha_r = \left(\sum_{k=1}^{K} \alpha_k \right)^2 - \sum_{k=1}^{K} \alpha_k^2 = L^2 - \sum_{k=1}^{K} \alpha_k^2.$$

By symmetry in k and r, it is obvious that the second and the fourth sums in the right member of (2.42) are equal. Using

$$\alpha_r(\alpha_r-1)\eta(r)^{\alpha_r-1}=\eta(r)\frac{\partial^2}{\partial\eta(r)^2}\eta(r)^{\alpha_r},$$

it follows that

$$\begin{aligned}\sum_{k\neq r}\left(\prod_{s\notin\{k,r\}}\eta(s)^{\alpha_s}\right)\alpha_r(\alpha_r-1)\eta(k)^{\alpha_k+1}\eta(r)^{\alpha_r-1}&=\sum_{k,r:k\neq r}\eta(k)\eta(r)\frac{\partial^2}{\partial\eta(r)^2}m(\eta)\\&=\sum_{k,r}\eta(k)\eta(r)\frac{\partial^2}{\partial\eta(r)^2}m(\eta)-\sum_{r=1}^{K}\eta(r)^2\frac{\partial^2}{\partial\eta(r)^2}m(\eta)\\&=N\sum_{r=1}^{K}\eta(r)\frac{\partial^2}{\partial\eta(r)^2}m(\eta)-\sum_{r=1}^{K}\eta(r)^2\frac{\partial^2}{\partial\eta(r)^2}m(\eta).\end{aligned}$$

The first summand in the last equality is an homogeneous polynomial of degree $L-1$ and the second one satisfies

$$-\sum_{r=1}^{K}\eta(r)^2\frac{\partial^2}{\partial\eta(r)^2}m(\eta)=-c_2\,m(\eta),$$

with

$$c_2=\sum_{r=1}^{K}\alpha_r(\alpha_r-1)=\sum_{r=1}^{K}\alpha_r^2-L.$$

As a conclusion, it comes from (2.42) that

$$\mathcal{A}_NV_m=-(c_1+c_2)V_m+V_q=-L(L-1)V_m+V_q,$$

where q is a polynomial of total degree strictly less than L, which proves (a).

□

2.C Proof of Lemma 2.1.7

First we prove Lemma 2.C.1 showing that the neutral multi-allelic Moran process driven by $\mathcal{Q}_{N,p}$ is reversible if and only if its mutation rate matrix can be written in the form of $Q_{\boldsymbol{\mu}}$, given by (2.11). We start by proving that when the neutral multi-allelic Moran process is reversible, then all the entries of the mutation matrix are positive and it can be written in the form of $Q_{\boldsymbol{\mu}}$, i.e. the *"only if part"*. Later, in Lemma 2.C.2 we prove that the process driven by $\mathcal{L}_{N,p}$ is reversible and we provide the explicit expression for its stationary distribution, i.e. we prove the *"if part"*. Actually, the results in Lemma 2.C.2 are proved for a more general Moral model with *selection at birth.*

Lemma 2.C.1. *If the process driven by the generator (2.2) is reversible, then $\mu_{i,j}=\mu_j>0$, for all $i\in[K]$, and every $j\in[K]$, $j\neq i$.*

Proof. We first prove that if the process is reversible, then all the entries of the mutation matrix are positive. Let us denote by $\nu_{N,p}$ the stationary probability measure of the process driven by $\mathcal{Q}_{N,p}$, which is assumed to be reversible. We denote $\mathcal{Q}_{N,p}[\eta,\xi]:=(\mathcal{Q}_{N,p}\delta_\xi)(\eta)$, for all $\eta,\xi\in\mathcal{E}_{K,N}$. Consider the states $\eta^{(1)}$ and $\eta^{(2)}$ defined as $\eta^{(1)}:=N\mathbf{e}_i$ and $\eta^{(2)}:=\eta^{(1)}-\mathbf{e}_i+\mathbf{e}_j$, for $i,j\in[K]$ such that $i\neq j$. Since the process is reversible, the measure ν_N satisfies the balance equation

$$\nu_{N,p}(\eta^{(1)})\mathcal{Q}_{N,p}[\eta^{(1)},\eta^{(2)}]=\nu_{N,p}(\eta^{(2)})\mathcal{Q}_{N,p}[\eta^{(2)},\eta^{(1)}],$$

see e.g. [Kel79, Thm1.3]. We have

$$\mathcal{Q}_{N,p}[\eta^{(1)}, \eta^{(2)}] = N\mu_{i,j}, \text{ and}$$
$$\mathcal{Q}_{N,p}[\eta^{(2)}, \eta^{(1)}] = \mu_{j,i} + p(N-1)/N > 0.$$

Furthermore, since the process is irreducible we have that $\nu_N(\eta) > 0$, for all $\eta \in \mathcal{E}_{K,N}$. Finally, the balance equation implies that $\mu_{i,j} > 0$, for all $i \neq j$.

Now, we prove that for every $j \in [K]$ we have $\mu_{i,j} = \mu_j > 0$, for all $i \in [K]$. For $K = 2$, there is nothing to prove. For $K \geq 3$, $N \geq 2$, let us consider a general model with a reversible stationary probability. Let i, j, k be three different indices on $[K]$ and consider the four states $\eta^{(1)}$, $\eta^{(2)}$, $\eta^{(3)}$ and $\eta^{(4)}$ in $\mathcal{E}_{K,N}$ defined by

$$\eta^{(1)} := N\mathbf{e}_i, \quad \eta^{(2)} := \eta^{(1)} - \mathbf{e}_i + \mathbf{e}_j, \quad \eta^{(3)} := \eta^{(1)} - 2\,\mathbf{e}_i + \mathbf{e}_j + \mathbf{e}_k, \quad \eta^{(4)} := \eta^{(1)} - \mathbf{e}_i + \mathbf{e}_k.$$

Note that

$$\begin{aligned}
&\mathcal{Q}_{N,p}[\eta^{(1)}, \eta^{(2)}] = N\mu_{i,j}, && \mathcal{Q}_{N,p}[\eta^{(2)}, \eta^{(1)}] = \mu_{j,i} + (N-1)p/N, \\
&\mathcal{Q}_{N,p}[\eta^{(2)}, \eta^{(3)}] = (N-1)\mu_{i,k}, && \mathcal{Q}_{N,p}[\eta^{(3)}, \eta^{(2)}] = \mu_{k,i} + (N-2)p/N, \\
&\mathcal{Q}_{N,p}[\eta^{(3)}, \eta^{(4)}] = \mu_{j,i} + (N-2)p/N, && \mathcal{Q}_{N,p}[\eta^{(4)}, \eta^{(3)}] = (N-1)\mu_{i,j}, \\
&\mathcal{Q}_{N,p}[\eta^{(4)}, \eta^{(1)}] = \mu_{k,i} + (N-1)p/N, && \mathcal{Q}_{N,p}[\eta^{(1)}, \eta^{(4)}] = N\mu_{i,k}.
\end{aligned}$$

Then,

$$\frac{\mathcal{Q}_{N,p}[\eta^{(4)}, \eta^{(1)}]}{N(N-1)} \prod_{r=1}^{3} \mathcal{Q}_{N,p}[\eta^{(r)}, \eta^{(r+1)}] = \mu_{i,j}\mu_{i,k}\left(\mu_{j,i} + p\frac{N-2}{N}\right)\left(\mu_{k,i} + p\frac{N-1}{N}\right),$$
$$\frac{\mathcal{Q}_{N,p}[\eta^{(1)}, \eta^{(4)}]}{N(N-1)} \prod_{r=1}^{3} \mathcal{Q}_{N,p}[\eta^{(r+1)}, \eta^{(r)}] = \mu_{i,k}\mu_{i,j}\left(\mu_{j,i} + p\frac{N-1}{N}\right)\left(\mu_{k,i} + p\frac{N-2}{N}\right).$$

Therefore, since the stationary probability is reversible, the *Kolmogorov cycle reversibility criterion* [Kel79, Thm. 1.8] holds:

$$\mathcal{Q}_{N,p}[\eta^{(4)}, \eta^{(1)}] \prod_{r=1}^{3} \mathcal{Q}_{N,p}[\eta^{(r)}, \eta^{(r+1)}] = \mathcal{Q}_{N,p}[\eta^{(1)}, \eta^{(4)}] \prod_{r=1}^{3} \mathcal{Q}_{N,p}[\eta^{(r+1)}, \eta^{(r)}],$$

and we get $p(N-1)\mu_{i,j}\mu_{i,k}(\mu_{j,i} - \mu_{k,i}) = 0$. We know that $\mu_{i,j} > 0$ for all $i, j \in [K]$, thus $\mu_{j,i} = \mu_{k,i}$, for all $j, k \in [K]$, with $j \neq k$, and every $i \in [K]$, with $i \notin \{j, k\}$. Denoting $\mu_j := \mu_{i,j}$ for any $i \in [K]$, with $i \neq j$, we prove that the mutation matrix is of the form of $Q_{\boldsymbol{\mu}}$ for a suitable vector $\boldsymbol{\mu}$.

□

It remains to prove that the stationary distribution of $\mathcal{L}_{N,p}$ is compound Dirichlet multinomial with suitable parameters. Actually, a more general version of Lemma 2.1.7 can be proved, where the values of the parameter p in (2.5.1) also depend on j, i.e. a model with selection at birth or fecundity selection [MW09]. Abusing notation, for two vectors $\boldsymbol{p} = (p_1, p_2, \ldots, p_K)$ and $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_K)$ such that $p_j, \mu_j > 0$, for all $j \in [K]$, let us denote by $\mathcal{L}_{N,\boldsymbol{p}}$ the infinitesimal generator satisfying

$$(\mathcal{L}_{N,\boldsymbol{p}}f)(\eta) := \sum_{i,j=1}^{K} \eta(i)\left(\mu_j + p_j\frac{\eta(j)}{N}\right)[f(\eta - \mathbf{e}_i + \mathbf{e}_j) - f(\eta)], \tag{2.43}$$

for every function f on $\mathcal{E}_{K,N}$ and all $\eta \in \mathcal{E}_{K,N}$. We define the *weighted Dirichlet-compound multinomial distribution* with parameters N, $\boldsymbol{\mu}$ and $\boldsymbol{p}$, denoted $\mathcal{WDM}(\cdot \mid N, \boldsymbol{\mu}, \boldsymbol{p})$, as follows

$$\mathcal{WDM}(\eta \mid N, \boldsymbol{\mu}, \boldsymbol{p}) := Z^{-1} \binom{N}{\eta} \prod_{k=1}^{K} p_k^{\eta(k)} (\alpha_k)_{(\eta(k))}, \tag{2.44}$$

for all $\eta \in \mathcal{E}_{K,N}$, where $\alpha_k = \mu_k / p_k$, for all $k \in [K]$ and Z is a normalisation constant satisfying

$$Z = \mathbb{E}\left[\left(\sum_{j=1}^{K} p_j X_j\right)^N\right], \tag{2.45}$$

where $(X_1, X_2, \ldots, X_K)$ follows a $\mathcal{DM}(\cdot \mid N, N\boldsymbol{\mu})$. Note that the measure defined by (2.44) with the normalisation constant (2.45) is a probability distribution. See [JKB97] and [NRA06] for more details about the weighted multinomial distributions.

Lemma 2.C.2 (Reversible probability of $\mathcal{L}_{N,\boldsymbol{p}}$)**.** *The process driven by (2.43) is reversible and its stationary distribution is $\mathcal{WDM}(\cdot \mid N, \boldsymbol{\alpha}, \boldsymbol{p})$, where $\alpha_k = N\mu_k$, for all $k \in [K]$.*

Remark 2.C.1. This result is known for multi-allelic Moran models with parent independent mutation. See e.g. [EG09, Section 3]. However, we have not found a proof in the literature. So, for the sake of completeness we provide a proof. When the vector $\boldsymbol{p}$ is constant we obtain the stationary distribution of the neutral case and we thus conclude the proof of Lemma 2.1.7.

Proof of Lemma 2.C.2. Let us define $q_k := p_k/N$, for $k \in [K]$ and, abusing notation, $\mathcal{L}_{N,\boldsymbol{p}}[\eta, \xi] := \mathcal{L}_{N,\boldsymbol{p}} \delta_\xi(\eta)$, for all $\eta, \xi \in \mathcal{E}_{K,N}$. Note that for $\eta, \xi \in \mathcal{E}_{K,N}$ with $\eta \neq \xi$, we have $\mathcal{L}_{N,\boldsymbol{p}}[\eta, \xi] \neq 0$ if and only if there exist $i, j \in [K]$, such that $i \neq j$, $\eta(i) > 0$ and $\xi = \eta - \mathbf{e}_i + \mathbf{e}_j$. In this case

$$\mathcal{L}_{N,\boldsymbol{p}}[\eta, \xi] = \eta(i)[\mu_j + \eta(j) q_j].$$

This implies that $\xi(j) = \eta(j) + 1 > 0$ and $\eta = \xi - \mathbf{e}_j + \mathbf{e}_i$. As a consequence

$$\mathcal{L}_{N,\boldsymbol{p}}[\xi, \eta] = \xi(j)[\mu_i + \xi(i) q_i] = (\eta(j) + 1)[\mu_i + (\eta(i) - 1) q_i].$$

Also $\eta(k) = \xi(k)$, for all $k \neq i,\ k \neq j$.

Therefore we get,

$$\begin{aligned} Z\, \mathcal{WDM}(\eta \mid N, \boldsymbol{\mu}, \boldsymbol{p}) \mathcal{L}_{N,\boldsymbol{p}}[\eta, \xi] &= \binom{N}{\eta} \left[\prod_{k=1}^{K} p_k^{\eta(k)} \left(\frac{\mu_k}{q_k}\right)_{(\eta(k))}\right] \eta(i)[\mu_j + \eta(j) q_j] \\ &= \frac{N!}{\prod_{k \notin \{i,j\}} \eta(k)!} \frac{1}{\eta(i)!\eta(j)!} \left[\prod_{k=1}^{K} \prod_{l=0}^{\eta(k)-1} (\mu_k + l\, q_k)\right] \eta(i)[\mu_j + \eta(j)\, q_j], \end{aligned} \tag{2.46}$$

where Z is the normalisation constant given by (2.45). Note that

$$\frac{N!}{\prod_{k \notin \{i,j\}} \eta(k)!} \prod_{k \notin \{i,j\}}^{K} \prod_{l=0}^{\eta(k)-1} (\mu_k + l\, q_k) = \frac{N!}{\prod_{k \notin \{i,j\}} \xi(k)!} \prod_{k \notin \{i,j\}}^{K} \prod_{l=0}^{\xi(k)-1} (\mu_k + l\, q_k), \tag{2.47}$$

because $\eta(k) = \xi(k)$, for $k \notin \{i, j\}$. Moreover,

$$\frac{1}{\eta(i)!\, \eta(j)!} \eta(i) = \frac{1}{(\eta(i) - 1)!\, \eta(j)!} = \frac{1}{\xi(i)!\, (\xi(j) - 1)!} = \frac{1}{\xi(i)!\, \xi(j)!}\, \xi(j), \tag{2.48}$$

because $\xi(i) = \eta(i) - 1$ and $\xi(j) = \eta(j) + 1$. In addition,

$$\prod_{l=0}^{\eta(i)-1} (\mu_i + l\, q_i) = \prod_{l=0}^{\xi(i)} (\mu_i + l\, q_i) = (\mu_i + \xi(i)\, q_i) \prod_{l=0}^{\xi(i)-1} (\mu_i + l\, q_i), \tag{2.49}$$

and

$$\left[\prod_{l=0}^{\eta(j)-1} (\mu_j + l\, q_j)\right] [\mu_j + \eta(j)\, q_j] = \prod_{l=0}^{\eta(j)} (\mu_j + l\, q_j) = \prod_{l=0}^{\xi(j)-1} (\mu_j + l\, q_j). \tag{2.50}$$

Using (2.47), (2.48), (2.49) and (2.50) in (2.46) gives

$$\begin{aligned} \mathcal{WDM}(\eta \mid N, \boldsymbol{\mu}, \boldsymbol{p}) \mathcal{L}_{N,\boldsymbol{p}}[\eta, \xi] &= Z^{-1} \binom{N}{\xi} \left[\prod_{k=1}^{K} p_k^{\xi(k)} \left(\frac{\mu_k}{p_k}\right)_{(\xi(k))}\right] \xi(j)[\mu_i + \xi(i) q_i] \\ &= \mathcal{WDM}(\xi \mid N, \boldsymbol{\mu}, \boldsymbol{p}) \mathcal{L}_{N,\boldsymbol{p}}[\xi, \eta], \end{aligned}$$

for all $\eta, \xi \in \mathcal{E}_{K,N}$. The distribution ν_N satisfies the detailed balance property, thus it is reversible for $\mathcal{L}_{N,\boldsymbol{p}}$, and it is the unique stationary measure, because the process generated by $\mathcal{L}_{N,\boldsymbol{p}}$ is irreducible.

□

Chapter 3

Quantitative bounds in a non-reversible example

This chapter is based on the paper [Cor21a], published in *Stochastic Processes and their Applications* in 2021.

Abstract: We study the neutral multi-allelic Moran model where the mutation scheme is given by an asymmetric random walk in the cycle graph. We show that this model has a remarkable exact solvability, despite the fact that it is non-reversible with non-explicit invariant distribution. Our main results include quantitative propagation of chaos and exponential ergodicity with explicit constants, as well as formulas for covariances at equilibrium in terms of the Chebyshev polynomials. We also obtain an explicit uniform in time bound for the convergence of the proportion of particles in each state when the number of particles goes to infinity.

3.1 Introduction

This chapter deals with a continuous-time Markov process describing the position of N particles moving around on the cycle graph. This type of model is usually known as Fleming – Viot process, or Moran type process [CT16b; EG09; FM07]. Consider a continuous-time Markov process on $E \cup \{\partial\}$, where E is finite and ∂ is an absorbing state. Briefly, the Fleming – Viot process consists of N particles moving in E as independent copies of the original process, until one of the particles gets absorbed. When this happens, the absorbed particle jumps instantaneously and uniformly to one of the positions of the other particles. The Fleming – Viot processes were originally and independently introduced by Del Moral, Guionnet, Miclo [DG99; DM00a] and Burdzy, Hołyst, March [BHM00] to approximate the law of a Markov process conditioned to non-absorption, and its *Quasi-Stationary Distribution* (QSD), which is the limit of this conditional law when $t \to \infty$. See e.g. the works of Méléard and Villemonais [MV12], Collet et al. [CMS13] and van Doorn et al. [DP13], excellent references for an introduction to the theory related to the QSD. For recent and quite general results about the convergence of Markov processes conditioned to non-absorption to a QSD, we refer the interested reader to [CV16], [CV17a] and [Ban+21].

The convergence of the empirical distributions induced by Fleming – Viot processes defined on discrete state spaces when the size of the population and the time increase have been assured under some assumptions. For example, Ferrari and Marić [FM07] and Asselah et al. [AFG11] study the convergence of the empirical distribution induced by the Fleming – Viot process to the unique QSD in countable and finite discrete space settings, respectively. With the aim to study the convergence of the particle process under the stationary distribution to the QSD, Lelièvre et al. [LPR18] proves a Central Limit Theorem for the finite state case. Additionally, Villemonais [Vil15] and Asselah et al. [Ass+16] study the convergence to the minimal QSD in a Galton – Watson type model and in a birth and death process, respectively. Similarly, Asselah and Thai [AT12] and Mariĉ [Mar15] address the study of the N-particle system associated to a random walk on $\mathbb{N}$ with a drift towards the origin, which is an absorbing state. In these scenarios there exist infinitely many QSD for each model, so it is important to ensure the ergodicity of the N-particle system and to determine to which QSD it converges. Additionally, Champagnat and Villemonais [CV21] study the convergence of the Fleming – Viot process to the minimal QSD under general conditions, providing also some specific examples.

In addition, some works have been devoted to the study of the speed of convergence when the number of particles and time tend to infinity. In particular, Cloez and Thai [CT16b] study the N-particle system in a discrete state space setting. They study the convergence of the empirical measure induced by the Fleming – Viot process when both $t \to \infty$ (ergodicity) and $N \to \infty$ (propagation of chaos), providing explicit bounds for the speed of convergence. Following the results in [CT16b], Cloez and Thai [CT16a] study two examples in details: the random walk on the complete graph with uniform killing and the random walk on the two-site graph. The simple geometries of the graphs of these models simplify the study of the N-particle dynamic and allows them to give explicit expressions for the stationary distributions of the N-particle processes and explicit bounds for its convergence to the QSD.

Consider the quantity λ defined in [CT16b] as

$$\lambda = \inf_{x,y} \Big(Q_{x,y} + Q_{y,x} + \sum_{s \neq x,y} Q_{x,s} \wedge Q_{y,s} \Big), \tag{3.1}$$

where $Q = \left(Q_{x,y}\right)_{x,y}$ is the infinitesimal generator matrix of the process until absorption. When $\lambda = 0$ some of the results of [CT16b] do not hold and most of the bounds given become too rough. Note that $\lambda > 0$ for the two examples studied in [CT16a], but λ is equal to zero for those models where there exist two vertices such that the distance between them is greater than two. The quantity λ is somehow related to the geometry of the graph associated to the Markov process. Hence, it becomes interesting to find explicit bounds for the speed of convergence of Fleming – Viot processes with more complex geometries.

In this chapter we focus on the random walk on the cycle graph $\mathbb{Z}/K\mathbb{Z}$ for $K \geq 3$. Note that for this graph it holds that $\lambda = 0$ when $K \geq 6$. For simplicity, we assume that the N particles jump to the absorbing state with the same rate, i.e., we consider a process with *uniform killing* (cf. [MV12]). Even if in this case the distribution of the conditional process is trivial, the study of the Fleming – Viot process becomes more complicated due to its non reversibility and the geometry of the cycle graph. We focus on providing bounds for the speed of the convergence of the empirical distribution induced by the particle system to the unique QSD when t and N tend to infinity. This example can be seen as a further step towards the study of the speed of convergence of Fleming – Viot process with more general geometry.

3.1.1 Model and notations

Consider a Markov process $(Z_t)_{t \geq 0}$ with state space $\mathbb{Z}/K\mathbb{Z} \cup \{\partial\}$, where $K \geq 3$ and ∂ is an absorbing state. Specifically, the infinitesimal generator of the process is given by

$$\mathcal{G}f(x) = f(x+1) - f(x) + \theta[f(x-1) - f(x)] + p[f(\partial) - f(x)],$$

where $x \in \mathbb{Z}/K\mathbb{Z}$, $\mathcal{G}f(\partial) = 0$, $\theta, p \in \mathbb{R}_+^*$ and f is a real function defined on $\mathbb{Z}/K\mathbb{Z} \cup \{\partial\}$. In words, $(Z_t)_{t \geq 0}$ is an asymmetric random walk on the K-cycle graph, which jumps with rates 1 and θ in the clockwise and the anti-clockwise directions, respectively. Also, with uniform rate p the process jumps to the absorbing state ∂, i.e., it is killed. Note that $\mathbb{Z}/K\mathbb{Z}$ is an irreducible class. The process generated by $\mathcal{G}$ is a particular case of the processes with *uniform killing in a finite state space* considered by Méléard and Villemonais [MV12, § 2.3].

Let $(X_t)_{t\geq 0}$ be the analogous asymmetric random walk on the cycle graph $\mathbb{Z}/K\mathbb{Z}$ without killing. The generator of this process, denoted by $\mathcal{H}$, is given by

$$\mathcal{H}f(x) = f(x+1) - f(x) + \theta[f(x-1) - f(x)], \quad \text{for all} \;\; x \in \mathbb{Z}/K\mathbb{Z}.$$

Note that, because of the uniform killing, the process $(Z_t)_{t\geq 0}$ could also be defined in the following way

$$Z_t = \begin{cases} X_t & \text{if} \quad t < \tau_p \\ \partial & \text{if} \quad t \geq \tau_p, \end{cases}$$

where τ_p is an exponential random variable with mean $1/p$ and independent of the random walk $(X_t)_{t\geq 0}$. This means that the law of the process $(Z_t)_{t\geq 0}$ conditioned to non-absorption is given by

$$\mathbb{P}_\mu[Z_t = k \,|\, t < \tau_p] = \mathbb{P}_\mu[X_t = k],$$

for $k \in \mathbb{Z}/K\mathbb{Z}$ and for every initial distribution μ on $\mathbb{Z}/K\mathbb{Z}$. As a consequence, the QSD of $(Z_t)_{t\geq 0}$, denoted by ν_{qs}, is the stationary distribution of $(X_t)_{t\geq 0}$, which is the uniform distribution on $\mathbb{Z}/K\mathbb{Z}$, as we will prove in Theorem 3.1.1.

Recall that the total variation norm of a signed measure μ defined on a discrete probability space E is given by $\|\mu\|_{\text{TV}} = \frac{1}{2}\|\mu\|_1$ where $\|\mu\|_p = (\sum_{x\in E}|\mu(x)|^p)^{1/p}$ is the p-norm, see for instance [LP17, § 4.1]. If (f_N) and (g_N) are two real sequences, $f_N \underset{N\to\infty}{\sim} g_N$ means $f_N - g_N = o\,(g_N)$.

Now, assume that we have N particles with independent behavior driven by the generator $\mathcal{G}$, until one of them jumps to the absorbing state. When this happens, the particle instantaneously and uniformly jumps to the positions of one of the other $N-1$ particles. We denote by $(\eta_t^{(N)})_{t\geq 0}$ the Markov process tracking the positions of N particles in the K-cycle graph at time t. Consider the state space $\mathcal{E}_{K,N}$ of this process, which is given by

$$\mathcal{E}_{K,N} = \left\{ \eta : \mathbb{Z}/K\mathbb{Z} \to \mathbb{N}, \sum_{k=0}^{K-1} \eta(k) = N \right\}.$$

At time t the system is in state $\eta_t = (\eta_t(0), \eta_t(1), \ldots, \eta_t(K-1))$ if there are $\eta_t(k)$ particles on site k, for $k = 0, 1, \ldots, K-1$. Note that the cardinality of $\mathcal{E}_{K,N}$ is equal to that of the set of nonnegative solutions of the integer equation $x_1 + x_2 + \cdots + x_K = N$, which is $\text{card}\,(\mathcal{E}_{K,N}) = \binom{K+N-1}{N}$, see e.g. [Com74, Thm. D, § 1.7].

The generator of the N-particle process $(\eta_t^{(N)})_{t\geq 0}$, denoted by $\mathcal{L}_{K,N}$, applied to a function f on $\mathcal{E}_{K,N}$ reads

$$(\mathcal{L}_{K,N}f)(\eta) = \sum_{i,j\in\mathbb{Z}/K\mathbb{Z}} \eta(i) \left(\mathbb{1}_{\{j=i+1\}} + \theta \mathbb{1}_{\{j=i-1\}} + p\frac{\eta(j)}{N-1} \right) [f(T_{i\to j}\eta) - f(\eta)], \tag{3.2}$$

where $\theta, p > 0$ and for every $\eta \in \mathcal{E}_{K,N}$ satisfying $\eta(i) > 0$, the configuration $T_{i\to j}\eta$ is defined as $T_{i\to j}\eta = \eta - \mathbf{e}_i + \mathbf{e}_j$, where $\mathbf{e}_i$ is the i-th canonical vector of $\mathbb{R}^K$. Under these dynamics, each of the N particles, no matter where it is, can jump to every site $j \in \mathbb{Z}/K\mathbb{Z}$ such that $\eta(j) > 0$. Note that the process $(\eta_t^{(N)})_{t\geq 0}$ is irreducible. Consequently, it has a unique stationary distribution denoted by ν_N.

For every $\eta \in \mathcal{E}_{K,N}$ the empirical distribution $m(\eta)$ associated to the configuration η is defined by

$$m(\eta) = \frac{1}{N}\sum_{k=0}^{K-1} \eta(k)\delta_{\{k\}},$$

where $\delta_{\{k\}}$ is the Dirac distribution at $k \in \mathbb{Z}/K\mathbb{Z}$.

The (random) empirical distribution $m(\eta_t^{(N)})$ approximates the QSD of the process $(Z_t)_{t\geq 0}$ (cf. [AFG11; FM07; Vil14]) which due to Theorem 3.1.1 below is the uniform distribution. We are interested in studying how fast $m(\eta_t^{(N)})$ converges to the uniform distribution on $\mathbb{Z}/K\mathbb{Z}$ when both t and N tend to infinity. Let us denote by $\eta_\infty^{(N)}$ a random variable with distribution ν_N, the stationary distribution of the process $(\eta_t^{(N)})_{t\geq 0}$. In this work we develop a similar analysis to that of the complete graph dynamics in [CT16a]. We focus on the convergences when both N and t tend to infinity, as shown in the following diagram

$$\begin{array}{ccc} m(\eta_t^{(N)}) & \xrightarrow[t\to\infty]{} & m(\eta_\infty^{(N)}) \\ {\scriptstyle N}\big\downarrow & & \big\downarrow{\scriptstyle N} \\ \mathcal{L}(Z_t \mid t < \tau_p) & \xrightarrow[t\to\infty]{} & \nu_{\mathrm{qs}} \end{array}$$

where the limits are in distribution. Theorem 3.1.1 provides lower and upper exponential bounds for the speed of convergence of $\mathcal{L}(Z_t \mid t < \tau_p)$ to ν_{qs} in the 2-norm, when $t \to \infty$. Likewise, Corollary 3.1.7 and Theorem 3.1.9 give bounds for the speed of convergence of $m(\eta_t^{(N)})$ to $\mathcal{L}(Z_t \mid t < \tau_p)$ and $m(\eta_\infty^{(N)})$ to ν_{qs}, when $N \to \infty$.

The quantitative long time behavior of the N-particle system in countable state spaces is studied in [CT16b]. Using a coupling technique and under certain conditions, an exponential bound is provided for the convergence of $\mathcal{L}(\eta_t^{(N)})$ to ν_N in the sense of a *Wasserstein distance* [CT16b, Thm. 1.1]. In particular, the parameter λ defined by (3.1) needs to be positive. As we said, this is not the case for the asymmetric random walk on the K-cycle graph with uniform killing, when $K \geq 6$. A study of this convergence can be carried out using the spectrum of the generator $\mathcal{L}_{K,N}$, which is obtained in the recent paper [Cor21b]. Indeed, using Example 3 in [Cor21b] we can get the following asymptotic expression for the profile of the convergence in total variation distance to stationarity:

$$\max_{\eta\in\mathcal{E}_{K,N}} \left\|\mathcal{L}_\eta\left(\eta_t^{(N)}\right) - \nu_N\right\|_{\mathrm{TV}} = \mathcal{O}\left(\mathrm{e}^{-\rho_K t}\right),$$

where $\rho_K = 2(1+\theta)\sin^2(\pi/K)$, $\mathcal{L}_\eta\left(\eta_t^{(N)}\right)$ stands for the law of the process generated by $\mathcal{L}_{K,N}$ at time t and with initial distribution concentrated at $\eta \in \mathcal{E}_{K,N}$, and where for a real positive function f we denote by $\mathcal{O}(f)$ another real positive function such that

$$C_1 f(t) \leq \mathcal{O}(f)(t) \leq C_2 f(t),$$

for two constants $0 < C_1 \leq C_2 < \infty$ and for all $t \geq T$, for $T > 0$ large enough. It would be interesting to get non asymptotic results, with explicit constants, for the speed of convergence of the process generated by $\mathcal{L}_{K,N}$ to stationarity. In order to do that, one possible alternative is to use the results in the recent paper of Villemonais [Vil20], for a suitable distance, to get upper bounds for the speed of convergence in the sense of a *Wasserstein distance*. In addition, the recent work of Hermon and Salez [HS19] offers clues to an alternative method for solving this problem: control the *Dirichlet form* of the Fleming−Viot process in terms of the *Dirichlet form* of a single particle. Moreover, it remains as an open question the study of the existence of a *cutoff phenomenon* when the number of particles N tends towards infinity. These are possible directions for future research.

3.1.2 Main results

We first prove that the uniform distribution on $\mathbb{Z}/K\mathbb{Z}$ is the QSD of $(Z_t)_{t\geq 0}$. We also establish exponential bounds in the 2-distance and the total variation distance between the distribution of this process at time t and its QSD.

Let us denote by $\mathcal{L}_\nu(Z_t \mid t < \tau_p)$ the distribution at time t of the asymmetric random walk on the cycle graph with initial distribution ν on $\mathbb{Z}/K\mathbb{Z}$ and conditioned to non-absorption up to time t. Let us denote by φ_ν the characteristic function of a distribution ν on $\mathbb{Z}/K\mathbb{Z}$, which satisfies

$$\varphi_\nu(t) = \mathbb{E}_\nu\left[\mathrm{e}^{\mathrm{i}tX}\right] = \sum_{k=0}^{K-1} \nu(k)\mathrm{e}^{\mathrm{i}tk},$$

for all $t \geq 0$ [Dur19, § 3.3]. Note that

$$\varphi_{\nu_{\mathrm{qs}}}(t) = \frac{1-\mathrm{e}^{\mathrm{i}tK}}{K(1-\mathrm{e}^{\mathrm{i}t})},$$

for all $t \geq 0$. Let us denote by $\mathrm{D}_2(t)$ and $\mathrm{D}_{\mathrm{TV}}(t)$ the maximum distances to stationarity in the 2-distance and in total variation at time t, respectively, which are defined as follows:

$$\begin{aligned}
\mathrm{D}_2(t) &= \max_\nu \left\|\mathcal{L}_\nu(Z_t \mid t < \tau_p) - \nu_{\mathrm{qs}}\right\|_2, \\
\mathrm{D}_{\mathrm{TV}}(t) &= \max_\nu \left\|\mathcal{L}_\nu(Z_t \mid t < \tau_p) - \nu_{\mathrm{qs}}\right\|_{\mathrm{TV}},
\end{aligned}$$

where the maximum runs over all possible initial distributions ν on $\mathbb{Z}/K\mathbb{Z}$. Since $\mathbb{Z}/K\mathbb{Z}$ is finite, we know that the convergence of $\mathcal{L}_\nu(Z_t \mid t < \tau_p)$ to ν_{qs} is exponential [DS67]. The following theorem gives exponential lower and upper bounds for this convergence.

Theorem 3.1.1 (Convergence in 2-distance and total variation distance)**.** *The QSD of the process $(Z_t)_{t\geq 0}$, ν_{qs}, is the uniform distribution on $\mathbb{Z}/K\mathbb{Z}$. Also, denoting*

$$\Delta_t(\mu,\nu) = \mathcal{L}_\nu(Z_t \mid t < \tau_p) - \mathcal{L}_\mu(Z_t \mid t < \tau_p),$$

we have, for for every initial distributions ν and μ on $\mathbb{Z}/K\mathbb{Z}$ and every $t \geq 0$,

$$\left|\varphi_\nu\left(\frac{2\pi}{K}\right) - \varphi_\mu\left(\frac{2\pi}{K}\right)\right| \mathrm{e}^{-\rho_K t} \leq \left\|\Delta_t(\mu,\nu)\right\|_2 \leq \left\|\nu-\mu\right\|_2 \mathrm{e}^{-\rho_K t}, \tag{3.3}$$

$$\frac{\sqrt{K}}{2}\left|\varphi_\nu\left(\frac{2\pi}{K}\right) - \varphi_\mu\left(\frac{2\pi}{K}\right)\right| \mathrm{e}^{-\rho_K t} \leq \left\|\Delta_t(\mu,\nu)\right\|_{\mathrm{TV}} \leq \frac{\sqrt{K}}{2}\left\|\nu-\mu\right\|_2 \mathrm{e}^{-\rho_K t}, \tag{3.4}$$

where

$$\rho_K = 2(1+\theta)\sin^2\left(\frac{\pi}{K}\right). \tag{3.5}$$

Furthermore, the convergence of $\mathcal{L}_\nu(Z_t \mid t < \tau_p)$ to ν_{qs} in the 2-distance and the total variation distance is exponential with rate $-\rho_K$. Indeed, for all $t \geq 0$,

$$\frac{1}{\sqrt{K}}\mathrm{e}^{-\rho_K t} \leq \mathrm{D}_2(t) \leq \sqrt{\frac{K-1}{K}}\mathrm{e}^{-\rho_K t}, \tag{3.6}$$

$$\frac{1}{2}\mathrm{e}^{-\rho_K t} \leq \mathrm{D}_{\mathrm{TV}}(t) \leq \frac{1}{2}\sqrt{K-1}\,\mathrm{e}^{-\rho_K t}. \tag{3.7}$$

In spite of its simplicity, we did not find this result in the literature. Therefore, for the sake of completeness, we provide a proof of this theorem in Section 3.2.

Consider the function $\phi : \mathcal{E}_{K,N} \to \mathcal{E}_{K,N}$ defined by

$$\phi(\eta_0, \eta_1, \ldots, \eta_{K-1}) = (\eta_1, \eta_2, \ldots, \eta_{K-1}, \eta_0) \tag{3.8}$$

and its l-composed $\phi^{(l)} = \phi \circ \phi \circ \cdots \circ \phi$ (l times) which acts on the cycle graph by rotating it l sites clockwise, for $l \in \{1, 2, \ldots, K-1\}$.

Even if the dynamics induced by $\mathcal{G}$ has some symmetry (in fact, it is symmetric when $\theta = 1$), we prove that $(\eta_t^{(N)})_{t\geq 0}$ is not reversible when $K \geq 4$ or when $K = 3$ and $\theta \neq 1$. However, we show that the stationary distribution of the N-particle process is rotation invariant. Using this invariance, we calculate the mean of the proportion of particles in each state under the stationary distribution.

Theorem 3.1.2 (Non-reversibility and rotation invariance)**.** *The N-particle system with generator given by (3.2) has the following properties*

a) It is not reversible, except when $K = 3$ and $\theta = 1$.

b) Its stationary distribution, denoted by ν_N, is invariant under rotations, i.e.

$$\nu_N = \nu_N \circ \phi^{(l)}, \quad l \in \{1, 2, \ldots, K-1\}.$$

c) Under the stationary dynamics, the empirical distribution of the N-particle system is an unbiased estimator of the QSD *of $(Z_t)_{t\geq 0}$, i.e.*

$$\mathbb{E}_{\nu_N}\left[\frac{\eta(k)}{N}\right] = \frac{1}{K}, \quad k \in \mathbb{Z}/K\mathbb{Z}.$$

Theorem 3.1.2 is proved in Section 3.3. Using parts b) and c) of Theorem 3.1.2, the following result is immediate.

Corollary 3.1.3 (Cyclic symmetry)**.** *For every $K \geq 3$ we have*

$$\mathrm{Cov}_{\nu_N}\left[\frac{\eta(0)}{N}, \frac{\eta(k)}{N}\right] = \mathrm{Cov}_{\nu_N}\left[\frac{\eta(0)}{N}, \frac{\eta(K-k)}{N}\right], \quad k \in \mathbb{Z}/K\mathbb{Z}.$$

Let T_n and U_n be the n-th degree Chebyshev polynomials of the first and the second kind, respectively, for $n \geq 1$. We recall that polynomials $(T_n)_{n\geq 0}$ and $(U_n)_{n\geq 0}$ satisfy both the recurrence relation

$$p_{n+1}(x) = 2x\, p_n(x) - p_{n-1}(x), \text{ for all } n \geq 1, \tag{3.9}$$

with initial conditions $T_0(x) = U_0(x) = 1$, $T_1(x) = x$ and $U_1(x) = 2x$, see e.g. [MH03]. We also extend the definition of the Chebyshev polynomials of the second kind for $n = -1$, by putting $U_{-1}(x) = 0$.

The following theorem provides explicit expressions for $\mathrm{Cov}_{\nu_N}[\eta(0)/N, \eta(k)/N]$ in terms of the Chebyshev polynomials of the first and the second kind, for $k \in \{0, 1, \ldots, K-1\}$ and the constant β_N, defined by

$$\beta_N = 2\left(1 + \frac{p}{(N-1)(1+\theta)}\right). \tag{3.10}$$

Theorem 3.1.4 (Explicit expressions for the covariances)**.** *We have*

- *If $K = 2K_2$, $K_2 \geq 2$,*

$$\mathrm{Var}_{\nu_N}\left[\frac{\eta(0)}{N}\right] = \frac{N-1}{KN}\frac{2}{\beta_N+2}\frac{T_{K_2}(\beta_N/2)}{U_{K_2-1}(\beta_N/2)} + \frac{1}{KN} - \frac{1}{K^2}, \tag{3.11}$$

$$\mathrm{Cov}_{\nu_N}\left[\frac{\eta(0)}{N}, \frac{\eta(k)}{N}\right] = \frac{N-1}{KN}\frac{2}{\beta_N+2}\frac{T_{K_2-k}(\beta_N/2)}{U_{K_2-1}(\beta_N/2)} - \frac{1}{K^2}, \tag{3.12}$$

for all $1 \leq k \leq K_2 - 1$.

- *If $K = 2K_2 + 1$, $K_2 \geq 1$,*

$$\mathrm{Var}_{\nu_N}\left[\frac{\eta(0)}{N}\right] = \frac{N-1}{KN}\frac{U_{K_2}(\beta_N/2) - U_{K_2-1}(\beta_N/2)}{U_{K_2}(\beta_N/2) + U_{K_2-1}(\beta_N/2)} + \frac{1}{KN} - \frac{1}{K^2}, \tag{3.13}$$

$$\mathrm{Cov}_{\nu_N}\left[\frac{\eta(0)}{N}, \frac{\eta(k)}{N}\right] = \frac{N-1}{KN}\frac{U_{K_2-k}(\beta_N/2) - U_{K_2-k-1}(\beta_N/2)}{U_{K_2}(\beta_N/2) + U_{K_2-1}(\beta_N/2)} - \frac{1}{K^2}, \tag{3.14}$$

for all $1 \leq k \leq K_2$.

Theorem 3.1.4 is proved in Section 3.3.2. Using previous result it is possible to show that the covariance between the proportions of particles under the stationary distribution in two different states decreases as a function of the graph distance between the states.

Corollary 3.1.5 (Geometry of the cycle graph and covariances)**.** *The covariance between two states under the stationary measure, ν_N, is decreasing as a function of the graph distance between these states, i.e. for all $k = 0, 1, \ldots, \lfloor \frac{K}{2} \rfloor - 1$ we have*

$$\mathrm{Cov}_{\nu_N}\left[\frac{\eta(0)}{N}, \frac{\eta(k)}{N}\right] \geq \mathrm{Cov}_{\nu_N}\left[\frac{\eta(0)}{N}, \frac{\eta(k+1)}{N}\right].$$

With the aim of proving the convergence of the proportion of particles in each state to $1/K$, we study the behavior of $\mathrm{Var}_{\nu_N}[\eta(0)/N]$ as a function of $1/N$ when N tends to infinity. Theorem 2 in [AFG11] states that these variances vanishes when N goes to infinity. We thus focus on the speed of this convergence. For this purpose, we find the asymptotic development of second order for $\mathrm{Cov}_{\nu_N}[\eta(0)/N, \eta(k)/N]$ as a function of $1/N$ when N tends to infinity, for $k \in \mathbb{Z}/K\mathbb{Z}$.

Theorem 3.1.6 (Asymptotic development of two-particle covariances)**.** *The asymptotic series expansion of order 2 when $N \to +\infty$ of $\mathrm{Cov}_{\nu_N}\left[\frac{\eta(0)}{N}, \frac{\eta(k)}{N}\right]$, for $k \in \mathbb{Z}/K\mathbb{Z}$, is given by*

$$\begin{aligned}\mathrm{Cov}_{\nu_N}\left[\frac{\eta(0)}{N}, \frac{\eta(k)}{N}\right] = &\frac{1}{KN}\left(\mathbb{1}_{\{k=0\}} - \frac{1}{K} + \frac{6k(k-K)+K^2-1}{6K}\frac{p}{1+\theta}\right) \\ &+ \frac{1}{K^2N^2}\frac{30k(K-k)[k(K-k)+2]-(K^2-1)(K^2+11)}{180}\left(\frac{p}{1+\theta}\right)^2 + o\left(\frac{1}{N^2}\right).\end{aligned} \tag{3.15}$$

The following result provides a bound for the speed of convergence, of the empirical distribution induced by the N-particle system, to the QSD when $N \to \infty$.

Corollary 3.1.7 (Convergence to the QSD)**.** *We have*

$$\mathbb{E}_{\nu_N}\left[\, \|m(\eta) - \nu_{\mathrm{qs}}\|_2 \,\right] \leq \sqrt{\frac{K-1}{N}}\sqrt{1 + \frac{p(K+1)}{6(1+\theta)}} + o\left(\frac{1}{\sqrt{N}}\right). \tag{3.16}$$

Theorem 3.1.6 and Corollary 3.1.7 are proved in Section 3.3.3. In particular, Corollary 3.1.7 implies the convergence at rate $1/\sqrt{N}$ under the stationary distribution of $m(\eta)$ towards the uniform distributions, when $N \to \infty$. Cloez and Thai [CT16a, Cor. 2.10] provide the same rate of convergence for the Fleming – Viot process in the K-complete graph. Moreover, Champagnat and Villemonais [CV21, Thm. 2.3] provide a general rate of convergence $1/N^{\alpha}$, with $\alpha \in (0, 1/2]$. However, as soon as the selection rate is not null, i.e. $p > 0$, one has $\alpha < 1/2$, which is actually not the optimal rate for the asymmetric random walk, killed at a uniform rate, studied in this chapter. To the best of our knowledge, there are no general results on Fleming – Viot process in discrete spaces assuring the rate of convergence $1/\sqrt{N}$, under the stationary distribution, of the empirical distribution to the QSD.

Finally, in Section 3.4 we study the convergence of the empirical distribution, $m(\eta_t)$, to the quasi-stationary distribution of $(Z_t)_{t\geq 0}$ when t tends to infinity. Let us denote by $\overline{m}(\eta_t^{(N)})$ the *mean empirical measure* induced by the N-particle process at time t, defined by

$$\overline{m}(\eta_t^{(N)})(k) = \mathbb{E}[m(\eta_t^{(N)})(k)] = \mathbb{E}[\eta_t^{(N)}(k)/N].$$

Using Theorem 3.1.1 we can prove the following two theorems.

Theorem 3.1.8 (Mean empirical distribution)**.** *Consider $\eta \in \mathcal{E}_{K,N}$ and $(\eta_t^{(N)})_{t\geq 0}$ the N-particle process with initial distribution concentrated at η. We have*

$$\overline{m}(\eta_t^{(N)}) = \mathcal{L}_{m(\eta)}(Z_t \mid t < \tau_p).$$

Furthermore, for every probability measure ν on $\mathbb{Z}/K\mathbb{Z}$ we obtain

$$\left|\varphi_{m(\eta)}\left(\frac{2\pi}{K}\right) - \varphi_{\nu}\left(\frac{2\pi}{K}\right)\right| \mathrm{e}^{-\rho_K t} \leq \left\|\overline{m}(\eta_t^{(N)}) - \mathcal{L}_{\nu}(Z_t \mid t < \tau_p)\right\|_2 \leq \|m(\eta) - \nu\|_2 \, \mathrm{e}^{-\rho_K t}, \tag{3.17}$$

where ρ_K are defined by (3.5), and $\varphi_{m(\eta)}$ and φ_ν denote the characteristic functions associated to the distributions $m(\eta)$ and ν, respectively.

Thus, the proportion of particles in each state is an unbiased estimator of the distribution of the conditioned process for all $t \geq 0$. Using [FM07, Thm. 1.2] we know that the variance of the proportion of particles in each state at time $t \geq 0$ vanishes when N goes to infinity, for every $t \geq 0$. The following result provides a bound for this convergence.

Theorem 3.1.9 (Convergence to the Conditioned Process)**.** *We have the following uniform upper bound for the variance of the proportion of particles in each state*

$$\max_{\substack{\eta\in\mathcal{E}_{K,N}\\ k\in\mathbb{Z}/K\mathbb{Z}}} \left|\mathrm{Var}_\eta\left[\frac{\eta_t^{(N)}(k)}{N}\right] - \mathrm{Var}_{\nu_N}\left[\frac{\eta(k)}{N}\right]\right| \leq C_{K,N}\frac{\mathrm{e}^{-p_N t} - \mathrm{e}^{-\rho_K t}}{\rho_K - p_N} + \mathrm{e}^{-p_N t}\,\mathrm{Var}_{\nu_N}\left[\frac{\eta(0)}{N}\right], \tag{3.18}$$

where ρ_K is given by (3.5) and

$$p_N = \frac{2p}{N-1}, \tag{3.19}$$

$$C_{K,N} = \frac{2}{N}\left(1 + \theta + \frac{p}{N-1} + \frac{pN(K+1)\sqrt{K-1}}{K\sqrt{K}(N-1)}\right). \tag{3.20}$$

Furthermore,

$$\begin{aligned}\left|\varphi_{m(\eta)}(t) - \varphi_\nu(t)\right|\mathrm{e}^{-\rho_K t} &\leq \mathbb{E}_\eta\left[\left\|m\left(\eta_t^{(N)}\right) - \mathcal{L}_\nu(Z_t \mid t \leq \tau_p)\right\|_2\right] \\ &\leq \sqrt{\frac{K}{N}}\left(D_K\frac{1-\mathrm{e}^{-\rho_K t}}{\rho_K} + E_K\right)^{1/2} + \mathrm{e}^{-\rho_K t}\|m(\eta) - \nu\|_2 + o\left(\frac{1}{\sqrt{N}}\right), \end{aligned} \tag{3.21}$$

for every $\eta \in \mathcal{E}_{K,N}$ and every initial distribution ν on $\mathbb{Z}/K\mathbb{Z}$, where ρ_K is given by (3.5), and

$$D_K = 2\left(1+\theta+\frac{p(K+1)\sqrt{K-1}}{K\sqrt{K}}\right), \quad E_K = \frac{K-1}{K^2} + \frac{K^2-1}{6K^2(1+\theta)}. \tag{3.22}$$

Theorems 3.1.8 and 3.1.9 is proved in Section 3.4. Similar results are proved in [CT16a] for the Fleming – Viot process on the complete graph and for the two-point process.

Remark 3.1.1 (Uniform bound). Note that the bound given by (3.18) tends exponentially towards zero when $t \to \infty$. In particular, the right side of (3.18) is bounded in t and can be used to obtain a uniform bound for the variance of the proportion of particles in each state of order $1/N$. Namely, using (3.18) and the inequality $(\mathrm{e}^{-p_N t} - \mathrm{e}^{-\rho_K t})/(\rho_K - p_N) \leq 1/\max(\rho_K, p_N)$, we obtain

$$\sup_{t\geq 0} \max_{\substack{\eta\in\mathcal{E}_{K,N}\\ k\in\mathbb{Z}/K\mathbb{Z}}} \operatorname{Var}_\eta\left[\frac{\eta_t^{(N)}(k)}{N}\right] \leq \frac{C_{K,N}}{\max(\rho_K, p_N)} + 2\operatorname{Var}_{\nu_N}\left[\frac{\eta(0)}{N}\right] = \left(\frac{D_K}{\rho_K} + 2E_K\right)\frac{1}{N} + o\left(\frac{1}{N}\right),$$

where ρ_K, p_N, $C_{K,N}$ and D_k and E_k, are given by (3.5), (3.19), (3.20) and (3.22), respectively.

Similar bounds are obtained for the convergence to the conditional distribution for Fleming – Viot process in discrete state spaces, see e.g. [DM00a, Thm. 1.1] and [Vil14, Thm. 2.2]. However, these results are not uniform in $t \geq 0$. Corollary 1.5 in [CT16b] does provide a uniform bound under certain conditions of order $1/N^\gamma$, with $\gamma < 1/2$, for the 1-distance between the empirical law associated to the Fleming – Viot process at time t and the law of the conditioned process. However, this result does not hold for the Fleming – Viot process on the K-cycle graph we study here, for $K \geq 6$, since the parameter λ given by (3.1) is null.

The rest of this chapter is organized as follows. Section 3.2 gives the proof of Theorem 3.1.1. In Section 3.3 we study the covariances of the proportions of particles in each state under the stationary distribution, and we thus prove Theorems 3.1.2, 3.1.4 and 3.1.6. Finally, Section 3.4 is devoted to the proof of Theorems 3.1.8 and 3.1.9 related to the variance of the proportion of particles in each site at a given time $t \geq 0$.

3.2 The asymmetric random walk on the cycle graph

We first prove that the QSD of $(Z_t)_{t\geq 0}$, denoted by ν_{qs}, which is the stationary distribution of $(X_t)_{t\geq 0}$, is the uniform distribution on $\mathbb{Z}/K\mathbb{Z}$. We also provide exponential bounds for the speed of convergence in the 2-distance and the total variation distance of $\mathcal{L}_\nu(Z_t \mid t < \tau_p)$ to ν_{qs}.

Recall that a square matrix C is called circulant if it takes the form

$$C = \begin{pmatrix} c_0 & c_1 & \cdots & c_{n-2} & c_{n-1} \\ c_{n-1} & c_0 & \ddots & c_{n-3} & c_{n-2} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ c_2 & c_3 & \ddots & c_0 & c_1 \\ c_1 & c_2 & \cdots & c_{n-1} & c_0 \end{pmatrix}. \tag{3.23}$$

It is evident that a circulant matrix is completely determined by its first row, therefore we will denote a circulant matrix with the form given by (3.23) by $C = \operatorname{circ}(c_0, c_1, \ldots, c_{n-1})$.

Let Q be the infinitesimal generator matrix of the process $(X_t)_{t\geq 0}$. Then, Q is circulant and it satisfies

$$Q = \operatorname{circ}(-(1+\theta), 1, 0, \ldots, 0, \theta). \tag{3.24}$$

Let us also denote by i the complex root of -1. Since the matrix Q is circulant, its spectrum is explicitly known, as follows in the next lemma.

Lemma 3.2.1 (Spectrum of Q). *The matrix Q satisfies $Q = F_K \Lambda F_K^\star$, where*

- *F_K is the K-dimensional Fourier matrix, i.e. the unitary matrix defined by*

$$[F_K]_{r,c} = \frac{1}{\sqrt{K}} (\omega_K)^{-r\,c}, \tag{3.25}$$

for each $r, c \in \{0, 1, \ldots, K-1\}$, where $\omega_K = \mathrm{e}^{\mathrm{i}\frac{2\pi}{K}}$,

- *$F_K^\star$ is the conjugate of F_K (and also its inverse because F_K is unitary and symmetric),*
- *Λ is the $K \times K$ diagonal matrix with $[\Lambda]_{k,k} = \lambda_k$, for all $0 \leq k \leq K-1$, where*

$$\lambda_k = -(1+\theta)\sin^2\left(\frac{\pi k}{K}\right) + \mathrm{i}(1-\theta)\sin\left(\frac{2\pi k}{K}\right),$$

for $k = 0, 1, \ldots, K-1$.

Proof of Lemma 3.2.1. Let us define the polynomial $p_Q : s \mapsto -(1+\theta) + s + \theta s^{K-1}$. Since Q is a circulant matrix, we can use [Dav79, Thm. 3.2.2] to diagonalize Q in the following way

$$Q = F_K \operatorname{Diag}(\lambda_0, \lambda_1, \ldots, \lambda_{K-1}) F_K^\star,$$

where F_K is the Fourier matrix defined by (3.25) and

$$\begin{aligned}
\lambda_k = p_Q(\mathrm{e}^{\mathrm{i}\frac{2k\pi}{K}}) &= -(1+\theta) + \mathrm{e}^{\mathrm{i}\frac{2k\pi}{K}} + \theta\left(\mathrm{e}^{\mathrm{i}\frac{2k\pi}{K}}\right)^{K-1} \\
&= -(1+\theta)\left[1 - \cos\left(\frac{2\pi k}{K}\right)\right] + \mathrm{i}(1-\theta)\sin\left(\frac{2\pi k}{K}\right) \\
&= -2(1+\theta)\sin^2\left(\frac{\pi k}{K}\right) + \mathrm{i}(1-\theta)\sin\left(\frac{2\pi k}{K}\right),
\end{aligned}$$

for $k = 0, 1, \ldots, K-1$. □

Remark 3.2.1 (Eigenvalues of Q). Note that $\frac{[\Re(\lambda_k)+(1+\theta)]^2}{(1+\theta)^2} + \frac{[\Im(\lambda_k)]^2}{(1-\theta)^2} = 1$, for all $\theta \neq 1$, where $\Re(\lambda_k)$ and $\Im(\lambda_k)$ are the real and the imaginary parts of λ_k, respectively, for $k = 0, 1, \ldots, K-1$. Thus, all the eigenvalues λ_k are on the ellipse with center $(0, -(1+\theta))$ and equation

$$\frac{(x+1+\theta)^2}{(1+\theta)^2} + \frac{y^2}{(1-\theta)^2} = 1.$$

Of course, for $\theta = 1$, since the matrix Q is symmetric, all the eigenvalues are real.

Also, the *second largest eigenvalue in modulus* (SLEM) of Q, denoted by ρ_K, is given by (3.5) and it is reached for $-\Re(\lambda_1)$ and $-\Re(\lambda_{K-1})$. The minimum of $\Re(\lambda_k)$ is reached for $\Re(\lambda_{K/2})$, if K is even, and for $\Re(\lambda_{(K-1)/2})$ and $\Re(\lambda_{(K+1)/2})$, if K is odd.

3.2.1 Proof of Theorem 3.1.1

Proof of Theorem 3.1.1. We know that $Q = F_K \Lambda F_K^\star$. Therefore $\mathrm{e}^{tQ} = F_K \mathrm{e}^{t\Lambda} F_K^\star$, and it follows that

$$\mathrm{e}^{tQ} = \sum_{k=0}^{K-1} \mathrm{e}^{\lambda_k t} F_K U_k F_K^\star = \sum_{k=0}^{K-1} \mathrm{e}^{\lambda_k t} \Omega_k,$$

where U_k, $0 \le k \le K-1$, is the $K \times K$ matrix with $[U_k]_{k,k} = 1$ and 0 elsewhere, and Ω_k is defined as $\Omega_k = F_K U_k F_K^\star$. In fact, Ω_k is the symmetric circulant matrix satisfying $[\Omega_k]_{r,c} = \frac{1}{K}\omega^{k(r-c)}$, for all $0 \le r, c \le K-1$ and for every $k \in \{0, 1, \ldots, K-1\}$. In particular $[\Omega_0]_{r,c} = \frac{1}{K}$ for all $0 \le r, c \le K-1$, and $\Omega_k\, \Omega_l = \mathbf{0}$, for all $k \neq l$. Then, for two probability measures μ and ν on $\{0, 1, \ldots, K-1\}$ we have

$$(\mu - \nu)\Omega_0 = \mathbf{0} \tag{3.26}$$

and therefore

$$(\mu - \nu)\mathrm{e}^{tQ} = \sum_{k=1}^{K-1} \mathrm{e}^{\lambda_k t} (\mu - \nu)\Omega_k. \tag{3.27}$$

Let us denote by $\langle \cdot, \cdot \rangle$ the usual inner product in $\mathbb{C}$ and for a matrix A let us denote by A^T its transpose. Note that for every K-dimensional vector $\mathbf{x}$ and $k \neq l$ we have

$$\langle \mathbf{x}\,\Omega_k,\ \mathbf{x}\,\Omega_l \rangle = \mathbf{x}\,\Omega_k \left[(\Omega_l^T)^\star \right] (\mathbf{x}^\star)^T = \mathbf{x}\,\Omega_k \Omega_l\, (\mathbf{x}^\star)^T = 0.$$

Thus, the set of vectors $(\mathbf{x}\Omega_k)_{k=1}^{K-1}$ are orthogonal in $(\mathbb{C}, \langle \cdot, \cdot \rangle)$. Now, using (3.27) and Pythagoras' theorem we have

$$\begin{aligned} \left\| (\mu - \nu)\mathrm{e}^{tQ} \right\|_2^2 &= \sum_{k=1}^{K-1} \left\| \mathrm{e}^{\lambda_k t} (\mu - \nu)\Omega_k \right\|_2^2 \\ &= \sum_{k=1}^{K-1} \mathrm{e}^{2\Re(\lambda_k)t} \left\| (\mu - \nu)\Omega_k \right\|_2^2. \end{aligned}$$

Since $\rho_K = -\max_{k=1,\ldots,K-1} \Re(\lambda_k)$ we obtain

$$\begin{aligned} \left\| (\mu - \nu)\mathrm{e}^{tQ} \right\|_2^2 &\le \mathrm{e}^{-2\rho_K t} \sum_{k=1}^{K-1} \left\| (\mu - \nu)\Omega_k \right\|_2^2 \\ &= \mathrm{e}^{-2\rho_K t} \left\| \sum_{k=1}^{K-1} (\mu - \nu)\Omega_k \right\|_2^2 \\ &= \mathrm{e}^{-2\rho_K t} \left\| \sum_{k=0}^{K-1} (\mu - \nu)\Omega_k \right\|_2^2 \\ &= \mathrm{e}^{-2\rho_K t} \|\mu - \nu\|_2^2. \end{aligned}$$

Note that the first equality holds due the Pythagoras' theorem, the second one uses (3.26) and the last one uses the fact that

$$\sum_{k=0}^{K-1} (\mu - \nu)\Omega_k = \mu - \nu.$$

Note that the upper bound in (3.4) is proved using the *Cauchy – Schwarz inequality*, which implies

$$\|\Delta_t(\mu, \nu)\|_{\mathrm{TV}} \le \frac{\sqrt{K}}{2} \|\Delta_t(\mu, \nu)\|_2\,,$$

where $\Delta_t(\mu, \nu)$ is as defined in the statement of Theorem 3.1.1, and the inequality holds for every pair of distributions ν and μ on $\mathbb{Z}/K\mathbb{Z}$, and for all $t \geq 0$.

To prove the lower bounds in (3.3) and (3.4) we recall the the r-norm of a function f on $\mathbb{Z}/K\mathbb{Z}$, allows the following characterization:

$$\|f\|_r = \max_g \frac{|\langle f, g\rangle|}{\|g\|_q},$$

where $q \in [1, \infty]$ is the conjugate of $r \in [1, \infty]$, i.e. $1/r + 1/q = 1$, and the maximum runs over all the functions on $\mathbb{Z}/K\mathbb{Z}$. Now, take $g : k \in \mathbb{Z}/K\mathbb{Z} \mapsto \frac{1}{\sqrt{K}}(\omega_K)^k$ as a test function, where $\omega_K = \mathrm{e}^{\frac{2\pi}{K}\mathrm{i}}$. Note that viewed as a column vector, g is equal to the last column of the Fourier matrix F_K. Then, g is a right eigenfuntion of Q with associated eigenvalue $-\rho_K$. Moreover, $\|g\|_2 = 1$ and $\|g\|_\infty = 1/\sqrt{K}$. Therefore,

$$\|\nu \mathrm{e}^{tQ} - \mu \mathrm{e}^{tQ}\|_2 \geq \frac{|\langle \nu \mathrm{e}^{tQ} - \mu \mathrm{e}^{tQ}, g\rangle|}{\|g\|_2} = \left|\varphi_\nu\left(\frac{2\pi}{K}\right) - \varphi_\mu\left(\frac{2\pi}{K}\right)\right| \mathrm{e}^{-\rho_K t},$$

$$\left\|\nu \mathrm{e}^{tQ} - \mu \mathrm{e}^{tQ}\right\|_{\mathrm{TV}} \geq \frac{|\langle \nu \mathrm{e}^{tQ} - \mu \mathrm{e}^{tQ}, g\rangle|}{2\|g\|_\infty} = \frac{\sqrt{K}}{2}\left|\varphi_\nu\left(\frac{2\pi}{K}\right) - \varphi_\mu\left(\frac{2\pi}{K}\right)\right| \mathrm{e}^{-\rho_K t},$$

To prove (3.6) first note that the 2-distance and the total variation distances satisfy

$$\begin{aligned} \mathrm{D}_2(t) &= \max_{k \in \mathbb{Z}/K\mathbb{Z}} \|\mathcal{L}_k(Z_t \mid t < \tau_p) - \nu_{\mathrm{qs}}\|_2\,, \\ \mathrm{D}_{\mathrm{TV}}(t) &= \max_{k \in \mathbb{Z}/K\mathbb{Z}} \|\mathcal{L}_k(Z_t \mid t < \tau_p) - \nu_{\mathrm{qs}}\|_{\mathrm{TV}}\,, \end{aligned}$$

which is a consequence of the convexity of these distances. Thus, the upper bounds in expression (3.6) and (3.7) are consequence of the equality $\|\delta_k - \nu_{\mathrm{qs}}\|_2 = \sqrt{\frac{K-1}{K}}$. The lower bounds in (3.6) and (3.7) is obtained using that $\varphi_{\nu_{\mathrm{qs}}}(2\pi/K) = 0$ and $\varphi_{\delta_k}(2\pi/K) = |g(k)| = 1/\sqrt{K}$, for every $k \in \mathbb{Z}/K\mathbb{K}$. □

3.3 Covariances of the proportions of particles under the stationary distribution

The following lemma gives us informations about the invariance of the generator $\mathcal{L}_{K,N}$, defined in (3.2), by the rotation function ϕ defined in (3.8).

Lemma 3.3.1 (Rotation invariance of the generator)**.** *The generator $\mathcal{L}_{K,N}$ of $(\eta_t^{(N)})_{t\geq 0}$ satisfies*

$$\mathcal{L}_{K,N}\mathbb{1}_\eta = \mathcal{L}_{K,N}\mathbb{1}_{\phi(\eta)} \circ \phi, \tag{3.28}$$

for every $\eta \in \mathcal{E}_{K,N}$.

Proof. Note that

$$(\mathcal{L}_{K,N}\mathbb{1}_\eta)(\eta') = \eta'(i)\left(\mathbb{1}_{\{j=i+1\}} + \theta\mathbb{1}_{\{j=i-1\}} + p\frac{\eta'(j)}{N-1}\right), \tag{3.29}$$

if $\eta = T_{i\to j}\eta'$, for some $i, j \in \mathbb{Z}/K\mathbb{Z}$, and it is null otherwise. Now, if $\eta = T_{i\to j}\eta'$, then we have $\phi(\eta) = T_{(i+1)\to(j+1)}\phi(\eta')$. Thus,

$$\left(\mathcal{L}_{K,N}\mathbb{1}_{\phi(\eta)}\right)(\phi(\eta')) = \phi(\eta')(i+1)\left(\mathbb{1}_{\{j=i+1\}} + \theta\mathbb{1}_{\{j=i-1\}} + p\frac{\phi(\eta')(j+1)}{N-1}\right). \tag{3.30}$$

Using (3.29) and (3.30) we can see that (3.28) holds, since $\eta'(i) = \phi(\eta')(i+1)$ and $\eta(j) = \phi(\eta)(j+1)$. □

3.3.1 Proof of Theorem 3.1.2

We will now prove Theorem 3.1.2, which describes some properties of ν_N, the stationary distribution of the N-particle process $(\eta_t^{(N)})_{t\geq 0}$.

Proof of Theorem 3.1.2.

a) *The process $(\eta_t^{(N)})_{t\geq 0}$ is not reversible, except when $K = 3$ and $\theta = 1$.*

For $K = 3$ and $N \geq 2$, let us consider the three states in $\mathcal{E}_{3,N}$,

$$\eta_1 = [N, 0, 0], \quad \eta_2 = [N-1, 1, 0], \quad \eta_3 = [N-1, 0, 1].$$

It is straightforward to verify that

$$\begin{array}{lll} (\mathcal{L}_{K,N}\mathbb{1}_{\eta_2})(\eta_1) = N, & (\mathcal{L}_{K,N}\mathbb{1}_{\eta_3})(\eta_1) = N\theta, & (\mathcal{L}_{K,N}\mathbb{1}_{\eta_1})(\eta_2) = p+\theta, \\ (\mathcal{L}_{K,N}\mathbb{1}_{\eta_3})(\eta_2) = 1, & (\mathcal{L}_{K,N}\mathbb{1}_{\eta_1})(\eta_3) = p+1, & (\mathcal{L}_{K,N}\mathbb{1}_{\eta_2})(\eta_3) = \theta. \end{array}$$

Moreover,

$$\begin{aligned} (\mathcal{L}_{K,N}\mathbb{1}_{\eta_1})(\eta_3)\cdot(\mathcal{L}_{K,N}\mathbb{1}_{\eta_2})(\eta_1)\cdot(\mathcal{L}_{K,N}\mathbb{1}_{\eta_3})(\eta_2) &= (p+1)N, \\ (\mathcal{L}_{K,N}\mathbb{1}_{\eta_3})(\eta_1)\cdot(\mathcal{L}_{K,N}\mathbb{1}_{\eta_2})(\eta_3)\cdot(\mathcal{L}_{K,N}\mathbb{1}_{\eta_1})(\eta_2) &= N\theta^2(p+\theta), \end{aligned}$$

the Kolmogorov cycle reversibility criterion, see [Kel79, Thm. 1.8], is not satisfied unless $\theta = 1$. Indeed, note that a necessary condition to have reversibility is that the polynomial

$$\alpha(\theta) = \theta^3 + p(N-1)\theta^2 - p(N-1) - 1 = (\theta-1)[\theta^2 + (\theta+1)(p+1)]$$

is equal to zero. Now, since $\theta^2 + (\theta+1)(p+1) > 0$ for all $\theta \geq 0$, the polynomial $\alpha(\theta)$ only has one positive root, which is $\theta = 1$.

For $K \geq 4$, $N \geq 2$ and $p > 0$, let us consider the two states in $\mathcal{E}_{K,N}$: $\eta_1 = [N, 0, \dots, 0]$ and $\eta_2 = [N-1, 0, 1, \dots, 0]$. Because $(\mathcal{L}_{K,N}\mathbb{1}_{\eta_2})(\eta_1) = 0$ and $(\mathcal{L}_{K,N}\mathbb{1}_{\eta_1})(\eta_2) = p \neq 0$, the detailed balanced property for a reversible process, see [Kel79, Thm. 1.3], $\nu_N(\eta_1)(\mathcal{L}_{K,N}\mathbb{1}_{\eta_2})(\eta_1) = \nu_N(\eta_2)(\mathcal{L}_{K,N}\mathbb{1}_{\eta_1})(\eta_2)$, is not satisfied.

Therefore, *a*) is proved except in the special case $K = 3$, $N \geq 2$ and $\theta = 1$. Note that in this case the model is a complete graph model, which was proved to be reversible in [CT16a, Thm. 2.4].

b) *The stationary distribution ν_N is invariant by rotation.*

Since ν_N is the unique stationary distribution of $(\eta_t^{(N)})_{t\geq 0}$, we know that $\nu_N(\mathcal{L}_{K,N}f) = 0$ for every function f on $\mathcal{E}_{K,N}$. Thus, in order to prove that ν_N is invariant by rotation, it is sufficient to prove that $\nu_N \circ \phi$ also satisfies $(\nu_N \circ \phi)(\mathcal{L}_{K,N}f) = 0$ for every function f on $\mathcal{E}_{K,N}$. Since $\mathcal{E}_{K,N}$ is finite, it is enough to consider the indicator functions $\mathbb{1}_\eta$, for every $\eta \in \mathcal{E}_{K,N}$. Using Lemma 3.3.1, we have

$$(\nu_N \circ \phi)(\mathcal{L}_{K,N}\mathbb{1}_\eta) = \nu_N\left(\mathcal{L}_{K,N}\mathbb{1}_\eta \circ \phi^{-1}\right) = \nu_N\left(\mathcal{L}_{K,N}\mathbb{1}_{\phi(\eta)}\right) = 0,$$

for every $\eta \in \mathcal{E}_{K,N}$, where the second equality holds due to (3.28) and the third is due to the fact that ν_N is stationary for $\mathcal{L}_{K,N}$. Consequently, by the uniqueness of the stationary distribution, we have $\nu_N = \nu_N \circ \phi$. The result trivially holds for any rotation $\phi^{(l)}$, $l \geq 1$.

c) *Mean of the proportion of particles in each state.*

Using part b) we have $\mathbb{E}_{\nu_N}[\eta(0)] = \mathbb{E}_{\nu_N}[\phi^{(k)}(\eta)(0)] = \mathbb{E}_{\nu_N}[\eta(k)]$, for all $k = 0, 1, \ldots, K-1$. Also, we know that $\eta(0) + \eta(1) + \cdots + \eta(K-1) = N$. Thus, $\mathbb{E}_{\nu_N}[\eta(k)] = \dfrac{N}{K}$, for all $k = 0, 1, \ldots, K-1$.

□

Let us define the functions f_k and $f_{k,l}$ on $\mathcal{E}_{K,N}$ as $f_k : \eta \mapsto \eta(k)$ and $f_{k,l} : \eta \mapsto \eta(k)\eta(l)$, for all $k, l \in \{0, 1, \ldots, K-1\}$. The following lemma provides explicit expressions for the evaluation of the generator of the N-particle process on these functions.

Lemma 3.3.2 (Dynamics of the N-particle process)**.** *We have that*

$$\begin{aligned}
\mathcal{L}_{K,N} f_k &= f_{k-1} - (1+\theta) f_k + \theta f_{k+1}, &(3.31)\\
\mathcal{L}_{K,N} f_{k,k} &= 2\left[f_{k-1,k} - \left(1+\theta+\frac{p}{N-1}\right) f_{k,k} + \theta f_{k,k+1}\right]\\
&\quad + f_{k-1} + \left(1+\theta+\frac{2pN}{N-1}\right) f_k + \theta f_{k+1}, &(3.32)\\
\mathcal{L}_{K,N} f_{k,k+1} &= -2\left(1+\theta+\frac{p}{N-1}\right) f_{k,k+1} + f_{k-1,k+1} + \theta f_{k+1,k+1} + f_{k,k} + \theta f_{k,k+2}\\
&\quad - f_k - \theta f_{k+1}, &(3.33)\\
\mathcal{L}_{K,N} f_{k,l} &= -2\left(1+\theta+\frac{p}{N-1}\right) f_{k,l} + f_{k-1,l} + \theta f_{k+1,l} + f_{k,l-1} + \theta f_{k,l+1}, &(3.34)
\end{aligned}$$

for all $k, l \in \mathbb{Z}/K\mathbb{Z}$ such that $|k-l| > 2$.

The proof of Lemma 3.3.2 is mostly technical and it is deferred to Section 3.A.

The expression (3.31) given by this lemma is used to study the behavior of the mean of the proportion of particles in each state. Also, (3.32), (3.33) and (3.34) are used to study the covariances of the number of particles when t and N tend to infinity.

Let us denote

$$s_k = \mathbb{E}_{\nu_N}\left[\frac{f_{l,l+k}(\eta)}{N^2}\right] = \mathbb{E}_{\nu_N}\left[\frac{f_{0,k}(\eta)}{N^2}\right] = \mathbb{E}_{\nu_N}\left[\frac{\eta(0)}{N}\frac{\eta(k)}{N}\right], \tag{3.35}$$

for all $k, l \in \mathbb{Z}/K\mathbb{Z}$. Note that the second equality comes from part b) of Theorem 3.1.2. Let us define the constant

$$\gamma_N = -2\left(1 + \frac{Np}{(N-1)(1+\theta)}\right). \tag{3.36}$$

The following two lemmas will be useful for obtaining explicit expressions for the quantities s_k, for $k = 0, 1, \ldots, K-1$.

Lemma 3.3.3. *Then, for $K \geq 3$, the values s_k, for $0 \leq k \leq K-2$, satisfy the following linear system:*

$$\begin{aligned}
-s_{K-1} + \beta_N\, s_0 - s_1 &= -\frac{\gamma_N}{KN}, &(3.37)\\
-s_0 + \beta_N\, s_1 - s_2 &= -\frac{1}{KN}, &(3.38)
\end{aligned}$$

and when $K \geq 4$:

$$-s_{l-1} + \beta_N\, s_l - s_{l+1} = 0, \tag{3.39}$$

for $2 \leq l \leq K-2$, where β_N and γ_N are defined by (3.10) and (3.36), respectively.

Proof of Lemma 3.3.3. Using (3.32) we have

$$\begin{aligned}\mathbb{E}_{\nu_N}[(\mathcal{L}_{K,N}f_{k,k})(\eta)] &= 2\left[\mathbb{E}_{\nu_N}[f_{k-1,k}(\eta)] - \left(1+\theta+\frac{p}{N-1}\right)\mathbb{E}_{\nu_N}[f_{k,k}(\eta)] + \theta\mathbb{E}_{\nu_N}[f_{k,k+1}(\eta)]\right] + \\ &\quad \mathbb{E}_{\nu_N}[f_{k-1}(\eta)] + \left(1+\theta+\frac{2pN}{N-1}\right)\mathbb{E}_{\nu_N}[f_k(\eta)] + \theta\mathbb{E}_{\nu_N}[f_{k+1}(\eta)].\end{aligned}$$

Since ν_N is the stationary distribution, we know that $\mathbb{E}_{\nu_N}[(\mathcal{L}_{K,N}f)(\eta)] = 0$, for all f on $\mathcal{E}_{K,N}$. Thus, using parts *a*) and *b*) of Theorem 3.1.2 and dividing by N^2, we have the equality

$$2(1+\theta)s_1 - 2\left(1+\theta+\frac{p}{N-1}\right)s_0 = -\frac{2}{KN}\left(1+\theta+\frac{pN}{N-1}\right).$$

Dividing by $(1+\theta)$, this last equality is equivalent to

$$\beta_N\, s_0 - 2s_1 = -\frac{\gamma_N}{KN}. \tag{3.40}$$

Note that $s_1 = s_{K-1}$ due to Corollary 3.1.3. Using this fact, we deduce that (3.40) is equivalent to (3.37).

Furthermore, using (3.33) we get

$$\begin{aligned}\mathbb{E}_{\nu_N}[(\mathcal{L}_{K,N}f_{k,k+1})(\eta)] &= -2\left(1+\theta+\frac{p}{N-1}\right)\mathbb{E}_{\nu_N}[f_{k,k+1}(\eta)] + \mathbb{E}_{\nu_N}[f_{k-1,k+1}(\eta)] \\ &\quad +\theta\,\mathbb{E}_{\nu_N}[f_{k+1,k+1}(\eta)] + \mathbb{E}_{\nu_N}[f_{k,k}(\eta)] + \theta\,\mathbb{E}_{\nu_N}[f_{k,k+2}(\eta)] \\ &\quad -\mathbb{E}_{\nu_N}[f_k(\eta)] - \theta\,\mathbb{E}_{\nu_N}[f_{k+1}(\eta)].\end{aligned}$$

In a similar way to the previous case we obtain the equation $-s_0 + \beta_N s_1 - s_2 = -1/KN$, which is equivalent to (3.38).

Similarly, using (3.34), the equality (3.39) is proved for all $2 \le l \le K-2$. □

Note that using Corollary 3.1.3 and formula (3.38) we can obtain the following relation

$$-s_{K-2} + \beta_N s_{K-1} - s_0 = -\frac{1}{KN}. \tag{3.41}$$

Let us define the $K \times K$ circulant matrix A_K and the K-vector $\mathbf{b}_K$ by

$$\begin{aligned}A_K &= \operatorname{circ}(\beta_N, -1, 0, \ldots, \ldots, 0, -1), \\ \mathbf{b}_K &= (\gamma_N, 1, 0, 0, \ldots, 0, 1)^T,\end{aligned}$$

for $K \ge 3$, where β_N and γ_N are defined by (3.10) and (3.36), respectively.

Using Equations (3.37), (3.38), (3.39) and (3.41), the quantities s_k, $0 \le k \le K-1$, defined in (3.35) are proved to verify the linear system of equations

$$A_K\,\mathbf{s}_K = -\frac{1}{KN}\mathbf{b}_K, \tag{3.42}$$

where $\mathbf{s}_K = (s_0, s_1, \ldots, s_{K-1})^T$ and β_N and γ_N are defined by (3.10).

Note that the vector $\mathbf{b}_K$ is *almost symmetric*, in the sense that $b_k = b_{K-k}$, $1 \le k \le K-1$, where $b_k, 0 \le k \le K-1$, are the K components of $\mathbf{b}_K$. Moreover, a vector $\mathbf{b}$ is almost symmetric if and only if the equality $J\mathbf{b} = \mathbf{b}$ holds, where

$$J = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & 0 & ⋰ & 1 & 0 \\ \vdots & \vdots & ⋰ & ⋰ & \vdots & \vdots \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ 0 & 1 & 0 & \ldots & 0 & 0 \end{pmatrix}.$$

In addition, any symmetric circulant matrix of size n can be expressed as follows

$$A = a_0 I + a_1 \Pi + a_2 \Pi^2 + \cdots + a_{n-1}\Pi^{n-1},$$

where $(a_0, a_1, \ldots, a_{n-1})$ is an almost symmetric vector and $\Pi = \mathrm{circ}(0, 1, 0, \ldots, 0)$.

The following result gives us information about the solution of a symmetric circulant system when the vector of constant terms is almost symmetric.

Proposition 3.3.4 (Circulant matrices)**.** *Let A be a n-dimensional invertible circulant symmetric matrix and let $\mathbf{b}$ be an almost symmetric vector of dimension n, then $\mathbf{x} = A^{-1}\mathbf{b}$, the solution of the linear system $A\mathbf{x} = \mathbf{b}$, is an almost symmetric vector.*

Proof. Since A is a invertible matrix, we know that $\mathbf{x}$ is the unique vector of dimension n satisfying $A\mathbf{x} = \mathbf{b}$ and this vector $\mathbf{x}$ is almost symmetric if and only if $\mathbf{x} = J\mathbf{x}$. So, it is sufficient to prove that $J\mathbf{x}$ is also a solution of the linear system, i.e. $A(J\mathbf{x}) = \mathbf{b}$. Since $\mathbf{b}$ is almost symmetric, the equation $A(J\mathbf{x}) = \mathbf{b}$ becomes equivalent to

$$JA(J\mathbf{x}) = \mathbf{b}. \tag{3.43}$$

It is sufficient to prove that $JAJ = A$. Note that the matrix J is an involutory matrix, i.e. $J^{-1} = J$, and

$$\begin{aligned} JAJ &= J\big(a_0 I + a_1\Pi + a_2\Pi^2 + \cdots + a_{n-1}\Pi^{n-1}\big)J \\ &= a_0 I + a_1 J\Pi J + a_2 J\Pi^2 J + \cdots + a_{n-1} J\Pi^{n-1} J. \end{aligned}$$

The matrix Π is orthogonal, satisfying $\Pi^{-1} = \Pi^T$. Moreover,

$$J\Pi J = J(\Pi J) = J \begin{pmatrix} 0 & 0 & \ldots & 0 & 0 & 1 \\ 0 & 0 & \ldots & 0 & 1 & 0 \\ 0 & 0 & \ldots & 1 & 0 & 0 \\ \vdots & \vdots & \iddots & \iddots & \vdots & \vdots \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} = \Pi^T,$$

which implies $J\Pi^n J = J\Pi J^2 \Pi^{n-1} J = \Pi^T J\Pi^{n-1} J = \cdots = \big(\Pi^T\big)^n$. Thus, we get

$$\begin{aligned} JAJ &= a_0 I + a_1 \Pi^T + a_2\big(\Pi^T\big)^2 + \cdots + a_{n-1}\big(\Pi^T\big)^{n-1} \\ &= \big(a_0 I + a_1\Pi + a_2(\Pi)^2 + \cdots + a_{n-1}(\Pi)^{n-1}\big)^T \\ &= A^T = A. \end{aligned}$$

Thus, (3.43) holds and hence $J\mathbf{x}$ is solution of the equation $A\mathbf{x} = \mathbf{b}$. By uniqueness of the solution we get $\mathbf{x} = J\mathbf{x}$, proving that $\mathbf{x}$ is almost symmetric. □

Because the $K \times K$ matrix A_K in (3.42) is a symmetric circulant matrix, it is possible to obtain explicit formulas for all its eigenvalues and eigenvectors using [Dav79, Thm. 3.2.2]. Since all its eigenvalues are non-null, we conclude that the matrix A_K is invertible. Thus, using Proposition 3.3.4, the linear system (3.42) has as its unique solution the vector $\mathbf{s}_K$, which is almost symmetric. In addition to its almost symmetry, the vector $\mathbf{b}_K$ satisfies $b_1 = b_{K-1}$, $b_k = 0$, $2 \le k \le K-2$. This simple structure of $\mathbf{b}_K$ allows us to deduce explicit expressions for s_k, $0 \le k \le K-1$, given in Theorem 3.1.4, which is proved below.

3.3.2 Proof of Theorem 3.1.4

Consider the four families of orthogonal polynomials $N_{\text{even},n}(x)$, $D_{\text{even},n}(x)$, $N_{\text{odd},n}(x)$, $D_{\text{odd},n}(x)$, $n \geq 0$, defined by

$$\begin{array}{llll} N_{\text{even},0}(x) = 2, & D_{\text{even},0}(x) = 0, & N_{\text{odd},0}(x) = 1, & D_{\text{odd},0}(x) = 1 \\ N_{\text{even},1}(x) = x, & D_{\text{even},1}(x) = x+2, & N_{\text{odd},1}(x) = x-1, & D_{\text{odd},1}(x) = x+1, \end{array}$$

satisfying all of them the recurrence relation

$$p_{n+1}(x) = x\,p_n(x) - p_{n-1}(x), \tag{3.44}$$

for all $n \geq 1$.

The next proposition will prove useful in the sequel.

Lemma 3.3.5. *The following relations hold, for all $n \geq 0$:*

$$\begin{aligned} 2N_{\text{even},\,n}(x) - xN_{\text{even},\,n+1}(x) + (x-2)D_{\text{even},\,n+1}(x) &= 0, & (3.45)\\ 2N_{\text{odd},\,n}(x) - xN_{\text{odd},\,n+1}(x) + (x-2)D_{\text{odd},\,n+1}(x) &= 0. & (3.46)\end{aligned}$$

Furthermore, we have the following identities involving the Chebyshev polynomials of the first and the second kind, for all $n \geq 0$:

$$\begin{aligned} N_{\text{even},\,n}(x) &= 2\,T_n(x/2), & (3.47)\\ D_{\text{even},\,n}(x) &= (x+2)\,U_{n-1}(x/2), & (3.48)\\ N_{\text{odd},\,n}(x) &= U_n(x/2) - U_{n-1}(x/2), & (3.49)\\ D_{\text{odd},\,n}(x) &= U_n(x/2) + U_{n-1}(x/2). & (3.50)\end{aligned}$$

Proof. Setting $P_n(x) = 2N_{\text{even},n}(x) - xN_{\text{even},n+1}(x) + (x-2)D_{\text{even},n+1}(x)$, for all $n \geq 0$, it follows from the definitions of $N_{\text{even},n}(x)$ and $D_{\text{even},n}(x)$ that $P_0(x) = 0, P_1(x) = 0$ and $P_n(x)$ satisfies the recurrence relation (3.44). Therefore $P_n(x) = 0$ for every $n \geq 0$ and (3.45) is proved. The proof of (3.46) is similar.

Now, note that the sequence of polynomials $(2\,T_n(x/2))_{n\geq 0}$ satisfy the recurrence relation (3.9). Furthermore, $2\,T_0(x/2) = 2 = N_{\text{even},\,0}(x)$ and $2\,T_1(x/2) = x = N_{\text{even},\,1}(x)$. Consequently, identity (3.47) is proved. Analogously, identities (3.48), (3.49) and (3.50) are proved. □

We now prove Lemma 3.3.6, which provides explicit expressions for s_k, $k \in \{0, 1, \ldots, K-1\}$, in terms of the polynomials $N_{\text{even},n}(x)$, $D_{\text{even},n}(x)$, $N_{\text{odd},n}(x)$ and $D_{\text{odd},n}(x)$.

Lemma 3.3.6 (Explicit formulas for s_k)**.** *The values of s_k, $0 \leq k \leq K-1$, are given by*

a) If $K = 2K_2$, $K_2 \geq 2$,

$$s_0 = \frac{N-1}{KN}\,\frac{N_{\text{even},K_2}(\beta_N)}{D_{\text{even},K_2}(\beta_N)} + \frac{1}{KN}, \tag{3.51}$$

$$s_k = \frac{N-1}{KN}\,\frac{N_{\text{even},K_2-k}(\beta_N)}{D_{\text{even},K_2}(\beta_N)},\ 1 \leq k \leq K_2, \tag{3.52}$$

$$s_{K-k} = s_k,\ 1 \leq k \leq K_2 - 1,$$

b) If $K = 2K_2 + 1$, $K_2 \geq 1$,

$$s_0 = \frac{N-1}{KN}\,\frac{N_{\text{odd},K_2}(\beta_N)}{D_{\text{odd},K_2}(\beta_N)} + \frac{1}{KN},$$

$$s_k = \frac{N-1}{KN}\,\frac{N_{\text{odd},K_2-k}(\beta_N)}{D_{\text{odd},K_2}(\beta_N)},\ 1 \leq k \leq K_2,$$

$$s_{K-k} = s_k,\ 1 \leq k \leq K_2,$$

where β_N is defined by (3.10).

Proof. We separate the proof into two cases: when K is even and when K is odd.

(a) When K is even, say $K = 2K_2$, Equation (3.42) is equivalent to the following linear system for s_k, $0 \le k \le K_2$,

$$\beta_N s_0 - 2s_1 = -\frac{1}{KN}\gamma_N, \tag{3.53}$$

$$-s_0 + \beta_N s_1 - s_2 = -\frac{1}{KN}, \tag{3.54}$$

$$-s_{k-1} + \beta_N s_k - s_{k+1} = 0, \tag{3.55}$$

for $2 \le k \le K_2 - 1$ and

$$\beta_N s_{K_2} - 2s_{K_2-1} = 0. \tag{3.56}$$

Note that (3.56) follows from the equality $s_{K_2-1} = s_{K_2+1}$.

Consider $A \in \mathbb{R}$ such that $s_{K_2} = 2A = N_{\text{even},0}(\beta_N)A$. Equation (3.56) implies

$$s_{K_2-1} = A\beta_N = AN_{\text{even},1}(\beta_N).$$

Equation (3.55) may be written as

$$s_{k-1} = \beta_N s_k - s_{k+1},$$

for $2 \le k \le K_2 - 1$. This proves that s_k, for k decreasing from K_2 to 1, may be written

$$s_k = AN_{\text{even},K_2-k}(\beta_N).$$

From Equation (3.54), we get

$$\begin{aligned} s_0 &= \beta_N s_1 - s_2 + \frac{1}{KN} \\ &= A\left[\beta_N N_{\text{even},K_2-1}(\beta_N) - N_{\text{even},K_2-2}(\beta_N)\right] + \frac{1}{KN} \\ &= AN_{\text{even},K_2}(\beta_N) + \frac{1}{KN}. \end{aligned} \tag{3.57}$$

Plugging (3.57) into Equation (3.53), we get

$$\begin{aligned} A\left[\beta_N N_{\text{even},K_2}(\beta_N) - 2N_{\text{even},K_2-1}(\beta_N)\right] &= -\frac{1}{KN}(\beta_N + \gamma_N) \\ &= \frac{1}{KN}\frac{2p}{1+\theta}. \end{aligned} \tag{3.58}$$

Using Equation (3.45) we get

$$\begin{aligned} A[\beta_N N_{\text{even},K_2}(\beta_N) - 2N_{\text{even},K_2-1}(\beta_N)] &= A(\beta_N - 2)D_{\text{even},K_2}(\beta_N) \\ &= A\frac{2p}{(N-1)(1+\theta)}D_{\text{even},K_2}(\beta_N). \end{aligned} \tag{3.59}$$

Thus, using (3.58) and (3.59), we obtain $A = \frac{1}{KN}\frac{N-1}{D_{\text{even},K_2}(\beta_N)}$, that achieves the proof of (3.52) for an even value of K.

(b) The proof when K is odd is similar. Indeed, for $K = 2K_2 + 1$, the linear system for s_k, with $0 \leq k \leq K_2$, is

$$\begin{aligned} \beta_N s_0 - 2s_1 &= -\frac{1}{KN}\gamma_N, && (3.60)\\ -s_0 + \beta_N s_1 - s_2 &= -\frac{1}{KN}, && (3.61)\\ -s_{k-1} + \beta_N s_k - s_{k+1} &= 0, && (3.62) \end{aligned}$$

for $2 \leq k \leq K_2 - 1$ and

$$-s_{K_2} + \beta_N s_{K_2} - s_{K_2-1} = 0. \tag{3.63}$$

Equation (3.63) may be written as

$$(\beta_N - 1)s_{K_2} = s_{K_2-1},$$

and so

$$s_{K_2} = B = BN_{\text{odd},0}(\beta_N),\ s_{K_2-1} = B(\beta_N - 1) = BN_{\text{odd},1}(\beta_N).$$

From Equations (3.46) and (3.62), it follows that

$$s_k = BN_{\text{odd},K_2-k}(\beta_N),\ 1 \leq k \leq K_2.$$

Then, from Equation (3.61), we get

$$\begin{aligned} s_0 &= \beta_N s_1 - s_2 + \frac{1}{KN}\\ &= B\left[\beta_N N_{\text{odd},K_2-1}(\beta_N) - N_{\text{odd},K_2-2}(\beta_N)\right] + \frac{1}{NK}\\ &= BN_{\text{odd},K_2}(\beta_N) + \frac{1}{KN}. \end{aligned}$$

From (3.60), it follows, using (3.46), that $B = \frac{N-1}{KND_{\text{odd},K_2}(\beta_N)}$. The proof of Lemma 3.3.6 is therefore complete. $\square$

We are now able to prove Theorem 3.1.4, which provides explicit expressions for the covariances of the proportions of particles in two states under the stationary distribution, in terms of the orthogonal Chebyshev polynomials of the first and the second kind.

Proof of Theorem 3.1.4. Using expressions (3.47), (3.48), (3.49) and (3.50), and Lemma 3.3.6 we obtain explicit expressions for s_k in terms of the Chebyshev polynomials of the first and the second kind, for $0 \leq k \leq K-1$. Since $\text{Cov}_{\nu_N}[\eta(0)/N, \eta(k)/N] = s_k - 1/K^2$, for all $0 \leq k \leq K-1$, we deduce that (3.11), (3.12), (3.13) and (3.14) hold. $\square$

Now, using Theorem 3.1.4 we are able to study the monotony of the covariance of the proportions of particles in two sites as a function of the graph distances between these two sites.

Proof of Corollary 3.1.5. Note that $\text{Cov}_{\nu_N}\left[\frac{\eta(0)}{N}, \frac{\eta(k)}{N}\right] \geq \text{Cov}_{\nu_N}\left[\frac{\eta(0)}{N}, \frac{\eta(k+1)}{N}\right]$ holds if and only if $s_k \geq s_{k+1}$, for all $k = 0, 1, \ldots, \lfloor \frac{K}{2} \rfloor$. So, for K even, using (3.11) and (3.12), it is sufficient to prove that $T_{k+1}(\beta_N/2) \geq T_k(\beta_N/2)$. Let us prove it by induction. We know that $T_1(\beta_N/2) = \beta_N/2 \geq 1 = T_0(\beta_N/2)$. Assume that $T_k(\beta_N/2) \geq T_{k-1}(\beta_N/2)$. Since $(T_n(x))_{n\geq 0}$ satisfies the recurrence relation (3.9) we have

$$T_{k+1}(\beta_N/2) - T_k(\beta_N/2) = (\beta_N - 1)T_k(\beta_N/2) - T_{k-1}(\beta_N/2) \geq T_k(\beta_N/2) - T_{k-1}(\beta_N/2) \geq 0,$$

where the first inequality is due to the inequality $\beta_N \geq 2$ and the second one because, by assumption, $T_k(\beta_N/2) \geq T_{k-1}(\beta_N/2)$. Then, $T_{k+1}(\beta_N/2) \geq T_k(\beta_N/2)$, for all $k \geq 0$.

Analogously, for K odd the inequality $\mathrm{Cov}_{\nu_N}\left[\frac{\eta(0)}{N}, \frac{\eta(k)}{N}\right] \geq \mathrm{Cov}_{\nu_N}\left[\frac{\eta(0)}{N}, \frac{\eta(k+1)}{N}\right]$ holds for all $k = 0, 1, \dots, \lfloor \frac{K}{2} \rfloor$ if

$$U_{k+1}(\beta_N/2) - U_k(\beta_N/2) \geq U_k(\beta_N/2) - U_{k-1}(\beta_N/2), \tag{3.64}$$

for all $k \geq 1$. For $k = 1$ we have that (3.64) is equivalent to $\beta_N^2 - 2\beta_N \geq 0$, which is trivially true since $\beta_N \geq 2$. Assume that (3.64) holds and let us prove the inequality for $k+1$. Indeed, using that $(U_n)_{n\geq 0}$ satisfies the recurrence relation (3.9), we have

$$U_{k+2}(\beta_N/2) - U_{k+1}(\beta_N/2) = (\beta_N - 1)U_{k+1}(\beta_N/2) - U_k(\beta_N/2) \geq U_{k+1}(\beta_N/2) - U_k(\beta_N/2).$$

Thus, (3.64) holds for all $k = 0, 1, \dots, K_2$. □

3.3.3 Proof of Theorem 3.1.6

Theorem 3.1.4 allows us to get a Taylor series expansion for s_k, $0 \leq k \leq K-1$ as a function of $\frac{1}{N}$, as soon as we are able to obtain such a series expansion for β_N, as a function of $1/N$, as well as for the polynomials $N_{\mathrm{odd},n}(x)$, $N_{\mathrm{even},n}(x)$, $D_{\mathrm{odd},n}(x)$, $D_{\mathrm{even},n}(x)$, $n \geq 0$ around $x = 2$, using their definitions by induction given in (3.44).

Lemma 3.3.7. *The polynomials $N_{\mathrm{odd},n}(x)$, $N_{\mathrm{even},n}(x)$, $D_{\mathrm{odd},n}(x)$, $D_{\mathrm{even},n}(x)$, for $n \geq 0$, satisfy the following Taylor series expansion of order 2 around $x = 2$:*

$$\begin{aligned}
N_{\mathrm{even},n}(x) &= 2 + n^2(x-2) + \frac{n^4 - n^2}{12}(x-2)^2 + o(x-2)^2, && (3.65)\\
D_{\mathrm{even},n}(x) &= 4n + \frac{2n^3 + n}{3}(x-2) + \frac{n^5 - n}{30}(x-2)^2 + o(x-2)^2, && (3.66)\\
N_{\mathrm{odd},n}(x) &= 1 + \frac{n^2 + n}{2}(x-2) + \frac{n^4 + 2n^3 - n^2 - 2n}{24}(x-2)^2 + o(x-2)^2, && (3.67)\\
D_{\mathrm{odd},n}(x) &= 2n + 1 + \frac{2n^3 + 3n^2 + n}{6}(x-2) + \frac{2n^5 + 5n^4 - 5n^2 - 2n}{120}(x-2)^2 + \\
&\quad + o(x-2)^2. && (3.68)
\end{aligned}$$

Proof. Assume $N_{\mathrm{even},n}(x) = a_0^{(n)} + a_1^{(n)}(x-2) + a_2^{(n)}(x-2)^2 + o(x-2)^2$, for all $n \geq 0$. Note that the polynomials $N_{\mathrm{even},n}(x)$ can also be defined as

$$\begin{aligned}
N_{\mathrm{even},0}(x) &= 2,\\
N_{\mathrm{even},1}(x) &= (x-2) + 2,\\
N_{\mathrm{even},n}(x) &= (x-2)N_{\mathrm{even},n-1}(x) + 2N_{\mathrm{even},n-1}(x) - N_{\mathrm{even},n-2}(x), \ n \geq 2. && (3.69)
\end{aligned}$$

Thus, the coefficients $(a_0^{(n)})_{n\geq 0}$ satisfy the recurrence relation $a_0^{(0)} = a_0^{(1)} = 2$ and $a_0^{(n)} = 2a_0^{(n-1)} - a_0^{(n-2)}$, for every $n \geq 2$, which yields $a_0^{(n)} = 2$, for all $n \geq 0$.

Also, using (3.69), the coefficients $(a_1^{(n)})_{n\geq 0}$ satisfy $a_1^{(0)} = 0$, $a_1^{(1)} = 1$ and

$$a_1^{(n)} = 2a_1^{(n-1)} - a_1^{(n-2)} + a_0^{(n-1)} = 2a_1^{(n-1)} - a_1^{(n-2)} + 2,$$

for all $n \geq 0$. Solving this recurrence gives $a_1^{(n)} = n^2$, for all $n \geq 2$.

Similarly, the coefficients $(a_2^{(n)})_{n\geq 0}$ satisfy $a_2^{(0)} = a_2^{(1)} = 0$ and

$$a_2^{(n)} = 2a_2^{(n-1)} - a_2^{(n-2)} + a_1^{(n-1)} = 2a_2^{(n-1)} - a_2^{(n-2)} + (n-1)^2,$$

for all $n \geq 0$. which yields $a_2^{(n)} = \dfrac{n^4 - n^2}{12}$, for all $n \geq 2$, proving (3.65).

The proofs of (3.66), (3.67) and (3.68) are similar. $\square$

We now prove Theorem 3.1.6, which provides a second order Taylor series expansion of the variance of the proportion of particles in each state, as a function of $1/N$, when N tends to infinity.

Proof of Theorem 3.1.6. Suppose K is even, say $K = 2K_2$. Using Lemma 3.3.6, we have

$$s_k = \frac{1}{K}\left(1 - \frac{1}{N}\right)\frac{N_{\text{even},K_2-k}(\beta_N)}{D_{\text{even},K_2}(\beta_N)},$$

for all $k = 1, 2, \ldots, K_2$. Note that β_N, defined by (3.10), tends to 2 when N tends to infinity, specifically

$$\beta_N - 2 = \frac{2p}{(N-1)(1+\theta)} = \frac{2p}{1+\theta}\left(\frac{1}{N} + \frac{1}{N^2}\right) + o\left(\frac{1}{N^2}\right).$$

Using (3.65) and (3.66), we have

$$\begin{aligned}
\frac{N_{\text{even},K_2-k}(\beta_N)}{D_{\text{even},K_2}(\beta_N)} &= \frac{2 + (K_2-k)^2(\beta_N-2) + \frac{(K_2-k)^4-(K_2-k)^2}{12}(\beta_N-2)^2 + o\left((\beta_N-2)^2\right)}{4K_2 + \frac{2K_2^3+K_2}{3}(\beta_N-2) + \frac{K_2^5-K_2}{30}(\beta_N-2)^2 + o\left((\beta_N-2)^2\right)} \\
&= \frac{1}{K} + \frac{(6k(k-K)+K^2-1)}{12K}(\beta_N-2) \\
&\quad + \frac{30k(K-k)[k(K-k)+2] - (K^2-1)(K^2+11)}{720K}(\beta_N-2)^2 \\
&\quad + o\left((\beta_N-2)^2\right),
\end{aligned} \tag{3.70}$$

where $K = 2K_2$.

Finally,

$$\begin{aligned}
s_k &= \frac{1}{K^2} + \left(-1 + \frac{(6k(k-K)+K^2-1)}{6}\frac{p}{1+\theta}\right)\frac{1}{K^2N} \\
&\quad + \frac{30k(K-k)[k(K-k)+2] - (K^2-1)(K^2+11)}{180}\left(\frac{p}{1+\theta}\right)^2\frac{1}{K^2N^2} + o\left(\frac{1}{N^2}\right).
\end{aligned}$$

Using (3.51), we get the following expression for s_0,

$$s_0 = \frac{1}{K^2} + \left(K - 1 + \frac{K^2-1}{6}\frac{p}{1+\theta}\right)\frac{1}{K^2N} + \frac{(K^2-1)(K^2+11)}{180}\left(\frac{p}{1+\theta}\right)^2\frac{1}{K^2N^2} + o\left(\frac{1}{N}\right).$$

Now, the expression (3.15) for $\mathrm{Cov}_{\nu_N}[\eta(0)/N, \eta(k)/N]$ with K even follows by noting that $\mathbb{E}_{\nu_N}\left[\frac{\eta(k)}{N}\right] = \frac{1}{K}$, for all $k = 0, 1, 2, \ldots, K-1$.

Considering K odd, specifically $K = 2K_2 + 1$, and using (3.67) and (3.68), we have

$$\begin{aligned}
\frac{N_{\text{odd},K_2-k}(\beta_N)}{D_{\text{odd},K_2}(\beta_N)} &= \frac{1}{K} + \frac{(6k(k-K)+K^2-1)}{12K}(\beta_N-2) \\
&\quad + \frac{30k(K-k)[k(K-k)+2] - (K^2-1)(K^2+11)}{720K}(\beta_N-2)^2 \\
&\quad + o\left((\beta_N-2)^2\right),
\end{aligned}$$

which is the same expression we get for $\dfrac{N_{\mathrm{even},K_2-k}(\beta_N)}{D_{\mathrm{even},K_2}(\beta_N)}$ in (3.70). So, the general result is proved. $\square$

Proof of Corollary 3.1.7. Using Jensen's inequality, we have

$$\begin{aligned}
\mathbb{E}_{\nu_N}\left[\|m(\eta)-\nu_{\mathrm{qs}}\|_2\right] &\leq \left(\mathbb{E}_{\nu_N}\|m(\eta)-\nu_{\mathrm{qs}}\|_2^2\right)^{1/2} \\
&= \left(\sum_{k=0}^{K-1}\mathrm{Var}_{\nu_N}\left[\frac{\eta(k)}{N}\right]\right)^{1/2} \\
&= \sqrt{K}\left(\mathrm{Var}_{\nu_N}\left[\frac{\eta(0)}{N}\right]\right)^{1/2}. \qquad (3.71)
\end{aligned}$$

Finally, (3.16) is proved using (3.71) and Theorem 3.1.6. $\square$

3.4 Covariances of the proportions of particles at a given time

3.4.1 Proof of Theorem 3.1.8

Proof of Theorem 3.1.8. Consider $\eta \in \mathcal{E}_{K,N}$ and the function $f_k : \eta \mapsto \eta(k)$, for $k \in \{0,1,\dots,K-1\}$. Using the expression of $\mathcal{L}_{K,N} f_k$, for $k=0,1,\dots,K-1$, given by (3.31), and the Kolmogorov equation, we get

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\,t}\mathbb{E}_\eta\left[\frac{f_k(\eta_t^{(N)})}{N}\right] &= \mathbb{E}_\eta\left[\frac{\mathcal{L}_{K,N}f_k(\eta_t^{(N)})}{N}\right] \\
&= \mathbb{E}_\eta\left[\frac{f_{k-1}(\eta_t^{(N)})}{N}\right] - (1+\theta)\mathbb{E}_\eta\left[\frac{f_k(\eta_t^{(N)})}{N}\right] \\
&\quad +\theta\,\mathbb{E}_\eta\left[\frac{f_{k+1}(\eta_t^{(N)})}{N}\right], \qquad (3.72)
\end{aligned}$$

for $k=0,1,\dots,K-1$.

Let us define $s_t(k) = \mathbb{E}_\eta\left[f_k(\eta_t^{(N)})/N\right] = \mathbb{E}_\eta\left[\eta_t^{(N)}(k)/N\right] = \overline{m}(\eta_t^{(N)})(k)$, for $k=0,1,\dots,K-1$, and the vector $\mathbf{s}_t = (s_t(0), s_t(1), \dots, s_t(K-1))^T$. Using (3.72), we get that $\mathbf{s}_t$ satisfies the differential equation

$$\frac{\mathrm{d}\,\mathbf{s}_t}{\mathrm{d}\,t} = \mathbf{s}_t Q,$$

where Q is the circulant infinitesimal rate matrix defined in (3.24), with initial condition $\mathbf{s}_0 = \eta/N$. Note that the solution of this differential equation is given by

$$\mathbf{s}_t = \frac{\eta}{N}\mathrm{e}^{tQ}.$$

Thus, $\overline{m}(\eta_t^{(N)})$ is actually equal to the distribution of the asymmetric random walk on the cycle graph $\mathbb{Z}/K\mathbb{Z}$ with infinitesimal generator matrix Q and initial distribution $m(\eta)$ at time $t=0$, which is $\mathcal{L}_{m(\eta)}(Z_t \mid t < \tau_p)$. So, the proof of formula (3.17) follows from (3.3) in Theorem 3.1.1. $\square$

3.4.2 Proof of Theorem 3.1.9

In order to study the convergence of the empirical distribution $m(\eta_t^{(N)})$ induced by the N-particle system, we will analyze the behavior of the covariance functions in time. Let $\eta \in \mathcal{E}_{K,N}$ be fixed and let us define the functions $s_t^{(2)}(k,r)$ as $s_t^{(2)}(k,r) = \mathbb{E}_\eta\left[f(k,r)/N^2\right] = \mathbb{E}_\eta\left[\eta(k)\eta(r)/N^2\right]$, for all $k, r \in \mathbb{Z}/K\mathbb{Z}$. Using (3.32), (3.33) and (3.34), we have

$$\begin{aligned}
\frac{\mathrm{d}\, s_t^{(2)}(k,k)}{\mathrm{d}\, t} &= 2\left[s_t^{(2)}(k,k-1) - \left(1+\theta+\frac{p}{N-1}\right)s_t^{(2)}(k,k) + \theta s_t^{(2)}(k,k+1)\right] \\
&\quad + \frac{1}{N}\left[s_t(k-1) + \left(1+\theta+2\frac{p}{N-1}\right)s_t(k) + \theta s_t(k+1)\right], \\
\frac{\mathrm{d}\, s_t^{(2)}(k,k+1)}{\mathrm{d}\, t} &= -2\left(1+\theta+\frac{p}{N-1}\right)s_t^{(2)}(k,k+1) + s_t^{(2)}(k-1,k+1) + \theta s_t^{(2)}(k+1,k+1) \\
&\quad + s_t^{(2)}(k,k) + \theta s_t^{(2)}(k,k+2) - \frac{1}{N}\left[s_t(k) + \theta s_t(k+1)\right] \\
\frac{\mathrm{d}\, s_t^{(2)}(k,k+l)}{\mathrm{d}\, t} &= -2\left(1+\theta+\frac{p}{N-1}\right)s_t^{(2)}(k,k+l) + s_t^{(2)}(k-1,k+l) + \theta s_t^{(2)}(k+1,k+l) \\
&\quad + s_t^{(2)}(k,k+l-1) + \theta s_t^{(2)}(k,k+l+1).
\end{aligned}$$

Consider the functions $g_t(k,r)$ defined as

$$g_t(k,r) = \mathrm{Cov}_\eta\left[\frac{\eta_t(k)}{N}, \frac{\eta_t(r)}{N}\right] = s_t^{(2)}(k,r) - s_t(k)s_t(r),$$

for all $k, r \in \mathbb{Z}/K\mathbb{Z}$.

Then, we obtain the following system of differential equations

$$\begin{aligned}
\frac{\mathrm{d}\, g_t(k,k)}{\mathrm{d}\, t} &= 2\left[g_t(k,k-1) - \left(1+\theta+\frac{p}{N-1}\right)g_t(k,k) + \theta g_t(k,k+1)\right] \\
&\quad + \frac{1}{N}\left[s_t(k-1) + \left(1+\theta+2\frac{p}{N-1}\right)s_t(k) + \theta s_t(k+1)\right] - \frac{2p}{N-1}s_t(k)^2, \\
\frac{\mathrm{d}\, g_t(k,k+1)}{\mathrm{d}\, t} &= -2\left(1+\theta+\frac{p}{N-1}\right)g_t(k,k+1) + g_t(k-1,k+1) + \theta g_t(k+1,k+1) \\
&\quad + g_t(k,k) + \theta g_t(k,k+2) - \frac{1}{N}\left[s_t(k) + \theta s_t(k+1)\right] - \frac{2p}{N-1}s_t(k)s_t(k+1), \\
\frac{\mathrm{d}\, g_t(k,l)}{\mathrm{d}\, t} &= -2\left(1+\theta+\frac{p}{N-1}\right)g_t(k,l) + g_t(k-1,l) + \theta g_t(k+1,l) \\
&\quad + g_t(k,l-1) + \theta g_t(k,l+1) - \frac{2p}{N-1}s_t(k)s_t(l).
\end{aligned}$$

Then, the K^2-dimensional vector $\mathbf{g}_t = \left(g_t(k,r)\right)_{k,r}$ satisfies the differential equation

$$\frac{\mathrm{d}\, \mathbf{g}_t}{\mathrm{d}\, t} = \mathbf{g}_t Q_p^{(2)} + \mathbf{w}_t, \tag{3.73}$$

where $Q_p^{(2)} = Q^{(2)} - 2\frac{p}{N-1}I$, I is the K^2-dimensional identity matrix, the matrix $Q^{(2)} \in M_{\mathbb{R}}(K^2)$ is defined as

$$Q^{(2)}_{(u,v),(k,r)} = \begin{cases} 1 & \text{if} \quad (k = u+1 \wedge r = v) \vee (k = u \wedge r = v+1), \\ \theta & \text{if} \quad (k = u-1 \wedge r = v) \vee (k = u \wedge r = v-1), \\ -2(1+\theta) & \text{if} \quad (k = u) \wedge (r = v). \end{cases} \tag{3.74}$$

and $\mathbf{w}_t = (w_t(k,r))_{k,r}$ is the K^2-vector defined by

$$w_t(k,r) = \begin{cases} \frac{1}{N}\left[s_t(k-1) + \left(1+\theta+2\frac{p}{N-1}\right)s_t(k) + \theta s_t(k+1)\right] - \frac{2p}{N-1}s_t(k)^2 & \text{if} \quad r=k \\ -\frac{1}{N}\left[s_t(k\wedge r) + \theta s_t(k\vee r)\right] - \frac{2p}{N-1}s_t(k)s_t(r) & \text{if} \quad |k-r|=1 \\ -\frac{2p}{N-1}s_t(k)s_t(r) & \text{if} \quad |k-r|>1, \end{cases}$$

for all $k, r \in \mathbb{Z}/K\mathbb{Z}$.

Note also that

$$\begin{aligned} g_0(k,r) &= 0, \\ g_\infty(k,r) &= \lim_{t\to\infty} g_t(k,r) = \mathrm{Cov}_{\nu_N}\left[\frac{\eta(k)}{N}, \frac{\eta(r)}{N}\right], \end{aligned}$$

and

$$w_\infty(k,r) = \lim_{t\to\infty} w_t(k,r) = \begin{cases} \frac{2}{KN}\left(1+\theta+\frac{p}{N-1}\right) - \frac{2p}{K^2(N-1)} & \text{if} \quad k=r, \\ -\frac{1}{KN}(1+\theta) - \frac{2p}{K^2(N-1)} & \text{if} \quad |k-r|=1, \\ -\frac{2p}{(N-1)}\frac{1}{K^2} & \text{if} \quad |k-r|>1, \end{cases}$$

for all $k, r \in \mathbb{Z}/K\mathbb{Z}$.

Let $A = (a_{r,c})$ and $B = (b_{r,c})$ be two matrices of dimensions $m \times n$ and $w \times q$, respectively. Recall that the Kronecker product of A and B, denoted by $A \otimes B$, is the $mw \times nq$ matrix defined as

$$A \otimes B = \begin{pmatrix} a_{0,0}B & a_{0,1}B & \dots & a_{0,n-1}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m-1,0}B & a_{m-1,1}B & \dots & a_{m-1,n-1}B \end{pmatrix}.$$

It is convenient to index the elements of $A \otimes B$ with two 2-dimensional index in the following way

$$(A\otimes B)_{(r_1,r_2),(c_1,c_2)} = (A\otimes B)_{r_1 m + r_2, c_1 n + c_2} = a_{r_1,c_1}\, b_{r_2,c_2},$$

for all $0 \le r_1 \le m-1, 0 \le r_2 \le w-1, 0 \le c_1 \le n-1, 0 \le c_2 \le q-1$. Now, consider that $m = n$ and $w = q$, i.e. A and B are square matrices of dimension n and q, respectively. The Kronecker sum of A and B, denoted by $A \oplus B$, is defined as $A \oplus B = A \otimes I_q + I_n \otimes B$, where I_q and I_n are the identity matrices of dimension q and n, respectively. It is well known that the exponential of matrices transforms Kronecker sums in Kronecker products as follows

$$\mathrm{e}^{A\oplus B} = \mathrm{e}^A \otimes \mathrm{e}^B. \tag{3.75}$$

See e.g. Chapter XIV of [Pea65] and [Dav79] for the proofs of these results and more details about the Kronecker product and sum of matrices.

Lemma 3.4.1. *The following properties hold:*

1. $Q^{(2)} = Q \oplus Q$,

2. $\mathrm{e}^{tQ^{(2)}} = \mathrm{e}^{tQ} \otimes \mathrm{e}^{tQ}$.

Consequently, the matrix $Q^{(2)}$ is the infinitesimal rate matrix of the independent coupling of two processes driven by the infinitesimal generator matrix Q.

Proof of Lemma 3.4.1. Note that using (3.74) for all $r_1, r_2, c_1, c_2 \in \{0, 1, \dots, K-1\}$, we have

$$Q^{(2)}_{(r_1,r_2),(c_1,c_2)} = Q_{r_1,c_1} I_{r_2,c_2} + I_{r_1,c_1} Q_{r_2,c_2} = (Q \oplus Q)_{(r_1,r_2),(c_1,c_2)},$$

where I is the K-dimensional identity matrix. Then, property 1 holds. Also, using (3.75) we can easily prove the property 2.

All the non-diagonal entries of matrix $Q^{(2)}$ are positive and the sum of each row is null, thus $Q^{(2)}$ is an infinitesimal matrix. Furthermore,

$$\mathrm{e}^{tQ^{(2)}}_{(r_1,r_2),(c_1,c_2)} = \mathrm{e}^{tQ}_{r_1,c_1} \mathrm{e}^{tQ}_{r_2,c_2},$$

which means that $Q^{(2)}$ is the infinitesimal rate matrix of the independent coupling of two processes driven by Q. $\square$

Note also that, when t goes to infinity in (3.73), we get $\mathbf{g}_\infty Q^{(2)}_p + \mathbf{w}_\infty = 0$. Since $Q^{(2)}$ is the infinitesimal matrix generator of a Markov process and $Q^{(2)}_p = Q^{(2)} - p_N I$, where $p_N = \frac{2p}{N-1}$, all the eigenvalues of $Q^{(2)}_p$ are strictly negative and thus, $Q^{(2)}_p$ is invertible. Then,

$$\mathbf{g}_\infty = -\mathbf{w}_\infty \left(Q^{(2)}_p\right)^{-1}. \tag{3.76}$$

We will now prove Theorem 3.1.9, which gives us the solution of the system of differential equations (3.73) and studies the convergence of the proportion of particles at time t in each state when t and N tend to infinity.

Proof of Theorem 3.1.9. The solutions of the system of differential equations (3.73) is given by

$$\begin{aligned}
\mathbf{g}_t &= \left(\int_0^t \mathbf{w}_u \mathrm{e}^{-uQ^{(2)}_p} \mathrm{d}u\right) \mathrm{e}^{tQ^{(2)}_p} \\
&= \left(\int_0^t (\mathbf{w}_u - \mathbf{w}_\infty) \mathrm{e}^{-uQ^{(2)}_p} \mathrm{d}u + \mathbf{w}_\infty \int_0^t \mathrm{e}^{-uQ^{(2)}_p} \mathrm{d}u\right) \mathrm{e}^{tQ^{(2)}_p} \\
&= \left(\int_0^t (\mathbf{w}_u - \mathbf{w}_\infty) \mathrm{e}^{-uQ^{(2)}_p} \mathrm{d}u + \mathbf{w}_\infty \left(Q^{(2)}_p\right)^{-1} \left(I - \mathrm{e}^{-tQ^{(2)}_p}\right)\right) \mathrm{e}^{tQ^{(2)}_p} \\
&= \int_0^t (\mathbf{w}_u - \mathbf{w}_\infty) \mathrm{e}^{(t-u)Q^{(2)}_p} \mathrm{d}u + \mathbf{g}_\infty \left(I - \mathrm{e}^{tQ^{(2)}_p}\right).
\end{aligned}$$

Note that the last equality comes from (3.76). Therefore, we have

$$\begin{aligned}
\|\mathbf{g}_t - \mathbf{g}_\infty\|_\infty &\leq \left\|\int_0^t (\mathbf{w}_u - \mathbf{w}_\infty) \mathrm{e}^{(t-u)Q^{(2)}_p} \mathrm{d}u\right\|_\infty + \left\|\mathbf{g}_\infty \left(\mathrm{e}^{tQ^{(2)}_p}\right)\right\|_\infty \\
&\leq \int_0^t \|\mathbf{w}_u - \mathbf{w}_\infty\|_\infty \left\|\mathrm{e}^{(t-u)Q^{(2)}_p}\right\|_\infty \mathrm{d}u + \|\mathbf{g}_\infty\|_\infty \left\|\mathrm{e}^{tQ^{(2)}_p}\right\|_\infty
\end{aligned} \tag{3.77}$$

We get

$$\left\|\mathrm{e}^{sQ^{(2)}_p}\right\|_\infty = \mathrm{e}^{-p_N s} \left\|\mathrm{e}^{sQ^{(2)}}\right\|_\infty = \mathrm{e}^{-p_N s}, \tag{3.78}$$

for all $s \geq 0$, where $p_N = \frac{2p}{N-1}$. Note that the second equality in (3.78) comes from the fact that the rows of $\mathrm{e}^{sQ^{(2)}}$ has sum equal to one, for all $s \geq 0$. Using Corollary 3.1.5, or the Cauchy–Schwarz inequality, we get

$$\|\mathbf{g}_\infty\|_\infty = \mathrm{Var}_{\nu_N}\left[\frac{\eta(0)}{N}\right]. \tag{3.79}$$

Using the inequality (3.6) we get

$$\left|s_t(k)-\frac{1}{K}\right| \leq \|\mathcal{L}_{m(\eta)}(Z_t \mid t<\tau_p)-\nu_{\mathrm{qs}}\|_2 \leq \sqrt{\frac{K-1}{K}}\mathrm{e}^{-\rho_K t},$$

for every $k\in\mathbb{Z}/K\mathbb{Z}$ and all $t\geq 0$. Therefore,

$$|\mathbf{w}_u(k,k)-\mathbf{w}_\infty(k,k)| \leq \frac{2}{N}\left(1+\theta+\frac{p}{N-1}\right)\mathrm{e}^{-\rho_K u}+\frac{2p}{N-1}\left|s_u(k)^2-\frac{1}{K^2}\right|.$$

But

$$\left|s_u(k)^2-\frac{1}{K^2}\right| = \left(s_u(k)+\frac{1}{K}\right)\left|s_u(k)-\frac{1}{K}\right| \leq \frac{K+1}{K}\sqrt{\frac{K-1}{K}}\mathrm{e}^{-\rho_K u}.$$

Thus,

$$|\mathbf{w}_u(k,k)-\mathbf{w}_\infty(k,k)| \;\leq\; \frac{2}{N}\left(1+\theta+\frac{p}{N-1}+\frac{p}{N-1}\frac{N(K+1)\sqrt{K-1}}{K\sqrt{K}}\right)\mathrm{e}^{-\rho_K u}. \tag{3.80}$$

Similarly we get,

$$|\mathbf{w}_u(k,k+1)-\mathbf{w}_\infty(k,k+1)| \;\leq\; \frac{2}{N}\left(1+\theta+\frac{p}{N-1}\frac{N(K+1)\sqrt{K-1}}{K\sqrt{K}}\right)\mathrm{e}^{-\rho_K u}, \tag{3.81}$$

$$|\mathbf{w}_u(k,l)-\mathbf{w}_\infty(k,l)| \;\leq\; \frac{2p}{N-1}\frac{(K+1)\sqrt{K-1}}{K\sqrt{K}}\mathrm{e}^{-\rho_K u}, \quad |k-l|\geq 2. \tag{3.82}$$

Inequalities (3.80), (3.81) and (3.82) imply that

$$\|\mathbf{w}_u-\mathbf{w}_\infty\|_\infty \leq C_{K,N}\,\mathrm{e}^{-\rho_K u}, \tag{3.83}$$

where $C_{K,N}$ is defined by (3.20). Plugging (3.78), (3.79) and (3.83) into (3.77), we obtain

$$\begin{aligned}
\|\mathbf{g}_t-\mathbf{g}_\infty\|_\infty &\leq C_{K,N}\int_0^t \mathrm{e}^{-\rho_K u}\mathrm{e}^{-p_N(t-u)}\mathrm{d}u+\mathrm{e}^{-p_N t}\|\mathbf{g}_\infty\|_\infty\\
&= C_{K,N}\mathrm{e}^{-p_N t}\int_0^t \mathrm{e}^{-(\rho_K-p_N)u}\mathrm{d}u+\mathrm{e}^{-p_N t}\,\mathrm{Var}_{\nu_N}\left[\frac{\eta(0)}{N}\right]\\
&= C_{K,N}\frac{\mathrm{e}^{-p_N t}-\mathrm{e}^{-\rho_K t}}{\rho_K-p_N}+\mathrm{e}^{-p_N t}\,\mathrm{Var}_{\nu_N}\left[\frac{\eta(0)}{N}\right]\\
&= C_{K,N}\frac{1-\mathrm{e}^{-\rho_K t}}{\rho_K}+\mathrm{Var}_{\nu_N}\left[\frac{\eta(0)}{N}\right]+o\left(\frac{1}{N}\right)\\
&= \frac{1}{N}\left\{D_K\frac{1-\mathrm{e}^{-\rho_K t}}{\rho_K}+E_K\right\}+o\left(\frac{1}{N}\right),
\end{aligned} \tag{3.84}$$

where D_K and E_K are given by (3.22). Note that (3.18) is obtained from (3.84).

In order to prove (3.21), note that for every initial distribution μ in $\mathbb{Z}/K\mathbb{Z}$ and any initial configuration $\eta\in\mathcal{E}_{K,N}$, we get

$$\|\overline{m}(\eta_t)-\mathcal{L}_\mu(Z_t \mid t\leq\tau_p)\|_2 \;\leq\; \mathbb{E}_\eta\left[\,\|m(\eta_t)-\mathcal{L}_\mu(Z_t \mid t\leq\tau_p)\|_2\,\right] \tag{3.85}$$

$$\leq\; \mathbb{E}_\eta\left[\,\|m(\eta_t)-\overline{m}(\eta_t)\|_2\,\right]+\|\overline{m}(\eta_t)-\mathcal{L}_\mu(Z_t \mid t\leq\tau_p)\|_2 \tag{3.86}$$

Inequality (3.85) is obtained using the convexity of the 2-norm and Jensen's inequality. Inequality (3.86) is proved using the triangular inequality. From Theorem 3.1.8 we know that for any initial configuration $\eta\in\mathcal{E}_{K,N}$, we obtain

$$\mathrm{e}^{-\rho_K t}\left|\varphi_{m(\eta)}\left(\frac{2\pi}{K}\right)-\varphi_\mu\left(\frac{2\pi}{K}\right)\right| \leq \|\overline{m}(\eta_t)-\mathcal{L}_\mu(Z_t \mid t\leq\tau_p)\|_2 \leq \mathrm{e}^{-\rho_K t}\,\|m(\eta)-\mu\|_2\,, \tag{3.87}$$

where ρ_K is given by (3.5). Also,

$$\begin{aligned}
\mathbb{E}_\eta[\,\|m(\eta_t)-\overline{m}(\eta_t)\|_2^2\,] &= \sum_{k=0}^{K-1}\mathrm{Var}_\eta\left[\frac{\eta_t(k)}{N}\right] \leq K\,\|g_t\|_\infty \\
&\leq \frac{2K}{N}\left(D_K\frac{1-\mathrm{e}^{-\rho_K t}}{\rho_K}+E_K\right)+o\left(\frac{1}{N}\right), \qquad (3.88)
\end{aligned}$$

where D_K and E_K are defined by (3.22). Finally, (3.21) is proved using (3.85), (3.86), (3.87), (3.88) and Jensen's inequality. □

3.A Proof of Lemma 3.3.2

In order to calculate $\mathcal{L}_{K,N}f_k$, note that

$$(\mathcal{L}_{K,N}f_k)(\eta) = \sum_{i,j}\eta(i)\left(\mathbb{1}_{\{j=i+1\}}+\theta\mathbb{1}_{\{j=i-1\}}+\eta(j)\frac{p}{N-1}\right)[f_k(T_{i\to j}\eta)-f_k(\eta)]\,.$$

But $f_k(T_{i\to j}\eta)=f_k(\eta)$ if $i\neq k$ and $j\neq k$. Thus,

$$\begin{aligned}
(\mathcal{L}_{K,N}f_k)(\eta) &= \eta(k)\sum_{j\neq k}\left(\mathbb{1}_{\{j=k+1\}}+\theta\mathbb{1}_{\{j=k-1\}}+\eta(j)\frac{p}{N-1}\right)[T_{k\to j}\eta(k)-\eta(k)] \\
&\quad+\sum_{i\neq k}\eta(i)\left(\mathbb{1}_{\{k=i+1\}}+\theta\mathbb{1}_{\{k=i-1\}}+\eta(k)\frac{p}{N-1}\right)[T_{i\to k}\eta(k)-\eta(k)] \\
&= -\eta(k)\left[1+\theta+p\frac{N-\eta(k)}{N-1}\right]+\eta(k-1)+\theta\eta(k+1)+p\,\eta(k)\frac{N-\eta(k)}{N-1} \\
&= \eta(k-1)-(1+\theta)\eta(k)+\theta\,\eta(k+1),
\end{aligned}$$

for all $\eta\in\mathcal{E}_{K,N}$. Thus, (3.31) is proved.

Now, for computing $\mathcal{L}_{K,N}f_{k,l}$ for all $1\leq k,l\leq K$, we separate the proof in three cases: $l=k$, $l=k+1$ and $l>k+1$, for all $0\leq k\leq K-2$.

Case $l=k$:

From (3.2) we have

$$(\mathcal{L}_{K,N}f_{k,k})(\eta) = \sum_{i,j}\eta(i)\left(\mathbb{1}_{\{j=i+1\}}+\theta\mathbb{1}_{\{j=i-1\}}+\eta(j)\frac{p}{N-1}\right)[f_{k,k}(T_{i\to j}\eta)-f_{k,k}(\eta)]\,,$$

for all $\eta\in\mathcal{E}_{K,N}$. Denote

$$S_{i,j}(\eta)=\eta(i)\left(\mathbb{1}_{\{j=i+1\}}+\theta\mathbb{1}_{\{j=i-1\}}+\eta(j)\frac{p}{N-1}\right)\left[T_{i\to j}\eta(k)^2-\eta(k)^2\right].$$

Note that if $\{i,j\}\cap\{k\}=\emptyset$, then we have $S_{i,j}(\eta)=0$. So,

$$(\mathcal{L}_{K,N}f_{k,k})(\eta)=\sum_{j\neq k}S_{k,j}(\eta)+\sum_{i\neq k}S_{i,k}(\eta).$$

Note that

$$\begin{aligned}
\sum_{j\neq k} S_{k,j}(\eta) &= \sum_{j\neq k} \eta(k)\left(\mathbb{1}_{\{j=k+1\}}+\theta\mathbb{1}_{\{j=k-1\}}+\eta(j)\frac{p}{N-1}\right)\left[T_{k\to j}\eta(k)^2-\eta(k)^2\right] \\
&= \eta(k)\left(1+\theta+\frac{p}{N-1}\sum_{j\neq k}\eta(j)\right)\left[(\eta(k)-1)^2-\eta(k)^2\right] \\
&= \left(\eta(k)+\theta\,\eta(k)+p\,\eta(k)\frac{N-\eta(k)}{N-1}\right)[-2\eta(k)+1], \qquad (3.89) \\
\sum_{i\neq k} S_{i,k}(\eta) &= \sum_{i\neq k} \eta(i)\left(\mathbb{1}_{\{k=i+1\}}+\theta\mathbb{1}_{\{k=i-1\}}+\eta(k)\frac{p}{N-1}\right)\left[T_{i\to k}\eta(k)^2-\eta(k)^2\right] \\
&= \left(\eta(k-1)+\theta\,\eta(k+1)+\eta(k)\frac{p}{N-1}\sum_{i\neq k}\eta(i)\right)\left[(\eta(k)+1)^2-\eta(k)^2\right] \\
&= \left(\eta(k-1)+\theta\,\eta(k+1)+p\,\eta(k)\frac{N-\eta(k)}{N-1}\right)[2\eta(k)+1]. \qquad (3.90)
\end{aligned}$$

Summing (3.89) and (3.90), we obtain

$$\begin{aligned}
(\mathcal{L}_{K,N}f_{k,k})(\eta) &= \sum_{j\neq k} S_{k,j}(\eta)+\sum_{i\neq k} S_{i,k}(\eta) \\
&= 2\eta(k)\left[\eta(k-1)-\eta(k)+\theta(\eta(k+1)-\eta(k))\right] \\
&\quad +(\eta(k)+\eta(k-1))+\theta(\eta(k+1)+\eta(k))+2p\,\eta(k)\frac{N-\eta(k)}{N-1} \\
&= 2\left[\eta(k-1)\eta(k)-\left(1+\theta+\frac{p}{N-1}\right)\eta(k)^2+\theta\eta(k)\eta(k+1)\right] \\
&\quad +\eta(k-1)+\left(1+\theta+\frac{2pN}{N-1}\right)\eta(k)+\theta\eta(k+1),
\end{aligned}$$

for all $\eta\in\mathcal{E}_{K,N}$. Thus, (3.32) holds.

Case $l=k+1$:

From (3.2), similarly to the previous case, we have

$$(\mathcal{L}_{K,N}f_{k,k+1})(\eta)=\sum_{i,j}\eta(i)\left(\mathbb{1}_{\{j=i+1\}}+\theta\mathbb{1}_{\{j=i-1\}}+\eta(j)\frac{p}{N-1}\right)[f_{k,k+1}(T_{i\to j}\eta)-f_{k,k+1}(\eta)].$$

Denote

$$R_{i,j}(\eta)=\eta(i)\left(\mathbb{1}_{\{j=i+1\}}+\theta\mathbb{1}_{\{j=i-1\}}+\eta(j)\frac{p}{N-1}\right)[T_{i\to j}\eta(k)\,T_{i\to j}\eta(k+1)-\eta(k)\eta(k+1)].$$

If $\{i,j\}\cap\{k,k+1\}=\emptyset$, then $R_{i,j}=0$. Thus,

$$(\mathcal{L}_{K,N}f_{k,k+1})(\eta)=\sum_{j\neq k}R_{k,j}(\eta)+\sum_{i\neq k,k+1}R_{i,k+1}(\eta)+\sum_{j\neq k+1}R_{k+1,j}(\eta)+\sum_{i\neq k,k+1}R_{i,k}(\eta).$$

Note that

$$
\begin{aligned}
\sum_{j\neq k} R_{k,j}(\eta) &= R_{k,k+1}(\eta) + \sum_{j\neq k,k+1} R_{k,j}(\eta) \\
&= \eta(k)[(\eta(k)-1)(\eta(k+1)+1) - \eta(k)\eta(k+1)]\left[1 + p\frac{\eta(k+1)}{N-1}\right] \\
&\quad + \sum_{j\neq k,k+1} \eta(k)[(\eta(k)-1)\eta(k+1) - \eta(k)\eta(k+1)] \\
&\quad \times \left(\mathbb{1}_{\{j=k+1\}} + \theta\mathbb{1}_{\{j=k-1\}} + \eta(j)\frac{p}{N-1}\right) \\
&= \eta(k)[\eta(k) - \eta(k+1) - 1]\left[1 + p\frac{\eta(k+1)}{N-1}\right] \\
&\quad -\eta(k)\eta(k+1)\left(\theta + \frac{p}{N-1}\sum_{j\neq k,k+1}\eta(j)\right) \\
&= \eta(k)[\eta(k)-1]\left[1 + p\frac{\eta(k+1)}{N-1}\right] \\
&\quad -\eta(k)\eta(k+1)(1+\theta) - p\,\eta(k)\eta(k+1)\frac{N-\eta(k)}{N-1}, \\
\sum_{i\neq k,k+1} R_{i,k+1}(\eta) &= \sum_{i\neq k,k+1} \eta(i)[\eta(k)(\eta(k+1)+1) - \eta(k)\eta(k+1)] \\
&\quad \times \left(\mathbb{1}_{\{k+1=i+1\}} + \theta\mathbb{1}_{\{k+1=i-1\}} + \eta(k+1)\frac{p}{N-1}\right) \\
&= \eta(k)\left(\theta\eta(k+2) + p\,\eta(k+1)\frac{\sum_{i\neq k,k+1}\eta(i)}{N-1}\right) \\
&= \theta\eta(k)\eta(k+2) + p\,\eta(k)\eta(k+1)\frac{N-\eta(k)-\eta(k+1)}{N-1}, \\
\sum_{j\neq k+1} R_{k+1,j}(\eta) &= R_{k+1,k}(\eta) + \sum_{j\neq k,k+1} R_{k+1,j}(\eta) \\
&= \eta(k+1)[(\eta(k)+1)(\eta(k+1)-1) - \eta(k)\eta(k+1)]\left[\theta + p\frac{\eta(k)}{N-1}\right] \\
&\quad + \sum_{j\neq k,k+1} \eta(k+1)[\eta(k)(\eta(k+1)-1) - \eta(k)\eta(k+1)] \\
&\quad \times \left(\mathbb{1}_{\{j=k+2\}} + \theta\mathbb{1}_{\{j=k\}} + p\frac{\eta(j)}{N-1}\right) \\
&= \eta(k+1)[\eta(k+1) - \eta(k) - 1]\left[\theta + p\frac{\eta(k)}{N-1}\right] \\
&\quad -\eta(k)\eta(k+1)\left(1 + \frac{p}{N-1}\sum_{j\neq k,k+1}\eta(j)\right) \\
&= \eta(k+1)[\eta(k+1)-1]\left[\theta + p\frac{\eta(k)}{N-1}\right] - \eta(k)\eta(k+1)(1+\theta) \\
&\quad -p\,\eta(k)\eta(k+1)\frac{N-\eta(k+1)}{N-1},
\end{aligned}
$$

$$\begin{aligned}\sum_{i\neq k,k+1} R_{i,k}(\eta) &= \sum_{i\neq k,k+1} \eta(i)[(\eta(k)+1)\eta(k+1)-\eta(k)\eta(k+1)]\\ &\quad\times\left(\mathbb{1}_{\{k=i+1\}}+\theta\mathbb{1}_{\{k=i-1\}}+\eta(k)\frac{p}{N-1}\right)\\ &= \eta(k+1)\left(\eta(k-1)+p\,\eta(k)\frac{N-\eta(k)-\eta(k+1)}{N-1}\right)\\ &= \eta(k-1)\eta(k+1)+p\,\eta(k)\eta(k+1)\frac{N-\eta(k)-\eta(k+1)}{N-1}.\end{aligned}$$

Then,

$$\begin{aligned}(\mathcal{L}_{K,N}f_{k,k+1})(\eta) &= -\eta(k)\eta(k+1)\left[2(1+\theta)+p\frac{2N-\eta(k)-\eta(k+1)-2[N-\eta(k)-\eta(k+1)]}{N-1}\right]\\ &\quad+\eta(k)[\eta(k)-1]\left(1+p\frac{\eta(k+1)}{N-1}\right)+\eta(k+1)[\eta(k+1)-1]\left(\theta+p\frac{\eta(k)}{N-1}\right)\\ &\quad+\eta(k-1)\eta(k+1)+\theta\eta(k)\eta(k+2)\\ &= -\eta(k)\eta(k+1)\left[2(1+\theta)+p\frac{\eta(k)+\eta(k+1)}{N-1}\right]\\ &\quad+\eta(k)[\eta(k)-1]\left(1+p\frac{\eta(k+1)}{N-1}\right)+\eta(k+1)[\eta(k+1)-1]\left(\theta+p\frac{\eta(k)}{N-1}\right)\\ &\quad+\eta(k-1)\eta(k+1)+\theta\eta(k)\eta(k+2)\\ &= -2\eta(k)\eta(k+1)(1+\theta)+\eta(k)[\eta(k)-1]+\theta\eta(k+1)[\eta(k+1)-1]\\ &\quad-2p\frac{\eta(k)\eta(k+1)}{N-1}+\eta(k-1)\eta(k+1)+\theta\eta(k)\eta(k+2)\\ &= -2\left(1+\theta+\frac{p}{N-1}\right)\eta(k)\eta(k+1)+\eta(k-1)\eta(k+1)\\ &\quad+\theta\eta(k+1)^2+\eta(k)^2+\theta\eta(k)\eta(k+2)-\eta(k)-\theta\eta(k+1),\end{aligned}$$

for all $\eta\in\mathcal{E}_{K,N}$, which is equivalent to (3.33).

Case $l>k+1$:

In this case we have

$$(\mathcal{L}_{K,N}f_{k,l})(\eta)=\sum_{i,j\in F}\eta(i)\left(\mathbb{1}_{\{j=i+1\}}+\theta\mathbb{1}_{\{j=i-1\}}+\eta(j)\frac{p}{N-1}\right)\left[f_{k,l}(T_{i\to j}\eta)-f_{k,l}(\eta)\right].$$

Denote

$$T_{i,j}(\eta)=\eta(i)\left(\mathbb{1}_{\{j=i+1\}}+\theta\mathbb{1}_{\{j=i-1\}}+\eta(j)\frac{p}{N-1}\right)[T_{i\to j}\eta(k)\,T_{i\to j}\eta(l)-\eta(k)\eta(l)].$$

Obviously, if $\{i,j\}\cap\{k,k+l\}=\emptyset$, then $T_{i,j}(\eta)=0$. Thus

$$(\mathcal{L}_{K,N}f_{k,l})(\eta)=\sum_{j\neq k}T_{k,j}(\eta)+\sum_{i\neq k,k+l}T_{i,k+l}(\eta)+\sum_{j\neq k+l}T_{k+l,j}(\eta)+\sum_{i\neq k,k+l}T_{i,k}(\eta).$$

Note that

$$
\begin{aligned}
\sum_{j\neq k} T_{k,j}(\eta) &= T_{k,k+l}(\eta) + \sum_{j\neq k,k+l} T_{k,j}(\eta) \\
&= \eta(k)[(\eta(k)-1)(\eta(k+l)+1) - \eta(k)\eta(k+l)]p\frac{\eta(k+l)}{N-1} \\
&\quad + \sum_{j\neq k,k+l} \eta(k)[(\eta(k)-1)\eta(k+l) - \eta(k)\eta(k+l)] \\
&\quad \times \left(\mathbb{1}_{\{j=k+1\}} + \theta\mathbb{1}_{\{j=k-1\}} + p\frac{\eta(j)}{N-1}\right) \\
&= \eta(k)[\eta(k)-\eta(k+l)-1]p\frac{\eta(k+l)}{N-1} \\
&\quad -\eta(k)\eta(k+l)\left(1+\theta+\frac{p}{N-1}\sum_{j\neq k,k+l}\eta(j)\right) \\
&= \eta(k)\eta(k+l)\left[\frac{p}{N-1}(\eta(k)-1) - (1+\theta) - p\frac{N-\eta(k)}{N-1}\right], \\
\sum_{i\neq k,k+l} T_{i,k+l}(\eta) &= \sum_{i\neq k,k+l} \eta(i)\left[\eta(k)(\eta(k+l)+1) - \eta(k)\eta(k+l)\right] \\
&\quad \times \left(\mathbb{1}_{\{k+l=i+1\}}\frac{1}{K} + \mathbb{1}_{\{k+l=i-1\}}\frac{\theta}{K} + \eta(k+l)\frac{p}{N-1}\right) \\
&= \eta(k)\left[\eta(k+l-1) + \theta\eta(k+l+1) + p\,\eta(k+l)\frac{N-\eta(k)-\eta(k+l)}{N-1}\right] \\
&= \eta(k)\eta(k+l-1) + \theta\eta(k)\eta(k+l+1) + p\,\eta(k)\eta(k+l)\frac{N-\eta(k)-\eta(k+l)}{N-1}, \\
\sum_{j\neq k+l} T_{k+l,j}(\eta) &= T_{k+l,k}(\eta) + \sum_{j\neq k,k+l} T_{k+l,j}(\eta) \\
&= \eta(k+l)\left[(\eta(k)+1)(\eta(k+l)-1) - \eta(k)\eta(k+l)\right]p\frac{\eta(k)}{N-1} \\
&\quad + \sum_{j\neq k,k+l} \eta(k+l)[\eta(k)(\eta(k+l)-1) - \eta(k)\eta(k+l)] \\
&\quad \times \left(\mathbb{1}_{\{j=k+l+1\}} + \theta\mathbb{1}_{\{j=k+l-1\}} + p\frac{\eta(j)}{N-1}\right) \\
&= \eta(k+l)\left[\eta(k+l)-\eta(k)-1\right]p\frac{\eta(k)}{N-1} \\
&\quad -\eta(k)\eta(k+l)\left[1+\theta+\frac{p}{N-1}\sum_{j\neq k,k+l}\eta(j)\right] \\
&= \eta(k+l)\left[\eta(k+l)-\eta(k)-1\right]p\frac{\eta(k)}{N-1} \\
&\quad -\eta(k)\eta(k+l)\left[1+\theta+p\frac{N-\eta(k)-\eta(k+l)}{N-1}\right] \\
&= \eta(k)\eta(k+l)\left[(\eta(k+l)-1)\frac{p}{N-1} - \frac{1+\theta}{N} - p\frac{N-\eta(k+l)}{N-1}\right], \\
\sum_{i\neq k,k+l} T_{i,k}(\eta) &= \sum_{i\neq k,k+l} \eta(i)[(\eta(k)+1)\eta(k+l) - \eta(k)\eta(k+l)] \\
&\quad \times \left(\mathbb{1}_{\{k=i+1\}} + \theta\mathbb{1}_{\{k=i-1\}} + p\frac{\eta(k)}{N-1}\right) \\
&= \eta(k+l)\left[\eta(k-1) + \theta\eta(k+1) + p\,\eta(k)\frac{N-\eta(k)-\eta(k+l)}{N-1}\right] \\
&= \eta(k-1)\eta(k+l) + \theta\eta(k+1)\eta(k+l) + p\,\eta(k)\eta(k+l)\frac{N-\eta(k)-\eta(k+l)}{N-1}.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
(\mathcal{L}_{K,N} f_{k,l})(\eta) &= \eta(k)\eta(k+l)\left(\frac{p}{N-1}[\eta(k)+\eta(k+l)-2]-2(1+\theta)\right.\\
&\quad\left.-\frac{p}{N-1}\left[2N-\eta(k)-\eta(k+l)-2[N-\eta(k)-\eta(k+l)]\right]\right)\\
&\quad+\eta(k)\left[\eta(k+l-1)+\theta\eta(k+l+1)\right]+\eta(k+l)\left[\eta(k-1)+\theta\eta(k+1)\right]\\
&= -2\eta(k)\eta(k+l)\left(1+\theta+\frac{p}{N-1}\right)+\eta(k)\left[\eta(k+l-1)+\theta\eta(k+l+1)\right]\\
&\quad+\eta(k+l)\left[\eta(k-1)+\theta\eta(k+1)\right],
\end{aligned}
$$

for all $\eta \in \mathcal{E}_{K,N}$, proving (3.34).

Appendix A

Notation

We will use the following notations for the sets

- $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$ of integers,
- $\mathbb{N} = \{1, 2, 3, \dots\}$ of positive integers,
- $\mathbb{N}_0 = \{0, 1, 2, 3, \dots\}$ of nonnegative integers,
- $\mathbb{R} = (-\infty, \infty)$ of real numbers,
- $\mathbb{R}_+ = [0, \infty)$ of nonnegative real numbers,
- $\mathbb{C}$ of complex numbers.

Given a discrete set Ω, we assume Ω is endowed with the sigma-algebra $\mathcal{P}(\Omega) = \{A : A \subset \Omega\}$ and we denote by

- $\mathcal{B}(\Omega)$ the set of (measurable) functions from Ω to $\mathbb{R}$,
- $\mathcal{B}_+(\Omega)$ the set of nonnegative (measurable) functions from Ω to $\mathbb{R}$,
- $\|\cdot\|$ the uniform norm defined by $\|f\| := \sup_{x\in\Omega} |f(x)| \in \mathbb{R} \cup \{+\infty\}$, for every $f \in \mathcal{B}(\Omega)$,
- $\mathcal{B}_b(\Omega)$ the set of bounded (measurable) functions from Ω to $\mathbb{R}$,
- $\mathcal{B}_1(\Omega)$ the set of (measurable) functions from Ω to $\mathbb{R}$ bounded by 1,
- $\mathcal{M}_1(\Omega)$ the set of probability measures on Ω,
- $\mu(f) := \sum_{x\in\Omega} f(x)\mu(x)$ the integral of f with respect to μ, for every $\mu \in \mathcal{M}_1(\Omega)$ and every function in $\mathcal{B}_+(\Omega)$, or every $f \in \mathcal{B}(\Omega)$ integrable with respect to μ, namely, such that $\mu(|f|) < \infty$,
- $\|\cdot\|_{\mathrm{TV}}$ the total variation norm, defined for all $\mu_1, \mu_2 \in \mathcal{M}_1(\Omega)$ by
$$\|\mu_1 - \mu_2\|_{\mathrm{TV}} := \sup_{A\subset\Omega} |\mu_1(A) - \mu_2(A)| = \frac{1}{2} \sup_{\phi\in\mathcal{B}_1(\Omega)} |\mu_1(\phi) - \mu_2(\phi)| = \frac{1}{2} \sum_{x\in\Omega} |\mu_1(x) - \mu_2(x)|.$$

Let $(x_n)_{n\in M}$, where $M \subset \mathbb{N}$ be an enumeration of the elements in Ω. We define the norm $\|\cdot\|_{\mathrm{w}}$ for all $\mu_1, \mu_2 \in \mathcal{M}_1(\Omega)$ as follows
$$\|\mu_1 - \mu_2\|_{\mathrm{w}} := \sum_{n\in M} 2^{-n} |\mu_1(x_n) - \mu_2(x_n)|.$$

Both norms $\|\cdot\|_{\mathrm{TV}}$ and $\|\cdot\|_{\mathrm{w}}$ metrize the convergence in law. Moreover, the sets $(\mathcal{B}_b(\Omega), \|\cdot\|)$, $(\mathcal{M}_1(\Omega), \|\cdot\|_{\mathrm{TV}})$ and $(\mathcal{M}_1(\Omega), \|\cdot\|_{\mathrm{w}})$ are complete spaces.

Markov chains

Given a discrete space Ω endowed with the sigma-algebra $\mathcal{P}(\Omega)$, a time homogeneous Markov chain with state space Ω is a family $\left(\Omega, (\mathcal{F}_t)_{t\geq 0}, (X_t)_{t\geq 0}, (P_t)_{t\geq 0}, (\mathbb{P}_x)_{x\in\Omega}\right)$ satisfying the conditions in [RW00, Def. III.1.1, p. 227], for $t \in \mathbb{R}_+$. We recall that $\mathbb{P}_x(X_0 = x) = 1$, that P_t is the transition function at time $t \geq 0$ of the process, and that the family $(P_t)_{t\geq 0}$ defines a semigroup of operators on the set $\mathcal{B}_b(\Omega)$.

For all $\mu \in \mathcal{M}_1(\Omega)$ and all $f \in \mathcal{B}_+(\Omega)$, we use the notations

$$\mathbb{P}_\mu(\cdot) := \sum_{x\in\Omega} \mathbb{P}_x(\cdot)\mu(x) \quad \text{and} \quad \mu\,\mathbb{P}_t f := \sum_{x\in\Omega} (P_t f)(x)\mu(x), \;\; \forall t \geq 0.$$

We denote by $\mathbb{E}_x$ and $\mathbb{E}_\mu$ the expectations associated to $\mathbb{P}_x$ and $\mathbb{P}_\mu$, respectively.

Absorbing Markov chains

Assume that $\left(X_t\right)_{t\geq 0}$ is an irreducible continuous-time Markov chain on a discrete space Ω whose generator is Q, which acts on every $\phi \in \mathcal{B}_b(E)$ as follows

$$(Q\phi) : x \to \sum_{y\in E} Q_{x,y}(\phi(y) - \phi(x)).$$

Let us denote by $\kappa \in \mathcal{B}_+(E)$, the killing rate function. Let us consider the absorbing Markov chain $(Y_t)_{t\geq 0}$ on $\Omega \cup \{\partial\}$, where $\partial \notin \Omega$ is an absorbing state, satisfying

$$Y_t = \begin{cases} X_t & \text{if } \int_0^t \kappa(X_s)\mathrm{d}s < \xi \\ \partial & \text{otherwise,} \end{cases}$$

where ξ is an exponential random variable with parameter 1, independent from $(X_t)_{t\geq 0}$. In words, $(Y_t)_{t\geq 0}$ evolves as $(X_t)_{t\geq 0}$ on E, and conditioned to be at $x \in E$, it jumps to ∂ (get absorbed) with rate $\kappa(x)$. Alternatively, $(Y_t)_{t\geq 0}$ can be defined as the process with generator $\tilde{Q}$ acting on every $\varphi \in \mathcal{B}_b(E \cup \{\partial\})$ as follows:

$$(\tilde{Q}\varphi) : x \to \kappa(x)(\varphi(\partial) - \varphi(x)) + \sum_{y\in E} Q_{x,y}(\varphi(y) - \varphi(x)),$$

for every $x \in E$, and we set $(\tilde{Q}\varphi)(\partial) = 0$. The absorption time τ_∂, defined as

$$\tau_\partial := \inf_{t\geq 0}\{X_t = \partial\},$$

is a stopping time, which means that for all $s \geq 0$, $X_s = \partial$ implies $X_t = \partial$, for all $t \geq s$. In the following we assume that the chain is absorbed almost surely:

$$\mathbb{P}_x(\tau_\partial < \infty) = 1, \;\; \forall x \in \Omega.$$

Bibliography

[Ald83] D. Aldous. “Random walks on finite groups and rapidly mixing Markov chains”. *Seminar on probability, XVII*. Vol. 986. Lecture Notes in Math. Springer, Berlin, 1983, pp. 243–297. DOI: 10.1007/BFb0068322. MR: 770418. Zbl: 0514.60067 (cit. on pp. 18, 43, 92).

[AD86] D. Aldous and P. Diaconis. “Shuffling cards and stopping times”. *Amer. Math. Monthly* 93.5 (1986), pp. 333–348. DOI: 10.2307/2323590. MR: 841111. Zbl: 0603.60006 (cit. on pp. 18, 43, 92).

[Ané+00] C. Ané et al. *Sur les inégalités de Sobolev logarithmiques.* French. Vol. 10. Panoramas et Synthèses [Panoramas and Syntheses]. With a preface by Dominique Bakry and Michel Ledoux. Société Mathématique de France, Paris, 2000, pp. xvi+217. MR: 1845806. Zbl: 0982.46026 (cit. on p. 67).

[AGJ21] L. Angeli, S. Grosskinsky, and A. M. Johansen. “Limit theorems for cloning algorithms”. *Stochastic Process. Appl.* 138 (2021), pp. 117–152. DOI: 10.1016/j.spa.2021.04.007. MR: 4256231 (cit. on pp. 11, 12, 36, 37, 56, 58–60, 66, 78).

[AD20] M. Arnaudon and P. Del Moral. “A duality formula and a particle Gibbs sampler for continuous time Feynman – Kac measures on path spaces”. *Electron. J. Probab.* 25 (2020), Paper No. 157, 54. DOI: 10.1214/20-ejp546. MR: 4193898. Zbl: 07373350 (cit. on pp. 56, 58, 67).

[AFG11] A. Asselah, P. A. Ferrari, and P. Groisman. “Quasistationary distributions and Fleming – Viot processes in finite spaces”. *J. Appl. Probab.* 48.2 (2011), pp. 322–332. DOI: 10.1239/jap/1308662630. MR: 2840302. Zbl: 1219.60081 (cit. on pp. 7, 8, 22, 24, 33, 47, 49, 53, 95, 96, 127, 130, 133).

[AT12] A. Asselah and M.-N. Thai. “A note on the rightmost particle in a Fleming – Viot process”. *arXiv e-prints* (Dec. 2012). arXiv: 1212.4168 [math.PR] (cit. on pp. 8, 34, 128).

[Ass+16] A. Asselah et al. “Fleming – Viot selects the minimal quasi-stationary distribution: the Galton–Watson case”. *Ann. Inst. Henri Poincaré, Probab. Stat.* 52.2 (2016), pp. 647–668. DOI: 10.1214/14-AIHP635. MR: 3498004. Zbl: 1342.60145 (cit. on pp. 7, 8, 33, 34, 53, 128).

[BCG20] V. Bansaye, B. Cloez, and P. Gabriel. “Ergodic behavior of non-conservative semigroups via generalized Doeblin’s conditions”. *Acta Appl. Math.* 166 (2020), pp. 29–72. DOI: 10.1007/s10440-019-00253-5. MR: 4077228. Zbl: 1442.47030 (cit. on p. 67).

[Ban+21] V. Bansaye et al. “A non-conservative Harris ergodic theorem”. Version 2. *arXiv e-prints* (Feb. 22, 2021). arXiv: 1903.03946v2 [math.AP] (cit. on pp. 67, 127).

[BLY06] J. Barrera, B. Lachaud, and B. Ycart. “Cut-off for n-tuples of exponentially converging processes”. *Stochastic Process. Appl.* 116.10 (2006), pp. 1433–1446. DOI: 10.1016/j.spa.2006.03.003. MR: 2260742. Zbl: 1103.60023 (cit. on pp. 18, 43, 92).

[BC15] M. Benaïm and B. Cloez. “A stochastic approximation approach to quasi-stationary distributions on finite spaces”. *Electron. Commun. Probab.* 20 (2015). Id/No 37, p. 13. DOI: 10.1214/ECP.v20-3956. MR: 3352332. Zbl: 1321.65009 (cit. on pp. 10, 22, 36, 47, 56, 65).

[Bré20] P. Brémaud. *Markov chains. Gibbs fields, Monte Carlo simulation and queues.* Vol. 31. Texts in Applied Mathematics. Cham: Springer, 2020, pp. xvi + 557. DOI: 10.1007/978-3-030-45982-6. MR: 4174390. Zbl: 1435.60003 (cit. on pp. 19, 43, 111).

[BHM00] K. Burdzy, R. Hołyst, and P. March. “A Fleming – Viot Particle Representation of the Dirichlet Laplacian”. *Comm. Math. Phys.* 214.3 (2000), pp. 679–703. DOI: 10.1007/s002200000294. MR: 1800866. Zbl: 0982.60078 (cit. on pp. 7, 33, 95, 127).

[Cér+20] F. Cérou et al. “A central limit theorem for Fleming – Viot particle systems”. *Ann. Inst. Henri Poincaré Probab. Stat.* 56.1 (2020), pp. 637–666. DOI: 10.1214/19-AIHP976. MR: 4059003. Zbl: 1447.82021 (cit. on pp. 7, 33, 53, 61).

[CV16] N. Champagnat and D. Villemonais. “Exponential convergence to quasi-stationary distribution and Q-process”. *Probab. Theory Related Fields* 164.1-2 (2016), pp. 243–283. DOI: 10.1007/s00440-014-0611-7. MR: 3449390. Zbl: 1334.60015 (cit. on pp. 7, 32, 63–65, 127).

[CV17a] N. Champagnat and D. Villemonais. “General criteria for the study of quasi-stationarity”. *arXiv e-prints* (2017). arXiv: 1712.08092v2 [math.PR] (cit. on pp. 67, 127).

[CV17b] N. Champagnat and D. Villemonais. “Uniform convergence to the Q-process”. *Electron. Commun. Probab.* 22 (2017). Id/No 33, p. 7. DOI: 10.1214/17-ECP63. MR: 3663104. Zbl: 1368.60079 (cit. on pp. 7, 32, 33, 60, 63, 64).

[CV20a] N. Champagnat and D. Villemonais. “Practical criteria for R-positive recurrence of unbounded semigroups”. *Electron. Commun. Probab.* 25 (2020), Paper No. 6, 11. DOI: 10.1214/20-ecp288. MR: 4066299. Zbl: 1434.60173 (cit. on p. 67).

[CV20b] N. Champagnat and D. Villemonais. “Erratum: Practical criteria for R-positive recurrence of unbounded semigroups”. *Electron. Commun. Probab.* 25 (2020), Paper No. 31, 2. DOI: 10.1214/20-ecp288. MR: 4089738. Zbl: 1441.60055 (cit. on p. 67).

[CV21] N. Champagnat and D. Villemonais. “Convergence of the Fleming – Viot process toward the minimal quasi-stationary distribution”. *ALEA, Lat. Am. J. Probab. Math. Stat.* 18.1 (2021), pp. 1–15. DOI: 10.30757/alea.v18-01. MR: 4198866 (cit. on pp. 7, 8, 33, 34, 53, 65, 67, 128, 134).

[Che06] G.-Y. Chen. “The cutoff phenomenon for finite Markov chains”. PhD thesis. Cornell University, 2006. URL: https://hdl.handle.net/1813/3047 (cit. on pp. 18, 20, 43, 45, 92, 93).

[CHS17] G.-Y. Chen, J.-M. Hsu, and Y.-C. Sheu. “The L^2-cutoffs for reversible Markov chains”. *Ann. Appl. Probab.* 27.4 (2017), pp. 2305–2341. DOI: 10.1214/16-AAP1260. MR: 3693527. Zbl: 1374.60130 (cit. on pp. 18, 43, 92).

[CK18] G.-Y. Chen and T. Kumagai. "Cutoffs for product chains". *Stochastic Process. Appl.* 128.11 (2018), pp. 3840–3879. DOI: 10.1016/j.spa.2018.01.002. MR: 3860012. Zbl: 1409.60106 (cit. on pp. 18, 43, 92).

[CS08] G.-Y. Chen and L. Saloff-Coste. "The cutoff phenomenon for ergodic Markov processes". *Electron. J. Probab.* 13 (2008), no. 3, 26–78. DOI: 10.1214/EJP.v13-474. MR: 2375599. Zbl: 1190.60007 (cit. on pp. 18–20, 43, 44, 92, 93).

[CS10] G.-Y. Chen and L. Saloff-Coste. "The L^2-cutoff for reversible Markov processes". *J. Funct. Anal.* 258.7 (2010), pp. 2246–2315. DOI: 10.1016/j.jfa.2009.10.017. MR: 2584746. Zbl: 1190.60007 (cit. on pp. 18, 43, 92).

[CC21] B. Cloez and J. Corujo. "Uniform in time propagation of chaos for a Moran model". Version 2. *arXiv e-prints* (July 22, 2021). arXiv: 2107.10794v2 [math.PR] (cit. on pp. 9, 35, 51).

[CT16a] B. Cloez and M.-N. Thai. "Fleming – Viot processes: two explicit examples". *ALEA Lat. Am. J. Probab. Math. Stat.* 13.1 (2016), pp. 337–356. DOI: 10.30757/alea.v13-14. MR: 3487076. Zbl: 1337.60241 (cit. on pp. 10, 17, 22, 23, 36, 43, 47, 56, 65, 90, 96, 128, 130, 134, 135, 139).

[CT16b] B. Cloez and M.-N. Thai. "Quantitative results for the Fleming – Viot particle system and quasi-stationary distributions in discrete space". *Stochastic Process. Appl.* 126.3 (2016), pp. 680–702. DOI: 10.1016/j.spa.2015.09.016. MR: 3452809. Zbl: 1333.60200 (cit. on pp. 7, 8, 13, 22, 24, 25, 33, 34, 38, 47, 49, 53, 59, 65, 66, 96, 97, 109, 127, 128, 130, 135).

[CMS13] P. Collet, S. Martínez, and J. San Martín. *Quasi-stationary distributions. Markov chains, diffusions and dynamical systems*. Probability and its Applications (New York). Springer, Heidelberg, 2013, pp. xvi+280. DOI: 10.1007/978-3-642-33131-2. MR: 2986807. Zbl: 1261.60002 (cit. on pp. 6, 32, 62, 127).

[Com74] L. Comtet. *Advanced combinatorics*. enlarged. The art of finite and infinite expansions. D. Reidel Publishing Co., Dordrecht, 1974, pp. xi+343. MR: 0460128. Zbl: 0283.05001 (cit. on p. 129).

[Con10] S. B. Connor. "Separation and coupling cutoffs for tuples of independent Markov processes". *ALEA Lat. Am. J. Probab. Math. Stat.* 7 (2010), pp. 65–77. MR: 2644042. Zbl: 1276.60004. URL: http://alea.impa.br/articles/v7/07-04.pdf (cit. on pp. 18, 43, 92).

[Cor17] F. Cordero. "The deterministic limit of the Moran model: a uniform central limit theorem". *Markov Process. Relat. Fields* 23.2 (2017), pp. 313–324. MR: 3701545. Zbl: 1379.92035 (cit. on pp. 2, 10, 28, 35, 53, 55, 65).

[Cor21a] J. Corujo. "Dynamics of a Fleming – Viot type particle system on the cycle graph". *Stochastic Process. Appl.* 136 (2021), pp. 57–91. DOI: 10.1016/j.spa.2021.02.001. MR: 4229456 (cit. on pp. 22, 24, 47, 49, 59, 96, 108, 109, 127).

[Cor21b] J. Corujo. "On the spectrum and ergodicity of a neutral multi-allelic Moran model". Version 2. *arXiv e-prints* (May 25, 2021). arXiv: 2010.08809v2 [math.PR] (cit. on pp. 14, 40, 65, 85, 130).

[CEZ11] D. Couty, J. Esterle, and R. Zarouf. "Décomposition effective de Jordan – Chevalley". French. *Gaz. Math.* 129 (2011), pp. 29–49. MR: 2850401. Zbl: 1251.15018 (cit. on p. 105).

[DS67] J. N. Darroch and E. Seneta. “On quasi-stationary distributions in absorbing continuous-time finite Markov chains”. *J. Appl. Probab.* 4 (1967), pp. 192–196. DOI: 10.2307/3212311. MR: 212866. Zbl: 0168.16303 (cit. on pp. 7, 32, 63, 131).

[Dav79] P. J. Davis. *Circulant matrices.* A Wiley-Interscience Publication, Pure and Applied Mathematics. John Wiley & Sons, New York-Chichester-Brisbane, 1979, pp. xv+250. MR: 543191. Zbl: 0418.15017 (cit. on pp. 101, 136, 142, 150).

[Del04] P. Del Moral. *Feynman – Kac formulae. Genealogical and interacting particle systems with applications.* New York, NY: Springer, 2004, pp. xviii + 555. DOI: 10.1007/978-1-4684-9393-1. MR: 2044973. Zbl: 1130.60003 (cit. on pp. 4, 5, 11, 30, 37, 54, 57).

[DG99] 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 (1999), pp. 429–434. DOI: 10.1016/S0764-4442(00)88619-X. MR: 1710091. Zbl: 0935.92001 (cit. on pp. 7, 33, 127).

[DG01] P. Del Moral and A. Guionnet. “On the stability of interacting processes with applications to filtering and genetic algorithms”. *Ann. Inst. Henri Poincaré, Probab. Stat.* 37.2 (2001), pp. 155–194. DOI: 10.1016/S0246-0203(00)01064-5. MR: 1819122. Zbl: 0990.60005 (cit. on pp. 12, 37, 58).

[DJ18] P. Del Moral and A. Jasra. “A note on random walks with absorbing barriers and sequential Monte Carlo methods”. *Stoch. Anal. Appl.* 36.3 (2018), pp. 413–442. DOI: 10.1080/07362994.2017.1412264. MR: 3784140. Zbl: 1390.82057 (cit. on pp. 22, 47).

[DM00a] P. Del Moral and L. Miclo. “A Moran particle system approximation of Feynman – Kac formulae”. *Stochastic Process. Appl.* 86.2 (2000), pp. 193–216. DOI: 10.1016/S0304-4149(99)00094-0. MR: 1741805. Zbl: 1030.65004 (cit. on pp. 7, 33, 127, 135).

[DM00b] P. Del Moral and L. Miclo. “Branching and interacting particle systems approximations of Feynman – Kac formulae with applications to nonlinear filtering”. *Séminaire de Probabilités XXXIV*. Vol. 1729. Lecture Notes in Math. Berlin: Springer, 2000, pp. 1–145. DOI: 10.1007/BFb0103798. MR: 1768060. Zbl: 0963.60040 (cit. on pp. 10, 12, 36, 37, 56, 59, 82, 95).

[DM02] P. Del Moral and L. Miclo. “On the stability of nonlinear Feynman – Kac semigroups”. *Annales de la Faculté des sciences de Toulouse : Mathématiques* Ser. 6, 11.2 (2002), pp. 135–175. MR: 1988460. Zbl: 02052899. URL: http://www.numdam.org/item/AFST_2002_6_11_2_135_0/ (cit. on p. 63).

[DM03] P. Del Moral and L. Miclo. “Particle approximations of Lyapunov exponents connected to Schrödinger operators and Feynman – Kac semigroups”. *ESAIM Probab. Stat.* 7 (2003), pp. 171–208. DOI: 10.1051/ps:2003001. MR: 1956078. Zbl: 1040.81009 (cit. on pp. 7, 33, 53, 61, 82).

[DPR11] P. Del Moral, F. Patras, and S. Rubenthaler. “Convergence of U-statistics for interacting particle systems”. *J. Theor. Probab.* 24.4 (2011), pp. 1002–1027. DOI: 10.1007/s10959-011-0355-6. MR: 2851242. Zbl: 1233.82025 (cit. on pp. 10, 36, 56).

[DP17] P. Del Moral and S. Penev. *Stochastic processes. From applications to theory.* Boca Raton, FL: CRC Press, 2017, pp. xlviii, 865. MR: 3618157. Zbl: 1368.60001 (cit. on pp. 19, 43, 111).

[DG14] P. Diaconis and R. Griffiths. "An introduction to multivariate Krawtchouk polynomials and their applications". *J. Statist. Plann. Inference* 154 (2014), pp. 39–53. DOI: 10.1016/j.jspi.2014.02.004. MR: 3258404. Zbl: 1306.60003 (cit. on pp. 18, 43, 92, 114).

[DG19] P. Diaconis and R. C. Griffiths. "Reproducing kernel orthogonal polynomials on the multinomial distribution". *J. Approx. Theory* 242 (2019), pp. 1–30. DOI: 10.1016/j.jat.2019.01.007. MR: 3915331. Zbl: 1428.33023 (cit. on pp. 21, 22, 47, 95).

[DS81] P. Diaconis and M. Shahshahani. "Generating a random permutation with random transpositions". *Z. Wahrsch. Verw. Gebiete* 57.2 (1981), pp. 159–179. DOI: 10.1007/BF00535487. MR: 626813. Zbl: 0485.60006 (cit. on pp. 18, 43, 92).

[DR00] P. Donnelly and E. R. Rodrigues. "Convergence to stationarity in the Moran model". *J. Appl. Probab.* 37.3 (2000), pp. 705–717. DOI: 10.1017/s002190020001593x. MR: 1782447. Zbl: 0968.60079 (cit. on pp. 21, 46, 94).

[Doo91] E. A. van Doorn. "Quasi-stationary distributions and convergence to quasi-stationarity of birth-death processes". *Adv. in Appl. Probab.* 23.4 (1991), pp. 683–700. DOI: 10.2307/1427670. MR: 1133722. Zbl: 0736.60076 (cit. on pp. 65, 66).

[DP13] E. A. van Doorn and P. K. Pollett. "Quasi-stationary distributions for discrete-state models". *European J. Oper. Res.* 230.1 (2013), pp. 1–14. DOI: 10.1016/j.ejor.2013.01.032. MR: 3063313. Zbl: 1317.60093 (cit. on pp. 6, 32, 62, 127).

[DMT95] D. Down, S. P. Meyn, and R. L. Tweedie. "Exponential and uniform ergodicity of Markov processes". *Ann. Probab.* 23.4 (1995), pp. 1671–1691. MR: 1379163. Zbl: 0852.60075. URL: https://www.jstor.org/stable/2244810 (cit. on p. 66).

[Dur08] R. Durrett. *Probability models for DNA sequence evolution*. Second. Probability and its Applications (New York). Springer, New York, 2008, pp. xii+431. DOI: 10.1007/978-0-387-78168-6. MR: 2439767. Zbl: 1311.92007 (cit. on pp. 1, 2, 27, 28, 53, 95, 96).

[Dur19] R. Durrett. *Probability—theory and examples*. Vol. 49. Cambridge Series in Statistical and Probabilistic Mathematics. Fifth edition of [MR1068527]. Cambridge University Press, Cambridge, 2019, pp. xii+419. DOI: 10.1017/9781108591034. MR: 3930614. Zbl: 1440.60001 (cit. on p. 131).

[Eth11] A. Etheridge. *Some mathematical models from population genetics*. Vol. 2012. Lecture Notes in Mathematics. Lectures from the 39th Probability Summer School held in Saint-Flour, 2009, École d'Été de Probabilités de Saint-Flour. [Saint-Flour Probability Summer School]. Springer, Heidelberg, 2011, pp. viii+119. DOI: 10.1007/978-3-642-16632-7. MR: 2759587. Zbl: 1320.92003 (cit. on pp. 2, 17, 28, 42, 53, 88, 90, 95, 96, 111).

[EG09] A. M. Etheridge and R. C. Griffiths. "A coalescent dual process in a Moran model with genic selection". *Theor. Popul. Biol.* 75.4 (2009), pp. 320–330. DOI: 10.1016/j.tpb.2009.03.004. Zbl: 1213.92038 (cit. on pp. 1, 27, 85, 86, 91, 97, 124, 127).

[EK93] S. N. Ethier and T. G. Kurtz. "Fleming–Viot processes in population genetics". *SIAM J. Control Optim.* 31.2 (1993), pp. 345–386. DOI: 10.1137/0331019. MR: 1205982. Zbl: 0774.60045 (cit. on pp. 91, 96).

[FMS14] E. A. Feinberg, M. Mandava, and A. N. Shiryaev. "On solutions of Kolmogorov's equations for nonhomogeneous jump Markov processes". *J. Math. Anal. Appl.* 411.1 (2014), pp. 261–270. DOI: 10.1016/j.jmaa.2013.09.043. MR: 3118483. Zbl: 1328.60192 (cit. on p. 69).

[Fen10] S. Feng. *The Poisson – Dirichlet distribution and related topics*. Probability and its Applications (New York). Models and asymptotic behaviors. Springer, Heidelberg, 2010, pp. xiv+218. DOI: 10.1007/978-3-642-11194-5. MR: 2663265. Zbl: 1214.60001 (cit. on pp. 7, 33, 95).

[FM07] P. Ferrari and N. Marić. "Quasi Stationary Distributions and Fleming – Viot processes in countable spaces". *Electron. J. Probab.* 12 (2007), no. 24, 684–702. DOI: 10.1214/EJP.v12-415. MR: 2318407. Zbl: 1127.60088 (cit. on pp. 7, 8, 33, 53, 57, 95, 96, 127, 130, 134).

[FV79] W. H. Fleming and M. Viot. "Some measure – valued Markov processes in population genetics theory". *Indiana Univ. Math. J.* 28.5 (1979), pp. 817–843. DOI: 10.1512/iumj.1979.28.28058. MR: 542340. Zbl: 0444.60064 (cit. on pp. 7, 33, 95).

[Fro12] G. Frobenius. "Über Matrizen aus nicht negativen Elementen". German. *Berl. Ber.* 1912 (1912), pp. 456–477. Zbl: 43.0204.09 (cit. on pp. 7, 32).

[Gri14] R. C. Griffiths. "The Λ-Fleming – Viot process and a connection with Wright – Fisher diffusion". *Adv. in Appl. Probab.* 46.4 (2014), pp. 1009–1035. DOI: 10.1239/aap/1418396241. MR: 3290427. Zbl: 1305.60038 (cit. on p. 88).

[GS13] R. C. Griffiths and D. Spanò. "Orthogonal polynomial kernels and canonical correlations for Dirichlet measures". *Bernoulli* 19.2 (2013), pp. 548–598. DOI: 10.3150/11-BEJ403. MR: 3037164. Zbl: 1281.60015 (cit. on pp. 114, 115).

[GJ13] P. Groisman and M. Jonckheere. "Simulation of quasi-stationary distributions on countable spaces". *Markov Process. Related Fields* 19.3 (2013), pp. 521–542. MR: 3156964. Zbl: 1321.60210 (cit. on pp. 7, 8, 33, 53).

[Gut13] A. Gut. *Probability: a graduate course*. Second. Springer Texts in Statistics. Springer, New York, 2013, pp. xxvi+600. DOI: 10.1007/978-1-4614-4708-5. MR: 2977961. Zbl: 1267.60001 (cit. on pp. 10, 36, 55).

[HM87] G. Harris and C. Martin. "The roots of a polynomial vary continuously as a function of the coefficients". *Proc. Amer. Math. Soc.* 100.2 (1987), pp. 390–392. DOI: 10.2307/2045978. MR: 884486. Zbl: 0619.30008 (cit. on p. 104).

[HS19] J. Hermon and J. Salez. "A version of Aldous' spectral-gap conjecture for the zero range process". *Ann. Appl. Probab.* 29.4 (2019), pp. 2217–2229. DOI: 10.1214/18-AAP1449. MR: 3984254 (cit. on pp. 95, 130).

[IX07] P. Iliev and Y. Xu. "Discrete orthogonal polynomials and difference equations of several variables". *Adv. Math.* 212.1 (2007), pp. 1–36. DOI: 10.1016/j.aim.2006.09.012. MR: 2319761. Zbl: 1133.47027 (cit. on p. 114).

[Ism05] M. E. H. Ismail. *Classical and quantum orthogonal polynomials in one variable*. Vol. 98. Encyclopedia of Mathematics and its Applications. With two chapters by Walter Van Assche, With a foreword by Richard A. Askey. Cambridge University Press, Cambridge, 2005, pp. xviii+706. DOI: 10.1017/CBO9781107325982. MR: 2191786. Zbl: 1082.42016 (cit. on p. 112).

[JS87] J. Jacod and A. N. Shiryaev. *Limit theorems for stochastic processes*. Vol. 288. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, 1987, pp. xviii+601. DOI: 10.1007/978-3-662-02514-7. MR: 959133. Zbl: 0635.60021 (cit. on p. 82).

[JKK05] N. L. Johnson, A. W. Kemp, and S. Kotz. *Univariate discrete distributions*. Third. Wiley Series in Probability and Statistics. Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, 2005, pp. xx+646. DOI: 10.1002/0471715816. MR: 2163227. Zbl: 1092.62010 (cit. on p. 91).

[JKB97] N. L. Johnson, S. Kotz, and N. Balakrishnan. *Discrete multivariate distributions*. Wiley Series in Probability and Statistics: Applied Probability and Statistics. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1997, pp. xxii+299. MR: 1429617. Zbl: 0868.62048 (cit. on p. 124).

[Kal21] O. Kallenberg. *Foundations of modern probability*. Third. Vol. 99. Probability Theory and Stochastic Modelling. Springer, Cham, 2021, p. 946. DOI: 10.1007/978-3-030-61871-1. MR: 4226142 (cit. on p. 70).

[KM65] S. Karlin and J. McGregor. "Ehrenfest urn models". *J. Appl. Probability* 2 (1965), pp. 352–376. DOI: 10.2307/3212199. MR: 184284. Zbl: 0143.40501 (cit. on pp. 18, 43, 92).

[KM75] S. Karlin and J. McGregor. "Linear growth models with many types and multidimensional Hahn polynomials". *Theory and application of special functions (Proc. Advanced Sem., Math. Res. Center, Univ. Wisconsin, Madison, Wis., 1975)*. 1975, 261–288. Math. Res. Center, Univ. Wisconsin, Publ. No. 35. MR: 0406574. Zbl: 0361.60071 (cit. on pp. 18, 43, 92).

[Kel79] F. P. Kelly. *Reversibility and stochastic networks*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Ltd., Chichester, 1979, pp. viii+230. MR: 554920. Zbl: 0422.60001 (cit. on pp. 123, 139).

[KZ09] K. Khare and H. Zhou. "Rates of convergence of some multivariate Markov chains with polynomial eigenfunctions". *Ann. Appl. Probab.* 19.2 (2009), pp. 737–777. DOI: 10.1214/08-AAP562. MR: 2521887. Zbl: 1171.60016 (cit. on pp. 18, 19, 21, 22, 43, 44, 47, 92, 93, 95, 114, 115).

[KLS10] R. Koekoek, P. A. Lesky, and R. F. Swarttouw. *Hypergeometric orthogonal polynomials and their q-analogues*. Springer Monographs in Mathematics. With a foreword by Tom H. Koornwinder. Springer-Verlag, Berlin, 2010, pp. xx+578. DOI: 10.1007/978-3-642-05014-5. MR: 2656096. Zbl: 1200.33012 (cit. on pp. 112, 113).

[Lac15] H. Lacoin. "A product chain without cutoff". *Electron. Commun. Probab.* 20 (2015), no. 19, 9. DOI: 10.1214/ECP.v20-3765. MR: 3320407. Zbl: 1321.60147 (cit. on pp. 18, 43, 92).

[LPR18] T. Lelièvre, L. Pillaud-Vivien, and J. Reygner. "Central limit theorem for stationary Fleming–Viot particle systems in finite spaces". *ALEA, Lat. Am. J. Probab. Math. Stat.* 15.2 (2018), pp. 1163–1182. DOI: 10.30757/alea.v15-43. MR: 3860819. Zbl: 1414.60016 (cit. on pp. 7, 8, 14, 22, 33, 34, 39, 47, 53, 61, 128).

[LP17] D. A. Levin and Y. Peres. *Markov chains and mixing times*. Second edition of [MR2466937], With contributions by Elizabeth L. Wilmer, With a chapter on "Coupling from the past" by James G. Propp and David B. Wilson. American Mathematical Society, Providence, RI, 2017, pp. xvi+447. DOI: 10.1090/mbk/107. MR: 3726904. Zbl: 1390.60001 (cit. on pp. 16, 18, 19, 41, 43, 89, 92, 101, 108, 109, 111, 120, 129).

[LSY99] Z. Li, T. Shiga, and L. Yao. "A reversibility problem for Fleming – Viot processes". *Electron. Commun. Probab.* 4 (1999), pp. 71–82. DOI: 10.1214/ECP.v4-1007. MR: 1711591. Zbl: 0926.60043 (cit. on p. 91).

[LPV03] L. Lovász, J. Pelikán, and K. Vesztergombi. *Discrete mathematics*. Undergraduate Texts in Mathematics. Elementary and beyond. Springer-Verlag, New York, 2003, pp. x+290. DOI: 10.1007/b97469. MR: 1952453. Zbl: 1059.00001 (cit. on p. 120).

[Mar15] N. Marić. "Fleming – Viot particle system driven by a random walk on $\mathbb{N}$". *J. Stat. Phys.* 160.3 (2015), pp. 548–560. DOI: 10.1007/s10955-015-1275-0. MR: 3366092. Zbl: 1360.82035 (cit. on pp. 8, 34, 128).

[MMV14] 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 (2014), pp. 756–768. DOI: 10.1239/jap/1409932672. MR: 3256225. Zbl: 1326.37005 (cit. on p. 65).

[MH03] J. C. Mason and D. C. Handscomb. *Chebyshev polynomials*. Chapman & Hall/CRC, Boca Raton, FL, 2003, pp. xiv+341. MR: 1937591. Zbl: 1015.33001 (cit. on p. 132).

[MV12] S. Méléard and D. Villemonais. "Quasi-stationary distributions and population processes". *Probab. Surv.* 9 (2012), pp. 340–410. DOI: 10.1214/11-PS191. MR: 2994898. Zbl: 1261.92056 (cit. on pp. 6, 7, 23, 32, 48, 62–65, 96, 127–129).

[MS19] M. Merle and J. Salez. "Cutoff for the mean-field zero-range process". *Ann. Probab.* 47.5 (2019), pp. 3170–3201. DOI: 10.1214/19-AOP1336. MR: 4021248. Zbl: 1448.60159 (cit. on pp. 10, 36, 56).

[Möh18] M. Möhle. "A spectral decomposition for the block counting process and the fixation line of the beta$(3, 1)$-coalescent". *Electron. Commun. Probab.* 23 (2018), Paper No. 102, 15. DOI: 10.1214/18-ECP203. MR: 3896840. Zbl: 1406.60104 (cit. on p. 95).

[Möh19] M. Möhle. "A spectral decomposition for a simple mutation model". *Electron. Commun. Probab.* 24 (2019), Paper No. 15, 14. DOI: 10.1214/19-ECP222. MR: 3933039. Zbl: 1412.60106 (cit. on p. 95).

[MP14] M. Möhle and H. Pitters. "A spectral decomposition for the block counting process of the Bolthausen-Sznitman coalescent". *Electron. Commun. Probab.* 19 (2014), no. 47, 11. DOI: 10.1214/ECP.v19-3464. MR: 3246966. Zbl: 1334.60157 (cit. on p. 95).

[Mor58] P. A. P. Moran. "Random processes in genetics". *Proc. Cambridge Philos. Soc.* 54 (1958), pp. 60–71. DOI: 10.1017/s0305004100033193. MR: 127989. Zbl: 0091.15701 (cit. on pp. 15, 40, 86, 95).

[Mos62] J. E. Mosimann. "On the compound multinomial distribution, the multivariate β-distribution, and correlations among proportions". *Biometrika* 49 (1962), pp. 65–82. DOI: 10.1093/biomet/49.1-2.65. MR: 143299. Zbl: 0105.12502 (cit. on p. 91).

[MW09] C. A. Muirhead and J. Wakeley. "Modeling multiallelic selection using a Moran model". *Genetics* 182.4 (2009), pp. 1141–1157. DOI: 10.1534/genetics.108.089474 (cit. on pp. 2, 3, 28, 53, 96, 123).

[NRA06] J. Navarro, J. M. Ruiz, and Y. del Aguila. "Multivariate weighted distributions: a review and some extensions". *Statistics* 40.1 (2006), pp. 51–64. DOI: 10.1080/02331880500439691. MR: 2207404. Zbl: 1098.62068 (cit. on p. 124).

[NO21] E. Nestoridi and S. Olesker-Taylor. "Limit Profiles for Markov Chains". Version 2. *arXiv e-prints* (Apr. 2021). arXiv: 2005.13437v2 [math.PR] (cit. on p. 117).

[NN14] F. Nielsen and R. Nock. "On the chi square and higher-order chi distances for approximating f-divergences". *IEEE Signal Processing Letters* 21.1 (2014), pp. 10–13. DOI: 10.1109/LSP.2013.2288355 (cit. on p. 92).

[Pea65] Marshall C. Pease III. *Methods of matrix algebra*. Mathematics in Science and Engineering. Vol. 16. Academic Press, New York-London, 1965, pp. xviii+406. MR: 0207719. Zbl: 0145.03701 (cit. on pp. 101, 150).

[Per07] O. Perron. "Zur Theorie der Matrices". German. *Math. Ann.* 64 (1907), pp. 248–263. DOI: 10.1007/BF01449896. MR: 1511438. Zbl: 38.0202.01 (cit. on pp. 7, 32).

[RL01] Yao-Feng Ren and Han-Ying Liang. "On the best constant in Marcinkiewicz–Zygmund inequality". *Statist. Probab. Lett.* 53.3 (2001), pp. 227–233. DOI: 10.1016/S0167-7152(01)00015-3. MR: 1841623. Zbl: 0991.60011 (cit. on p. 84).

[RW00] L. C. G. Rogers and David Williams. *Diffusions, Markov processes, and martingales. Vol. 1*. Cambridge Mathematical Library. Foundations, Reprint of the second (1994) edition. Cambridge University Press, Cambridge, 2000, pp. xx+386. DOI: 10.1017/CBO9781107590120. MR: 1796539. Zbl: 0949.60003 (cit. on p. 160).

[Rou06] M. Rousset. "On the control of an interacting particle estimation of Schrödinger ground states". *SIAM J. Math. Anal.* 38.3 (2006), pp. 824–844. DOI: 10.1137/050640667. MR: 2262944. Zbl: 1174.60045 (cit. on pp. 10–12, 36, 37, 56–59, 61, 66, 70, 71, 75, 78).

[Sal21] J. Salez. *Temps de mélange des chaînes de Markov*. French. Accessed 21 July 2021. 2021. URL: https://www.ceremade.dauphine.fr/~salez/mixing.pdf (cit. on pp. 18, 43).

[Sal97] L. Saloff-Coste. "Lectures on finite Markov chains". *Lectures on probability theory and statistics (Saint-Flour, 1996)*. Vol. 1665. Lecture Notes in Math. Springer, Berlin, 1997, pp. 301–413. DOI: 10.1007/BFb0092621. MR: 1490046. Zbl: 0885.60061 (cit. on pp. 18, 43, 92, 111).

[SH17] D. Schrempf and A. Hobolth. "An alternative derivation of the stationary distribution of the multivariate neutral Wright–Fisher model for low mutation rates with a view to mutation rate estimation from site frequency data". *Theor. Popul. Biol.* 114 (2017), pp. 88–94. Zbl: 1369.92075 (cit. on p. 119).

[Ser10] D. Serre. *Matrices*. Second. Vol. 216. Graduate Texts in Mathematics. Theory and applications. Springer, New York, 2010, pp. xiv+289. DOI: 10.1007/978-1-4419-7683-3. MR: 2744852. Zbl: 1206.15001 (cit. on pp. 103–105).

[SRW15] O. Szehr, D. Reeb, and M. M. Wolf. "Spectral convergence bounds for classical and quantum Markov processes". *Comm. Math. Phys.* 333.2 (2015), pp. 565–595. DOI: 10.1007/s00220-014-2188-5. MR: 3296158. Zbl: 1323.60095 (cit. on p. 108).

[Vil14] D. Villemonais. "General approximation method for the distribution of Markov processes conditioned not to be killed". *ESAIM Probab. Stat.* 18 (2014), pp. 441–467. DOI: 10.1051/ps/2013045. MR: 3333998. Zbl: 1310.82032 (cit. on pp. 7, 8, 10, 33, 34, 36, 53, 56, 130, 135).

[Vil15] D. Villemonais. "Minimal quasi-stationary distribution approximation for a birth and death process". *Electron. J. Probab.* 20 (2015), no. 30, 18. DOI: 10.1214/EJP.v20-3482. MR: 3325100. Zbl: 1376.37019 (cit. on pp. 8, 34, 67, 128).

[Vil20] D. Villemonais. "Lower bound for the coarse Ricci curvature of continuous-time pure-jump processes". *J. Theoret. Probab.* 33.2 (2020), pp. 954–991. DOI: 10.1007/s10959-019-00918-9. MR: 4091580. Zbl: 1451.60025 (cit. on pp. 96, 130).

[Wat61] G. A. Watterson. "Markov chains with absorbing states: A genetic example". *Ann. Math. Statist.* 32 (1961), pp. 716–729. DOI: 10.1214/aoms/1177704967. MR: 125633. Zbl: 0108.30803 (cit. on p. 106).

[Yag47] A. M. Yaglom. "Certain limit theorems of the theory of branching random processes". Russian. *Doklady Akad. Nauk SSSR (N.S.)* 56 (1947), pp. 795–798. MR: 0022045. Zbl: 0041.45602 (cit. on pp. 6, 32).

[Yca99] B. Ycart. "Cutoff for samples of Markov chains". *ESAIM Probab. Statist.* 3 (1999), pp. 89–106. DOI: 10.1051/ps:1999104. MR: 1716128. Zbl: 0932.60077 (cit. on pp. 18, 43, 92).

[Zho08] H. Zhou. "Examples of Multivariate Markov Chains with Orthogonal Polynomial Eigenfunctions". PhD Thesis. Stanford University, 2008 (cit. on pp. 19, 44, 88, 93, 106, 119).

[ZL09] H. Zhou and K. Lange. "Composition Markov chains of multinomial type". *Adv. in Appl. Probab.* 41.1 (2009), pp. 270–291. DOI: 10.1239/aap/1240319585. MR: 2514954. Zbl: 1161.60023 (cit. on pp. 18, 43, 86, 87, 91, 92, 95, 101, 119).

RÉSUMÉ

L'objectif principal de cette thèse est d'étudier l'évolution, en temps long et pour une grande taille de population, des modèles de Moran multi-alléliques, qui sont des processus de Markov à temps continu et à espace discret, inspirés de modèles mathématiques pour la biologie. Nous nous intéressons à l'étude, entre autres aspects, de la relation entre le processus de Moran, compris comme un système de particules en interaction, et la théorie des distributions quasi-stationnaires. Plus précisément, nous exhibons des phénomènes de propagation du chaos lorsque la taille de la population est grande, et nous établissons des contrôles quantitatifs de la convergence en temps long vers l'équilibre. Les principaux résultats sont divisés en trois chapitres. Dans le premier chapitre, nous montrons que la mesure de probabilité empirique induite par le système de particules converge, lorsque la taille de la population est grande, vers la loi d'une chaîne de Markov absorbante conditionnée à ne pas être absorbée. De plus, nous établissons un contrôle de cette convergence, en prouvant une propagation du chaos uniforme en temps. Nous prouvons également la normalité asymptotique du biais et nous fournissons une expression explicite pour la variance asymptotique, utilisée ensuite pour définir un autre système de particules avec une erreur quadratique plus petite. Dans le deuxième chapitre, nous considérons un modèle plus simple où l'espace d'état est fini et le taux de mortalité est uniforme. Dans ce contexte, nous trouvons une expression explicite pour le spectre du générateur du système de particules en termes de spectre de la matrice des taux de mutation. De plus, nous étudions l'ergodicité du processus et, pour un schéma particulier de mutation, mutation indépendante des parents, nous sommes en mesure de prouver l'existence de phénomènes de cutoff pour les distances de variation totale et chi-deux. Le troisième chapitre est consacré à l'étude d'un cas particulier, où le processus de mutation correspond à une marche aléatoire asymétrique sur le graphe cyclique. Nous montrons que ce modèle possède une solvabilité remarquable, malgré le fait qu'il soit non-réversible avec une distribution invariante non-explicite.

MOTS CLÉS

Processus de Moran multi-allélique; systèmes de particules de type Fleming-Viot ; distribution quasi-stationnaire ; propagation du chaos ; ergodicité ; phénomène de cutoff

ABSTRACT

The main goal of this thesis is to study the evolution of a multi-allelic Moran model, which is a continuous-time discrete state Markov process, inspired by biological applications. We study, among many other aspects, the relation between the Moran process, understood as an interacting particle system, and the theory of quasi-stationary distributions. More precisely, we prove the existence of a propagation of chaos phenomenon when the population size is large, and we study the quantitative control for the long time convergence to stationarity by spectral arguments. The main results are divided in three chapters. In the first chapter we show that the empirical probability measure induced by the particle system converges, when the number of particles goes to infinity, to the law of an absorbing Markov process conditioned to non-absorption. Furthermore, we establish a control on this convergence, by proving a uniform in time propagation of chaos. We also prove the asymptotic normality of the bias and we provide an explicit expression for the asymptotic variance, which is later used to define another particle system with smaller quadratic error. In the second chapter, we consider a simpler model where the state space is finite and the killing rate is uniform. In this context we find an explicit expression for the spectrum of the particle system generator in terms of the spectrum of the mutation rate matrix. Moreover, we study the ergodicity of the process and, for a particular mutation scheme, which is the parent independent mutation, we are able to prove the existence of cutoff phenomena in the total variation and chi-square distances. The third chapter is devoted to the study of a particular case, where the mutation process is driven by an asymmetric random walk on the cycle graph. We show that this model has a remarkable exact solvability, despite the fact that it is non-reversible with non-explicit invariant distribution.

KEYWORDS

Moran process; Fleming-Viot particle system; quasi-stationary distribution; propagation of chaos; ergodicity; cutoff phenomenon