Manuscrit présenté pour l'obtention de

l'Habilitation à Diriger des Recherches

Convergence exponentielle vers une distribution quasi-stationnaire et applications

---

Denis Villemonais

Membres du jury

Michel Benaïm, Université de Neuchâtel (examinateur)

Jean Bertoin, Université de Zürich (examinateur)

Pierre Collet, École Polytechnique (rapporteur)

Anne Gégout-Petit, Université de Lorraine (examinatrice)

Bénédicte Haas, Université Paris 13 (rapporteuse)

Tony Lelièvre, École des Ponts ParisTech (examinateur)

Sylvie Méléard, École Polytechnique (examinatrice)

Ross Pinsky, Technion (rapporteur)

Soutenance le 28 novembre 2019
à l'Institut Élie Cartan de Lorraine
Université de Lorraine, Site de Nancy, France

Contents

Remerciements

En premier lieu, je tiens à remercier Bénédicte Haas, Pierre Collet et Ross Pinsky qui ont tous trois pris le temps et fait l'effort de rapporter ce manuscrit. C'est un grand plaisir et un grand honneur d'avoir pu profiter de leurs précieux conseils et remarques. Merci également à Michel Benaïm, Jean Bertoin, Anne Gégout-Petit, Tony Lelièvre et Sylvie Méléard pour avoir accepté de faire partie du jury. Merci également à Nicolas Champagnat, qui m'a fait l'honneur de parrainer cette habilitation. Merci enfin à David Steinsaltz pour sa participation à ce projet.

La théorie des distributions quasi-stationnaires, que j'ai étudiée et continue d'étudier avec passion, m'a été dévoilée par Sylvie Méléard durant ma thèse. Je la remercie de m'avoir initié a ce sujet et pour son soutien toutes ses années. Merci également à Pierre Del Moral, Servet Martínez et Jaime San Martín pour leurs explications sur ce sujet alors que je terminais tout juste ma thèse il y a 8 ans. Ce document résume des travaux qui ont pour une large part été réalisés en collaboration avec Nicolas Champagnat, que je remercie pour sa présence et les heures innombrables que nous avons passées à discuter de mathématiques et, souvent, de distributions quasi-stationnaires. Je le remercie donc pour tout ce qu'il m'a apporté et continue de m'apporter, car il est certain que sa présence a été un facteur prépondérant dans le bon déroulement de ma recherche en mathématique. Je souhaite également remercier mon collaborateur William Oçafrain, le premier étudiant avec lequel j'ai travaillé pendant et suite à son parcours recherche à l'École des Mines de Nancy, et dont le travail de recherche aujourd'hui confirme les dispositions mathématiques que j'avais cru déceler chez lui il y a quelques années. Merci également à Camille Coron qui m'a patiemment expliqué certains aspects des probabilités génétiques. Merci aux membres de l'université Bath, dont Cécile Mailler qui m'a introduit aux modèles d'urnes, à Alexandre Cox, Andreas Kyprianou et Emma Horton qui m'ont accueilli dans leur exploration de l'équation de transport neutronal : ils m'ont apporté une ouverture nouvelle vers les processus de branchement et Emma m'a fait le plaisir de me rejoindre cette année en post-doc à Nancy. Merci aussi à Michel Benaïm pour la générosité dont il a fait preuve lors de nos rencontres à Neuchâtel et dans les Alpes suisses, pour m'avoir patiemment expliqué le fonctionnement des algorithmes d'approximation stochastique et, enfin et surtout, pour m'avoir recommandé son brillant étudiant Édouard Strickler, qui nous a fait le plaisir de venir collaborer avec Nicolas et moi dans le cadre d'un post-doctorat à l'IECL. Merci également à Koléhè Coulibaly-Pasquier pour sa collaboration et son aide en probabilités, et pour son éternelle bonne humeur qui éclaire nos journées à l'IECL. Merci également à René Schott, pour m'avoir introduit aux algorithmes distribués et aux applications possibles des distributions quasi-stationnaires dans ce domaine. Ces dernières années ont également été pour moi l'occasion de me pencher sur des problématiques nouvelles, en statistique et médecine notamment, en rejoignant un projet de travail sur des données de longueurs de télomères chez l'être humain. C'est une expérience qui a été et continue d'être enrichissante, je remercie pour cela Samy Tindel et Éliane Albuisson, sans qui je n'aurais jamais eu connaissance de cette problématique, Athanasios Benetos et Simon Toupance, qui ont pris le temps de discuter pendant de nombreuses heures de leurs données et de la mécanique des télomères, Anne Gégout-Petit, qui a porté le projet tout en lui apportant sa précieuse expérience en statistique, et enfin nos étudiants et post-docs : Daphné Germain, étudiante de l'École des Mines de Nancy avec qui ce projet a commencé et qui a su oeuvrer pour le rapprochement des deux communautés, Sylvain Chabanet, étudiant de l'École des Mines de Nancy qui a abordé l'étude d'une cohorte plus nombreuse pour tenter d'étayer nos résultats avec des méthodes pointues qu'il a brillament mises en oeuvre, Olivier Coudray, stagiaire en Master 2 qui a sû très vite s'adapter et a pu mettre en place des outils théoriques essentiels pour l'étude de modèles de populations branchantes, et Lionel Lenotre, qui nous a accompagné pendant une courte année avant d'accepter un poste de maître de conférences à Mulhouse et qui nous a apporté son expertise dans le domaine de l'analyse des semi-groupes positifs.

Je tiens également à remercier les membres de l'équipe INRIA TOSCA. Merci en premier lieu à Madalina Deaconu pour la direction de cette équipe qu'elle assurait déjà avant mon arrivée à Nancy. Sa présence au sein de ce qui a été ma première structure d'accueil après ma thèse, m'a aidé à déterminer mes premières directions de recherche de manière autonome, tout en me sentant entouré et soutenu. Je tiens également à remercier Antoine Le Jay, pour sa présence et son rôle tant au sein de cette équipe projet, qu'en tant que responsable de l'équipe de probabilités et statistiques à l'IECL, et enfin et surtout en tant que collaborateur pour le co-encadrement de mon premier doctorant Alexis Anagnostakis. Que ce dernier, qui subit ma supervision en recherche et en enseignement, soit ici remercié pour ses efforts, son endurance et ses progrès. Au sein de cette équipe, j'ai également pris un grand plaisir à échanger et travailler avec Oana Lopalescu, Khaled Salhi et Coralie Fritsch.

Le poste de maître de conférences que j'occupe à l'Institut Élie Cartan de Lorraine et à l'École des Mines de Nancy m'a été attribué en l'année 2012. Depuis, ce laboratoire et cette école m'ont apporté en continu tout le support et les encouragements que j'étais en droit d'espérer et qui m'ont permis de m'épanouir dans le métier de mathématicien, dans ses missions d'enseignement, d'encadrement et de recherche. Je tiens à ce titre à remercier tout son personnel, et en particulier leurs directeurs actuels Xavier Antoine et François Rousseau, ainsi que les personnels de soutien, dont le sérieux et l'efficacité ont permis le déroulement sans encombre de toute cette période d'activité. Mon métier n'aurait pas été le même sans la patience et l'aide de Nathalie Benito, Élodie Cunat, Didier Gemmerlé, Laurence Quirot et Paola Schneider à l'IECL, ou sans le support de Françoise Laurent et l'efficacité redoutable de Thierry Umoret à l'École des Mines de Nancy. Je les remercie ici pour tout leur travail et leur présence. Cette page est bien sûr pour moi l'occasion de remercier Antoine Henrot, pour sa direction généreuse et humaine du département GIMA, ainsi que Céline Lacaux, qui m'a guidé dans mes premiers pas à l'École des Mines de Nancy. Tous deux m'ont offert par leur présence des perspectives nouvelles tant dans mes missions d'enseignement que de recherche. Enfin, je remercie pour leur bienveillance et leur amabilité les membres de l'IECL et du département GIMA, qualités qui ont été, au jour le jour, au coeur de la satisfaction que m'a apporté le métier de mathématicien à Nancy. Qu'on me permette enfin de rendre hommage à toutes les mathématiciennes et tous les mathématiciens dont les travaux m'ont accompagnés et qui ont nourri de leurs idées originales mes travaux de recherches.

C'est alors que toutes ces collaborations et rencontres suivaient leur chemin, qu'un grand et heureux événement a changé le cours de ma vie le 26 mai 2017 : ma compagne Sophie a donné naissance à ma fille Léonore et toutes deux m'accompagnent depuis pour mon plus grand bonheur. Le rayonnement de leur amour et de leurs sourires m'a accompagné, entre autres, tout au long de la rédaction de ce manuscrit et de la préparation de cette habilitation.

Résumé

Ce manuscrit est le mémoire de mon dossier de candidature à l'*habilitation à diriger des recherches*. À ce titre, il contient un exposé représentatif de mes travaux de recherche depuis la fin de ma thèse. Son objectif est également de présenter sous une forme plus accessible les outils théoriques principaux développés dans ces travaux, ainsi que leur application à divers modèles.

Mes recherches, depuis l'obtention de ma thèse en novembre 2011, portent principalement sur l'étude du comportement quasi-stationnaire de processus avec absorption et de leurs applications. En particulier, je me suis attaché à développer des critères suffisants pour l'existence et la convergence des lois conditionnelles d'un processus vers une distribution quasi-stationnaire. Ces travaux sont exposés dans les première et seconde parties de ce manuscrit, qui contiennent respectivement les critères abstraits et leurs application à des modèles classiques. Ils correspondent aux publications et pré-publications suivantes[1], écrites dans leur plus grande partie dans le cadre d'une collaboration fructueuse avec Nicolas Champagnat à l'Institut Élie Cartan de Lorraine à Nancy.

- S. Martínez, J. San Martín, and D. Villemonais. Existence and uniqueness of a quasi- stationary distribution for Markov processes with fast return from infinity. *J. Appl. Probab.*, 51(3):756–768, 2014.

- D. Villemonais. Minimal quasi-stationary distribution approximation for a birth and death process. *Electron. J. Probab.*, 20:no. 30, 18, 2015.

- N. Champagnat and D. Villemonais. Exponential convergence to quasi-stationary distribution and Q-process. *Probab. Theory Related Fields*, 164(1):243–283, 2016.

- N. Champagnat and D. Villemonais. Uniform convergence of conditional distributions for absorbed one-dimensional diffusions. *Adv. in Appl. Probab.*, 50(1):178–203, 2017.

- N. Champagnat and D. Villemonais. General criteria for the study of quasi-stationarity. ArXiv e-prints, Dec. 2017.

- N. Champagnat and D. Villemonais. Lyapunov criteria for uniform convergence of conditional distributions of absorbed Markov processes. ArXiv e-prints, Apr. 2017.

- N. Champagnat and D. Villemonais. Uniform convergence of penalized time-inhomogeneous Markov processes. *ESAIM Probab. Stat.*, 22:129–162, 2018.

[1]Toutes mes publications sont accessibles depuis ma page web `https://www.normalesup.org/~villemonais/`

- N. Champagnat, K. A. Coulibaly-Pasquier, and D. Villemonais. Criteria for Exponential Convergence to Quasi-Stationary Distributions and Applications to Multi-Dimensional Diffusions. *Séminaire de Probabilités*, XLIX:165–182, 2018.

Ces résultats ont des conséquences qui vont au delà de la seule théorie des distributions quasi-stationnaires et mènent naturellement à l'étude de leurs applications aux propriétés quasi-ergodiques, à la R-positivité de semi-groupes non-bornés, aux processus de Pólya à valeur mesure et aux processus auto-renforcés. Ces applications sont développées dans la troisième partie de ce manuscrit et correspondent aux publications et pré-publications suivantes.

- N. Champagnat and D. Villemonais. Uniform convergence to the Q-process. *Electron. Commun. Probab.*, 22:7 pp., 2017.

- N. Champagnat and D. Villemonais. Practical criteria for R-positive recurrence of unbounded semigroups. ArXiv e-prints, Apr. 2019.

- C. Mailler and D. Villemonais. Stochastic approximation on non-compact measure spaces and application to measure-valued Pólya processes. ArXiv e-prints, Sep. 2018.

- M. Benaïm, N. Champagnat, and D. Villemonais. Stochastic approximation of quasi- stationary distributions for diffusion processes in a bounded domain. ArXiv e-prints, Apr. 2019.

Enfin, d'autres problématiques seront exposées, dont certaines ont nourri des encadrements et co-encadrements de stages de recherche long, d'une thèse et d'un post-doctorat. Ils correspondent aux publications suivantes.

- D. Villemonais. Lower bound for the coarse ricci curvature of continuous-time pure-jump processes. *J. Theoret. Probab. Probability*, May 2019.

- C. Coron, S. Méléard, and D. Villemonais. Impact of demography on extinction/fixation events. *J. Math. Biol.*, 78(3):549–577, Feb 2019.

- S. Toupance, D. Villemonais, D. Germain, A. Gegout-Petit, E. Albuisson, and A. Benetos. The individual's signature of telomere length distribution. *Sci. Rep.*, 9(1):685, 2019.

Certains travaux post-thèse ne sont pas abordés dans ce manuscrit. En particulier, les travaux sur les processus de type Fleming-Viot en tant qu'outils d'approximation des distributions quasi-stationnaires ne sont pas présentés, car dans la continuité directe de mes travaux de thèse. D'autres sujets portent sur des modèles dont la seule définition demanderait une introduction trop technique. Il s'agit d'une proportion conséquente des articles précédemments cités et de la totalité des publications et pré-publications suivantes.

- W. Oçafrain and D. Villemonais. Convergence of a non-failable mean-field particle system. *Stoch. Anal. Appl.*, 35(4):587–603, 2017.

- N. Champagnat and D. Villemonais. Exponential convergence to quasi-stationary distribution for absorbed one-dimensional diffusions with killing. *ALEA Lat. Am. J. Probab. Math. Stat.*, 14(1):177–199, 2017.

- N. Champagnat, R. Schott, and D. Villemonais. Probabilistic Non-asymptotic Analysis of Distributed Algorithms. ArXiv e-prints, Feb. 2018.
- N. Champagnat and D. Villemonais. Convergence of the Fleming-Viot process toward the minimal quasi-stationary distribution. ArXiv e-prints, Oct. 2018.
- E. Horton, A. E. Kyprianou, and D. Villemonais. Stochastic Methods for the Neutron Transport Equation I: Linear Semigroup asymptotics. ArXiv e-prints, Oct. 2018.
- A. M. G. Cox, E. L. Horton, A. E. Kyprianou, and D. Villemonais. Stochastic Methods for Neutron Transport Equation III: Generational many-to-one and $\mathtt{k}_{eff}$. ArXiv e-prints, Sept. 2019.

Enfin les travaux suivants ont été développés entièrement ou en grande partie durant ma thèse et ne sont donc pas exposés en détail dans ce mémoire. Il s'agit des publications suivantes.

- D. Villemonais. Interacting particle systems and yaglom limit approximation of diffusions with unbounded drift. *Electron. J. Probab.*, 16:1663–1692, 2011.
- S. Méléard and D. Villemonais. Quasi-stationary distributions and population processes. *Probab. Surv.*, 9:340–410, 2012.
- D. Villemonais. Uniform tightness for time-inhomogeneous particle systems and for conditional distributions of time-inhomogeneous diffusion processes. *Markov Process. Related Fields*, 19(3):543–562, 2013.
- D. Villemonais. General approximation method for the distribution of Markov processes conditioned not to be killed. *ESAIM Probab. Stat.*, 18:441–467, 2014.
- P. Del Moral and D. Villemonais. Exponential mixing properties for time inhomogeneous diffusion processes with killing. *Bernoulli*, 24(2):1010–1032, 2018.

Notons pour conlure que certains des résultats de ce mémoire sont présentés sous une forme différente de leur première publication, voire sont originaux. Toutefois, il s'agit de développements incrémentaux et les preuves ne présentent pas d'intérêt particulier. Afin de ne pas alourdir le manuscrit, les preuves de ces résultats, ainsi que celles déjà publiées, ne seront pas incluses.

Abstract

This manuscript is the dissertation of my application to the *habilitation à diriger des recherches* degree. As such, its content reflects my research activities since the end of my PhD thesis. Its purpose is also to present under a more accessible form the main theoretical tools developed in these works, as well as their applications to various models.

Since my PhD defence in November 2011, my researches focus mainly on the quasi-stationary behaviour of absorbed Markov processes and their applications. In particular, I devoted myself on developing sufficient criteria for the existence and convergence of the conditional laws of a process towards a quasi-stationary distribution. These works are described in the first and second parts of this manuscript, which respectively present the abstract criteria and their application to different models. They correspond to the following publications and pre-publications[2], most of them written as part of a fruitful collaboration with Nicolas Champagnat at the Institut Élie Cartan de Lorraine in Nancy.

- S. Martínez, J. San Martín, and D. Villemonais. Existence and uniqueness of a quasi- stationary distribution for Markov processes with fast return from infinity. *J. Appl. Probab.*, 51(3):756–768, 2014.
- D. Villemonais. Minimal quasi-stationary distribution approximation for a birth and death process. *Electron. J. Probab.*, 20:no. 30, 18, 2015.
- N. Champagnat and D. Villemonais. Exponential convergence to quasi-stationary distribution and Q-process. *Probab. Theory Related Fields*, 164(1):243–283, 2016.
- N. Champagnat and D. Villemonais. Uniform convergence of conditional distributions for absorbed one-dimensional diffusions. *Adv. in Appl. Probab.*, 50(1):178–203, 2017.
- N. Champagnat and D. Villemonais. General criteria for the study of quasi-stationarity. ArXiv e-prints, Dec. 2017.
- N. Champagnat and D. Villemonais. Lyapunov criteria for uniform convergence of conditional distributions of absorbed Markov processes. ArXiv e-prints, Apr. 2017.
- N. Champagnat and D. Villemonais. Uniform convergence of penalized time-inhomogeneous Markov processes. *ESAIM Probab. Stat.*, 22:129–162, 2018.
- N. Champagnat, K. A. Coulibaly-Pasquier, and D. Villemonais. Criteria for Exponential Convergence to Quasi-Stationary Distributions and Applications to Multi-Dimensional Diffusions. *Séminaire de Probabilités*, XLIX:165–182, 2018.

[2]All my publications can be accessed from my web-page `https://www.normalesup.org/~villemonais/`

The consequences of these results go beyond the theory of quasi-stationary distributions and naturally lead to the study of quasi-ergodic properties, of R-positive recurrence of unbounded semi-groups, to measure valued Pólya processes and to reinforced processes. These applications are developed in the second part of this manuscript and correspond to the following publications and pre-publications.

- N. Champagnat and D. Villemonais. Uniform convergence to the Q-process. *Electron. Commun. Probab.*, 22:7 pp., 2017.
- N. Champagnat and D. Villemonais. Practical criteria for R-positive recurrence of unbounded semigroups. ArXiv e-prints, Apr. 2019.
- C. Mailler and D. Villemonais. Stochastic approximation on non-compact measure spaces and application to measure-valued Pólya processes. ArXiv e-prints, Sep. 2018.
- M. Benaïm, N. Champagnat, and D. Villemonais. Stochastic approximation of quasi- stationary distributions for diffusion processes in a bounded domain. ArXiv e-prints, Apr. 2019.

Then I present other subjects of study in the fourth part, some of which have been supported by or lead to long-term research internships and the co-supervising of a PhD-thesis and a post-doctorate. They correspond to the following publications.

- D. Villemonais. Lower bound for the coarse ricci curvature of continuous-time pure-jump processes. *J. Theoret. Probab. Probability*, May 2019.
- C. Coron, S. Méléard, and D. Villemonais. Impact of demography on extinction/fixation events. *J. Math. Biol.*, 78(3):549–577, Feb 2019.
- S. Toupance, D. Villemonais, D. Germain, A. Gegout-Petit, E. Albuisson, and A. Benetos. The individual's signature of telomere length distribution. *Sci. Rep.*, 9(1):685, 2019.

Some of my post-PhD thesis works are not covered in this manuscript. In particular, works on Fleming-Viot type processes as approximation tools for quasi-stationary distributions are not presented, because they are in the direct continuity of my PhD thesis. Other subjects relate to models whose mere definition would require a very technical introduction. This represents a substantial proportion of the articles cited above and all of the subsequent publications and pre-publications.

- W. Oçafrain and D. Villemonais. Convergence of a non-failable mean-field particle system. *Stoch. Anal. Appl.*, 35(4):587–603, 2017.
- N. Champagnat and D. Villemonais. Exponential convergence to quasi-stationary distribution for absorbed one-dimensional diffusions with killing. *ALEA Lat. Am. J. Probab. Math. Stat.*, 14(1):177–199, 2017.
- N. Champagnat, R. Schott, and D. Villemonais. Probabilistic Non-asymptotic Analysis of Distributed Algorithms. ArXiv e-prints, Feb. 2018.
- N. Champagnat and D. Villemonais. Convergence of the Fleming-Viot process toward the minimal quasi-stationary distribution. ArXiv e-prints, Oct. 2018.

- E. Horton, A. E. Kyprianou, and D. Villemonais. Stochastic Methods for the Neutron Transport Equation I: Linear Semigroup asymptotics. ArXiv e-prints, Oct. 2018.
- A. M. G. Cox, E. L. Horton, A. E. Kyprianou, and D. Villemonais. Stochastic Methods for Neutron Transport Equation III: Generational many-to-one and $\mathtt{k}_{eff}$. ArXiv e-prints, Sept. 2019.

Finally, the following works have been developed entirely or for a large part during my PhD thesis and are therefore not exposed in detail in this memory. They correspond to the following publications.

- D. Villemonais. Interacting particle systems and yaglom limit approximation of diffusions with unbounded drift. *Electron. J. Probab.*, 16:1663–1692, 2011.
- S. Méléard and D. Villemonais. Quasi-stationary distributions and population processes. *Probab. Surv.*, 9:340–410, 2012.
- D. Villemonais. Uniform tightness for time-inhomogeneous particle systems and for conditional distributions of time-inhomogeneous diffusion processes. *Markov Process. Related Fields*, 19(3):543–562, 2013.
- D. Villemonais. General approximation method for the distribution of Markov processes conditioned not to be killed. *ESAIM Probab. Stat.*, 18:441–467, 2014.
- P. Del Moral and D. Villemonais. Exponential mixing properties for time inhomogeneous diffusion processes with killing. *Bernoulli*, 24(2):1010–1032, 2018.

To conclude, note that some of the results of this memoir are presented in a form that differs from their first publication, and some of them are even original. However, these are incremental developments and the proofs are not of particular interest. In order not to increase unnecessarily the size of the manuscript, the proofs of these results, as well as those already published, will be omitted.

Notations

We will use the following notations for the sets

- $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$ of integers,
- $\mathbb{Z}_+ = \{0, 1, 2, 3, \dots\}$ of non-negative integers,
- $\mathbb{N} = \{1, 2, 3, \dots\}$ of positive integers,
- $\mathbb{R} = (-\infty, +\infty)$ of real numbers,
- $\mathbb{R}_+ = [0, +\infty)$ of non-negative real numbers.

Given a measurable set $(E, \mathscr{E})$, we denote by

- $\mathscr{B}(E)$ the set of measurable functions from E to $\mathbb{R}$,
- $\mathscr{B}_b(E)$ the set of bounded measurable functions from E to $\mathbb{R}$,
- $\mathscr{B}_+(E)$ the set of non-negative measurable functions from E to $\mathbb{R}_+$,
- $\|\cdot\|_\infty$ the uniform norm on $\mathscr{B}_b(E)$,
- $\mathscr{M}(E)$ the set of non-negative measures on E,
- $\mathscr{M}_1(E)$ the set of probability measures on E,
- $\mu(f)$ the integral of f with respect to μ, defined for all $\mu \in E$ and all $f \in \mathscr{B}_+(E)$ or all $f \in \mathscr{B}(E)$ such that f is integrable with respect to μ (i.e. $\mu(|f|) < +\infty$),
- $\|\cdot\|_{TV}$ the *total variation norm*, defined for all $\mu_1, \mu_2 \in \mathscr{M}_1(E)$ by

$$\left\|\mu_1 - \mu_2\right\|_{TV} = 2\sup_{A \in \mathscr{E}} |\mu_1(A) - \mu_2(A)| = \sup\left\{|\mu_1(f) - \mu_2(f)| : \ f \in \mathscr{B}_b(E), \ \|f\|_\infty \leq 1\right\}.$$

Given a measurable set $(E, \mathscr{E})$ and a positive measurable function $\psi : E \to (0, +\infty)$, we define

- $\mathscr{M}(\psi) := \left\{\mu \in \mathscr{M}_1(E) : \ \mu(\psi) < +\infty\right\}$
- $\|\cdot\|_{\mathscr{M}(\psi)}$ the *weighted total variation norm*, defined for all $\mu_1, \mu_2 \in \mathscr{M}(\psi)$ by

$$\left\|\mu_1 - \mu_2\right\|_{\mathscr{M}(\psi)} := \sup\left\{|\mu_1(f) - \mu_2(f)| : \ |f| \leq \psi\right\}.$$

- $L^\infty(\psi) := \left\{f \in \mathscr{B}(E) : |f/\psi| \in \mathscr{B}_b(E)\right\}$,
- $\|f\|_{L^\infty(\psi)} := \sup_{x\in E} \frac{|f(x)|}{\psi(x)}$ for all $f \in L^\infty(\psi)$.

The sets $\left(\mathscr{M}_1(E), \|\cdot\|_{TV}\right)$, $\left(\mathscr{M}(\psi), \|\cdot\|_{\mathscr{M}(\psi)}\right)$ and $\left(L^\infty(\psi), \|\cdot\|_{L^\infty(\psi)}\right)$ are complete spaces, and $\mathscr{B}_b(E) = L^\infty(\mathbf{1}_E)$ and $\|\cdot\|_\infty = \|\cdot\|_{L^\infty(\mathbf{1}_E)}$.

Given a Polish space (E,d), we define

- $\mathscr{M}^d(E) := \left\{\mu \in \mathscr{M}(E),\ \int d(x,y)\mu(dy) < \infty\right\}$,
- the *Wasserstein distance* $\mathscr{W}_d$ between two probability measures μ and ν on E belonging to $\mathscr{M}^d(E)$, as
$$\mathscr{W}_d(\mu,\nu) = \inf_\pi \int_{E\times E} d(x,y)\,\pi(dx,dy),$$
where the infimum is taken over all probability measures π on $E\times E$ such that $\pi(\cdot,E) = \mu(\cdot)$ and $\pi(E,\cdot) = \nu(\cdot)$ (π is called a *coupling measure* for μ and ν),
- the Wasserstein distance $\mathscr{W}_d$ between two measures in $\mathscr{M}^d(E)$ with the same mass: for all $\alpha > 0$ and any probability measures μ,ν on E belonging to $\mathscr{M}^d(E)$, we set
$$\mathscr{W}_d(\alpha\mu,\alpha\nu) = \inf_\pi \int_{E\times E} d(x,y)\,\pi(dx,dy) = \alpha\mathscr{W}_d(\mu,\nu),$$
where the infimum is taken over all measures π on $E\times E$ with mass α and such that $\pi(\cdot,E^N) = \mu(\cdot)$ and $\pi(E^N,\cdot) = \nu(\cdot)$. Note that if a coupling π realizes the minimum in the definition of $\mathscr{W}_d(\mu,\nu)$, then $\alpha\pi$ realizes the minimum in the definition of $\mathscr{W}_d(\alpha\mu,\alpha\nu)$.

We emphasize that the state space $\left(\mathscr{P}_d(E^N),\mathscr{W}_d\right)$ is a complete state space (see for instance Lemma 5.2 and Theorem 5.4 in [69]).

Given a measurable set $(E,\mathscr{E})$, a *time homogeneous Markov process with state space E* is a family $(\Omega, (\mathscr{F}_t)_{t\geq 0}, (X_t)_{t\geq 0}, (P_t)_{t\geq 0}, (\mathbb{P}_x)_{x\in E\cup\{\partial\}})$ satisfying the conditions of [234, Definition III.1.1], in the discrete or continuous time settings (i.e. $t \in \mathbb{Z}_+$ or $t\in\mathbb{R}_+$). 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 semi-group of operators on the set $\mathscr{B}_b(E)$. For all $\mu \in \mathscr{M}_1(E)$ and all $f \in \mathscr{B}_+(E)$, we will use the notations
$$\mathbb{P}_\mu(\cdot) := \int_{E\cup\{\partial\}} \mathbb{P}_x(\cdot)\mu(dx) \quad \text{and} \quad \mu P_t f := \int_{E\cup\partial} P_t f(x)\,\mu(dx),\ \forall t \geq 0.$$
We shall denote by $\mathbb{E}_x$ (resp. $\mathbb{E}_\mu$) the expectation corresponding to $\mathbb{P}_x$ (resp. $\mathbb{P}_\mu$). We will say that the process is *absorbed* at a point ∂ if, for all all $s \geq 0$, $X_s = \partial$ implies $X_t = \partial$ for all $t \geq s$. This implies in particular that the absorption time, defined as
$$\tau_\partial := \inf\{t \geq 0, X_t = \partial\},$$
is a stopping time. In this work, we will often assume (although not systematically) that, for all $t \geq 0$ and $\forall x \neq \partial$, $\mathbb{P}_x(t < \tau_\partial) > 0$.

Part I

Criteria for the exponential convergence to a quasi-stationary distribution

Chapter 1

A necessary and sufficient condition for uniform exponential convergence to a quasi-stationary distribution

This chapter is dedicated to the presentation of a necessary and sufficient condition for uniform exponential convergence to a quasi-stationary distribution in the total variation norm. We first recall the definition of a quasi-stationary distribution in Section 1.1, state the abstract results in Section 1.2 and give a first application to irreducible Markov chains on finite state spaces in Section 1.4.

1.1 Definition

Let $(X_t)_{t\geq 0}$ be a time-homogeneous Markov process with state space $E\cup\{\partial\}$ which is absorbed at $\partial\notin E$, in discrete or continuous time settings. A *quasi-stationary distribution* is a probability measure ν_{QSD} on E such that

$$\mathbb{P}_{\nu_{QSD}}(X_t\in A\mid t<\tau_\partial)=\nu_{QSD}(A),\quad \forall t\geq 0,\ A\in\mathscr{E}, \tag{1.1}$$

where we recall that $\tau_\partial=\inf\{t\geq 0,\ X_t=\partial\}$ is the absorption time of X. We refer the reader to the book [80] and the surveys [198, 203, 252] for several properties, analysis and historical notes on the concept of quasi-stationary distributions. In particular, it is known that a probability measure ν_{QSD} on E is a quasi-stationary distribution if and only if there exists a probability measure μ on E such that

$$\nu_{QSD}(A)=\lim_{t\to+\infty}\mathbb{P}_\mu(X_t\in A\mid t<\tau_\partial),\quad \forall A\in\mathscr{E}. \tag{1.2}$$

For a given quasi-stationary distribution ν_{QSD}, the set of probability measures μ such that (1.2) holds is called the *domain of attraction of* ν_{QSD}. It is non-empty since it contains at least ν_{QSD} and may contains an infinite number of elements. In particular, when the limit in (1.2) exists for any $\mu=\delta_x$, $x\in E$, and doesn't depend on the initial position x, then ν_{QSD} is called the *Yaglom limit* or the *minimal quasi-stationary distribution*. Thus the minimal quasi-stationary distribution, when it exists, is the unique quasi-stationary distribution whose domain of attraction contains $\{\delta_x,\ x\in E\}$.

The study of quasi-stationary distributions can be traced back to the works of Yaglom [267] on Galton-Watson processes. Later, birth and death processes have been studied in [164, 236], and finite state space processes in [86, 87]. A L^2 spectral approach was developed in [224] for the study of multi-dimensional diffusion processes and in [79] for one-dimensional diffusion processes, whose principles were later used in a long range of successful papers studying quasi-stationary distributions for diffusion processes (see for instance [10, 239, 44, 46, 175, 184, 146]). New advances based on the theory of R-positive recurrent processes have been developed in [8, 212, 213, 243], with a nice practical criterion exhibited later in [119]. Other approaches have been developed, such as the use of h-transforms in [130], an original renewal method in [120], the use of intrinsic ultra-contractivity properties in [172, 174]. Several other methods have been used to describe the quasi-stationary behaviour of stochastic models with absorption, see in particular [135] where quasi-stationary distributions with rescaling are considered and in [240] where small noise limits of quasi-stationary distributions are studied. We also refer the reader to the large bibliography established by Pollett [227] for more than 450 references on the theory of quasi-stationary distributions, classified by topics. In the present manuscript, a different approach is exposed, inspired by the classical theory of stationary processes [204]: we will use in particular modified Doblin conditions see Section 1.2, and drift criteria based on Lyapunov type functions (see Chapter 2).

Although there are similarities with the classical notion of stationary distributions (which is, in fact, a particular instance of quasi-stationary distribution), some important differences remain. For instance, the linear combination of quasi-stationary distributions is not necessarily a quasi-stationary distribution. Also, there may exist several (a continuous infinite number of) quasi-stationary distributions even for simple irreducible regular processes (this is the case for instance for linear birth and death processes [246]).

It is well known (see for instance [203]) that when ν_{QSD} is a quasi-stationary distribution, there exists $\lambda_0 \geq 0$ such that, for all $t \geq 0$,

$$\mathbb{P}_{\nu_{QSD}}(t < \tau_\partial) = e^{-\lambda_0 t} \quad \text{and} \quad e^{\lambda_0 t}\nu_{QSD}P_t = \nu_{QSD}. \tag{1.3}$$

An other remarkable fact is that the absorption times and the absorption position are independent under $\mathbb{P}_{\nu_{QSD}}$ [80]. Because of these properties and others, the concept of quasi-stationary distribution has found a wide range of applications, see for instance [98, 149, 134, 35, 262, 263].

1.2 Main result

Let us consider a time-homogeneous Markov process X with state space $E \cup \{\partial\}$ which is absorbed at $\partial \notin E$. We are interested in a necessary and sufficient condition for the existence of a unique quasi-stationary distribution ν_{QSD} on E for the process $(X_t)_{t\geq 0}$, where, in addition, the convergence in (1.2) is exponential and uniform with respect to μ and A.

Our base assumption is the following one.

Assumption A. There exists a probability measure ν on E such that

A1. there exist $t_0, c_1 > 0$ such that for all $x \in E$,

$$\mathbb{P}_x(X_{t_0} \in \cdot \mid t_0 < \tau_\partial) \geq c_1 \nu(\cdot);$$

A2. there exist $c_2 > 0$ such that for all $x \in E$ and $t \geq 0$,

$$\mathbb{P}_\nu(t < \tau_\partial) \geq c_2 \mathbb{P}_x(t < \tau_\partial).$$

In Section 1.4, we will consider the finite state space case as a simple illustrative example and show that Assumption A is satisfied as soon as the process is irreducible. More advanced applications will be provided in following chapters, namely to birth and death processes, to one dimensional diffusion processes, and to multi-dimensional diffusion processes, in Chapters 3, 4 and 5 respectively.

For conservative Markov processes (i.e. when $\mathbb{P}_x(\tau_\partial = +\infty) = 1$), one recognises in Condition A1 a Doblin condition and in c_1 a Dobrushin coefficient. In this case, the following theorem is already well known (see for instance [102, Theorem 18.2.4.] and [102, Section 18.7]) and can be proved using coupling methods. Natural extension of these methods are presented in a pedagogical way in [181]. Assumption A is thus an extension of these Doblin criteria to conditioned processes. As thus, it suffers similar drawbacks. Namely, it only applies to processes that *come down from infinity* [44, 201]. In the next chapter, we provide a criterion inspired by Meyn and Tweedie's works [205, 206, 207]), which is sufficient for the non-uniform exponential convergence to a quasi-stationary distribution.

The following result is proved in [56] with additional refinements, including equivalent assertions comparable to those provided in [204, Chapter 16]. Its proof, sketched in Section 1.3, immediately extends to the time-inhomogeneous setting (which were natively handled in [92, 93, 94]), as detailed in [63, 20]. Prior works implying the uniform exponential convergence of normalised semigroups can be found in the bibliography of Del Moral (see [92, 93] and references therein), as well as in my earlier works in [201] and[94] and in [37, 174].

Theorem 1.1. *Assumption A implies the existence of a probability measure ν_{QSD} on E such that, for any initial distribution $\mu \in \mathcal{M}_1(E)$,*

$$\left\| \mathbb{P}_\mu(X_t \in \cdot \mid t < \tau_\partial) - \nu_{QSD}(\cdot) \right\|_{TV} \leq 2(1 - c_1 c_2)^{\lfloor t/t_0 \rfloor}, \tag{1.4}$$

where $\lfloor \cdot \rfloor$ is the integer part function and $\| \cdot \|_{TV}$ is the total variation norm.

Conversely, if there exists a probability measure ν_{QSD} and positive constants γ, C such that, for all probability measures μ on E,

$$\left\| \mathbb{P}_\mu(X_t \in \cdot \mid t < \tau_\partial) - \nu_{QSD}(\cdot) \right\|_{TV} \leq C\,\mathrm{e}^{-\gamma t}, \quad \forall t \geq 0, \tag{1.5}$$

then Assumption A holds true.

In this case, for all probability measures μ_1, μ_2 on E, and for all $t > 0$,

$$\left\| \mathbb{P}_{\mu_1}(X_t \in \cdot \mid t < \tau_\partial) - \mathbb{P}_{\mu_2}(X_t \in \cdot \mid t < \tau_\partial) \right\|_{TV} \leq \frac{(1 - c_1 c_2)^{\lfloor t/t_0 \rfloor}}{c_2(\mu_1) \vee c_2(\mu_2)} \|\mu_1 - \mu_2\|_{TV},$$

here $c_2(\mu)$ is a positive constant which only depends on μ.[1]

The constants c_1 and c_2 may seem difficult to compute explicitly and it is not clear at first glance if the above quantitative rates are of practical interest (besides the fact that they provide an exponential speed of convergence). However, in [64], the authors succeed in proving appropriate time scales for the convergence to quasi-stationary distribution by estimating the parameters c_1 and

[1] This is proved in [56] with $c_2(\mu_1) \wedge c_2(\mu_2)$ instead of $c_2(\mu_1) \vee c_2(\mu_2)$, however the proof of the result with the latter stronger estimate is almost identical.

c_2. Similarly, the authors of [28] used a time inhomogeneous version of Assumption A and, estimating c_1 and c_2 in this situation, were able to prove a sub-exponential convergence rate toward a quasi-stationary distribution for a model with two communication classes.

Assumption A also has the following consequences.

Proposition 1.2. *Assume that Assumption A holds true. Then there exists a non-negative function η on $E \cup \{\partial\}$, positive on E and vanishing on ∂, defined by*

$$\eta(x) = \lim_{t\to\infty} \frac{\mathbb{P}_x(t < \tau_\partial)}{\mathbb{P}_{\nu_{QSD}}(t < \tau_\partial)} = \lim_{t\to+\infty} e^{\lambda_0 t}\mathbb{P}_x(t < \tau_\partial),$$

where the convergence holds for the uniform norm on $E \cup \{\partial\}$ and $\nu_{QSD}(\eta) = 1$. More precisely, there exists a positive constant a_1 such that

$$\left| e^{\lambda_0 t}\mathbb{P}_x(t < \tau_\partial) - \eta(x) \right| \le a_1\, e^{\lambda_0 t}\mathbb{P}_x(t < \tau_\partial)(1 - c_1 c_2)^{t/t_0}, \tag{1.6}$$

Furthermore, the function η is bounded, belongs to the domain of the infinitesimal generator L of the semi-group $(P_t)_{t\ge 0}$ on $(\mathscr{B}_b(E \cup \{\partial\}), \|\cdot\|_\infty)$ and

$$L\eta = -\lambda_0 \eta.$$

In the irreducible case, exponential ergodicity is known to be related to a spectral gap property (see for instance [180]). Our results imply a similar property under Assumption A for the infinitesimal generator L of the semi-group on $(\mathscr{B}_b(E \cup \{\partial\}), \|\cdot\|_\infty)$.

Proposition 1.3. *Suppose that Assumption A holds true. If $f \in \mathscr{B}_b(E \cup \{\partial\})$ is a right eigenfunction for L for an eigenvalue λ, then either*

1. *$\lambda = 0$ and f is constant,*
2. *or $\lambda = -\lambda_0$ and $f = \nu_{QSD}(f)\eta$,*
3. *or $\lambda \le -\lambda_0 - \gamma$, $\nu_{QSD}(f) = 0$ and $f(\partial) = 0$.*

We conclude this section with an original result concerning a refinement of the speed of convergence of the conditional distribution of the process toward its quasi-stationary distribution. Its proof is a simple adaptation of the proof of Theorem 1.1 and is omitted here. Note that $\nu(\eta)/\|\eta\|_\infty \ge c_2$ and thus the rate is an improvement over (1.4).

Proposition 1.4. *Suppose that Assumption A holds. Then there exists a constant $C > 0$ such that*

$$\left\| \mathbb{P}_\mu(X_t \in \cdot \mid t < \tau_\partial) - \nu_{QSD}(\cdot) \right\|_{TV} \le C \left(1 - c_1 \frac{\nu(\eta)}{\|\eta\|_\infty} \right)^{t/t_0}.$$

1.3 Sketch of the proof for the sufficient condition

In order to sketch the proof of the direct implication "Assumption A ⇒ exponential convergence", let us assume that X satisfies Assumption A with $t_0 = 1$ (the extension to any t_0 is immediate).

Following an idea that goes back at least to [93] and already used in [201] in the context of quasi-stationary distributions, we define, for all $0 \le s \le t \le T$, the linear operator $R^T_{s,t}$ by

$$R^T_{s,t} f(x) = \mathbb{E}_x(f(X_{t-s}) \mid T - s < \tau_\partial) = \mathbb{E}(f(X_t) \mid X_s = x,\ T < \tau_\partial),\ \forall f \in \mathscr{B}_b(E),$$

by the Markov property. For any $T > 0$, the family $(R^T_{s,t})_{0\le s\le t\le T}$ is a Markov (time-inhomogeneous) semi-group: we have, for all $0 \le u \le s \le t \le T$ and all $f \in \mathscr{B}_b(E)$,

$$R^T_{u,s}(R^T_{s,t} f)(x) = R^T_{u,t} f(x).$$

The main idea of the proof is to check that this conservative semi-group satisfies a Doblin condition (see Step 1): for all $T \ge 1$ and all $0 \le t \le T-1$, there exists a probability measure ν_{T-t} on E such that, for all measurable sets $A \subset E$ and all $x \in E$,

$$R^T_{t,t+1}(A) = \mathbb{P}_x(X_1 \in A \mid T - t < \tau_\partial) \ge c_1 c_2 \nu_{T-t}(A). \tag{1.7}$$

Once this is proved, one deduces (as in the classical time-uniform conservative case) a uniform mixing property for the conservative semi-groups R^T and then for the conditional distributions:

$$\left\| \mathbb{P}_{\mu_1}(X_T \in \cdot \mid T < \tau_\partial) - \mathbb{P}_{\mu_2}(X_T \in \cdot \mid T < \tau_\partial) \right\|_{TV} \le 2(1 - c_1 c_2)^{\lfloor T \rfloor}, \quad \forall \mu_1, \mu_2 \in \mathscr{M}_1(E). \tag{1.8}$$

This immediately implies that there is at most one quasi-stationary distribution and implies in particular that the sequence $\mathbb{P}_{\mu_1}(X_T \in \cdot \mid T < \tau_\partial)_{T\ge 0}$ is a Cauchy sequence and hence that it converges to some probability ν_{QSD} (recall that the set of probability measures endowed with the total variation norm is complete). By [203], ν_{QSD} is a quasi-stationary distribution.

Step 1: Doblin condition (1.7)
Let us show that, for all $t \ge 1$, there exists a probability measure ν_t on E such that (1.7) holds true. First, one can check that Assumption A1 and Markov property imply that

$$\mathbb{P}_x(X_1 \in A \text{ and } t < \tau_\partial) \ge c_1 \nu(\mathbf{1}_A(\cdot)\mathbb{P}_\cdot(t-1 < \tau_\partial))\, \mathbb{P}_x(1 < \tau_\partial).$$

Dividing both sides by $\mathbb{P}_x(t < \tau_\partial)$, we deduce that

$$\mathbb{P}_x(X_1 \in A \mid t < \tau_\partial) \ge c_1 \nu(\mathbf{1}_A(\cdot)\mathbb{P}_\cdot(t-1 < \tau_\partial)) \frac{\mathbb{P}_x(1 < \tau_\partial)}{\mathbb{P}_x(t < \tau_\partial)}.$$

But, using again the Markov property, we have

$$\mathbb{P}_x(t < \tau_\partial) \le \mathbb{P}_x(1 < \tau_\partial) \sup_{y\in E} \mathbb{P}_y(t-1 < \tau_\partial),$$

so that

$$\mathbb{P}_x(X_1 \in A \mid t < \tau_\partial) \ge c_1 \frac{\nu(\mathbf{1}_A(\cdot)\mathbb{P}_\cdot(t-1 < \tau_\partial))}{\sup_{y\in E} \mathbb{P}_y(t-1 < \tau_\partial)}.$$

Now Assumption (A2) implies that the non-negative measure

$$B \mapsto \frac{\nu(\mathbf{1}_B(\cdot)\mathbb{P}_\cdot(t-1 < \tau_\partial))}{\sup_{y\in E} \mathbb{P}_y(t-1 < \tau_\partial)}$$

has a total mass greater than c_2. Therefore (1.7) holds with the probability measure

$$\nu_t : B \mapsto \frac{\nu\left(\mathbf{1}_B(\cdot)\mathbb{P}_\cdot(t-1<\tau_\partial)\right)}{\mathbb{P}_\nu(t-1<\tau_\partial)}$$

Step 2: exponential contraction for the conditional distributions
Using the semi-group property of $(R^T_{s,t})_{s,t}$, we deduce that, for any $x, y \in E$ and all $0 \le t \le T$,

$$\|\delta_x R^T_{0,t} - \delta_y R^T_{0,t}\|_{TV} \le 2\,(1-c_1 c_2)^{\lfloor t \rfloor}.$$

By definition of $R^T_{0,T}$, this inequality immediately implies that

$$\left\|\mathbb{P}_x(X_T \in \cdot \mid T < \tau_\partial) - \mathbb{P}_y(X_T \in \cdot \mid T < \tau_\partial)\right\|_{TV} \le 2(1-c_1c_2)^{\lfloor T \rfloor}.$$

Since, in general, $\mathbb{P}_\mu(X_T \in \cdot \mid T < \tau_\partial)$ is not linear in μ, it is not immediate that this inequality extends to any pair of initial probability measures μ_1, μ_2 on E. However, this is easily overcome by the following computations. Let μ_1 be a probability measure on E and $x \in E$. We have

$$\begin{aligned}
&\|\mathbb{P}_{\mu_1}(X_T \in \cdot \mid T < \tau_\partial) - \mathbb{P}_x(X_T \in \cdot \mid T < \tau_\partial)\|_{TV} \\
&= \frac{1}{\mathbb{P}_{\mu_1}(T<\tau_\partial)} \|\mathbb{P}_{\mu_1}(X_T \in \cdot) - \mathbb{P}_{\mu_1}(T<\tau_\partial)\mathbb{P}_x(X_T \in \cdot \mid T<\tau_\partial)\|_{TV} \\
&\le \frac{1}{\mathbb{P}_{\mu_1}(T<\tau_\partial)} \int_{y\in E} \|\mathbb{P}_y(X_T \in \cdot) - \mathbb{P}_y(T<\tau_\partial)\mathbb{P}_x(X_T \in \cdot \mid T<\tau_\partial)\|_{TV}\, d\mu_1(y) \\
&\le \frac{1}{\mathbb{P}_{\mu_1}(T<\tau_\partial)} \int_{y\in E} \mathbb{P}_y(T<\tau_\partial)\|\mathbb{P}_y(X_T \in \cdot \mid T<\tau_\partial) - \mathbb{P}_x(X_T \in \cdot \mid T<\tau_\partial)\|_{TV}\, d\mu_1(y) \\
&\le \frac{1}{\mathbb{P}_{\mu_1}(T<\tau_\partial)} \int_{y\in E} \mathbb{P}_y(T<\tau_\partial) 2(1-c_1c_2)^{\lfloor T\rfloor}\, d\mu_1(y) \\
&\le 2(1-c_1c_2)^{\lfloor T \rfloor}.
\end{aligned}$$

The same computation, replacing δ_x by any probability measure, leads to (1.8).

Using the fact that $\mathcal{M}_1(E)$ endowed with the total variation norm is a complete space, this easily leads to (1.4).

1.4 The finite state space case

The problem of existence and uniqueness of a quasi-stationary distribution in the finite state space setting has been studied by Darroch and Seneta [86, 87]. They completely solved this problem in the irreducible state space case using Perron-Frobenius theorem, inspiring a long and rich lineage of developments for the study of quasi-stationary distributions based on spectral theoretical tools. The aim of this section is to give an application of Theorem 1.1 in a simple situation, recovering this classical result with additional explicit bounds on the rate of convergence.

Let $(X_t)_{t\in\mathbb{Z}_+}$ be a discrete time Markov process on a finite state space $E \cup \partial$, where $\partial \notin E$ is absorbing. We say that X is irreducible and aperiodic if there exists $t_0 \in \mathbb{N}$ such that, for all $x, y \in E$, $\mathbb{P}_x(X_{t_0} = y) > 0$. Darroch and Seneta obtained in [86] (see [87] for its continuous time version) that there exist two positive constants such that the exponential convergence (1.5) holds true, with γ being the second spectral gap of the transition matrix.

The following convergence result is not focused toward optimality, but rather aims at illustrating how to check Assumption A in a simple case. We observe that, associated with Proposition 1.3, it provides an explicit lower bound for the second spectral gap of the matrix P.

Proposition 1.5. *Let X be an irreducible and aperiodic Markov chain on a finite state space E with transition matrix $(P_{x,y})_{x,y\in E}$. Let $t_0 \in \mathbb{N}$ be such that P^{t_0} has positive entries and set*

$$c_1 = \sum_{y\in E} \inf_{x\in E} \frac{P^{t_0}_{x,y}}{\sum_{z\in E} P^{t_0}_{x,z}} \quad \textit{and} \quad c_2 = \inf_{x,y\in E} \frac{P^{t_0}_{x,y}}{\sum_{z\in E} P^{t_0}_{x,z}}.$$

Then X satisfies Assumption A with the constants c_1, c_2 and t_0.

We refer the reader to [198, 252] for a survey on different properties of quasi-stationary distributions in this finite state space/discrete time setting. Extension of this result to reducible discrete time Markov chains on a finite state space is developed in [54, 251]. In [77, 99], the authors consider the problem of the stochastic comparison between convergence toward a quasi-stationary distribution and convergence toward a stationary distribution for an ad hoc conservative process. Probabilistic representations of the Perron-Frobenius theorem are provided in [48, 49, 128] for finite state space Markov chains.

Since the aim of this section is to illustrate the application of Theorem 1.1, we detail the elementary proof of the above proposition.

Proof. We define the probability measure ν on E by

$$\nu(\{y\}) = \inf_{x\in E} \frac{P^{t_0}_{x,y}}{c_1 \sum_{z\in E} P^{t_0}_{x,z}}, \quad \forall y \in E.$$

We have for all $x, y \in E$,

$$\mathbb{P}_x(X_{t_0} = y \mid t_0 < \tau_\partial) = \frac{P^{t_0}_{x,y}}{\sum_{z\in E} P^{t_0}_{xz}} \geq c_1 \nu(\{y\}),$$

which entails Assumption A1. Now, for all $x \in E$ and all $n \geq 2$,

$$\mathbb{P}_\nu(n < \tau_\partial) \geq \nu(\{x\}) \mathbb{P}_x(n < \tau_\partial) \geq c_2 \mathbb{P}_x(n < \tau_\partial),$$

which implies Assumption A2. □

Chapter 2

Non-uniform convergence toward a quasi-stationary distribution

In this chapter, we present a sufficient criterion ensuring the exponential convergence of the conditional distribution of Markov processes toward a quasi-stationary distribution. Contrarily to the criteria of the previous part, we obtain non-uniform convergence with respect to the initial distribution. This allows to derive new existence and convergence results for a far larger class of processes, since it does not require the conditioned process to come back from infinity. In particular, this result applies to birth and death chains and Galton-Watson processes, which do not enter the general settings of Chapter 1. In general, our results also apply to processes admitting several quasi-stationary distributions, which is known to happen in a variety of specific cases, even for processes irreducible in E (including branching processes [236, 12, 177, 188], one-dimensional birth and death processes [246, 121, 120, 259] and one-dimensional diffusion processes [185, 197]).

The results presented below, but in Section 2.2, first appeared in [58].

2.1 Main results

We present here our main assumption and main results, in the discrete time and continuous time settings. In order to illustrate them, we develop in Section 2.2 a simple application to birth and death chains. More involved applications, including comparison to the results of [119] based on R-positive matrix theory, application to Galton-Watson processes, and application to perturbed dynamical systems are presented in Sections 2.3, 2.4 and 2.5 respectively. Applications to continuous time processes are presented in the next chapters.

2.1.1 Discrete time models

Let $(X_n)_{n\in\mathbb{Z}_+}$ be a Markov process in $E\cup\{\partial\}$ where E is a measurable space and $\partial\not\in E$ In this section, we study the sub-Markovian transition semigroup of X denoted $(P_n)_{n\in\mathbb{Z}_+}$ and defined as

$$P_n f(x) = \mathbb{E}_x\left(f(X_n)\mathbf{1}_{n<\tau_\partial}\right),\ \forall n\in\mathbb{Z}_+,$$

for all bounded or nonnegative measurable function f on E and all $x\in E$. We recall the notations

$$\mu P_n f = \mathbb{E}_\mu\left(f(X_n)\mathbf{1}_{n<\tau_\partial}\right) = \int_E P_n f(x)\,\mu(dx),$$

for all probability measures μ on E and all bounded measurable f. We make the following assumption[1].

Assumption E. There exist a positive integer n_1, positive real constants $\theta_1, \theta_2, c_1, c_2, c_3$, two functions $\varphi_1, \varphi_2 : E \to \mathbb{R}_+$ and a probability measure ν on a measurable subset $K \subset E$ such that

E1. *(Local A1-A2).* $\forall x \in K$,

$$\mathbb{P}_x(X_{n_1} \in \cdot) \geq c_1 \nu(\cdot \cap K) \quad \text{and} \quad \sup_{n \in \mathbb{Z}_+} \frac{\sup_{y \in K} \mathbb{P}_y(n < \tau_\partial)}{\inf_{y \in K} \mathbb{P}_y(n < \tau_\partial)} \leq c_2.$$

E2. *(Global Lyapunov criterion).* We have $\theta_1 < \theta_2 \leq 1$ and

$$\begin{aligned}
&\inf_{x\in E} \varphi_1(x) \geq 1, \ \sup_{x \in K} \varphi_1(x) < \infty \\
&\inf_{x\in K} \varphi_2(x) > 0, \ \sup_{x \in E} \varphi_2(x) \leq 1, \\
&P_1 \varphi_1(x) \leq \theta_1 \varphi_1(x) + c_3 \mathbf{1}_K(x), \ \forall x \in E \\
&P_1 \varphi_2(x) \geq \theta_2 \varphi_2(x), \ \forall x \in E.
\end{aligned}$$

E3. *(Aperiodicity).* For all $x \in K$, there exists $n_4(x)$ such that, for all $n \geq n_4(x)$,

$$\mathbb{P}_x(X_n \in K) > 0.$$

Remark 2.1. The construction of Lyapunov functions such as φ_1 is rather classical. On the contrary, finding functions such as φ_2 may seem at first more challenging. In fact, they are many ways to construct such a function, as illustrated by the numerous applications of the original paper. For instance, if Assumption E1 holds true and if there exists $\theta_2 \in (0,1)$ such that $\theta_2^{-n} \mathbb{P}_x(X_n \in K) \to +\infty$ when $n \to +\infty$, then, for any n_0 large enough, the function $\varphi_2(x) := \sum_{k=0}^{n_0} \theta_2^{-k} \mathbb{P}_x(X_k \in K)$ satisfies (up to renormalisation) condition E2.

In the rest of this section, we state our main results. We start with the exponential contraction in total variation of the conditional marginal distributions of the process given non-absorption (refinements and extensions of the following results are detailed in the original article [58]).

In the following result, $\left(\mathcal{M}(\varphi_1), \|\cdot\|_{\mathcal{M}(\varphi_1)}\right)$ is the complete space defined p. xv.

Theorem 2.1. *Assume that Condition E holds true. Then there exist a constant $C > 0$, a constant $\alpha \in (0,1)$, and a probability measure ν_{QSD} on E such that*

$$\left\| \frac{\mu P_n}{\mu P_n \mathbf{1}_E} - \nu_{QSD} \right\|_{\mathcal{M}(\varphi_1)} \leq C \alpha^n \frac{\mu(\varphi_1)}{\mu(\varphi_2)}, \tag{2.1}$$

for all probability measures μ on E such that $\mu(\varphi_1) < \infty$ and $\mu(\varphi_2) > 0$. Moreover, ν_{QSD} is the unique quasi-stationary distribution of X such that $\nu_{QSD}(\varphi_1) < \infty$ and $\nu_{QSD}(\varphi_2) > 0$. In addition $\nu_{QSD}(K) > 0$.

[1] The presentation is slightly different, although equivalent, to the one in the original article [58]. Our aim here is to make appear clearly the fact that Condition A from Chapter 1 is assumed to hold locally.

We define

$$E' := \{x \in E : \exists k \geq 0 \text{ s.t. } P_k \mathbf{1}_K(x) > 0\} = \left\{x \in E : \exists k \geq 0 \text{ s.t. } P_k \varphi_2(x) > 0\right\}.$$

We refer the reader to the original article for the equality between the two sets above.

Corollary 2.2. *Assume that Condition E holds true. Then the domain of attraction of ν_{QSD} contains all the probability measures μ on E such that $\mu(E') > 0$ and $\mu(\varphi_1^{1/p}) < \infty$ for some $p < {}^{\log\theta_1}/_{\log\theta_2}$.*

In particular, if φ_1 is bounded and $E' = E$, there exists a unique quasi-stationary distribution which attracts all the initial distributions.

We focus now on the asymptotic behaviour of the absorption probabilities and on the existence of an eigenfunction for P_1 associated to the eigenvalue θ_0, where $\theta_0 \in (0,1]$ is such that

$$\mathbb{P}_{\nu_{QSD}}(n < \tau_\partial) = \theta_0^n, \quad \forall n \in \mathbb{Z}_+.$$

We recall that the existence of θ_0 is a classical general result for quasi-stationary distributions [80, 203, 252]. In the following result, $(L^\infty(\psi), \|\cdot\|_{L^\infty(\psi)})$ is the Banach space defined p. xvi.

Proposition 2.3. *Assume that Condition E holds true. Then, there exists a function $\eta : E \to \mathbb{R}_+$ such that*

$$\eta(x) = \lim_{n\to+\infty} \frac{\mathbb{P}_x(n < \tau_\partial)}{\mathbb{P}_{\nu_{QSD}}(n < \tau_\partial)} = \lim_{n\to+\infty} \theta_0^{-n} \mathbb{P}_x(n < \tau_\partial), \quad \forall x \in E, \tag{2.2}$$

where the convergence is geometric in $L^\infty(\varphi_1^{1/p})$ for all $p \in [1, \log\theta_1/\log\theta_0)$. In addition, $\inf_{y\in K} \eta(y) > 0$, $E' = \{x \in E : \eta(x) > 0\}$, $\nu_{QSD}(\eta) = 1$,

$$P_1 \eta = \theta_0 \eta \quad \text{and} \quad \theta_0 \geq \theta_2 > \theta_1.$$

Remark 2.2. The last result implies that, when η is bounded, one can actually take $\varphi_2 = \eta/\|\eta\|_\infty$ in Condition (E2). This property can be adapted to the case where η is not bounded, using for instance the approach of Chapter 7.

Remark 2.3. Similarly to Chapter 1, Assumption E also entails the existence of a spectral gap between θ_0 and the next eigenvalue. It also implies that $\eta \in L^\infty\left(\varphi_1^{\log\theta_0/\log\theta_1}\right)$ (see Corollary 2.6 in [58]).

2.1.2 Continuous time models

We consider in this section an absorbed Markov process $(X_t)_{t\in\mathbb{R}_+}$ in the continuous time setting.

Assumption F. There exist positive real constants $\gamma_1, \gamma_2, c_1, c_2, c_3, t_1$ and t_2, a measurable function $\psi_1 : E \to [1, +\infty)$, and a probability measure ν on a measurable subset $L \subset E$ such that

F0. *(A strong Markov property).* Defining

$$\tau_L := \inf\{t \in \mathbb{R}_+ : X_t \in L\}, \tag{2.3}$$

assume that for all $x \in E$, $X_{\tau_L} \in L$, $\mathbb{P}_x$-almost surely on the event $\{\tau_L < \infty\}$, and, for all $t > 0$ and all $f \in \mathscr{B}_b(E \cup \{\partial\})$,

$$\mathbb{E}_x\left[f(X_t)\mathbf{1}_{\tau_L \leq t < \tau_\partial}\right] = \mathbb{E}_x\left[\mathbf{1}_{\tau_L \leq t\wedge\tau_\partial} \mathbb{E}_{X_{\tau_L}}\left[f(X_{t-u})\mathbf{1}_{t-u<\tau_\partial}\right]\Big|_{u=\tau_L}\right].$$

F1. *(Local A1 and A2).* $\forall x \in L$,

$$\mathbb{P}_x(X_{t_1} \in \cdot) \geq c_1 \nu(\cdot \cap L) \text{ and } \sup_{t \in \mathbb{R}_+} \frac{\sup_{y \in L} \mathbb{P}_y(t < \tau_\partial)}{\inf_{y \in L} \mathbb{P}_y(t < \tau_\partial)} \leq c_2.$$

F2. *(Global Lyapunov criterion).* We have $\gamma_1 < \gamma_2$ and

$$\begin{aligned}
&\mathbb{E}_x(\psi_1(X_{t_2}) \mathbf{1}_{t_2 < \tau_L \wedge \tau_\partial}) \leq \gamma_1^{t_2} \psi_1(x), \ \forall x \in E \\
&\mathbb{E}_x(\psi_1(X_t) \mathbf{1}_{t < \tau_\partial}) \leq c_3, \ \forall x \in L, \ \forall t \in [0, t_2], \\
&\gamma_2^{-t} \mathbb{P}_x(X_t \in L) \xrightarrow[t \to +\infty]{} +\infty, \ \forall x \in L.
\end{aligned}$$

Theorem 2.4. *Under Assumption F, $(X_t)_{t \in I}$ admits a quasi-stationary distribution ν_{QSD}, which is the unique one satisfying $\nu_{QSD}(\psi_1) < \infty$ and $\mathbb{P}_{\nu_{QSD}}(X_t \in L) > 0$ for some $t \in I$. Moreover, there exist constants $\alpha \in (0,1)$ and $C > 0$ such that, for all probability measures μ on E satisfying $\mu(\psi_1) < \infty$ and $\mu(\psi_2) > 0$,*

$$\left\| \mathbb{P}_\mu(X_t \in \cdot \mid t < \tau_\partial) - \nu_{QSD} \right\|_{TV} \leq C \alpha^t \frac{\mu(\psi_1)}{\mu(\psi_2)}, \ \forall t \in I, \tag{2.4}$$

where $\psi_2(x) = \sum_{k=0}^{n_0} \gamma_2^{-k t_2} \mathbb{P}_x(X_{k t_2} \in L)$ for $n_0 \geq 1$ and $t_2 \in \mathbb{R}_+$ large enough. In addition, there exists a constant $\lambda_0 \geq 0$ such that $\lambda_0 \leq \log(1/\gamma_2) < \log(1/\gamma_1)$ and $\mathbb{P}_{\nu_{QSD}}(t < \tau_\partial) = e^{-\lambda_0 t}$ for all $t \geq 0$, and there exists a function η such that

$$\eta(x) = \lim_{t \to +\infty} e^{\lambda_0 t} \mathbb{P}_x(t < \tau_\partial), \quad \forall x \in E, \tag{2.5}$$

where the convergence is exponential in $L^\infty(\psi_1^{1/p})$ for all $p \in [1, \log(1/\gamma_1)/\lambda_0)$, and $P_t \eta(x) = e^{-\lambda_0 t} \eta(x)$ for all $x \in E$ and $t \in \mathbb{R}_+$.

In particular, if η is bounded and setting $\eta(\partial) = 0$, then the function η defined on $E \cup \{\partial\}$ belongs to the domain of the infinitesimal generator $\mathscr{L}$ of X and $\mathscr{L}\eta = -\lambda_0 \eta$.

Remark 2.4. The main point of the proof is to check that Assumption F entails Assumption E for the sub-Markovian semigroup $(P_n)_{n \geq 0}$ of the absorbed Markov process $(X_{n t_2})_{n \in \mathbb{Z}_+}$, with the functions $\varphi_1 = \psi_1$ and $\varphi_2 = \frac{\gamma_2^{-t_2} - 1}{\gamma_2^{-(n_0+1)t_2} - 1} \psi_2$, any $\theta_1 \in (\gamma_1^{t_2}, \gamma_2^{t_2})$, $\theta_2 = \gamma_2^{t_2}$ and the set

$$K = \left\{ y \in E, \ \mathbb{P}_y(\tau_L \leq t_2)/\psi_1(y) \geq (\theta_1 - \gamma_1^{t_2})/c_2 \right\} \supset L.$$

Remark 2.5. The first two lines of F2 can be deduced by classical Foster-Lyapunov inequalities (cf. [207]). Indeed, denoting by $\mathscr{L}$ the infinitesimal generator of the process X, if

$$\mathscr{L}\psi_1(x) \leq -\lambda_1 \psi_1(x) + C \mathbf{1}_L(x), \quad \forall x \in E, \tag{2.6}$$

then (formally, assuming one can apply Dynkin's formula) $\mathbb{E}_x[\mathbf{1}_{1 \leq \tau_L \wedge \tau_\partial} \psi_1(X_1)] \leq e^{-\gamma_1} \psi_1(x)$ and $\mathbb{E}_x[\psi_1(X_t) \mathbf{1}_{t < \tau_\partial}] \leq e^{Ct} \psi_1(x)$. However, a function ψ_1 satisfying (2.6) does not necessarily belong to the domain of the infinitesimal generator $\mathscr{L}$, so one needs to extend the notion of infinitesimal generator as in [207, 59].

An other approach has been considered in [194], with conditions involving Lyapunov functions for the family of time-inhomogeneous processes, defined as the process X conditioned not to be absorbed before time t, where t runs over $\mathbb{R}_+$. We also refer the reader to [253], where the authors provide criteria based on hitting time controls. Such controls are of course related to the existence of Lyapunov functions, as explained in [58, Lemma 3.6].

2.2 Birth and death Markov chains

Let $(\lambda_x)_{x\in\mathbb{N}}$ and $(\mu_x)_{x\in\mathbb{N}}$ be families of positive numbers such that $\mu_x + \lambda_x = 1$ for all $x \in \mathbb{N}$. We consider the birth and death chain on $E \cup \{\partial\}$, $E = \mathbb{N}$ and $\partial = 0$, with the following transition probabilities

$$\mathbb{P}_x(X_1 = y) = \begin{cases} \lambda_x & \text{if } y = x+1, \\ \mu_x & \text{if } y = x-1, \\ 0 & \text{otherwise.} \end{cases}$$

These models are known as birth and death chains, and we refer the reader to [120, 24, 81, 99, 9] and references therein for earlier studies of their quasi-stationary behaviour.

We assume that the process X is aperiodic, that $\inf_{x\geq 1} \mu_x > 0$, and that $\lambda_x \to 0$ when $x \to +\infty$. We define

$$\varphi_1(x) = \mathbf{1}_{x\geq 1} e^{ax} \quad \text{with } a > 0 \text{ such that } e^{-a} \leq \frac{1-\mu_1}{4} \quad \text{and} \quad \varphi_2(x) = \mathbf{1}_{x\geq 1}, \quad \forall x \geq 0,$$

and show that Assumption E is satisfied in this case.

Since X is assumed to be aperiodic, there exists $x_0 \geq 1$ such that $\lambda_{x_0} + \mu_{x_0} < 1$ (otherwise it would be 2-periodic) and this entails Condition E3. Since the birth and death coefficients $(\lambda_x)_{x\in\mathbb{N}}$ and $(\mu_x)_{x\in\mathbb{N}}$ are positive, for any finite set $K = \{1, \dots, x_K\}$, there exists n_K such that $\inf_{x,y\in K} \mathbb{P}_x(X_{n_K} = y) > 0$. Similarly to the proof of (A1-A2) in Section 1.4 of Chapter 1, this entails Condition E1.

It only remains to check that E2 holds true, for some finite set K as above. Set $\theta_2 = \lambda_1$ and $\theta_1 = \theta_2/2$. For all $x \geq 1$, we have

$$\mathbb{E}_x(\varphi_2(X_1)) = \lambda_1 = \theta_2 \varphi_2(x).$$

and

$$\begin{aligned} \mathbb{E}_x(\varphi_1(X_1)) &= \lambda_x e^{a(x+1)} + \mu_x e^{a(x-1)} = \varphi_1(x) \left(\lambda_x e^a + \mu_x e^{-a} + (1 - \lambda_x - \mu_x)\right) \\ &\leq \varphi_1(x) \left(\lambda_x e^a + \theta_1/2\right) \leq \theta_1 \varphi_1(x) + C \mathbf{1}_{x\leq n}, \end{aligned}$$

where C and n are chosen large enough. This concludes the proof.

2.3 Theory of R-positive matrices

We consider a Markov chain $(X_n)_{n\in\mathbb{Z}_+}$ in a countable state space $E \cup \{\partial\}$ absorbed at $\partial \notin E$ and with irreducible transition probabilities in E, i.e. such that for all $x, y \in E$, there exists $n = n(x, y) \geq 1$ such that $\mathbb{P}_x(X_n = y) > 0$. In this case, one of the most general criterion for existence and convergence to a quasi-stationary distribution is provided in [119]. In this paper, the authors obtain a convergence result similar to the one of Theorem 2.1 restricted to Dirac initial distributions, and the pointwise convergence to η (as defined in Proposition 2.3), using the powerful theory of R-positive matrices (see [7, 6, 118] for recent applications). In this section, we show how our criterion allows to recover these results, providing in addition the characterisation of a non-trivial subset of the domain of attraction and a stronger convergence to η.

We assume that the absorption time τ_∂ is almost surely finite. Without loss of generality, we will assume that the process is aperiodic, meaning that $\mathbb{P}_x(X_n = y) > 0$ for all $x, y \in E$ provided n is large enough; the extension to general periodic processes is routine, as observed in [119].

Let us now recall the statement of [119, Theorem 1]. Here R is defined as

$$1/R = \lim_{n\to\infty} \left[\mathbb{P}_x(X_n = y)\right]^{1/n}, \tag{2.7}$$

for some and hence any value of $x, y \in E$ (see e.g. [254]).

Theorem 2.5 (Theorem 1, [119])**.** *Assume that the Markov chain satisfies the following conditions:*

(a) *there exist a nonempty set $U_1 \subset E$ and two positive constants ε_0, C_1 such that, for all $x \in U_1$ and all $n \geq 0$,*

$$\mathbb{P}_x(\tau_\partial > n,\ \text{but } X_\ell \notin U_1 \text{ for all } 1 \leq \ell \leq n) \leq C_1(R+\varepsilon_0)^{-n},$$

(b) *there exist a state $x_0 \in U_1$ and a positive constant C_2 such that, for all $x \in U_1$ and $n \geq 0$,*

$$\mathbb{P}_x(n < \tau_\partial) \leq C_2 \mathbb{P}_{x_0}(n < \tau_\partial),$$

(c) *there exist a finite set $U_2 \subset E$ and constants $0 \leq n_0 < \infty$, $C_3 > 0$, such that for all $x \in U_1$,*

$$\mathbb{P}_x(X_n \in U_2 \text{ for some } n \leq n_0) \geq C_3.$$

Then X is R-positive-recurrent and there exists a probability measure ν_{QSD} on E such that

$$\lim_{n\to+\infty} \mathbb{P}_x(X_n = y \mid n < \tau_\partial) = \nu_{QSD}(y), \quad \forall x, y \in E$$

and a positive function η on E such that

$$\lim_{n\to+\infty} R^n \mathbb{P}_x(n < \tau_\partial) = \eta(x), \quad \forall x \in E.$$

The main result of this section is the following

Proposition 2.6. *The assumptions of [119, Theorem 1] imply Assumption E.*

In the settings of this section, Assumption E is actually equivalent to the conditions of [119, Theorem 1]. Besides the additional properties provided in Chapter 2, one of our main contribution in this particular setting is to provide a different, sometimes more tractable criterion, through the use of Lyapunov functions. This is illustrated in the next subsection, with an application to population processes, extending to the multi-dimensional case some models studied in [132]. The application of [119] is not "impractical for such models of biological population extinction" as claimed in [132, p. 262], but it would be a little bit more involved.

2.4 Application to the extinction of biological populations dominated by Galton-Watson processes

A Markov process $(Z_n)_{n\in\mathbb{Z}_+}$ evolving in $\mathbb{Z}_+^d = E \cup \{\partial\}$ absorbed at $\partial = 0$ is called a Galton-Watson process with d types if, for all $n \geq 0$ and all $i \in \{1, \ldots, d\}$,

$$Z_{n+1}^i = \sum_{k=1}^{d} \sum_{\ell=1}^{Z_n^k} \zeta_{k,i}^{(n,\ell)}, \tag{2.8}$$

where the random variables $(\zeta_{k,1}^{(n,\ell)},\dots,\zeta_{k,d}^{(n,\ell)})_{n,\ell,k}$ in $\mathbb{Z}_+$ are assumed independent and such that, for all $k\in\{1,\dots,d\}$, $(\zeta_{k,1}^{(n,\ell)},\dots,\zeta_{k,d}^{(n,\ell)})_{n,\ell}$ is an i.i.d. family. We define the matrix $M=(M_{k,i})_{1\leq k,i\leq d}$ of mean offspring as

$$M_{k,i}=\mathbb{E}(\zeta_{k,i}^{(n,\ell)}),\quad \forall k,i\in\{1,\dots,d\},$$

and assume that $M_{k,i}<+\infty$ and that there exists $n\geq 1$ such that $[M^n]_{k,i}>0$ for all $k,i\in\{1,\dots,d\}$.

Using the classical formalism of [137], we consider a positive right eigenvector v of the matrix M of mean offspring and we denote by $\rho(M)$ its spectral radius. The sub-critical case corresponds to $\rho(M)<1$. It is well-known [159] (see also [144, 12]) that this implies the existence of a quasi-stationary distribution whose domain of attraction contains all Dirac measures (a so-called *Yaglom limit* or *minimal quasi-stationary distribution*). The authors also prove that $\nu_{QSD}(|\cdot|)<\infty$ if and only if $\mathbb{E}[|Z_1|\log(|Z_1|)\mid Z_0=(1,\dots,1)]<\infty$. While the following result makes the stronger assumption that $\mathbb{E}[|Z_1|^{q_0}\mid Z_0=(1,\dots,1)]<\infty$ for some $q_0>1$, we obtain a stronger form of convergence (in total variation norm with exponential speed), a non-trivial subset of the domain of attraction of the minimal quasi-stationary distribution and stronger moment properties for this quasi-stationary distribution.

Proposition 2.7. *If $(Z_n)_{n\in\mathbb{Z}_+}$ is a d-type irreducible, aperiodic sub-critical Galton-Watson process, and if, for some $q_0>1$,*

$$\mathbb{E}[|Z_1|^{q_0}\mid Z_0=(1,\dots,1)]<\infty,$$

then Condition (E) holds true with $\varphi_1(z)=|z|^q$ for any $q\in(1,q_0]$. In particular, the domain of attraction of ν_{QSD} contains all the probability measures such that $\mu(|\cdot|^q)<\infty$ for some $q>1$.

We focus now on population processes dominated by population-dependent Galton-Watson processes. More precisely, we consider an aperiodic and irreducible Markov population process $(Z_n)_{n\in\mathbb{N}}$ on $\mathbb{Z}_+^d=E\cup\{\partial\}$ absorbed at $\partial=0$ such that, for all $n\geq 0$,

$$\|Z_{n+1}\|\leq\sum_{i=1}^{|Z_n|}\xi_{i,n}^{(Z_n)},\tag{2.9}$$

where $\|\cdot\|$ is a norm on $\mathbb{R}^d$ and $|z|=z_1+\ldots+z_d$ for all $z\in\mathbb{Z}_+^d$ and, for all $n\geq 0$, the nonnegative random variables $(\xi_{i,n}^{(Z_n)},1\leq i\leq|Z_n|)$ are assumed independent (but not necessarily identically distributed) given Z_n.

We assume that

$$\mathbb{E}\left(\sum_{i=1}^{|z|}\xi_{i,n}^{(z)}\right)\leq m\|z\|,\quad \forall z\in\mathbb{Z}_+^d \text{ such that } |z|\geq n_0,\tag{2.10}$$

for some $m<1$ and $n_0\in\mathbb{N}$. This means that the population size has a tendency to decrease (in mean) when it is large. This also implies that $\tau_\partial<\infty$ a.s.

In the following theorem, $R>0$ is the limiting value defined in (2.7).

Proposition 2.8. *Assume that $(Z_n)_{n\in\mathbb{Z}_+}$ is aperiodic irreducible, that it satisfies the assumptions* (2.9) *and* (2.10) *and that, for some $q_0>\frac{\log R}{\log(1/m)}\vee 1$,*

$$\sup_{n\geq 0,\ z\in\mathbb{Z}_+^d,\ 1\leq i\leq|z|}\mathbb{E}[(\xi_{i,n}^{(z)})^{q_0}]<\infty,$$

Then Condition E holds true with $\varphi_1(x)=|x|^q$, for all $q\in\left(\frac{\log R}{\log(1/m)}\vee 1,q_0\right]$.

Remark 2.6. This result easily applies if $\sup_{n\geq 0,\ z\in\mathbb{Z}_+^d,\ 1\leq i\leq |z|}\mathbb{E}[(\xi_{i,n}^{(z)})^q]<\infty$ for all $q>0$. In other cases, one needs a upper bound for $R>0$ in order to check the validity of the assumptions of Proposition 2.8. For instance, one may use the fact that $R\leq 1/\sup_{z\in\mathbb{Z}_+^d}\mathbb{P}_z(Z_1=z)$.

The above theorem applies for instance when Z is obtained from a Galton-Watson multi-type process with an additional population-dependent death rate. Typically, one can assume that additional death events may affect a fraction of the population, modelling global death events. Note that, in this case and contrary to the Galton-Watson case, the independence between the progeny of individuals breaks down and the classical approach based on generating functions is rendered helpless.

Another situation covered by the above result is the case where the domain of absorption of Z is a larger set than 0, for example the process may be absorbed when it reaches one edge of $\mathbb{Z}_+^d$ (i.e. when one type disappears).

Another typical application of Proposition 2.8 is the case of population-dependent Galton-Watson processes, i.e. of processes such that, given Z_n, Z_{n+1} is the sum of $|Z_n|$ independent random variables whose law may depend on Z_n. In this situation, Proposition 2.8 and its consequences stated in Chapter 2 generalise the results of [132] to the multi-type models and provides finer results on the domain of attraction of the minimal quasi-stationary distribution. The reducible cases considered in [132] can also be recovered using the approach of Section 2.6. Of course, the above specifications may be combined.

2.5 Perturbed dynamical systems

We consider the following perturbed dynamical system

$$X_{n+1}=f(X_n)+\xi_n,$$

where $f:\mathbb{R}^d\to\mathbb{R}^d$ is a measurable function and $(\xi_n)_{n\in\mathbb{N}}$ is an i.i.d. sequence in $\mathbb{R}^d$. The quasi-stationary behaviour of such processes has been studied under different constraints in [117, 31, 23, 147]. We assume that the process evolves in a measurable set D of $\mathbb{R}^d$ with positive Lebesgue measure, and that it is immediately sent to $\partial\not\in\mathbb{R}^d$ as soon as $X_n\not\in D$. We consider here the case where the law of the random variables ξ_n have support $\mathbb{R}^d$ and, more precisely, admit a positive bounded density with respect to the Lebesgue measure[2]. This includes the particular case of dynamical systems perturbed by a Gaussian noise, which is considered in Example 2.1 below. In this setting, the perturbed dynamical system $X_{n+1}=f(X_n)+\xi_n$ with $(\xi_i)_{i\in\mathbb{Z}_+}$ i.i.d. Gaussian, absorbed when it leaves a given measurable set D of $\mathbb{R}^d$ with positive Lebesgue measure, admits a quasi-stationary distribution as soon as $|x|-|f(x)|\to+\infty$ when $|x|\to+\infty$.

Proposition 2.9. *Assume that f is locally bounded, that the law of ξ_n has a bounded density $g(x)$ with respect to Lebesgue's measure, that*

$$\inf_{|x|\leq R} g(x)>0,\quad \forall R>0,$$

[2] In the original article [58], we also consider situations where the random variables ξ_n does not admits a bounded density with respect to Lebesgue's measure. The same arguments would also work, at the expense of additional technical difficulties, if $X_{n+1}=f(X_n)+\xi_n(X_n)$, where the sequence of random maps $(x\mapsto\xi_n(x))_{n\geq 0}$ are i.i.d.

and that there exists a locally bounded function $\varphi : \mathbb{R}^d \to [1, +\infty)$ *such that* $x \mapsto \mathbb{E}(\varphi(x+\xi_1))$ *is locally bounded on* $\mathbb{R}^d$ *and*

$$\limsup_{|x|\to+\infty,\, x\in D} \frac{\mathbb{E}(\varphi(f(x)+\xi_1))}{\varphi(x)} = 0. \tag{2.11}$$

Then Condition E is satisfied with $\varphi_1 = \varphi$ *and* φ_2 *positive on* D.

Let us illustrate this proposition with three examples.

Example 2.1. If there exists $\alpha > 0$ such that $\mathbb{E}e^{\alpha|\xi_1|} < +\infty$ and if $|x| - |f(x)| \to +\infty$ when $|x| \to +\infty$, then Proposition 2.9 applies. Indeed, choosing $\varphi(x) = \exp(\alpha|x|)$, we have

$$\frac{\mathbb{E}\varphi(|f(x)+\xi_1|)}{\varphi(x)} \le e^{\alpha(|f(x)|-|x|)}\mathbb{E}e^{\alpha|\xi_1|} \xrightarrow[|x|\to+\infty]{} 0.$$

For instance, this covers the case of Gaussian perturbations.

Example 2.2. If there exists $p > 0$ such that $\mathbb{E}(\xi_1^p) < +\infty$ and if $|f(x)| = o(|x|)$ when $|x| \to +\infty$, then Proposition 2.9 applies. Indeed, choosing $\varphi(x) = (1+|x|)^p$, we have

$$\frac{\mathbb{E}\varphi(|f(x)+\xi_1|)}{\varphi(x)} \le \frac{(1+|f(x)|)^p}{(1+|x|)^p}\mathbb{E}[(1+|\xi_1|)^p] \xrightarrow[|x|\to+\infty]{} 0.$$

Example 2.3. If $\mathbb{E}\log(1+|\xi_1|) < \infty$ and $|f(x)| \le C|x|^{\varepsilon(x)}$ for some $C > 0$ and some $\varepsilon(x) \to 0$ when $|x| \to +\infty$, then Proposition 2.9 applies. Indeed, choosing $\varphi(x) = \log(e+|x|)$, we have

$$\frac{\mathbb{E}\varphi(|f(x)+\xi_1|)}{\varphi(x)} \le \frac{\log(e+C)+\varepsilon(x)\log(e+|x|)}{\log(1+|x|)} + \frac{\mathbb{E}\log(1+|\xi_1|)}{\log(e+|x|)}.$$

The condition on f is true for example if $|f(x)| \le C\exp\sqrt{\log(1+|x|)}$ for some constant C.

2.6 Reducible models

The study of quasi-stationary behaviour of models with multiple communication classes has been conducted in [216, 132, 55, 54, 252, 28]. Our criteria provide new practical tools to tackle this problem.

In Subsection 2.6.1, we consider a general setting with three successive sets. In Subsection 2.6.2, we consider a birth and death process with a countable infinity of communication classes.

2.6.1 Three successive sets

Consider a discrete time Markov process $(X_n)_{n\in\mathbb{Z}_+}$ evolving in a measurable set $E\cup\{\partial\}$ with absorption at $\partial \notin E$. We assume that the transition probabilities of X satisfy the structure displayed in Figure 2.1 : one can find a partition $\{D_1, D_2, D_3\}$ of E such that the process starting from D_1 can access $D_1\cup D_2\cup D_3\cup\{\partial\}$, the process starting from D_2 can only access $D_2\cup D_3\cup\{\partial\}$, and the process starting from D_3 can only access $D_3\cup\{\partial\}$. More formally, we assume that $\mathbb{P}_x(T_{D_3}\wedge\tau_\partial < T_{D_1}) = 1$ for all $x\in D_2$ and that $\mathbb{P}_x(\tau_\partial < T_{D_1\cup D_2}) = 1$ for all $x \in D_3$, where we recall that, for any measurable set $A\subset E$, $T_A = \inf\{n\in\mathbb{Z}_+,\ X_n\in A\}$.

Our aim is to provide sufficient conditions ensuring that X satisfies Assumption E. In order to do so, we assume that Assumption E is satisfied by the process X before exiting D_2. This corresponds to the following assumption.

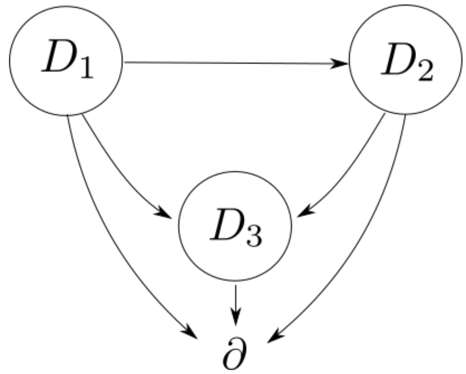

Figure 2.1: Transition graph displaying the relation between the sets D_1, D_2, D_3 and ∂.

Assumption H1. The absorbed Markov process Y evolving in $D_2 \cup \{\partial\}$, defined by

$$Y_n = \begin{cases} X_n & \text{if } n < T_{D_1 \cup D_3 \cup \{d\}}, \\ \partial & \text{if } n \geq T_{D_1 \cup D_3 \cup \{d\}}, \end{cases}$$

satisfies Assumption E. In what follows, we denote the objects related to Y with a superscript Y, for instance, the constants of Assumption E for Y are denoted by $\theta_1^Y > 0$, $\theta_2^Y > 0$.

We also assume that the exit times from D_1 and D_3 for the process X admit exponential moments of sufficiently high order, as stated by the following assumption.

Assumption H2. There exists a positive constant $\gamma < \theta_0^Y$ such that, for all $x \in D_1$,

$$\mathbb{E}_x\left(\gamma^{-T_{D_2}} \varphi_1^Y\left(X_{T_{D_2}}\right) \mathbf{1}_{T_{D_2} < T_{D_3} \wedge \tau_\partial}\right) < +\infty, \quad \mathbb{E}_x\left(\gamma^{-T_{D_3} \wedge \tau_\partial} \mathbf{1}_{T_{D_3} \wedge \tau_\partial < T_{D_2}}\right) < +\infty,$$

and such that

$$\sup_{x \in D_3} \mathbb{E}_x\left(\gamma^{-\tau_\partial}\right) < +\infty.$$

We are now able to state the main result of this section.

Proposition 2.10. *Under Assumptions H1 and H2, the process X satisfies Assumption E with $K = K^Y$,*

$$\varphi_1(x) = \mathbb{E}_x\left(\gamma^{-T_K \wedge \tau_\partial}\right) \quad \textit{and} \quad \varphi_2(x) \geq c\mathbf{1}_{x \in K}, \ \forall x \in E.$$

In particular, it admits a unique quasi-stationary distribution ν_{QSD} such that $\nu_{QSD}(\varphi_1) < \infty$ and $\nu_{QSD}(\varphi_2) > 0$. Moreover, there exist two constants $C > 0$ and $\alpha \in (0,1)$ such that, for all probability measures μ on E such that $\mu(\varphi_1) < \infty$ and $\mu(\varphi_2) > 0$,

$$\left\| \mathbb{P}_\mu(X_n \in \cdot \mid n < \tau_\partial) - \nu_{QSD} \right\|_{TV} \leq C\alpha^n \frac{\mu(\varphi_1)}{\mu(\varphi_2)}.$$

Finally, $\theta_0 = \theta_0^Y$, $\nu_{QSD}(D_1) = 0$ and the function η of Proposition 2.3 vanishes on D_3.

In particular, one deduces from the last property that $E' \subset D_1 \cup D_2$, where we recall that $E' = \{x \in E : \exists n \in \mathbb{N},\ P_n \mathbf{1}_K(x) > 0\}$.

Remark 2.7. 1. The fact that there are three different sets D_1, D_2 and D_3 in the decomposition of E is not restrictive on the number of communication classes. Indeed, the three sets can contain several communication classes.

2. A similar result can be obtained for continuous time processes, based on Assumption F instead of E, with the additional technical assumption that the exit times of D_1 and D_2 are stopping times.

3. Beside the exponential moment assumption, there is no additional requirement on the behaviour of the Markov process in D_1 and D_3. In these sets, the process might be periodic or deterministic for instance.

4. The quasi-stationary distribution of this process may not be unique, for instance if the process restricted to D_3 also admits a quasi-stationary distribution.

2.6.2 Countably many communication classes

We consider now a particular case of a continuous time càdlàg Markov process $(X_t)_{t\in\mathbb{R}_+}$ with a countable infinity of communication classes and we show that the process admits a quasi-stationary distribution.

More precisely, we assume that X evolves in the state space $\mathbb{N}\times\mathbb{Z}_+$ (the first component is the index of the communication class and the second is the position of the process in this communication class) and, denoting $N_t\in\mathbb{N}$ and $Y_t\in\mathbb{Z}_+$ the two components of X_t for all $t\in\mathbb{R}_+$, that there exist three positive functions $b,d,f:\mathbb{N}\to(0,+\infty)$ such that

- N is a Poisson process with intensity 1,
- Y is a process such that, at time t,

$$Y \text{ jumps from } Y_t \text{ to } y\in\mathbb{Z}_+ \text{ with rate } \begin{cases} f(N_t)\,b(Y_t) & \text{if } y=Y_t+1 \text{ and } Y_t\geq 1,\\ f(N_t)\,d(Y_t) & \text{if } y=Y_t-1 \text{ and } Y_t\geq 1,\\ 0 & \text{otherwise.}\end{cases}$$

The set $\mathbb{N}\times\{0\}$ is absorbing for X and we are interested in the quasi-stationary behaviour of X conditioned to not hit this set. Note that, in this case, each set $\{n\}\times\mathbb{N}$ is a communication class.

This process can be used to model the evolution of the vitality of an individual (for example a bacterium) whose metabolic efficiency (for example its ability to consume resources) changes with time, due to ageing [238]. Here Y is the vitality of the individual, who dies when its vitality hits 0, and $f(N)$ is the metabolic rate of the individual.

This process can also be used to model the accumulation of deleterious mutations in a population under the assumption that mutations do not overlap, i.e. that when a mutant succeeds to invade the population (either because they are advantaged or due to genetic drift for deleterious mutations), other types of mutants disappear rapidly. Here Y represents the size of the population and N the number of mutations (see e.g. [84, 82]).

In both cases, it is relevant to assume that f is decreasing on $\{1,2,\dots,n_0\}$ and increasing to $+\infty$ on $\{n_0,n_0+1,\dots\}$, which we do from now on. We also assume that $(d(y)-b(y))/y\to+\infty$ when $y\to+\infty$ or that there exists $\delta>1$ such that $d(y)-\delta\,b(y)\to+\infty$.

Proposition 2.11. *Under the above assumptions, the process X satisfies Assumption F and admits a quasi-stationary distribution ν_{QSD} whose domain of attraction contains all Dirac measures $\delta_{n,y}$, with $n \leq n_0$ and $y \in \mathbb{N}$.*

Part II

Application of the criteria of Part I to classical models

Chapter 3

Birth and death processes

In this chapter, we focus on the application of the results of Chapters 1 and 2 to birth and death processes. In Section 3.1, we recall the classical result on quasi-stationary distributions for birth and death processes. In Section 3.2, we consider the case of one-dimensional birth and death processes with entrance boundary at infinity, and, in Section 3.3, the case of one-dimensional birth and death processes which are λ_0-positive recurrent. In Section 3.4, we consider the case of multi-dimensional birth and death processes.

The results of this chapter first appeared in the articles [201, 259, 56, 58]

3.1 Quasi-stationary distributions for one-dimensional birth and death processes

Let $(X_t)_{t\in\mathbb{R}_+}$ be a birth and death process on $\mathbb{Z}_+$ with birth rates $(b_x)_{x\in\mathbb{Z}_+}$ and death rates $(d_x)_{x\in\mathbb{Z}_+}$. We assume that $b_x > 0$ and $d_x > 0$ for any $x \in \mathbb{N}$ and $b_0 = d_0 = 0$. The stochastic process X is a $\mathbb{Z}_+$-valued pure jump process whose only absorption point is 0 and whose transition rates from any point $x \geq 1$ are given by

$$\begin{aligned} &x \to x+1 \text{ with rate } b_x,\\ &x \to x-1 \text{ with rate } d_x,\\ &x \to y \text{ with rate } 0, \text{ if } y \notin \{x-1, x+1\}. \end{aligned}$$

It is well known (see e.g. [203, Theorem 10 and Proposition 12]) that X is stable, conservative and hits 0 in finite time almost surely (for any initial distribution) if and only if

$$\sum_{n=1}^{\infty} \frac{d_1 d_2 \cdots d_n}{b_1 b_2 \cdots b_n} = +\infty, \tag{3.1}$$

which we shall assume from now on.

Such processes are extensively studied because of their conceptual simplicity and pertinence as demographic models. Concerning the study of their quasi-stationary behaviour, see for instance [164, 165, 131, 47, 246, 121, 122, 171, 170, 248, 247, 249].

From a demographic point of view, the study of the minimal quasi-stationary distribution of a birth and death process aims at answering the following question: *knowing that a population isn't extinct after a long time t, what is the probability that its size is equal to n at time t?*

For these processes, van Doorn [246] gave the following picture of the situation: a birth and death process can have no quasi-stationary distribution, one unique quasi-stationary distribution or an infinity (in fact a continuum) of quasi-stationary distributions. In order to determine whether a birth and death process has 0, one or an infinity of quasi-stationary distributions, one define inductively the sequence of polynomials $(Q_n(x))_{n\geq 0}$ for all $x\in\mathbb{R}$ by

$$\begin{aligned}
&Q_1(x)=1,\\
&b_1\,Q_2(x)=b_1+d_1-x \text{ and}\\
&b_n\,Q_{n+1}(x)=(b_n+d_n-x)\,Q_n(x)-d_{n-1}\,Q_{n-1}(x),\ \forall n\geq 2.
\end{aligned}\tag{3.2}$$

As recalled in [246, eq. (2.13)], one can uniquely define the non-negative number λ_0 satisfying

$$x\leq\lambda_0 \iff Q_n(x)>0,\ \forall n\geq 1. \tag{3.3}$$

Also, the useful quantity

$$S:=\sup_{x\geq 1}\mathbb{E}_x(T_0),$$

can be easily computed (see [5, Section 8.1]), since, for any $z\geq 1$,

$$\sup_{x\geq z}\mathbb{E}_x(T_z)=\sum_{k\geq z+1}\frac{1}{d_k\pi_k}\sum_{l\geq k}\pi_l,$$

with $\pi_k=\left(\prod_{i=1}^{k-1}b_i\right)/\left(\prod_{i=2}^{k}d_i\right)$. The following theorem answers the question of existence and uniqueness of a quasi-stationary distribution for birth and death processes.

Theorem 3.1 (van Doorn, 1991 [246]). *Let X be a birth and death process satisfying* (3.1).

1. *If $\lambda_0=0$, there is no quasi-stationary distribution.*

2. *If $S<+\infty$, then $\lambda_0>0$ and the Yaglom limit is the unique quasi-stationary distribution.*

3. *If $S=+\infty$ and $\lambda_0>0$, then there is a continuum of quasi-stationary distributions, given by the one parameter family $(\rho_a)_{0<a\leq\lambda_0}$:*

$$\rho_a(x)=\frac{\pi_x}{d_1}\,a\,Q_x(a),\ \forall x\geq 1,$$

and the minimal quasi-stationary distribution is given by ρ_{λ_0}.

Theorem 3.1 is quite remarkable since it describes completely the possible outcomes of the existence and uniqueness problem for quasi-stationary distributions. However, it only partially answers the crucial problem of finding the domain of attraction of the existing quasi-stationary distributions. The aim of this chapter is to show how the theory exposed in the two previous chapters entail new result for this class of processes. We first look at birth and death processes with entrance boundary at infinity (i.e. $S<+\infty$) in Section 3.2 and then to λ_0-positive birth and death processes in Section 3.3.

3.2 Birth and death processes with entrance boundary at infinity

Theorem 3.1 tells us that a birth and death process admits a unique quasi-stationary distribution if and only if $S < +\infty$. Building on the methods and results of [246], Zhang and Zhu [268] proved that, in this setting, the limit (1.2) holds true for all probability measures μ on $\mathbb{N}$. However, the spectral theory tools used in these publications are not well suited to study total variation convergence to the quasi-stationary distribution. The following result, first proved in [201], completes the picture offered in [246] on the quasi-limiting behaviour of birth and death processes with an entrance boundary at infinity. It shows that $S < +\infty$ if and only if Assumption A of Chapter 1 holds true.

Proposition 3.2. *A birth and death process X admits a unique quasi-stationary distribution if and only if there exist two constants $C, \gamma > 0$ and a probability measure ν_{QSD} on $\mathbb{N}$ such that, for any initial distribution μ on $\mathbb{N}$,*

$$\|\mathbb{P}_\mu(X_t \in \cdot \mid t < T_0) - \nu_{QSD}\|_{TV} \le C e^{-\gamma t}, \ \forall t \ge 0. \tag{3.4}$$

In this case, ν_{QSD} is the unique quasi-stationary distribution associated to X.

In [56, Section 4.1.1], we extended this result to birth and death processes with catastrophes. The existence of quasi-stationary distributions for similar processes was already studied in [250]. The settings are the following. Let X^c be a birth and death process on $\mathbb{Z}_+$ with birth rates $(b_x)_{x\ge 0}$ and death rates $(d_x)_{x\ge0}$ with $b_0 = d_0 = 0$ and $b_x, d_x > 0$ for all $x \ge 1$, and allow the process to jump to 0 from any state $x \ge 1$ at rate $a_x \ge 0$. In particular, the jump rate from 1 to 0 is $a_1 + d_1$. This process is absorbed in $\partial = 0$ at time $T_0^c := \inf\{t \ge 0,\ X_t^c = 0\}$.

Proposition 3.3. *Assume that* $\sup_{n\ge1} a_n < \infty$. *Then $S < +\infty$ if and only if Assumption A is satisfied.*

We conclude this section with an original extension of the above proposition to birth and death processes with (possibly large) negative jumps. These processes are called skip-free to the right in [169], where the quasi-stationary behaviour of skip-free to the left processes are studied. Let X^{sf} be a process on $\mathbb{Z}_+$ with transition rate matrix $(Q(x,y))_{x,y\in\mathbb{Z}_+}$ of X given by

$$Q(x,y) = \begin{cases} 0, \text{ if } x = 0 \text{ or } y \ge x+2, \\ b_x, \text{ if } x \ge 1 \text{ and } y = x+1, \\ d_x = d_x^{(1)}, \text{ if } x \ge 2 \text{ and } y = x-1, \\ d_x^{(2)}, \text{ if } x \ge 3 \text{ and } y = x-2, \\ \vdots \\ d_x^{(k)}, \text{ if } x \ge k \text{ and } y = x-k, \\ \vdots \\ d_x^{(x)}, \text{ if } y = 0. \end{cases}$$

Here we assume that the families $(b_x)_{x\ge1}$ and $(d_x^{(k)})_{x\ge1,k\ge1}$ are positive and $b = 0$, so that the process is irreducible on $\mathbb{N}$ and with only absorbing state 0, whose hitting time is denoted by $T_0^{sf} := \inf\{t \ge 0,\ X_t^{sf} = 0\}$. The proof of this result can be obtained as a straightforward adaptation of the arguments of [56, Section 4.1.1] and is omitted here.

Proposition 3.4. *Assume that $S < \infty$ and that* $\sup_x d_x^{(x)} < \infty$. *Then Assumption A is satisfied.*

3.3 The case of λ_0-positive recurrent birth and death processes

We consider now the more involved situation where ∞ is not an entrance boundary. This case, where there may exist an infinity of quasi-stationary distributions, is trickier and can be partially solved, as we will show, when the birth and death process is λ_0-positive recurrent, as defined below. We refer the reader to [138, 248, 247, 249] for several properties and examples of λ_0-positive birth and death processes.

The quantity λ_0 in Theorem 3.1 is equal to the decay parameter (see Theorem 3.3 in [245]), usual to the theory of λ_0-positive semigroups (see for instance [8] and references therein) and defined as follows: for all $x \in \mathbb{N}$,

$$\lambda_0 = \inf\left\{\lambda > 0, \text{ s.t. } \liminf_{t\to+\infty} e^{\lambda t}\mathbb{P}_x(X_t = x) > 0\right\}. \tag{3.5}$$

Definition 3.1. The birth and death process X is said to be λ_0*-positive recurrent* if the decay parameter λ_0 is positive and if, for some $x \in \mathbb{N}$ and hence for all $x \in \mathbb{N}$, we have

$$\lim_{t\to\infty} e^{\lambda_0 t}\mathbb{P}_x(X_t = x) > 0.$$

In the following theorem (proved in [259], we assume that the process is λ_0-positive recurrent and we exhibit a subset of the domain of attraction for the minimal quasi-stationary distribution.

Theorem 3.5. *Let X be a λ_0-positive recurrent birth and death process as in Section 3.1. Then the domain of attraction of the minimal quasi-stationary distribution of X contains the set $\mathcal{D}$ defined by*

$$\mathcal{D} = \left\{\mu \in \mathcal{M}_1(\mathbb{N}),\ \sum_{i=1}^{\infty} \mu_i Q_i(\lambda_0) < +\infty\right\}.$$

Assume moreover that there exist $C > 0$, $\lambda_1 > \lambda_0$ and $\varphi : \mathbb{Z}_+ \to [1, +\infty)$ such that $\varphi(i)$ goes to infinity when $i \to \infty$ and

$$b_i(\varphi(i+1) - \varphi(i)) + d_i(\varphi(i-1) - \varphi(i)) \le -\lambda_1\varphi(i) + C,\ \forall i \ge 1. \tag{3.6}$$

Then the domain of attraction of the minimal quasi-stationary distribution of X contains the set $\mathcal{D}_\varphi$ defined by

$$\mathcal{D}_\varphi = \left\{\mu \in \mathcal{M}_1(\mathbb{N}),\ \sum_{i=1}^{\infty} \mu_i \varphi(i) < +\infty\right\}.$$

As shown in [259], we have $\mathcal{D}_\varphi \subset \mathcal{D}$ for all function φ satisfying the assumptions of Theorem 3.5. However, $Q_\cdot(\lambda_0)$ cannot be computed explicitly but in few situations. On the contrary, in many situations, it is possible to guess a function φ satisfying the Lyapunov criterion of the above theorem. In fact, in this situation, the results of Chapter 2 apply, as stated in the next proposition, and hence improve the description of the convergence.

Importantly, we do not assume that the process is λ_0-positive recurrent in the next statement. This is a key step for the generalisation developed afterwards, since in these cases the classical theory is lacking practical criteria for of λ_0-positive recurrence (for general considerations on these properties, we refer the reader to [166, 167]).

Proposition 3.6. *Assume that X is a birth and death process as in Section 3.1 and that there exist $C > 0$, $\lambda_1 > \lambda_0$ and $\varphi : \mathbb{Z}_+ \to [1, +\infty)$ such that*

$$b_i(\varphi(i+1) - \varphi(i)) + d_i(\varphi(i-1) - \varphi(i)) \le -\lambda_1 \varphi(i) + C, \ \forall i \ge 1. \tag{3.7}$$

Then there exist positive constants C', γ and a probability measure ν_{QSD} on E such that

$$\left\| \mathbb{P}_\mu(X_t \in \cdot \mid t < \tau_\partial) - \nu_{QSD} \right\|_{TV} \le C' \frac{\mu(\varphi)}{\mu(\eta)} e^{-\gamma t}, \tag{3.8}$$

where $1 \le \eta \le \varphi$ is the right-eigenfunction of the infinitesimal generator of X with eigenvalue $-\lambda_0$.

The above proposition appeared in [58] in a more general but weaker form, since $\mu(\varphi)/\mu(\eta)$ was replaced by $\mu(\varphi)/\mu(\varphi')$, where $\varphi' : E \to \mathbb{R}_+$ is a positive bounded function (while η is lower bounded away from 0 and may be unbounded). However the extension can be obtained in several ways: either as a consequence of the non-uniform exponential ergodic of the Q-process[1] (using the fact that η is lower bounded) or extending [21]. Both methods use the existence of an eigenfunction (already established in [246]), but the results may also be obtained directly by using the arguments of Chapter 7 and hence extended to more general processes, as those described in the next sections and chapters.

Finally, note that the subset of the domain of attraction provided by Proposition 3.6 is included (often strictly) in $\mathscr{D}$ of Theorem 3.5.

3.4 Multi-dimensional birth and death processes

We focus now on the extension of the above results to the case of multi-dimensional birth and death processes, as studied in [59, 58]. We focus first on the general case (see Subsection 3.4.1) and show that, in the case of Lotka-Volterra type parameters, we obtain uniform convergence to the quasi-stationary distribution with respect to the initial distribution (see Subsection 3.4.2).

3.4.1 General processes in discrete state space and continuous time

Let X be a non-explosive[2] Markov process in a countable state space $E \cup \{\partial\}$ absorbed in ∂, with infinitesimal generator $\mathscr{L}$ acting on nonnegative real functions f on $E \cup \{\partial\}$ such that $\sum_{y \in E \cup \{\partial\}} q_{x,y} f(y) < \infty$ for all $x \in E$ as

$$\mathscr{L} f(x) = \sum_{y \neq x \in E \cup \{\partial\}} q_{x,y}(f(y) - f(x)), \quad \forall x \in E, \quad Lf(\partial) = 0, \tag{3.9}$$

where $q_{x,y}$ is the jump rate of X from x to $y \neq x$ and $\sum_{y \in E \cup \{\partial\} \setminus \{x\}} q_{x,y} < \infty$ for all $x \in E$.

Theorem 3.7. *Assume that there exists a finite subset D_0 of E such that $\mathbb{P}_x(X_1 = y) > 0$ for all $x, y \in D_0$, so that the constant*

$$\lambda_0 := \inf \left\{ \lambda > 0, \ s.t. \ \liminf_{t \to +\infty} e^{\lambda t} \mathbb{P}_x(X_t = x) > 0 \right\}$$

[1] See Chapter 6 for the definition and properties of Q-processes

[2] One could actually consider the case of explosive Markov processes, but then τ_∂ shall be defined as the infimum between the first hitting time of ∂ and the explosion time.

is finite and independent of $x \in D_0$. If in addition there exist constants $C > 0$, $\lambda_1 > \lambda_0$, a function $\varphi : E \cup \{\partial\} \to \mathbb{R}_+$ such that $\varphi_{|E} \geq 1$, $\varphi(\partial) = 0$, $\sum_{y\in E\setminus\{x\}} q_{x,y}\varphi(y) < \infty$ for all $x \in E$ and such that

$$\mathscr{L}\varphi(x) \leq -\lambda_1\varphi(x) + C\mathbf{1}_{x\in D_0},\ \forall x \in E, \tag{3.10}$$

then Assumption F is satisfied with $L = D_0$, $\gamma_1 = e^{-\lambda_1}$, any $\gamma_2 \in (e^{-\lambda_1}, e^{-\lambda_0})$ and $\psi_1 = \varphi_{|E}$. In addition, $\mathbb{P}_{\nu_{QSD}}(t < \tau_\partial) = e^{-\lambda_0 t}$ for all $t \geq 0$, the function η of Proposition 2.3 satisfies $P_t\eta = e^{-\lambda_0 t}\eta$ for all $t \geq 0$ and $\sum_{y\in E\setminus\{x\}} q_{x,y}\eta(y) < \infty$ and $\mathscr{L}\eta(x) = -\lambda_0\eta(x)$ for all $x \in E$.

Among earlier studies of the quasi-stationary behaviour of general continuous time Markov chains, we refer the reader to [211, 228, 229, 112] and the survey [252] and references therein.

Example 3.1. We consider general multitype birth and death processes in continuous time, taking values in a connected (in the sense of the nearest neighbours structure of $\mathbb{Z}^d$) subset E of $\mathbb{Z}_+^d$ for some $d \geq 1$, with transition rates

$$q_{x,y} = \begin{cases} b_i(x) & \text{if } y = x + e_i, \\ d_i(x) & \text{if } y = x - e_i, \\ 0 & \text{otherwise,} \end{cases}$$

with $e_i = (0,\dots,0,1,0,\dots,0)$ is the i^{th} element of the canonical basis and with the convention that the process is sent instantaneously to ∂ when it jumps to a point $y \notin E$ according to the previous rates. To ensure irreducibility, it is sufficient (although not optimal) to assume that $b_i(x) > 0$ and $d_i(x) > 0$ for all $1 \leq i \leq d$ and $x \in E$.

We show below that Theorem 3.7 applies under the assumption that

$$\frac{1}{|x|}\sum_{i=1}^d (d_i(x) - b_i(x)) \xrightarrow[x\in E,\ |x|\to+\infty]{} +\infty. \tag{3.11}$$

or that there exists $\delta > 1$ such that

$$\sum_{i=1}^d (d_i(x) - \delta\, b_i(x)) \xrightarrow[x\in E,\ |x|\to+\infty]{} +\infty. \tag{3.12}$$

Let us first show that (3.11) implies that the assumptions of Theorem 3.7 are satisfied. In order to do so, we define $\varphi(x) = |x| = x_1 + \ldots + x_d$ and $\varphi(\partial) = 0$ and obtain

$$\mathscr{L}\varphi(x) = \sum_{i=1}^d (b_i(x) - d_i(x)) = -\varphi(x)\frac{\sum_{i=1}^d (d_i(x) - b_i(x))}{|x|}$$

The proof is concluded by setting $D_0 = \left\{x \in E,\ \text{s.t. } \frac{\sum_{i=1}^d (d_i(x)-b_i(x))}{|x|} \geq \lambda_0 + 1\right\}$.

Let us now show that (3.12) implies that the assumptions of Theorem 3.7 are satisfied. Setting $\varphi(x) = \exp\langle a, x\rangle$ for a given $a \in (0,\infty)^d$ and $\varphi(\partial) = 0$, we obtain

$$\mathscr{L}\varphi(x) \leq -\varphi(x)\left(\sum_{i=1}^d (1 - e^{-a_i})d_i(x) + (1 - e^{a_i})b_i(x)\right).$$

Choosing $a = (\varepsilon,\dots,\varepsilon)$ with ε small enough, we have

$$\liminf_{x\in E,\ |x|\to+\infty} \sum_{i=1}^d (1 - e^{-a_i})d_i(x) + (1 - e^{a_i})b_i(x) = +\infty.$$

Taking $D_0 = \left\{x \in E,\ \text{s.t. } \sum_{i=1}^d (1 - e^{-a_i})d_i(x) + (1 - e^{a_i})b_i(x) \geq \lambda_0 + 1\right\}$ allows us to conclude the proof.

3.4.2 Birth and death processes with Lotka-Volterra type parameters

The following result was obtained in [59] as part of a more general result based on Lyapunov type criterion. Since this general criterion is rather technical, we focus here on its application to Lotka-Volterra birth and death processes.

A Lotka-Volterra birth and death process in dimension $d \geq 2$ is a Markov process $(X_t)_{t\in\mathbb{R}_+}$ on $\mathbb{Z}_+^d$ with transition rates $q_{n,m}$ from $n = (n_1, \dots, n_d) \in \mathbb{Z}_+^d$ to $m \neq n$ in $\mathbb{Z}_+^d$ given by

$$q_{n,m} = \begin{cases} n_i(\lambda_i + \sum_{j=1}^d \gamma_{ij} n_j) & \text{if } m = n + e_i, \text{ for some } i \in \{1, \dots, d\} \\ n_i(\mu_i + \sum_{j=1}^d c_{ij} n_j) & \text{if } m = n - e_i, \text{ for some } i \in \{1, \dots, d\} \\ 0 & \text{otherwise.} \end{cases}$$

We have $q_{n,n-e_i} = 0$ if $n_i = 0$, so that the process remains in the state space $\mathbb{Z}_+^d$. Since in addition $q_{n,m} = 0$ for all n such that $n_i = 0$ and m such that $m_i \geq 1$, the set $\partial = \mathbb{Z}_+^d \setminus \mathbb{N}^d$ is absorbing for the process. We make the usual convention that

$$q_{n,n} := -q_n := -\sum_{m \neq n} q_{n,m}.$$

From the biological point of view, the constant $\lambda_i > 0$ is the birth rate per individual of type $i \in \{1, \dots, d\}$, the constant $\mu_i > 0$ is the death rate per individual of type i, $c_{ij} \geq 0$ is the rate of death of an individual of type i from competition with an individual of type j, and $\gamma_{ij} \geq 0$ is the rate of birth of an individual of type i from cooperation with (or predation of) an individual of type j. In general, a Lotka-Volterra process could be explosive if some of the γ_{ij} are positive, but the assumptions of the next theorem ensure that it is not the case and that the process is almost surely absorbed in finite time.

Proposition 3.8. *Consider a competitive Lotka-Volterra birth and death process $(X_t)_{t\in\mathbb{R}_+}$ in $\mathbb{Z}_+^d$ as above. Assume that the matrix $(c_{ij} - \gamma_{ij})_{1\leq i,j\leq d}$ defines a positive operator on $\mathbb{R}_+^d$ in the sense that, for all $(x_1, \dots, x_d) \in \mathbb{R}_+^d \setminus \{0\}$, $\sum_{ij} x_i(c_{ij} - \gamma_{ij})x_j > 0$. Then the process has a unique quasi-stationary distribution ν_{QSD} and there exist constants $C, \gamma > 0$ such that, for all probability measures μ on $\mathbb{N}^d$,*

$$\left\| \mathbb{P}_\mu(X_t \in \cdot \mid t < \tau_\partial) - \nu_{QSD} \right\|_{TV} \leq Ce^{-\lambda t}, \quad \forall t \geq 0.$$

Note that the existence of a quasi-stationary distribution for this kind of multi-dimensional birth and death processes can also be obtained using the theory of positive matrices, as exposed in [119], or using the result of the preceding section. However, neither approach provides the uniform convergence with respect to the initial distribution.

Chapter 4

One-dimensional diffusion processes

In this chapter, we consider general one-dimensional diffusion processes and expose necessary and sufficient conditions for the exponential convergence to a quasi-stationary distribution. After reminders on general one-dimensional diffusion processes in Section 4.1, we study their quasi-stationary distributions when there is no natural boundaries in Section 4.2. We consider the case of stochastic differential equations (SDEs) with possibly natural boundaries in Section 4.3.

Convergence of conditioned one-dimensional diffusion processes has received a lot of attention in the past decades and general results have been obtained, using in general spectral theoretic arguments (self-adjoint operators, Sturm-Liouville theory, intrinsic ultra-contractivity), which proved to be extremely powerful (see for instance [239, 44, 184, 175, 146, 210]). Our main contribution to the theory of quasi-stationary distributions for one-dimensional diffusion processes, concerns the question of speed of convergence with respect to the initial distribution and the relaxation of the regularity of the coefficients. Moreover, our original approach can be easily adapted to other models, such as diffusion processes with jumps and time inhomogeneous diffusion processes (see Section 4.4 where several examples are provided).

4.1 Some reminders on general diffusion processes

In this section, we recall the definition and first properties of a general one-dimensional diffusion process $(Y_t, t \geq 0)$ on (a,b), $-\infty \leq a < b \leq +\infty$, up to its exit time of (a,b) defined by $\tau_\partial = \inf\{t \geq 0,\ \limsup_{s\to t} Y_s = b \text{ or } \liminf_{s\to t} Y_s = a\}$. The process Y is sent to a cemetery point $\partial \notin (a,b)$ for all $t \geq \tau_\partial$. Its distribution given $Y_0 = x \in (a,b)$ will be denoted $\mathbb{P}_x$. We refer the reader to [107, 154, 163, 127, 50] for additional developments and proofs of the following properties.

A stochastic process $(Y_t, t \geq 0)$ on (a,b) is called a *diffusion process* if it has a.s. continuous paths in (a,b) up to time τ_∂, satisfies the strong Markov property and is *regular*. By regular, we mean that for all $x, y \in (a,b)$, $\mathbb{P}_x(T_y < \infty) > 0$, where T_y is the first hitting time of y by the process Y. This notion is closely related to the concept of irreducibility of the set (a,b), since this open interval cannot be decomposed into strict subsets from which the process Y cannot evade.

To such a process, one can associate a continuous and strictly increasing function s on (a,b) such that, for all $l < x < r \in (a,b)$,

$$\mathbb{P}_x(T_l < T_r) = \frac{s(x) - s(l)}{s(r) - s(l)}.$$

The function s is called a *scale function* of Y and is unique up to affine transformation. If $s(x) = x$ is a scale function for Y, then we say that Y is *on natural scale*.

It can also be proved that, to any one dimensional diffusion process Y, one can associate a unique locally finite positive measure $m_Y(dx)$ on $(0,+\infty)$, called the *speed measure* of Y, which gives positive mass to any open subset of $(0,+\infty)$ and such that, for all $l < x < r \in (a,b)$ and for all measurable functions $f : (a,b) \to \mathbb{R}_+$

$$\mathbb{E}_x\left(\int_0^{T_l\wedge T_r} f(Y_s)\,ds\right) = \int_{(l,r)} G_{l,r}(x,y)\,f(y)\,m_Y(dy), \tag{4.1}$$

where G denotes the green function

$$G_{l,r}(x,y) = 2\frac{(s(x)\wedge s(y) - s(l))(s(r) - s(x)\vee s(y))}{s(r)-s(l)}.$$

This formula can be extended to $l = a$ or $r = b$ by letting l and r tend to a and b respectively in the above expressions. For instance, if $s(a) = 0$ and $s(b) = +\infty$, equation (4.1) remains valid with $l = a$, $r = b$, and $G_{a,b}(x,y) = 2\,s(x)\wedge s(y)$.

The meaning of the speed measure m_Y is somewhat counter-intuitive given its name: the process slows down in subsets with higher speed measure. For instance, the sticky Brownian motion with parameter $\theta > 0$, which is the general diffusion on natural scale on $(-\infty,+\infty)$ and speed measure $\Lambda + \theta\delta_0$ stays longer in 0 if θ is larger (where Λ is the Lebesgue measure on $\mathbb{R}$). It is possible to give a precise meaning of this fact by looking as the construction of diffusion processes on natural scale as time changed Brownian motion (since this considerations are further away from the presentation at hand, we refer the reader to the original paper [60] and references therein).

It is also well known that one can classify the boundaries a and b of the state space as *exit*, *regular*, *natural* or *entrance*. Informally, they respectively mean (exit) that the process can reach the boundary and cannot come back to (a,b); (regular) that the process can reach the boundary and may be constructed so that it comes back to (a,b), depending on the boundary conditions; (natural) that the process cannot reach the boundary and that, when it starts near the boundary, it stays a long time near the boundary; (entrance) that the process cannot reach the boundary and that, when it starts near the boundary, it can reach any compact subsets of (a,b) in finite time. A boundary is said to be *reachable* if it is exit or regular, and *non-reachable* otherwise.

These definitions correspond to the following parameters (see for instance Section 5.11 of Itô's book [153]). Set, for some fixed $c > 0$,

$$I_a = \int\int_{a<y<x<c} m_Y(dx)\,s(dy), \quad II_a = \int\int_{a<y<x<c} m_Y(dy)\,s(dx),$$
$$I_b = \int\int_{c<x<y<b} m_Y(dx)\,s(dy), \quad II_b = \int\int_{c<x<y<b} m_Y(dy)\,s(dx).$$

Then, for $\gamma = a$ or b, we have

- γ is an exit boundary if $I_\gamma < \infty$ and $II_\gamma = \infty$,
- γ is a regular boundary if $I_\gamma < \infty$ and $II_\gamma < \infty$,
- γ is a natural boundary if $I_\gamma = \infty$ and $II_\gamma = \infty$,
- γ is an entrance boundary if $I_\gamma = \infty$ and $II_\gamma < \infty$.

Also one checks that, if $\gamma \in \{a,b\}$ is a reachable boundary, then $s(\gamma) \in (-\infty,+\infty)$ and that, if $\gamma \in \{a,b\}$ is an entrance boundary, then $s(\gamma) \in \{-\infty,+\infty\}$.

Remark 4.1. The process $X_t = s(Y_t)$ is a diffusion on natural scale on $(s(a),s(b))$, whose speed measure is given by

$$m(dx) = m_Y \circ s^{-1}(dx), \tag{4.2}$$

where $m_Y \circ s^{-1}$ is the push-forward measure of m_Y through the function s (this follows for instance from [232, Thm. VII.3.6]). Moreover, the boundaries $s(a)$ and $s(b)$ for X have respectively the same nature as a and b for Y. Since it is straightforward that the diffusion processes X and Y have the same quasi-stationary behaviour, we restricted most of the original article to diffusion processes X on natural scale.In the present manuscript, we present theses results in the general situation of diffusion processes that may not be on natural scale.

The case of solutions to stochastic differential equations.

One of the most widespread case of diffusion processes in the literature concerns the solutions to stochastic differential equations. Let Y be the solution to the general SDE on (a,b) with $-\infty < a < b \leq +\infty$

$$dY_t = \sigma(Y_t)dB_t + \beta(Y_t)dt,\ Y_0 = x \in (a,b), \tag{4.3}$$

where $(1+|\beta|)/\sigma^2 \in L^1_{\text{loc}}((a,b))$ (which ensures the existence of a weak solution [163, Ch. 23] up to the exit time of (a,b)). In this case, one obtains the following semi-explicit expressions for the scale function and the speed measure (see for instance [232, Chapter VII, Section 3])

$$s(x) = \int_{(c,x)} \exp\left(-\int_{(c,y)} 2\beta(z)\sigma^{-2}(z)\,dz\right) dy \quad \text{and} \quad m(dx) = \frac{2dx}{s'(x)\,\sigma^2(x)} \tag{4.4}$$

for any arbitrarily fixed point $c \in (a,b)$.

4.2 General one-dimensional diffusion processes without natural boundaries

We first consider the situation where none of the boundary is natural. This case is already well understood in the literature for solution to SDEs with regular coefficients. For instance, combining the works of [239, 44, 184, 175, 146], one deduces that there exists a unique quasi-stationary distribution for solutions to SDE's on $(0,+\infty)$ with smooth coefficients, 0 reachable and ∞ entrance. These results also imply that the convergence toward a quasi-stationary distribution hold true pointwisely, for any initial distribution on $(0,+\infty)$ (some of the above cited works require the initial measures to be compactly supported, but the extension to non-compactly supported measures is not difficult). Of course, they can be extended without to much difficulty to the situation where $(0,+\infty)$ is replaced by an arbitrary non-empty open set (a,b). For general one-dimensional diffusion processes, Miura [210] used the abstract and powerful analytical results on the density of diffusion processes proved by Mastumoto in [202] in order to prove that there exists a unique quasi-stationary distribution attracting all initial distribution. The conditions of Matsumoto are satisfied in most practical situations.

Our two main aims are thus 1/ to extend the existing results by covering all the edge cases not covered by [202] and 2/ to precise the convergence results, using the results of Chapters 1 and 2.

We begin with the following result, which answers to the first aim above. It can be obtained using the approach of Section 4 in [58].

Proposition 4.1. *Assume that a is reachable or entrance, and that b is reachable or entrance. Then Condition F of Chapter 2 is satisfied with $\psi_1 = 1$.*

The main drawback of this result is that it does not imply uniform convergence with respect to the original distribution, while this is a very common situation as we explain now.

The following result was obtained in [60] for diffusion processes on natural scale evolving in $(0,+\infty)$ and absorbed at 0. Here, we translate this result to the general case of diffusion process that may not be on natural scale. The extension of this result to one-dimensional diffusions with killing is absorbed in [57]. In the following result, $t_c > 0$ is an arbitrarily fixed time.

Theorem 4.2. *We have equivalence between*

(i) *The boundary a is reachable and*

$$\mathbb{P}(t_c < \tau_\partial \mid Y_0 = y) \le As(y), \quad \forall y \in (a,b),$$

or the boundary a is an entrance boundary. The boundary b is reachable and

$$\mathbb{P}(t_c < \tau_\partial \mid Y_0 = y) \le A(s(b) - s(y)), \quad \forall y \in (a,b),$$

or the boundary b is an entrance boundary.

(ii) *Assumption A of Chapter 1 is satisfied.*

In this case, if either a or b is reachable, then there exists $\lambda_0 > 0$ such that $\mathbb{P}_{\nu_{QSD}}(t < \tau_\partial) = e^{-\lambda_0 t}$ and the unique quasi-stationary distribution ν_{QSD} of the process is absolutely continuous with respect to m_Y with Radon-Nikodym density

$$\frac{d\nu_{QSD}}{dm_Y}(x) = 2\int_a^b G_{a,b}(x,y)\nu_{QSD}(dy),$$

where $G_{a,b}$ is the Green function defined in the previous section. In addition,

$$\int_a^b \mathfrak{s}(y)\nu_{QSD}(dy) < \infty,$$

where $\mathfrak{s}(y) = s(y) - s(a)$ if a is reachable and b entrance, $\mathfrak{s}(y) = s(b) - s(y)$ if a is entrance and b is reachable and $\mathfrak{s}(y) = (s(y) - s(a))(s(b) - s(y))$ if both a and b are entrance.

Remark 4.2. The Green function $G_{a,b}$ is well defined only if $s(a)$ or $s(b)$ is finite, which is the case when a or b is reachable. When both a and b are entrance boundaries, the classical theory of conservative Markov processes applies and one checks that $\alpha_Y = m_Y / m_Y(a,b)$ and $\lambda_0 = 0$.

Of course, in order to apply this result in practical situations, one needs to be able to check whether a given general diffusion process satisfies condition (i) above. This is the subject of Propositions 4.3. Its proof, detailed in [60] only uses elementary tools. In the next result, c is an arbitrary fixed point in (a,b).

Proposition 4.3. *Assume that b is reachable or entrance, and that, for all $x \in (a,c)$,*

$$I(x) := \int_{(a,x)} (s(y) - s(a))\, m_Y(dy) \le C(s(x) - s(a))^\rho \tag{4.5}$$

for some constants $C > 0$ and $\rho > 0$. Then, for all $t > 0$, there exists $A_t < \infty$ such that

$$\mathbb{P}_x(t < \tau_\partial) \le A_t\,(s(x) - s(a)),\ \forall x > 0.$$

We immediately deduce the next corollary.

Corollary 4.4. *Assume that a is regular or entrance, and that b is regular or entrance. Then the conditions of Theorem 4.2 are satisfied.*

Proposition 4.3 also covers many situations where the boundary is exit, although it does not cover all cases. To understand the generality of this criterion, let us consider the case of stochastic diffusion processes without drift terms. One easily checks that a diffusion process X evolving in $(0,+\infty)$ and solution to the SDE

$$dX_t = \sigma(X_t)dB_t$$

satisfies (4.5) as soon as $\sigma(x) \ge C' x^{1-\varepsilon}$ in the neighbourhood of 0 for some constants $\varepsilon > 0$ and $C' > 0$. We recall that, if $\sigma(x) \le C'x$ in a neighbourhood of 0, the boundary 0 is not reachable. As a consequence, this simple criterion covers most practical situations. However, it does not cover all cases: for instance, if $\sigma(x) = x \log x^{1/2-\varepsilon}$, then 0 is a reachable (exit) boundary, but (4.5) is not satisfied. This case is covered by a more involved criterion described in the original article [60] (which, however, does not cover all reachable boundary cases either). This leads to the following natural open question (which admits a positive answer for birth and death processes).

Open Question. Is it true that any diffusion process on (a,b) with a and b either entrance or reachable satisfies Condition A of Chapter 1?

This open question can also be translated into an open problem from strict martingale theory, as we detail in the original article [60].

Construction of a diffusion process with prescribed quasi-stationary distribution

Our goal here is to give a sufficient condition on a given positive measure α_Y ensuring that it is the unique quasi-stationary distribution of a diffusion process on (a,b) with identified speed measure and scale function. The problem of finding diffusion processes with prescribed quasi-stationary distributions have been independently studied by other authors since the publication of the original article. New advances with original applications to Monte Carlo methods can be found for instance in [230, 262].

Proposition 4.5. *Fix $-\infty \le a < b \le +\infty$, $\lambda_0 > 0$, a continuous strictly increasing function s on (a,b) and a probability measure α on (a,b). Assume that s and $\frac{1}{\mathfrak{t}(x)\wedge 1}\alpha(dx)$ are respectively the scale function and the speed measure of a regular diffusion process X on $(0,+\infty)$ satisfying the conditions of Theorem 4.2 with a or b reachable, where*

$$\mathfrak{t}(x) = \begin{cases} s(x) - s(a) & \text{if } a \text{ is reachable and } b \text{ is entrance for } X,\\ s(b) - s(x) & \text{if } a \text{ is entrance and } b \text{ is reachable for } X,\\ (s(x)-s(a))(s(b)-s(x)) & \text{if both } a \text{ and } b \text{ are reachable for } X.\end{cases}$$

Then the diffusion process Y with scale function s and speed measure

$$m_Y(dx) = \frac{\alpha(dx)}{\lambda_0 \int_0^\infty G_{a,b}(x,y)\alpha(dy)}, \quad \forall x \in (0,\infty)$$

satisfies the conditions of Theorem 4.2, its unique quasi-stationary distribution is α *and it satisfies*

$$\mathbb{P}_\alpha(t < \tau_\partial) = e^{-\lambda_0 t}, \quad \forall t \geq 0.$$

4.3 One dimensional SDEs with possibly natural boundaries

In this section, we consider the case of one-dimensional diffusion processes solution to a SDE and admitting a natural boundary. The case of general one-dimensional diffusion processes can be handled using a similar approach, but the construction of Lyapunov functions compatible with their infinitesimal generator is more tricky [154]. The question of existence and convergence to a quasi-stationary distribution for such diffusion processes is largely solved when the coefficients are regular (see in particular [79, 196, 197, 239, 44, 175, 146]). However, the question of finding a non-trivial subset of the domain of attraction of the quasi-stationary distribution is less understood (although some partial answers can be found in [79]). Our main goal is thus to extend those result to less regular coefficients and to give a criterion ensuring that Assumption F of Chapter 2 holds true, providing precisions on the domain of attraction and the rate of convergence. Note that we cannot recover the sub-exponential rate of convergence of [79].

Let X be the solution in (a,b), where $-\infty \leq a < b \leq +\infty$, to the SDE

$$dX_t = \sigma(X_t)\,dB_t + \beta(X_t)\,dt, \quad X_0 \in (a,b),$$

where $\sigma : (a,b) \to (0,+\infty)$ and $\beta : (a,b) \to \mathbb{R}$ are measurable functions such that $(1+|\beta|)/\sigma^2$ is locally integrable on (a,b). We assume that the process is sent to a cemetery point ∂ when it exits the set (a,b) and that it is subject to an additional killing rate $\kappa : (a,b) \to \mathbb{R}_+$ which is measurable and locally integrable w.r.t. Lebesgue's measure. This assumption implies that the killed process is regular.

We define λ_0 as

$$\lambda_0 := \inf\left\{\lambda > 0, \text{ s.t. } \liminf_{t\to+\infty} e^{\lambda t}\,\mathbb{P}_x(X_t \in [c,d]) > 0\right\} \tag{4.6}$$

for some $x \in [c,d] \subset (a,b)$. The fact that λ_0 does not depend on x nor $[c,d]$ is a consequence of the regularity of the process. This defines the *decay parameter*, analogously to (3.5) for birth and death processes.

Let $\delta : (a,b) \to \mathbb{R}_+$ and $s : (a,b) \to \mathbb{R}$ be defined by

$$\delta(x) = \exp\left(-2\int_c^x \frac{b(u)}{\sigma(u)^2}\,du\right) \quad \text{and} \quad s(x) = \int_c^x \delta(u)\,du,$$

for some arbitrary $c \in (a,b)$. We recall that s is the scale function of X (unique up to an affine transformation).

Theorem 4.6. *Assume that one among the following conditions* (i), (ii) *or* (iii) *holds true:*

(i) *a and b are reachable boundaries and* $\varphi = 1$*;*

(ii) *a is reachable and there exist $\lambda_1 > \lambda_0$, a $\mathscr{C}^2((a,b))$ function $\varphi : (a,b) \to [1,+\infty)$ and $x_1 \in (a,b)$ such that, for all $x \in [x_1, b)$,*

$$\frac{\sigma(x)^2}{2}\varphi''(x) + \beta(x)\varphi'(x) - \kappa(x)\varphi(x) \leq -\lambda_1\varphi(x); \tag{4.7}$$

(iii) *there exist $\lambda_1 > \lambda_0$, a $\mathscr{C}^2((a,b))$ function $\varphi : (a,b) \to [1,+\infty)$ and $x_0 < x_1 \in (a,b)$ such that* (4.7) *holds true for all $x \in (a, x_0) \cup (x_1, b)$.*

Then X admits a quasi-stationary distribution ν_{QSD} which satisfies $\nu_{QSD}(\varphi^{1/p}) < +\infty$ for all $p > 1$. Moreover, for all $p \in (1, \lambda_1/\lambda_0)$, there exist a constant $\alpha_p \in (0,1)$, a constant C_p and a positive function $\varphi_{2,p} : (a,b) \to (0,+\infty)$ uniformly bounded away from 0 on compact subsets of (a,b) such that, for all $\mu \in \mathscr{M}(\varphi_1^{1/p})$,

$$\left\| \mathbb{P}_\mu(X_t \in \cdot \mid t < \tau_\partial) - \nu_{QSD} \right\|_{TV} \leq C_p \alpha_p^t \frac{\mu(\varphi^{1/p})}{\mu(\varphi_{2,p})}, \ \forall t \in [0,+\infty).$$

In particular, ν_{QSD} is the only quasi-stationary distribution of X which satisfies $\nu_{QSD}(\varphi^{1/p}) < +\infty$ for at least one value of $p \in (1, \lambda_1/\lambda_0)$.

In order to apply this result in practice, one needs to find computable estimates for λ_0 and candidates for φ. One may for instance use the sharp bounds for the first eigenvalue of the (Dirichlet) infinitesimal generator of $(X_t)_{t\in\mathbb{R}_+}$ obtained in a L^2 (symmetric) setting in [226, 265, 266], as observed in [175]. We propose also in [58] two different upper bounds for λ_0 which follow from the characterisation (4.6) of the eigenvalue λ_0 and Dynkin's formula.

4.4 Examples

Example 4.1. In the settings of the last section, assume that $(a,b) = (0,+\infty)$, κ is locally bounded and that X is solution to the SDE in (a,b)

$$dX_t = \sqrt{X_t}\,dB_t - X_t\,dt.$$

Then 0 is reachable for X and since

$$\frac{\sigma(x)^2\delta(x)^2}{8s(x)^2} \xrightarrow[x\to+\infty]{} +\infty,$$

we deduce from [58, Proposition 4.7] and Theorem 4.6 that X admits a quasi-stationary distribution ν_{QSD} and, for all $p \geq 1$, there exist positive constants C_p, γ_p and a positive function $\varphi_{2,p}$ on $(0,+\infty)$ such that

$$\left\| \mathbb{P}_\mu(X_t \in \cdot \mid t < \tau_\partial) - \nu_{QSD} \right\|_{TV} \leq C_p \frac{\int_{(0,+\infty)} \exp(x/p)\,\mu(dx)}{\mu(\varphi_{2,p})}\, e^{-\gamma_p t},$$

for all probability measures μ on $(0,+\infty)$. In particular, one deduces that the domain of attraction ν_{QSD} contains any initial distribution μ admitting a finite exponential moment. Note that, in the case where $\kappa \equiv 0$, the process X is a continuous state branching process (Feller diffusion), for which quasi-stationarity was already studied (see [177] and the references therein).

Example 4.2. Assume that $(a,b)=\mathbb{R}$, that $\beta\equiv 0$ and σ is bounded measurable on $\mathbb{R}$. Assume also that the absorption of X is due to the killing rate $\kappa(x)=\kappa_0\left(1-\frac{1}{1+|x|}\right)$ for some constant $\kappa_0>0$. We deduce from [58, Proposition 4.7] that

$$\lambda_0 \le \frac{\pi^2\|\sigma\|_\infty^2}{8\mathfrak{b}^2}+\kappa_0\left(1-\frac{1}{1+\mathfrak{b}}\right)\le\kappa_0\left(1-\frac{1}{1+2\mathfrak{b}}\right)$$

for $\mathfrak{b}$ large enough. Moreover, choosing $\varphi=1$ and $x_0=-3\mathfrak{b}$, $x_1=3\mathfrak{b}$, one deduces that, for all $x\not\in[-x_1,x_1]$,

$$\frac{\sigma(x)^2}{2}\varphi''(x)-\kappa(x)\varphi(x)\le-\kappa_0\left(1-\frac{1}{1+3\mathfrak{b}}\right)\varphi(x).$$

Hence Theorem 4.6 implies that there exists a unique quasi-stationary distribution ν_{QSD} for X and that it attracts all probability measures μ on $\mathbb{R}$.

Example 4.3. We consider the case $(a,b)=(0,+\infty)$, $\sigma(x)=1$, $\beta(x)=x\sin x$, and $\kappa(x)=\kappa_0\left(1-\frac{1}{1+x}\right)$ for some constant $\kappa_0>\pi^2+3$. This corresponds to a SDE $dX_t=dB_t+\nabla U(X_t)dt$ where the potential $U(x)=\sin x-x\cos x$ has infinitely many wells with arbitrarily large depths, meaning that the process X without killing has a tendency to be “trapped” away from zero for large initial conditions. Nevertheless, thanks to the killing, we are able to prove convergence to a unique quasi-stationary distribution. Indeed, using [58, Proposition 4.7], we obtain

$$\lambda_0\le\sup_{x\in(0,1)}\frac{\pi^2}{2}+\frac{\sin x+x\cos x+x^2\sin^2 x}{2}+\kappa_0\left(1-\frac{1}{1+x}\right)\le\frac{\pi^2}{2}+\frac{3}{2}+\kappa_0/2.$$

Moreover, 0 is a reachable boundary for X and, taking $\varphi=1$, one has, for all $x_1>0$ and all $x>x_1$,

$$\frac{\sigma(x)^2}{2}\varphi''(x)+b(x)\varphi'(x)-\kappa(x)\varphi(x)\le-\kappa_0\left(1-\frac{1}{1+x_1}\right)\varphi(x)$$

Hence, since we assumed that $\kappa_0>\pi^2+3$, one deduces that there exists a unique quasi-stationary distribution ν_{QSD} for X and that it attracts all probability measures μ on $(0,+\infty)$.

Example 4.4 (Sticky Brownian motion absorbed at -1 and $+1$). We recall that a diffusion process on $\mathbb{R}$ with speed measure $\Lambda+\delta_0$ is called a *sticky Brownian motion* [154, 4], where Λ is the Lebesgue measure on $\mathbb{R}$. It is called “sticky” because it slows down at 0, giving the impression that the trajectory of the process is glued to 0. We consider here a diffusion process X evolving as a sticky Brownian motion in $(-1,1)$ and absorbed at -1 and 1. This means that X is a diffusion on natural scale with speed measure $m(dx)=\Lambda(dx)+\delta_0(dx)$ on $(-1,1)$, absorbed at -1 and 1.

In this case, the conditions of Theorem 4.2 are satisfied since both boundaries -1 and 1 are regular for X (see Corollary 4.4). Moreover, since the unique quasi-stationary distribution ν_{QSD} of X satisfies

$$\frac{d\nu_{QSD}}{dm}(x)=\lambda_0\int_{-1}^{1}(x\wedge y+1)(1-x\vee y)\,\nu_{QSD}(dy),\ \forall x\in(-1,1),$$

careful computations show that

$$\nu_{QSD}(dx)=\frac{\gamma^*}{2}\sin\left(\gamma^*(1+x)\wedge(1-x)\right)m(dx),$$

where γ^* is the unique solution in $(0,\pi]$ of cotan $\gamma=\gamma/2$.

Example 4.5 (Diffusion process with multiple sticking points). We consider here a diffusion process X on $(0,\infty)$ on natural scale and with absorption at 0, where ∞ is an entrance boundary and which "sticks" at the points $a_1, a_2, \ldots$, where $(a_i)_{i\geq 1}$ is decreasing, converges to 0 and $a_1 < 1$. Denoting by m the speed measure of this diffusion, this means that $\int_1^\infty y\, m(dy) < \infty$ and

$$m_{|(0,1)} = \Lambda_{|(0,1)} + \sum_{i\geq 1} \delta_{a_i},$$

where Λ is the Lebesgue measure on $\mathbb{R}$.

Assuming that there exist constants $C, \rho > 0$ such that for all $j \geq 1$,

$$\sum_{i\geq j} a_i \leq C a_j^\rho, \tag{4.8}$$

then for all $x \in (0,1)$, defining $i_0 := \inf\{j \geq 1 : a_j < x\}$,

$$\int_{(0,x)} y\, m(dy) = \frac{x^2}{2} + \sum_{i\geq i_0} a_i \leq \frac{x^2}{2} + C a_{i_0}^\rho \leq \frac{x^2}{2} + C x^\rho,$$

and we can apply Proposition 4.3. For example, the choice $a_i = i^{-\frac{1}{1-\rho}}$, for all $i \geq 1$, satisfies (4.8).

Example 4.6 (A simple model with jumps). The following example uses a simple extension of the results presented above.

We consider a diffusion process $(X_t)_{t\in\mathbb{R}_+}$ on $(0,\infty)$ with speed measure m, on natural scale, and satisfying the conditions of Theorem 4.2. Let us denote by $\mathscr{L}$ the infinitesimal generator of X. We consider the Markov process $(\widetilde{X}_t)_{t\in\mathbb{R}_+}$ with infinitesimal generator

$$\widetilde{\mathscr{L}} f(x) = \mathscr{L} f(x) + (f(x+1) - f(x))\mathbf{1}_{x\geq 1},$$

for all f in the domain of $\mathscr{L}$. In other words, we consider a càdlàg process following a diffusion process on natural scale with speed measure m between jump times, which occur at the jump times of an independent Poisson process $(N_t)_{t\in\mathbb{R}_+}$ of rate 1, with jump size +1 if the process is above 1, and 0 otherwise. We denote by $\widetilde{\tau}_\partial$ its first hitting time of 0.

In this situation, we show in [60] that $\widetilde{X}$ admits a unique quasi-stationary distribution ν_{QSD} on $(0,\infty)$ and that there exist two constants $C, \gamma > 0$ such that, for all initial distribution μ on $(0,\infty)$,

$$\left\| \mathbb{P}_\mu(\widetilde{X}_t \in \cdot \mid t < \widetilde{\tau}_\partial) - \nu_{QSD}(\cdot) \right\|_{TV} \leq C e^{-\gamma t},\ \forall t \geq 0. \tag{4.9}$$

Example 4.7 (One-dimensional diffusion processes with time-inhomogeneous coefficients). We consider a time inhomogeneous diffusion process X on $(0,+\infty)$ stopped when it exits $(0,+\infty)$ at time $T_0^X = \inf\{t \geq 0,\ X_{t-} = 0\}$, which is assumed finite almost surely, and solution, for all $s \geq 0$, on $[s, T_0^X)$ to

$$dX_t = \sigma(t, X_t)\, dB_t, \quad X_s \in (0,+\infty), \tag{4.10}$$

where B is a standard one-dimensional Brownian motion and σ is a measurable function on $(0,+\infty)\times(0,+\infty)$ to $(0,+\infty)$. We assume that

$$\sigma_*(x) \leq \sigma(t,x) \leq \sigma^*(x), \tag{4.11}$$

for some measurable functions σ^* and σ_* from $(0,+\infty)$ to $(0,+\infty)$ satisfying

$$\int_{(0,+\infty)} \frac{x\,dx}{\sigma_*(x)^2} < \infty \quad \text{and} \quad \int_{(a,b)} \frac{dx}{\sigma^*(x)^2} > 0,\ \forall 0 < a < b < \infty.$$

We also assume that $\sigma_*(x) \geq C x \log^{\frac{1+\varepsilon}{2}} \frac{1}{x}$ for some constants $C > 0$ and $\varepsilon > 0$ in a neighbourhood of the boundary 0.

Under the above assumptions, for all probability measures μ_1 and μ_2 on E, and for all $t \geq 0$, we show in [63] using coupling with the diffusion studied in the previous sections, that

$$\left\| \mathbb{P}_{\mu_1}(X_t \in \cdot \mid t < \tau_\partial) - \mathbb{P}_{\mu_2}(X_t \in \cdot \mid t < \tau_\partial) \right\| \leq C e^{-\gamma t}$$

for some positive constants $C, \gamma > 0$.

A more complete result is provided in [63], where we also state a general criterion for time-inhomogeneous semi-groups and provide applications to birth and death processes in random environments. Note that the classical (without absorption) version of these results (e.g. when both 0 and ∞ are an entrance boundary for σ_* and σ^*), can be obtained using existing time-inhomogeneous Doblin type criteria, known to apply since decades [76].

Chapter 5

Multi-dimensional diffusion processes

In this chapter, we consider the application of the results of Chapters 1 and 2 to diffusion processes absorbed at the boundary of a domain. We give a general criterion in Section 5.1 and apply it to uniformly elliptic diffusions in Section 5.2 and to an example with vanishing diffusion coefficient at the boundary of the domain in Section 5.3. Of course the study of the quasi-stationary behaviour of multi-dimensional diffusion processes is not new and has been largely understood in many situation, see for instance [224, 89, 130, 174, 94, 53]. Our main contribution is to strongly reduce regularity requirements of these works (both on the boundary of the domain and on the coefficients), to prove exponential convergence in total variation norm and to provide a non-trivial subset of the domain of attraction.

5.1 A general criterion

The results of this section and of the following one first appeared in [58].

We consider a diffusion process X on a connected, open domain $D\subset\mathbb{R}^d$ for some $d\geq 1$, solution to the SDE

$$dX_t = b(X_t)\,dt+\sigma(X_t)\,dB_t, \tag{5.1}$$

where B is a standard, r-dimensional Brownian motion and $b:D\to\mathbb{R}^d$ and $\sigma:D\to\mathbb{R}^{d\times r}$ are locally Hölder functions, such that σ is locally uniformly elliptic in D, i.e.

$$\forall K\subset D \text{ compact},\quad \inf_{x\in K}\inf_{s\in\mathbb{R}^d\setminus\{0\}}\frac{s^*\sigma(x)\sigma^*(x)s}{|s|^2}>0,$$

where $|\cdot|$ is the standard Euclidean norm on $\mathbb{R}^d$. We assume that the process is immediately absorbed [1] at some cemetery point $\partial\not\in D$ at its first exit time of D, denoted τ_∂. The existence and basic properties of this process are detailed in Subsection 12.1 of the original article [58]. We can observe that, for all $k\geq 1$, defining the compact set

$$K_k=\left\{x\in D:|x|\leq k \text{ and } d(x,D^c)\geq 1/k\right\},$$

[1] The study of diffusion processes with additional soft killing can also be derived from the same lines, see for instance [58, Section 4.4].

a weak solution to (5.1) can be constructed up to the first exit time $\tau_{K_k^c}$ of K_k. The proper definition of the absorption time τ_∂ is then

$$\tau_\partial = \sup_{k\geq 1} \tau_{K_k^c}. \tag{5.2}$$

We introduce the differential operator associated to the SDE (5.1), related to the infinitesimal generator of the process X: for all $f \in \mathscr{C}^2(D)$, we define for all $x \in D$

$$\mathscr{L} f(x) := \sum_{i=1}^d b_i(x) \frac{\partial f}{\partial x_i}(x) + \frac{1}{2} \sum_{i,j=1}^d \sum_{k=1}^r \sigma_{ik}(x)\sigma_{jk}(x) \frac{\partial^2 f}{\partial x_i \partial x_j}(x). \tag{5.3}$$

We also define the constant

$$\lambda_0 := \inf\left\{\lambda > 0, \text{ s.t. } \liminf_{t\to+\infty} e^{\lambda t}\mathbb{P}_x(X_t \in B) > 0\right\} \tag{5.4}$$

for some $x \in D$ and some open ball B such that $\overline{B} \subset D$. It is standard to prove using Harnack inequalities that, under the previous assumptions, $\lambda_0 < +\infty$ and its value is independent of the choice of $x \in D$ and of the non-empty, open ball B such that $\overline{B} \subset D$.

We obtain the following result.

Theorem 5.1. *Assume that there exist some constants $C > 0$, $\lambda_1 > \lambda_0$, a $\mathscr{C}^2(D)$ function $\varphi : D \to [1,+\infty)$ and a subset $D_0 \subset D$ closed for the relative topology on D such that $\sup_{x\in D_0} \varphi(x) < +\infty$ and*

$$\mathscr{L}\varphi(x) \leq -\lambda_1 \varphi(x) + C\mathbf{1}_{x\in D_0}, \ \forall x \in D. \tag{5.5}$$

Assume also that there exists a time $s_1 > 0$ such that

$$\sup_{x\in D_0} \mathbb{P}_x(s_1 < \tau_{K_k} \wedge \tau_\partial) \xrightarrow[k\to\infty]{} 0. \tag{5.6}$$

Then X admits a quasi-stationary distribution ν_{QSD} which satisfies $\nu_{QSD}(\varphi^{1/p}) < +\infty$ for all $p > 1$. Moreover, for all $p \in (1, \lambda_1/\lambda_0)$, there exist a constant $\alpha_p \in (0,1)$, a constant C_p and a positive function $\varphi_{2,p} : D \to (0,+\infty)$ uniformly bounded away from 0 on compact subsets of D such that, for all probability measures μ on E satisfying $\mu(\varphi^{1/p}) < \infty$,

$$\left\|\mathbb{P}_\mu(X_t \in \cdot \mid t < \tau_\partial) - \nu_{QSD}\right\|_{TV} \leq C_p \alpha_p^t \frac{\mu(\varphi^{1/p})}{\mu(\varphi_{2,p})}, \ \forall t \in [0,+\infty).$$

In particular, ν_{QSD} is the only quasi-stationary distribution of X which satisfies $\nu_{QSD}(\varphi^{1/p}) < +\infty$ for at least one value of $p \in (1, \lambda_1/\lambda_0)$.

Remark 5.1. The assumptions of Theorem 5.1 do not ensure the non-explosion of the Markov process X. In the event of an explosion, the absorption time τ_∂ is equal to the explosion time.

The last result has other consequences of interest, gathered in the next corollary.

Corollary 5.2. *Under the assumptions of Theorem 5.1, the infimum defining the constant λ_0 in (5.4) is actually a minimum and it satisfies $\mathbb{P}_{\nu_{QSD}}(t < \tau_\partial) = e^{-\lambda_0 t}$ for all $t \geq 0$. In addition, the function η of Theorem 2.4 satisfies $P_t \eta = e^{-\lambda_0 t}\eta$ for all $t \geq 0$. In particular, η belongs to the domain of the infinitesimal generator of the semigroup of the process X defined as acting on the Banach space $L^\infty(\varphi_1)$, and it is an eigenfunction for the eigenvalue $-\lambda_0$. In addition, $\eta \in \mathscr{C}^2(D)$ and $\mathscr{L}\eta(x) = -\lambda_0 \eta(x)$ for all $x \in D$.*

It is natural to ask if $\varphi_{2,p}$ may be replaced by η in the conclusion of Theorem 5.1. It is not too difficult to see that this is the case if η is bounded (this comes down to the fact that one can take $\varphi_2 = \eta$ in Assumption E when η is bounded). Actually, it is also the case if η is not bounded, in which case one needs to apply the above criteria to the $\varphi_{1,p}$-transform of the diffusion process X, following the same approach as in Chapter 7.

5.2 Application to uniformly elliptic diffusion processes

We consider the case where σ can be extended as a locally uniformly elliptic matrix to $\mathbb{R}^d$. We emphasise that, contrary to previous results on the existence of quasi-stationary distributions for diffusions in a domain (see [224, 89, 130, 174, 94, 53]), no regularity on the boundary of D is required.

Corollary 5.3. *Let D be an open connected subset of $\mathbb{R}^d$, $d \geq 1$. Let X be solution to the SDE*

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t,\ t < \tau_\partial,$$

where $b : \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times r}$ are locally Hölder continuous in $\mathbb{R}^d$ and σ is locally uniformly elliptic on $\mathbb{R}^d$. Recall the definition (5.4) *of λ_0 and assume that there exist constants $C > 0$, $\lambda_1 > \lambda_0$, a $\mathscr{C}^2(D)$ function $\varphi : D \to [1, +\infty)$ and a bounded subset $D_0 \subset D$ closed in D such that*

$$\mathscr{L}\varphi(x) \leq -\lambda_1\varphi(x) + C\mathbf{1}_{x\in D_0},\ \forall x \in D. \tag{5.7}$$

Then the process X absorbed at the boundary of D satisfies the assumptions of Theorem 5.1.

Again, we do not assume that φ is a norm-like function, hence the process X may be explosive (see Remark 5.1). We give now three examples of application.

Example 5.1. Assume that D is bounded. Then, one can choose $D_0 = D$ and $\varphi_1 = 1$ in Corollary 5.3. Hence the process X has a unique quasi-stationary distribution ν_{QSD} whose domain of attraction is the whole set of probability measures on D. Since η is bounded in this case, one can actually prove that

$$\left\|\mathbb{P}_\mu(X_t \in \cdot \mid t < \tau_\partial) - \nu_{QSD}\right\|_{TV} \leq \frac{C}{\mu(\eta)}\,\alpha^t,\ \forall t \in [0, +\infty),$$

This completely solve the question of existence, uniqueness and convergence to a quasi-stationary distribution for such diffusion processes. However, finding the correct speed of convergence remains an open problem, since the factor $1/\mu(\eta)$ is not optimal. Indeed, as we will see in Section 5.4 under regularity assumptions on the boundary of D, the convergence can be proved uniform in μ.

Example 5.2. Assume that $D \subset \mathbb{R}^d_+$ is open connected and that

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t$$

in D, where $b : \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times r}$ are locally Hölder continuous in $\mathbb{R}^d$, σ is locally uniformly elliptic on $\mathbb{R}^d$ and

$$\frac{\langle b(x), 1\rangle}{\langle x, 1\rangle} \xrightarrow[|x|\to+\infty]{} -\infty,$$

where $\langle\cdot,\cdot\rangle$ is the standard Euclidean product in $\mathbb{R}^d$ and $|\cdot|$ is the associated norm. Then (5.7) is satisfied for $\varphi(x) = 1 + x_1 + \ldots + x_d$ and hence the process X absorbed at the boundary of D satisfies the assumptions of Theorem 5.1.

Example 5.3. Assume that $D \subset \mathbb{R}^d$ is open connected and that

$$dX_t = b(X_t)\,dt + dB_t$$

in D, where $b:\mathbb{R}^d \to \mathbb{R}^d$ is locally Hölder continuous in $\mathbb{R}^d$ and

$$\limsup_{|x|\to+\infty} \frac{\langle b(x), x\rangle}{|x|} < -\frac{3}{2}\sqrt{\lambda_0}, \tag{5.8}$$

where $\langle\cdot,\cdot\rangle$ is the standard Euclidean product in $\mathbb{R}^d$ and λ_0 is defined in (5.4). Then the process X absorbed at the boundary of D satisfies the assumptions of Theorem 5.1.

To apply this criterion, one can use *a priori* bounds on λ_0. If the best available bound is $\lambda_0 < +\infty$, one may still apply the above criterion using that (5.8) is implied by

$$\lim_{|x|\to+\infty} \frac{\langle b(x), x\rangle}{|x|} = -\infty.$$

5.3 Feller diffusion with competition

We study now the case of a diffusion matrix σ that cannot be extended out of D as a locally uniformly elliptic matrix. This example deals with Feller diffusions with competition and is motivated by models of population dynamics with d species in interaction, where absorption corresponds to the extinction of one of the populations [46].

Assume that $D = (0,\infty)^d$ and

$$dX_t^i = \sqrt{\gamma_i X_t^i}\,dB_t^i + X_t^i b_i(X_t)\,dt,$$

where $\gamma_i > 0$ for all $1 \le i \le d$, $B^1,\dots,B^d$ are independent standard Brownian motions and b_i are locally Hölder in $(0,\infty)^d$ and locally bounded in $\mathbb{R}_+^d$.

Proposition 5.4. *Assume that there exist constants $c_0, c_1 > 0$ such that*

$$\sum_{i=1}^d \frac{x_i b_i(x)}{\gamma_i} \le c_0 - c_1|x|, \quad \forall x \in (0,\infty)^d.$$

Then the process X absorbed at the boundary of D satisfies the assumptions of Theorem 5.1.

In order to prove Proposition 5.4, one shows that the assumptions of Theorem 5.1 hold true with $\varphi(x) = \exp(c(x_1/\gamma_1 + \ldots + x_n/\gamma_n))$, where $c = c_1 \min_i \gamma_i/\sqrt{d}$. See Section 4.3 in [58] for details.

Compared to the existing literature on multi-dimensional Feller diffusions (see [46]), this result covers cases where the process does not come down from infinity, e.g. $b_i(x) = r_i - \sum_{j=1}^d c_{ij}\frac{x_j}{1+x_j}$, for some positive constants r_i and c_{ij} such that $r_i < c_{ii}$ for all $1 \le i \le d$. Also, the case considered in [46] is restricted to (transformations of) Kolmogorov diffusions where the drift derives from a potential ($b = \nabla V$), which allows the authors to use a spectral theoretic approach as in the one-dimensional case [44]. In the case of logistic Feller diffusions, where $b_i(x) = r_i - \sum_{j=1}^d c_{ij}x_j$, this requires the additional assumption that the matrix $(c_{ij}\gamma_j)_{1\le i,j\le d}$ is symmetric. While our results on existence and convergence to quasi-stationary distributions are more general than those of [46], we do not recover the fine results they obtain on the spectrum of the process, such as its discreteness.

Multidimensional Feller diffusions absorbed when one of the coordinates hits 0. A competitive Lotka-Volterra Feller diffusion process in dimension $d \geq 2$ is a Markov process $(X_t)_{t\in\mathbb{R}_+}$ on $\mathbb{R}_+^d$, where $X_t = (X_t^1, \ldots, X_t^d)$, is solution to the stochastic differential equation

$$dX_t^i = \sqrt{\gamma_i X_t^i} dB_t^i + X_t^i \left(r_i - \sum_{j=1}^{d} c_{ij} X_t^j \right) dt, \quad \forall i \in \{1, \ldots, d\}, \tag{5.9}$$

where $B^1, \ldots, B^d$ are independent standard Brownian motions. The Brownian terms and the linear drift terms correspond to classical Feller diffusions, and the quadratic drift terms correspond to Lotka-Volterra interactions between coordinates of the process. The variances per individual γ_i is a positive number, and the growth rates per individual r_i can be any real number, for all $1 \leq i \leq d$. The competition parameters c_{ij} are assumed nonnegative for all $1 \leq i, j \leq d$, which corresponds to competitive Lotka-Volterra interaction. While this model enters the settings of Proposition 5.4, Theorem 5.1 cannot provide uniform convergence toward a quasi-stationary distribution. However, this can be done for competitive Lotka-Volterra diffusion processes using the criterion of Chapter 1, through the specialised drift condition of [59], where we obtain the following

Proposition 5.5. *Consider a competitive Lotka-Volterra Feller diffusion* $(X_t)_{t\in\mathbb{R}_+}$ *in* $\mathbb{R}_+^d$ *as above. Assume that* $c_{ii} > 0$ *for all* $i \in \{1, \ldots, d\}$. *Then the process has a unique quasi-stationary distribution* ν_{QSD} *and there exist constants* $C, \gamma > 0$ *such that, for all probability measures* μ *on* $(0, \infty)^d$,

$$\left\| \mathbb{P}_\mu(X_t \in \cdot \mid t < \tau_\partial) - \nu_{QSD} \right\|_{TV} \leq C e^{-\lambda t}, \quad \forall t \geq 0. \tag{5.10}$$

5.4 Uniform convergence using gradient estimates

The results of this section first appeared in [53].

We state here that gradient estimates on the semi-group of the an absorbed Markov process $(X_t)_{t\in\mathbb{R}_+}$ can imply the exponential convergence in the situation described in Example 5.1.

Let X be a diffusion process[2] in a compact manifold with boundary M absorbed at the boundary ∂M. We assume that one of the two following assumptions $S1$ or $S2$ is satisfied:

S1. M is a bounded, connected and closed C^2 Riemannian manifold with C^2 boundary ∂M and the infinitesimal generator of the diffusion process X is given by $L = \frac{1}{2}\Delta + Z$, where Δ is the Laplace-Beltrami operator and Z is a C^1 vector field.

S2. M is a compact subset of $\mathbb{R}^d$ with non-empty, connected interior and C^2 boundary ∂M and X is solution to the SDE $dX_t = \sigma(X_t) dB_t + b(X_t) dt$, where B is a r-dimensional Brownian motion, $b : M \to \mathbb{R}^d$ is bounded and continuous and $\sigma : M \to \mathbb{R}^{d\times r}$ is continuous, $\sigma\sigma^*$ is uniformly elliptic and for all $r > 0$,

$$\sup_{x,y\in M,\ |x-y|=r} \frac{|\sigma(x) - \sigma(y)|^2}{r} \leq g(r) \tag{5.11}$$

for some function g such that $\int_0^1 g(r) dr < \infty$. For instance, (5.11) is satisfied as soon as σ is uniformly α-Hölder on M for some $\alpha > 0$.

[2] General processes with gradient estimates are studied in the original article [53].

In both situations, one can use the gradient estimates obtained by Wang in [264] and Priola and Wang in [231]: there exists $t_1 > 0$ such that the process satisfies a gradient estimate of the form

$$\|\nabla P_{t_1} f\|_\infty \le C\|f\|_\infty,\ \forall f \in \mathscr{B}_b(M), \tag{5.12}$$

where $P_t f(x) = \mathbb{E}_x(f(X_t)\mathbf{1}_{t<\tau_\partial})$ denotes the Dirichlet semi-group of X and the notation $\|\nabla P_{t_1} f\|_\infty$ has to be understood as

$$\|\nabla P_{t_1} f\|_\infty := \sup_{x,y\in E\cup\{\partial\}} \frac{|P_{t_1} f(x) - P_{t_1} f(y)|}{\rho(x,y)}.$$

Theorem 5.6. *Assume that the diffusion process $(X_t)_{t\in\mathbb{R}_+}$ satisfies Assumption S1 or Assumption S2. Then Condition A of Chapter 1 and hence* (1.5) *are satisfied. Moreover, there exist two constants $B,\gamma > 0$ such that, for any initial distributions μ_1 and μ_2 on E,*

$$\left\|\mathbb{P}_{\mu_1}(X_t \in \cdot \mid t < \tau_\partial) - \mathbb{P}_{\mu_2}(X_t \in \cdot \mid t < \tau_\partial)\right\|_{TV} \le \frac{Be^{-\gamma t}}{\mu_1(\rho_\partial) \vee \mu_2(\rho_\partial)} \|\mu_1 - \mu_2\|_{TV}, \tag{5.13}$$

where ρ_∂ is the Euclidean distance to the boundary.

Remark 5.2. The gradient estimates of [231] are proved for diffusion processes with space-dependent killing rate $V : M \to [0,\infty)$. More precisely, they consider infinitesimal generators of the form

$$\mathscr{L} = \frac{1}{2}\sum_{i,j=1}^{d} [\sigma\sigma^*]_{ij}\partial_i\partial_j + \sum_{i=1}^{d} b_i\partial_i - V$$

with V bounded measurable. The proof proposed in [53] can be adapted to this setting.

5.5 Time inhomogeneous diffusion processes

Let D be a bounded open subset of $\mathbb{R}^d$ $(d \ge 1)$ whose boundary ∂D is of class C^2 and consider the stochastic differential equation

$$dZ_t = \sigma(t, Z_t)dB_t + b(t, Z_t)dt,\ Z_0 \in D, \tag{5.14}$$

where B is a standard d dimensional Brownian motion. We assume that the functions

$$\sigma : \begin{array}{l} [0,+\infty[\times\mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d \\ (t,x) \mapsto \sigma(t,x) \end{array} \quad \text{and } b : \begin{array}{l} [0,+\infty[\times\mathbb{R}^d \to \mathbb{R}^d \\ (t,x) \mapsto b(t,x) \end{array}$$

are continuous on $[0,+\infty[\times\mathbb{R}^d$. Moreover, we assume that they are time-periodic and Lipschitz in $x \in D$ uniformly in $t \in [0,+\infty[$. This means that there exist two constants $\Pi > 0$ and $C_0 > 0$ such that, for all $x, y \in D$ and $t \ge 0$,

$$\begin{aligned} &\sigma(t+\Pi, x) = \sigma(t,x) \text{ and } b(t+\Pi, x) = b(t,x), \\ &\|\sigma(t,x) - \sigma(t,y)\| + |b(t,x) - b(t,y)| \le k_0|x-y|. \end{aligned} \tag{5.15}$$

We also assume that the coefficients are elliptic, meaning here that there exists a constant $c_0 > 0$ such that

$$c_0|y| \le |\sigma(t,x)y|,\ \forall (t,x,y) \in [0,+\infty[\times D \times \mathbb{R}^d.$$

For all $s>0$ and any probability distribution μ on D, we denote by $(\mathcal{Z}^{\mu}_{s,t})_{t\geq s}$ the unique solution to this stochastic differential equation starting at time $s>0$ with distribution μ, killed when it hits the boundary and killed with a rate $\kappa(t,\mathcal{Z}^{\mu}_{s,t})\geq 0$, where $\kappa:[0,+\infty[\times D\to\mathbb{R}_+$ is a uniformly bounded non-negative measurable function which is also time-periodic (with period Π).

This model was studied in [94], where mixing properties for the conditioned process were proved, using, as in the preceding section, the gradient estimates proved in [231] and adapted to the time inhomogeneous setting, but where we assumed an additional involved differentiability assumption in order to make use of an involved tightness result elaborated in [258]. This last assumption proved to be useless in [53] by controlling the distance to the boundary with a coupling argument involving one-dimensional drifted Brownian motions and the fact that the distance to the boundary is C^2 in a vicinity of the boundary (see e.g. [95, Chapter 7, Section 8]). This coupling argument can be repeated here (the details are left to the reader), allowing us to state the following result, which is thus a direct improvement over [94].

Theorem 5.7. *Under the assumptions of this section, there exist two constants $C>0$ and $\gamma>0$ such that, for all $0\leq s\leq t$, and all probability measures μ_1 and μ_2 on D,*

$$\sup_{\mu_1,\mu_2\in\mathcal{M}_1(D)}\left\|\mathbb{P}(\mathcal{Z}^{\mu_1}_{s,t}\in\cdot\mid t<\tau_\partial)-\mathbb{P}(\mathcal{Z}^{\mu_2}_{s,t}\in\cdot\mid t<\tau_\partial)\right\|_{TV}\leq Ce^{-\gamma(t-s)}.$$

Additional general results on time-inhomogeneous processes conditioned not to be absorbed are also presented in [63]. They make use of the simple fact that the coupling methods used in the proof of Theorem 1.1 in Chapter 1 can be effortlessly adapted to the time inhomogeneous setting. We also refer the reader to [214, 215] for processes in time inhomogeneous environments.

Part III

Application of the criteria of Part I to the study of other problems

Chapter 6

The Q-process and quasi-ergodic properties

6.1 Definition and existence

The Q-process is obtained as a Markov process conditioned to never be absorbed. The event "*the process is never absorbed*" has probability 0 in most practical cases, where we have $\mathbb{P}_x(\tau_\partial < +\infty) = 1$. Hence conditioning on this event is ill defined and the Q-process is instead obtained here as the limit, when $T \to +\infty$, of the process conditioned not be absorbed before time T. We show in [56, 58] that it is well defined under Assumption A of Chapter 1 and under Assumption E of Chapter 2. It is also well defined under Assumption F. Earlier studies of Q-processes go back at least to [86, 87], see also [224, 42, 44] and [155], where unusual behaviour of processes conditioned to survive are exhibited. These processes belong to the more general class of penalized Markov processes, as studied in [235], used in order to derive the following result (stated here both for discrete and continuous time models).

Proposition 6.1. *Assume that either Assumption A, or Assumption E or Assumption F is satisfied. Then there exists a family* $(\mathbb{Q}_x)_{x\in E}$ *of probability measures on* Ω *defined by*

$$\mathbb{Q}_x(A) = \lim_{t\to+\infty} \mathbb{P}_x(A \mid t < \tau_\partial) \tag{6.1}$$

for all $\mathscr{F}_s$*-measurable set A. It defines an* E'*-valued homogeneous Markov process*[1]*. Moreover, if* X *is a strong Markov process under* $\mathbb{P}$*, then so is* X *under* $\mathbb{Q}$*.*

Remark 6.1. There are other ways of conditioning a Markov process to never be absorbed. For instance, for a non-explosive one dimensional diffusion process evolving in $(0, +\infty)$ and absorbed when it reaches 0, one may define a process as the limit, when $A \to +\infty$, of the process conditioned to reach A before reaching 0 (see [221] for a detailed investigation of this situation); the resulting conservative process is then different from the Q-process. Other interesting limiting processes may be studied when the quasi-stationary behaviour of the process is known, such as the *two-side taboo limit* introduced and studied in [129].

[1] E' is the set defined in 2 under Assumption E (and similarly under Assumption F). We define it as equal to the state space E under Assumption A.

In fact, the convergence (6.1) holds in total variation norm, uniformly over the initial distribution under Assumption A. (A weak converse result is also provided in the original paper.) We even obtain an exponential speed of convergence, which is useful to deduce quasi-ergodic properties of the process. In the following result, η is the right-eigenfunction whose existence has been established under Assumption A, E or F in Chapters 1 and 2. The result was proved in [61] under Assumption A and can be generalised as follows (the proof of the generalisation is omitted here, but it follows the same lines as the original one).

Proposition 6.2. *Assume that Assumption A or Assumption E or Assumption F holds true. Then there exist two constants $C,\gamma>0$ such that, for all $0\leq t\leq T$ and for all $\Gamma\in\mathscr{F}_t$,*

$$\|\mathbb{Q}_x(\Gamma)-\mathbb{P}_x(\Gamma\mid T<\tau_\partial)\|_{TV}\leq C\,\mu\left(\frac{f}{\eta}\right)e^{-\gamma(T-t)}, \tag{6.2}$$

where $f=\eta$ under Assumption A, $f=\varphi_1$ under Assumption E and $f=\psi_1$ under Assumption F.

One of the main feature of the Q-process as described above is that it is a η-transform of the original absorbed process. Note that the η transform includes a source term $e^{\lambda_0 t}$, so that it is conservative. Without this term, it would be a sub-Markov process with killing rate λ_0. In the case of diffusion processes, this is a general long time feature of h-transforms, as proved in [90, 225].

Proposition 6.3. *Under the assumptions of the above proposition, for all measurable set $A\subset E$ and $t\geq 0$,*

$$\mathbb{Q}_x(X_t\in\cdot)=\frac{e^{\lambda_0 t}}{\eta(x)}\mathbb{E}_x\left(\eta(X_t)\mathbf{1}_A(X_t)\right), \tag{6.3}$$

where $\lambda_0\geq 0$ is such that $\mathbb{P}_{\nu_{QSD}}(t<\tau_\partial)=e^{-\lambda_0 t}$ for all $t\geq 0$.

This feature is particularly interesting and has been used on several occasions. For instance, the notion of Q-process can be used to derive results on stochastic representations of the eigenvectors of sub-Markovian semi-groups. This is the subject of Proposition 2.4 in [198] (although the cited study is restricted to the finite dimensional case for simplicity, most of the results and proofs of the authors extend directly to models with infinite state space). Interestingly, the Q-process approach of this paper can also be used to derive very simply the main result of [48], where the authors obtain a “Markov chain representation of the normalised Perron-Frobenius eigenvector”.

The above representation of the Q-process naturally suggests the following use of h-transforms to deduce quasi-stationary properties. Assume that there exists a positive eigenfunction h for the semi-group $(P_t)_{t\geq 0}$ of an absorbed[2] process X, associated to some eigenvalue $e^{-\lambda_0 t}\in(0,1]$. Then one can define the h-transform of P as follows:

$$Q_t f(x)=\frac{e^{\lambda_0 t}}{h(x)}P_t(f(x)h(x)).$$

This defines the semi-group of a Markov process (without absorption) to which classical results for convergence to a stationary distribution can be used. Then one recovers the asymptotic behaviour of $(P_t)_{t\geq 0}$ through the formula

$$\mathrm{e}^{\lambda_0 t}P_t f(x)=Q_t(f/h)(x).$$

[2] actually, this development can also be used to derive the asymptotic stability of unbounded semi-groups, in which case λ_0 may be negative.

The difficulties are two-fold: first the existence of h must be pre-established, second the limiting behaviour of $Q_t(f/h)(x)$ must be obtained, although $1/h$ is typically unbounded. This general method has been used successfully for instance in [130, 210, 241, 123] and may be used in several other cases, in particular when explicit formulas or at least good estimates are available for the eigenfunction h.

6.2 Exponential ergodicity

We are interested now in the ergodic behaviour of the Q-process. These results were proved under Assumption A and under Assumption E in [56] and [58] respectively, and the extension to Assumption F is straightforward.

Proposition 6.4. *Assumption A, E or F implies that the probability measure β on E' defined by $\beta(dx) = \eta(x)\nu_{QSD}(dx)$ is the unique invariant distribution of the Markov process X under $(\mathbb{Q}_x)_{x\in E'}$. Moreover, there exist constants $C > 0$ and $\alpha \in (0,1)$ such that, for all initial distributions μ on E' such that $\mu(f/\eta) < \infty$,*

$$\left\|\mathbb{Q}_\mu(X_t \in \cdot) - \beta(\cdot)\right\|_{\mathcal{M}(f/\eta)} \le C\alpha^t \mu\big(f/\eta\big), \quad \forall t \ge 0, \tag{6.4}$$

where $f = \eta$ under Assumption A, $f = \varphi_1$ under Assumption E and $f = \psi_1$ under Assumption F.

Note that one can reformulate the last result as follows: for all $g \in L^\infty(f)$,

$$\left\|e^{-\lambda_0 t}P_t g(\cdot) - \eta(\cdot)\nu_{QSD}(g)\right\|_{L^\infty(f)} \le C\alpha^t f(x)\,\|g\|_{L^\infty(f)},\ \forall x \in E,\ \forall t \ge 0.$$

6.3 Quasi-ergodic properties

Quasi-ergodic theorems go back at least to [42]. We refer the interested reader to [65, 269, 142, 66, 67, 141] for further developments. The proof of the following result is developed in [61] under Assumption A. The proof of its extension to Assumptions E or F follows the same lines and is omitted here.

Corollary 6.5. *Assume that Assumption A or Assumption E or Assumption F holds true. Then there exists a positive constant C such that, for all $T > 0$ and all bounded measurable functions $g : E' \to \mathbb{R}$,*

$$\left|\mathbb{E}_x\left(\frac{1}{T}\int_0^T g(X_t)\,dt \mid T < \tau_\partial\right) - \int_E g\,d\beta\right| \le \frac{C\|g\|_\infty f(x)}{T\eta(x)},\ \forall x \in E', \tag{6.5}$$

where $f = \eta$ under Assumption A, $f = \varphi_1$ under Assumption E and $f = \psi_1$ under Assumption F.

One month before the release of the first preprint version of [61], where the above result has been announced, the quasi-ergodic result (6.5) has been obtained by He, Zhang and Zu [143, Thm. 2.1] under Assumption A, without the convergence estimate in $1/T$.

Recently, double quasi-ergodic properties have been developed by Zhang, Li and Song in [269, Theorem 3.2]. They can also be obtained under Assumption A, E or F, following very similar proofs, and are stated as follows.

Proposition 6.6. *Assume that Assumption A or Assumption E or Assumption F holds true. Then there exists a constant $C > 0$ such that, for all bounded measurable functions $g_1, g_2 : E' \to \mathbb{R}$ and constants $0 < p < q < 1$, we have*

$$\left|\mathbb{E}_x\left(\frac{1}{T}\int_0^T g_1(X_{pt})g_2(X_{qt})\,dt \mid T < \tau_\partial\right) - \beta(g_1)\beta(g_2)\right| \leq \frac{C\|g_1\|_\infty\|g_2\|_\infty f(x)}{\eta(x)T},\ \forall x \in E',$$

where $f = \eta$ under Assumption A, $f = \varphi_1$ under Assumption E and $f = \psi_1$ under Assumption F.

In [269, Theorem 3.6], the authors prove a nice consequence to the above double quasi-ergodic result, whose proof holds true here without modification. We thus obtain, using their approach combined with the above results, the following

Corollary 6.7. *Assume that Assumption A or Assumption E or Assumption F holds true. Then, for all $\varepsilon > 0$ and all bounded measurable functions $g : E' \to \mathbb{R}$,*

$$\mathbb{P}_x\left(\left|\frac{1}{T}\int_0^T g(X_t)\,dt - \beta(g)\right| \geq \varepsilon \mid T < \tau_\partial\right) \xrightarrow[T\to+\infty]{} 0,\ \forall x \in E'.$$

Chapter 7

R-positive recurrence of unbounded semi-groups

The aim of this chapter is to show how the results of Chapter 2 can be used to deduce effortlessly general criteria for the geometric convergence of normalised unbounded semigroups. This natural extension provides practical criteria for the R-positive recurrence of unbounded semigroups as developed in [213, Section 6.2] and [212]. It has applications to penalized Markov processes [92, 93], to the study of the long time behaviour of Markov branching processes (see for instance [150, 151, 152, 36, 156, 74, 34, 32, 33]), of non-conservative PDEs (see e.g. [20, 21] and references therein).

Let E be a measurable space and $(P_n)_{n\in\mathbb{Z}_+}$ be a positive semigroup on the set of bounded measurable functions on E. We shall consider cases where there exists a measurable (possibly unbounded) function $\psi_1 : E \to (0,+\infty)$ such that $P_1\psi_1 \leq c\psi_1$ for some constant c, so that $P_n f$ is naturally defined for all measurable $f \in L^\infty(\psi_1)$ and all positive measure μ such that $\mu(\psi_1) < +\infty$ (this corresponds to the settings described in [213, Section 6.2]). In this settings, the recent article [21] makes use of the methods developed in [56, 61] (which correspond to Chapters 1 and 2) to give a necessary and sufficient condition for the existence of a positive eigenfunction η of P_1 with eigenvalue θ_0 and the geometric convergence of $\theta_0^{-n}\mu P_n f$. We show below that this result can be strengthened as an immediate corollary of the results of Chapter 2 applied to the sub-Markov semigroup $\frac{P_n(\cdot\psi_1)}{c^n\psi_1}$ for the sufficient condition, and standard results on ergodicity of Markov processes applied to a well-chosen h-transform of P_n for the necessary condition.

Section 7.1 is devoted to the general statement of this result. We then explain in Section 7.2 how large classes of semigroups satisfying our hypotheses can be deduced from those studied in [61]. We focus on two applications: penalized semigroups associated to perturbed (discrete-time) dynamical systems in Subsection 7.2.1 and diffusion processes in Subsection 7.2.2.

7.1 Main result

We first introduce the assumptions on which our results are based. We state them following the same structure as Assumption E in Chapter 2 to emphasise their similarity.

Condition G. There exist positive real constants $\theta_1, \theta_2, c_1, c_2, c_3$, an integer $n_1 \geq 1$, two functions $\psi_1 : E \to (0,+\infty)$, $\psi_2 : E \to \mathbb{R}_+$ and a probability measure ν on a measurable subset K of E such that

G1 *(Local A1-A2).* $\forall x \in K$ and all measurable $A \subset K$,

$$P_{n_1}(\psi_1 \mathbf{1}_A)(x) \geq c_1 \nu(A)\psi_1(x) \quad \text{and} \quad \sup_{n\in\mathbb{Z}_+} \frac{\sup_{y\in K} P_n\psi_1(y)/\psi_1(y)}{\inf_{y\in K} P_n\psi_1(y)/\psi_1(y)} \leq c_2.$$

G2 *(Global Lyapunov criterion).* We have $\theta_1 < \theta_2$ and

$$\inf_{x\in K} \psi_2(x)/\psi_1(x) > 0, \ \sup_{x\in E} \psi_2(x)/\psi_1(x) \leq 1,$$
$$P_1\psi_1(x) \leq \theta_1\psi_1(x) + c_3\mathbf{1}_K(x)\psi_1(x), \ \forall x \in E$$
$$P_1\psi_2(x) \geq \theta_2\psi_2(x), \ \forall x \in E.$$

G3 *(Aperiodicity).* For all $x \in K$, there exists $n_4(x)$ such that for all $n \geq n_4(x)$,

$$P_n(\mathbf{1}_K\psi_1) > 0.$$

Theorem 7.1. *Assume that Condition G holds true. Then there exists a positive measure ν_P on E such that $\nu_P(\psi_1) = 1$ and $\nu_P(\psi_2) > 0$, and two constants $C < +\infty$ and $\alpha \in (0,1)$ such that, for all $f \in L^\infty(\psi_1)$ and all positive measure μ on E such that $\mu(\psi_1) < +\infty$ and $\mu(\psi_2) > 0$,*

$$\left|\frac{\mu P_n f}{\mu P_n\psi_1} - \nu_P(f)\right| \leq C\alpha^n \frac{\mu(\psi_1)}{\mu(\psi_2)}, \quad \forall n \in \mathbb{Z}_+. \tag{7.1}$$

In addition, there exists $\theta_0 > 0$ such that $\nu_P P_n = \theta_0^n \nu_P$ and a function $\eta : E \to \mathbb{R}_+$ such that $\theta_0^{-n} P_n\psi_1$ converges uniformly and geometrically toward η in $L^\infty(\psi_1)$ and such that $P_1\eta = \theta_0\eta$ and $\nu_P(\eta) = \nu_P(\psi_1) = 1$. Moreover, there exist two constants $C' > 0$ and $\beta \in (0,1)$ such that, for all $f \in L^\infty(\psi_1)$ and all positive measures μ on E such that $\mu(\psi_1) < +\infty$,

$$\left|\theta_0^{-n}\mu P_n f - \mu(\eta)\nu_P(f)\right| \leq C'\beta^n\mu(\psi_1). \tag{7.2}$$

Remark 7.1. One can check that replacing ψ_1 by ψ_2 in G1 and/or G3 give equivalent versions of Condition G. In [21], a similar result is obtained, but with the additional assumptions that $\psi_2 > 0$ on E and $n_1 = 1$. In this restricted case, one can easily check that their assumptions on the discrete-time semigroup are equivalent to ours. The fact that ψ_2 can vanish in Assumption G allows to deal with reducible processes (as in Section 2.6).

The proof of the above theorem is straightforward, since the semi-group $(Q_n)_{n\in\mathbb{N}}$ defined by

$$Q_n(f) = \frac{P_n(f\psi_1)}{(\theta_1 + c_2)^n\psi_1}, \ \forall n \geq 0, \ \|f\|_\infty \leq 1$$

satisfies Condition E in Chapter 2 with $\varphi_1 = 1$ and $\varphi_2 = \psi_2/\psi_1$, using $\theta_1/(\theta_1 + c_2)$ in place of θ_1, $\theta_2/(\theta_1 + c_2)$ in place of θ_2 and $c_1/(\theta_1 + c_2)^{n_1}$ in place of c_1. See the original paper [62] for the details.

Remark 7.2. The elementary method consisting in studying the sub-Markov semi-group (Q_n) instead of (P_n) is neither new nor specific to our approach. It can also be used to derive immediately sufficient criteria for the convergence of unbounded semi-groups from the abundant theory of sub-Markovian semi-groups, as developed for instance in [80, 78, 253, 123, 175, 146]. Note that a similar approach has been used in [34] to describe the asymptotic behaviour of the growth-fragmentation equation using Krein-Rutman theorem and other criteria for R-positivity.

Whether Assumption G is necessary for (7.1) is still an open problem up to our knowledge. However, if one assumes that there exists a positive eigenfunction η such that (7.2) holds true, then one recovers easily Assumption G by applying the classical counterpart of Forster-Lyapunov criteria for conservative semigroups. Here, the conservative semigroup is the one associated to the η-transform of P_n defined by $R_n f := \frac{\theta_0^{-n}}{\eta} P_n(\eta f)$ (which corresponds to the Q-process in the sub-Markovian case, cf. Chapter 6). The only difficulty in the proof of the following theorem is that η may vanish on some subset of E.

Proposition 7.2. *Assume that there exist a positive function $\psi : E \to (0,+\infty)$ and a non-negative eigenfunction $\eta \in L^\infty(\psi)$ of P_1 for the eigenvalue $\theta_0 > 0$, such that*

$$\left|\theta_0^{-n} P_n f(x) - \eta(x)\nu_P(f)\right| \le \zeta_n \psi(x),\ \forall x \in E,\ f \in L^\infty(\psi_1), \tag{7.3}$$

where $(\zeta_n)_{n\ge 0}$ is some positive sequence converging to 0. Then Assumption G is satisfied with $\psi_2 = \eta$ and with some function $\psi_1 \in L^\infty(\psi)$ such that $\psi \in L^\infty(\psi_1)$.

Remark 7.3. A similar partial counterpart to Proposition 7.2 was proven in [21], where the authors assume that ζ_n is geometrically decreasing, that η is positive and use the approach of [56] to conclude.

For continuous time semigroups $(P_t)_{t\in[0,+\infty)}$, the conclusions of Theorem 7.1 can be easily deduced from properties on the discrete skeleton $(P_{nt_0})_{n\in\mathbb{N}}$ (similar properties where already observed in Theorem 5 of [243] and in [61]). In the following result, the function η and the positive measure ν_P are the one of Theorem 2.1 applied to the discrete skeleton $(P_{nt_0})_{n\in\mathbb{N}}$.

Corollary 7.3. *Let $(P_t)_{t\in[0,+\infty)}$ be a continuous time semigroup. Assume that there exists $t_0 > 0$ such that $(P_{nt_0})_{n\in\mathbb{N}}$ satisfies Assumption G, $\left(\frac{P_t\psi_1}{\psi_1}\right)_{t\in[0,t_0]}$ is upper bounded by a constant $\bar{c} > 0$ and $\left(\frac{P_t\psi_2}{\psi_2}\right)_{t\in[0,t_0]}$ is lower bounded by a constant $\underline{c} > 0$. Then there exist some constants $C'' > 0$ and $\gamma > 0$ such that, for all $f \in L^\infty(\psi_1)$ and all positive measure μ on E such that $\mu(\psi_1) < +\infty$ and $\mu(\psi_2) > 0$,*

$$\left|\frac{\mu P_t f}{\mu P_t \psi_1} - \nu_P(f)\right| \le C'' e^{-\gamma t}\frac{\mu(\psi_1)}{\mu(\psi_2)},\quad \forall t \in [0,+\infty), \tag{7.4}$$

In addition, there exists $\lambda_0 \in \mathbb{R}$ such that $\nu_P P_t = e^{\lambda_0 t}\nu_P$ for all $t \ge 0$, and $e^{-\lambda_0 t}P_t\psi_1$ converges uniformly and exponentially toward η in $L^\infty(\psi_1)$ when $t \to +\infty$. Moreover, there exist some constants $C''' > 0$ and $\gamma' > 0$ such that, for all $f \in L^\infty(\psi_1)$ and all positive measures μ on E such that $\mu(\psi_1) < +\infty$,

$$\left|e^{-\lambda_0 t}\mu P_t f - \mu(\eta)\nu_P(f)\right| \le C''' e^{-\gamma' t}\mu(\psi_1),\quad \forall t \in [0,+\infty). \tag{7.5}$$

Remark 7.4. These results can be seen as an extension to bounded non-conservative semigroups of criteria of convergence for semigroups associated to Markov processes (in particular, Harris theorem and all its extensions based on Doblin's conditions and Foster-Lyapunov criteria, see e.g. [204, 102]) and as a practical alternative to R-recurrent Markov chains theory [243, 213, 212]. In particular, it provides an alternative to spectral theoretic results dealing with existence of eigenfunctions and convergence to them (e.g. Krein-Rutman theorem, spectral theory of symmetric operators, or the theorem of convergence of normalised semigroups of Birkhoff [37] and its extensions).

7.2 Some applications

Given a positive semigroup P acting on measurable functions on E, one can try to directly check Assumption (G) by finding appropriate functions ψ_1 and ψ_2. Another natural and equivalent strategy is to find a function ψ such that the semigroup defined by $Q_n f = \frac{P_n(\psi f)}{c^n \psi}$ is sub-Markovian and check that it satisfies Assumption E of Chapter 2. The main advantage of this last approach is that Q has a probabilistic interpretation as the semigroup of a sub-Markov process. As such, one can apply all the criteria developed in the first part of this manuscript and, more generally, use the intuitions and toolboxes of the theory of stochastic processes. Since both approaches are equivalent, this is rather a question of taste.

7.2.1 Perturbed dynamical systems

Let $f : \mathbb{R}^d \to \mathbb{R}^d$ be a locally bounded measurable function and consider the perturbed dynamical system $X_{n+1} = f(X_n) + \xi_n$ with $(\xi_i)_{i\in\mathbb{Z}_+}$ i.i.d. non-degenerate Gaussian random variables. We are interested in the asymptotic behaviour of the associated Feynman-Kac semigroup

$$P_n f(x) = \mathbb{E}_x\left(\prod_{k=1}^n G(X_k)\mathbf{1}_{X_k\in E} f(X_n)\right),$$

where E is a measurable subset of $\mathbb{R}^d$ with positive Lebesgue measure and $G : E \to (0, +\infty)$ is a measurable function.

Proposition 7.4. *Assume that $1/G$ is locally bounded, that $G(x) \le C \exp(|x|)$ for all $x \in E$ and some constant $C > 0$, and there exists $p > 1$ such that $|x| - p|f(x)| \to +\infty$ when $|x| \to +\infty$, then the semigroup $(P_n)_{n\in\mathbb{N}}$ satisfies Assumption G.*

7.2.2 Diffusion processes

Let $(X_t)_{t\in\mathbb{R}_+}$ be solution to the SDE

$$dX_t = dB_t + b(X_t)\,dt, \quad X_0 \in (0, +\infty)^d, \tag{7.6}$$

where B is a standard d-dimensional Brownian motion and $b : \mathbb{R}^d \to \mathbb{R}^d$ is locally Hölder. Let $r : (0, +\infty)^d \to \mathbb{R}$ be locally bounded and consider the semigroup $(P_t)_{t\in\mathbb{R}_+}$ defined by

$$P_t f(x) = \mathbb{E}_x\left(e^{\int_0^t r(X_u)\,du} f(X_t)\,\mathbf{1}_{X_s\in(0,+\infty)^d,\ \forall s\in[0,t]}\right). \tag{7.7}$$

The term $\mathbf{1}_{X_s\in(0,+\infty)^d,\ \forall s\in[0,t]}$ above corresponds to a killing at the boundary of the domain $(0, +\infty)^d$. Note that the solution to (7.6) may explode in finite time if b does not satisfy the linear growth condition. However, we assume by convention that $X_t \not\in (0, +\infty)^d$ after the explosion time, so that (7.7) makes sense. We refer to [61, Sections 4.1 and 12.1] for the precise construction of the process.

One motivation for the study of this semigroup comes from the Feynman-Kac formula. Indeed, when the coefficients are smooth enough (see for instance [222, Section 1.3.3]), this semigroup is solution to the Cauchy linear parabolic partial differential equation

$$\begin{aligned} &r v - \frac{\partial v}{\partial t} + \mathscr{L} v = 0, \text{ on } [0, +\infty) \times (0, +\infty)^d \\ &v(0, \cdot) = f, \text{ on } (0, +\infty)^d, \end{aligned}$$

where $\mathscr{L}$ is the differential operator of second order

$$\mathscr{L}\varphi(x) = \frac{1}{2}\Delta\varphi(x) + b(x)\cdot\nabla\varphi(x), \quad \forall\varphi\in C^2(\mathbb{R}^d),$$

with Dirichlet boundary conditions.

Proposition 7.5. *Assume that*

$$r(x) + \sum_{i=1}^{d} b_i(x) \xrightarrow[|x|\to\infty,\ x\in(0,\infty)^d]{} -\infty. \tag{7.8}$$

Then the semigroup $(P_t)_{t\in[0,+\infty)}$ *satisfies the assumptions of Corollary 7.3.*

Chapter 8

Measure-valued Pólya urn processes

Measure-valued Pólya processes (MVPPs) are a generalisation of Pólya urns to the infinitely-many-colour case. Pólya urns date back to Pólya & Eggenberger [110], and have been thoroughly studied since then; highlights include, e.g., the seminal works of Athreya & Karlin [11] and Janson [157]. Although the question of generalising Pólya urns to infinitely-many colours was posed in 2004 in [157], MVPPs were only introduced recently by Bandyopadhyay & Thacker [19] and Mailler & Marckert [190].

In this chapter, we focus on almost-sure convergence of a large class of MVPPs using stochastic-approximation methods (in the spirit of Duflo [104] and Benaïm [25]). The main difficulty comes from the fact that the stochastic-approximation algorithm that we consider is defined on the space of measures on a *non-compact* space. All the results below originate from a collaboration with Cécile Mailler [191].

The stochastic-approximation approach is a classical method for the study of Pólya urn processes when the colour-set is finite, see for instance [25], [179], [178], Zhang [270] and the survey paper [220]. Our main contribution from the stochastic-approximation point of view is to prove convergence of a stochastic-approximation algorithm defined on a non-compact space, namely the set of probability measures on the colour-space (being an arbitrary Polish space). To our knowledge, very little is known for measure valued stochastic-approximation algorithm on non-compact spaces, with some exceptions such as [158] and [189].

Our main contribution to the theory of MVPPs is to prove almost-sure convergence for a large class of MVPPs (instead of the convergence in probability shown in [19, 190]). Furthermore, we generalise the definition of measure-valued Pólya processes to allow different colours to have different "weights", and to allow the so-called "replacement rule" to be random (two features that are classical in the context of Pólya urns). We are also able to treat the "non-balanced" case, which was not treated at all by Bandyopadhyay & Thacker [19] or Mailler & Marckert [190].

The link between Pólya urns and quasi-stationary distributions already exists in the literature; for example, Aldous, Flannery and Palacios [2] apply the convergence results of Athreya and Karlin [11] to approximating quasi-stationary distributions on a finite state space. Our main result generalises this work to the case of measure-valued Pólya processes.

8.1 Definition of the model and main result

Let E be a Polish space endowed with its Borel sigma-field. A measure-valued Pólya process (MVPP) is a Markov chain $(m_n)_{n\geq 0}$ taking values in $\mathscr{M}(E)$. It depends on three parameters: its *initial composition* m_0 a non-zero non-negative measure on E, a sequence of i.i.d. *replacement* kernels[1] $(R^{(n)})_{n\geq 1}$ on E, and a non-negative *weight* kernel P on E. We assume that

$T_{>0}$. almost surely, for all $x \in E$, $R^{(n)}_x$ is a non-negative measure.

Given m_n, we define m_{n+1} as follows: pick a random element Y_{n+1} of E according to the probability distribution proportional to $m_n P$, i.e., for all Borel set A of E,

$$\mathbb{P}(Y_{n+1} \in A \,|\, m_n) = \frac{\int_E P_x(A)\, dm_n(x)}{\int_E P_x(E)\, dm_n(x)}, \tag{8.1}$$

and then set

$$m_{n+1} = m_n + R^{(n+1)}_{Y_{n+1}}.$$

Measure-valued Pólya processes were originally introduced by [19] and [190], as a generalisation of d-colour Pólya urns, although they did not consider "weighted" MVPPs (they always had $P_x = \delta_x$ for all $x \in E$). Several examples are developed in Section 8.2. For now, let us recall the definition of a Pólya urn and show why MVPPs generalise this model. A d-colour Pólya urn is a Markov process $(U(n))_{n\geq 0}$ on $\mathbb{N}^d$ that depends on three parameters: the initial composition vector $U(0)$, the replacement matrix M, and weights $w_1, \dots, w_d \in (0,\infty)$. The vector $U(n)$ represents the content of an urn that contains balls of d different colours; balls of colour i all have weight w_i. Given $U(n)$, one defines $U(n+1)$ by picking a ball at random in the urn with probability proportional to its weight, denoting the colour of this random ball ξ_{n+1}, and setting $U(n+1) = U(n) + M_{\xi_n}$, where $M_1, \dots, M_d$ are the lines of M.

If we let $E = \{1, \dots, d\}$ and $m_n = \sum_{i=1}^d U_i(n)\delta_i$ for all $n \geq 0$, then m_n is a measure-valued Pólya process with replacement kernel

$$R^{(n)}_x = \sum_{i=1}^d M_{x,i}\delta_i \quad (\text{ almost surely for all } n \geq 0, 1 \leq x \leq d),$$

and weight kernel $P_x = w_x \delta_x$ for all $1 \leq x \leq d$.

Therefore, the MVPP process $(m_n)_{n\geq 0}$ can be thought of as a composition measure on a set E of colours, and the random variable Y_{n+1} can be seen as the colour of the "ball" drawn at time $n+1$. The main advantage of this wider model is that one can consider Pólya urns defined on an infinite, and even uncountable, set.

Our main result is to prove almost-sure convergence of the sequence $(m_n/m_n(E))_{n\geq 0}$ to a deterministic measure under the following assumptions. We denote by R the common expectation of the $R^{(n)}$'s and set $Q^{(n)} = R^{(n)}P$ for all $n \geq 1$, and $Q = RP$, meaning that, for all $x \in E$ and all Borel set $A \subseteq E$,

$$Q^{(n)}_x(A) = \int_E P_y(A)\, dR^{(n)}_x(y) \quad \text{and} \quad Q_x(A) = \int_E P_y(A)\, dR_x(y).$$

We assume that

[1] A kernel (resp. a non-negative kernel) on E is, by definition, a function from E into the set of measures (resp. non-negative measures) on E. In particular, for all $x \in E$, $R^{(n)}_x$ is a measure on E almost surely.

C1. for all $x \in E$, we have $Q_x(E) \le 1$, and there exists a probability measure μ on $\mathbb{R}$ with positive mean such that, for all $x \in E$, the law of $Q_x^{(i)}(E)$ stochastically dominates μ. In particular, setting $c_1 = \int_0^\infty x \,\mathrm{d}\mu(x)$,

$$0 < c_1 \le \inf_{x\in E} Q_x(E) \le \sup_{x\in E} Q_x(E) \le 1;$$

C2. there exists a locally bounded function $V : E \to [1,+\infty)$ such that,

(i) for all $N \ge 0$, the set $\{x \in E\colon V(x) \le N\}$ is relatively compact;

(ii) there exist two constants $\theta \in (0, c_1)$ and $K \ge 0$ such that

$$Q_x \cdot V \le \theta V(x) + K, \quad \forall x \in E,$$

(iii) and that there exist three constants $r > 1$, $p > \frac{\ln\theta}{\ln(\theta/c_1)} \vee 2$, $A > 0$ such that

$$\mathbb{E}\left[R_x^{(1)}(E)^r\right] \vee \mathbb{E}\left[Q_x^{(1)}(E)^p\right] \le AV(x), \quad \forall x \in E.$$

Under Assumption C1, Q is a non-negative kernel such that $\sup_x Q_x(E) \le 1$, so that $Q - I$ is the jump kernel (or infinitesimal generator) of a unique sub-Markovian transition kernel $(P_t)_{t\ge0}$ on E. We consider the continuous-time pure-jump Markov process $(X_t)_{t\ge0}$ on $E \cup \{\partial\}$, where $\partial \notin E$ is an absorbing state, with Markovian transition kernel $P_t + (1 - P_t(E))\delta_\partial$.

C3. the continuous-time pure jump Markov process X with sub-Markovian jump kernel $Q - I$ admits a quasi-stationary distribution $\nu_{QSD} \in \mathcal{M}_1(E)$. We further assume that the convergence of $\mathbb{P}_\alpha(X_t \in \cdot \mid X_t \ne \partial)$ holds uniformly with respect to the total variation norm on $\{\alpha \in \mathcal{M}_1(E) \mid \alpha \cdot V^{1/q} \le C\}$, for each $C > 0$, where $q = p/(p-1)$.

Finally, we need the following technical assumption:

C4. for all bounded continuous functions $f : E \to \mathbb{R}$, $x \in E \mapsto R_x f$ and $x \in E \mapsto Q_x f$ are continuous.

Under these assumptions, we are able to prove almost-sure convergence of the normalised MVPP $\tilde{m}_n := m_n/m_n(E)$. This result is applied to concrete models in Section 8.2

Theorem 8.1. *Under Assumptions $T_{>0}$ and C1– 4, if $m_0 \cdot V < \infty$ and $m_0 P \cdot V < \infty$, then the sequence of random measures $(m_n/n)_{n\ge0}$ converges almost surely to νR with respect to the topology of weak convergence. Moreover, $\sup_n\{m_n P \cdot V^{1/q}/n\} < +\infty$ almost surely, where $q = p/(p-1)$.*

Furthermore, if $\nu R(E) > 0$, then $(\tilde{m}_n)_{n\in\mathbb{N}}$ converges almost surely to $\nu R/\nu R(E)$ with respect to the topology of weak convergence.

Remark 8.1. Several refinements and precisions are provided on this result in the original article [191]. In particular, our more general result includes the possibility to remove ball from the urns, which is useful in several applications. This leads to several additional technicalities, both in the proof of our results and in the exposition of our assumptions.

Remark 8.2. If $R = Q$, then the quasi-stationary distribution ν is a left eigenfunction for R, with associated eigenvalue $\theta_0 \in (0,1]$. In particular, Theorem 8.1 implies that the average mass of m_n, i.e. $m_n(E)/n$, converges almost surely to θ_0.

8.2 Examples

8.2.1 The finite state space case

To illustrate how this theorem applies, let us first consider the simple case of a classical d-colour Pólya urn of random replacement matrix $M^{(n)}$ with no weights, where $(M^{(n)})_n$ is a sequence of i.i.d. random matrices with non-negative entries and mean M. We assume that $\sum_{i=1}^d M_{x,i} > 0$ for all $1 \le x \le d$ and that M is irreducible. Let $S = \max_{x=1}^d \sum_{i=1}^d M_{x,i}$, and let $m_n = \frac{1}{S}\sum_{i=1}^d U_i(n)\delta_i$, where $U_i(n)$ is the number of balls of colour i in the urn at time n. One can check that $(m_n)_{n\ge 0}$ is an MVPP with replacement kernel $R_x^{(n)} = \frac{1}{S}\sum_{i=1}^d M_{x,i}^{(n)}\delta_i$ on $E = \{1,\dots,d\}$, for all $n \ge 0$ and $1 \le x \le d$, such that $R = M/S$.

Note that, since we have no weights, $R = Q$. Let μ be the distribution of $\min_{x\in\{1,\dots,d\}} Y_x$, where $Y_1,\dots,Y_d$ are independent random variables respectively distributed as $Q_1^{(1)}(E),\dots,Q_d^{(1)}(E)$. Assumption C1 is satisfied since μ has positive mean $c_1 \le Q_x(E) \le 1$ for all $1 \le x \le d$. Assumption C2 is automatically satisfied since the colour space E is compact. Consider the process X on $E \cup \{\partial\}$ absorbed at ∂ and whose jump matrix restricted to E is given by $M/S - I$. Then, since M/S is irreducible, the process X conditioned on not hitting ∂ has a unique quasi-stationary distribution $\nu_{QSD} = \sum_{i=1}^n v_i\delta_i$, which is given by the unique non-negative left eigenvector v of $M/S - I$ and hence of M. It is also known (see e.g. Darroch & Seneta[87]) that there exists $C, \delta > 0$ such that $\|\mathbb{P}_\alpha(X_t \in \cdot | X_t \neq \partial) - \nu_{QSD}\|_{TV} \le Ce^{-\delta t}$ for all $\alpha \in \mathcal{M}_1(E)$, which thus implies C3. Finally, Assumption C4 is trivially satisfied since E is discrete.

Thus, Theorem 8.1 applies, and we get that, almost surely when n tends to infinity, $\tilde{m}_n \to \nu_{QSD}R/\nu_{QSD}R(E) = \nu_{QSD}$ (with respect to the topology of weak convergence), and thus, $U(n)/n \to v$, a result that dates back to Athreya & Karlin's work on generalised Pólya urns [11].

8.2.2 Ergodic Markov chains

In [190], the following example is treated: take $E = \mathbb{Z}_+$, fix $0 < \lambda < \mu$, and set

$$R_x = \frac{\lambda}{x\mu + \lambda}\delta_{x+1} + \frac{x\mu}{x\mu+\lambda}\delta_{x-1},$$

for all $x \neq 0$, and $R_0 = \delta_1$. This example is not weighted, meaning that $P_x = \delta_x$ for all $x \in E$, and balanced since $R_x(E) = 1$ for all $x \in E$. Note that the Markov chain of transition kernel R is the $M/M/\infty$ queue. The authors proved that that this MVPP satisfies

$$n^{-1} m_n \to \gamma \quad \text{in probability,}$$

where γ is the stationary measure of the $M/M/\infty$ queue, i.e.

$$\gamma(x) = \left(\frac{\lambda}{\mu}\right)^x \frac{e^{-\lambda/\mu}}{x!} \quad \forall x \in \mathbb{Z}_+.$$

Our result also applies to this situation and thus implies the stronger almost-sure convergence of this MVPP.

Since this example is simple, let us detail how one show that our result applies. First note that, in this example, the $R^{(i)}$ are deterministic and equal to R, $P_x = \delta_x$; therefore, $Q^{(i)} = Q = R$ ($\forall i \ge 1$).

Since $R_x(E) = 1$ for all $x \in \mathbb{N}$, then C1 is satisfied (we can take $\mu = \delta_1$, and thus, $c_1 = 1$). Assumption C2 also holds: one can take $V(x) = \mathrm{e}^x$, implying that

$$R_x \cdot V = \frac{\lambda \mathrm{e}^{x+1} + \mu x \mathrm{e}^{x-1}}{\lambda + \mu x} = \frac{\lambda \mathrm{e}^2 + \mu x}{\lambda + \mu x} \mathrm{e}^{x-1} = \frac{\lambda \mathrm{e}^2 + \mu x}{\mathrm{e}(\lambda + \mu x)} V(x).$$

Note that

$$\frac{\lambda \mathrm{e}^2 + \mu x}{\mathrm{e}(\lambda + \mu x)} < \frac{2}{\mathrm{e}} \Leftrightarrow x > \frac{\lambda(\mathrm{e}^2 - 2)}{\mu},$$

therefore,

$$R_x \cdot V \leq \theta\, V(x) + K,\ \forall x \in \mathbb{Z}_+$$

where $\theta = \frac{2}{\mathrm{e}} \in (0, c_1)$ and $K = \sup_{x \leq \lambda(\mathrm{e}^2-2)/\mu} R_x \cdot V$. Also note that, for all $r, p > 1$, we have

$$\mathbb{E} R_x^{(1)}(E)^r \vee \mathbb{E} Q_x^{(1)}(E)^p = R_x(E)^r \vee R_x(E)^p = 1,$$

implying that C2-(iii) holds, while the rest of C2 is clear. Since the queue $M/M/\infty$ is ergodic with stationary distribution ν_{SD}, we can infer that the continuous-time Markov process of generator $R - I$ is also ergodic. Moreover, one can show that, for any $q > 1$, $Q_x \cdot V^{1/q} \leq \theta^{1/q} V(x) + K^{1/q}$, where $\theta^{1/q} < 1$. This and the Foster-Lyapunov type criteria of [207] provide the uniform convergence to ν_{SD} required in Assumption C3. Finally, since $\mathbb{Z}_+$ is discrete, C4 is trivially satisfied. Thus, Theorem 8.1 applies and we can conclude that if $\sum_{k\geq 0} \mathrm{e}^k m_0(k)$ is finite, then

$$n^{-1} m_n \to \nu_{SD} \quad \text{almost surely when } n \to \infty.$$

8.2.3 Quasi-ergodic Markov chains

Let us now consider the more general case where $E = \mathbb{Z}_+$ and, for all $x \in E$,

$$R_x = \lambda_x \delta_{x+1} + \mu_x \delta_{x-1},$$

where $(\lambda_x)_x$ and $(\mu_x)_x$ are families of positive numbers such that $\mu_0 = 0$, $\lambda_0 > 0$, $\inf_{x\geq 1} \mu_x > 0$, $\sup_x \mu_x < \infty$ and $\lambda_x = o(\mu_x)$ when $x \to +\infty$. In this situation, the MVPP is not weighted, so that $P_x = \delta_x$ and $Q_x = R_x$ for all $x \in E$, and it is not balanced (hence the results of [19] and [190] do not apply).

We assume, without loss of generality, that $\sup_x(\lambda_x + \mu_x) = 1$, so that $Q_x(E) \leq 1$ for all $x \in E$. The situation is reminiscent of the simple example of Section 2.2 and the calculus are similar. In particular, we deduce that, for the irreducible process X with infinitesimal generator $Q - I$, there exist a quasi-stationary distribution ν_{QSD} for X and two positive constants Cst, $\delta > 0$ such that, for all probability measures $\alpha \in E$, satisfying $\alpha \cdot V^{1/q} < +\infty$, where $V = \exp(ax)$ with $a > 0$ large enough,

$$\|\mathbb{P}_\alpha(X_t \in \cdot \mid t < \tau_\partial) - \nu_{QSD}\|_{TV} \leq \mathrm{Cst}\, \alpha \cdot V^{1/q} \mathrm{e}^{-\delta t},$$

which entails Assumption C3 and provides a candidate for the long time behaviour of the MVPP $m_n/m_n(E)$. One can then easily check that the other assumptions of Theorem 8.1 hold true and hence that

$$\frac{m_n}{m_n(E)} \xrightarrow[n\to+\infty]{a.s.} \frac{\nu_{QSD} R}{\nu_{QSD} R(E)} = \nu_{QSD}.$$

with respect to the topology of weak convergence, as soon as $m_0(V) < +\infty$.

8.2.4 Random trees

As discussed in Janson [157, Examples 7.5 and 7.6], infinitely-many-colour urns are particularly useful for the study of some functionals of random trees. We give below two examples where our main result applies, and gives stronger convergence results.

Definition 8.1 (Outdegree profiles)**.** We define the out-degree profile of a rooted tree τ as

$$\mathrm{Out}(\tau) = \sum_{v\in\tau} \delta_{\mathrm{outdeg}(v)},$$

where for all nodes v in τ, outdeg(v) is the out-degree of v (i.e. its number of children).

Out-degree profile in the random recursive tree. The random recursive tree $(\mathrm{RRT}_n)_{n\geq 1}$ is a sequence of random rooted trees defined recursively as follows:

- RRT_1 has one node (the root);
- we build RRT_{n+1} from RRT_n by choosing a node of RRT_n uniformly at random, and adding a child to this node.

It is straightforward to see that the sequence $(\mathrm{Out}(\mathrm{RRT}_n))_{n\geq 1}$ of the out-degree profile of the random recursive tree is a MVPP on $\mathbb{Z}_+$ of initial composition $m_1 = \delta_0$, and replacement kernel

$$R_x = -\delta_x + \delta_0 + \delta_{x+1}, \quad \forall x \in \mathbb{Z}_+.$$

Note that the replacement measures R_x are not positive, but the process satisfies the additional assumptions detailed in Section 1.4 of [191] for unbalanced MVPPs. In this case, $P_x = \delta_x$, and $R^{(i)} = R = Q$ almost surely for all $i \geq 1$. Note that $Q_x(\mathbb{Z}_+) = 1$ for all $x \in \mathbb{Z}_+$, and, therefore, Assumption C1 holds with $\mu = \delta_1$ and $c_1 = 1$.

Choosing $\varepsilon \in (0, 1/2)$ and setting $V(x) = (2-\varepsilon)^x$ for all $x \geq 0$, we show in [191] that Theorem 8.1 applies and that

$$n^{-1}\mathrm{Out}(\mathrm{RRT}_n) \to \nu \quad \text{weakly, almost surely when } n \to \infty, \tag{8.2}$$

where $\nu_x = 2^{-x-1}$, for all $x \in \mathbb{Z}_+$.

Different versions of this result can be found in the literature: Bergeron, Flajolet & Salvy [30, Corollary 4] prove it using generating functions, Mahmoud & Smythe [186] prove a joint central limit theorem for the number of nodes of out-degree 0, 1 and 2, Janson [157, Example 7.5] extends this result by considering out-degrees $0, 1, \ldots, M$ for all $M \geq 0$, which implies (8.2). The approach of [186] and [157] relies on the remarkable fact that, in that particular example, one can reduce the problem to finitely many types.

Our main contribution for this example is to prove the convergence in a stronger sense, and thus answer a question of Janson (see Remark 1.2 [158]). Indeed, Theorem 8.1 also gives that, for all $q \in (1,2)$,

$$\sup_n \frac{\mathrm{Out}(\mathrm{RRT}_n)}{n} \cdot V^{1/q} < +\infty,$$

since $P_x = \delta_x$ for all x, in this example. This leads to the following proposition.

Proposition 8.2. *For all $\varepsilon \in (0, 1/2)$, for all $q \in (1,2)$, for all functions $f : \mathbb{Z}_+ \to \mathbb{R}$ such that $f(x) = o\big((2-\varepsilon)^{x/q}\big)$ when $x \to \infty$, we have*

$$\frac{1}{n}\int f \, d\mathrm{Out}(\mathrm{RRT}_n) \to \sum_{x=0}^{\infty} 2^{-x-1} f(x), \text{ almost surely when } n \to \infty.$$

Our approach also has the advantage of providing a framework that can be easily generalised, as, for example, in the next application to which Janson's finitely-many-types approach wouldn't apply.

Out-degree profile in a random recursive forest with multiple children. Let us now consider the following generalisation of the random recursive tree studied above. The random recursive forest $(\mathrm{RRF}_n)_{n\geq 1}$ with multiple children is defined as a sequence of random rooted forests defined recursively as follows: consider a probability measure α on $\{-1\} \cup \{1,2,\ldots\}$ (with $0 < \alpha_{-1} < 1$) and a probability measure β on $\{1,2,\ldots\}$;

- RRF_1 has one node (the root);
- we build RRF_{n+1} from RRF_n by choosing a node of RRF_n uniformly at random, and, if this node has at least one child,
  - with probability α_{-1}, remove the edge between the node and one of his children (hence forming an other tree in the forest),
  - with probability α_k ($k \geq 1$), add k children to this node,

  while, if this node has 0 child, with probability β_k ($k \geq 1$), add k children to this node.

We define $\mathrm{Out}(\mathrm{RRF}_n)$ as the sum of the out-degree profiles (see Definition 8.1) of the trees composing the forest RRF_n and obtain the following result.

Proposition 8.3. *Assume that α and β both admit an exponential moment of order λ, for some fixed $\lambda > 0$. Then there exists a probability distribution ν_{QSD} such that, for all $q \in (1,2)$, for all $a > 0$ satisfying*

$$\sum_{k=1}^{+\infty} \alpha_k e^{ak} < 2\sum_{k=1}^{\infty} \alpha_k,$$

and for all function $f : \mathbb{Z}_+ \to \mathbb{R}$ such that $f(x) = o(e^{ax/q})$ when $x \to \infty$, we have

$$\int f \frac{d\mathrm{Out}(\mathrm{RRF}_n)}{\mathrm{Out}(\mathrm{RRF}_n)(E)} \to \int f \, d\nu_{QSD}, \text{ almost surely when } n \to \infty. \tag{8.3}$$

Protected nodes in the random recursive tree. A node v of a tree τ is 2-protected if the closest leaf is at distance at least 2 from v; in a social network, 2-protected nodes can be users who used to invite new users to the network but have not done so recently. The proportion of such nodes in different models of random trees have been studied in the literature: Motzkin trees in Cheon & Shapiro [72], random binary search tree in Bóna [40], and more recently in the m-ary search tree in Holmgren, Janson & Šileikis [148]. Devroye & Janson [96] show how results of Aldous [1] about fringe trees can be used to study this question with a unified approach for different models of random trees, including simply generating trees and the random recursive tree. We show here how our main result allows to get information about protected nodes in random trees.

For all $n \geq 1$ and $x \geq 0$, let us denote by $X_{n,x}$ the number of internal nodes in RRT_n having exactly x leaf-children. The random measure

$$m_n = \sum_{x\in\mathbb{N}} X_{n,x}\delta_x$$

is a MVPP $E = \mathbb{Z}_+$ of initial composition $m_0 = \delta_1$. The replacement kernel of $(m_n)_{n\geq 0}$ is (for all $i \geq 1$ and $x \geq 1$)

$$R_0^{(i)} = -\delta_0 + \delta_1 \quad \text{and} \quad R_x^{(i)} = B_{1/x+1}^{(i)}\delta_{x+1} + \big(1 - B_{1/x+1}^{(i)}\big)(\delta_{x-1} + \delta_1) - \delta_x,$$

where $\big(B_{1/x+1}^{(i)}\big)$ is a sequence of i.i.d. random Bernoulli-distributed variables of parameters $1/x+1$ for all $x \geq 1$. The weight kernel of $(m_n)_{n\geq 0}$ is $P_x = (x+1)\delta_x$, for all $x \in \mathbb{Z}_+$. We therefore have

$$R_0 = -\delta_0 + \delta_1 \quad \text{and} \quad R_x = \frac{1}{x+1}\delta_{x+1} + \frac{x}{x+1}(\delta_{x-1} + \delta_1) - \delta_x,$$

and

$$Q_x = \frac{x+2}{x+1}\delta_{x+1} + \frac{x}{x+1}(x\delta_{x-1} + 2\delta_1) - (x+1)\delta_x,$$

for all $x \geq 0$. Note that $Q_x(\mathbb{Z}_+) = 1$ for all $x \in \mathbb{Z}_+$.

We prove in [191] that the assumptions of Theorem 8.1 holds true and obtain the following

Proposition 8.4. *For all $x \geq 1$, the proportion $p_{n,x}$ of internal nodes having exactly x leaf-children in the n-node random recursive tree converges almost surely to*

$$\frac{2}{\mathrm{e}} \sum_{i\geq x+1} \frac{1}{i!}.$$

The proportion $p_{n,0}$ of protected internal nodes converges almost surely to $1 - 2/\mathrm{e}$. Moreover, for all $q \in (1,2)$ and all function $f : \mathbb{Z}_+ \to \mathbb{R}$ such that $f(x) = o\big(\prod_{i=2}^{x}(i-\varepsilon)^{1/q}\big)$ for some $\varepsilon > 0$ when $x \to \infty$, we have

$$\sum_{i\geq 0} p_{n,i} f(i) \to (1 - 2/\mathrm{e}) f(0) + \frac{2}{\mathrm{e}} \sum_{i\geq 1} f(i) \sum_{j\geq i+1} \frac{1}{j!}$$

almost surely when $n \to \infty$.

Using this result, one can show for instance that the proportion of protected internal nodes converges almost surely to $1/2 - 1/\mathrm{e}$, improving on the convergence in probability already established by Ward & Mahmoud [187].

Chapter 9

Reinforced processes

In this chapter, we study a random process with reinforcement, which evolves following the dynamics of a given absorbed Markov process and is resampled according to its occupation measure when it reaches the absorption point. We show in different situations that its occupation measure converges to the minimal quasi-stationary distribution of the absorbed Markov process.

Let X be a time homogeneous Markov process with state space $E \cup \{\partial\}$, where $\partial \notin E$ is an absorbing state for the process. We assume that $\mathbb{P}_x(\tau_\partial < \infty) = 1$ and $\mathbb{P}_x(t < \tau_\partial) > 0$ for all $t \geq 0$ and $\forall x \in E$.

We consider a random process $(Y_t)_{t\geq 0}$ with reinforcement, which evolves following the dynamic of X when it lies in E and which is resampled according to its occupation measure when it reaches ∂. More precisely, given a probability measure μ on E, we set

$$Y_t = \sum_{k=1}^{\infty} \mathbf{1}_{t\in[\theta_{k-1},\theta_k)} X^{(k)}_{t-\theta_{k-1}}, \quad \forall t \geq 0,$$

where $\theta_0 = 0$,

- $(X^{(1)}_t)_{t\in\mathbb{R}_+}$ is a realisation of the process $(X_t)_{t\in\mathbb{R}_+}$ with $X^{(1)}_0 \sim \mu$ (i.e. under $\mathbb{P}_\mu$) and the stopping time θ_1 is defined as $\theta_1 = \tau^{(1)}_\partial$ the first hitting time of ∂ by $X^{(1)}$,
- given $X^{(1)}$, $(X^{(2)}_t)_{t\in\mathbb{R}_+}$ is a realisation of the process $(X_t, t \geq 0)$ with $X^{(2)}_0 \sim \mu_{\theta_1}$, where

  $$\mu_{\theta_1} = \frac{1}{\theta_1}\int_0^{\theta_1} \delta_{Y_s}\, ds$$

  and $\theta_2 - \theta_1 = \tau^{(2)}_\partial$ the first hitting time of ∂ by $X^{(2)}$,
- for all $k \geq 1$, given $X^{(1)}, X^{(2)}, \ldots, X^{(k)}$,$(X^{(k+1)}_t)_{t\in\mathbb{R}_+}$ is a realisation of the process $(X_t)_{t\in\mathbb{R}_+}$ with $X^{(k+1)}_0 \sim \mu_{\theta_k}$, where

  $$\mu_{\theta_k} = \frac{1}{\theta_k}\int_0^{\theta_k} \delta_{Y_s}\, ds$$

  and $\theta_{k+1} - \theta_k = \tau^{(k+1)}_\partial$ the first hitting time of ∂ by $X^{(k+1)}$.

We also define, for all $t \in \mathbb{R}_+$,

$$\mu_t = \frac{1}{t}\int_0^t \delta_{Y_s}\, ds, \quad \text{i.e.} \quad \mu_t(f) = \frac{1}{t}\int_0^t f(Y_s)\, ds, \quad \forall f \in \mathscr{B}_b(E).$$

This process has been studied in several situations, with the main goal of proving an almost sure convergence result for the occupation measure μ_t when $t \to +\infty$. In the finite state space case and in a discrete time setting, Aldous, Flannery and Palacios [2] solved this problem by showing that the proportion of colours in a Pólya urn type process converges almost surely to the left eigenfunction of the replacement matrix, which was also identified as the quasi-stationary distribution of a corresponding Markov chain. Under a similar setting but using stochastic approximation techniques, Benaïm and Cloez [27] and Blanchet, Glynn and Zheng[39] independently proved the almost sure convergence of the occupation measure μ_t toward the quasi-stationary distribution of X. These works have since been generalised to the compact state space case by Benaïm, Cloez and Panloup [28] under general criteria for the existence of a quasi-stationary distribution for X. The case of continuous time diffusion processes with smooth bounded killing rate on compact Riemannian manifolds has been recently solved by Wang, Roberts and Steinsaltz [263], who show that a similar algorithm with weights also converges toward the quasi-stationary distribution of the underlying diffusion process.

In Section 9.1, we solve the question of convergence of the occupation measure toward the quasi-stationary distribution of X when this process is a uniformly elliptic diffusion process evolving in an open bounded connected open set D with C^2 boundary ∂D, with hard killing when the process hits the boundary. This answers positively the open problem stated in Section 8 of [28].

In section 9.2, we state such a convergence result for processes with smooth and bounded killing rate evolving in unbounded spaces using a measure-valued Pólya process representation of this reinforced algorithm. This result strongly relies on the convergence of Measure-valued Pólya processes as stated in Chapter 8.

9.1 Stochastic approximation of a quasi-stationary distributions for diffusion processes in a bounded domain

Let $(X_t)_{t\in\mathbb{R}_+}$ be a diffusion process in a connected bounded open set D of $\mathbb{R}^d$, $d \geq 2$ with C^2 boundary ∂D and absorbed at ∂D. We assume that X is solution to the SDE

$$dX_t = \sigma(X_t)dB_t + b(X_t)dt, \tag{9.1}$$

where B is a r-dimensional Brownian motion, $b : D \to \mathbb{R}^d$ is bounded and continuous and $\sigma : D \to \mathbb{R}^{d\times r}$ is continuous, $\sigma\sigma^*$ is uniformly elliptic and for all $\rho > 0$,

$$\sup_{x,y\in D,\ |x-y|=\rho} \frac{|\sigma(x)-\sigma(y)|^2}{\rho} \leq g(\rho) \tag{9.2}$$

for some function g such that $\int_0^1 g(r)dr < \infty$. Note that, in this case, the process $(Y_t, \mu_t)_{t\geq 0}$ described in the introduction is well-defined since one can prove that $\theta_k \to +\infty$ a.s. [28, Lemma 8.1].

Recall that, the results of Chapter 5 entail that, under the above regularity assumptions, the killed diffusion process X satisfies Assumption A of Chapter 1 and hence that it admits a unique quasi-stationary distribution α. We denote by λ_0 the positive constant such that $\mathbb{P}_\alpha(t < \tau_\partial) = \exp(-\lambda_0 t)$ for all $t \geq 0$.

Remark 9.1. All the results of this chapter, and in particular the next one, can be extended to the one-dimensional diffusion processes studied in Chapter 4 and to the diffusion processes on compact manifolds studied in Chapter 5.

We obtain the following

Theorem 9.1. *For all bounded measurable function $f : D \to \mathbb{R}$, one has*

$$\mu_t f \xrightarrow[t\to+\infty]{} \alpha f \quad a.s.$$

Moreover, $\theta_n / n \to 1/\lambda_0$ almost surely when $n \to +\infty$.

The proof of this result, detailed in [26], relies, among other methods, on the properties of the Green operator A on $\mathscr{B}_b(E)$ for X, defined as

$$Af(x) = \mathbb{E}_x\left[\int_0^{\tau_\partial} f(X_s)\,ds\right] = \int_0^\infty P_s f(x)\,ds. \tag{9.3}$$

Assuming that X satisfies Conditions A of Chapter 1 (and hence in the present situation), one easily checks that this operator is bounded on $(\mathscr{B}_b(E), \|\cdot\|_\infty)$. For all $\mu \in \mathscr{M}_1(E)$, we also define the notation $\mu A f = \int_E Af(x)\,\mu(dx)$. We obtain the following

Proposition 9.2. *Assume that Condition A of Chapter 1 is satisfied. Then, for all $\mu \in \mathscr{M}_1(E)$, all $f \in \mathscr{B}_b(E)$ and all $n \geq 1$, we have*

$$\left|\mu A^n f - \frac{\alpha(f)\mu(\eta)}{\lambda_0^n}\right| \leq \|f\|_\infty \frac{C}{(\lambda_0+\gamma)^n}, \tag{9.4}$$

for some positive constant $C, \gamma > 0$. We also have for some constant $B > 0$

$$\left\|\frac{\mu A^n}{\mu A^n \mathbf{1}} - \alpha\right\|_{TV} \leq \frac{B}{\mu(\eta)}\left(\frac{\lambda_0}{\lambda_0+\gamma}\right)^n \tag{9.5}$$

and, for all $t \geq 0$,

$$\left\|\frac{\mu e^{tA}}{\mu e^{tA}\mathbf{1}} - \alpha\right\|_{TV} \leq \frac{B}{\mu(\eta)} e^{-t\frac{\gamma}{\lambda_0(\lambda_0+\gamma)}}. \tag{9.6}$$

9.2 Stochastic approximation of a quasi-stationary distributions for diffusion processes with soft killing

Let $(X_t)_{t\in\mathbb{R}_+}$ be the solution in $E = \mathbb{R}^d$ to the stochastic differential equation

$$\mathrm{d}X_t = \mathrm{d}B_t + b(X_t)\mathrm{d}t,$$

where B is a standard d-dimensional Brownian motion and $b : \mathbb{R}^d \mapsto \mathbb{R}^d$ is locally Hölder continuous in $\mathbb{R}^d$. We assume that X is subject to an additional soft killing $\kappa : x \mapsto [0, +\infty)$, which is continuous, uniformly bounded and such that $\kappa \geq 1$. Note that the quasi-stationary distribution of X with killing rate κ is the same as the quasi-stationary distribution of X with a killing rate $\kappa - 1$.

We also assume that

$$\limsup_{|x|\to+\infty} \frac{\langle b(x), x\rangle}{|x|} < -\frac{3}{2}\|\kappa\|_\infty^{1/2},$$

so that the process X admits a unique quasi-stationary distribution ν_{QSD} such that $\nu_{QSD} \cdot V < +\infty$, where $V : x \in \mathbb{R}^d \mapsto \exp(\|\kappa\|_\infty^{1/2}|x|)$ (this is an application of the results of Chapter 5).

We consider the self-interacting process $(Y_t)_{t\geq 0}$ evolving with the same dynamic of X but, at rate κ, instead of being killed, it jumps to a new position chosen accordingly to its empirical occupation measure, as described in the beginning of this chapter.

Proposition 9.3. *The empirical occupation measure* $\frac{1}{t}\int_0^t \delta_{Y_s}\,\mathrm{d}s$ *converges almost-surely when* $t \to +\infty$*, with respect to the topology of weak convergence, to the unique quasi-stationary distribution* ν_{QSD} *of* X *such that* $\nu_{QSD}(V) < \infty$.

The proof of this result uses the theory of Measure-Valued Pólya processes exposed in Chapter 8. More precisely, it derives from a larger class of models, called *sample paths Pólya-Urns*, whose study is developed in [191] for continuous and discrete-time models.

Part IV

Some other works

Chapter 10

Coarse Ricci curvature

Let (E,d) be a Polish space. Fix $N \geq 1$ and consider a continuous time pure jump particle system of N particles $(\bar{X}_t)_{t\geq 0} = (X^1_t,\dots,X^N_t)_{t\geq 0}$ evolving in E^N. We assume that the process is non-explosive and that its infinitesimal generator $\mathscr{L}$ is given, for all $\bar{x} = (x_1,\dots,x_N) \in E^N$ and any bounded measurable function $f : E^N \to \mathbb{R}$, by

$$\mathscr{L} f(\bar{x}) = \sum_{i=1}^{N} \int_E \big(f(x_1,\dots,x_{i-1},y,x_{i+1},\dots,x_N) - f(x_1,\dots,x_n)\big)\, F_i\left(x_i,\bar{x},dy\right),$$

where the terms $F_i(x_i,\bar{x},\cdot)$ are finite non-negative measures on E, measurable with respect to x_i and $\bar{x}$ and such that, for some (and hence for all) $\bar{x} \in E^N$, $\int d(x_i,y) F_i(x_i,\bar{x},dy) < \infty$. Our aim is to provide, using coupling methods, a lower bound for the coarse Ricci curvature of $\bar{X}$ evolving in E^N endowed with the metric

$$d(\bar{x},\bar{y}) = \frac{1}{N}\sum_{i=1}^{N} d(x_i,y_i),\ \forall \bar{x} = (x_1,\dots,x_N),\ \bar{y} = (y_1,\dots,y_N) \in E^N.$$

We recall that the coarse Ricci curvature of the continuous-time Markov process $\bar{X}$ is the largest constant σ satisfying, for all $t \geq 0$,

$$\mathscr{W}_d\left(\mathbb{P}(\bar{X}_t \in \cdot \mid \bar{X}_0 = \bar{x}), \mathbb{P}(\bar{X}_t \in \cdot \mid \bar{X}_0 = \bar{y})\right) \leq e^{-\sigma t}\, d(\bar{x},\bar{y}),\ \forall \bar{x},\bar{y} \in E^N,$$

where $\mathscr{W}_d$ denotes the Wasserstein distance. All the results of this chapter originally appeared in [260].

A lower bound on σ provides a measure of the instantaneous convergence rate to a unique stationary distribution (see for instance [68]). This concept is closely related to the optimal coupling theory developed by Chen (see for instance [68, 70]). It also entails spectral gap inequalities and concentration inequalities (see [217, 162, 160, 161, 256, 257, 100]). We refer the reader to [3] for a different approach, based on Kantorovich potentials. We also refer the reader to [51, Section 3.2] for a link between coarse Ricci curvature and functional inequalities. For general state space processes and for diffusion processes, we refer the reader to the works of Veysseire, where a systematic study of the coarse Ricci curvature has been conducted (see [256, 257]) with nice implications on concentration inequalities and spectral gap estimates. Let us also mention that estimates on the coarse Ricci curvature of a continuous time process immediately provide estimates for the curvature of its discrete time included Markov chain, which also implies several interesting properties (see the works of Ollivier [217, 218] and references therein).

Reconstituting transport distance bounds on Markov chains on product spaces from the behaviour of marginals via suitable couplings was already used by Talagrand and Marton, see for instance [199, 200] and references therein. These methods also apply to Markov processes that are not of pure jump types and to cost functions d that are not distance functions. For diffusion processes, we refer the reader to [71] and to [261, Corollary 1.4] for necessary and sufficient conditions in the case where the drift derives from a potential. We also refer the reader to [108, 109] with an introduction to parallel coupling and the construction of *ad hoc* distances on the state space. Computation of the coarse Ricci curvature for diffusion processes on manifold has also been studied by Veysseire [257]. For piecewise deterministic processes, we refer the reader to [75, Lemma 5.2] and [52, Theorem 2.3]. Original coupling approaches are also provided in [193, 192, 45].

Below, we state our main result in Section 10.1, provide applications to the case $N = 1$ in Section 10.2, and conclude with an application to a simple model of agents in interaction in Section 10.3. We refer the reader to the original paper [260] for the proofs, additional details and references. Therein, we also compute a lower bound for the coarse Ricci curvature of other interacting particle systems, including zero range dynamics, Fleming-Viot type systems and some of their natural extensions, birth and death processes in mean-field type interaction, and finally systems of particles whose jump measures admit a density with respect to the Lebesgue measure or the counting measure.

10.1 Main result

One of the difficulties of the continuous time setting, compared to the discrete time setting [217], is that the jump measures do not, in general, share the same mass. In order to overcome this difficulty, we introduce the family of functions $(J_d^{x,y})_{x,y\in E}$ from $\mathscr{M}^d(E)^2$ to $\mathbb{R}$, defined for all $m_1, m_2 \in \mathscr{M}^d(E)$ by

$$J_d^{x,y}(m_1, m_2) = \mathscr{W}_d(m_1 + m_2(E)\delta_x, m_2 + m_1(E)\delta_y) - (m_1(E) + m_2(E))d(x, y),$$

where δ_x denotes the Dirac measure at point x and $m_2(E)\delta_x$ is the product of the scalar $m_2(E)$ by δ_x. Note that the finite measures m_1 and m_2 can have different masses. Proper generalisations of the Wasserstein distance between measures with different masses already exist in the literature (such as the flat metric [103] and the generalised $W_1^{1,1}$ Wasserstein distance [223], see also the recent developments in [73, 176, 182] with applications to convergence of measure valued dynamical systems), but are not directly relevant in our context.

Theorem 10.1. *Consider the Markov process $\bar{X}$ with generator $\mathscr{L}$. Then there exists a coupling operator $\mathscr{L}^c$ of $\mathscr{L}$ such that, for all $\bar{x}, \bar{y} \in E^N$,*

$$\mathscr{L}^c d(\bar{x}, \bar{y}) = \frac{1}{N}\sum_{i=1}^{N} J_d^{x_i,y_i}(F_i(x_i, \bar{x}, \cdot), F_i(y_i, \bar{y}, \cdot)).$$

In particular, the coarse Ricci curvature σ of the process $(\bar{X}_t)_{t\geq 0}$ satisfies

$$\sigma \geq - \sup_{\bar{x},\bar{y}\in E^N} \frac{\frac{1}{N}\sum_{i=1}^{N} J_d^{x_i,y_i}(F_i(x_i, \bar{x}, \cdot), F_i(y_i, \bar{y}, \cdot))}{d(\bar{x}, \bar{y})}.$$

One main feature of this result is that one does not need to build an explicit coupling between processes to conclude. This is particularly useful for involved jump matrices.

Fix $x, y \in E$. We now provide some interesting properties of the functional $J_d^{x,y}$, which are also useful to derive upper bounds and hence to apply Theorem 10.1.

Proposition 10.2. *For all $m_1, n_1, m_2, n_2 \in \mathcal{M}^d(E)$ and all $\alpha > 0$, we have*

$$J_d^{x,y}(\alpha m_1, \alpha m_2) = \alpha J_d^{x,y}(m_1, m_2) \tag{10.1}$$

and

$$J_d^{x,y}(m_1 + n_1, m_2 + n_2) \leq J_d^{x,y}(m_1, m_2) + J_d^{x,y}(n_1, n_2). \tag{10.2}$$

The following inequality is in general a crude estimate, but it is in some cases useful and sharp (such as for one dimensional birth and death processes, see Example 10.2).

Proposition 10.3. *We have, for all $m_1, m_2 \in \mathcal{M}^d(E)$,*

$$J_d^{x,y}(m_1, m_2) \leq \int_E [d(u, y) - d(x, y)]\, m_1(du) + \int_E [d(x, v) - d(x, y)]\, m_2(dv).$$

The following property implies in particular that, if m_1 and m_2 are two probability measures, then $J_d^{x,y}(m_1, m_2)$ is smaller than $\mathcal{W}_d(m_1, m_2) - d(x, y)$. It also implies that, for measures m_1 and m_2 on E such that $m_1(E) \geq m_2(E)$, then $J_d^{x,y}(m_1, m_2) \leq \mathcal{W}_d(m_1, m_2 + (m_1(E) - m_2(E))\delta_y) - m_1(E)d(x, y)$.

Proposition 10.4. *We have, for all $m_1, m_2 \in \mathcal{M}^d(E)$,*

$$J_d^{x,y}(m_1, m_2) = \min_{a,b} \mathcal{W}_d(m_1 + a\delta_x, m_2 + b\delta_y) - (m_1(E) + a)\, d(x, y),$$

where a, b are taken in the set of real numbers such that $m_1 + a\delta_x$ and $m_2 + b\delta_y$ are non-negative measures on E with equal mass, i.e. such that $m_1(E) + a \geq 0$, $m_2(E) + b \geq 0$ and $m_1(E) + a = m_2(E) + b$. In addition, the minimum is attained for all $a \geq m_2(E)$ (or equivalently $b \geq m_1(E)$).

10.2 First applications in the particular case $N = 1$

In this section, we state our result in the simpler case $N = 1$. The following corollary is an immediate consequence of Theorem 10.1.

Corollary 10.5. *Let L be the infinitesimal generator of a pure jump non-explosive Markov process on E defined, for any bounded measurable function $f : E \to \mathbb{R}$, by*

$$Lf(x) = \int_E (f(u) - f(x))\, q(x, du), \ \forall x \in E,$$

where $(q(x, du))_{x \in E}$ is a jump kernel of finite non-negative measures. Then the coarse Ricci curvature σ of the Markov process generated by L satisfies

$$\sigma \geq -\sup_{x,y \in E} \frac{J_d^{x,y}\left(q(x,\cdot), q(y,\cdot)\right)}{d(x, y)}.$$

For continuous time birth and death processes, Mielke [208] recently computed a lower bound for an other notion of discrete Ricci curvature, related to the fact that the evolution of the law of a continuous time birth and death process can be described through a gradient flow system. To relate both definitions is still an open problem, but the lower bound obtained in Mielke's work has a similar expression (see Section 5 in [208] and Example 10.2 below) and may be a good starting point to compare both approaches. This example has also been considered by Fathi and Maas in [116, Theorem 4.1] in the setting of Entropic Ricci curvature.

Example 10.1. In this example, d is the trivial distance on E (so that $\mathcal{W}_d$ is the total variation distance). Assume that there exist a non-negative measure ζ on E and a measurable function $\alpha : E\times E \to \mathbb{R}_+$ such that

$$q(x,dz)=\alpha(x,z)\zeta(dz),\ \forall x\in E.$$

Then one carefully checks that

$$J_d^{x,y}(q(x,\cdot),q(y,\cdot)) = -\int_E \alpha(x,z)\wedge\alpha(y,z)\zeta(dz) - \alpha(y,x)\zeta(\{x\}) - \alpha(x,y)\zeta(\{y\}).$$

In particular, the coarse Ricci curvature σ of the process satisfies

$$\sigma \geq \inf_{x\neq y}\left[\int_E \alpha(x,z)\wedge\alpha(y,z)\zeta(dz) + \alpha(y,x)\zeta(\{x\}) + \alpha(x,y)\zeta(\{y\})\right].$$

Example 10.2. Consider the particular case where $E=\mathbb{Z}_+$ and L is the infinitesimal generator of a birth and death process with birth rates $(b_x)_{x\in\mathbb{Z}_+}$ and death rates $(d_x)_{x\in\mathbb{Z}_+}$, all positive but $d_0=0$. In this case, for all $x,y\in\mathbb{Z}_+$,

$$q(x,y)=\begin{cases} b_x & \text{if } y=x+1\\ d_x & \text{if } x\geq 1 \text{ and } y=x-1\\ 0 & \text{otherwise}\end{cases}$$

We also assume that the distance d is given by $d(x,y)=\left|\sum_{k=0}^{x-1}u_k-\sum_{k=0}^{y-1}u_k\right|$, where $(u_k)_{k\geq 0}$ is a sequence of positive numbers. After careful computations, we deduce from Proposition 10.3 and Corollary 10.5 that the coarse Ricci curvature σ of the process satisfies

$$\sigma \geq \inf_{x\in\mathbb{Z}_+} b_x + d_{x+1} - d_x\frac{u_{x-1}}{u_x} - b_{x+1}\frac{u_{x+1}}{u_x}.$$

In [68], [161] and [51], it is shown that there is equality in the above equation. This implies that, at least in some cases, Corollary 10.5 and hence Theorem 10.1 are sharp. Note that, in this case, Proposition 10.3 provides an explicit expression for the quantity $J_d^{x,y}(q(x,\cdot),q(y,\cdot))$.

Example 10.3. The choice of the classical coupling (i.e. the use of Proposition 10.3) in the previous example was judicious because the measures involved for a birth and death process are stochastically ordered. This is not the case in the present example, where we assume that

$$q(x,y)=\begin{cases} b_x & \text{if } y=x+2\\ d_x & \text{if } y=x-1,\\ 0 & \text{otherwise.}\end{cases}$$

In this case, using a slight extension of [244], one obtain that the coarse Ricci curvature of the process satisfies

$$\sigma \geq \inf_{x\in\mathbb{Z}_+} b_x + d_{x+1} - d_x \frac{u_{x-1}}{u_x} - |b_{x+1} - b_x| \frac{u_{x+1}}{u_x} - b_{x+1} \frac{u_{x+2}}{u_x}.$$

10.3 A model of interacting agents

We study now a simple model of interacting agents whose individual behaviour is influenced in a non-linear way by the behaviour of the other agents: each agent wanders randomly in a complete graph and also changes its position to a new one, depending on a function of the number of agents in this position. This dynamic is modelled by a system of N particles evolving in the complete finite graph E of size $\#E \geq 2$: we assume that there exist $T > 0$ and a function $f : [0,1] \to \mathbb{R}_+$ such that any agent jumps from state x to $y \in E$ with the following rate

$$x \to y \text{ with rate } \frac{T}{\#E} + f\left(\frac{\text{Number of agents in y}}{N}\right). \tag{10.3}$$

In this model, T is the temperature of the system and f is a preference function. For instance, with an increasing function f with high convexity, the agents will give higher preferences to positions that are already favoured by many other agents; with a larger temperature T, the agents act more independently. Our aim is to determine characteristics of f and values of T for which a herd behaviour occurs or not in this model. By a herd behaviour, we mean a meta-stable state of the whole particle system where a majority of the agents share the same position for a long time. Note that this model can be written in the settings of the present paper, by setting, for all $x, y \in E$ and $\bar{x} \in E^N$,

$$F_i(x, \bar{x}, \{y\}) = \frac{T}{\#E} + f\left(\frac{\sum_{i=1}^N \mathbf{1}_{x_i = y}}{N}\right), \ \forall y \in E.$$

This process is exponentially ergodic and the marginal of its empirical stationary distribution is the uniform probability measure on E (this is an immediate consequence of the symmetry of the state space and of the dynamic of the particles). The existence of the phase without herd behaviour is obtained using the results of Section **??**, while the existence of the phase with herd behaviour is proved using large deviation results obtained in [105, 106]. Since the publication of the original article, Erbar, Fathi and Schlichting [115] have studied this model (among others) in the settings of Entropic Ricci curvature.

In this first proposition, we assume that f is Lipschitz and provide a coarse Ricci curvature's lower bound independent of N.

Proposition 10.6. *Assume that f is a Lipschitz function and define the Lipschitz constant of f as $\|f\|_{Lip} = \sup_{u\neq v\in[0,1]} |f(u) - f(v)|/|u - v|$. Then the coarse Ricci curvature σ of the particle system described above satisfies*

$$\sigma \geq T - 2\|f\|_{Lip} + \inf_{\mu} \sum_{x\in E} f(\mu(x)),$$

where the infimum is taken over the probability measures μ on E. Moreover, if f is monotone, then

$$\sigma \geq T - \|f\|_{Lip} + \inf_{\mu} \sum_{x\in E} f(\mu(x)).$$

In the next proposition, we assume that f is a non-decreasing strictly convex function and show that, for small values of T, the process exhibits a meta-stable state, so that the agents have a herd behaviour for large values of N : if all the agents start with the same choice $x \in E$, then, during a time of order $\exp(cN)$, where $c > 0$, x is favoured by the majority of the agents. This happens despite the fact that, during this interval of time, most agents have changed their choices at multiple times.

Proposition 10.7. *Assume that f is a strictly convex function such that $f(0) = 0$, let $z_* \in (1/2, 1)$ such that*

$$z_* = argmax_{z\in[1/2,1]} f(z) - z(f(z) + f(1-z))$$

and set

$$m_* = f(z_*) - z_*(f(z_*) + f(1-z_*)) > 0.$$

If the temperature is sufficiently small, namely if $0 \le T < \frac{m_* \#E}{z_* \#E - 1}$, then there exists a positive constant $\delta > 0$ such that, for all $x \in E$,

$$\frac{1}{N}\log\mathbb{P}\left(\exists s\in[0,t],\ \mu_s^N(x)\le z^*\right) =_{N\to+\infty} \mathcal{O}\left(\min\left(\delta(\mu_0^N(x)-z_*)_+^2, \delta\bar{\varepsilon}^2 - \frac{\log t}{N}\right)\right),$$

uniformly in $t \ge 0$ and where $\mu_s^N = \frac{1}{N}\sum_{i=1}^N \delta_{X_s^i}$.

In order to check that $m_* > 0$ in the above result, one simply uses the fact that f is strictly convex with $f(0) = 0$, so that, for all $z \in (1/2, 1)$, $f(1-z)/(1-z) < f(z)/z$.

Example 10.4. Assume that f is an affine function : $f(x) = ax + b$ for some $a \in \mathbb{R}$ and $b \ge 0$ such that $a + b \ge 0$. Then f is Lipschitz with $\|f\|_{Lip} = |a|$ and $\sum_{x\in E} f(\mu(x)) = a + b\#E$ for any probability measure μ on E. Hence Proposition 10.6 implies that the Wasserstein curvature of the process is bounded from below by $T + b\#E + a - |a|$. In particular, it is positive since

$$T + b\#E + a - |a| = \begin{cases} T + b\#E > 0 & \text{if } a \ge 0,\\ T + b(\#E-2) + 2(b+a) > 0 & \text{if } a < 0,\end{cases}$$

and hence the system of agents does not exhibit a herd behaviour.

Example 10.5. Assume that $f(x) = x^2$. Then $\|f\|_{Lip} = 2$ and

$$\inf_{\mu}\sum_{x\in E} f(\mu(x)) = \frac{1}{\#E},$$

Moreover,

$$z_* = \text{argmax}_{z\in[1/2,1]} z^2 - z(z^2 + (1-z)^2) = \frac{1}{2} + \frac{1}{\sqrt{12}}$$

and

$$m_* = z_*^2 - z_*(z_*^2 + (1-z_*)^2) = \frac{1}{6\sqrt{3}}.$$

Hence we deduce from Proposition 10.6 and Proposition 10.7 that

- if $T > 2 - 1/\#E$, then the Wasserstein curvature of the particle system is positive (bounded from below by $T - 2 + 1/\#E$) and the system of agents does not exhibits a herd behaviour;
- if $0 \le T < \#E/((3+3\sqrt{3})\#E - 6\sqrt{3})$, then the system of agents exhibits a herd behaviour.

Chapter 11

Maintenance of biodiversity and perpetual integrals

Often, demogenetics model are obtained from a specific scaling of the parameters in the individual-based model, leading to a stochastic differential equation with a diffusion term proportional to the square root of the size of the population (Feller type diffusion processes). Other scaling will lead to different coefficients and we refer the reader to [82, 22, 83] for an in-depth discussion of such models.

Our aim here is to emphasise the importance of this diffusion term in fixation problems. Our main question is whether, in a given demogenetic model, one allele gets fixed almost surely before the population goes extinct. In a collaboration with Camille Coron and Sylvie Méléard [85], we prove that this is the case almost surely for Feller type diffusion coefficients. We also show that, in fact, it depends on the behaviour of the diffusion coefficient near extinction in the equation satisfied by the population size, as detailed below. The next theorem notably highlights the major effect of the demography on the maintenance of genetic diversity by giving a necessary and sufficient criterion ensuring almost sure fixation before extinction. The main tool of the proof has its own interest, since it derives from finiteness criteria for *perpetual integrals*, which we detail at the end of this chapter.

11.1 Demography and maintenance of biodiversity

Let us consider the process $(N_t, X_t)_{t\in\mathbb{R}_+}$ solution to the system of stochastic differential equations

$$\begin{cases} dN_t &= \sigma(N_t)\,dB_t + N_t(\rho-\alpha N_t)dt,\ N_0>0, \alpha>0 \\ dX_t &= \sqrt{\frac{X_t(1-X_t)}{f(N_t)}}\,dW_t \end{cases},\quad t<T^N_{0+}, \tag{11.1}$$

where B, W are independent one-dimensional Brownian motions, $\sigma:(0,+\infty)\to(0,+\infty)$ is locally Lipschitz and $f:(0,+\infty)\to(0,+\infty)$ is locally bounded away from 0 and where

$$T^N_{0+} := \lim_{n\to+\infty} T^N_{1/n}$$

denotes the extinction time of the population. The system admits a pathwise unique strong solution up to the extinction time and we denote by

$$T^f := \inf\{t\geq 0,\ X_t\in\{0,1\}\}$$

the fixation time of the process.

Theorem 11.1. *Fixation occurs before extinction with probability one if and only if*

$$\int_{0+} \frac{y}{\sigma^2(y) f(y)}\, dy = +\infty. \tag{11.2}$$

Consider the particular case where f is the identity function. Whereas for the usual demographic term $\sigma(N) = \sqrt{N}$, fixation occurs almost surely before extinction, a small perturbation of this diffusion term, taking for example $\sigma(N) = N^{(1-\varepsilon)/2}$, $\varepsilon > 0$, leads to extinction before fixation with positive probability. An example of trajectory for which fixation does not occur before extinction is given in Figure 11.1, and the effect of ε on the probability of extinction before fixation is numerically studied in Figure 11.2.

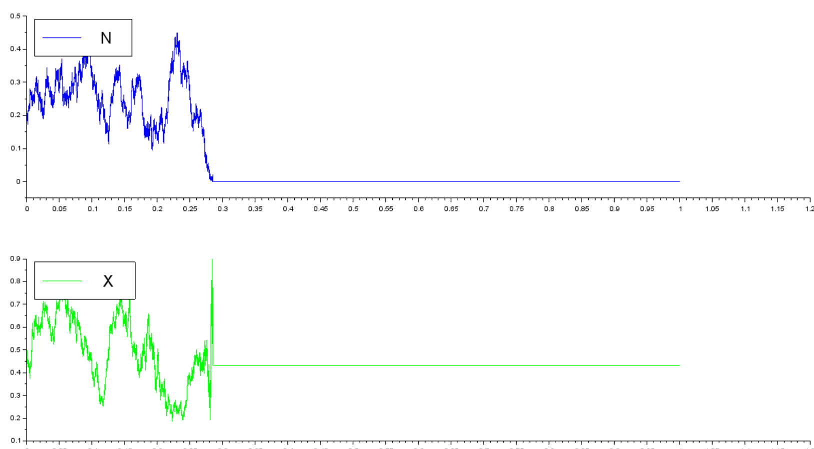

Figure 11.1: We plot a trajectory of the 2-dimensional diffusion process (N, X) such that $dN_t = \sqrt{N_t^{(1-\varepsilon)}}dB_t + N_t(\rho - \alpha N_t)dt$ and $dX_t = \sqrt{\frac{X_t(1-X_t)}{N_t}}dW_t$, with $\varepsilon = 0.4$, $\rho = -1$ and $\alpha = 0.1$. For this trajectory, fixation does not occur before extinction.

11.2 Integrability properties for diffusion processes

We state a result implying that, depending on the behaviour of the diffusion and drift coefficients near absorption, the integral of the paths of diffusion processes are either almost surely finite or almost surely infinite. This 0-1 law criterion has already been proved by various methods, using a combination of the local time formula and Ray-Knight theorem [114, 209, 168] (see also [113, 126] for proofs in particular settings). In [85], we give a simpler proof of this criterion, which also provides explicit bounds for the moments of perpetual integrals and can be easily extended to more general one dimensional Markov processes.

11.2.1 General diffusion processes on $[0, +\infty)$

Let us consider a general one-dimensional diffusion process $(Z_t)_{t\in\mathbb{R}_+}$ (see Chapter 4) with values in $(0, +\infty)$. Let us denote by $\mathbb{P}_z$ the law of Z starting from z. We assume that Z is regular ($\forall z \in$

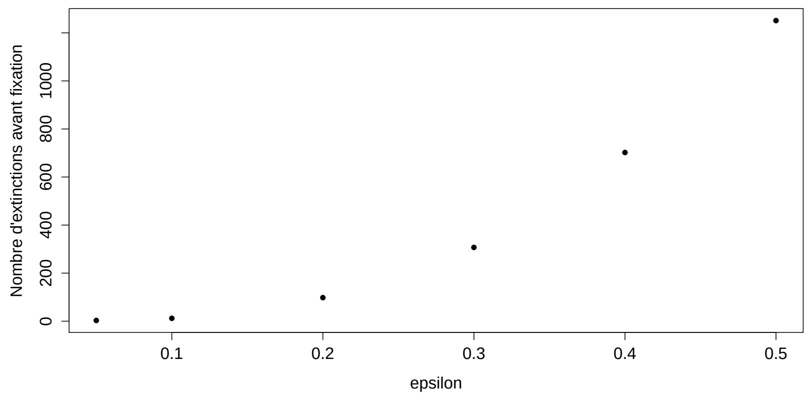

Figure 11.2: For different values of ε, we simulate 10000 trajectories of the 2-dimensional diffusion process (N, X) such that $dN_t = \sqrt{N_t^{(1-\varepsilon)}}dB_t^1 + N_t(r - cN_t)dt$ and $dX_t = \sqrt{\frac{X_t(1-X_t)}{N_t}}$, with $r = -1$ and $c = 0.1$. We plot the number of simulations for which fixation does not occur before extinction.

$(0,+\infty), \forall y \in (0,+\infty)$, $\mathbb{P}_z(T_y < +\infty) > 0)$. This implies that for any $a < b \in (0,+\infty)$ and $a \le z \le b$, $\mathbb{E}_z(T_a \wedge T_b) < +\infty$ and we can associate with Z a scale function s and a locally finite speed measure m on $[0,+\infty)$ (see Chapter 4). We moreover assume that for all $z \in (0,+\infty)$,

$$\mathbb{P}_z(T_0 = T_0 \wedge T_e < +\infty) = 1, \tag{11.3}$$

where T_e is the explosion time (this is equivalent to $s(+\infty) = +\infty$, $s(0) > -\infty$ and $\int_{0+}(s(y)-s(0))\, m(dy) < +\infty$). Since the function s is defined up to a constant, we choose by convention $s(0) = 0$ as soon as $s(0) > -\infty$.

Theorem 11.2. *Let $(Z_t)_{t\in\mathbb{R}_+}$ be a regular diffusion process on $[0,+\infty)$ with scale function s and speed measure m on $(0,+\infty)$ satisfying (11.3). Let also f be a non-negative locally integrable function on $(0,+\infty)$. Then, for all $z > 0$ and all $n \ge 1$,*

$$\mathbb{E}_z\left[\left(\int_0^{T_0} f(Z_s)\, ds\right)^n\right] \le n!\left(\int_0^{\infty} s(y)\, f(y)\, m(dy)\right)^n$$

and

$$\int_{0^+} s(y)\, f(y)\, m(dy) < +\infty \iff \int_0^{T_0} f(Z_s)\, ds < +\infty \quad \mathbb{P}_z - \text{almost surely}$$
$$\int_{0^+} s(y)\, f(y)\, m(dy) = +\infty \iff \int_0^{T_0} f(Z_s)\, ds = +\infty \quad \mathbb{P}_z - \text{almost surely.}$$

Let us give two examples for population size processes.

Example 11.1 (Branching process with immigration). Let us consider the solution of the stochastic differential equation $dN_t = \sigma\sqrt{N_t}dB_t + \beta dt$, $\beta > 0$. Computing s and m as in (4.4) of Chapter 4, we easily obtain that (11.3) $\iff \beta/\sigma^2 < 1/2$. Hence

$$\int_0^{T_0} \frac{1}{(N_s)^\alpha}\, ds = +\infty \quad a.s. \iff \alpha \ge 1\,;\ \int_0^{T_0} \frac{1}{(N_s)^\alpha}\, ds < +\infty \quad a.s. \iff \alpha < 1. \tag{11.4}$$

In the particular case $\alpha = 1$, the authors of [126] propose an other approach based on self-similarity properties.

Example 11.2. Logistic diffusion process. Let us consider the process

$$dN_t = \sqrt{N_t}\,dB_t + N_t\,(b - c\,N_t)\,dt\,;\ N_0 > 0,$$

where $c > 0$. Then $s(y) = \int_0^y e^{cz^2 - 2bz}dz$ and $m(dy) = \dfrac{2e^{-cy^2+2by}}{y}dy$ and $\int_{0^+} s(y)m(dy) < +\infty$, since $\frac{s(y)}{s'(y)\,y} \to_{y\to 0} 1$. (Note that if $c = 0$, the condition $s(+\infty) = +\infty$ is not satisfied). It is immediate to check that (11.4) also holds.

Chapter 12

The individual's signature of telomere length distribution.

In a recent collaboration with Éliane Albuisson (CHRU of Nancy and IECL), Athanase Benetos (CHRU of Nancy), Anne Gégout-Petit (IECL), Daphné Germain (former student at École des Mines de Nancy) and Simon Toupance (CHRU of Nancy), we studied the evolution of telomere length distribution over time in adults. This statistical study was published in [242].

12.1 A short introduction to telomere lengths

Telomeres are specialised non-coding double-stranded repetitive DNA-protein complexes that form protective caps on the ends of chromosomes. They safeguard their extremity and maintain genomic integrity by allowing cells to distinguish telomeres from sites of DNA damage[133, 88]. Telomere length displays progressive shortening in replicating somatic cells with age[183, 136]. Eventually cells will acquire critically short and dysfunctional telomeres that, consequently, activate a DNA damage response and growth arrest known as replicative senescence[237, 97]. Therefore, all somatic cells have limited cell proliferation capacity called the Hayflick limit[140, 219].

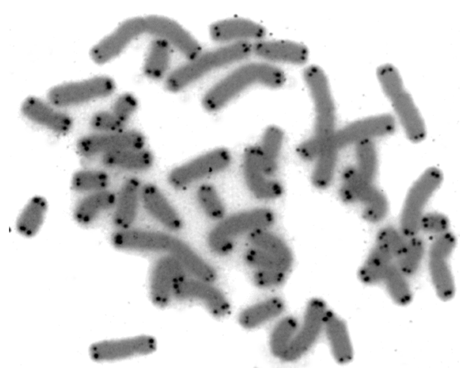

Figure 12.1: Human chromosomes (grey) capped by telomeres (black). Wikipedia.

Short leukocyte telomere length is associated with many degenerative diseases linked to ageing and with higher mortality risk. Epidemiological studies use leukocyte telomere length to examine the potential role of telomere length in health and disease. It is known that leukocyte telomere length decreases with age and thus is considered as a biomarker of chronological ageing [233, 38, 16]. In humans, a long leukocyte telomere length is associated with better survival in the elderly [17, 173, 124, 91] and a recent meta-analysis has indicated a strong relation between short telomeres and mortality risk, particularly at younger ages [41]. A shorter leukocyte telomere length is associated with many degenerative diseases linked to ageing such as cardiovascular disease [139, 101], neurodegenerative disease [43, 125] and metabolic diseases [101, 18].

Several methods have been developed to measure the length of telomere repeats from cells or extracted DNA [13]. The most used are: Southern blot analysis of the terminal restriction fragments length (TRF); quantitative PCR (qPCR) amplification of telomeric DNA, expressed as the ratio of telomere repeats relative to a single copy gene; single telomere length analysis (STELA), a PCR and Southern blot combining method that measures telomere lengths from individual chromosomes; and fluorescent in situ hybridisation (FISH) techniques, quantitative FISH (qFISH) based on microscopy and flow FISH using flow cytometry. There are ongoing debates as to which method is the most suitable to measure telomere's length in clinical studies, particularly between TRF and qPCR [111, 195, 255]. The advantages of the qPCR method are the high throughput and low cost but TRF is considered the "gold standard" since it displays less variability [14], gives absolute values of telomere length and gives access to telomere length distribution [15]. However, the vast majority of studies using TRF measurements only use mean telomere length and not distribution as a result. Assessing telomere length distribution with TRF measurement can give access to new information concerning telomere length dynamics since a cell's shortest telomeres and the "load" (amount) of short telomeres appear to play a role independently of mean telomere length [29, 145]. Mean telomere attrition rates do not capture changes in telomere distribution that may play a role in pathology development (see Figure 12.2).

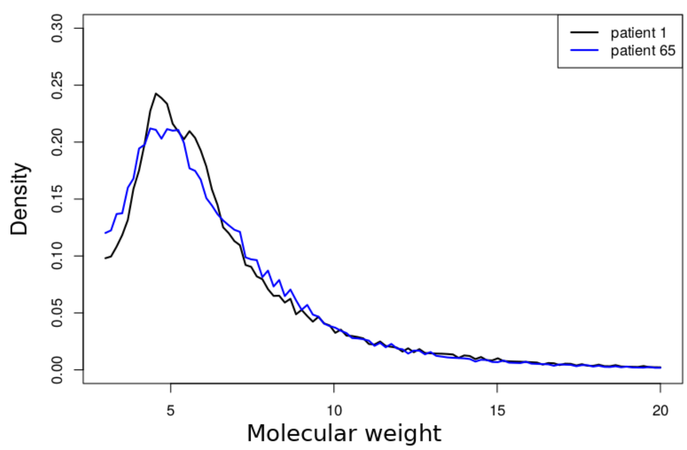

Figure 12.2: Telomere length distribution of two patients. The mean of the two distributions is the same, although the distributions differs.

The aim of the study described in this chapter was to analyse the distribution of telomere lengths and its evolution in time, using data generated by TRF in a longitudinal study in which two sequential measurements of telomeres were performed at the beginning and the end of the study corresponding to a mean time distance of 8 years.

12.2 Telomere length signature

A quick look at the datas at the beginning and at the end of the study, suggests that the telomere length distribution conserves the same shape at base line and at follow up. To study the conservation of the shape, the leukocyte telomere length distribution have then been translated in order to have the same median, and then drowned (see Figure 12.3).

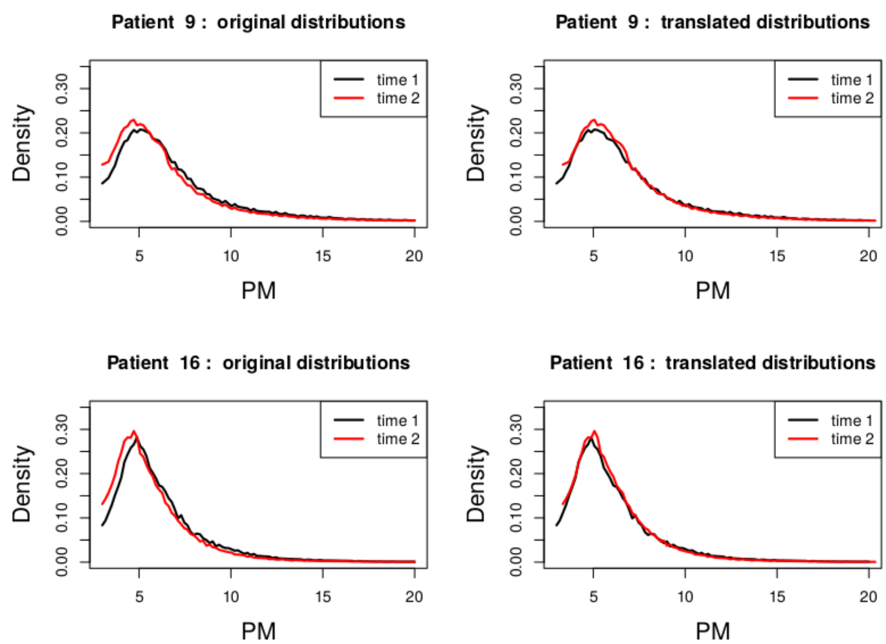

Figure 12.3: Telomere length distribution of two patients and their translation.

For most patients (although not systematically), the translated distributions at time 1 and at time 2 coincide very well.

In order to measure numerically the distance between distributions, we used the Kolmogorov distance between distributions, defined as the infinite norm distance between their cumulative functions.

To show how the shape of the distributions of one Subject is well conserved between the two times, we have computed the Kolmogorov distances between successive distributions for each of the 72 Subjects. The distribution of these 72 distances is given in the normalised histogram in red of the Figure 12.4. We have also computed the 72×71 Kolmogorov distances between the translated telomere length distribution at time 1 for one Subject and time 2 for another one. The normalised histogram is given in blue at Figure 12.4. We see clearly that intra-subject distances seem to be lower than the inter-subjects distances.

We have performed a T-test to confirm the tendency: it strongly rejects the equality between the observed mean of the intra-Subject distances ($mean_1 = 0.0258$) and the mean of the inter-Subject distances that we consider like a theoretical expectancy ($mean_2 = 0.0639$). We can say that the shapes of the leukocyte telomere length distributions of one Subject are significantly closer than two shapes of two different Subjects.

Our conclusion is that leukocyte telomere length distribution characterises an individual and

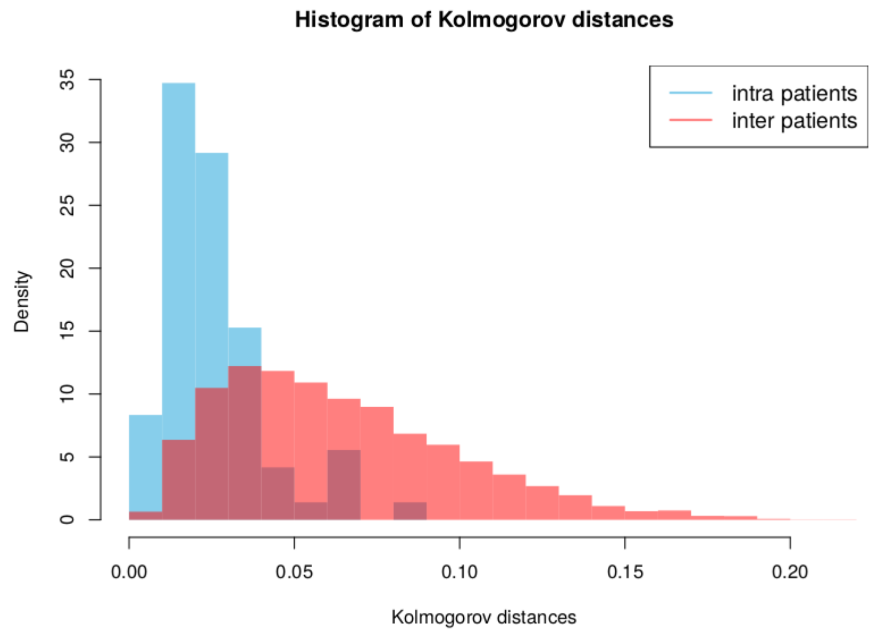

Figure 12.4: Histogram of Kolmogorov distances

we introduce the concept that telomere length distribution represents an individual signature in humans that remains stable over the adult life.

12.3 Discussion

The contribution of this study is to show a strong stability over time not only of ranking but also of the telomere length distribution. The telomere length signature could become a new criterion to describe patients since we have seen that two individuals can have the same mean leukocyte telomere length but different distributions. We also acknowledge limitations of this study. First, the sample size is modest with less than hundred Subjects. Second, our cohort comprised participants who were all above 60 years of age at the beginning of the study. Third, the follow-up duration was only 8 years and we can't conclude on variation on longer period. However, the clear results obtained on only 72 patients are strong enough to overcome difficulties and sources of errors linked to this type of studies: blood samples have been taken 8 years apart by different nurses and DNA extracted 8 years apart by different researchers. In the future, in clinical studies, maybe that telomere length signature could capture new associations of telomere dynamics with clinical parameters or disease markers or help to better clusterise patients.

In an ongoing collaboration, we are currently working on a larger cohort (around 500 individuals of all ages), which allowed us to confirm the finding of this study and to provide new insights on the evolution of telomere length distributions over time. Our next step will be to relate, when possible, the telomere length signature of the patients to their medical records.

Bibliography

[1] D. Aldous. Asymptotic fringe distributions for general families of random trees. *The Annals of Applied Probability*, pages 228–266, 1991.

[2] D. Aldous, B. Flannery, and J. L. Palacios. Two applications of urn processes the fringe analysis of search trees and the simulation of quasi-stationary distributions of markov chains. *Probability in the engineering and informational sciences*, 2(3):293–307, 1988.

[3] A. Alfonsi, J. Corbetta, and B. Jourdain. Evolution of the Wasserstein distance between the marginals of two Markov processes. *ArXiv e-prints*, June 2016. To appear in Bernoulli Journal.

[4] M. Amir. Sticky brownian motion as the strong limit of a sequence of random walks. *Stochastic Processes and their Applications*, 39(2):221 – 237, 1991.

[5] W. J. Anderson. *Continuous-time Markov chains: an applications-oriented approach*. Springer, 1991.

[6] E. Andjel, F. Ezanno, P. Groisman, and L. T. Rolla. Subcritical contact process seen from the edge: convergence to quasi-equilibrium. *Electron. J. Probab.*, 20:no. 32, 16, 2015.

[7] E. D. Andjel. Convergence in distribution for subcritical 2D oriented percolation seen from its rightmost point. *Ann. Probab.*, 42(3):1285–1296, 2014.

[8] E. Arjas, E. Nummelin, and R. L. Tweedie. Semi-markov processes on a general state space: α-theory and quasi-stationarity. *Journal of the Australian Mathematical Society. Series A. Pure Mathematics and Statistics*, 30(2):187–200, 1980.

[9] J. Artalejo and M. Lopez-Herrero. Quasi-stationary and ratio of expectations distributions: A comparative study. *Journal of theoretical biology*, 266(2):264–274, 2010.

[10] A. Asselah and P. A. Ferrari. Regularity of quasi-stationary measures for simple exclusion in dimension $d \geq 5$. *Ann. Probab.*, 30(4):1913–1932, 2002.

[11] K. B. Athreya and S. Karlin. Embedding of urn schemes into continuous time Markov branching processes and related limit theorems. *Ann. Math. Statist.*, 39:1801–1817, 1968.

[12] K. B. Athreya and P. E. Ney. *Branching processes*. Springer-Verlag, New York, 1972. Die Grundlehren der mathematischen Wissenschaften, Band 196.

[13] G. Aubert, M. Hills, and P. M. Lansdorp. Telomere length measurement—caveats and a critical assessment of the available technologies and tools. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 730(1-2):59–67, 2012.

[14] A. Aviv, S. C. Hunt, J. Lin, X. Cao, M. Kimura, and E. Blackburn. Impartial comparative analysis of measurement of leukocyte telomere length/dna content by southern blots and qpcr. *Nucleic acids research*, 39(20):e134–e134, 2011.

[15] D. M. Baird. New developments in telomere length analysis. *Experimental gerontology*, 40(5):363–368, 2005.

[16] D. M. Baird and D. Kipling. The extent and significance of telomere loss with age. *Annals of the New York Academy of Sciences*, 1019(1):265–268, 2004.

[17] S. L. Bakaysa, L. A. Mucci, P. E. Slagboom, D. I. Boomsma, G. E. McClearn, B. Johansson, and N. L. Pedersen. Telomere length predicts survival independent of genetic influences. *Aging cell*, 6(6):769–774, 2007.

[18] M. Balasubramanyam, A. Adaikalakoteswari, S. F. Monickaraj, and V. Mohan. Telomere shortening & metabolic/vascular diseases. *Indian Journal of Medical Research*, 125(3):441–450, 2007.

[19] A. Bandyopadhyay and D. Thacker. A New Approach to Pólya Urn Schemes and Its Infinite Color Generalization. *Arxiv:1606.05317*, June 2016.

[20] V. Bansaye, B. Cloez, and P. Gabriel. Ergodic behavior of non-conservative semigroups via generalized Doeblin's conditions. *arXiv e-prints*, page arXiv:1710.05584, Oct 2017.

[21] V. Bansaye, B. Cloez, P. Gabriel, and A. Marguet. A non-conservative Harris' ergodic theorem. *arXiv e-prints*, page arXiv:1903.03946, Mar 2019.

[22] V. Bansaye and S. Méléard. *Stochastic models for structured populations*, volume 1 of *Mathematical Biosciences Institute Lecture Series. Stochastics in Biological Systems.* Springer, Cham; MBI Mathematical Biosciences Institute, Ohio State University, Columbus, OH, 2015. Scaling limits and long time behavior.

[23] M. Baudel and N. Berglund. Spectral theory for random poincaré maps. *SIAM J. Appl. Math.*, 49(6):4319–4375, 2017.

[24] N. Bean, L. Bright, G. Latouche, C. Pearce, P. Pollett, and P. G. Taylor. The quasi-stationary behavior of quasi-birth-and-death processes. *The Annals of Applied Probability*, pages 134–155, 1997.

[25] M. Benaïm. Dynamics of stochastic approximation algorithms. In *Séminaire de Probabilités, XXXIII*, volume 1709 of *Lecture Notes in Math.*, pages 1–68. Springer, Berlin, 1999.

[26] M. Benaïm, N. Champagnat, and D. Villemonais. Stochastic approximation of quasi-stationary distributions for diffusion processes in a bounded domain. *ArXiv e-prints*, Apr. 2019.

[27] M. Benaïm and B. Cloez. A stochastic approximation approach to quasi-stationary distributions on finite spaces. *Electronic Communications in Probability*, 20, 2015.

[28] M. Benaim, B. Cloez, and F. Panloup. Stochastic approximation of quasi-stationary distributions on compact spaces and applications. *The Annals of Applied Probability*, 28(4):2370–2416, 2018.

[29] L. Bendix, P. B. Horn, U. B. Jensen, I. Rubelj, and S. Kolvraa. The load of short telomeres, estimated by a new method, universal stela, correlates with number of senescent cells. *Aging cell*, 9(3):383–397, 2010.

[30] F. Bergeron, P. Flajolet, and B. Salvy. Varieties of increasing trees. In *Colloquium on Trees in Algebra and Programming*, pages 24–48. Springer, 1992.

[31] N. Berglund and D. Landon. Mixed-mode oscillations and interspike interval statistics in the stochastic FitzHugh-Nagumo model. *Nonlinearity*, 25(8):2303–2335, 2012.

[32] J. Bertoin. On a Feynman-Kac approach to growth-fragmentation semigroups and their asymptotic behaviors. *arXiv e-prints*, page arXiv:1804.04905, Apr 2018.

[33] J. Bertoin and A. Watson. The strong Malthusian behavior of growth-fragmentation processes. *arXiv e-prints*, page arXiv:1901.07251, Jan 2019.

[34] J. Bertoin and A. R. Watson. A probabilistic approach to spectral analysis of growth-fragmentation equations. *Journal of Functional Analysis*, 274(8):2163 – 2204, 2018.

[35] A. Bianchi and A. Gaudillière. Metastable states, quasi-stationary distributions and soft measures. *Stochastic Process. Appl.*, 126(6):1622–1680, 2016.

[36] J. D. Biggins and A. E. Kyprianou. Measure change in multitype branching. *Adv. in Appl. Probab.*, 36(2):544–581, 2004.

[37] G. Birkhoff. Extensions of Jentzsch's theorem. *Trans. Amer. Math. Soc.*, 85:219–227, 1957.

[38] E. H. Blackburn, E. S. Epel, and J. Lin. Human telomere biology: a contributory and interactive factor in aging, disease risks, and protection. *Science*, 350(6265):1193–1198, 2015.

[39] J. Blanchet, P. Glynn, and S. Zheng. Analysis of a stochastic approximation algorithm for computing quasi-stationary distributions. *Advances in Applied Probability*, 48(3):792–811, 2016.

[40] M. Bóna. k-protected vertices in binary search trees. *Advances in Applied Mathematics*, 53:1–11, 2014.

[41] J. J. Boonekamp, M. J. Simons, L. Hemerik, and S. Verhulst. Telomere length behaves as biomarker of somatic redundancy rather than biological age. *Aging cell*, 12(2):330–332, 2013.

[42] L. A. Breyer and G. O. Roberts. A quasi-ergodic theorem for evanescent processes. *Stochastic Process. Appl.*, 84(2):177–186, 1999.

[43] Z. Cai, L.-J. Yan, and A. Ratka. Telomere shortening and alzheimer's disease. *Neuromolecular Med*, 15(1):25–48, 2013.

[44] P. Cattiaux, P. Collet, A. Lambert, S. Martínez, S. Méléard, and J. San Martín. Quasi-stationary distributions and diffusion models in population dynamics. *Ann. Probab.*, 37(5):1926–1969, 2009.

[45] P. Cattiaux and A. Guillin. Semi log-concave Markov diffusions. In *Séminaire de Probabilités XLVI*, volume 2123 of *Lecture Notes in Math.*, pages 231–292. Springer, Cham, 2014.

[46] P. Cattiaux and S. Méléard. Competitive or weak cooperative stochastic Lotka-Volterra systems conditioned to non-extinction. *J. Math. Biol.*, 60(6):797–829, 2010.

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

[48] R. Cerf and J. Dalmau. A Markov chain representation of the normalized Perron-Frobenius eigenvector. *Electron. Commun. Probab.*, 22:Paper No. 52, 6, 2017.

[49] R. Cerf and J. Dalmau. A probabilistic proof of Perron's theorem. *arXiv e-prints*, page arXiv:1801.05252, Jan 2018.

[50] U. Cetin. An introduction to markov processes and their applications in mathematical economics. Lecture notes, 2014.

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

[52] D. Chafaï, F. Malrieu, and K. Paroux. On the long time behavior of the TCP window size process. *Stochastic Process. Appl.*, 120(8):1518–1534, 2010.

[53] N. Champagnat, K. A. Coulibaly-Pasquier, and D. Villemonais. Criteria for exponential convergence to quasi-stationary distributions and applications to multi-dimensional diffusions. *Séminaire de Probabilités*, XLIX:165–182, 2018.

[54] N. Champagnat, P. Diaconis, and L. Miclo. On Dirichlet eigenvectors for neutral two-dimensional Markov chains. *Electron. J. Probab.*, 17:no. 63, 41, 2012.

[55] N. Champagnat and S. Rœlly. Limit theorems for conditioned multitype Dawson-Watanabe processes and Feller diffusions. *Electron. J. Probab.*, 13:no. 25, 777–810, 2008.

[56] N. Champagnat and D. Villemonais. Exponential convergence to quasi-stationary distribution and Q-process. *Probab. Theory Related Fields*, 164(1):243–283, 2016.

[57] N. Champagnat and D. Villemonais. Exponential convergence to quasi-stationary distribution for absorbed one-dimensional diffusions with killing. *ALEA Lat. Am. J. Probab. Math. Stat.*, 14(1):177–199, 2017.

[58] N. Champagnat and D. Villemonais. General criteria for the study of quasi-stationarity. *ArXiv e-prints*, Dec. 2017.

[59] N. Champagnat and D. Villemonais. Lyapunov criteria for uniform convergence of conditional distributions of absorbed Markov processes. *ArXiv e-prints*, Apr. 2017.

[60] N. Champagnat and D. Villemonais. Uniform convergence of conditional distributions for absorbed one-dimensional diffusions. *Adv. in Appl. Probab.*, 50(1):178–203, 2017.

[61] N. Champagnat and D. Villemonais. Uniform convergence to the Q-process. *Electron. Commun. Probab.*, 22:7 pp., 2017.

[62] N. Champagnat and D. Villemonais. Practical criteria for R-positive recurrence of unbounded semigroups. *ArXiv e-prints*, Apr. 2019.

[63] Champagnat, Nicolas and Villemonais, Denis. Uniform convergence of penalized time-inhomogeneous markov processes. *ESAIM Probab. Stat.*, 22:129–162, 2018.

[64] J.-R. Chazottes, P. Collet, and S. Méléard. On time scales and quasi-stationary distributions for multitype birth-and-death processes. *ArXiv e-prints*, Feb. 2017.

[65] J. Chen and S. Jian. Some limit theorems of killed Brownian motion. *Sci. China Math.*, 56(3):497–514, 2013.

[66] J. Chen and S. Jian. Quasi-ergodicity for absorbing Markov processes *via* deviation inequality. *ESAIM Probab. Stat.*, 21:159–167, 2017.

[67] J. Chen and S. Jian. A deviation inequality and quasi-ergodicity for absorbing Markov processes. *Ann. Mat. Pura Appl. (4)*, 197(2):641–650, 2018.

[68] M.-F. Chen. Optimal Markovian couplings and applications. *Acta Math. Sinica (N.S.)*, 10(3):260–275, 1994. A Chinese summary appears in Acta Math. Sinica **3**8 (1995), no. 4, 575.

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

[70] M.-F. Chen. *Eigenvalues, inequalities, and ergodic theory.* Probability and its Applications (New York). Springer-Verlag London, Ltd., London, 2005.

[71] M.-F. Chen and S. F. Li. Coupling methods for multidimensional diffusion processes. *Ann. Probab.*, 17(1):151–177, 1989.

[72] G.-S. Cheon and L. W. Shapiro. Protected points in ordered trees. *Applied Mathematics Letters*, 21(5):516–520, 2008.

[73] L. Chizat, G. Peyré, B. Schmitzer, and F.-X. Vialard. Unbalanced optimal transport: dynamic and Kantorovich formulations. *J. Funct. Anal.*, 274(11):3090–3123, 2018.

[74] B. Cloez. Limit theorems for some branching measure-valued processes. *Advances in Applied Probability*, 49(2):549–580, 2017.

[75] B. Cloez and M. Hairer. Exponential ergodicity for Markov processes with random switching. *Bernoulli*, 21(1):505–536, 2015.

[76] H. Cohn. On a paper by Doeblin on nonhomogeneous Markov chains. *Adv. in Appl. Probab.*, 13(2):388–401, 1981.

[77] P. Collet, F. J. López, and S. Martínez. Order relations of measures when avoiding decreasing sets. *Statist. Probab. Lett.*, 65(3):165–175, 2003.

[78] P. Collet, S. Martínez, S. Méléard, and J. San Martín. Quasi-stationary distributions for structured birth and death processes with mutations. *Probability Theory and Related Fields*, 151:191–231, 2011. 10.1007/s00440-010-0297-4.

[79] P. Collet, S. Martínez, and J. San Martín. Asymptotic laws for one-dimensional diffusions conditioned to nonabsorption. *Ann. Probab.*, 23(3):1300–1314, 1995.

[80] P. Collet, S. Martínez, and J. San Martín. *Quasi-stationary distributions*. Probability and its Applications (New York). Springer, Heidelberg, 2013. Markov chains, diffusions and dynamical systems.

[81] P. Coolen-Schrijner and E. A. van Doorn. Quasi-stationary distributions for a class of discrete-time Markov chains. *Methodol. Comput. Appl. Probab.*, 8(4):449–465, 2006.

[82] C. Coron. *Stochastic modeling and eco-evolution of a diploid population*. PhD thesis, Palaiseau, Ecole polytechnique, 2013.

[83] C. Coron. Slow-fast stochastic diffusion dynamics and quasi-stationarity for diploid populations with varying size. *J. Math. Biol.*, 72(1-2):171–202, 2016.

[84] C. Coron, S. Méléard, E. Porcher, and A. Robert. Quantifying the mutational meltdown in diploid populations. *The American Naturalist*, 181(5):623–636, 2013.

[85] C. Coron, S. Méléard, and D. Villemonais. Impact of demography on extinction/fixation events. *J. Math. Biol.*, 78(3):549–577, Feb 2019.

[86] J. N. Darroch and E. Seneta. On quasi-stationary distributions in absorbing discrete-time finite Markov chains. *J. Appl. Probability*, 2:88–100, 1965.

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

[88] T. de Lange. How telomeres solve the end-protection problem. *Science*, 326(5955):948–952, 2009.

[89] R. D. DeBlassie. The lifetime of conditioned brownian motion in certain lipschitz domains. *Probability Theory and Related Fields*, 75(1):55–65, May 1987.

[90] R. D. DeBlassie. Doob's conditioned diffusions and their lifetimes. *The Annals of Probability*, pages 1063–1083, 1988.

[91] J. Deelen, M. Beekman, V. Codd, S. Trompet, L. Broer, S. Hägg, K. Fischer, P. E. Thijssen, H. E. D. Suchiman, I. Postmus, et al. Leukocyte telomere length associates with prospective mortality independent of immune-related parameters and known genetic markers. *Int J Epidemiol*, 43(3):878–886, 2014.

[92] P. Del Moral. *Feynman-Kac formulae*. Probability and its Applications (New York). Springer-Verlag, New York, 2004. Genealogical and interacting particle systems with applications.

[93] P. Del Moral. *Mean field simulation for Monte Carlo integration*, volume 126 of *Monographs on Statistics and Applied Probability*. CRC Press, Boca Raton, FL, 2013.

[94] P. Del Moral and D. Villemonais. Exponential mixing properties for time inhomogeneous diffusion processes with killing. *Bernoulli*, 24(2):1010–1032, 2018.

[95] M. C. Delfour and J.-P. Zolésio. *Shapes and geometries*, volume 4 of *Advances in Design and Control*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2001. Analysis, differential calculus, and optimization.

[96] L. Devroye and S. Janson. Protected nodes and fringe subtrees in some random trees. *Electronic Communications in Probability*, 19, 2014.

[97] F. d. di Fagagna, P. M. Reaper, L. Clay-Farrace, H. Fiegler, P. Carr, T. Von Zglinicki, G. Saretzki, N. P. Carter, and S. P. Jackson. A dna damage checkpoint response in telomere-initiated senescence. *Nature*, 426(6963):194, 2003.

[98] G. Di Gesù, T. Lelièvre, D. Le Peutrec, and B. Nectoux. Jump Markov models and transition state theory: the quasi-stationary distribution approach. *Faraday Discussions*, 195:469–495, 2016.

[99] P. Diaconis and L. Miclo. On times to quasi-stationarity for birth and death processes. *J. Theoret. Probab.*, 22(3):558–586, 2009.

[100] H. Djellout, A. Guillin, and L. Wu. Transportation cost-information inequalities and applications to random dynamical systems and diffusions. *Ann. Probab.*, 32(3B):2702–2732, 2004.

[101] M. J. D´Mello, S. A. Ross, M. Briel, S. S. Anand, H. Gerstein, and G. Paré. Association between shortened leukocyte telomere length and cardiometabolic outcomes: systematic review and meta-analysis. *Circulation: cardiovascular genetics*, 8(1):82–90, 2015.

[102] R. Douc, E. Moulines, P. Priouret, and P. Soulier. *Markov chains*. Springer Series in Operations Research and Financial Engineering. Springer, Cham, 2018.

[103] R. M. Dudley. *Real analysis and probability*, volume 74 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2002. Revised reprint of the 1989 original.

[104] M. Duflo. *Random iterative models*, volume 34 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 1997. Translated from the 1990 French original by Stephen S. Wilson and revised by the author.

[105] P. Dupuis, R. S. Ellis, and A. Weiss. Large deviations for Markov processes with discontinuous statistics. I. General upper bounds. *Ann. Probab.*, 19(3):1280–1297, 1991.

[106] P. Dupuis, K. Ramanan, and W. Wu. Large Deviation Principle For Finite-State Mean Field Interacting Particle Systems. *ArXiv e-prints*, Jan. 2016.

[107] E. B. Dynkin. *Markov processes. Vols. I, II*, volume 122 of *Translated with the authorization and assistance of the author by J. Fabius, V. Greenberg, A. Maitra, G. Majone. Die Grundlehren der Mathematischen Wissenschaften, Bände 121*. Academic Press Inc., Publishers, New York, 1965.

[108] A. Eberle. Reflection coupling and Wasserstein contractivity without convexity. *C. R. Math. Acad. Sci. Paris*, 349(19-20):1101–1104, 2011.

[109] A. Eberle. Reflection couplings and contraction rates for diffusions. *Probab. Theory Related Fields*, 166(3-4):851–886, 2016.

[110] F. Eggenberger and G. Pólya. Über die statistik verketetter vorgäge. *Zeitschrift für Angewandte Mathematik und Mechanik*, 1:279–289, 1923.

[111] C. C. Elbers, M. E. Garcia, M. Kimura, S. R. Cummings, M. A. Nalls, A. B. Newman, V. Park, J. L. Sanders, G. J. Tranah, S. A. Tishkoff, et al. Comparison between southern blots and qpcr analysis of leukocyte telomere length in the health abc study. *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences*, 69(5):527–531, 2013.

[112] S. Elmes, P. Pollett, and D. Walker. Further results on the relationship between μ-invariant measures and quasi-stationary distributions for absorbing continuous-time Markov chains. *Math. Comput. Modelling*, 31(10-12):107–113, 2000. Stochastic models in engineering, technology, and management (Gold Coast, 1996).

[113] H. J. Engelbert and T. Senf. On functionals of a Wiener process with drift and exponential local martingales. In *Stochastic processes and related topics (Georgenthal, 1990)*, volume 61 of *Math. Res.*, pages 45–58. Akademie-Verlag, Berlin, 1991.

[114] H.-J. Engelbert and G. Tittel. Integral functionals of strong Markov continuous local martingales. In *Stochastic processes and related topics (Siegmundsburg, 2000)*, volume 12 of *Stochastics Monogr.*, pages 111–132. Taylor & Francis, London, 2002.

[115] M. Erbar, M. Fathi, and A. Schlichting. Entropic curvature and convergence to equilibrium for mean-field dynamics on discrete spaces. *arXiv e-prints*, page arXiv:1908.03397, Aug 2019.

[116] M. Fathi and J. Maas. Entropic Ricci curvature bounds for discrete interacting systems. *Ann. Appl. Probab.*, 26(3):1774–1806, 2016.

[117] M. Faure and S. J. Schreiber. Quasi-stationary distributions for randomly perturbed dynamical systems. *Ann. Appl. Probab.*, 24(2):553–598, 2014.

[118] A. Ferrari, P. Groisman, and K. Ravishankar. Tumor growth, R-positivity, Multitype branching and Quasistationarity. *arXiv e-prints*, page arXiv:1906.08446, Jun 2019.

[119] P. A. Ferrari, H. Kesten, and S. Martínez. R-positivity, quasi-stationary distributions and ratio limit theorems for a class of probabilistic automata. *Ann. Appl. Probab.*, 6(2):577–616, 1996.

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

[121] P. A. Ferrari, S. Martínez, and P. Picco. Some properties of quasi-stationary distributions in the birth and death chains: a dynamical approach. In *Instabilities and nonequilibrium structures, III (Valparaíso, 1989)*, volume 64 of *Math. Appl.*, pages 177–187. Kluwer Acad. Publ., Dordrecht, 1991.

[122] P. A. Ferrari, S. Martínez, and P. Picco. Existence of nontrivial quasi-stationary distributions in the birth-death chain. *Adv. Appl. Probab.*, 24(4):795–813, 1992.

[123] G. Ferré, M. Rousset, and G. Stoltz. More on the long time stability of Feynman-Kac semigroups. *arXiv e-prints*, page arXiv:1807.00390, Jul 2018.

[124] A. L. Fitzpatrick, R. A. Kronmal, M. Kimura, J. P. Gardner, B. M. Psaty, N. S. Jenny, R. P. Tracy, S. Hardikar, and A. Aviv. Leukocyte telomere length and mortality in the cardiovascular health study. *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences*, 66(4):421–429, 2011.

[125] D. A. Forero, Y. González-Giraldo, C. López-Quintero, L. J. Castro-Vega, G. E. Barreto, and G. Perry. Telomere length in parkinson's disease: a meta-analysis. *Exp Gerontol*, 75:53–55, 2016.

[126] C. Foucart and O. Hénard. Stable continuous-state branching processes with immigration and Beta-Fleming-Viot processes with immigration. *Electron. J. Probab.*, 18:no. 23, 21, 2013.

[127] D. Freedman. *Brownian motion and diffusion*. Springer-Verlag, New York-Berlin, second edition, 1983.

[128] P. W. Glynn and P. Y. Desai. A Probabilistic Proof of the Perron-Frobenius Theorem. *arXiv e-prints*, page arXiv:1808.04964, Aug 2018.

[129] P. W. Glynn and H. Thorisson. Two-sided taboo limits for Markov processes and associated perfect simulation. *Stochastic Process. Appl.*, 91(1):1–20, 2001.

[130] G. L. Gong, M. P. Qian, and Z. X. Zhao. Killed diffusions and their conditioning. *Probab. Theory Related Fields*, 80(1):151–167, 1988.

[131] P. Good. The limiting behavior of transient birth and death processes conditioned on survival. *J. Austral. Math. Soc.*, 8:716–722, 1968.

[132] F. Gosselin. Asymptotic behavior of absorbing Markov chains conditional on nonabsorption for applications in conservation biology. *Ann. Appl. Probab.*, 11(1):261–284, 2001.

[133] J. Griffith, L. Comeau, S. Rosenfield, R. Stansel, A. Bianchi, M. H., and de Lange T. Mammalian telomeres end in a large duplex loop. *Cell.*, 97(4):503–514, 1999.

[134] B. Haas. Asymptotic behavior of solutions of the fragmentation equation with shattering: An approach via self-similar markov processes. *Ann. Appl. Probab.*, 20(2):382–429, 04 2010.

[135] B. Haas and V. Rivero. Quasi-stationary distributions and Yaglom limits of self-similar Markov processes. *Stochastic Process. Appl.*, 122(12):4054–4095, 2012.

[136] C. Harley, A. Futcher, and C. Greider. Telomeres shorten during ageing of human fibroblasts. *Nature*, 345(6274):458–460, 1990.

[137] T. E. Harris. *The theory of branching processes*. Die Grundlehren der Mathematischen Wissenschaften, Bd. 119. Springer-Verlag, Berlin; Prentice-Hall, Inc., Englewood Cliffs, N.J., 1963.

[138] A. G. Hart, S. Martínez, and J. San Martín. The λ-classification of continuous-time birth-and-death processes. *Advances in Applied Probability*, 35(4):1111–1130, 2003.

[139] P. C. Haycock, E. E. Heydon, S. Kaptoge, A. S. Butterworth, A. Thompson, and P. Willeit. Leucocyte telomere length and risk of cardiovascular disease: systematic review and meta-analysis. *BMJ*, 349:g4227, 2014.

[140] L. Hayflick and P. S. Moorhead. The serial cultivation of human diploid cell strains. *Experimental cell research*, 25(3):585–621, 1961.

[141] G. He. A note on the quasi-ergodic distribution of one-dimensional diffusions. *C. R. Math. Acad. Sci. Paris*, 356(9):967–972, 2018.

[142] G. He and H. Zhang. On quasi-ergodic distribution for one-dimensional diffusions. *Statist. Probab. Lett.*, 110:175–180, 2016.

[143] G. He, H. Zhang, and Y. Zhu. On the quasi-ergodic distribution of absorbing Markov processes. *Statist. Probab. Lett.*, 149:116–123, 2019.

[144] C. R. Heathcote, E. Seneta, and D. Vere-Jones. A refinement of two theorems in the theory of branching processes. *Teor. Verojatnost. i Primenen.*, 12:341–346, 1967.

[145] M. T. Hemann, M. A. Strong, L.-Y. Hao, and C. W. Greider. The shortest telomere, not average telomere length, is critical for cell viability and chromosome stability. *Cell*, 107(1):67–77, 2001.

[146] A. Hening and M. Kolb. Quasistationary distributions for one-dimensional diffusions with singular boundary points. *Stochastic Processes and their Applications*, 129(5):1659 – 1696, 2019.

[147] G. Hinrichs, M. Kolb, and V. Wachtel. Persistence of one-dimensional AR(1)-sequences. *ArXiv e-prints*, Jan. 2018.

[148] C. Holmgren, S. Janson, and M. Šileikis. Multivariate normal limit laws for the numbers of fringe subtrees in m-ary search trees and preferential attachment trees. *ArXiv:1603.08125*, 2016.

[149] W. Huisinga, S. Meyn, and C. Schütte. Phase transitions and metastability in Markovian and molecular systems. *Ann. Appl. Probab.*, 14(1):419–458, 2004.

[150] N. Ikeda, M. Nagasawa, and S. Watanabe. Branching Markov processes. I. *J. Math. Kyoto Univ.*, 8:233–278, 1968.

[151] N. Ikeda, M. Nagasawa, and S. Watanabe. Branching Markov processes. II. *J. Math. Kyoto Univ.*, 8:365–410, 1968.

[152] N. Ikeda, M. Nagasawa, and S. Watanabe. Branching Markov processes. III. *J. Math. Kyoto Univ.*, 9:95–160, 1969.

[153] K. Itō. *Essentials of stochastic processes*, volume 231. American Mathematical Soc., 2006.

[154] K. Itô and H. P. McKean, Jr. *Diffusion processes and their sample paths*, volume 125 of *Die Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, Berlin, 1974. Second printing, corrected.

[155] S. Jacka and J. Warren. Examples of convergence and non-convergence of Markov chains conditioned not to die. *Electron. J. Probab.*, 7:no. 1, 22, 2002.

[156] P. Jagers. General branching processes as Markov fields. *Stochastic Process. Appl.*, 32(2):183–212, 1989.

[157] S. Janson. Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stochastic Processes and Applications*, 110(2):177–245, 2004.

[158] S. Janson. Asymptotic degree distribution in random recursive trees. *Random Structures & Algorithms*, 26(1-2):69–83, 2005.

[159] A. Joffe and F. Spitzer. On multitype branching processes with $\rho \leq 1$. *J. Math. Anal. Appl.*, 19:409–430, 1967.

[160] A. Joulin. Poisson-type deviation inequalities for curved continuous-time Markov chains. *Bernoulli*, 13(3):782–798, 2007.

[161] A. Joulin. A new Poisson-type deviation inequality for Markov jump processes with positive Wasserstein curvature. *Bernoulli*, 15(2):532–549, 2009.

[162] A. Joulin and Y. Ollivier. Curvature, concentration and error estimates for Markov chain Monte Carlo. *Ann. Probab.*, 38(6):2418–2442, 2010.

[163] O. Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.

[164] S. Karlin and J. McGregor. The classification of birth and death processes. *Trans. Amer. Math. Soc.*, 86:366–400, 1957.

[165] S. Karlin and J. L. McGregor. The differential equations of birth-and-death processes, and the Stieltjes moment problem. *Trans. Amer. Math. Soc.*, 85:489–546, 1957.

[166] G. Kersting. Strong ratio limit property and R-recurrence of reversible Markov chains. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 30:343–356, 1974.

[167] G. Kersting. A note on R-recurrence of Markov chains. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 35(4):355–358, 1976.

[168] D. Khoshnevisan, P. Salminen, and M. Yor. A note on a.s. finiteness of perpetual integral functionals of diffusions. *Electron. Comm. Probab.*, 11:108–117, 2006.

[169] M. Kijima. Quasi-limiting distributions of markov chains that are skip-free to the left in continuous time. *Journal of Applied Probability*, 30(3):509–517, 1993.

[170] M. Kijima, M. G. Nair, P. K. Pollett, and E. A. van Doorn. Limiting conditional distributions for birth-death processes. *Adv. in Appl. Probab.*, 29(1):185–204, 1997.

[171] M. Kijima and E. Seneta. Some results for quasi-stationary distributions of birth-death processes. *J. Appl. Probab.*, 28(3):503–511, 1991.

[172] P. Kim and R. Song. Intrinsic ultracontractivity of non-symmetric diffusion semigroups in bounded domains. *Tohoku Math. J. (2)*, 60(4):527–547, 2008.

[173] M. Kimura, J. v. B. Hjelmborg, J. P. Gardner, L. Bathum, M. Brimacombe, X. Lu, L. Christiansen, J. W. Vaupel, A. Aviv, and K. Christensen. Telomere length and mortality: a study of leukocytes in elderly danish twins. *American journal of epidemiology*, 167(7):799–806, 2008.

[174] R. Knobloch and L. Partzsch. Uniform conditional ergodicity and intrinsic ultracontractivity. *Potential Anal.*, 33:107–136, 2010.

[175] M. Kolb and D. Steinsaltz. Quasilimiting behavior for one-dimensional diffusions with killing. *Ann. Probab.*, 40(1):162–212, 2012.

[176] S. Kondratyev, L. Monsaingeon, and D. Vorotnikov. A new optimal transport distance on the space of finite Radon measures. *Adv. Differential Equations*, 21(11-12):1117–1164, 2016.

[177] A. Lambert. Quasi-stationary distributions and the continuous-state branching process conditioned to be never extinct. *Electron. J. Probab.*, 12:no. 14, 420–446, 2007.

[178] S. Laruelle and G. Pagès. Nonlinear randomized urn models: a stochastic approximation viewpoint. *Arxiv:1311.7367*, 2013.

[179] S. Laruelle and G. Pagès. Randomized urn models revisited using stochastic approximation. *The Annals of Applied Probability*, 23(4):1409–1436, 2013.

[180] M. Ledoux. *The concentration of measure phenomenon*, volume 89 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2001.

[181] D. A. Levin, Y. Peres, and E. L. Wilmer. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, 2009. With a chapter by James G. Propp and David B. Wilson.

[182] M. Liero, A. Mielke, and G. Savaré. Optimal entropy-transport problems and a new Hellinger-Kantorovich distance between positive measures. *Invent. Math.*, 211(3):969–1117, 2018.

[183] J. Lindsey, N. McGill, L. Lindsey, D. Green, and H. Cooke. In vivo loss of telomeric repeats with age in humans. *Mutat Res.*, 256(1):45–48, 1991.

[184] J. Littin C. Uniqueness of quasistationary distributions and discrete spectra when ∞ is an entrance boundary and 0 is singular. *J. Appl. Probab.*, 49(3):719–730, 2012.

[185] M. Lladser and J. San Martin. Domain of attraction of the quasi-stationary distributions for the Ornstein-Uhlenbeck process. *J. Appl. Probab.*, 37(2):511–520, 2000.

[186] H. M. Mahmoud and R. T. Smythe. Asymptotic joint normality of outdegrees of nodes in random recursive trees. *Random Structures & Algorithms*, 3(3):255–266, 1992.

[187] H. M. Mahmoud and M. D. Ward. Asymptotic properties of protected nodes in random recursive trees. *Journal of Applied Probability*, 52(1):290–297, 2015.

[188] P. Maillard. The λ-invariant measures of subcritical bienaymé–galton–watson processes. *Bernoulli*, 24(1):297–315, 02 2018.

[189] P. Maillard and E. Paquette. Choices and intervals. *Israel Journal of Mathematics*, 212(1):337–384, 2016.

[190] C. Mailler and J.-F. Marckert. Measure-valued pólya urn processes. *Electron. J. Probab.*, 22:33 pp., 2017.

[191] C. Mailler and D. Villemonais. Stochastic approximation on non-compact measure spaces and application to measure-valued P\'olya processes. *ArXiv e-prints*, Sept. 2018.

[192] M. B. Majka. Coupling and exponential ergodicity for stochastic differential equations driven by L\'{e}vy processes. *ArXiv e-prints*, Sept. 2015.

[193] M. B. Majka. Transportation inequalities for non-globally dissipative SDEs with jumps via Malliavin calculus and coupling. *ArXiv e-prints*, Oct. 2016.

[194] A. Marguet. A law of large numbers for branching Markov processes by the ergodicity of ancestral lineages. *arXiv e-prints*, page arXiv:1707.07993, Jul 2017.

[195] C. M. Martin-Ruiz, D. Baird, L. Roger, P. Boukamp, D. Krunic, R. Cawthon, M. M. Dokter, P. Van Der Harst, S. Bekaert, T. De Meyer, et al. Reproducibility of telomere length assessment: an international collaborative study. *International journal of epidemiology*, 44(5):1673–1683, 2014.

[196] S. Martínez and J. San Martín. Rates of decay and h-processes for one dimensional diffusions conditioned on non-absorption. *Journal of Theoretical Probability*, 14(1):199–212, Jan 2001.

[197] S. Martínez and J. San Martín. Classification of killed one-dimensional diffusions. *Ann. Probab.*, 32(1A):530–552, 2004.

[198] S. Martínez and J. San Martín. Some properties of quasi-stationary distributions for finite Markov chains. In *Surveys in stochastic processes*, EMS Ser. Congr. Rep., pages 59–66. Eur. Math. Soc., Zürich, 2011.

[199] K. Marton. A measure concentration inequality for contracting Markov chains. *Geom. Funct. Anal.*, 6(3):556–571, 1996.

[200] K. Marton. Logarithmic Sobolev inequalities in discrete product spaces: a proof by a transportation cost distance. *arXiv e-prints*, page arXiv:1507.02803, Jul 2015.

[201] S. Martínez, J. San Martín, and D. Villemonais. Existence and uniqueness of a quasistationary distribution for markov processes with fast return from infinity. *J. Appl. Probab.*, 51(3):756–768, 2014.

[202] H. Matsumoto. Coalescing stochastic flows on the real line. *Osaka J. Math.*, 26(1):139–158, 1989.

[203] S. Méléard and D. Villemonais. Quasi-stationary distributions and population processes. *Probab. Surv.*, 9:340–410, 2012.

[204] S. Meyn and R. L. Tweedie. *Markov chains and stochastic stability*. Cambridge University Press, Cambridge, second edition, 2009. With a prologue by Peter W. Glynn.

[205] S. P. Meyn and R. L. Tweedie. Stability of Markovian processes i: Criteria for discrete-time chains. *Advances in Applied Probability*, 24(3):542–574, 1992.

[206] S. P. Meyn and R. L. Tweedie. Stability of Markovian processes. II. Continuous-time processes and sampled chains. *Adv. in Appl. Probab.*, 25(3):487–517, 1993.

[207] S. P. Meyn and R. L. Tweedie. Stability of Markovian processes. III. Foster-Lyapunov criteria for continuous-time processes. *Adv. in Appl. Probab.*, 25(3):518–548, 1993.

[208] A. Mielke. Geodesic convexity of the relative entropy in reversible Markov chains. *Calc. Var. Partial Differential Equations*, 48(1-2):1–31, 2013.

[209] A. Mijatović and M. Urusov. Convergence of integral functionals of one-dimensional diffusions. *Electron. Commun. Probab.*, 17:no. 61, 13, 2012.

[210] Y. Miura. Ultracontractivity for Markov semigroups and quasi-stationary distributions. *Stoch. Anal. Appl.*, 32(4):591–601, 2014.

[211] M. Nair and P. Pollett. On the relationship between μ-invariant measures and quasi-stationary distributions for continuous-time markov chains. *Advances in Applied Probability*, 25(1):82–102, 1993.

[212] S. Niemi and E. Nummelin. On nonsingular renewal kernels with an application to a semigroup of transition kernels. *Stochastic Process. Appl.*, 22(2):177–202, 1986.

[213] E. Nummelin. *General irreducible Markov chains and nonnegative operators*, volume 83 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1984.

[214] W. Oçafrain. Quasi-stationarity and quasi-ergodicity for discrete-time markov chains with absorbing boundaries moving periodically. *ArXiv e-prints*, Mar. 2017.

[215] W. Oçafrain. Q-process and asymptotic properties of Markov processes conditioned not to hit moving boundaries. *ArXiv e-prints*, Mar. 2018.

[216] Y. Ogura. Asymptotic behavior of multitype Galton-Watson processes. *J. Math. Kyoto Univ.*, 15(2):251–302, 1975.

[217] Y. Ollivier. Ricci curvature of Markov chains on metric spaces. *J. Funct. Anal.*, 256(3):810–864, 2009.

[218] Y. Ollivier. A survey of Ricci curvature for metric spaces and Markov chains. In *Probabilistic approach to geometry*, volume 57 of *Adv. Stud. Pure Math.*, pages 343–381. Math. Soc. Japan, Tokyo, 2010.

[219] A. M. Olovnikov. A theory of marginotomy: the incomplete copying of template margin in enzymic synthesis of polynucleotides and biological significance of the phenomenon. *Journal of theoretical biology*, 41(1):181–190, 1973.

[220] R. Pemantle. A survey of random processes with reinforcement. *Probab. Surv.*, 4:1–79, 2007.

[221] N. Perkowski and J. Ruf. Conditioned martingales. *Electron. Commun. Probab.*, 17:no. 48, 12, 2012.

[222] H. Pham. *Continuous-time stochastic control and optimization with financial applications*, volume 61 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2009.

[223] B. Piccoli and F. Rossi. On properties of the generalized Wasserstein distance. *Arch. Ration. Mech. Anal.*, 222(3):1339–1365, 2016.

[224] R. G. Pinsky. On the convergence of diffusion processes conditioned to remain in a bounded region for large time to limiting positive recurrent diffusion processes. *Ann. Probab.*, 13(2):363–378, 1985.

[225] R. G. Pinsky. The lifetimes of conditioned diffusion processes. *Ann. Inst. H. Poincaré Probab. Statist.*, 26(1):87–99, 1990.

[226] R. G. Pinsky. Explicit and almost explicit spectral calculations for diffusion operators. *Journal of Functional Analysis*, 256(10):3279–3312, 2009.

[227] P. Pollett. Quasi-stationary distributions : a bibliography. http://www.maths.uq.edu.au/~pkp/papers/qsds/qsds.pdf.

[228] P. Pollett. The determination of quasistationary distributions directly from the transition rates of an absorbing markov chain. *Mathematical and Computer Modelling*, 22(10):279 – 287, 1995.

[229] P. K. Pollett. Quasistationary distributions for continuous time Markov chains when absorption is not certain. *J. Appl. Probab.*, 36(1):268–272, 1999.

[230] M. Pollock, P. Fearnhead, A. M. Johansen, and G. O. Roberts. The Scalable Langevin Exact Algorithm: Bayesian Inference for Big Data. *ArXiv e-prints*, Sept. 2016.

[231] E. Priola and F.-Y. Wang. Gradient estimates for diffusion semigroups with singular coefficients. *J. Funct. Anal.*, 236(1):244–264, 2006.

[232] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.

[233] S. Rizvi, S. T. Raza, and F. Mahdi. Telomere length variations in aging and age-related diseases. *Current aging science*, 7(3):161–167, 2014.

[234] L. C. G. Rogers and D. Williams. *Diffusions, Markov processes, and martingales. Vol. 1*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2000. Foundations, Reprint of the second (1994) edition.

[235] B. Roynette, P. Vallois, and M. Yor. Some penalisations of the Wiener measure. *Jpn. J. Math.*, 1(1):263–290, 2006.

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

[237] J. W. Shay, W. E. Wright, and H. Werbin. Defining the molecular mechanisms of human cell immortalization. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer*, 1072(1):1–7, 1991.

[238] D. Steinsaltz and S. N. Evans. Markov mortality models: Implications of quasistationarity and varying initial conditions. *Theo. Pop. Bio.*, 65(65):319–337, 2004.

[239] D. Steinsaltz and S. N. Evans. Quasistationary distributions for one-dimensional diffusions with killing. *Trans. Amer. Math. Soc.*, 359(3):1285–1324 (electronic), 2007.

[240] E. Strickler. *Persistance de processus de Markov déterministes par morceaux*. PhD thesis, Université de Neuchâtel, 2019.

[241] M. Takeda. Existence and uniqueness of quasi-stationary distributions for symmetric markov processes with tightness property. *Journal of Theoretical Probability*, Jan 2019.

[242] S. Toupance, D. Villemonais, D. Germain, A. Gegout-Petit, E. Albuisson, and A. Benetos. The individual's signature of telomere length distribution. *Sci. Rep.*, 9(1):685, 2019.

[243] P. Tuominen and R. L. Tweedie. Exponential decay and ergodicity of general Markov processes and their discrete skeletons. *Adv. in Appl. Probab.*, 11(4):784–803, 1979.

[244] S. S. Vallender. Calculations of the Vasseršteĭn distance between probability distributions on the line. *Teor. Verojatnost. i Primenen.*, 18:824–827, 1973.

[245] E. A. van Doorn. Conditions for exponential ergodicity and bounds for the decay parameter of a birth-death process. *Advances in Applied Probability*, 17(3):514–530, 1985.

[246] E. A. van Doorn. Quasi-stationary distributions and convergence to quasi-stationarity of birth-death processes. *Adv. Appl. Probab.*, 23(4):683–700, 1991.

[247] E. A. van Doorn. Birth–death processes and associated polynomials. *Journal of Computational and Applied Mathematics*, 153(1):497 – 506, 2003. Proceedings of the 6th International Symposium on Orthogonal Poly nomials, Special Functions and their Applications, Rome, Italy, 18-22 June 2001.

[248] E. A. van Doorn. On associated polynomials and decay rates for birth–death processes. *Journal of Mathematical Analysis and Applications*, 278(2):500 – 511, 2003.

[249] E. A. van Doorn. On the α-classification of birth-death and quasi-birth-death processes. *Stochastic Models*, 22(3):411–421, 2006.

[250] E. A. van Doorn. Conditions for the existence of quasi-stationary distributions for birth-death processes with killing. *Stochastic Process. Appl.*, 122(6):2400–2410, 2012.

[251] E. A. van Doorn and P. K. Pollett. Quasi-stationary distributions for reducible absorbing Markov chains in discrete time. *Markov Process. Related Fields*, 15(2):191–204, 2009.

[252] E. A. van Doorn and P. K. Pollett. Quasi-stationary distributions for discrete-state models. *European J. Oper. Res.*, 230(1):1–14, 2013.

[253] A. Velleret. Unique Quasi-Stationary Distribution, with a possibly stabilizing extinction. *arXiv e-prints*, page arXiv:1802.02409, Feb 2018.

[254] D. Vere-Jones. Ergodic properties of nonnegative matrices. I. *Pacific J. Math.*, 22:361–386, 1967.

[255] S. Verhulst, E. Susser, P. R. Factor-Litvak, M. J. Simons, A. Benetos, T. Steenstrup, J. D. Kark, and A. Aviv. Commentary: The reliability of telomere length measurements. *International journal of epidemiology*, 44(5):1683–1686, 2015.

[256] L. Veysseire. Coarse Ricci curvature for continuous-time Markov processes. *ArXiv e-prints*, Feb. 2012.

[257] L. Veysseire. *Coarse Ricci curvature of Markov processes*. PhD Thesis, École normale supérieure de Lyon - ENS LYON, July 2012.

[258] D. Villemonais. Uniform tightness for time-inhomogeneous particle systems and for conditional distributions of time-inhomogeneous diffusion processes. *Markov Process. Related Fields*, 19(3):543–562, 2013.

[259] D. Villemonais. Minimal quasi-stationary distribution approximation for a birth and death process. *Electron. J. Probab.*, 20:no. 30, 18, 2015.

[260] D. Villemonais. Lower bound for the coarse ricci curvature of continuous-time pure-jump processes. *J. Theoret. Probab.*, May 2019.

[261] M.-K. von Renesse and K.-T. Sturm. Transport inequalities, gradient estimates, entropy, and Ricci curvature. *Comm. Pure Appl. Math.*, 58(7):923–940, 2005.

[262] A. Q. Wang, M. Kolb, D. Steinsaltz, and G. O. Roberts. Theoretical Properties of Quasistationary Monte Carlo Methods. *ArXiv e-prints*, July 2017.

[263] A. Q. Wang, G. O. Roberts, and D. Steinsaltz. An Approximation Scheme for Quasistationary Distributions of Killed Diffusions. *ArXiv e-prints*, Aug. 2018.

[264] F.-Y. Wang. Gradient estimates of Dirichlet heat semigroups and application to isoperimetric inequalities. *Ann. Probab.*, 32(1A):424–440, 2004.

[265] J. Wang. First eigenvalue of one-dimensional diffusion processes. *Electronic Communications in Probability*, 14:232–244, 2009.

[266] J. Wang. Sharp bounds for the first eigenvalue of symmetric markov processes and their applications. *Acta Mathematica Sinica, English Series*, 28(10):1995–2010, 2012.

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

[268] H. Zhang and Y. Zhu. Domain of attraction of the quasistationary distribution for birth-and-death processes. *J. Appl. Probab.*, 50(1):114–126, 2013.

[269] J. Zhang, S. Li, and R. Song. Quasi-stationarity and quasi-ergodicity of general Markov processes. *Sci. China Math.*, 57(10):2013–2024, 2014.

[270] L.-X. Zhang. Central limit theorems of a recursive stochastic algorithm with applications to adaptive designs. *The Annals of Applied Probability*, 26(6):3630–3658, 2016.