Cell dynamics of multitype populations in oncology and Invasion probability of cooperative parasites in structured host populations

Vianney Brouard

UNIVERSITÉ DE LYON

ENS DE LYON

Thèse

en vue de l'obtention du grade de Docteur, délivrée par

l'École Normale Supérieure de Lyon

École Doctorale N° 512

École Doctorale en Informatique et Mathématiques de Lyon

Discipline : Mathématiques

Soutenue publiquement le 12/07/2024, par :

Vianney BROUARD

Cell dynamics of multitype populations in oncology

&

Invasion probability of cooperative parasites in structured host populations

---

Devant le jury composé de :

Vincent Calvez,	Directeur de recherche CNRS à l'Université de Brest,	Co-directeur de thèse,
Loren Coquille,	Maîtresse de conférences à l'Institut Fourier, Grenoble,	Examinatrice,
Jasmine Foo,	Northrop Professor at the University of Minnesota-Twin Cities,	Rapportrice,
Amaury Lambert,	Professeur à l'ENS Paris,	Examinateur,
Hélène Leman,	Chargée de recherche INRIA à l'ENS de Lyon,	Directrice de thèse,
Marcel Ortgiese,	Senior Lecturer in Probability at the University of Bath,	Rapporteur,
Chi Tran,	Professeur à l'Université Gustave Eiffel,	Examinateur.

Abstract

This thesis focuses on the study of two stochastic models related to medical problems. The first one lies on understanding infection spread of cooperating bacteriophages on a structured multi-drug resistant bacterial host population. Motivated by this example, we introduce an epidemiological model where infections are generated by cooperation of parasites in a host population structured on a configuration model. We analysed the invasion probability for which we obtain a phase transition depending on the connectivity degree of the vertices and the offspring number of parasites during an infection of a host. At the critical scaling, the invasion probability is identified as the survival probability of a Galton-Watson process.

With the aim to get a biological more relevant model, we analysed a similar model where a spatial structure is added for the host population using a random geometric graph. We have shown that such spatial structure facilitates cooperation of parasites. A similar phase transition occurs where at the same critical scaling the invasion probability is upper and lower bounded by the survival probabilities of two discrete branching processes with cooperation.

The second medical question deals with understanding the evolution of the genetic composition of a tumour under carcinogenesis, using multitype birth and death branching process models on a general finite trait space. In the case of neutral and deleterious cancer evolution, we provide first-order asymptotics results on all mutant subpopulation sizes. In particular such results capture the randomness of all cell trait sizes when a tumour is clinically observed, and mostly it allows to characterise the effective evolutionary pathways, providing information on the past, present, and future of tumour evolution.

Moving beyond this restrictive neutral and deleterious cancer evolution framework, we provide a new method to understand the first selective mutant trait size.

Résumé

Cette thèse porte sur l'étude de deux modèles stochastiques liés à des problèmes médicaux. Le premier vise à comprendre le processus épidémique généré par des bactériophages coopératifs dans une population de bactéries résistantes aux antibiotiques. Pour cela, nous introduisons un modèle épidémiologique où les infections sont générées par la coopération de parasites dans une population d'hôtes structurée selon un modèle de configuration. Une transition de phase est observée pour la probabilité d'invasion dépendant du degré de connectivité des sommets et du nombre de parasites générés lors d'une infection d'un hôte. Au seuil critique, la probabilité d'invasion est identifiée comme la probabilité de survie d'un processus de Galton-Watson.

Dans le but d'obtenir un modèle biologiquement plus pertinent, nous avons analysé un modèle similaire où une structure spatiale est ajoutée à la population d'hôtes en utilisant un "random geometric graph". Nous avons montré qu'une telle structure spatiale facilite la coopération des parasites. Une transition de phase similaire se produit où au seuil critique, des bornes supérieure et inférieure sont obtenues pour la probabilité d'invasion en tant que probabilités de survie de deux processus de branchement avec coopération.

La deuxième question médicale concerne la compréhension de l'évolution de la composition génétique d'une tumeur en formation, en utilisant des processus de naissance et de mort multi-types branchants sur un espace de traits fini. Considérant une évolution neutre et délétère, nous fournissons des résultats au premier ordre asymptotique pour toutes les tailles des sous-populations mutantes. En particulier, nous capturons la stochasticité associée aux tailles des sous-populations mutantes lorsqu'une tumeur est observée cliniquement, et surtout nous caractérisons les chemins évolutifs effectifs, fournissant des informations sur le passé, le présent et le futur de l'évolution tumorale.

Au-delà de ce cadre restrictif d'évolution neutre et délétère, nous proposons une nouvelle méthode pour comprendre le premier ordre asymptotique du premier trait mutant sélectif.

Remerciements

Mes tous premiers remerciements iront naturellement à Hélène. À commencer bien avant la thèse, durant mon M1, ton année d'arrivée à l'UMPA, où j'étais un jeune étudiant en mathématiques bien paumé, considérant sérieusement l'option d'abandonner l'ENS et les mathématiques qui me paraissaient froides et austères, pour recommencer à zéro en étude de médecine. J'étais loin d'imaginer à l'époque à quel point assister à un exposé de mathématiques appliquées à la biologie, le tien en l'occurrence, serait aussi déterminant dans mes choix futurs pour en arriver aujourd'hui à soutenir une thèse dans ce domaine et y envisager une carrière scientifique. Pour cette raison, qui nous dépasse tous les deux, je t'en serai jamais assez reconnaissant, notamment d'avoir pris le temps de m'écouter à ce moment pour m'orienter vers le bon M2, d'avoir accepté de m'encadrer en stage puis de me faire suffisamment confiance pour me superviser en thèse (et donc t'engager à passer ton HDR), et également pour m'avoir aiguillé sur les recherches de mon stage prédoctoral à l'étranger avec Cornelia.

Durant la thèse, je voudrais commencer par te remercier pour la confiance que tu m'as directement attribuée en me laissant beaucoup d'indépendance et de liberté dans ma manière d'aborder la recherche, en m'incitant à avoir confiance en mes qualités scientifiques et à me sentir à ma place. Se savoir soutenu dès le départ m'a vraiment permis d'attaquer sereinement la recherche. J'ai également beaucoup de gratitude pour tout le cadre scientifique et humain que tu m'as apporté tout au long de la thèse. Ils ont sans conteste étaient les conditions de mon épanouissement scientifique, et personnel du coup. Je pense notamment à tes divers encouragements, à la valorisation de mes idées, à ta bienveillance émotionnelle, ainsi qu'à tes différents conseils mathématiques, que ce soit en matière de rédaction, de présentation orale, ou de réflexion sur la modélisation. Je pense aussi au cadre scientifique riche dans lequel tu m'as permis d'évoluer, notamment grâce au groupe de travail Math-Bio de Lyon, à nos discussion d'équipe Proba-Bio, à toutes les réflexions scientifiques que nous avons partagées en travaillant ensemble, ainsi qu'à l'ensemble des conférences auxquelles j'ai assistées et que tu m'as toujours encouragé à suivre. Pour toutes ces raisons, je te suis sincèrement très reconnaissant, et je peux dire que je pars serein à Bath pour de nouvelles aventures scientifiques. Merci pour tout.

Je tiens également à exprimer ma gratitude envers Vincent. Déjà pour avoir accepté de nous suivre avec Hélène lorsque nous t'avons contacté pour réaliser une codirection en modélisation en cancérologie. Puis surtout pour toutes les discussions scientifiques en modélisation déterministe, un domaine dont je ne connaissais finalement pas grand chose en début de thèse si ce n'est les connaissances classiques en sortant d'un M2 généraliste en modélisation. J'ai beaucoup apprécié en apprendre davantage, notamment via les différents débuts de projets que nous avions commencés en début de thèse, en faisant de multiples calculs au tableau noir, j'en garde un excellent souvenir. J'ai surtout compris la richesse scientifique qu'il en ressort de faire dialoguer ensemble les approches

différentes de modélisation que sont les EDPs et les probabilités, pour y voir leurs ressemblances et différences ainsi que leurs forces et faiblesses respectives. Pour cette richesse scientifique diverse, ta curiosité et ta facilité à transmettre la passion, je voudrais grandement te remercier. Tout comme Hélène, tu m'as totalement fait confiance dès le départ, et je ne peux que t'en remercier amplement, ça m'a permis d'aborder de nouvelles questions scientifiques avec plus de facilité mais surtout avec beaucoup d'envie. Finalement il n'y aura pas une once d'EDPs dans ce manuscrit, mais ça ne reflète pas tout ce que j'ai appris à ton contact sur la modélisation déterministe en biologie. En toute honnêteté, je ne peux pas dire que les EDPs ne m'ont pas conquises. Je n'en ai pas vraiment fait dans les trois dernières années, mais je n'exclue pas d'en faire un jour (donc ne soit pas surpris si je viens toquer à ta porte un jour), et pour cela merci.

I am deeply grateful to Jasmine Foo and Marcel Ortgiese for their willingness to review my PhD thesis. I recognize this as a significant and generous favor, and I appreciate their valuable, insightful and constructive feedback. I hope that my efforts to ensure clarity in writing my manuscript have somewhat made up for its length. Je tiens à exprimer ma profonde gratitude à Loren Coquille, Amaury Lambert et Chi Tran pour avoir accepté de jouer le rôle d'examinateur·rices de ma thèse et pour consacrer leur temps précieux à ma soutenance.

I would like to express my deepest gratitude to Cornelia. My first steps in research were with you. I have excellent memories of the year I spent in Frankfurt with your team, your warm welcome, and your enthusiasm for working together. You helped me build my confidence in mathematics, reaffirmed my decision to pursue a Ph.D., and most importantly, helped me do so with ease. I truly enjoyed learning more about the modeling of cooperative behaviours of infectious agents, a research topic I am particularly passionate about, inspired by your love for sharing mathematics. I also discovered German culture for the first time alongside you, and I hope to learn even more in the future. Thank you as well for continuing to work together throughout the Ph.D., especially for all the trips to Frankfurt and the short stay in Bonn. I am eager to continue this wonderful collaboration by visiting you in Lübeck.

I would also like to thank Marco, my other German collaborator, for trusting Cornelia and me to delve into the biological modeling of our cooperation questions. It has been a real pleasure working together on this wonderful project.

Il est temps de faire mes immenses remerciements à Céline, sans qui la vie au labo et ces trois années de thèse n'auraient pas été les mêmes, pour sûr. Aujourd'hui j'ai bien du mal à imaginer qu'il y a trois ans nous n'étions presque que des inconnu.es. Ce sont les mathématiques qui ont permis de nous rencontrer, mais c'est bien au delà du cadre professionnel strict que notre amitié s'est développée. J'aime bien dire que tu es comme une grande soeur, ou ma *partner in crime* au labo et dans la vie. Tu as toujours été là pour me soutenir, m'encourager, et me donner envie de persister en recherche, tout comme j'ai développé un véritable amour pour la modélisation en cancérologie en t'écoutant m'en parler. Au-delà des maths, j'ai bien grandi personnellement en trois ans, et je peux dire avec certitude que la confiance et compréhension mutuelles que nous avons construites ensemble ont été le terreau fertile à cet épanouissement. Pour toutes ces raisons, je te remercie infiniment. La vie dans une autre ville sans toi sera bien différente, c'est certain. Mais j'y vais avec beaucoup de joie, très peu d'appréhension, et beaucoup d'espoir de construire une vie riche que j'aurai hâte de te partager au plus vite quand tu viendras me rendre visite.

Je voudrais aussi particulièrement remercier David. Notamment pour ta bonne humeur et ton humour qui me fait toujours incroyablement rire, tu as une flexibilité à manier les mots et leurs sens qui m'impressionnera toujours. Mais aussi pour ta présence et ton écoute constante au labo, savoir que tu es là prêt à discuter est toujours très réconfortant.

Un merci s'impose pour le travail formidable de Jessica, Magalie et Virginia tout au long de ces années. J'ai tout de même fait pas mal de missions, et elles se sont toujours très bien déroulées notamment grâce à votre travail en amont et en aval. Et au-delà du travail administratif pur, c'est toujours un plaisir de discuter avec vous autour d'un café ou d'un repas, votre bonne humeur fait beaucoup de bien au quotidien, et pour ça merci beaucoup.

Merci également à Nicolas Champagnat et Yohann de Castro pour avoir fait parti de mon comité de suivi de thèse. Vos divers retours positifs, ainsi que vos remarques détaillées sur mon travail m'ont vraiment encouragé à poursuivre mes projets en y confrontant de nouveaux points de vues. Je voudrais également vous remercier pour vos différents conseils, aussi bien en terme de proposition de littérature que sur l'orientation de ma recherche et des compétences scientifiques à acquérir.

Je remercie maintenant mes différents cobureaux de ces trois dernières années: Clément, Elise et Alexis. Clément, quand j'y pense ça fait maintenant sept ans qu'on se connaît, et on peut dire qu'on a (un poil) grandi ensemble à l'ENS pendant ces sept années. Je me souviens qu'on était tous les deux bien plus timides début L3, je crois que j'ai de la tendresse pour les très jeunes personnes que nous étions, un peu perdues (plus moi que toi sur ce point) en sortant de prépa. Ça a vraiment été un plaisir de partager tout ces moments avec toi, aussi bien pendant la vie étudiante de l'ENS que dans notre bureau commun pendant la thèse. On aura quand même bien ri, parfois des petits pétages de câbles communs, mais surtout je retiendrai une belle écoute, à s'entraider mutuellement à respecter les deadlines et les attentes de la thèse. Discuter avec toi, quasi au quotidien (bon ok je suis peut-être trop parti en conf' pour parler de quotidien), était vraiment très agréable et c'était aussi très rassurant de se savoir soutenu. Elise et Alexis, ça aura été plus court, un an seulement. Mais j'ai beaucoup aimé partager un bureau commun avec vous. Je me souviens notre première rencontre Elise, j'étais tout gêné de te voir assise à mon bureau et de devoir te dire gentiment que tu avais en fait le pire bureau des trois. J'étais assez content, et en même temps avec un peu d'appréhension, d'accueillir une troisième personne dans le bureau, et sans mentir le match s'est bien fait, j'ai passé une super année avec toi (même si apparemment c'est moins le cas de ton côté, à partir dans un autre bureau...). Alexis on a pas eu beaucoup de temps pour vraiment se rencontrer, entre mes déplacements, la rédaction, et ta présence au labo qu'on pourrait dire assez éparse (même si je vois que tu t'y sens de mieux en mieux). Je te souhaite de continuer de t'épanouir durant tes deux dernières années de thèse, tu sembles déjà bien parti, et j'espère que tout se passera bien avec tes deux nouveaux cobureaux l'année prochaine.

Au labo, il y a encore pleins de personnes que je voudrais remercier particulièrement, pour tous les bons moments passés ensemble, les discussions en salle passerelle, les sorties dans des bars, les soirées posées chez les un.es et les autres. Vous avez toustes contribué à faire que j'apprécie venir tous les jours au labo, le quotidien sans vous aurait été définitivement plus terne, le départ sera certainement difficile et vous y êtes pour quelque chose. Je pense notamment à Alice (un an seulement qu'on se connaît, mais tu es un véritable coup de coeur amical, réciproque je crois bien, j'ai déjà hâte de te revoir, notamment pour une seconde édition du date leboncoin hehe), Basile (j'aurais voulu avoir plus de temps pour te rencontrer davantage en vrai! Merci pour toutes les différentes découvertes culturelles, notamment le cinéma de Wong Kar-wai. Je pose ici une invitation à venir à Bath dans les trois prochaines années, t'as plus le choix), Charlie (oublie pas d'écouter Britpop d'A.G. toi! et l'Angleterre c'est la terre natale de l'hyperpop, so u know viens y écouter les meilleurs sets avec moi, et puis je t'en ferai un spécial pour toi), Léo (tu as un peu été comme un grand frère de thèse, avec des conseils très utiles, une écoute attentive, des encouragements, et une grande curiosité pour le travail des un.es et des autres), William, Antoine, Thuy, Josué, Vanessa et Pablo.

Je tiens également à remercier le laboratoire dans son ensemble pour l'accueil chaleureux et le bon vivre en son sein, ainsi que la MMI (Olivier, Nina, Camille et Charlotte) pour tous les bons moments de médiations scientifiques passés ensemble.

En dehors du labo, je vais évidement commencer par te remercier Sonia. Qui de mieux placée que toi pour comprendre toute ma trajectoire mathématique et personnelle. Je crois bien que personne, y compris et surtout nous-même, il y a sept ans n'aurait parié que nous en serions arrivé.es là aujourd'hui, à soutenir nos thèses après avoir passé de belles années mathématiques, à être rempli.es de curiosité et d'appétence intellectuelle scientifique, et à être prêt.es à partir vers de nouvelles aventures et découvertes mathématiques, au Brésil pour toi et en Angleterre pour moi. C'est toujours très vertigineux de penser à tous nos souvenirs communs, tellement on a partagé des moments d'une grande sincérité, majoritairement remplis de joies intenses mais aussi de sublimes entraides fondées sur une écoute attentive, une empathie et une présence active lors de nos épreuves respectives. Je tiens évidemment à te remercier de la manière la plus sincère possible, pour tout ce qu'on a vécu ensemble, mais également pour tout ce que nous vivrons, j'ai hâte de pouvoir partager d'autant plus tout au long de nos vies (mathématiques).

Ensuite c'est à Marwane que je vais exprimer mes remerciements. Je n'avais aucun doute après notre petit périple commun à Vienne en M1 que nous deviendrions très proche en coloc' sur Lyon, mais je ne pensais pas que nous allions vivre aussi intensément ces deux belles années. J'ai vraiment l'impression qu'en deux ans à vivre ensemble, on a fondé une amitié qui s'étale sur beaucoup de temps, grâce à toutes les joies et malheurs qu'on a partagé.es de manière condensée. Je te remercie également pour cette dernière année de thèse, nous ne vivions certes plus ensemble, mais on a réussi à toujours être présent l'un pour l'autre, et nos petits moments à deux ont été très précieux.

Je tiens tout particulièrement à remercier mes ami.es qui m'ont soutenu pendant ces trois années, notamment Matthieu, Émilie, Claire, Virgile, Jérémy, Cécile, Emma, Noëmi, Enguerrand, Alex, Louis, Axel, Gauthier, Hubert, et Alice.

Je vais terminer ces remerciements par ma famille, spécifiquement mes parents et mes frères et soeurs Clément, Alban et Flavie. Votre présence constante et votre soutien tout au long de mes études ont été d'une aide inestimable. Vous m'avez toujours encouragé à poursuivre la voie scientifique que j'ai choisie, même dans les moments de doutes, en me rappelant avec beaucoup de tendresse mon amour des mathématiques. Je suis certain que plus d'une fois vous vous seriez bien passé.es de m'écouter parler de mes travaux, mais vous avez toujours joué le jeu et montré de l'intérêt malgré l'incompréhension de ce langage que sont les mathématiques. Pouvoir partager avec vous mes joies, mes détresses, et parfois mes exaltations, scientifiques ont été des moments précieux que j'ai beaucoup chéris. Durant ces années nous n'étions certes pas proche géographiquement, mais revenir vous voir à Bordeaux, apprécier la chaleur familiale, sortir un peu des mathématiques pour vivre des moments remplis de joie, d'écoute, de bienveillance et d'amour, sont des moments essentiels que j'affectionne énormément. Je m'en vais encore plus loin l'année prochaine, ce sera probablement d'autant plus difficile de se voir, mais ce sera très certainement pour une belle expérience de vie remplie de nouvelles découvertes. J'ai hâte que vous me rendiez visite, vous verrez l'Angleterre est un très beau pays. Mon tout dernier remerciement ira spécifiquement à toi Maman, tu es sans conteste la personne qui a suivi le plus assidûment tout ce voyage scientifique et personnel effectué pendant la thèse. Je ne compte plus tous nos appels à te raconter dans les détails tous mes périples. Le soutien et l'amour que tu m'as donnés ont eu une grande influence sur l'aboutissement de ce travail de thèse et l'être humain que je suis aujourd'hui. Pour cela je te remercie et t'en serai toujours immensément reconnaissant.

À mes deux Parents, Marie et René.

Contents

Chapter 1

Introduction

This thesis focuses on the study of two stochastic models related to medical problems, with the aim of providing quantitative results. The first one concerns the understanding of infection spread of cooperative bacteriophages on a multi-drug resistant bacterial host population structured on random graphs, through analysis of the invasion probability. The second medical question deals with understanding the evolutionary process of a tumour under carcinogenesis, using multitype birth and death branching process models on a general finite trait space.

Invasion of Cooperative Bacteriophages

In the evolutionary arms race with CRISPR–Cas, a defence mechanism of bacteria, bacteriophages (phages for short) have developed diverse strategies to block or evade bacterial host immunity. One common evasion mechanism known as anti-CRISPRs relies on the cooperation of phages. Motivated by this example we introduce an epidemiological model where infections are generated by cooperation of parasites in a hosts population structured on random graphs. The number of vertices and typical number of direct neighbours grow to infinity with respect to the scaling parameter N. In the initial chapter, we study the configuration model with N vertices and d_N direct neighbours for each vertex such that $d_N = \Theta\left(N^\beta\right)$, for some $0 < \beta < 1$. This modelling choice reflects the idea that among a host population, many hosts are accessible from any host, but only a negligible part of the total number. However, the biological relevance of spatially structured host populations is not reflected when considering the configuration model. In the second chapter, we address this spatial feature by studying cases where host populations are structured on a random geometric graph on $[0,1]^n$, with $n \in \mathbb{N}$. The number of vertices is Poisson(N)-distributed and the typical number of direct neighbours d_N for any vertex is $d_N = N^\beta$. Particularly, using the same scaling for the number of vertices and the number of direct neighbours in both chapters allows for comparison of the results. Quite naturally, adding a spatial structure facilitates cooperation of parasites.

The number of parasite offspring generated after the infection of a host, denoted by v_N, scales to infinity too. A host is deterministically infected when attacked by at least 2 parasites. We show that the natural scaling for observing pairs of parasites attacking the same host occurs when the number of parasite offspring scales as the square root of the typical number of neighbours of a vertex (as for the birthday problem), i.e., $v_N \sim a\sqrt{d_N}$, for some $a > 0$.

In the second chapter, we show that a phase transition occurs at the critical scaling $v_N \in \Theta\left(\sqrt{d_N}\right)$, where we use the notation $f_N = \Theta(g_N)$ if f_N grows asymptotically as fast as g_N, i.e.

$$0 < \liminf_{N\to\infty} \frac{f_N}{g_N} \leq \limsup_{N\to\infty} \frac{f_N}{g_N} < \infty.$$

More precisely, we demonstrate that when $v_N = o(d_N)$, the invasion probability is asymptotically null, since cooperation of parasites is unlikely. At the critical scaling, the invasion probability is identified as the survival probability of a Galton-Watson process. Above such critical scaling, when $\sqrt{d_N} = o(v_N)$, the invasion probability is asymptotically 1. This chapter corresponds to the article [1], entitled *Invasion of cooperative parasites in moderately structured host populations*, published in Stochastic Processes and their Applications, and written with Cornelia Pokalyuk.

In the third chapter, we demonstrate that a phase transition also occurs at the same critical scaling for the random geometric graph. We obtain lower and upper bounds on the invasion probability as the survival probabilities of two Discrete Branching Process with Cooperation (DBPC). Essentially, a DBPC is defined as follows: conditioned on knowing the number of individuals at a generation, every individual and every pair of individuals generate a certain number of new individuals according to offspring and cooperation laws, independently from each other. This Chapter corresponds to the article [2], entitled *Spatial invasion of cooperative parasites*, published in Theoretical Population Biology for the special issue for the 60th birthday of Alison Etheridge, and written with Cornelia Pokalyuk, Marco Seiler, and Hung Tran for the simulations.

The asymptotic results for the invasion probabilities at the critical scaling differs for the complete graph and the random geometric graph, in the sense that they are characterised using the survival probabilities of two different classes of processes: Galton-Watson processes for the configuration model and DBPC for the random geometric graph. This difference comes from the fact that the spatial structure facilitates the cooperation of parasites. More precisely, cooperation of parasites generated from different infected hosts effectively spread the infection directly at the beginning of the epidemic process on the random geometric graph, whereas such events become likely in the configuration model only when the invasion is already decided with high probability.

Genetic Composition of a Tumour

In a second part, we tackle another biological question that is tumour formation. More precisely, the aim is to provide a tractable probabilistic model from which theoretical results on the evolution of the genetic composition of a tumour during carcinogenesis can be obtained. Indeed, understanding the evolutionary history of a tumour can be of great clinical help. It provides a lot of information regarding its future evolution, and consequently, on clinical prognosis. Additionally, it sheds light on decisions regarding suitable adaptive therapeutic treatments to combat resistant subpopulations within it. The objective is, therefore, to study a simple probabilistic mathematical model taking into account the following different mechanisms governing tumour formation: cell division and death, in addition to mutation during cell division of one or both daughter cells towards a mutant trait. We decided to study a continuous-time multitype branching model on $\mathbb{N}_0^V$ where cells are represented by a phenotypic trait distributed as the vertices of a graph (V, E) where edges represent mutational pathways during cell division. More precisely, during a cell division event, each daughter cell, independently of each other, can mutate according to a probability kernel towards a neighbour trait of their mother cell trait. Initially, the process starts with only one wild-type cell, whose trait is supposed to be supercritical.

Tumours are typically detected when they reach a large size of cells, depending on cancer type, but approximately of order $[10^6, 10^9]$ cells. The mutation rate per base pair per cell division is generally estimated to be of order $[10^{-9}, 10^{-6}]$ [3, 4, 5]. This naturally leads to consider the framework of *large population and power law mutation rates regime*. It is a classical stochastic regime studied in [6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20]. A scaling parameter $n \in \mathbb{N}$ is employed to quantify both the decrease in mutation probabilities, as negative powers of n, and the typical size of the population, which depends on n as a positive power, at which we are interested in understanding the genetic composition. A label is assigned to each edge to quantify the power law scaling of the mutation under consideration: the larger the label, the lower the probability of mutation. The aim is to derive first-order asymptotics on all mutant subpopulation sizes as n tends to infinity, and to mathematically characterise the effective evolutionary pathways of tumour progression. In Chapter 4, we examine neutral and deleterious cancer evolution, while in Chapter 5, we focus on selective cancer evolution. By neutral and deleterious cancer evolution, we refer to cases where the growth rate of any mutant trait is smaller than that of the wild-type subpopulation. Here, mutations are biologically termed *passenger mutations*, as they do not confer any selective advantage; all *driver mutations*, mutations conferring selective advantage, were already present in the initiating cancer cell. Conversely, in cases allowing for selective mutation, where a mutation has a growth rate strictly higher than that of the wild-type subpopulation, any such selective mutant trait represents a driver mutation, along with potentially numerous passenger mutations included in the model.

In the fourth chapter, we narrow the mutational framework to neutral and deleterious mutations. We furnish first-order asymptotic results on the mutant subpopulation sizes. This approach not only enables to capture the stochastic nature of mutant subpopulation sizes but also allows for the characterisation of effective evolutionary pathways. We demonstrate that the first-order asymptotic randomness of all mutant subpopulation sizes is entirely determined by the random variable describing the long-term behaviour of the *primary subpopulation*, which is the lineage of wild-type cell issued from the initial cell. Effective evolutionary pathways are characterised based on their lengths, defined as the sum of the labels on its edges, and their numbers of neutral mutations. This Chapter corresponds to a major revision of the preprint [21], entitled ≪Genetic Composition of Supercritical Branching Populations under Power Law Mutation Rates≫ for the Annals of Applied Probability.

In the fifth chapter, we investigate the first-order asymptotics of the first-selective mutant trait. The method developed in the previous chapter no longer applies once a selective mutation is encountered. We propose a new proof method to get asymptotic results for the first selective mutation along a mono-directional finite graph. This method lies on breaking down the selective mutant population into a sum of subpopulations for which different probabilistic techniques are applied. We obtain that well renormalized, the selective mutant population behaves asymptotically as a sum of a compound Luria-Delbrück distribution with infinite expectation and of the large-time limit of an approximate model with less stochasticity, independent from n. This model is introduced and studied in [22, 23, 24]. This chapter is an ongoing work in collaboration with Hélène Leman.

In the rest of the Introduction, we give more details about the biological contexts, the model definitions, and the mathematical results. We motivate them all using heuristics and furnish sketches of the proofs, alongside some perspectives.

1.1 Invasion probability of cooperative parasites in structured host populations

This work falls within the field of mathematical epidemiology on random graphs. The uniqueness of this study lies in infections being transmitted via a population of cooperative infectious agents, known as parasites, moving within the host population undergoing infection. Together with Cornelia Pokalyuk and Marco Seiler, we have developed mathematical tools and techniques existing in epidemiology on random graphs to adapt them to this specific context of infection via cooperation, in order to estimate invasion probabilities. We start by giving a biological context to motivate the microscopic model and then we give details on the quantitative results we obtained alongside heuristics, sketches of the proofs, biological remarks and generalisations.

1.1.1 Biological context

Understanding the dynamics and mechanisms of infection processes is a highly relevant and dynamic research area in biology. In Chapters 2 and 3, our focus lies on modelling infections based on observations of bacteriophage populations, viruses that infect bacteria. Interest in phages has surged in recent years due to the rise in multi-drug resistant bacteria. Phage therapy, an alternative to antibiotics, involves inoculating infected hosts with phages to eliminate the pathogenic bacterial population [25]. Phages exert selective pressure on bacterial populations, prompting the development of various defence mechanisms. One such mechanism is CRISPR-Cas (Clustered Regularly Interspaced Short Palindromic Repeats-associated proteins), a well-studied RNA-guided adaptive immune system. The CRISPR–Cas immune response begins with acquiring short DNA fragments, called protospacers, from invading phages. These protospacers are inserted as spacers within the CRISPR array to establish infection memory as well as heredity [26]. The CRISPR array is then expressed as a long transcript, processed into small, mature CRISPR RNAs (crRNAs) carrying spacer sequences. Interference complexes, consisting of a crRNA and one or more Cas proteins, patrol the bacterial cell to detect and destroy complementary nucleic acid targets, often adjacent to a short protospacer-adjacent motif, in phage genetic material previously encountered by the bacterial cell (or one of its ancestors) and stored in the bacterial genome. This mechanism leads to the destruction of encountered phages and is illustrated in Figure 1.1.

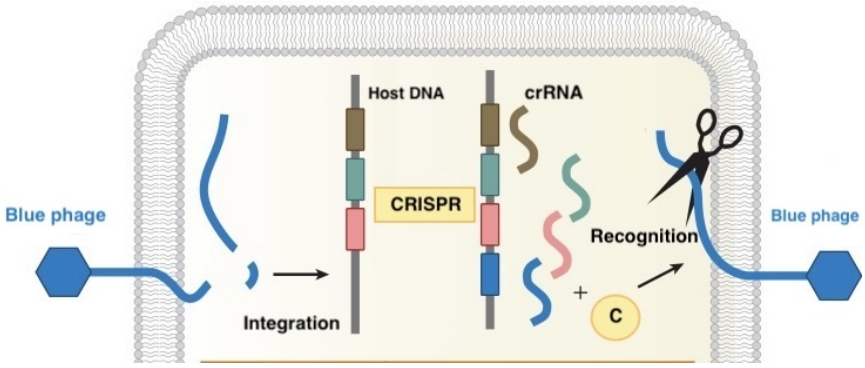

Figure 1.1: CRISPR-Cas defence mechanism, © Westra et al. (2012) Annu Rev Genet

In the evolutionary arms race with CRISPR–Cas, phages have developed diverse strategies to block or evade bacterial host immunity. One common evasion mechanism involves protein-based CRISPR–Cas inhibitors known as anti-CRISPRs (ACRs), which mainly interact directly with Cas proteins [27]. ACR defence mechanism relies on the cooperation of ACRs-phages [28]. Indeed, when a CRISPR-resistant bacterium is attacked by a single ACRs-phage, the phage often dies. However, when several phages attack a bacterium simultaneously or subsequently, they have a higher chance of replication [29],[28]. The observation stems from the fact that the time period required for an ACRs-phage to inhibit all interference complexes present in the bacterial host is long enough to allow the bacterial host to destroy the phage. However, during a subsequent infection, the remaining number of interference complexes is low enough to enable the infecting ACRs-phage to block all of them. The ACR mechanism is illustrated in Figure 1.2.

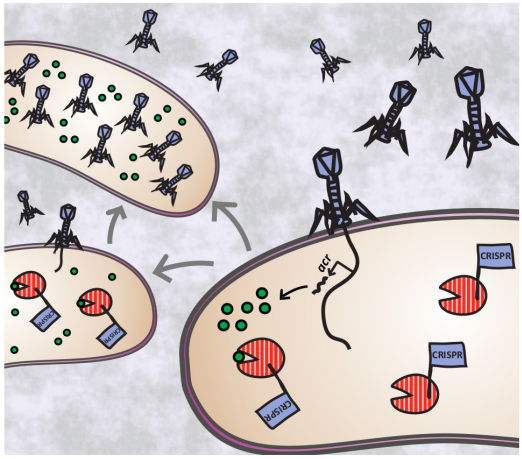

Figure 1.2: ACRs defence mechanism, © S. van Houte

While the primary biological motivation of this study lies in understanding the cooperative infection of phages within a bacterial population, other relevant "host-parasite" systems with cooperative infectious mechanisms also exist. One such example is virotherapy, where oncolytic viruses are emerging as potential treatment options for cancer. Both natural and genetically engineered viruses exhibit various antitumour mechanisms, such as direct cytolysis, immune system potentiation through antigen release, activation of inflammatory responses, or interference with elements in the tumour microenvironment. They can also indirectly affect tumour cells by modifying energy metabolism and displaying antiangiogenic action. However, a significant challenge in oncolytic virotherapy is the efficient delivery of the virus into tumour cells and its ability to bind to various biological and non-biological vectors. Despite this challenge, oncolytic viruses have demonstrated efficacy in eliminating cancer cells resistant to standard treatments in numerous clinical trials across various cancers, including melanoma, lung, and hepatic cancers [30]. It is believed that infecting cancer cells with multiple viruses simultaneously enhances effectiveness, as this strategy enables the virus to better counteract the interferon-based antiviral response of the host, see [31].

1.1.2 Microscopic model of Chapters 2 and 3

Consider a population of hosts and a population of parasites both located on a random graph. A scaling parameter $N \in \mathbb{N}$ tending to infinity is used for quantifying the number of vertices b_N, the typical number of neighbours a vertex has d_N, and the number of offspring parasites generated at a host infection v_N. Initially, on each vertex a single host is placed and they do not move on the graph during the epidemic spread. The infection process starts by the infection of a host with a parasite. We say that parasites infect a host, when the infecting parasites replicate in the host. At replication $v_N \in \mathbb{N}$ offspring parasites are generated (independent on the number of infecting parasites) and the host as well as the infecting parasite(s) die(s). The infection process continues in discrete generations according to the following scheme. At the beginning of each generation, parasites move independently to nearest neighbouring vertices. If a vertex to which a parasite moves to is still occupied with a host the parasite attacks this host. If a host is attacked by at least two parasites, they cooperate and infect the host. If, however, a host is attacked by a single parasite, the parasite replicates and infects the host only with a small probability ρ_N. Otherwise (with probability $1 - \rho_N$), the parasite dies and the host survives. If a parasite moves to a vertex that is no longer occupied by a host, two options are considered depending on the random graph, either it stays there and moves further in the next generation or it dies.

Notice that cooperative infections are of two types: either the infecting parasites comes from the same previously infected vertex or not. We are going to see in the mathematical analysis that they imply different epidemic spread speeds. Cooperation from parasites generated at the same vertex infection is responsible for the exponential growth of the infection process, where cooperation from parasites generated on different vertex infections allows super-exponential growth, provided that such events are likely.

Given a sequence of parameters $(b_N, d_N, v_N, \rho_N)_{N\in\mathbb{N}}$ we denote for each $N \in \mathbb{N}$ by

$$\mathcal{I}^{(N)} = \left(I_n^{(N)}\right)_{n\in\mathbb{N}_0},$$

the process that counts the number of infected hosts in the generations $n \in \mathbb{N}_0$ with $I_0^{(N)} = 1$. Introduce also

$$\overline{\mathcal{I}}^{(N)} = \left(\overline{I}_n^{(N)}\right)_{n\in\mathbb{N}_0}, \text{ with } \overline{I}_n^{(N)} := \sum_{i=0}^{n} I_i^{(N)}$$

the process that counts the number of hosts infected till generation $n \in \mathbb{N}_0$. We are interested in the probability that the parasite population invades the host population. More precisely, we consider the following events.

Definition 1.1.1 (Invasion of parasites)**.** *Consider the above host-parasite model with parameters* $(b_N, d_N, v_N, \rho_N)_{N\in\mathbb{N}}$*. Let* $u \in (0, 1]$ *and denote by*

$$E_u^{(N)} := \left\{\exists n \in \mathbb{N}_0 : \overline{I}_n^{(N)} \geq u \cdot b_N\right\},$$

the event that the parasites invade the host population (at least) to a proportion u*.*

The different graphs under consideration are the *configuration model* in Chapter 2, and the *random geometric graph* on $[0, 1]^n$ for $n \in \mathbb{N}$, alongside the *complete graph*, on Chapter 3. The

configuration model is constructed by pairing uniformly at random the d_N half-edges of each vertex (assuming that $d_N b_N$ is even). The complete graph satisfies $b_N = d_N$. The random geometric graph has intensity $N\lambda^n(\cdot)$, where λ^n denotes the Lebesgue measure on $\mathbb{R}^n$. Moreover introduce the parameter $r_N > 0$ such that each vertex $x \in [0,1]^n$ on the graph is connected to all vertices located on $B_x^{(N)}(r_N) \cap [0,1]^n$, the intersection between the ball centred around x of radius r_N for the L^∞-norm and the considered space $[0,1]^n$. Choose r_N that satisfies $(2r_N)^n N = d_N$ such that the mean number of direct neighbours of a vertex, not too close to the border, is d_N.

We consider the following parameter regimes:

- **Number of vertices:** The host population is initially large, that is $b_N \underset{N\to\infty}{\longrightarrow} \infty$. More specifically, $b_N = N$ for the configuration model, $b_N \sim \text{Pois}(N)$ for the random geometric graph, and $b_N \in \Theta\left(N^\beta\right)$, for some $\beta \in (0,1)$, for the complete graph.

- **Typical number of neighbours:** Assume $d_N \in \Theta\left(N^\beta\right)$ for the configuration model, and $d_N = N^\beta$ for the random geometric graph, for some $\beta \in (0,1)$, implying that $r_N = \frac{1}{2^n} N^{\frac{\beta-1}{n}}$. Such scaling implies that from each host many other hosts are direct neighbours but the population is not well mixed (compared to the complete graph). In particular taking the same scaling for the two models allows for comparison of results, i.e. understanding the influence of a spatial structure on the epidemic spread.

- **Number of offspring parasites:** Many offspring parasites are produced at infection of a host, i.e. $v_N \to \infty$. The results on the invasion probabilities will depend on the scaling of v_N compared to $\sqrt{d_N}$.

- **Single parasite infection:** The contribution of parasites attacking a host alone is at most critical in the sense that the expected number of offspring $v_N\rho_N$ generated at such attacks converges at most to 1, i.e. $v_N\rho_N \underset{N\to\infty}{\longrightarrow} x \in [0,1]$. For simplicity, when the host population is structured on the random geometric graph or the complete graph, we take $\rho_N = 0$. This means that in the latter cases infections are generated only through cooperation of parasites. Such restrictive choice facilitates the mathematical analysis, but it could have been possible not to consider it without modifying the results on the invasion probability.

The choice of the initially infected vertex is done uniformly at random for the configuration model and the complete graph. In the case of the random geometric graph, choose the infected vertex to be the closest to the central point $\left(\frac{1}{2}, \cdots, \frac{1}{2}\right)$. Moreover, a parasite moving to an empty vertex (where the host has been infected and killed in a previous generation) survives for the configuration model, whereas it dies for the random geometric graph and the complete graph. Though, results on the invasion probability would be similar allowing survival and next movement of the parasite.

In Theorems 1.1.2, 1.1.5 and 1.1.6 we identify the phase transition that occurs at the critical scaling $v_N \in \Theta\left(\sqrt{d_N}\right)$, at which invasion of the host population turns from an unlikely to an asymptotically almost sure event, for all the random graphs under consideration. Moreover, at the critical scaling, the invasion probability is characterised for the configuration model, respectively the complete graph, as the survival probability of a Galton-Watson process, respectively of a Discrete Branching Process with Cooperation (DBPC). For the random geometric graph, upper and lower bounds are derived for the invasion probability as the survival probabilities of two DBPCs.

1.1.3 Results of Chapter 2

In this subsection, the results from Chapter 2 are given with explanations, heuristics, sketches of the proofs and some generalisations. The main result deals with the invasion probability of cooperative parasites on host populations structured according to a configuration model, as mentioned in the previous subsection. Under the scaling $b_N = N$, $d_N \in \Theta\left(N^\beta\right)$ for some $\beta \in (0,1)$ and $\rho_N v_N \to x \in [0,1]$, a phase transition occurs for the invasion probability at the critical scaling $v_N \in \Theta\left(\sqrt{d_N}\right)$. More precisely, below the critical scaling the invasion is unlikely, above it happens with high probability and at the critical scaling the invasion probability is characterised as the survival probability of a Galton-Watson process (GWP for short) with Poisson$\left(\frac{a^2}{2} + x\right)$ offspring distribution. Such critical scaling, the identification of the invasion probability as the survival probability of a GWP, and the specific offspring distribution of this GWP can heuristically be derived.

Heuristics

Critical scaling: To derive the critical scaling of v_N with respect to d_N, we take into account only infections generated due to cooperation of parasites and forget the single parasite infection process, i.e. $\rho_N = 0$. Cooperation of parasites becomes likely only when the critical scaling is reached, that is $v_N \in \Theta\left(\sqrt{d_N}\right)$. Indeed, consider an infected vertex such that v_N parasites located on it are going to move to one of its d_N neighbours, and assume that all of these neighbours have a host located on them. Counting the number of infections generated due to the movement of these v_N parasites is similar to the so-called *birthday problem*: consider the experiment where you have d days and v persons, the questions is what is the scaling of v with respect to d for which it becomes likely that at least two persons share a birthday, when the birthdays are distributed uniformly and independently at random on the days. The answer is exactly $v \in \Theta\left(\sqrt{d}\right)$. Performing computations, in the scaling $v \sim a\sqrt{d}$ for some $a > 0$, the asymptotic distribution of the number of shared birthdays is a Poisson$\left(\frac{a^2}{2}\right)$ distribution. To be even more precise, asymptotically only the days where exactly two persons share a birthday are contributing when considering the scaling $v \in \Theta\left(\sqrt{d}\right)$, in other words, at the limit days with more than two people sharing a birthday disappear.

Now if you consider back the single parasite infection process, due to the scaling $\rho_N v_N \to x \in [0,1]$ this infection process is (sub)critical, in the denomination of GWP, in the sense that the mean number of infections $v_N \rho_N$ generated by this process converges at most to 1. That explains why for the scaling $v_N \in o\left(\sqrt{d_N}\right)$, so when infections from cooperation are not likely, the invasion probability is asymptotically null.

Offspring distribution of the GWP: Consider both the critical scaling $v_N \sim a\sqrt{d_N}$ for some $a > 0$, and the scaling for the single parasite infection process $\rho_N v_N \to x \in [0,1]$. Counting the number of infections generated by v_N parasites located on a vertex and moving to d_N non-empty vertices is an alternative version of the birthday problem. In this situation consider the sum of the number of days where at least two persons share a birthday and the number of single birthdays which are selected independently from each other with a successful probability ρ_N. Performing computations gives that asymptotically this sum has a Poisson$\left(\frac{a^2}{2} + x\right)$ distribution. Indeed, under the scaling $v_N \to \infty$, the number of host infections triggered by pairs of parasites and the number

of single parasite infections are asymptotically independent. They are asymptotically distributed respectively as Poisson$\left(\frac{a^2}{2}\right)$, derived by the classical birthday problem, and Poisson(x), derived as the limit of a Bin(v_N, ρ_N) random variable under the scaling $v_N\rho_N \to x$. As a consequence their sum is Poisson$\left(\frac{a^2}{2} + x\right)$ distributed.

Identification of the invasion probability as the survival probability of the previous GWP: Moreover, the events that make the infection process deviating from a GWP are unlikely at least up to the random generation at which a specific level, tending to infinity, is reached for the infection process. Such events are typically cooperation from parasites generated from two different infected vertices and re-hitting of a vertex by at least a pair of parasites or a single parasite with (normally) successful host infection. In particular, this observation allows for comparing the infection model, up to it reaches the aforementioned level, by a coupling argument with a GWP whose offspring distribution is sufficiently close to a Poisson$\left(\frac{a^2}{2} + x\right)$ distribution such that its survival probability is identified to the one of the GWP whose offspring distribution is exactly a Poisson distribution with parameter $\frac{a^2}{2} + x$. In Proposition 1.1.3 we explicit the mathematical meaning of closeness for offspring distribution allowing asymptotic identification of the survival probability for a sequence of GWPes. That explains why the natural candidate for the invasion probability is the survival probability of this latter GWP. To prove it, it suffices to show that the aforementioned level grows to infinity sufficiently quickly such that the invasion of the host population is with high probability already decided when this level is reached for the infection process.

Main Results

Now we can state the result about the invasion probability of cooperative parasites in host populations structured on a configuration model in the next Theorem.

Theorem 1.1.2 (Invasion probability for the configuration model)**.** *Assume $d_N \in \Theta\left(N^\beta\right)$ for some $0 < \beta < 1$, and $\rho_N v_N \to x$ for some $x \in [0, 1]$. A phase transition for the invasion probability occurs depending on the scaling of the parasite offspring parameter v_N:*

(i) Assume $v_N \in o\left(\sqrt{d_N}\right)$. Then for all $0 < u \leq 1$

$$\lim_{N\to\infty} \mathbb{P}\left(E_u^{(N)}\right) = 0.$$

(ii) Assume $v_N \sim a\sqrt{d_N}$ for some $a > 0$. Denote by $\pi(a, x)$ the survival probability of a GWP with Poisson$\left(\frac{a^2}{2} + x\right)$-offspring distribution. Then for all $0 < u \leq 1$

$$\lim_{N\to\infty} \mathbb{P}\left(E_u^{(N)}\right) = \pi(a, x).$$

(iii) Assume $\sqrt{d_N} \in o(v_N)$. Then

$$\lim_{N\to\infty} \mathbb{P}\left(E_1^{(N)}\right) = 1.$$

Before giving some sketches of the proof, we are formulating some remarks on this theorem:

- **Invasion probability:** Notice that even in the setting of Theorem 1.1.2 (ii), the invasion is not ensured with a strictly positive probability. Indeed, for $\frac{a^2}{2} + x \leq 1$ the considered GWP is (sub)critical meaning that $\pi(a, x) = 0$. In the scaling $v_N \rho_N \to x$, if we allowed $x > 1$, asymptotically the single parasite infection process is supercritical, meaning that it would have been possible for the setting of Theorem 1.1.2 (i) to see invasion with a strictly positive probability, but not due to cooperation of parasites, explaining why we focused only on $x \leq 1$.

- **Population viscosity:** In the model we are considering that the spatial structure of the host population is passed on to the parasite population that profits from this structure to cooperate. It has been shown that population viscosity, i.e. limited dispersal of individuals, is generally beneficial for cooperation, see [32]. Thus, in host-parasite systems the host population may on the one hand profit from a spatial structure by enhancing cooperation of hosts, but on the other hand spatial structure may reduce the fitness of the host population because parasite populations may benefit from the spatial structure as well, as demonstrated in Theorem 1.1.2.

- **Death of parasites:** In our model we assume that parasites that hit empty vertices keep moving further and hosts are not reproducing. These parasite have only a negligible impact on the fate of the parasite population. Hence, the statements of Theorem 2.2.2 remain valid, if we assume that parasites die (or die with a certain probability) when hitting an empty vertex.

- **Time to invasion:** During the proof of Theorem 1.1.2 (ii) we also show an upper bound on the time till total invasion, conditioned on a parasite outbreak. We obtained the upper bound $\frac{\left(1-\frac{3}{4}\beta+\varepsilon\right)\log N}{\log\left(\frac{a^2}{2}+x\right)}$ for any $\varepsilon > 0$. Indeed to prove Theorem 1.1.2 (ii) we approximate $\mathcal{I}^{(N)}$ by a GWP from below, that is truncated from time to time but grows at the same speed as an ordinary GWP, with (asymptotically) mean offspring number $\frac{a^2}{2} + x$, until the infection level $N^{1-\frac{3}{4}\beta+\varepsilon}$ is reached, for some $\varepsilon > 0$ sufficiently small. Afterwards we show that the host population gets killed with high probability within two more generations due to cooperation of parasites generated from infections of different vertices. From this follows immediately that the host population is with high probability killed after time $\frac{(1-\frac{3}{4}\beta+\varepsilon)\log N}{\log\left(\frac{a^2}{2}+x\right)}$ for any $\varepsilon > 0$ in case of invasion of the parasite population. With some more effort we expect that it is possible to show that in the setting of Theorem 1.1.2 (ii), conditioned on invasion of the host population, it happens with high probability before $\frac{(1-\beta+\varepsilon)\log N}{\log\left(\frac{a^2}{2}+x\right)}$ generations, for any $\varepsilon > 0$. Infection by cooperation of parasites generated from different vertex infections takes over when the number of infected hosts exceeds the level $N^{1-\beta+\varepsilon}$, subsequently the host population should be killed with high probability in a finite number of generations. Such efforts are actually made in Chapter 3, where they are necessary for understanding the invasion probability.

In Chapter 2, in addition to results on the invasion probability, we are also proving interesting results on asymptotic survival probability of sequences of GWPes. In particular the more refined result states that reaching any arbitrary level b_N tending to infinity for the total size of a sequence of GWPes, under the condition of uniform convergence for the generating functions to the generating function of a GWP, is asymptotically equals to the survival probability of the latter GWP. Such

criterion comes from the identification of the extinction probability of a GWP as the fixed point of the generating function of the offspring distribution.

Proposition 1.1.3. *Consider a sequence of GWPes* $\left(Z^{(N)}\right)_{N\in\mathbb{N}}$ *and a GWP* Z*. Denote by* $\left(\Phi^{(N)}\right)_{N\in\mathbb{N}}$ *and* $\left(\pi^{(N)}\right)_{N\in\mathbb{N}}$ *the sequences of respectively the generating functions of the offspring distributions and the survival probabilities of* $\left(Z^{(N)}\right)_{N\in\mathbb{N}}$*, and by* Φ *and* π *the ones of* Z*. Assume that the generating functions* $\left(\Phi^{(N)}\right)_{N\in\mathbb{N}}$ *converge uniformly in* $[0,1]$ *to* Φ*. Furthermore, let* $(b_N)_{N\in\mathbb{N}}$ *be an* $\mathbb{N}$*-valued sequence with* $b_N \to \infty$*. Then the following holds:*
i)

$$|\pi^{(N)} - \pi| \to 0,$$

ii)

$$\mathbb{P}\left(Z^{(N)}_{b_N} = 0\right) \to 1 - \pi,$$

iii)

$$\mathbb{P}\left(\exists\ n \in \mathbb{N}_0 : Z^{(N)}_n \geq b_N\right) \to \pi,$$

iv)

$$\mathbb{P}\left(\exists\ n \in \mathbb{N}_0 : \overline{Z}^{(N)}_n \geq b_N\right) \to \pi,$$

where $\overline{Z}^{(N)}_n := \sum_{i=0}^{n} Z^{(N)}_i$ *for all* $n \in \mathbb{N}$*.*

Sketch of the Proof

We proceed by sketching the proof of Theorem 1.1.2:

- **Theorem 1.1.2 (i):** Under the scaling $v_N \in o\left(\sqrt{d_N}\right)$, parasites are unlikely to cooperate, as mentioned in the heuristics with the birthday problem. Hence, invasion could only be achieved by the single parasite infection process. But since we are considering the parameters regime $\rho_N v_N \to x \leq 1$, such process is asymptotically (sub)critical, meaning that having a successful single parasite infection is asymptotically too rare for allowing the invasion of the host population. Consequently, only a negligible proportion of the host population is infected by the epidemic process before it stops. So for any $u \in (0,1]$ the invasion probability is $o(1)$. Concretely, we show that for any $a > 0$ the infection process $\mathcal{I}^{(N)}$ can be coupled for N large enough with the infection process $\mathcal{I}^{(N)}(a)$ defined according to the same model but with the scaling $a\sqrt{d_N}$ for the number of offspring parasites, such that $\overline{\mathcal{I}}^{(N)}$ is bounded from above by $\overline{\mathcal{I}}^{(N)}(a)$. Then, applying the same proof techniques as in Theorem 1.1.2 (ii) one obtains that the invasion probability of the process $\mathcal{I}^{(N)}$ is asymptotically bounded from above by $\pi(a,x)$ for any $a > 0$. Since $\pi(a,x) \underset{a\to 0}{\to} 0$, it implies that the invasion probability of the infection process $\mathcal{I}^{(N)}$ is asymptotically null.
- **Theorem 1.1.2 (ii)** Consider the critical scaling $v_N \sim a\sqrt{d_N}$ for some $a > 0$. We identify the invasion probability as the survival probability $\pi(a,x)$ of a GWP whose offspring distribution is $\mathrm{Pois}\left(\frac{a^2}{2} + x\right)$ distributed using coupling arguments from above and from below with well-chosen sequences of upper and lower (truncated) GWPes such that both their survival probabilities asymptotically are $\pi(a,x)$.

Upper bound: More precisely, the invasion probability is naturally upper bounded by the probability to reach any level $\ell_N \to \infty$, satisfying $\ell_N \leq N$, for the total number of host infections. Thus we couple from above with high probability the total number of infected hosts $\overline{\mathcal{I}}^{(N)}$ by the total size of an upper GWP $\overline{Z}_u^{(N)}$ until it reaches a certain level $\ell_N \to \infty$ or it dies out. This level is chosen to grow to infinity sufficiently slowly such that the coupling with a branching structure works. The sequence of offspring distributions is shown to be sufficiently close to the $\mathrm{Pois}\left(\frac{a^2}{2} + x\right)$ distribution such that assumption of Proposition 1.1.3 is satisfied. In particular it implies, from Proposition 1.1.3 iv), that reaching the level ℓ_N for the total size of this sequence of GWPes is asymptotically $\pi(a, x)$. Combining these arguments give that the invasion probability is asymptotically upper bounded by $\pi(a, x)$.

Lower bound: Deriving a lower bound on the invasion probability is more complex. The proof is done in two steps: first we show that reaching the level $N^{1-\frac{3}{4}\beta+\varepsilon}$ for the total number of infected hosts $\overline{\mathcal{I}}^{(N)}$, for $\varepsilon > 0$ small enough, occurs asymptotically with probability $\pi(a, x)$, then we argue that when this specific level is reached, the total population of hosts is killed with high probability in the next two generations due to infections generated by parasites issued from different previously infected hosts. Indeed, if $I_n^{(N)} \in \Theta\left(N^\gamma\right)$ in some generation n for some $\gamma > 0$ $\left(\text{and } \overline{I}_n^{(N)} \ll N\right)$, an order of $\Theta\left(v_N N^\gamma\right)$ offspring parasites are generated. From these parasites, $\Theta\left(N^{2\gamma+\beta}\right)$ pairs of parasites can be formed. The majority of these pairs consists of parasites that have been generated on different vertices. The probability that such a pair of parasites attacks the same vertex is approximately $\frac{1}{N}$. For $\gamma > 1 - \beta$ we have $2\gamma + \beta - 1 > \gamma$. Hence, when $\Theta\left(N^\gamma\right)$ hosts are infected for some $1 - \beta < \gamma < 1$, more hosts get infected by pair of parasites generated from different infected hosts than by pairs of parasites coming from the same host infection. Furthermore, for $1 - \frac{3\beta}{4} < \gamma < 1 - \frac{\beta}{2}$, after one generation, $\Theta\left(N^{2\gamma+\beta-1}\right)$ hosts get infected and since $2\gamma + \beta - 1 > 1 - \frac{\beta}{2}$ and $2\left(1 - \frac{\beta}{2}\right) + \beta - 1 = 1$, after another generation on average all hosts get killed.

For the first step of the proof, a dichotomy occurs depending on the value of β with respect to $\frac{4}{7}$.

Case $\beta > 4/7$: The easiest case consists of $\beta > \frac{4}{7}$, where we couple from below with high probability the total number of infected hosts $\overline{\mathcal{I}}^{(N)}$ by the total size of a lower GWP $\overline{Z}_l^{(N)}$ until no further hosts are killed or the total number of hosts that got infected exceeds the threshold N^α, for $0 < \alpha < \beta$. As for the upper GWPes, the sequence of offspring distributions is shown to be sufficiently close to the $\mathrm{Pois}\left(\frac{a^2}{2} + x\right)$ distribution such that reaching the level N^α for the total size of this sequence of lower GWPes is asymptotically $\pi(a, x)$, applying Proposition 1.1.3 iv). Hence the probability to reach the level N^β for the total infection process $\overline{\mathcal{I}}^{(N)}$ is asymptotically lower bounded by $\pi(a, x)$. Moreover $\beta > \frac{4}{7} \iff 1 - \frac{3}{4}\beta < \beta$, meaning that in particular the coupling can be applied with $\alpha = 1 - \frac{3}{4}\beta + \varepsilon$, for $0 < \varepsilon < \frac{7}{4}\beta - 1$. Thus, combining the two steps of the proof gives the $\pi(a, x)$ as a lower bound for the invasion probability in the case $\beta > \frac{4}{7}$.

Case $\beta \leq 4/7$: In the case $\beta \leq \frac{4}{7}$ the argument is slightly more involved, since in this case it is not possible to approximate with high probability $\overline{\mathcal{I}}^{(N)}$ from below by the total size of the lower GWP $\mathcal{Z}_l^{(N)}$ until $N^{1-\frac{3\beta}{4}+\varepsilon}$ hosts get infected (or the infection dies out), for any $\varepsilon > 0$.

Indeed, if the number of infected hosts exceeds the level N^β, then with non-trivial probability an edge is attacked from both ends simultaneously by pairs of parasites or parasites with successful infection probability. In this case none of these parasites cause an infection of a host, because the vertices to which these parasites are heading to are already empty. However, we can derive an upper bound on the number of parasites involved in such events and remove the corresponding branches in the lower GWP. Since these parasites make up only a vanishing proportion of the total parasite population, the growth of the corresponding truncated GWP is asymptotically the same as that of the original GWP. Hence, for the truncated GWP essentially the same techniques can be applied to finish the proof concerning the probability of invasion in the case $\beta \leq \frac{4}{7}$.

- **Theorem 1.1.2 (iii):** Under the scaling $\sqrt{d_N} \in o(v_N)$, the infection of a single host leads to an asymptotically infinite number of further host infections. Consequently, at least one of the infected hosts triggers the invasion of the host population with high probability. Concretely, for any $a > 0$, we couple the infection process $\mathcal{I}^{(N)}$ with the infection process $\mathcal{I}^{(N)}(a)$ defined according to the same model but with the scaling $a\sqrt{d_N}$ for the number of offspring parasites, such that $\overline{\mathcal{I}}^{(N)}$ is bounded from below by $\overline{\mathcal{I}}^{(N)}(a)$. From this fact, we obtain that the invasion probability of the infection process $\mathcal{I}^{(N)}$ is asymptotically lower bounded by $\pi(a, x)$, for any $a > 0$. Since $\pi(a, x) \underset{a\to\infty}{\to} 1$, it implies that the invasion probability of the infection process $\mathcal{I}^{(N)}$ is asymptotically 1.

Generalisations

The model under study is a toy model for quantitatively understanding the impact of cooperative mechanism on the infection of host populations by parasites. Theorem 1.1.2, stating the asymptotics of the invasion probability, can be extended to more general biological settings. Next we point out some of these and explain why such generalisations are still tractable models from which we can obtain similar results for the invasion probability:

(i) **More randomness:** In order to make the study simple we assume that the number d_N of half-edges per vertex and the number v_N of parasite offspring, as well as the probability ρ_N are deterministic. However, biologically such quantities could vary from vertex/host/parasite. One could draw these numbers in an i.i.d. manner according to some distributions $\mathcal{D}^{(N)}$, $\mathcal{V}^{(N)}$ and $\mathcal{P}^{(N)}$, in order to add more heterogeneity in the model. If such distributions are sufficiently concentrated, the model remains tractable by simple proof adaptations, and Theorem 1.1.2 remains valid in this context. For instance, a sufficient condition is to have i.i.d. random variables $\left(Y_i^{(N)}\right)_{N\in\mathbb{N}}$ distributed as $\mathcal{D}^{(N)}$, $\mathcal{V}^{(N)}$, with corresponding expectation μ_N, satisfying that it exists some $c_N \in o(\mu_N)$ such that

$$\mathbb{P}\left(\bigcap_{i=1}^{N}\left\{\left|Y_i^{(N)} - \mu_N\right| \leq c_N\right\}\right) = \left(1 - \mathbb{P}\left(\left|Y_i^{(N)} - \mu_N\right| > c_N\right)\right)^N \to 1, \tag{1.1.1}$$

and given the total number of parasites that can be generated is M_N, to have i.i.d. random

variables $\left(Y_i^{(N)}\right)_{N\in\mathbb{N}}$ distributed as $\mathcal{P}^{(N)}$ such that

$$\mathbb{P}\left(\bigcap_{i=1}^{M_N}\left\{\left|Y_i^{(N)}-\mu_N\right|\leq c_N\right\}\right)=\left(1-\mathbb{P}\left(\left|Y_i^{(N)}-\mu_N\right|>c_N\right)\right)^{M_N}\to 1.$$

For example, this is fulfilled for $\mathcal{D}^{(N)}$ if $Y_1^{(N)}$ is distributed as a discretized normal distribution with mean $\mu_N\in\Theta\left(N^\beta\right)$ and variance $\sigma_N^2\in o\left(N^{2\beta-\delta}\right)$ for some $\delta>0$ or if it is $Pois\left(N^\beta\right)$ distributed. If $\left(Y_i^{(N)}\right)_{N\in\mathbb{N}}$ has a heavy-tailed distribution with mean $\mu_N=N^\beta$ and $Y_i^{(N)}-\mu_N$ has a Pareto-tail, then Equation (1.1.1) is fulfilled, if the tail is of order $\tau>\frac{1}{\beta}$. Similar distributions can be chosen for $\mathcal{V}^{(N)}$ and $\mathcal{P}^{(N)}$.

(ii) **Longer lasting blocking of bacteria:** In our model we implicitly assume that bacteria get blocked only for a single generation after a phage attack. In reality this blocking may last for a longer time. In this case the result of the asymptotic invasion probability remains the same. Indeed, recall that as long as $\overline{I}_n^{(N)}\leq N^\gamma$ for some $\gamma<\frac{1}{2}(1-\beta)<1-\beta$ the number of vertices attacked from parasites generated on different infected hosts is negligible with respect to the number of newly infected vertices. Assuming $\overline{I}_n^{(N)}=N^\gamma$ for some $\gamma<\frac{1}{2}(1-\beta)$ we also have with high probability that $I_n^{(N)}=\Theta\left(N^\gamma\right)$ and the probability that a blocked host (which number is of order $N^\gamma v_N$) is also attacked by another parasite in generation n is of order $\mathcal{O}\left(N^\gamma v_N\frac{N^\gamma v_N}{N}\right)$. In the scaling $v_N\in\Theta\left(\sqrt{d_N}\right)$, this probability is non-trivial for $\gamma\geq\frac{1}{2}(1-\beta)$. Since invasion of the host population is already decided if the frequency of infected host reaches N^ε for some $\varepsilon>0$, at this stage of the epidemic invasion of the host population occurs anyway with probability $1-o(1)$.

(iii) **Heterogeneity of the phage population:** Among the phage population, some heterogeneity can exist. Some phages can be ACRs-phages, and others are not. Including this heterogeneity on the model would not change the result of Theorem 1.1.2. Indeed, phages that are not able to block CRISPR-resistant bacteria may have a chance to replicate in bacteria that have been blocked by ACRs-phages before. However, by a similar reasoning as in item (ii) of these remarks, this is only likely when the amount of this type of phages is of order $N^{1-\beta/2}$, that is this type of phages must be much more frequent than ACRs-phages initially.

(iv) **Dependence of offspring parasites on the number of infecting parasites:** In reality the number of offspring parasites generated during an infection could depend on the number of parasites infecting a host. In the scaling of Theorem 1.1.2 (ii) the probability that a host gets infected by k parasites, for $k\geq 3$, from a set of parasites of size v_N located on the same vertex scales as $N^{-\frac{(k-2)\beta}{2}}$. Consequently, such events are becoming likely when the total infection process reaches the level $\Theta\left(N^{\frac{\beta}{2}}\right)$, where cooperation of 3 parasites can asymptotically be observed. Thus these kind of reproduction events have only a negligible impact on the initial spread of the parasite population. Moreover when the level $\Theta\left(N^{\frac{\beta}{2}}\right)$ is reached, the invasion probability is already decided with high probability. Meaning that giving $v_N^{(k)}$ offspring generated at reproduction of k parasites infecting a host, such that

$v_N = o\left(v_N^{(k)}\right)$, is not enhancing the invasion probability but only speeds up the infection spread after reaching the level $\Theta\left(N^{\frac{\beta}{2}}\right)$.

(v) **Reproducing hosts:** Considering host reproductions during the time scale of the parasite infection is biologically relevant. For instance, one could include in the model that hosts may reproduce on empty nearest-neighbour spots with random offspring numbers per host that are bounded. With this consideration, Theorem 1.1.2 remains valid. Indeed, for $\varepsilon > 0$ small enough, the probability that at least $N^{1-\beta+\varepsilon}$ hosts get infected is asymptotically independent on the state of the vertices on which hosts have been killed already, because the probability to re-hit these vertices is small when the overall number of infected hosts is $\ll N$. After reaching the level $N^{1-\beta+\varepsilon}$, cooperation of parasites generated on different infected hosts makes the parasite population expands faster with every generation. It confers a super exponential speed to the epidemic process, implying that in only a finite number of generations the host population gets killed with high probability. Host reproduction cannot curb down this strong parasite expansion, at least when the offspring host numbers are bounded.

1.1.4 Results of Chapter 3

In this subsection, the results from Chapter 3 are explained alongside with some sketches of the proofs, heuristics and some simulations. Two distinct results are presented in this chapter. Both of them deals with the invasion probability of cooperative parasites on hosts populations. The first one considers a host population structured on a complete graph where the second one focuses on a random geometric graph (RGG for short), as mentioned in Subsection 1.1.2.

With the RGG, we are taking the same order of vertices and typical number of direct neighbours as for the configuration model in Chapter 2, in order to compare the two structures, where a spatial component with more biological insights is added when considering the RGG. With the complete graph, there is less biological relevance, but its analysis is used as an elementary brick for the study on the RGG, and has its own mathematical interest. Before giving details on the results, we are making a proper definition of a DBPC, which is the mathematical object arising for analysing the invasion probability of cooperative parasites both for the complete graph and the RGG structures.

Definition 1.1.4 (Discrete-time Branching Process with Cooperation)**.** *Let $\mathcal{L}_o$ and $\mathcal{L}_c$ be two probability distributions on $\mathbb{N}_0$. A discrete-time branching process with cooperation $\mathbf{Z} = (Z_g)_{g\geq 0}$ with offspring distribution $\mathcal{L}_o$ and cooperation distribution $\mathcal{L}_c$ is recursively defined as follows. Assume $Z_0 = k$ a.s. for some $k \in \mathbb{N}$, then for any $g \geq 1$, Z_g is defined as*

$$Z_g := \sum_{i=1}^{Z_{g-1}} X_{g,i} + \sum_{i,j=1, i>j}^{Z_{g-1}} Y_{g,i,j},$$

where $(X_{g,i})_{(g,i)\in\mathbb{N}^2}$ and $(Y_{g,i,j})_{(g,i,j)\in\mathbb{N}^3}$ are sequences of i.i.d. random variables distributed according to $\mathcal{L}_o$ and $\mathcal{L}_c$ respectively. We denote by $\overline{\mathbf{Z}} = \left(\overline{Z}_g\right)_{g\geq 0}$ the total size process, i.e.

$$\overline{Z}_g = \sum_{i=0}^{g} Z_i.$$

In the following we will denote the probability weights of the distributions $\mathcal{L}_o$ and $\mathcal{L}_c$ by $(p_{k,o})_{k\in\mathbb{N}_0}$ and $(p_{k,c})_{k\in\mathbb{N}_0}$ respectively.

In the case of the complete graph, under the scaling $b_N = d_N \in \Theta\left(N^\beta\right)$ for some $\beta \in (0,1)$, a phase transition occurs for the invasion probability at the critical scaling $v_N \in \Theta\left(\sqrt{d_N}\right)$. More precisely, below the critical scaling the invasion is unlikely, above it happens with high probability and at the critical scaling, the invasion probability is characterised as the survival probability of a DBPC with $\mathrm{Pois}\left(\frac{a^2}{2}\right)$ and $\mathrm{Pois}\left(a^2\right)$ offspring and cooperation distributions respectively.

In the case of the RGG, under the scaling $r_N = \frac{1}{2^n} N^{\frac{\beta-1}{n}}$, such that the mean number of direct neighbours a vertex (not too close to the border of $[0,1]^n$, in the sense with an asymptotically non-vanishing distance to its boundary) has is $d_N = (2r_N)^n N = N^\beta$ for some $\beta \in (0,1)$, a phase transition also occurs for the invasion probability at the same critical scaling $v_N \in \Theta\left(\sqrt{d_N}\right)$. At the critical scaling, upper and lower bounds on the invasion probability are given. They are both characterised as survival probabilities of two distinct DBPCs. The offspring and cooperation distributions are respectively $\mathrm{Pois}\left(\frac{a^2}{2}\right)$ and $\mathrm{Pois}\left(a^2\right)$ distributed for the upper bound, and respectively $\mathrm{Pois}\left(\frac{a^2}{2^{n+1}}\right)$ and $\mathrm{Pois}\left(\frac{a^2}{2^n}\right)$ distributed for the lower bound.

Such critical scaling, characterisation of the invasion probability, respectively upper and lower bounds on the invasion probability, as the survival probability of a DBPC for the complete graph, respectively as the survival probabilities of two distinct DBPCs for the RGG, and the offspring and cooperation distributions of these specific DBPCs are next heuristically derived.

Heuristics for the Complete Graph

Critical scaling: Concerning the derivation of the critical scaling of v_N with respect to d_N, the same heuristics as for the configuration model can be performed. Indeed, in order to spread the epidemic, the v_N parasites located on the first infected vertex needs to cooperate otherwise the epidemic process would stop. Each vertex has $d_N - 1$ neighbours in the complete graph, meaning that the situation, similar to the birthday problem, implies the critical scaling $v_N \in \Theta\left(\sqrt{d_N}\right)$ for allowing cooperation. Such heuristics also gives that under the subcritical scaling $v_N \in o\left(\sqrt{d_N}\right)$, with high probability the v_N parasites generated by the first infected host do not cooperate. Thus, with high probability the epidemic process stops at the first generation, leaving $b_N - 1$ hosts uninfected. Hence, the invasion probability is asymptotically null under this subcritical scaling.

Offspring and cooperation distribution of the DBPC: Under the critical scaling $v_N \sim a\sqrt{d_N}$, for some $a > 0$, the epidemic spread on the complete graph differs from the one on the configuration model, even at the initial phase of the infection process. Where interactions between two infected hosts are negligible in the case of the configuration model at least as long as the epidemic process has not reached a certain level, they are not directly from the first generation in the case of the complete. This phenomenon comes from the share of $d_N - 2 \in \Theta\left(N^\beta\right)$ common neighbours on the complete graph for two infected vertices, making cooperation of the two sets of the v_N generated parasites on these vertices likely. Indeed, to quantify this fact, a more complex version of the birthday problem can be used, with two distinct groups of persons. For simplicity, consider the following experiment: assume there are d boxes and two sets of v red and blue balls that are thrown independently and uniformly at random into the boxes. Under the scaling $v \sim a\sqrt{d}$ for some $a > 0$, asymptotically the number of boxes with at least two blue/red balls, respectively one of each colour, into it is distributed according to a $\mathrm{Pois}\left(\frac{a^2}{2}\right)$, respectively $\mathrm{Pois}\left(a^2\right)$, distribution. More precisely, these three limited random variables are independent from each other, and asymp-

totically only the configurations where there are no boxes with at least three balls on them are effectively contributing. Notice in particular that the sum is Poisson$\left(2a^2\right)$ distributed, as in the birthday problem with d days and $2v$ persons. Consequently, such heuristics between cooperation of the generated parasites from two infected vertices motivate the asymptotic comparison between the infection process and a DBPC whose offspring and cooperation distribution are Poisson distributed with parameters $\frac{a^2}{2}$ and a^2 respectively.

Identification of the invasion probability as the survival probability of a DBPC: Moreover, the events that make the infection process deviating from a DBPC are unlikely at least up to the generation at which a specific level, tending to infinity, is reached for the infection process. Such events typically are the re-hitting of a vertex, in the sense that at least two pairs of parasites go to the same vertex, at the same generation or not. In particular, this observation allows for comparing the infection process, up to it reaches the aforementioned level, with a DBPC whose offspring and cooperation distributions are asymptotically close to Poisson$\left(\frac{a^2}{2}\right)$ and Poisson$\left(a^2\right)$ distributions, such that its survival probability is asymptotically identified to the one of the DBPC whose offspring and cooperation distributions are exactly Poisson distributed with parameters $\frac{a^2}{2}$ and a^2 respectively. That explains why the natural candidate for the invasion probability is the survival probability of this latter DBPC. To prove it, it suffices to show that the aforementioned level grows to infinity sufficiently quickly such that the invasion of the host population is with high probability already decided when this level is reached for the infection process.

Heuristics for the RGG

Critical scaling: Concerning the derivation of the critical scaling, the situation is not different from the one of the complete graph. In a ball centred around an infected vertex, there are a typical number of vertices $d_N = N(2r_N)^n = N^\beta$ at which this infected vertex is directly connected to according to the definition of a RGG. Thus, cooperation of parasites becomes likely exactly at the same critical scaling $v_N \in \Theta\left(\sqrt{d_N}\right)$.

Offspring and cooperation distribution of the upper and lower DBPCs: However, cooperation of parasites generated from different infected vertices differs from the situation on the complete graph. Due to the distance between two infected vertices located on x and y in $[0,1]^n$, the area composed of the intersection of the balls of radius r_N centred around these two infected vertices is at most of volume the one of a ball of radius r_N. This area is composed exactly of the vertices that are both direct neighbours of the two infected vertices. Hence, the number of infections generated by cooperation of parasites generated from these two infected vertices are asymptotically stochastically dominated by a Pois$\left(a^2\right)$ distribution, and depend on the coordinates of the vector $x - y$. Such dependence on the spatial localisation implies that a direct comparison between the infection process and a DBPC is actually not straightforward. The spatial component of the infection process resulting from the spatial structure of the RGG can not asymptotically disappear. Nevertheless, an identification could be done as the survival probability of a more complex model than a DBPC, typically where individuals are spatially located and the cooperation distribution depends on the distance between them.

Though, asymptotic upper and lower bounds of the invasion probability can be obtained as the survival probabilities of two well chosen DBPCs, by constructing couplings for which the spatial component is not an issue anymore. Indeed, cooperation of parasites from two infected vertices

decreases with the distance between these two infected vertices. Then, naturally an upper bound is obtained if one gives the maximal distribution for these cooperation, as if the two infected vertices where located exactly at the same spot. With this idea, one can show that the invasion probability is asymptotically upper bounded by the survival probability of a DBPC whose offspring and cooperation distributions are $\mathrm{Pois}\left(\frac{a^2}{2}\right)$ and $\mathrm{Pois}\left(a^2\right)$ distributed respectively (as in the case of the complete graph). Deriving a lower bound is a bit more tricky, where one have to use that locally a RGG looks like a complete graph. Indeed, all the vertices located on the ball of radius $\frac{r_N}{2}$ centred around the initially infected vertex are connected altogether, because they are all at a distance less than r_N from each others. Consequently, if one consider the sub-epidemic process resulting from infections of vertices located on this ball generated by cooperation of parasites coming from infected vertices from this sub-epidemic process, a natural coupling from below between the actual infection process and this sub-epidemic process can be performed. In this ball, the typical number of vertices is $\widetilde{d}_N := r_N^n N = \frac{1}{2^n} d_N$. Moreover, when a vertex is infected, among the v_N generated parasites the typical number of them that are actually moving to one of the vertices located on this ball is $\widetilde{v}_N := \frac{1}{2^n} v_N \sim \frac{a}{2^n}\sqrt{d_N} = \frac{a}{\sqrt{2^n}}\sqrt{\widetilde{d}_N}$. Thus, infections generated by cooperation of parasites coming from the same infected vertex is asymptotically $\mathrm{Pois}\left(\frac{a^2}{2^{n+1}}\right)$ distributed, and the ones from cooperation of parasites coming from different infected vertices is asymptotically $\mathrm{Pois}\left(\frac{a^2}{2^n}\right)$ distributed. Consequently, the natural identification of the lower bound for the invasion probability is the survival probability of a DBPC whose offspring and cooperation distributions are the latter ones. To prove it, it suffices to show that it exists a level tending to infinity for which the sub-epidemic process is well approximated with the aforementioned DBPC up to it reaches this level, and that the actual invasion of the hosts population is with high probability decided when it is reached.

Main Results

Now we can state the results about the invasion probability of cooperative parasites on hosts populations structured on a complete graph and on a RGG in the next two theorems.

Theorem 1.1.5 (Invasion probability for the complete graph)**.** *Assume the host population is structured on a complete graph with $b_N = d_N \in \Theta\left(N^\beta\right)$, for some $0 < \beta < 1$, vertices. A phase transition for the invasion probability occurs depending on the scaling of the parasite offspring parameter v_N:*

(i) Assume $v_N \in o\left(\sqrt{d_N}\right)$. Then for all $0 < u \leq 1$

$$\lim_{N\to\infty} \mathbb{P}\left(E_u^{(N)}\right) = 0.$$

(ii) Assume $v_N \sim a\sqrt{d_N}$ for some $a > 0$. Denote by $\pi_c(a)$ the survival probability of a DBPC with $\mathrm{Poisson}\left(\frac{a^2}{2}\right)$ and $\mathrm{Poisson}\left(a^2\right)$ offspring and cooperation distributions respectively. Then for all $0 < u \leq 1$

$$\lim_{N\to\infty} \mathbb{P}\left(E_u^{(N)}\right) = \pi_c(a) > 0.$$

(iii) Assume $\sqrt{d_N} \in o(v_N)$. Then

$$\lim_{N\to\infty} \mathbb{P}\left(E_1^{(N)}\right) = 1.$$

Simulations of the invasion probability at the critical scaling are presented in Figure 1.3 for a complete graph with $\beta = 0.7$, as well as the asymptotic survival probability of the corresponding DBPC for graphical representation of Theorem 1.1.5 (ii).

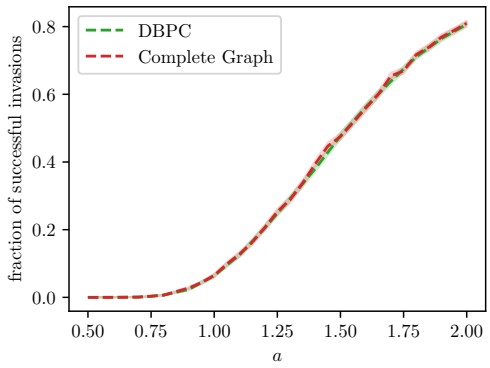

Figure 1.3: Simulated invasion probabilities with a host population structured on a complete graph for $N = 10^6$ and $\beta = 0.7$ as well as simulated survival probabilities $\pi_c(a)$. The number of simulations per date point is 10^4.

Theorem 1.1.6 (Invasion probability for the RGG)**.** *Assume the host population is structured on a RGG on* $[0,1]^n$ *with intensity* $N\lambda^n(\cdot)$ *where* λ^n *denotes the Lebesgue measure on* $\mathbb{R}^n$*, and where vertices at distance smaller than* $r_N = \frac{1}{2^n} N^{\frac{\beta-1}{n}}$ *for the* L^∞*-norm share an edge, for some* $0 < \beta < 1$*. Introduce the mean number of neighbours a vertex (with an asymptotically non-vanishing distance to the boundary) has as* $d_N = N(2r_N)^n = N^\beta$*. A phase transition for the invasion probability occurs depending on the scaling of the parasite offspring parameter* v_N*:*

(i) Assume $v_N \in o\left(\sqrt{d_N}\right)$*. Then for all* $0 < u \leq 1$

$$\lim_{N\to\infty} \mathbb{P}\left(E_u^{(N)}\right) = 0.$$

(ii) Assume $v_N \sim a\sqrt{d_N}$ *for some* $0 < a < \infty$*. Then for all* $0 < u \leq 1$

$$\pi_c\left(\frac{a}{\sqrt{2^n}}\right) \leq \liminf_{N\to\infty} \mathbb{P}\left(E_u^{(N)}\right) \leq \limsup_{N\to\infty} \mathbb{P}\left(E_u^{(N)}\right) \leq \pi_c(a).$$

(iii) Assume $\sqrt{d_N} \in o(v_N)$*. Then*

$$\lim_{N\to\infty} \mathbb{P}\left(E_1^{(N)}\right) = 1.$$

Moreover when $v_N \sim a\sqrt{d_N}$ *for some* $a > 0$*. Denote by*

$$T^{(N)} := \inf\left\{g \in \mathbb{N} \middle| \overline{I}_g^{(N)} = b_N\right\}.$$

Then

$$\lim_{N\to\infty} \mathbb{P}\left(\left\lfloor \frac{1}{2r_N}\right\rfloor \leq T^{(N)} \leq \left\lceil \frac{1}{2r_N}\right\rceil + \mathcal{O}\left(\max\left(\log(\log(N)), \frac{\varepsilon_N}{r_N^2}\right)\right) \middle| T^{(N)} < \infty\right) = 1,$$

with $\varepsilon_N = \left(N^{\frac{\frac{\beta}{2}-1}{n}+\delta}\right)$, *for any* $\delta > 0$.

Simulations of the invasion probability at the critical scaling are presented in Figure 1.4 for a RGG on $[0,1]$ and $[0,1]^2$ both for $\beta = 0.7$. In particular, the lower and upper bounds are also added, for graphical representation of Theorem 1.1.6 (ii). The lower bound is not sharp at all (and become even more worse with the increasing dimension), whereas the upper bound is a better approximation, even if there is still not an identification with the invasion probability.

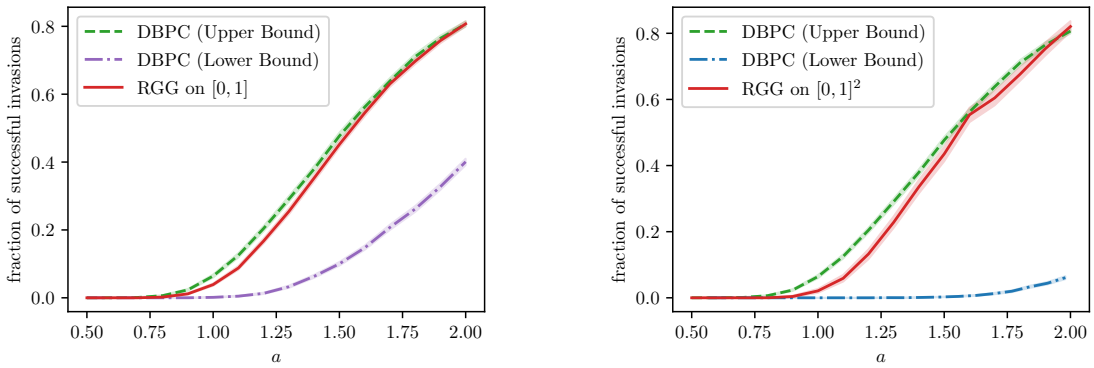

Figure 1.4: Simulated invasion probabilities with a host population structured on a RGG on $[0,1]$ (left) and $[0,1]^2$ (right) for $N = 10^6$ and $\beta = 0.7$ as well as simulated survival probabilities $\pi_c\left(\frac{a}{\sqrt{2^n}}\right)$ and $\pi_c(a)$.

Before giving some sketch of the proof, we are formulating some remarks on this theorem:

- **Strictly positive invasion probability:** One shows by application of Proposition 1.1.8 that for any $a > 0$ the survival probability $\pi_c\left(\frac{a}{\sqrt{2^n}}\right)$, of a DBPC with Poisson$\left(\frac{a^2}{2^{n+1}}\right)$ and Poisson$\left(\frac{a^2}{2^n}\right)$ offspring and cooperation distributions, is strictly positive. Therefore, in Theorem 1.1.6 (ii), for any $a > 0$ the invasion probability is strictly positive. This contrasts with the host population structured on a configuration model, as in Chapter 2, where for $\frac{a^2}{2} \leq 1$ (considering $x = 0$ to perform accurate model comparison), the invasion probability is asymptotically 0, see Theorem 1.1.2 (ii).

- **Riemannian manifold structure:** In Theorem 1.1.6 we are looking at a RGG structured on $[0,1]^n$ with the L^∞-norm. This choice comes from making the mathematical analysis the simplest possible, although it has a limited biological modelling interest. Indeed, most of the time, a bacterial population does not structured as the unit cube, but has a more complex shape, that can be modelled through a Riemannian manifold. However, in Subsection 3.2.3

of Chapter 3, we give a brief heuristic justification why our results should also carry over to a setting where the unit cube $[0,1]^n$ is replaced by a Riemannian manifold, and the L^∞-norm by the metric on the considered Riemannian manifold (induced by the Riemannian metric). Such setting choice allows for a more adequate biological representation of a bacterial population.

- **Generalisations:** As for Chapter 2, the model on the RGG can be generalised to take into account more biological insights, but still being tractable. In particular, as for the study on the configuration model, a success probability ρ_N of infection from a single parasite can be considered under the scaling $\rho_N v_N \underset{N\to\infty}{\to} x \in [0,1]$. The invasion probability under the subcritical or supercritical scaling won't be modified. In the case of the critical scaling, the lower bound, respectively the upper bound, on the invasion probability would now be identified as the survival probability of a DBPC with offspring and cooperation distributions $\text{Poisson}\left(\frac{a^2}{2^{n+1}} + \frac{x}{2^n}\right)$ and $\text{Poisson}\left(\frac{a^2}{2^n}\right)$, respectively $\text{Poisson}\left(\frac{a^2}{2} + x\right)$ and $\text{Poisson}\left(a^2\right)$. One could also consider the case where parasites moving to an empty vertex survive and can move again in the next generation, such a generalisation won't change the result. And as for the first remark of Theorem 1.1.2, some randomness on the number v_N of generated parasites could be considered.

Sketches of the Proofs

We start by sketching the proof of Theorem 1.1.5:

- **Theorem 1.1.5 (i):** For the subcritical scaling $v_N \in o\left(\sqrt{d_N}\right)$, as previously mentioned in the heuristics, it is unlikely for the v_N parasites generated on the first infected vertex to cooperate such that with high probability the epidemic stops at the first generation with a total of only one infected host. Hence the invasion probability is asymptotically null.

- **Theorem 1.1.5 (ii):** For the critical scaling $v_N \sim a\sqrt{d_N}$, to arrive at an upper bound on the invasion probability we couple with high probability the total infection process $\overline{\mathcal{I}}^{(N)}$ from above with the total size of a DBPC $\overline{\mathbf{Z}}_u^{(N)}$ until it remains constant or it reaches at least the level ℓ_N for a sequence ℓ_N with $\ell_N \to \infty$ sufficiently slowly for the coupling to hold. The offspring and cooperation distributions of the latter DBPC are sufficiently close to Poisson distributions with parameters $\frac{a^2}{2}$ and a^2 respectively, such that the sequence of the survival probabilities of these approximating DBPCs asymptotically converges to $\pi_c(a)$, see Proposition 1.1.10. Moreover the probability to reach the level ℓ_N for the process $\overline{\mathbf{Z}}_u^{(N)}$ is asymptotically equal to $\pi_c(a)$, since for a DBPC reaching a level tending to infinity implies survival with high probability. In case the level ℓ_N is reached for the total infection process $\overline{\mathcal{I}}^{(N)}$, that is asymptotically upper bounded by $\pi_c(a)$ as just mentioned before, we upper bound the invasion probability by 1, meaning that afterwards also the remaining hosts get infected.

  For the lower bound we couple with high probability $\overline{\mathcal{I}}^{(N)}$ from below with the total size of a DBPC $\overline{\mathbf{Z}}_\ell^{(N)}$ until it remains constant or it reaches the level N^ε for $\varepsilon > 0$ small enough for the coupling to hold. As for the upper bound, the offspring and cooperation distributions of the latter DBPC are chosen sufficiently close to Poisson distributions with parameters $\frac{a^2}{2}$ and a^2 such that its survival probability is asymptotically equal to $\pi_c(a)$. As for the upper bound,

the probability to reach the level N^{ε} for $\overline{\mathbf{Z}}_{\ell}^{(N)}$ is asymptotically equal to $\pi_c(a)$. When the level N^{ε} is reached for the total infection process $\overline{\mathcal{I}}^{(N)}$ we show that with high probability such process grows in a constant number of generations to a level $N^{\frac{\beta}{2}+\widetilde{\varepsilon}}$ for some small $\widetilde{\varepsilon} > 0$, due to the superexponential epidemic spread triggered by cooperation of parasites generated on different infected vertices. After reaching this particular level the remaining uninfected hosts get killed with high probability in a single generation.

- **Theorem 1.1.5 (iii):** Under the supercritical scaling $\sqrt{d_N} \in o(v_N)$, we couple from below with high probability the total infection process $\overline{\mathcal{I}}^{(N)}$ with the total size of a GWP whose offspring distribution is approximately a Poisson distribution with parameter $\frac{a^2}{2}$, until N^{α} hosts get infected or the parasite population dies out for any $0 < \alpha < \beta$ and any $a > 0$. This is done by ignoring infection due to cooperation of parasites generated on different vertices. The offspring distribution of the latter GWP is chosen sufficiently close to a Poisson distribution with parameter $\frac{a^2}{2}$ implying that its survival probability asymptotically converges to $\pi(a)$, the survival probability of a GWP with exactly Poisson$\left(\frac{a^2}{2}\right)$ offspring distribution. For a GWP, reaching any level tending to infinity implies survival with high probability. Thus the probability for the total infection process to reach the level N^{α} is asymptotically lower bounded by $\pi(a)$, due to the latter property on GWPes combined with the coupling from below. By choosing $\alpha > \frac{\beta}{2}$, we show that once the level N^{α} is reached for the total infection process $\overline{\mathcal{I}}^{(N)}$, with high probability after one more generation the remaining uninfected hosts get killed, a consequence of cooperation of parasites generated on different infected vertices triggering superexponential epidemic spread. Since the result is obtained for any $a > 0$ and that $\pi(a) \underset{a\to\infty}{\to} 1$, the invasion probability is asymptotically 1.

Now we deal with a sketch of the proof of Theorem 1.1.6:

- **Theorem 1.1.6 (i):** Such heuristic does not differ from the one for the complete graph.
- **Theorem 1.1.6 (ii):** Consider the critical scaling $v_N \sim a\sqrt{d_N}$ for some $a > 0$. For the derivation of the upper bound on the survival probability, we couple (as in the case of the complete graph) the total infection process $\overline{\mathcal{I}}^{(N)}$ with a DBPC with offspring and cooperation distributions that are close to Poisson$\left(\frac{a^2}{2}\right)$ and Poisson$\left(a^2\right)$ respectively, until a certain level ℓ_N of total infection is reached, for a sequence $\ell_N \to \infty$ sufficiently slowly for the coupling to hold. The offspring and cooperation distributions of the latter DBPC are sufficiently close to Poisson distributions with parameters $\frac{a^2}{2}$ and a^2 respectively, such that the sequence of the survival probabilities of these approximating DBPCs asymptotically converges to $\pi_c(a)$. Combining the coupling, the property that for a DBPC reaching a level tending to infinity implies survival with high probability, and the previous argument, the probability to reach the level ℓ_N for the total infection process is asymptotically upper bounded by $\pi_c(a)$. In case the level ℓ_N is reached for the total infection process $\overline{\mathcal{I}}^{(N)}$, we upper bound the invasion probability by 1, meaning that afterwards also the remaining hosts get infected.

  For the lower bound we couple with high probability $\overline{\mathcal{I}}^{(N)}$ from below with the total size of a DBPC until it remains constant or reaches the level N^{ε} for $0 < \varepsilon < \beta$ small enough for the coupling to hold. The offspring and cooperation distributions of the latter DBPC are

chosen sufficiently close to Poisson distributions with parameters $\frac{a^2}{2^{n+1}}$ and $\frac{a^2}{2^n}$ such that its survival probability is asymptotically equal to $\pi_c\left(\frac{a}{\sqrt{2^n}}\right)$, the one of the DBPC with offspring and cooperation distribution exactly Poisson distributed with parameters $\frac{a^2}{2^{n+1}}$ and $\frac{a^2}{2^n}$ respectively. Once N^ε many hosts are infected we show that with high probability after at most $\mathcal{O}(\log(\log(N)))$ many further generations the infection process expands by a distance $r_N(1-o(1))$ per generation. On the other hand the invasion time is lower bounded by $\frac{1}{2r_N}$, since parasites can move in any generation at most at a distance r_N and the infection starts to the closest vertex of the centre of the cube.

- **Theorem 1.1.6 (iii):** Under the supercritical scaling $\sqrt{d_N} \in o(v_N)$, again as for the complete graph, we couple with high probability from below the total infection process $\overline{\mathcal{I}}^{(N)}$ with the total size process of a GWP with approximately Poisson$\left(\frac{a^2}{2}\right)$ offspring distribution until N^ε hosts get infected or the parasite population dies out for any $0 < \varepsilon < \beta$ and any $a > 0$. In addition we can show that when the level N^ε is reached for the total infection process, there exists a ball of diameter r_N which contains at least $N^\varepsilon \log^{-1}(N)$ infected hosts. By choosing $\varepsilon > \beta/2$ we show that once the level N^ε is reached, after one more generation the remaining hosts in this ball get infected with high probability. Afterwards with high probability the infection expands by a distance $r_N(1-o(1))$ in every generation (similar as in Theorem 1.1.6 (ii)) until the remaining hosts are all infected. Combining that for any GWP the probability to reach a level tending to infinity is asymptotically equal to its survival probability, and that the survival probability of the latter GWP is asymptotically equal to $\pi(a)$, we obtain that the invasion probability is lower bounded by $\pi(a)$ for any $a > 0$. Hence, it is asymptotically 1 since $\pi(a) \underset{a\to\infty}{\to} 1$.

Results on Discrete Branching Processes with Cooperation

For proving Theorem 1.1.5 and Theorem 1.1.6 we are deriving some useful results on DBPCs that are interesting enough to be stated below. Remember the notation we introduced in Definition 1.1.4. We start with the extinction-explosion principle, which is well-known for GWP and also holds for DBPCs.

Proposition 1.1.7. *(Extinction-explosion principle for DBPCs) Let* $\mathbf{Z}$ *be a DBPC satisfying* $p_{1,o} \neq 1$ *and* $(p_{0,o}, p_{1,c}) \neq (1,1)$. *Then there exists a nullset* $\mathcal{N}$ *such that*

$$\{\forall g \in \mathbb{N}_0 : Z_g > 0\} \subseteq \{\forall i \in \mathbb{N}, \exists g_0 \in \mathbb{N}_0, \forall g \geq g_0 : Z_g \geq i\} \cup \mathcal{N}.$$

The conditions $p_{1,o} \neq 1$ and $(p_{0,o}, p_{1,c}) \neq (1,1)$ are necessary to exclude two pathological cases where the statement does obviously not hold. The first case is where the process stays constant at one individual in each generation, and the second condition ensures that the DBPC with three or less individuals can further increase and is not stuck below three.

In contrast to GWPes, a DBPC always has a positive survival probability, except for pathological cases. We denote the expectation and the variance of the offspring distribution by μ_o and ν_o and for the cooperation distribution by μ_c and ν_c.

Proposition 1.1.8. *Let* $\mathbf{Z}$ *be a DBPC with* $\mu_o, \mu_c, \nu_o, \nu_c \in (0,\infty)$ *and* $p_{0,o} + p_{1,o} < 1$. *Suppose* $Z_0 = x \in \mathbb{N}$, *then* $\mathbf{Z}$ *has a positive survival probability, i.e.*

$$\mathbb{P}_x(Z_g > 0 \,\forall\, g > 0) > 0.$$

The superexponential growth speed of a DBPC in case of survival can be derived. More precisely, the next Proposition shows that for any sequence $b_N \to \infty$, reaching the level b_N or dying out for a DBPC, or a sequence of DBPCs under the next conditions, is at most of order $\log\log(b_N)$, which contrasts with GWPes. Let $\left(\mathbf{Z}^{(N)}\right)_{N\in\mathbb{N}_0}$ be a sequence of DBPCs for which $\mu_o^{(N)}, \mu_c^{(N)}, \nu_o^{(N)}, \nu_c^{(N)}$ denote the expectations and the variances of the offspring and cooperation distributions. We assume that

$$\mu_o^{(N)}, \mu_c^{(N)}, \nu_o^{(N)}, \nu_c^{(N)} \underset{N\to\infty}{\longrightarrow} \mu_o, \mu_c, \nu_o, \nu_c \in (0,\infty), \tag{1.1.2}$$

and we introduce $\mu := \min(\mu_o, \mu_c) > 0$ and $\nu := \max(\nu_o, \nu_c) > 0$.

Proposition 1.1.9. *Let $\left(\mathbf{Z}^{(N)}\right)_{N\in\mathbb{N}_0}$ be a sequence of DBPCs which satisfy Equation (1.1.2). Assume $Z_0^{(N)} = x$ for some $x \in \mathbb{N}$. Furthermore, assume that there exists an $N_0 > 0$ such that*

$$\inf_{N\geq N_0} \mathbb{P}\left(Z_1^{(N)} > L | Z_0^{(N)} = x\right) > 0,$$

where $L =: \lceil \mu^{-1}(8+\nu)^2 \rceil$. Let $(b_N)_{N\in\mathbb{N}}$ be a $\mathbb{N}$-valued sequence with $b_N \underset{N\to\infty}{\to} \infty$ and denote by

$$\tau_{b_N,0} := \inf\left\{ g \in \mathbb{N} : Z_g^{(N)} \geq b_N \text{ or } Z_g^{(N)} = 0 \right\}.$$

Then there exists a constant $C > 0$ such that

$$\mathbb{P}_x\left(\tau_{b_N,0} \leq C \log\log(b_N)\right) \underset{N\to\infty}{\longrightarrow} 1.$$

The following Proposition states that, as for GWPes, when a DBPC, or a sequence of DBPCs under the latter conditions, reaches a certain level tending to infinity at some generation or up to some generation then survival is ensured with high probability, the counterpart of Proposition 1.1.3 for DBPCs.

Proposition 1.1.10. *Consider a sequence of DBPC $\left(\mathbf{Z}^{(N)}\right)_{N\in\mathbb{N}}$ with offspring and cooperation distributions $\left(p_{k,o}^{(N)}\right)_{k\in\mathbb{N}_0}$ and $\left(p_{k,c}^{(N)}\right)_{k\in\mathbb{N}_0}$ respectively, which satisfies Equation (1.1.2). Furthermore, let $\mathbf{Z}^\infty$ be a DBPC with offspring and cooperation distribution $(p_{k,o})_{k\in\mathbb{N}_0}$ and $(p_{k,c})_{k\in\mathbb{N}_0}$. Assume that $p_{k,o}^{(N)} \to p_{k,o}$ and $p_{k,c}^{(N)} \to p_{k,c}$ as $N \to \infty$ for all $k \geq 0$. Then for any $\mathbb{N}$-valued sequence $(b_N)_{N\in\mathbb{N}}$ with $b_N \to \infty$ it holds that*

$$\begin{aligned}
\lim_{N\to\infty} \mathbb{P}\left(\forall g \in \mathbb{N}_0 : Z_g^{(N)} > 0\right) &= \lim_{N\to\infty} \mathbb{P}\left(\exists g \in \mathbb{N}_0 : Z_g^{(N)} \geq b_N\right) \\
&= \lim_{N\to\infty} \mathbb{P}\left(\exists g \in \mathbb{N}_0 : \overline{Z}_g^{(N)} \geq b_N\right) \\
&= \pi,
\end{aligned}$$

where π denotes the survival probability of $\mathbf{Z}^\infty$.

1.1.5 Perspectives

Many refinements on the understanding of the invasion probability of cooperated parasites on structured host populations are still possible. First, we obtain that, at the critical scaling $v_N \in \Theta\left(\sqrt{d_N}\right)$, the invasion probability is identified as the survival probability $\varphi(a)$ of a Galton-Watson process with $\mathrm{Pois}\left(\frac{a^2}{2}\right)$ offspring distribution when the host population is structured on a configuration model (see Theorem 1.1.2, taking $x = 0$), whereas we obtain a lower bound when it is structured on a random geometric graph as the survival probability of a DBPC whose offspring and cooperation distributions are $\mathrm{Pois}\left(\frac{a^2}{2^{n+1}}\right)$ and $\mathrm{Pois}\left(\frac{a^2}{2^n}\right)$ distributed (see Theorem 1.1.6). Hence direct comparison of such results can not be achieved since the offspring distributions of the limiting processes are not the same. Nevertheless, one can show in the case of the random geometric graph that the invasion probability is also lower bounded by $\varphi(a)$. To get this lower bound, it suffices to neglect cooperation of parasites generated by different infected hosts at the beginning of the infection process. It implies that having a finite-dimensional spatial structure of the population increases the invasion probability compared to scenario where the host population is structured by a configuration model leading to an infection graph with a locally tree like structure.

Moreover, adapting the proof techniques for deriving the lower bound in the case of the random geometric graph, one can show that the invasion probability is lower bounded by the survival probability of 2^n independent DBPCs with offspring and cooperation distributions that are $\mathrm{Pois}\left(\frac{a^2}{2^{n+1}}\right)$ and $\mathrm{Pois}\left(\frac{a^2}{2^n}\right)$ distributed. To get such result, it suffices to consider the sub-epidemic processes located on the 2^n non-intersecting balls of radius $\frac{r_N}{2}$ that are contained in the ball of radius r_N centred around the initially infected vertex. One can apply for each of the 2^n balls similar reasoning as made in Chapter 4 when deriving the lower bound, by coupling from below each of the 2^n sub-epidemic processes with the total sizes of DBPCs whose offspring and cooperation distributions are close to $\mathrm{Pois}\left(\frac{a^2}{2^{n+1}}\right)$ and $\mathrm{Pois}\left(\frac{a^2}{2^n}\right)$ distributions respectively. One has also to show that the couplings can be made independently from each others by controlling interaction events between the sub-epidemic processes. The last step is to prove that asymptotically the survival probability of the sum of these lower DBPCs is identified as the survival probability of a sum of 2^n independent DBPCs with exactly $\mathrm{Pois}\left(\frac{a^2}{2^{n+1}}\right)$ and $\mathrm{Pois}\left(\frac{a^2}{2^n}\right)$ offspring and cooperation distributions, by application of Proposition 1.1.10.

Another perspective concerning Chapter 3 when the host population is structured on the random geometric graph, is to derive the exact asymptotic invasion probability (instead of different upper and lower bounds) when the critical scaling $v_N \sim a\sqrt{d_N}$ is considered. Simulations of Figures 3.1 and 3.2 demonstrate that the invasion probability is much more closer to the upper bound compared to the lower bound. This can be easily understood because, in deriving the lower bound, we consider an epidemic model centred around the initially infected vertex, which considers infections only generated within the ball of radius $r_N/2$. This means that many infections are not considered in this process compared to the original process, resulting in an inadequate lower bound. In the model analysis, we show that the invasion probability is actually determined by what occurs at the beginning of the epidemic. Once a certain minor outbreak is reached (meaning that a certain negligible order of vertices is infected), global invasion is asymptotically achieved with probability 1. Therefore, to identify the invasion probability, one should study an approximate model that could be referred to as *branching with cooperation random walk*. Then, it should be shown that the spatial infection process behaves sufficiently close to this approximate process, at least for a

sufficient number of generations, such that the previous minor outbreak is obtained. More precisely, by branching with cooperation random walk, we mean that at each generation, each individual (or pair of individuals) gives birth to a random number of offspring, which are spatially distributed around the parent (or parents), according to a certain distribution and independently from each other. An example of study of such model is made in [33] where the authors considered a nearest neighbour cooperative branching-coalescing random walk on $\mathbb{Z}$.

Adding some dynamical behaviours to the hosts could be included in the model and being interesting for more biological relevance. For instance, in a framework at the intersection of the two biological themes of this thesis, such as modelling virotherapy with cooperative viruses. Indeed as mentioned in Subsection 1.1.1, it is believed that the infection of cancer cells with oncolytic viruses is more effective if a cancer cell is hit by several viruses simultaneously, because in this manner the virus can cope better with the (interferon-based) anti-viral response of the host, see [31]. Here, the parasites could represent viruses designed to recognise cancer cells, and the host population would be a mixture of cancerous and wild-type cells with interactions between them.

1.2 Cell dynamics of multitype populations in oncology

In the next subsections we are presenting the model studied for understanding the genetic composition during cancer evolution, motivated by some biological context, as well as the results obtained with heuristics, sketches of the proofs and some perspectives.

1.2.1 Biological context

Cancer is a genetic disease that results from accumulation of subsequent *driver* mutations. They trigger activation of oncogenes conferring a selective growth advantage to tumour cells [34] resulting in abnormal growth. Other mutations called *neutral* or *passenger* are also present in cancer evolution. Contrary to the driver ones, they do not provide any selective advantage to the growing cancer population but are retained either because they are necessary mutations to later obtain a selective one, such as inactivation of the first copy of a tumour suppressor gene, or either by chance during repeated rounds of cell division and clonal expansion. The number of driver mutations triggering cancer formation is heterogeneous across cancer types. For instance, in the case of solid tumours, typically more than one driver mutation is required for the development of malignancy, while a single genetic alteration may be sufficient to cause certain types of leukaemia [35].

Advanced sequencing technology applied to protein-coding genes has led to the identification of specific driver genes implicated in carcinogenesis, including oncogenes, tumour suppressor genes, and DNA repair genes. For instance, in colorectal cancer, commonly mutated driver genes such as APC, TP53, and KRAS are noteworthy examples [36, 37, 38]. Similarly, the fusion gene BCR-ABL has been linked to chronic myeloid leukaemia [39]. These studies, based on statistical analysis, have also found that tumours mostly contain a large number of neutral mutations compared to the number of selective ones driving the initiation, progression and maintenance of the tumour. Historically tumour dynamics has only been seen under the prism of clonal expansion of selective mutations, without even considering the paradigm of neutral cancer evolution. Such paradigm, which has been considered only recently, see [40, 41, 3, 42, 43], advocates that the driver mutations are already present in the initial cell and that the occurring mutations are neutral ones. Indeed, the genetic heterogeneity inside a tumour could be explained only considering neutral mutations.

Various statistical methods are developed to infer the evolutionary history of tumours, including test of neutral evolution, see [44, 45, 46] for details about that.

Understanding the evolutionary trajectory of a tumour holds clinical significance since prognosis hinges on its future evolutionary path, while the effectiveness of treatment is influenced by the emergence of resistant subpopulations. Hence, some of the key questions in cancer research involve uncovering the identities and the effects of mutations on tumorigenesis. Subsequently considerable effort has been invested to quantitatively understand evolutionary dynamics in exponentially expanding populations, in order to answer the typical questions: when, how and how may. More precisely, how many cells exist with a set of mutations, via which sequence of steps is it most likely to emerge and how long does it take for these cells to appear? With this aim of characterising disease progression, an increasing body of probabilistic research has been developed providing valuable insights on cancer genetic evolution. More specifically on clonal expansion [47, 48, 22, 49, 50, 6, 23, 24, 8, 10, 51, 52, 53, 54, 55], on assessing recurrence timing and the genetic composition of a tumour at recurrence timing [56, 18, 19, 20], on drug resistance [57, 58, 59], and on understanding a classical summary statistics for the genetic composition of a tumour that is the *site frequency spectrum* [60, 52, 46, 61, 18, 62, 63, 64].

In Chapter 4 and Chapter 5, we propose a multitype birth and death branching process model on a general finite trait space to derive quantitative results on cancer evolution. More precisely, under neutral (and deleterious) cancer evolution, we obtain first-order asymptotics for the mutant subpopulation sizes capturing stochasticity on the genetic composition and allowing to characterise effective evolutionary pathways. These results are given in Chapter 4. Concerning the selective case, the study is made up to the first selective mutant trait and is given in Chapter 5. Generalisation to derive the first-order asymptotics when considering selective mutations on a general finite directed labelled graph is a work in progress in collaboration with Hélène Leman.

1.2.2 An individual-based toy model for carcinogenesis

The biological motivation of the model under study in Chapter 4 and Chapter 5 is to capture the dynamics over time of the genetic composition of a population of cells during carcinogenesis. Mathematically, the first-order asymptotics of all the different mutant subpopulations is captured over time for neutral and deleterious cancer evolution in Chapter 4 and up to the first selective mutant subpopulation in Chapter 5.

Cells are represented by a phenotypic trait, modelled through a finite set V with $0 \in V$. Cells with trait 0 are called *wild-type cells*, and all cells with trait $v \in V \backslash \{0\}$ are called *mutant cells*. For all $v \in V$, we denote by $(Z_v(t))_{t\in\mathbb{R}^+}$ the number of cells of trait v at time t in the population. For modelling carcinogenesis, we consider the initial condition

$$\forall v \in V, Z_v(0) = \mathbb{1}_{\{v=0\}}, \text{ almost surely.}$$

The evolutionary process is modelled using a continuous-time branching process on $\mathbb{N}_0^V$, where cell division, death and mutations are taking into account. More precisely, cells divide to give birth to two daughter cells and die with rates depending only on their phenotypic trait. The birth, death and growth rate functions are respectively

$$\begin{aligned} \alpha &: V \longrightarrow \mathbb{R}^+, \\ \beta &: V \longrightarrow \mathbb{R}^+, \\ \lambda &:= \alpha - \beta. \end{aligned}$$

The growth rate of the wild-type subpopulation is assumed to be strictly positive $\lambda(0) > 0$, otherwise the wild-type subpopulation won't survive almost surely. During a division event of a cell of trait $v \in V$, independent mutations over the two daughter cells are considered. Mutation landscape across traits is encoded via a graph structure (V, E) on the trait space. $E \subset \{(v,u), \forall v, u \in V^2\}$ is a set of ordered pairs over V satisfying for all $v \in V$, $(v,v) \cap E = \emptyset$, and such that for all trait v there exists a path from 0 to v. In other words, (V, E) is a finite directed graph without self-loop for which each vertex is on the directed-connected component of trait 0. Mutation from trait v to trait u is possible if and only if $(v,u) \in E$. Let $\mu : E \longrightarrow [0,1]$ be a mutation kernel satisfying

$$\forall v \in V, \overline{\mu}(v) := \sum_{u \in V:(v,u)\in E} \mu(v,u) \leq 1.$$

A daughter cell mutates from trait v to trait u with probability $\mu(v,u)$, meaning that $\overline{\mu}(v)$ is its total mutation probability.

Finally the exact transition rates from a state $z = (z_v)_{v\in V} \in \mathbb{N}_0^V$ of the evolutionary process are

$$z \mapsto \begin{cases} z - \delta_v, \text{ at rate } z_v \beta(v), \\ z - \delta_v + \delta_u + \delta_w, \text{ at rate } 2 z_v \alpha(v) \mu(v,u) \mu(v,w) \mathbb{1}_{\{(v,u)\in E\}} \mathbb{1}_{\{(v,w)\in E\}} \mathbb{1}_{\{u \neq w\}}, \\ z - \delta_v + 2\delta_u, \text{ at rate } z_v \alpha(v) \mu(v,u)^2 \mathbb{1}_{\{(v,u)\in E\}}, \\ z + \delta_v, \text{ at rate } z_v \alpha(v) \left(1 - \overline{\mu}(v)\right)^2 + 2 \sum_{u \in V:(u,v)\in E} z_u \alpha(u) \mu(u,v) \left(1 - \overline{\mu}(u)\right), \end{cases}$$

where $\forall v \in V, \delta_v = \left(\mathbb{1}_{\{u=v\}}\right)_{u\in V}$.

Following typical parameter values in cancer evolution, we consider the framework of *power law mutation rates limit*. A scaling parameter $n \in \mathbb{N}$ is used to quantify both the decrease of the mutation probabilities, as negative powers of n, and also the typical size of the population, depending on n as positive power of n, at which we are interested in understanding the genetic composition. We aim to obtain first-order asymptotics of all mutant subpopulation sizes when n goes to infinity. To be more precise, let $L := \{\ell(v,u) \in \mathbb{R}_+^*, \forall (v,u) \in E\}$ be a set of strictly positive labels on the edges of the graph. Introduce a sequence of the latter model where for each $n \in \mathbb{N}$, the mutation kernel under consideration is $\mu^{(n)} : E \longrightarrow [0,1]$ satisfying

$$\forall (v,u) \in E, n^{\ell(v,u)} \mu^{(n)}(v,u) \underset{n\to\infty}{\longrightarrow} \mu(v,u) \in \mathbb{R}^+. \tag{1.2.1}$$

We are going to study the cell populations on the $\log(n)$-accelerated time scale, the natural one for exponentially growing models to reach powers of n, as well as the random time scales at which the wild-type subpopulation or the total cell population reach powers of n.

1.2.3 Results of Chapter 4

In Chapter 4 we obtain first-order asymptotics of the size of the mutant subpopulations on the deterministic $\log(n)/\lambda(0)$-accelerated time scale, and on the random time scale at which the total population, respectively the wild-type subpopulation, reaches powers of n, for neutral and delete-

rious cancer evolution, see Theorem 1.2.4. For that, introduce for $t > 0$ the time scales

$$\begin{aligned}
\mathfrak{t}_t^{(n)} &:= t\frac{\log(n)}{\lambda_0},\\
\eta_t^{(n)} &:= \inf\left\{u \in \mathbb{R}^+ : Z_0^{(n)}(u) \geq n^t\right\},\\
\sigma_t^{(n)} &:= \inf\left\{u \in \mathbb{R}^+ : \sum_{v\in V} Z_v^{(n)}(u) \geq n^t\right\},\\
\rho_t^{(n)} &:= \eta_t^{(n)} \text{ or } \sigma_t^{(n)}.
\end{aligned} \tag{1.2.2}$$

Under neutral and deleterious cancer evolution, the stopping times $\eta_t^{(n)}$ and $\sigma_t^{(n)}$ are asymptotically the same (see Proposition 4.3.12). For simplicity, when stating the results, we introduce the notation $\rho_t^{(n)}$. Two different biological interpretations can be made to motivate the study on the random time scales. For instance, when considering metastasis, the wild-type subpopulation $Z_0^{(n)}$ may represent the primary tumour, and the mutant subpopulations $Z_v^{(n)}$, for all $v \in V\backslash\{0\}$, may correspond to secondary tumours. As it is size and not age of a tumour that clinicians have access to, it is biologically relevant to estimate the genetic composition of the secondary tumours when the primary one has a given size. This is mathematically encoded in looking at the first-order asymptotics of $Z_v^{(n)}\left(\eta_t^{(n)}\right)$ for all $v \in V\backslash\{0\}$. Another biological setting is when the total population $Z_{tot}^{(n)}$ represents one tumour. It is appropriate to obtain theoretical results about the size of the mutant subpopulations $Z_v^{(n)}$ for all $v \in V\backslash\{0\}$ when the tumour has reached a given size. This corresponds exactly to looking at the first-order asymptotics of $Z_v^{(n)}\left(\sigma_t^{(n)}\right)$.

We also obtain asymptotic results for the stochastic exponents (or also called stochastic Hopf-Cole transform) of the mutant subpopulations, defined as

$$X_v^{(n)}(t) := \frac{Z_v^{(n)}\left(\mathfrak{t}_t^{(n)}\right)}{\log(n)/\lambda(0)}, \tag{1.2.3}$$

without imposing the framework of neutral or deleterious cancer evolution, see Theorem 1.2.5.

We are going to perform some heuristics for the results on the first-order asymptotics under the assumption

$$\forall v \in V, \lambda(v) \leq \lambda(0). \tag{1.2.4}$$

Heuristics

The heuristics for understanding Theorem 1.2.4 are constructed in three steps. The first one aims to understand the first-order asymptotics of the size of a direct neighbour of the wild-type subpopulation by studying the case of the simplest graph one can think of, i.e. a two-traits model where only mutations from wild-type to the one and only mutant trait is considered. Then, we explain how such heuristics on a two-traits model is used as an elementary brick for getting heuristics on a finite monodirectional graph. Finally, we end by quantitatively comparing for all traits $v \in V\backslash\{0\}$ the asymptotic contribution to the subpopulation size of trait v of all the different walks on the trait space graph structure linking trait 0 to v :

- **Heuristics for a two-traits model:** Consider the two-traits model (V, E, L) with $V = \{0,1\}$, $E = \{(0,1)\}$ and $L = \{\ell(0,1)\}$. We start the heuristics by deriving that the first time on the $\log(n)/\lambda(0)$-accelerated time scale for mutations to be likely is actually $\ell(0,1)$. Under the power law mutation rates regime, the inner birth and death rate of the wild-type subpopulation are so close to $\alpha(0)$ and $\beta(0)$ (only differs by a factor of order $n^{-\ell(0,1)}$) that the growth of the wild-type subpopulation is approximately an exponential growth with rate $\lambda(0)$. Then in case of survival, at time $\mathfrak{t}_t^{(n)}$, its size is of order $\Theta(n^t)$. With a mutation probability which scales as $n^{-\ell(0,1)}$, the total mutation probability up to time $\mathfrak{t}_t^{(n)}$ scales as $\int_0^t n^u n^{-\ell(0,1)} d\left(u\frac{\log(n)}{\lambda(0)}\right) = \frac{n^{-\ell(0,1)}}{\lambda(0)}(n^t - 1)$ which is of order 1 for $t = \ell(0,1)$. This heuristics is formalised by D. Cheek and T. Antal in [8, 10].

  Then we explain how to predict the size of the mutant subpopulation at time $\mathfrak{t}_t^{(n)}$, for all $t \geq \ell(0,1)$. An illustration of these heuristics can be found in Figure 1.5. Let $\ell(0,1) \leq u \leq t$, the number of new mutations generated at time $\mathfrak{t}_u^{(n)}$ scales as

$$\Theta(n^u \cdot n^{-\ell(0,1)}) = \Theta\left(\exp\left(\lambda(0)(u - \ell(0,1))\frac{\log(n)}{\lambda(0)}\right)\right).$$

  The remaining time for these new mutant cells to grow exponentially at rate $\lambda(1)$ until time $\mathfrak{t}_t^{(n)}$ is $\mathfrak{t}_{t-u}^{(n)}$. This implies that their lineages have at time $\mathfrak{t}_t^{(n)}$ a size of order

$$\begin{aligned}\Theta\left(\exp\left(\lambda(0)(u - \ell(0,1))\frac{\log(n)}{\lambda(0)}\right) \cdot \exp\left(\lambda(1)(t-u)\frac{\log(n)}{\lambda(0)}\right)\right) \\ = \Theta\left(\exp\left(\left[\lambda(1)t + (\lambda(0) - \lambda(1))u - \lambda(0)\ell(0,1)\right]\frac{\log(n)}{\lambda(0)}\right)\right).\end{aligned} \tag{1.2.5}$$

  Then two scenarios are possible depending on if the mutant trait is neutral or deleterious:

  (i) If $\lambda(1) < \lambda(0)$: Equation (1.2.5) is maximal for $u = t$ and equal to $\Theta\left(n^{t-\ell(0,1)}\right)$. This means that the dynamics of the mutant subpopulation is driven by the mutation from the wild-type subpopulation and not from its inner growth. More precisely, its size order at time $\mathfrak{t}_t^{(n)}$ is fully given by the mutations generated at this time -and so is of order $\mathcal{O}\left(n^{t-\ell(0,1)}\right)$- and not from the lineages issued from mutations generated at strictly previous time.

  (ii) If $\lambda(1) = \lambda(0)$: Equation (1.2.5) is independent of u and equal to $\Theta\left(n^{t-\ell(0,1)}\right)$ for any $\ell(0,1) \leq u \leq t$. This means that these lineages have the same size order at time $\mathfrak{t}_t^{(n)}$ than any other lineages of mutant cells generated from mutational events at any other time between $\mathfrak{t}_{\ell(0,1)}^{(n)}$ and $\mathfrak{t}_t^{(n)}$. In the dynamics of the mutant subpopulation there is a balance between the contribution of mutations and its inner growth. This is a consequence of assuming $\lambda(1) = \lambda(0)$. Hence to capture the total size of the mutant subpopulation at time $\mathfrak{t}_t^{(n)}$, it remains to integrate all the lineages issued from mutational events over time $\mathfrak{t}_u^{(n)}$, for $\ell(0,1) \leq u \leq t$. This exactly gives the order $\Theta\left((t - \ell(0,1))\log(n)n^{t-\ell(0,1)}\right)$.

  To sum up, for this simple graph, the mutant subpopulation scales after time $\mathfrak{t}_{\ell(0,1)}^{(n)}$ as

$$\Theta\left(n^{t-\ell(0,1)}\left[\mathbb{1}_{\{\lambda(0)>\lambda(1)\}} + \mathbb{1}_{\{\lambda(0)=\lambda(1)\}}(t - \ell(0,1))\log(n)\right]\right). \tag{1.2.6}$$

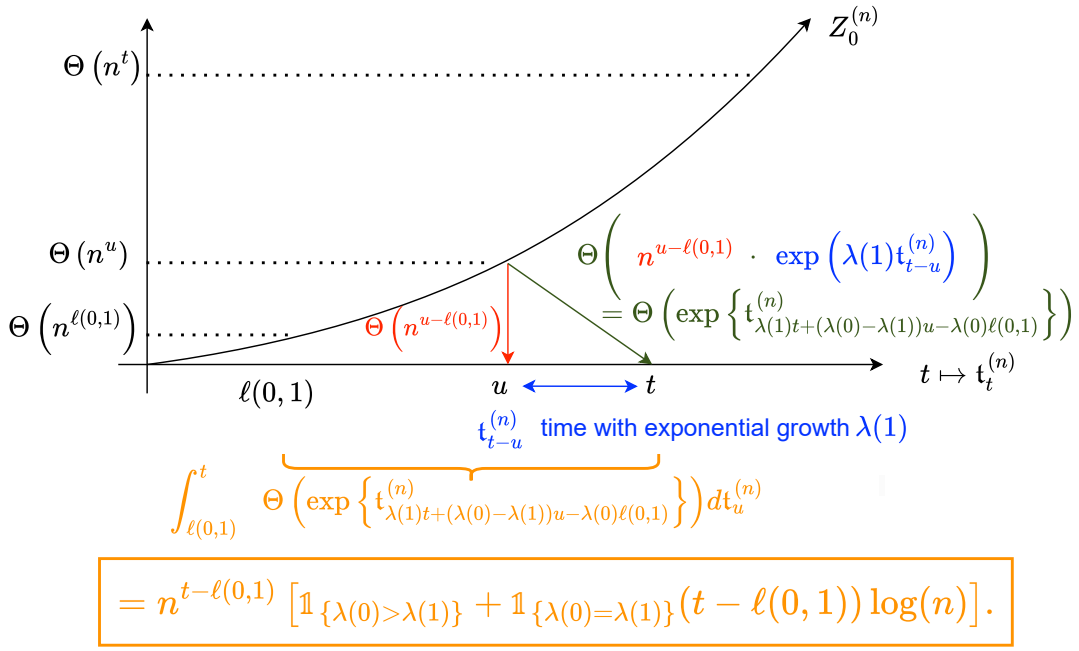

Figure 1.5: Heuristics for the size of the mutant subpopulation after time $\mathfrak{t}_{\ell(0,1)}^{(n)}$

In any case, the mutant subpopulation has an exponential growth at rate $\lambda(0)$ after time $\mathfrak{t}_{\ell(0,1)}^{(n)}$, given by the factor $n^{t-\ell(0,1)}$. A supplementary multiplicative factor of order $\log(n)$ is captured in case of neutral mutation.

- **Heuristics for a finite mono-directional graph:** These heuristics on this simple two-traits model can be used as an elementary brick for getting some on a finite monodirectional graph by iterations to get both the first-occurrence time for seeing mutations in the $\log(n)/\lambda(0)$-accelerated time scale, as well as the exponential growth and the multiplicative factor of $\log(n)$. Consider a path from wild-type trait 0 to a trait $v \in V\backslash\{0\}$. We obtained that after the first-occurrence time for mutations to be generated for the neighboured trait of the wild-type trait, which is the label on the edge, this mutant subpopulation has an exponential growth at rate $\lambda(0)$. Thus the time u to wait for seeing a cell of trait v, on the time scale $\mathfrak{t}_u^{(n)}$, generated due to this specific mutational walk is the sum of the labels of the edges of this walk, called the *length* of this walk. Then, after this time, this subpopulation of cells of trait v due to this walk grows exponentially fast at rate $\lambda(0)$. Moreover, as seen in (1.2.6), when a neutral mutation is encountered a multiplicative factor of order $\log(n)$ is captured in the asymptotic size, meaning that for any neutral mutation on the walk leading to v a supplementary multiplicative factor of order $\log(n)$ is captured on the size order. These two facts combined give that after time the length of this walk, the subpopulation of trait v grows exponentially fast at rate $\lambda(0)$ and has a multiplicative factor $\log(n)$ to the power the number of neutral mutations there are on this walk.

- **Heuristics for a general finite directed labelled graph:** Considering a vertex $v \in V\backslash\{0\}$, there are potentially many mutational walks from the initial vertex 0 to v on the trait space graph structure. Then one needs to understand which ones are involved in the size order of the mutant subpopulation of trait v, by using the previous heuristics on a monodirectional graph. First, the subpopulation of trait v starts having cells after a time which is the minimum of the lengths over the walks from 0 to v. Second, after this time only the walks whose lengths are equal to the latter minimum might contribute to the size order of the mutant cells of trait v. This is due to the fact that having a time delay, on the $\log(n)/\lambda(0)$-accelerated time scale, creates an exponential delay of order a power of n in the size order. This discrimination over the walks that might asymptotically contribute is asymptotically captured in Theorem 1.2.5 where results on the asymptotic limits of the stochastic exponents suffice. Thirdly, the supplementary multiplicative factors of order $\log(n)$ due to the neutral mutations implies that over the walks from 0 to v satisfying that their lengths are equal to the latter minimum, only those with the maximal number of neutral mutations are actually contributing to the size order of the mutant subpopulation of trait v. More specifically with a factor of $\log(n)$ at the power this maximal number of neutral mutations. Such characterisation is asymptotically captured by results on the first-order asymptotics, presented in Theorem 1.2.4, where results only in the stochastic exponents are not sufficient. Moreover for any of these admissible walks, at each neutral mutation a supplementary time integral is obtained, as seen in Equation (1.2.6). An illustration with an example is given in Figure 1.6.

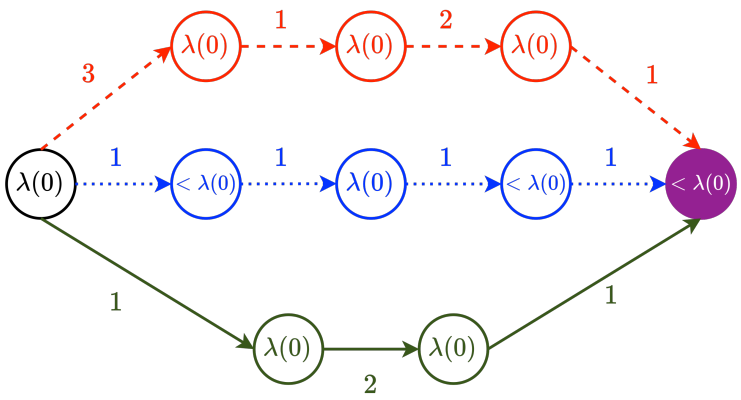

Figure 1.6: Heuristics for the contribution of walks to the size order of the plain purple mutant subpopulation: in this example, the dashed red walk has a length of 7, while the dotted blue and plain green walks have a length of 4. Therefore, only the two latter walks may contribute to the size order of the plain purple mutant subpopulation, making them sub-admissible walks. However, the dotted blue walk has only one neutral mutation, whereas the plain green walk has two neutral mutations. As a result, only the plain green walk will ultimately contribute to the size order of the purple mutant subpopulation. For $t \geq 4$, at time $\mathfrak{t}_t^{(n)}$, it will grow as $\log^2(n)e^{\lambda(0)\mathfrak{t}_{t-4}^{(n)}}$. Notice, in particular, that the dashed red walk has the maximal number of neutral mutations, which is 3. However, since it is not a sub-admissible walk, the multiplicative factor of $\log(n)$ remains 2 instead of 3.

Definitions

Now, the natural definitions derived from these heuristics are formally established before presenting the results.

Definition 1.2.1 (Deleterious and neutral vertices)**.** *A vertex $v \in V$ is called a neutral vertex if $\lambda(v) = \lambda(0)$, and a deleterious vertex if $\lambda(v) < \lambda(0)$.*

The following definition provides a structured framework to analyse the contribution of evolutionary pathways to the growth of mutant subpopulations. It does so by introducing the adapted vocabulary, for the neutral and deleterious evolutionary context of the model, associated with walks in labelled graphs. We use the term *'walk'* here according to the standard nomenclature of graph theory.

Definition 1.2.2 (Walk in the graph)**.** *A walk $\gamma = (v(0), \cdots, v(k))$ in the graph (V, E) is defined as a sequence of vertices linking $v(0)$ to $v(k)$ such that for all $0 \leq i \leq k, v(i) \in V$, and for all $0 \leq i \leq k-1, (v(i), v(i+1)) \in E$. We will sometimes use the term 'path' to refer to a walk that visits only distinct vertices. Given a walk $\gamma = (v(0), v(1), \cdots v(k))$ in the labelled graph (V, E, L), we define:*

- *The sum of the labels of the edges and the sum over the first i edges of the walk γ, respectively:*

$$t(\gamma) := \sum_{i=0}^{k-1} \ell(v(i), v(i+1)) \text{ and for all } i \leq k, t_\gamma(i) := \sum_{j=0}^{i-1} \ell(v(j), v(j+1)).$$

- *The subset of neutral heads of the edges of the walk γ and its cardinality:*

$$\gamma_{neut} = \{v(i), 1 \leq i \leq k : \lambda(v(i)) = \lambda(0)\} \text{ and } \theta(\gamma) := |\gamma_{neut}|.$$

- *The weights $w_{neut}(\gamma)$ and $w_{del}(\gamma)$ associated with the neutral and deleterious vertices of the walk γ, respectively:*

$$w_{neut}(\gamma) := \prod_{1 \leq i \leq k, \lambda(v(i)) = \lambda(0)} \frac{2\alpha(v(i-1))\mu(v(i-1), v(i))}{\lambda(0)},$$
$$w_{del}(\gamma) := \prod_{1 \leq i \leq k, \lambda(v(i)) < \lambda(0)} \frac{2\alpha(v(i-1))\mu(v(i-1), v(i))}{\lambda(0) - \lambda(v(i))}.$$

  Along a walk, the constant of the asymptotic contribution of a vertex- depending on its parameters and those of the upstream vertex- takes a distinct form based on whether the vertex is neutral or deleterious. This distinction motivates the use of the separate weights $w_{neut}(\gamma)$ and $w_{del}(\gamma)$.

- *The time dependence associated with the neutral vertices: Let σ be an increasing function from $\{1, \cdots, \theta(\gamma)\}$ to $\{1, \cdots, k\}$, such that $v(\sigma_i)$ is the i-th neutral vertex of the walk γ. For all $t > 0$, define the multiple integral $I_\gamma(t)$ as*

$$I_\gamma(t) := \int_{t_\gamma(\sigma_{\theta(\gamma)})}^{t \vee t_\gamma(\sigma_{\theta(\gamma)})} \int_{t_\gamma(\sigma_{\theta(\gamma)-1})}^{u_1} \cdots \int_{t_\gamma(\sigma_{\theta(\gamma)-k})}^{u_k} \cdots \int_{t_\gamma(\sigma_1)}^{u_{\theta(\gamma)-1}} du_{\theta(\gamma)} \cdots du_1.$$

Along a walk, for each neutral vertex that is visited, an additional integral over the time parameter appears in the asymptotic limit, as described in the heuristics. This motivates the definition of $I_\gamma(t)$.

- *The weight of the walk γ at time t:*

$$w_\gamma(t) := w_{del}(\gamma) w_{neut}(\gamma) I_\gamma(t).$$

This expression captures the total weight of a walk γ at time t, accounting for both the deleterious and neutral visited vertices, and the integrals over the time parameters associated with these neutral vertices.

Definition 1.2.3 (Admissible walks)**.** *For all $v \in V$, let $P(v)$ denote the set of all walks γ in the graph (V, E) that link the vertex 0 to the vertex v. We define the:*

- *The minimum total label sum among all walks from vertex 0 to vertex v:*

$$t(v) := \min_{\gamma \in P(v)} t(\gamma).$$

- *The maximum number of neutral vertices among the shortest walks from vertex 0 to vertex v:*

$$\theta(v) := \max_{\gamma \in P(v), t(\gamma) = t(v)} \theta(\gamma).$$

- *The set of admissible walks from vertex 0 to vertex v:*

$$A(v) := \{\gamma \in P(v) : t(\gamma) = t(v) \text{ and } \theta(\gamma) = \theta(v)\}.$$

In the previous definition, the set $A(v)$ is referred to as the set of admissible walks because, as indicated by the heuristics, only walks belonging to $A(v)$ contribute to the growth dynamics of the mutant subpopulation of trait v. This is formally established in Theorem 1.2.4.

Main results for neutral and deleterious cancer evolution

In the next theorem we state the first-order asymptotic results for the size of all the mutant subpopulations structured on a general finite directed labelled graph for the trait space, under neutral and deleterious cancer evolution.

Theorem 1.2.4. *Assume that the general finite directed labelled graph (V, E, L) satisfies both the power law mutation rates regime described in (1.2.1) and the non-increasing growth rate graph condition given in (1.2.4). Let $h_n = \frac{\log(n)}{\log(\log(n))\theta_{\max} + \varphi_n}$, where $\varphi_n \underset{n\to\infty}{\to} \infty$ such that $h_n \underset{n\to\infty}{\to} \infty$ and where $\theta_{max} := \max_{v \in V \setminus \{0\}} \theta(v)$. Let also ψ_n such that $\sqrt{\log(n)} = o(\psi_n)$. Define for all $(t, s) \in \mathbb{R}^+ \times \mathbb{R}$,*

$$\begin{aligned} d_v^{(n)}(t,s) := & \mathbb{1}_{\{t \in [0, t(v) - h_n^{-1})\}} + \mathbb{1}_{\{t \in [t(v) - h_n^{-1}, t(v))\}} \psi_n \log^{\theta(v)-1}(n) \\ & + \mathbb{1}_{\{t \in [t(v), \infty)\}} n^{t - t(v)} \log^{\theta(v)}(n) e^{\lambda(0) s}. \end{aligned} \tag{1.2.7}$$

Let $(T, M) \in \left(\mathbb{R}_+^*\right)^2$ *and* $0 < T_1 < T_2$. *Using the mathematical definition of the model given in Section 4.4 (see (4.4.1), (4.4.2), (4.4.3), (4.4.4) and (4.4.5)), there exists a random variable* W, *properly defined in (4.4.6), satisfying*

$$W \overset{law}{:=} Ber\Big(\frac{\lambda(0)}{\alpha(0)}\Big) \otimes Exp\Big(\frac{\lambda(0)}{\alpha(0)}\Big),$$

such that for all $v \in V\backslash\{0\}$, *we obtain the convergence results in probability in* $L^\infty([0,T]\times[-M,M])$ *for Equation (1.2.8) and in* $L^\infty([T_1,T_2]\times[-M,M])$ *for Equations (1.2.9), (1.2.10) and (1.2.11):*

(i) ***Deterministic time scale:***
If $\lambda(v) = \lambda(0)$, *then*

$$\frac{Z_v^{(n)}\Big(\mathfrak{t}_t^{(n)} + s\Big)}{d_v^{(n)}(t,s)} \underset{n\to\infty}{\longrightarrow} W \sum_{\gamma\in A(v)} w_\gamma(t). \tag{1.2.8}$$

If $\lambda(v) < \lambda(0)$, *then*

$$\frac{Z_v^{(n)}\Big(\mathfrak{t}_{t(v)+t}^{(n)} + s\Big)}{n^t \log^{\theta(v)}(n) e^{\lambda(0)s}} \underset{n\to\infty}{\longrightarrow} W \sum_{\gamma\in A(v)} w_\gamma(t(v)+t). \tag{1.2.9}$$

(ii) ***Random time scales:*** *Consider* $\left(\rho_t^{(n)}\right)_{t\in\mathbb{R}^+}$ *as defined in (1.2.2).*
If $\lambda(v) = \lambda(0)$, *then*

$$\frac{Z_v^{(n)}\Big(\rho_t^{(n)} + s\Big)}{d_v^{(n)}(t,s)} \underset{n\to\infty}{\longrightarrow} \mathbb{1}_{\{W>0\}} \sum_{\gamma\in A(v)} w_\gamma(t). \tag{1.2.10}$$

If $\lambda(v) < \lambda(0)$, *then*

$$\frac{Z_v^{(n)}\Big(\rho_{t(v)+t}^{(n)} + s\Big)}{n^t \log^{\theta(v)}(n) e^{\lambda(0)s}} \underset{n\to\infty}{\longrightarrow} \mathbb{1}_{\{W>0\}} \sum_{\gamma\in A(v)} w_\gamma(t(v)+t). \tag{1.2.11}$$

The proof of Theorem 1.2.4 relies on a martingale approach using Doob's and Maximal Inequalities. The initial step involves controlling the growth of the lineage of wild-type cells originated from the initial cell, for both the deterministic and random time scales. For any vertex $v \in V\backslash\{0\}$, there may be several mutational walks in the graph (V, E) that start from 0 and lead to v. Understanding the contribution of each of these walks to the first-order asymptotics of the size of the mutant subpopulation of trait v is essential. The proof proceeds in 2 steps. First, we consider an infinite mono-directional graph under Assumption (1.2.4) and we establish the result for this specific graph. Performing this step for an infinite graph is particularly helpful in handling cycles (such as backward mutations) in a general finite directed graph. Then, we identify and exclude walks from the initial vertex 0 to v that do not contribute to the first-order asymptotics of the size of the mutant subpopulation of trait v.

Before stating the result obtained on the stochastic exponent when allowing selective mutations, many remarks about Theorem 1.2.4 are performed:

(i) **Supplementary $\log(n)$ factor:** Notice that a multiplicative factor of $\log^{\theta(v)}(n)$ is captured after time $\mathfrak{t}^{(n)}_{t(v)}$, see Equations (1.2.7), (1.2.8), (1.2.9), (1.2.10) and (1.2.11). Obtaining a result on the stochastic exponents, as in Theorem 1.2.5, does not capture such a factor. For instance with the two-traits model used for the heuristics, if $\lambda(1) = \lambda(0)$, Theorem 1.2.4 gives that after time $\ell(0,1)$, $Z^{(n)}_1\big(\mathfrak{t}^{(n)}_t\big)$ behaves approximately as $\log(n)e^{\lambda(0)\mathfrak{t}^{(n)}_{t-\ell(0,1)}}$. However, what is captured with $X^{(n)}_1(t)$ after time $\ell(0,1)$ is asymptotically $\lambda(0)(t-\ell(0,1))$, see Theorem 1.2.5.

(ii) **Asymptotic stochasticity:** The random variable W is explicitly defined as the random variable that quantifies the randomness over the long time of the lineage of wild-type cells issued from the initial cell. Due to the power law mutation rates regime, mutations arise after a long time, so that the stochasticity of this lineage is already captured by W. Notice that under Assumption (1.2.4), the randomness in the first-order asymptotics of any mutant subpopulation size is described completely by W. This means that the stochasticity of these subpopulations is driven primarily by the randomness in the growth of the wild-type subpopulation rather than by the one of the mutational process or of any lineage of mutant cells. In particular, if the process starts with a large number of wild-type cells instead of just one, the first-order asymptotics of the size of the mutant subpopulations would be entirely deterministic.

(iii) **Selective mutations:** It seems quite natural not to obtain such a result when considering selective mutation ($\lambda(v) > \lambda(0)$). Indeed, a selective mutation imply that any time advantage translates directly into a growth advantage. Thus, the stochasticity of the mutational process, as well as the randomness in the lineages of the mutant cells, cannot be ignored. Therefore, expecting to control the stochasticity of the mutant subpopulation solely by controlling the randomness in the wild-type subpopulation, without also accounting for the randomness in the mutational process and the mutant lineages, is vain. More precisely, using a martingale approach to derive the first-order asymptotics cannot be successful for a selective mutation. Technically, this is because the expected size of the selective mutant subpopulation is of a higher order than its typical asymptotic size. Indeed, the rare event of the initial cell mutating to the selective trait extremely quickly, an event that asymptotically vanishes, is responsible for this discrepancy between the expected value and the typical asymptotic size of the selective mutant subpopulation. Nevertheless, when examining the stochastic exponent, the martingale approach allows us to obtain convergence results as given in Theorem 1.2.5. This is because the aforementioned rare event contributes only a factor proportional to its probability to the expected value of the stochastic exponent, meaning it actually asymptotically neither contributes to the typical size nor to the expected value of the stochastic exponent of the selective mutant subpopulation. Moreover, in Chapter 5 we develop another approach to get the first-order asymptotics for the first mutant trait. Generalisation to derive the first-order asymptotics when considering selective mutations for a general finite directed graph is a work in progress.

(iv) **Result comparison between neutral and deleterious traits:** Considering the time scale $\mathfrak{t}^{(n)}_t$, the result slightly differs depending on whether the vertex is neutral or deleterious. Indeed, when looking at the asymptotic behaviour for a deleterious vertex v, the result holds strictly after time $t(v)$, whereas, in the case of a neutral vertex, the entire trajectory from the initial time can be analysed. Mathematically, this difference arises from the additional

multiplicative factor of $\log(n)$ in the first-order asymptotics when considering a neutral mutation. This factor allows us to control the quadratic variation at time $t(v)$ for the martingale associated to the mutant subpopulation. Three distinct regimes are obtained, as indicated by (1.2.7) and (1.2.8) :

(i) Up to time $t(v) - h_n^{-1}$: with high probability, no mutational pathway from 0 to v has generated a mutant cell of trait v. Since $h_n \to \infty$ and satisfies $h_n = o(\log(n))$, $t(v)$ can be interpreted as the first time -when considering the time scale accelerated by $\log(n)$- at which it becomes asymptotically possible to observe the first occurrence of a mutant cell of trait v. This result is also true for deleterious mutations.

(ii) For $t \in \big[t(v) - h_n^{-1}, t(v)\big)$: in this time interval, some mutant cells of trait v are produced, but the interval's length is insufficient to achieve any power of n for the size of the mutant subpopulation of trait v. We succeed to dominate its growth by $\psi_n \log^{\theta(v)-1}(n)$, with a well-chosen ψ_n. Heuristically, the total number of mutant cells of trait v resulting from a mutational event up to time t is of order $\Theta\big(\log^{\theta(v)-1}(n)\big)$. With the remaining time for these mutant cells' lineages to grow, we manage to control the size of the mutant subpopulation of trait v by at most $\sqrt{\log(n)}\log^{\theta(v)-1}(n)$. Consequently, dividing by any function ψ_n satisfying $\sqrt{\log(n)} = o(\psi_n)$ results in an asymptotic limits of 0. This result also holds for deleterious mutations. The $\sqrt{\log(n)}$ factor in the growth control comes from a mathematical analysis using a martingale approach, particularly considering the time scale accelerated by $\log(n)$. With further refinement, we conjecture that the actual size of the mutant subpopulation at time $t(v)$ is of order $\Theta\big(\big(\mathbb{1}_{\{\lambda(0)=\lambda(v)\}}\log(\log(n)) + \mathbb{1}_{\{\lambda(0)>\lambda(v)\}}\big)\log^{\theta(v)-1}(n)\big)$.

(iii) For $t \in [t(v), \infty)$: with high probability, the number of mutant cells of trait v grows exponentially at rate $\lambda(0)$. A supplementary multiplicative factor $\log^{\theta(v)}(n)$ is present due to the neutral mutations on the walks in $A(v)$. Thus, the growth scales globally as $n^{(t-t(v))}\log^{\theta(v)}(n)w_v(t)$.

(v) **Results comparison between time scales:** When comparing point (i) and (ii) of Theorem 1.2.4, notice that the result transitions from the deterministic time scale $\mathfrak{t}_t^{(n)}$ to the random time scale $\rho_t^{(n)}$ merely by switching W to $\mathbb{1}_{\{W>0\}}$. This seemingly surprising fact can be explained by the essential role of W. As mentioned in point (ii) of these remarks, W encodes the long-term stochasticity of the lineage of wild-type cells originating from the initial cell. By showing that the time scale $\mathfrak{t}_t^{(n)}$ serves as the correct deterministic approximation of $\rho_t^{(n)}$, it follows that obtaining an asymptotic result on time scale $\mathfrak{t}_t^{(n)}$ also yields a result for the time scale $\rho_t^{(n)}$. This idea is formalised using a technique similar to that in [56, Lemma 3]. The switch from W to $\mathbb{1}_{\{W>0\}}$ in the result occurs because the time scale $\rho_t^{(n)}$ inherently carries the stochasticity of the random variable W. Consequently, the only remaining randomness that needs to be considered is the survival of the lineage from the initial cell, which is asymptotically given by $\mathbb{1}_{\{W>0\}}$.

Stochastic exponents for selective cancer evolution

The next theorem does not require the non-increasing growth rate condition of Equation (1.2.4). Without this assumption, a martingale approach, as used for proving Theorem 1.2.4 fails to obtain

the first-order asymptotics of the mutant subpopulation sizes. However, the stochastic exponents of the mutant subpopulations, as defined in (1.2.3), can be uniformly tracked over time. In Chapter 5, we actually derive the first-order asymptotics of the size of the mutant subpopulations, as in Theorem 1.2.4, but only for the very specific trait space that is mono-directional leading to the first selective mutant trait. In the next theorem, we show for a general finite directed graph that, under the event $\{W > 0\}$, the limits are positive deterministic non-decreasing piecewise linear continuous functions. Such limits are defined via a recursive algorithm tracking their slopes over time. More precisely, we show that the slopes can only increase and take values from the growth rate function λ. This result is obtained through an adaptation of the techniques developed in the large population and power mutation rates regime, see in particular [6, 11, 12] to the current multitype birth and death branching process.

In the tracking algorithm, two different kinds of updates can be made:

- **Birth of a new trait:** The first update is the birth of a new trait which takes as its slope the maximum between its inner growth rate and the slope of the subpopulation that gave birth to it. In fact, it could also happen that many subpopulations give birth to it at the same time; in this case it is the maximum of their slopes that is compared to the inner growth rate of the born trait. Such a comparison on the growth rates indicates which mechanism is driving the subpopulation growth: either its inner growth if this subpopulation is selective compared to the subpopulation(s) that is/are giving birth to it, or conversely the mutational process if it is deleterious. The neutral case corresponds to a balance of these two mechanisms, as previously mentioned in Theorem 1.2.4.

- **Growth driven by another trait:** The second kind of update is when a live trait v increases its slope because another live trait u among its incoming neighbours, with a higher slope, has reached its typical size so that the mutational contribution from trait u now drives the growth of trait v. Consequently trait v now takes the slope of trait u. Again potentially many traits u among the incoming neighbours of trait v can reach at the same time the typical size for the mutational contribution to drive the growth of trait v; in this case the growth of trait v is driven by the trait u with the maximal slope. This kind of update encodes the possibility in the evolutionary process that the driving mechanism of a subpopulation can change over time, always triggering an increase in the actual growth of the subpopulation.

How these two different kinds of updates happen in the tracking algorithm is made formal in the following theorem. Moreover, they can happen at the same time for different vertices. The complexity of such an algorithm comes mostly from the generality both on the growth rate function and on the trait structure. Under the non-increasing growth rate condition (1.2.4), the limiting functions $(x_v)_{x\in V}$ have an explicit form, see Corollary 1.2.7; this is also true when the graph structure is mono-directional, see Corollary 1.2.6.

Theorem 1.2.5. *Let $0 < T_1 < T_2$. The stochastic exponents defined in (1.2.3) satisfy*

$$\Big(\big(X_v^{(n)}(t)\big)_{v\in V}\Big)_{t\in[T_1,T_2]} \underset{n\to\infty}{\longrightarrow} \mathbb{1}_{\{W>0\}}\Big(\big(x_v(t)\big)_{v\in V}\Big)_{t\in[T_1,T_2]}$$

in probability in $L^\infty[T_1, T_2]$. For each $v \in V$, x_v is a positive deterministic non-decreasing piecewise linear continuous function obtained via a recursive approach tracking its slope over time. In particular there exist $k^* \in \mathbb{N}$ and $0 = \Delta_0 < \Delta_1 < \cdots < \Delta_{k^*} < \infty$ such that the slopes of $(x_v)_{v\in V}$ change only at the times $(\Delta_j)_{j\in\{0,\cdots,k^*\}}$. For $j \in \{0, \cdots, k^*\}$, at time Δ_j two kinds of updates in the slopes

can occur: (i) either a new trait starts to grow or (ii) an already growing trait increases its slope due to a growth driven now by another more selective trait. The algorithm tracks the following quantities for all $j \in \{0, \cdots, k^*\}$ *at time* Δ_j*:*

- *the set of alive traits,* A_j*,*
- *the set of not-yet-born traits,* U_j*,*
- *the slope of* x_v*,* $\lambda_j(v)$*,*
- *and the set of traits whose growth is driven by trait* v*,* $C_j(v)$*.*

Initialisation: *Set* $A_0 = \{0\}$*,* $U_0 = V \setminus \{0\}$ *and for all* $v \in V$

$$x_v(0) = 0, \lambda_0(v) = \lambda(0)\mathbb{1}_{\{v=0\}}, \text{ and } C_0(v) = \emptyset.$$

Induction: *Let* $j \in \{0, \cdots, k^* - 1\}$*. Assume that there exist times* $0 = \Delta_0 < \Delta_1 < \cdots < \Delta_j < \infty$ *such that* $(x_v)_{v\in V}$ *are positive deterministic non-decreasing piecewise linear continuous functions defined on* $[0, \Delta_j]$*, where changes of slopes occur only on the discrete set* $\{\Delta_1, \cdots, \Delta_j\}$*. Also assume that there exist* $\lambda_j(v)$*,* A_j*,* U_j*, and* $C_j(v)$*, respectively the slope of* x_v*, the set of alive vertices and not-yet-born vertices, and the set of vertices whose growth is driven by* v*, everything at time* Δ_j*.*

Then there exists $\Delta_{j+1} \in (\Delta_j, \infty)$ *such that* $(x_v)_{v\in V}$ *are constructed during the time period* $[\Delta_j, \Delta_{j+1}]$ *according to the following. For all* $v \in V$ *and for all* $t \geq \Delta_j$ *let*

$$y_v(t) = (t - \Delta_j)\lambda_j(v) + x_v(\Delta_j).$$

For all $v \in U_j$ *define*

$$\begin{aligned}
&\forall u \in A_j \text{ such that } (u, v) \in E, \delta_{u,v} := \inf\{t \geq \Delta_j : y_u(t) \geq \lambda(0)\ell(u, v)\},\\
&\delta_v := \inf_{u\in A_j:(u,v)\in E} \delta_{u,v},\\
&\nu(v) := \{u \in A_j : (u, v) \in E \text{ and } \delta_{u,v} = \delta_v\}.
\end{aligned}$$

For all $v \in A_j$ *define*

$$\begin{aligned}
&B_j(v) := \{u \in A_j : (v, u) \in E \text{ and } \lambda_j(v) > \lambda_j(u)\},\\
&\forall u \in B_j(v), \delta_{v,u} := \inf\{t \geq \Delta_j : y_v(t) \geq y_u(t) + \lambda(0)\ell(v, u)\},\\
&\delta_v := \inf_{u\in B_j(v)} \delta_{v,u},\\
&\nu(v) := \{u \in B_j(v) : \delta_{v,u} = \delta_v\}.
\end{aligned}$$

Then define $\Delta_{j+1} := \inf_{v\in V} \delta_v$ *and* $\nu_{j+1} := \{v \in V : \delta_v = \Delta_{j+1}\}$*. Then proceed to the following updates:*

- *Let* $A_{j+1} := A_j \cup (\nu_{j+1} \cap U_j)$ *and* $U_{j+1} = U_j \setminus (\nu_{j+1} \cap U_j)$*. Also let* $\forall v \in U_{j+1}$*,* $\lambda_{j+1}(v) = \lambda_j(v) = 0$*,* $C_{j+1}(v) = C_j(v) = \emptyset$*.*
- *For all* $v \in \nu_{j+1} \cap A_j$*, introduce the set* $\nu^{(-)}(v) := \{u \in \nu(v) : \exists w \in \nu_{j+1} \cap A_j, \lambda_j(w) > \lambda_j(v), \text{ and } u \in \nu(w)\}$*.*

  Then let $C_{j+1}(v) := C_j(v) \cup \bigcup_{u\in\nu(v)\setminus\nu^{(-)}(v)} (\{u\} \cup C_j(u))$*. For all* $u \in \nu(v)\setminus\nu^{(-)}(v)$ *and* $w \in C_j(u)$*,* $\lambda_{j+1}(u) = \lambda_{j+1}(w) = \lambda_j(v)$*.*

- *For all $v \in A_j$ whose slope has not been updated yet, let $\lambda_{j+1}(v) = \lambda_j(v)$. And for all $v \in A_j$ whose set $C_j(v)$ has not been updated yet, let $C_{j+1}(v) := C_j(v)$.*
- *For all $v \in \nu_{j+1} \cap U_j$, let $\lambda_{j+1}(v) := \max\left(\lambda(v), \max_{u\in\nu(v)} \lambda_{j+1}(u)\right)$, and $C_{j+1}(v) = C_j(v) = \emptyset$. If $\lambda_{j+1}(v) \geq \lambda(v)$, introduce the following set $\nu^+(v) := \{u \in \nu(v) : \lambda_{j+1}(u) = \max_{w\in\nu(v)} \lambda_{j+1}(w)\}$, and for all $u \in \nu^+(v)$, $C_{j+1}(u) := C_{j+1}(u) \cup \{v\}$.*

When considering a(n) (infinite) mono-directional graph, the structure of such a graph is sufficiently simple to allow for an explicit form of the limiting functions $(x_v)_{v\in V}$, see the next corollary. In particular, there is only one possible slope change that can happen at a time. More specifically, when a not-yet-born trait becomes alive due to the previous trait reaching the typical size allowing for mutations. When this happens, the new born trait takes the slope the maximum between its inner growth rate or the current slope of the previous trait (as mentioned in the first point of the heuristics preceding Theorem 1.2.5). Any alive trait cannot update its slope because no backward mutation is permitted with this graph structure. Moreover, only a single trait becomes alive at each time, due to the scaling labels $\ell(i, i+1)$ being positive.

Corollary 1.2.6 (Theorem 1.2.5 applied to a mono-directional graph)**.** *Assume the graph is infinite and mono-directional, i.e. $(V, E) = (\mathbb{N}_0, \{(i, i+1), i \in \mathbb{N}_0\})$ and that $\ell^* := \inf\{\ell(i, i+1), i \in \mathbb{N}_0\} > 0$. Then the limiting functions $(x_i)_{i\in\mathbb{N}_0}$ of Theorem 1.2.5 have the following simplified form:*

$$\forall t \in \mathbb{R}^+, x_i(t) := \lambda_{\max}(i)(t - \widetilde{t}(i))_+,$$

where $\lambda_{\max}(i) = \max_{j\in\{0,\cdots,i\}} \lambda(j)$ and $\widetilde{t}(i) := \sum_{j=0}^{i-1} \frac{\ell(j,j+1)\lambda(0)}{\lambda_{\max}(j)}$.

Theorem 1.2.5 is more general than Theorem 1.2.4 in the sense that there is no assumption on the growth rate function, but it is a less refined result. We are going to do a full comparison of Theorem 1.2.4 and 1.2.5 on the example of Figure 1.6.

Comparison between Theorems 1.2.4 and 1.2.5:

The asymptotic function x obtained through Theorem 1.2.5 for the plain purple trait is $x(t) = \mathbb{1}_{\{t\geq 4\}}\lambda(0)(t-4)$. In the caption of Figure 1.6, it is already made explicit that only the plain green walk will contribute to the size order of the plain purple mutant subpopulation. If one denotes respectively by 1, 2 and 3 the vertices on the plain green walk such that this walk is exactly $(0, 1, 2, 3)$, where 3 is the plain purple vertex, the asymptotic limits for vertex 3, captured by Theorem 1.2.4, is for all $t \geq 4$,

$$\begin{aligned}
&\frac{2\alpha(0)\mu(0,1)}{\lambda(0)} \cdot \frac{2\alpha(1)\mu(1,2)}{\lambda(0)} \cdot \frac{2\alpha(2)\mu(2,3)}{\lambda(0)-\lambda(3)} W \int_3^t \left(\int_1^u ds\right) du \cdot n^{t-4} \log^2(n) \\
&\qquad = \left(\frac{t^2}{2} - t - \frac{3}{2}\right) \frac{8\alpha(0)\alpha(1)\alpha(2)\mu(0,1)\mu(1,2)\mu(2,3)}{\lambda^2(0)\left(\lambda(0)-\lambda(3)\right)} W n^{t-4} \log^2(n).
\end{aligned}$$

In particular, Theorem 1.2.5 captures only the power of n which is $t-4$ whereas Theorem 1.2.4 captures the stochasticity W, a supplementary scaling factor $\log^2(n)$, a time polynomial $\frac{t^2}{2} - t - \frac{3}{2}$ and also a constant depending only on the parameters of the visited vertices $\frac{8\alpha(0)\alpha(1)\alpha(2)\mu(0,1)\mu(1,2)\mu(2,3)}{\lambda^2(0)(\lambda(0)-\lambda(3))}$.

Now we make explicit the form of the limiting functions $(x_v)_{v\in V}$ in the special case where we assume the non-increasing growth rate condition. Under this condition, the limiting functions

$(x_v)_{v\in V}$ take a very simple form. The only quantity one has to consider is the time $t(v)$ at which trait v becomes alive, where $t(v)$ is defined in Definition 1.2.3. Then after this time, trait v grows at speed $\lambda(0)$ due to the non-increasing growth rate condition. This is made formal in the next corollary.

Corollary 1.2.7 (Theorem 1.2.5 applied with the non-increasing growth rate condition of (1.2.4))**.** *Assume the non-increasing growth rate condition of* (1.2.4)*. Then the limiting functions* $(x_v)_{v\in V}$ *of Theorem 1.2.5 have the following simplified form:*

$$\forall t \in \mathbb{R}^+, x_v(t) = \lambda(0)\,(t - t(v))_+\,,$$

where $\forall x \in \mathbb{R}, x_+ := x\mathbb{1}_{\{x\in\mathbb{R}^+\}}$.

1.2.4 Results of Chapter 5

The scope of Chapter 5 is to extend the results of Theorem 1.2.4 to a more general cancer evolution framework that allows for selective mutant traits. A novel approach is introduced making progress in obtaining first-order asymptotics results for the mutant subpopulation sizes when selective mutations are considered in the trait space. This chapter is an ongoing work with Hélène Leman. Up to this point, we have achieved the case of a finite labelled mono-directional graph, where there is a unique selective mutation corresponding to the last one. The result, stated in Theorem 1.2.12, reveals several biologically and mathematically intriguing results, which contrast with the previous findings in the context where only neutral and deleterious mutations are considered.

The model corresponds to the one described in Subsection 1.2.2 where we are considering the specific case of a finite labelled mono-directional graph with $k+1$ vertices, for some $k \geq 1$. More precisely, we have

$$(V, E, L) = \Big(\big\{0, \cdots, k\big\}, \big\{(i, i+1), i \in \{0, \cdots, k-1\}\big\}, \big\{\ell(i), i \in \{0, \cdots, k-1\}\big\}\Big). \quad (1.2.12)$$

We denote respectively by $\alpha_v, \beta_v, \lambda_v$ for the birth, death, and growth rates of a trait $v \in V$ and use the simplified notations $\mu_i^{(n)} := \mu^{(n)}(i, i+1)$ giving the following power law mutation rates regime

$$\forall 0 \leq i \leq k-1, n^{\ell(i)}\mu_i^{(n)} \underset{n\to\infty}{\longrightarrow} \mu_i < \infty.$$

Assume that the first selective mutant trait is the last one, which mathematically means

$$\forall 0 \leq i \leq k-1, \lambda_i \leq \lambda_0 \text{ and } \lambda_k > \lambda_0. \quad (1.2.13)$$

With this setting three different scenarios can happen during a division event of a cell of trait $i \in \{0, \cdots, k-1\}$:

- with probability $\left(1 - \mu_i^{(n)}\right)^2$ each daughter cell keeps the trait i of its mother cell,
- with probability $2\mu_i^{(n)}\left(1 - \mu_i^{(n)}\right)$ exactly one of the daughter cell mutates to the next trait $i+1$ when the second daughter cell keeps the trait i of its mother cell,
- with probability $\left(\mu_i^{(n)}\right)^2$ both of the daughter cells mutate to the next trait $i+1$.

For $i = k$, during a division event of a cell of trait k, both daughter cells keeps the trait k. Instead of employing the deterministic $\log(n)/\lambda_0$-accelerated time scale, which is appropriate for neutral and deleterious cancer evolution since all subpopulations grow exponentially fast at rate λ_0 according to Theorem 1.2.4, we opt for the $\log(n)$-accelerated time scale $t \mapsto t\log(n)$ in this context. To be fully consistent, we define again, for this new time scale, $t(i)$ for all $i \in V$.

Definition 1.2.8. *Define for all* $1 \leq i \leq k$

$$t(i) := \frac{1}{\lambda_0} \sum_{j=0}^{i-1} \ell(j),$$

as the sum of the labels on the edges from trait 0 *to trait* i *renormalized by* λ_0.

We rewrite in the following proposition the asymptotic results from Theorem 1.2.4 in the $\log(n)$-accelerated time scale instead of in the $\log(n)/\lambda_0$-accelerated time scale, for the subpopulations of traits $i \in \{1, \cdots, k-1\}$.

Proposition 1.2.9. *Let* (V, E, L) *be a finite labelled mono-directional graph with* $k+1$ *vertices, for some* $k \geq 1$ *as defined in* (1.2.12). *Assume that the first selective mutation on the graph is the last one as in* (1.2.13). *Let* $M > 0$ *and* $0 < T_1 < T_2$. *Using the mathematical definition of the model given in Section 5.3, see* (5.3.2), (5.3.3) *and* (5.3.4), *there exists a random variable* W *properly defined in* (5.3.10) *such that for all* $i \in \{1, \cdots, k-1\}$

$$\left((t,s) \mapsto \frac{Z_i^{(n)}\left((t(i)+t)\log(n)+s\right)}{n^{t\lambda_0}\log^{\theta(i)}(n)e^{\lambda_0 s}} \right) \underset{n\to\infty}{\longrightarrow} W\widetilde{w}_{(0,\cdots,i)}(t(i)+t),$$

in probability in $L^\infty([T_1,T_2]\times[-M,M])$ *and with* $\widetilde{w}_{(0,\cdots,i)}(t) := w_{(0,\cdots,i)}(t\lambda_0)$. *For any other mathematical description from the one of Section 5.3, the convergence is at least in distribution in* $\mathbb{D}([T_1,T_2]\times[-M,M])$.

Main results and biological interpretation

In this subsection, the result regarding the first-order asymptotics of the first selective mutant subpopulation size is stated. Moreover, some mathematical and biological remarks about this result, as well as a sketch of the proof, based on heuristics, are provided.

When considering a selective mutation, the first intuitive idea that arises from the term "selective" is that potentially only the lineages of the first generated mutant cells will asymptotically contribute. Indeed, any time delay for a lineage results in an exponential growth loss for its size, providing some qualitative weight to this intuition. However, the challenge lies precisely in quantifying this intuition. Specifically, it is necessary to determine how each lineage of the generated mutant cells contributes asymptotically to the subpopulation size. This raises natural questions, such as whether only a finite number of surviving lineages contribute asymptotically, and if so, how many, as well as whether it is the first lineage to reach a certain level that contributes, among potentially many other questions. Answering them is not possible solely with results on the stochastic exponents. However, Theorem 1.2.12 quantitatively answers these natural questions on the contribution of all the selective mutant cell lineages on the first-order asymptotics of the selective mutant subpopulation size. For doing that, we are going to separate the process $Z_k^{(n)}$ into 3 different processes having different mathematical analysis, asymptotic contribution and biological meaning:

(i) The first one, denoted by $Z_{k,1}^{(n)}$, corresponds to the process composed of the mutant cells and their lineages generated up to time

$$t_{-}^{(n)}(k) := t(k)\log(n) - \frac{\theta(k-1)}{\lambda_0}\log(\log(n)).$$

This time corresponds exactly to the first deterministic time at which it becomes likely to observe mutational events from trait $k-1$ to trait k. Indeed, the typical order of the subpopulation $Z_{k-1}^{(n)}$ at time $(t(k-1)+t)\log(n)$, which is $n^{t\lambda_0}\log^{\theta(k-1)}(n)$ according to Proposition 1.2.9, reaches $n^{\ell(k-1)}$, the typical size allowing for mutations, at time $t_{-}^{(n)}(k)$. In Chapter 5 we will show that well renormalized, the subpopulation $Z_{k,1}^{(n)}$ at time $(t(k)+t)\log(n)$ asymptotically follows a compound Luria-Delbrück distribution, in the vein of the works of Cheek and Antal [8, 10].

(ii) The second one, denoted by $Z_{k,2}^{(n)}$, corresponds to the process composed of the mutant cells and their lineages generated between times $\left[t_{-}^{(n)}(k), t(k)\log(n)\right]$, that is during a time scale of order $\log(\log(n))$. Notice that this subpopulation $Z_{k,2}^{(n)}$ is null if there is no neutral mutation up to trait $k-1$, meaning that $t_{-}^{(n)}(k) = t(k)\log(n)$. In the case where there is at least one neutral mutation, we will show that well renormalized, the subpopulation $Z_{k,2}^{(n)}$ at time $(t(k)+t)\log(n)$ asymptotically follows the large time distribution of an approximate model with less stochasticity, that is independent from n. This approach is inspired by the works of Durrett and Moseley [22], Nicholson and Antal [23] and Nicholson, Cheek and Antal [24]. The previous authors directly study this approximate model. The novelty of our approach is to show that the subpopulation $Z_{k,2}^{(n)}$ is sufficiently close to the previous approximate model to allow for great control of $Z_{k,2}^{(n)}$. We developed a martingale approach to get such result.

(iii) The last one, denoted by $Z_{k,3}^{(n)}$, corresponds to the process composed of the mutant cells and their lineages generated after time $t(k)\log(n)$. We will show, using a similar martingale approach as the one of Chapter 4, that asymptotically this process becomes negligible with respect to the total mutant population $Z_k^{(n)}$.

A graphical representation of these three processes is given in Figure 1.7. As aforementioned, for stating the first-order asymptotics of the mutant subpopulation $Z_{k,1}^{(n)}$ and $Z_{k,2}^{(n)}$, we introduce a proper definition of the Luria-Delbrück distribution, the compound Luria-Delbrück distribution, as well as the approximate model under consideration.

Definition 1.2.10 (Luria-Delbrück distribution)**.** *Let $(\xi_i)_{i\in\mathbb{N}}$ be an i.i.d. sequence of exponentially distributed random variables with parameter λ. Let $(Y_i)_{i\in\mathbb{N}}$ be an i.i.d. sequence of birth and death branching processes with rates α and β respectively, satisfying almost surely the initial condition $Y_i(0) = 1$. Let K be a Poisson random variable with parameter ω. The random variables and processes $(\xi_i)_{i\in\mathbb{N}}, (Y_i)_{i\in\mathbb{N}}, K$ are mutually independent. The Luria-Delbrück distribution with parameters $(\lambda, \alpha, \beta, \omega)$, that may be chosen randomly, is defined as the distribution of*

$$B = \sum_{i=1}^{K} Y_i(\xi_i).$$

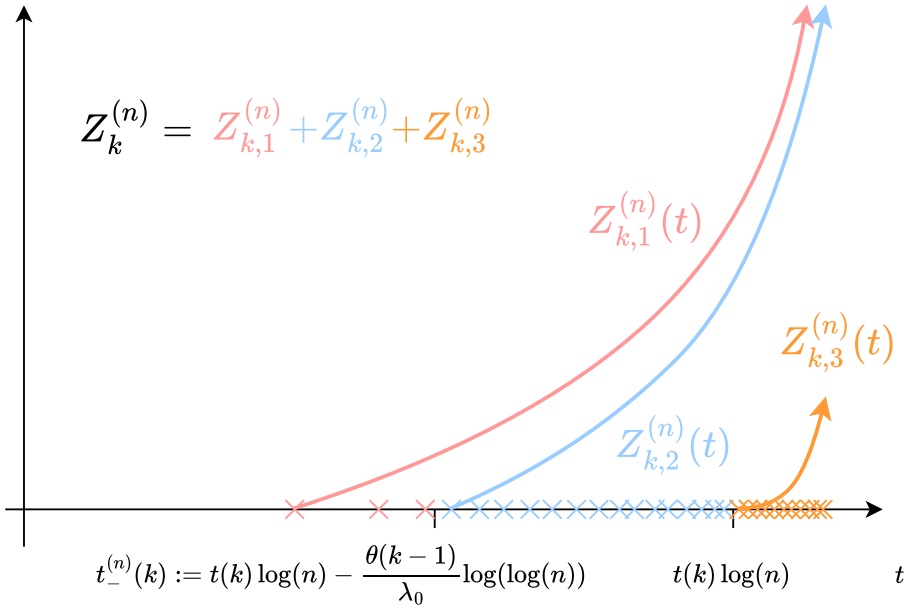

Figure 1.7: Graphical representation of the 3 sub-processes composing $Z_k^{(n)}$

The compound Luria-Delbrück distribution with parameters $(\lambda, \alpha, \beta, \omega)$ *and associated with the distribution* $\mathcal{U}$ *is defined as the distribution of*

$$Z = \sum_{i=1}^{B} U_i,$$

with $(U_i)_{i\in\mathbb{N}}$ *a sequence of i.i.d. random variables following* $\mathcal{U}$ *that is independent from* B*.*

Definition 1.2.11 (Approximate Model)**.** *The approximate model with parameters* $(\lambda, \mu, \alpha, \beta)$ *is defined as the distribution of*

$$Z(t) = \sum_{i=1}^{K(t)} Y_i(t - T_i),$$

where $K(t) = N\left(\int_0^t \mu e^{\lambda s} ds\right)$*, with* N *a Poisson process with intensity 1, independent from the i.i.d. sequence* $(Y_i)_{i\in\mathbb{N}}$ *of birth and death branching processes with rates* α *and* β*, and* $T_i = \inf\{t \geq 0 : K(t) \geq i\}$*. Notice that for all* $t > 0$ *and for all* $i \leq K(t)$*, we have* $t - T_i \geq 0$*. Define* $\widetilde{\lambda} := \alpha - \beta > \lambda$*, then we have*

$$e^{-\widetilde{\lambda}t} Z(t) \underset{t\to\infty}{\longrightarrow} Z^{\infty}, \tag{1.2.14}$$

almost surely where $Z^\infty \in L^1$ *satisfies* $\mathbb{E}[Z^\infty] = \mu/\left(\widetilde{\lambda} - \lambda\right)$ *and its Laplace transform follows the equation*

$$L(\xi,t) := \mathbb{E}\left[e^{-\xi e^{-\widetilde{\lambda}t} Z(t)}\right] \underset{t\to\infty}{\longrightarrow} \mathbb{E}\left[\exp\left(-\xi\frac{\mu}{\widetilde{\lambda}}\Phi\left(-\xi\frac{\alpha}{\widetilde{\lambda}}, 1, \frac{\widetilde{\lambda}-\lambda}{\widetilde{\lambda}}\right)\right)\right],$$

where Φ *is the Lerch transcendent defined as* $\forall\ \Re(s) > 0,\ \Re(a) > 0$

$$\Phi(z,s,a) = \frac{1}{\Gamma(s)}\int_0^\infty \frac{t^{s-1}e^{-at}}{1-ze^{-t}}dt.$$

Equation (1.2.14) is obtained in [24], Proposition 2 for the exact same context, but similar reasoning can be found in [22, 8]. Now we have all the material to state the result of the first order asymptotics for the first selective mutant subpopulation size.

Theorem 1.2.12 (First-order asymptotics for the first selective mutant subpopulation size)**.** *Let* (V,E,L) *be a finite labelled mono-directional graph with* $k+1$ *vertices, for some* $k \geq 1$*, as defined in* (1.2.12)*. Assume that the first selective mutation on the graph is the last one as in* (1.2.13)*. Let* $0 < T_1 < T_2$*. Using the mathematical definition of the model given in Section 5.3, see* (5.3.2)*,* (5.3.3)*,* (5.3.4)*,* (5.3.5)*,* (5.3.6)*,* (5.3.7)*,* (5.3.8)*, and* (5.3.9)*, there exists two random variables* $Z^\infty_{k,1}$ *and* $Z^\infty_{k,2}$ *properly defined in Propositions 5.4.3 and 5.4.4 respectively, that are independent conditioning on* W *(properly defined in* (5.3.10)*) such that*

$$\left(\frac{Z_k^{(n)}\left((t(k)+t)\log(n)\right)}{n^{t\lambda_k}\log^{\theta(k-1)\frac{\lambda_k}{\lambda_0}}(n)}\right)_{t\in[T_1,T_2]} \underset{n\to\infty}{\longrightarrow} Z^\infty_{k,1} + Z^\infty_{k,2},$$

in probability in $L^\infty([T_1,T_2])$*. The random variable* $Z^\infty_{k,1}$ *follows a compound Luria-Delbrück distribution with parameters*

$$\left(\lambda_0, \alpha_k, \beta_k, \frac{2\alpha_{k-1}\mu_{k-1}}{\lambda_0}w_{k-1}W\right)$$

where $w_{k-1} := \widetilde{w}_{(0,\cdots,k-1)}(t(k))$*, and associated with* $\mathcal{U}_k$*, the distribution of the almost sure large limit of the natural martingale associated to a birth and death branching process with rates* α_k *and* β_k *respectively. More precisely,*

$$\mathcal{U}_k := Ber\left(\frac{\lambda_k}{\alpha_k}\right) \otimes Exp\left(\frac{\lambda_k}{\alpha_k}\right).$$

The random variable $Z^\infty_{k,2}$ *follows the distribution of the asymptotic large time limit of the population defined by the approximate model from Definition 1.2.11 with parameters*

$$(\lambda_0, 2\alpha_{k-1}\mu_{k-1}w_{k-1}W, \alpha_k, \beta_k).$$

For any other mathematical description, the convergence is at least in distribution in $\mathbb{D}([T_1,T_2])$*.*

The proof of this theorem involves employing distinct proof techniques for the three subpopulations $Z^{(n)}_{k,1}$, $Z^{(n)}_{k,2}$ and $Z^{(n)}_{k,3}$ that are summed up here, alongside some heuristics and biological interpretations:

- **Heuristics concerning the convergence of $Z_{k,1}^{(n)}$:** The random variable $Z_{k,1}^{\infty}$ is the asymptotic limit of the well-renormalized subpopulation $Z_{k,1}^{(n)}$ composed of the mutant cell lineages issued from mutational events generated before time $t_{-}^{(n)}(k)$. The analysis is made in two steps. First we adapt a proof from [8] or [10] to demonstrate that at time $t_{-}^{(n)}(k)$, the number of mutant cells of trait k asymptotically follows a Luria-Delbrück distribution with parameters $\left(\lambda_0, \alpha_k, \beta_k, \frac{2\alpha_{k-1}\mu_{k-1}}{\lambda_0} w_{k-1} W\right)$, which has a infinite expectation. Such proof uses the Poissonian structure of the process of mutation from trait $k-1$ to trait k as well as that conditioning on the number of mutant cells generated due to mutational events up to time $t_{-}^{(n)}(k)$, the vector of the remaining time for each of the mutant clone to grow up to time $t_{-}^{(n)}(k)$ is asymptotically the order statistic of a vector of i.i.d. exponential random variables with parameter λ_0. Subsequently, we establish control over the size of all mutant cell lineages present at time $t_{-}^{(n)}(k)$ by time $(t(k)+t)\log(n)$ using a martingale approach. These lineages experience exponential growth at rate λ_k over a duration of time $(t(k)+t)\log(n) - t_{-}^{(n)}(k) = t\log(n) + \frac{\theta(k-1)}{\lambda_0}\log(\log(n))$, resulting in a typical size of order $\mathcal{O}\left(n^{t\lambda_k}\log^{\theta(k-1)\frac{\lambda_k}{\lambda_0}}(n)\right)$. This subpopulation corresponds to the lineages of the possible mutant cells generated due to random mutational events appeared before the deterministic time $t_{-}^{(n)}(k)$. Due to the selectiveness of the mutation, their lineages have at time $t_{-}^{(n)}(k)$ asymptotically infinite expectancies. Moreover, the fact that they asymptotically contribute to the size order of trait k agrees with the natural intuition behind considering a selective mutation.

- **Heuristics concerning the convergence of $Z_{k,2}^{(n)}$:** Dealing with the appropriate normalisation of $Z_{k,2}^{(n)}$ is more complex. For the mutant cells generated between times $t_{-}^{(n)}(k)$ and $t(k)\log(n)$, and their lineages, we approximate this subpopulation with the simpler model of Definition 1.2.11 with the adequate parameters that are $(\lambda_0, 2\alpha_{k-1}\mu_{k-1}w_{k-1}W, \alpha_k, \beta_k)$, to reduce stochasticity and facilitate obtaining its asymptotic limit at time $t(k)\log(n)$. We rigorously establish by a martingale argument using a L^1 convergence that at the first-order, this subpopulation behaves asymptotically akin to the considered approximate model up to time $t(k)\log(n)$. In particular, such approximate model grows exponentially fast at rate λ_k during a time $t(k)\log(n) - t_{-}^{(n)}(k) = \frac{\theta(k-1)}{\lambda_0}\log(\log(n))$, implying that at time $t(k)\log(n)$ the subpopulation $Z_{k,2}^{(n)}$ is of order $\mathcal{O}\left(\log^{\theta(k-1)\frac{\lambda_k}{\lambda_0}}(n)\right)$. Then we obtain its asymptotic limit at time $(t(k)+t)\log(n)$ using a law of large numbers argument. More precisely, after time $t(k)\log(n)$ there are no longer mutational event from trait $k-1$ to trait k that are counted in this subpopulation anymore. Because at this time we have an order of $\mathcal{O}\left(\log^{\theta(k-1)\frac{\lambda_k}{\lambda_0}}(n)\right)$ mutants cells, and that their lineages grow exponentially fast at rate λ_k during a time $t\log(n)$, this results in a subpopulation of order $\mathcal{O}\left(\log^{\theta(k-1)\frac{\lambda_k}{\lambda_0}}(n) n^{t\lambda_k}\right)$ at time $(t(k)+t)\log(n)$. This subpopulation encapsulates the idea that an asymptotically infinite number of lineages (but not all of them) significantly contribute to the size of the selective mutant trait subpopulation, despite being characterised by a finite mean random variable. However, these infinite number of lineages are all produced on the slower $\log(\log(n))-$accelerated time scale, meaning that they are asymptotically all condensed at the same time point for the $\log(n)-$accelerated time scale. Such result contrasts a bit with the first biological intuition previously mentioned, be-

cause we obtained that an infinite number of selective mutant cell lineages are asymptotically contributing to the selective mutant subpopulation size, but not so much in the sense that all these contributing mutant cells are actually produced at times that are asymptotically condensed around $t(k)\log(n)$, in agreement with the intuition under selectiveness.

- **Heuristics concerning the convergence of $Z_{k,3}^{(n)}$:** Dealing with $Z_{k,3}^{(n)}$, the subpopulation containing solely mutant cells generated after time $t(k)\log(n)$ along with their lineages, is done using a similar martingale method as the one used in Chapter 4, to get that it scales as $n^{t\lambda_k}\log^{\theta(k-1)}(n)$. Hence, asymptotically it becomes negligible and vanishes in the limit. Heuristically, the selectiveness of the trait implies that only the one generated around time $t(k)\log(n)$ are asymptotically contributing at time $(t(k)+t)\log(n)$. At time $t(k)\log(n)$, the number of mutant cells of trait k generated from mutational events from trait $k-1$ is of order $\mathcal{O}\left(\log^{\theta(k-1)}(n)\right)$. Then their lineages are growing exponentially fast at rate λ_k during a time $t\log(n)$, giving a size of order $\mathcal{O}\left(\log^{\theta(k-1)}(n)n^{t\lambda_k}\right)$ at time $(t(k)+t)\log(n)$. Biologically meaning, after time $t(k)\log(n)$ the generated selective mutant cells and their lineages have too much time delay compared to the first generated mutant cells such that their contribution appears negligible.

Some natural remarks about the result of Theorem 1.2.12 can be made:

- **First-order asymptotics:** In Theorem 1.2.5, one obtains that the stochastic exponent associated to the mutant subpopulation of trait k asymptotically converges, conditioning on $\{W>0\}$, to $t\lambda_k$, which corresponds to the power of n captured in Theorem 1.2.12. Notably, such result means that in the case of a selective mutant trait, growth stems not from the mutational process but rather from the inherent expansion of selective mutant cell lineages. This stands in contrast to deleterious mutations, where growth is actually driven by the mutational process, and neutral mutations, where growth represents a balance between mutational process and inherent lineage expansions. Moreover the power of $\log(n)$ is also captured by this theorem, which is a generalisation of Theorem 1.2.4 allowing for a first selective mutant trait. An interesting pattern for such power occurs. For deleterious and neutral mutations, we previously obtained that a neutral one increases such power by 1 where a deleterious one has no effect. For a selective mutation, it is a bit more intricate: the power obtained for the previous trait is accelerated by the ratio between the growth rate of the selective trait and the actual one of the previous trait (that is the one of the wild-type subpopulation). In particular if there are only deleterious mutations before the first selective one, then the power of $\log(n)$ stays equal to 0 for the first selective trait.

- **Time dependence:** An interesting characteristic of the limit for the selective mutant trait, that contrasts with neutral and deleterious mutations, is its independence from the time parameter t. In Theorem 1.2.4 we have shown that alongside a mono-directional graph composed of neutral or deleterious mutations only, the asymptotic limit of a trait depends on the time parameter t as a polynomial function of degree the number of neutral mutations up to this considered trait. With Theorem 1.2.12 we show that this time dependence is lost when the first selective mutant trait is encountered in the mutational pathway. In a certain sense, one can say that it resets the time dependence. Observing such differences depending on the type of the mutation seems more than intuitively natural. Indeed, with a neutral mutation, the balance between the mutational process and the inherent growth makes clear that the lineages

have equal contributions on the size order, generating a dependence on the time parameter via an extra integral operator from the time dependence of the previous trait. For a deleterious mutation, driven by the mutational process, the time dependence comes from taking into account only clones generated at the considered time, meaning that the dependence is identically forwarded from the one of the previous trait. For a selective mutation, the inherent growth prevails the mutational process, implying that only the lineages generated around the typical time to observe such mutant trait are asymptotically contributing, and so all the time dependency from the previous trait is completely lost.

- **Stochasticity of the limit:** In the case of neutral and deleterious cancer evolution, we obtained in Theorem 1.2.4 that the stochasticity captured by the first order asymptotics of the mutant subpopulation sizes is fully given by the random variable W, which quantifies the large time stochasticity associated to the primary subpopulation. In the remarks made on Theorem 1.2.4 point (iii), we argued that such result could not be obtained for a selective mutation, explaining the failure of the martingale approach. In Theorem 1.2.12 we show that the latter conjecture is true. In the limiting random variables, $Z^{\infty}_{k,1}$ and $Z^{\infty}_{k,2}$, the stochasticity of W is present, but another layer of stochasticity is also given due to the growth of the lineages of the selective mutant cells. The independence, when conditioning on W, of these two limiting random variables comes from the fact that they correspond to the asymptotic limits of two subpopulations of mutant cell lineages that are independent due to the mutational process which is Poissonian.

- **Dealing with the next mutant trait:** Asymptotically we obtain that the limiting random variable $Z^{\infty}_{k,1}$ has an infinite expectation, coming from the Luria-Delbrück part of the compound random variable, a consequence of the selectiveness of the considered mutation. Such specificity entails mathematical difficulties, explaining why the martingale approach developed in Chapter 4 fails. Indeed, taking an expectation balances the probability of an event with the number of cells seen if the event occurs. Consequently, unlikely events, such as rapid mutations to the selective trait, can lead to extremely large population sizes far exceeding typical asymptotic sizes implying the expectation order is given by these unlikely event.

  Moreover, it hampers from dealing with the next mutant trait. In particular the cells of trait $k+1$, and their lineages, produced by $Z^{(n)}_{k,1}$ are, for now, untractable. More specifically, dealing with the latter subpopulation of trait $k+1$ using the same approach as for capturing the asymptotic limit of trait k as in Chapter 5, the martingale approach developed to show that $Z^{(n)}_{k+1,2}$ and the adapted approximate model are sufficiently close to transfer the limiting behaviour from the approximate model to $Z^{(n)}_{k+1,2}$ fails. Indeed, a L^1 approach is needed for this step. Instead of approximating a mutant trait by its asymptotic limit to tackle the next mutant trait, a potential solution to deal with such difficulty is to directly deal with all the mutant subpopulations together without using the approximations given by the limits at each step. This procedure would normally allow for overstepping the infinite expectation of $Z^{\infty}_{k,1}$.

- **Two steps limit:** Instead of being interested in the double limit given by the large population and power law mutation rates regime, another interesting approach is to consider a two steps asymptotics where first the large time limit is obtained with fixed mutation rates and then the limit when these rates tend to 0 is obtained. In [24], Nicholson, Cheek and Antal are studying this two steps limit. An interesting fact is that the results are different, meaning

that there is no equivalence between performing the double limit directly or in two steps. But some similarities are actually obtained for the structure of the asymptotic limits. Notably, both asymptotic results can be decomposed into the product of a time-independent random variable (which are different) and a simple time-dependent deterministic function controlled by the growth rate of the selective mutant trait (with the same deterministic function but taken at different times). More precisely, in their case, the stochasticity of the asymptotic limit that is captured (see [24], Theorem 1) is a Mittag-Leffler distribution with tail parameter the ratio between the growth rate of the wild-type subpopulation and the one of the first selective mutant trait, and a scale parameter satisfying a specific equation depending on the birth, growth, and mutation rates of the preceding mutant traits. Such asymptotic random variable is not captured by Theorem 1.2.12, as mentioned above. Moreover, the deterministic function is in both cases the exponential function whose growth rate is the one of the selective mutant trait, but taken at different times. In their case, when looking at the limit when $t \to \infty$, the time that is taken for the deterministic function is actually t, where in our case, at time $(t(k) + t)\log(n)$ it is $t\log(n)$ that is taken.

1.2.5 Perspectives

Many refinements on the model of carcinogenesis developed in Chapters 4 and 5 are still possible. Regarding Chapter 5 the immediate next step is to adapt the current approach in order to capture the first order asymptotics when the graph is a mono-directional one allowing for more than one selective mutant trait. Then, using an approach based on quantifying the contribution of the mutational paths on the graph, as in Subsection 4.4 when allowing only neutral and deleterious mutations, one can obtain the first order asymptotics for a general finite directed labelled graph. Such result would give characterisation of evolutionary effective pathways, the ones that contribute asymptotically. Currently, we have developed a method capable of dealing with up to the first selective mutant trait. The next step is therefore to find a way to handle subsequent mutations after this first selective mutant trait. We will start by looking at a simpler model with 2 mutations where the first one is selective. The challenge in this generalisation lies in the fact that, compared to the asymptotic limits for neutral and deleterious traits, the asymptotic limit for the selective mutant trait has an infinite expectation. This specificity implies that all the techniques used to deal with a new mutant trait, regardless of its selective advantage, no longer work. One potential approach to address this singularity is to consider the asymptotic limits of mutant traits all together, rather than relying solely on the asymptotic behaviour of the previous mutant subpopulation to deal with the next mutant trait.

Moreover, in Theorem 1.2.12, the asymptotic limit of the first selective mutant trait is given as the sum of two (more or less complex) random variables. One could try to get a simpler identification of the asymptotic distribution. A promising strategy is to introduce a time parameter $C > 0$ such that the selective mutant population $Z_k^{(n)}$ is separated into two subpopulations $Z_{k,C,1}^{(n)}$ and $Z_{k,C,2}^{(n)}$ composed respectively of the lineages of the mutant cells generated before time $t_-^{(n)}(k) + C$ (instead of $t_-^{(n)}(k)$ in the current approach), respectively after the later time. Adapting the proof techniques of Chapter 5, one can show that both the subpopulations converge to some random variables $Z_{k,C,1}^{\infty}$ and $Z_{k,C,2}^{\infty}$, whose law depends on the parameter C. By identification of the limit, the distribution of $Z_{k,1}^{\infty} + Z_{k,2}^{\infty}$ is the same as the distribution of $Z_{k,C,1}^{\infty} + Z_{k,C,2}^{\infty}$ for all $C > 0$, allowing to look at the limit when C goes to infinity. Intuitively, when C is growing the contribution of $Z_{k,C,1}^{(n)}$ becomes

more important since more lineages of trait k are included in this subpopulation. One could show that when taking the limit C goes to infinity, the distribution of $Z^{\infty}_{k,C,2}$ converges to 0, since all the lineage contributions are incorporated in the limit of the distribution of $Z^{\infty}_{k,C,1}$. Then, using the Laplace transform could be an option for characterising the limiting distribution of $Z^{\infty}_{k,C,1}$ when C goes to infinity.

The main insight provided by Theorem 1.2.4 in the asymptotic limits concerns the identification of distinct effective evolutionary pathways, a facet not previously elucidated in the literature, such as in the stochastic exponent results outlined in Theorem 1.2.5. A natural progression would be to leverage this newfound information to conduct inference on the graph structure, particularly in contexts where genetic compositions over time are accessible. This refinement holds significant appeal for biologists as it promises insights into future tumour evolution behaviours, thereby facilitating the development of adaptive therapeutic strategies. Furthermore, the asymptotic limits afford valuable insights into model parameters, offering potential avenues for parameter estimation based on theoretical findings. Another possibility entails utilising information on neutral (and deleterious) cancer evolution to conduct precise neutrality cancer evolution tests.

The branching hypothesis for cancer formation becomes inadequate as the cell population increases, necessitating consideration of cell-cell interactions. Introducing a model that incorporates competition between individuals, with a carrying capacity K that scales as a specific power of the parameter n, becomes imperative. The environment's carrying capacity is a measure of the maximal population size that the environment can sustain for a long time. The mathematical challenge lies in obtaining first-order asymptotic results for mutant subpopulations, akin to Chapters 4 and 5. This endeavour represents an extension of [12], where the authors explore a finite directed graph trait space within the context of large population and power law mutation rates, incorporating a competitive kernel between individuals and deriving asymptotic results for the stochastic exponents of mutant subpopulations. To achieve this, one must first perform the first-order asymptotic results for growing subpopulations and the resident population around any change in slope in the algorithm described in Theorem 1.2.5. Subsequently, one must address first-order asymptotics around any change in slope, arising from either the emergence of a new trait (by a similar analysis as in Chapters 4 and 5) or a shift in the resident population.

When investigating cancer evolution through genetic composition analysis, the ultimate aim is to devise an adaptive therapeutic strategy. Insights gleaned from the past and present genetic composition of the tumour provide valuable foresight into its future evolution, enabling the formulation of strategies to impede tumour growth. In the context of adaptive therapy, one pertinent consideration involves the strategic switching between different treatments. A key biological question revolves around determining the optimal frequency of such switches to effectively control tumour growth. This can be achieved by incorporating various therapies into the model through the introduction of different environments. The growth rate function becomes contingent not only on the trait but also on the prevailing environment. One fundamental mathematical inquiry pertains to identifying, given the trait space, mutational regime, and growth rate function, the optimal strategy for switching environments to control tumour growth. One potential criterion could involve ensuring the tumour does not exceed a certain size threshold or, if unavoidable, delaying this occurrence for as long as possible. Consequently, the frequency of switches may not be deterministic. The consideration of periodic switching of environment is studied in [17] for a different model but under the large population and power law mutation rates regime.

Some criticisms can be raised regarding the mutation regime. While we typically fix the number of genetic traits, it could be argued that the number of sites on the genome where mutations can

occur is sufficiently high relative to the observed tumour size, thus scaling with the parameter n, rather than remaining constant. This perspective would lead us to consider the *large population, large trait space, and power law mutation rates regime* instead of solely focusing on the large population and power law mutation rates regime. To illustrate, consider the trait space represented by the set of vertices $V = \{0,1\}^S$, where $S \in \mathbb{N}$ denotes the number of sites on the genome, and the set of edges $E = \{(u,v) \in V^2, u \neq v\}$. In this scenario, each cell's phenotypic traits are represented by a finite sequence of 0s and 1s across the S sites on the genome, with each cell capable of mutating to any other cell trait. Specifically, the mutation probability among sites could be independent of each other, as observed in [8, 10]. In this particular context, the large population, large trait space, and power law mutation rates regime means

$$\begin{aligned}
&\exists (s_n)_{n\in\mathbb{N}}, s_n \underset{n\to\infty}{\longrightarrow} \infty \text{ and } \left|S^{(n)}\right| = s_n,\\
&\forall i \in \mathbb{N}, \forall j \in \{0,1\}, \exists \ell_{i,j} \in \mathbb{R}^+, n^{\ell_{i,j}} \mu_{i,j}^{(n)} \underset{n\to\infty}{\longrightarrow} \mu_{i,j} \in \mathbb{R}^+,
\end{aligned}$$

where $\mu_{i,j}^{(n)}$ represents the mutation probability on site i from trait j to trait $\overline{j} := \{0,1\}\backslash j$. While such considerations are mentioned in [10], the studied mutation regimes first focus on power law mutation rate regimes before considering the scenario where the number of sites on the genome tends to infinity.

In Theorem 1.2.5, when considering a specific trait space (V,E,L) and a given growth rate function λ, the evolutionary pathways of the tumour observed by this theorem are entirely deterministic. This means that these data alone suffice to pre-determine tumour evolution from a deterministic standpoint. The results on first-order asymptotics, as presented in Theorem 1.2.4, do not introduce additional complexity for the evolutionary process. Instead, they precisely characterise the deterministic nature of the evolutionary pathways, which are functions of (V,E,L) and λ. The random processes captured through first-order asymptotics at the limit are deterministic functions of the set of evolutionary pathways. To introduce a model that allows for evolutionary processes not deterministically pre-determined by the parameters of the model, one could initially consider introducing stochasticity into the growth rate function λ. This randomness would create variability into the evolutionary pathways, thereby departing from strict determinism.

Part I

Invasion of Cooperative Parasites

Chapter 2

Invasion of cooperative parasites in moderately structured host populations

This chapter corresponds to the publication [1] "Invasion of cooperative parasites in moderately structured host populations", written in collaboration with Cornelia Pokalyuk, which has been published in *Stochastic Processes and their Applications*.

Abstract:
Certain defence mechanisms of phages against the immune system of their bacterial host rely on cooperation of phages. Motivated by this example we analyse invasion probabilities of cooperative parasites in moderately structured host populations. We assume that hosts occupy the vertices of a configuration model and offspring parasites move to neighbouring sites to infect new hosts. Parasites (usually) reproduce only when infecting a host simultaneously and then generate many offspring. In this regime we identify and analyse the spatial scale of the population structure at which invasion of parasites turns from being an unlikely to a highly probable event.
Keywords: host, parasite, cooperation, invasion probability, configuration model.
MSC2020: 92D30, 60J80, 92D25.

2.1 Introduction

We analyse the invasion probability of parasites in moderately structured host populations. The motivation of this study stems from observations of phage populations. Phages are viruses infecting bacteria. The interest in phages has been growing in recent years because of the growing incidence of multi-drug resistant bacteria. As an alternative to antibiotics, in phage therapy the infected host is inoculated with a population of phages to eliminate the pathogenic bacterial population [65].

Bacteria own various mechanisms to defend against phages, one of these is CRISPR-Cas. This mechanism relies on certain complexes of proteins, that are patrolling in the bacterial cell to detect (and subsequently distroy) genetic material of phages (that the bacterial cell or its ancestors encountered previously and stored at the so called CRISPR-locus in the bacterial genome), see [26]. Some phages can block these complexes with mechanisms called anti-CRISPR (ACR) which

relies essentially on cooperation of ACR-phages [28]. Indeed, when a CRISPR-resistant bacterium is attacked by a single ACR-phage, the phage often dies, whereas when several phages attack a bacterium simultaneously or subsequently, they have a good chance to replicate [29], [28].
The models that have been investigated so far to understand the underlying growth dynamics of ACR-phages and CRISPR-resistant bacterial populations are deterministic models that map the behaviour of well-mixed phage and bacterial populations [28]. In these models one starts with a relatively large phage population, for which simultaneous or rapid subsequent attacks of phages are likely.

Here we consider a phage population that is initially small. In this setting stochastic effects cannot be ignored. We are interested in the probability that the phage population manages to invade the bacterial population, in the sense that a non-trivial proportion of the bacterial population gets infected and subsequently killed by the phages.

We assume that offspring phages attach to neighbouring bacteria. If the bacterial population is well-mixed, offspring numbers of phages need to be very large for simultaneous infections of neighbouring bacteria to be likely. However, many bacterial populations are spatially structured, e.g. in biofilms, see [25]. In this case bacteria are only adjacent to a relatively small part of the bacterial population and co-infections of bacteria are common even when offspring numbers of phages are moderate. Consequently, invasion of phages should be more likely in spatially structured bacterial populations than in well-mixed populations.

Population dynamics involving cooperation have been mainly studied from the perspective of a single population that is divided into defectors and cooperators. In these studies one often is interested if cooperators may prevail or coexist with the population of the defectors, see e.g. [66], [67]. Here we consider only cooperators. The survival of the population of cooperators is nonetheless non-trivial, because the capability of the individuals to cooperate depends on the population structure of another population, the host population.

Even though the motivations of this project come from phages, we think that our results might be also relevant for other host-parasite systems. For example it is believed that the infection of cancer cells with oncolytic viruses, that is viruses that attack cancer cells, is more effective, if a cancer cells are hit by several viruses simultaneously, because in this manner the virus can cope better with the (interferon-based) anti-viral response of the host, see [31].

In order to put our study into a general context in the following we will consider instead of a population of phages and bacteria a population of cooperative parasites and hosts. Even though viruses (and in particular phages) are not regarded as parasites by biologists we think it is appropriate to call the involved individuals parasites and hosts, because the population dynamics of the phage population is characterised by the fact that phages are only capable to reproduce in their host, the main feature of parasites.

Spread of parasites or pathogens in finite host populations has been analysed mainly with respect to epidemiological models, in which only the host population is modelled. Hosts are either susceptible, infected or recovered and the host population is placed on the complete graph or the configuration model, see [68], [69] or [70]. Here we consider both populations.

We model the spatial structure of the host population by placing hosts on the vertices of a random graph of size N formed according to the configuration model. We assume that each host is neighboured by d_N hosts, where $1 \ll d_N \ll N$, and hosts are placed on vertices of a random graph whose edges are arranged according to the configuration model. Initially a single host gets infected by a parasite and v_N offspring parasites are produced. Thereafter the populations evolve in discrete generations. At the beginning of each generation parasites move randomly to neighbouring

hosts. Whenever a host gets attacked by at least two parasites the parasites reproduce. If a host gets infected only by a single parasite, the infection is successful only with some small probability ρ_N. At parasite reproduction v_N parasites are generated. We show that at the scale $v_N \sim c\sqrt{d_N}$, for some $c > 0$, the number of neighbouring hosts that is attacked simultaneously by offspring parasites is approximately Poisson distributed with parameter $c^2/2$. Furthermore, in the regime $v_N\rho_N \sim x$, for some $0 \leq x \leq 1$ the number of hosts that get successfully infected by single parasites is approximately Poisson distributed as well this time with parameter x. (The assumption $x \leq 1$ guarantees that invasion due to infections by single parasites is unlikely.)

We explore the spread of the parasite population within the host population (guided by the analysis of epidemics on random graphs, see [69], Part III, as well as [68]) by couplings with (truncated) Galton-Watson processes (GWP) until N^α hosts get infected for some $\alpha > 0$ sufficiently large. In this phase the invasion process is essentially driven by pairs of parasites originating from the same vertex and attacking neighbouring hosts simultaneously as well as parasites attacking hosts alone successfully in the case $\rho_N v_N \to x$ with $x > 0$. Once the number of infected hosts per generation exceeds the level N^α, with high probability in a finite number of generations the remaining hosts get infected due to parasites attacking hosts simultaneously from different edges. Hence, the invasion probability of the parasite population, that is the probability that the host population eventually gets killed, is in the critical scale $v_N \sim c\sqrt{d_N}$ asymptotically equal to the survival probability of a Galton-Watson process with an offspring distribution that is given by the sum of independent Pois$(c^2/2)$ and Pois(x)-distributed random variables.

2.2 A host-parasite model with cooperative parasites

2.2.1 Model description and main results

Consider a population of hosts and a population of parasites both located on a random graph. The graph has N vertices and each vertex has d_N half-edges. We assume that $d_N N$ is even and half-edges are matched according to the configuration model, i.e. half-edges are paired uniformly at random.

Initially, on each vertex a single host is placed. We start the infection process by infecting a randomly chosen host with a parasite. We say that parasites infect a host, when the infecting parasites replicate in the host. At replication v_N offspring parasites are generated (independent on the number of infecting parasites) and the host as well as the infecting parasite(s) die(s).

The infection process continues in discrete generations according to the following scheme. At the beginning of each generation, parasites move independently to nearest neighbouring vertices. If a vertex to which a parasite moves to is still occupied with a host the parasite attacks this host. If a host is only attacked by a *single parasite*, the parasite replicates only with a small probability ρ_N. In this case v_N offspring parasites are generated and the reproducing parasite as well as the host die. Otherwise (with probability $1 - \rho_N$), the parasite dies and the host survives. If, however, at least two parasites attack a host simultaneously, the parasites cooperate, they produce (with probability 1) in total v_N offspring parasites and the infecting parasites and the host die. If a parasite moves to a vertex that is no longer occupied by a host, it stays there and moves further in the next generation. Hosts do not move on the graph during the infection process. See Figure 2.1 for an illustration of the infection process.

Given a sequence of parameters $(N, d_N, v_N, \rho_N)_{N\in\mathbb{N}}$ we denote for each $N \in \mathbb{N}$ by

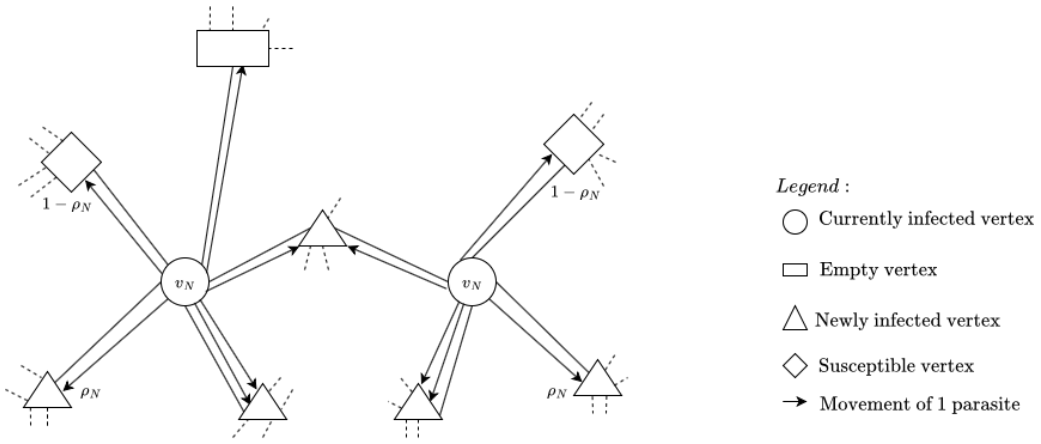

Figure 2.1: Illustration of different infection types

$$\mathcal{I}^{(N)} = (I_n^{(N)})_{n\in\mathbb{N}_0},$$

the process that counts the number of infected hosts in the generations $n \in \mathbb{N}_0$ and by

$$\overline{\mathcal{I}}^{(N)} = (\overline{I}_n^{(N)})_{n\in\mathbb{N}_0},$$

with

$$\overline{I}_n^{(N)} := \sum_{i=0}^{n} I_i^{(N)},$$

the process that counts the number of hosts infected till generation $n \in \mathbb{N}_0$.
We are interested in the probability that the parasite population invades the host population. More precisely, we consider the following events.

Definition 2.2.1 (Invasion of parasites)**.** *Consider the above host-parasite model with parameters* $(N, d_N, v_N, \rho_N)_{N\in\mathbb{N}}$. *Let* $u \in (0, 1]$ *and denote by*

$$E_u^{(N)} := \Big\{\exists n \in \mathbb{N}_0 : \overline{I}_n^{(N)} \geq uN\Big\},$$

the event that the parasites invade the host population (at least) to a proportion u.

In the following we consider parameter regimes for which the host population is initially large, that is $N \to \infty$. We will write $\to$ for $\overset{N\to\infty}{\to}$ throughout the paper, unless otherwise specified. We assume that from each host many other hosts can be reached, i.e. $d_N \to \infty$, but the population is not well mixed, in the sense that $d_N \in o(N)$. Furthermore many offspring parasites are produced at infection of a host, i.e. $v_N \to \infty$, and the contribution of parasites attacking a host alone is at most critical, in the sense that the expected number of offspring $v_N\rho_N$ generated at such attacks is at most 1. In Theorem 2.2.2 we identify the critical scaling of v_N and d_N, at which invasion of a non-trivial proportion of the host population turns from an improbable to a very likely event.

Theorem 2.2.2. *Assume $d_N \in \Theta(N^\beta)$ for some $0 < \beta < 1$, and $\rho_N v_N \to x$ for some $0 \le x \le 1$. Depending on the order of v_N we obtain the following invasion regimes:*
(i) Assume $v_N \in o(\sqrt{d_N})$. Then for all $0 < u \le 1$

$$\lim_{N\to\infty} \mathbb{P}\left(E_u^{(N)}\right) = 0.$$

(ii) Assume $v_N \sim c\sqrt{d_N}$ for $c > 0$. Denote by $\pi(c, x)$ the survival probability of a Galton-Watson process with Pois$(\frac{c^2}{2} + x)$-offspring distribution. Then the invasion probability of parasites satisfies for all $0 < u \le 1$

$$\lim_{N\to\infty} \mathbb{P}(E_u^{(N)}) = \pi(c, x).$$

(iii) Assume $\sqrt{d_N} \in o(v_N)$. Then

$$\lim_{N\to\infty} \mathbb{P}\left(E_1^{(N)}\right) = 1.$$

After Remark 2.2.3 we will sketch the proof of Theorem 2.2.2 in Subsection 2.2.2 and discuss some generalisations of the model and the results in Subsection 2.2.3. A rigorous proof of Theorem 2.2.2 will be given in Section 2.6 after preparing auxiliary results in Sections 2.3 - 2.5. In Table 2.1 notation that is frequently used in the manuscript is summarized.
We will often write *whp* for *with high probability* to indicate that an event occurs with a probability that is asymptotically 1 as $N \to \infty$.

Remark 2.2.3. *(i) In the setting of Theorem 2.2.2 (ii) for $\frac{c^2}{2} + x \le 1$ we have $\pi(c, x) = 0$, which means that whp parasites do not invade the host population.*

(ii) We assume $v_N \rho_N \to x \le 1$, that is the capability for reproduction of parasites hitting a host alone is subcritical or critical (in the terminology of branching processes).

(iii) It has been shown that population viscosity, i.e. limited dispersal of individuals, is generally beneficial for cooperation, see [32]. Here we see an example at which the spatial structure of the host population is passed on to the parasite population that profits from this structure as well. Consequently, in host-parasite systems the host population may on the one hand profit from a spatial structure by enhancing cooperation of hosts, but on the other hand spatial structure may reduce the fitness of the host population because parasite populations may benefit from the spatial structure as well.

(iv) The proof of Theorem 2.2.2 (ii) yields that the time till the entire host population gets infected is upper bounded by $\frac{(1-\frac{3}{4}\beta+\varepsilon)\log N}{\log(c^2/2+x)}$ for any $\varepsilon > 0$, conditioned on a parasite outbreak. Indeed to prove Theorem 2.2.2(ii) we approximate $\mathcal{I}^{(N)}$ by a Galton-Watson process from below, that is truncated from time to time but grows at the same speed as an ordinary Galton-Watson process (with asymptotic offspring mean $c^2/2 + x$), until the level $N^{1-\frac{3}{4}\beta+\delta}$ is reached, for some $\delta > 0$ sufficiently small. Afterwards the host population gets killed whp within two more generations. From this follows immediately that the host population is whp killed after time $\frac{(1-\frac{3}{4}\beta+\varepsilon)\log N}{\log(c^2/2+x)}$ for any $\varepsilon > 0$ in case of invasion of the parasite population. Similarly, in the setting of Case (iii) it follows directly from the proof (in which couplings between infection processes from Case (iii) and Case (ii) are established, see Section 2.6 for more details) that the time till extinction of the host population is whp $o(\log(N))$.

With some more effort we expect that it is possible to show that in the setting of Theorem 2.2.2(ii) invasion of the host population ends whp after $\frac{(1-\beta+\varepsilon)\log N}{\log(c^2/2+x)}$ *generations. Infection by cooperation of parasites attacking vertices from different edges takes over when the number of infected hosts exceeds the level* $N^{1-\beta+\varepsilon}$*, see (the sketch of) the proof of Theorem 2.2.2 for more details, subsequently the host population should be killed whp in a finite number of generations.*

Furthermore, depending on the size of the ratio $\frac{v_N^2}{d_N}$ *invasion of the host population is considerably faster than* $\log(N)$ *in Case (iii). One shows for example easily that the host population gets whp killed after finitely many generations, if* $\frac{v_N^2}{d_N} \sim N^\gamma$ *for some* $\gamma > 0$*.*

2.2.2 Sketch of the proof of Theorem 2.2.2

In the following we will use an adaptation of the classical notation for SIR epidemics on a configuration model (see e.g. [69], Part III). Define the set of *susceptible hosts* $S_n^{(N)}$ as the set of hosts which have not been infected until generation n, the set of *infected hosts* $I_n^{(N)}$ as the set of hosts which get infected (and killed) at generation n, and the set of *removed hosts* $R_n^{(N)}$ as the set of hosts which got infected (and killed) strictly before generation n. Since each host is uniquely related to a vertex, we will sometimes also speak of susceptible vertices and infected vertices instead of susceptible and infected hosts. In addition we will call vertices whose hosts have been removed *empty* vertices.

We explore the random network of hosts while the parasites are spreading in the population. We start at the vertex that got infected initially and build up an edge between two vertices once the edge gets occupied by at least one parasite, see Figure 2.2. Half-edges and edges along which parasites move to neighbouring vertices we call *occupied half-edges* and *occupied edges*, respectively. While an half-edge can get occupied only from a single side (at which it is connected to the vertex), edges can get occupied from two sides. Half-edges and edges that have not been explored yet are called *free half-edges* and *free edges*, respectively.

We proceed by sketching first the proof in the critical parameter regime $v_N \sim c\sqrt{d_N}$ for some $c > 0$, as defined in Theorem 2.2.2 (ii). In this scaling at the beginning the number of new infections generated by hosts that got infected in the previous generation is closely related to the birthday problem. When the number of parasites is relatively small, offspring parasites from different hosts whp do not interfere and hosts get mainly infected by cooperating parasites that have been generated in the same host and move along the *same* edge, as well as by single parasites attacking successfully neighbouring hosts in the case $x > 0$. (In the following we will refer to these single parasites as *successful single parasites*.) Only at a later stage of the epidemic, when the number of infected and removed hosts exceed the level $N^{1-\beta}$ it gets likely that hosts are infected by parasites that attack the host from *different* edges. Recall that by assumption at parasite reproduction, v_N offspring parasites are generated and a host is connected over d_N half-edges to (roughly) d_N different neighbours. Hence, at the beginning the number of new infections occurring due to cooperation of parasites is for each infected host roughly given by the number of days at which at least two persons share a birthday, when the birthdays of v_N persons are independently and randomly distributed on d_N days.

If $v_N \sim c\sqrt{d_N}$ for some $c > 0$ the number of days at which at least two persons share a birthday is asymptotically $\mathrm{Pois}(\frac{c^2}{2})$-distributed. Furthermore, the number of infections initiated by successful single parasites is asymptotically $\mathrm{Pois}(x)$-distributed, if $x > 0$. Since $v_N \to \infty$, the number of host infections triggered by pairs of parasites moving along the same edge and the number of infections

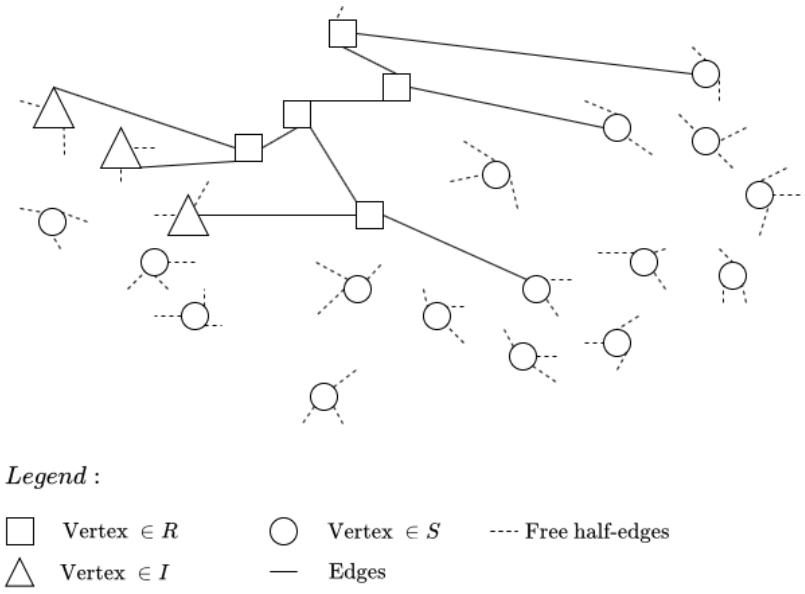

Figure 2.2: Illustration of the graph structure

generated by successful single parasites are asymptotically independent. Hence, when the number of infected hosts is still small by each infected host roughly Pois$(\frac{c^2}{2}+x)$ many new host infections are generated.
Furthermore, offspring parasites of different hosts whp do not interfere at the beginning, hence, for some time the total number of removed and infected hosts can be estimated from above and below by the total sizes of Galton-Watson-processes with offspring distributions that are close to a Pois $\left(\frac{c^2}{2}+x\right)$-distribution, see Definition 2.3.1 and 2.4.5 for a rigorous definition of these processes.

To obtain an upper bound on the invasion probability it suffices to prove that whp the total number of removed and infected hosts can be stochastically dominated by the total size of the upper Galton-Watson process until a level ℓ_N is reached, for some level ℓ_N with $\ell_N \to \infty$. Since the upper Galton-Watson process reaches any level ℓ_N with $\ell_N \to \infty$ with the probability $\pi(c,x)+o(1)$, see Proposition 2.3.3, the probability to invade the host population up to level u for $0 < u \leq 1$ is upper bounded by $\pi(c,x)+o(1)$ as well.

To derive a lower bound on the invasion probability we couple first $\mathcal{I}^{(N)}$ with a Galton Watson process $\mathcal{Z}_l^{(N)}$, such that $\overline{\mathcal{I}}$ is whp bounded from below by the total size of $\mathcal{Z}_l^{(N)}$ until no further hosts are killed or the total number of removed and infected hosts exceeds the threshold N^α, for $0 < \alpha < \beta$. As for the upper bound, the probability that the total size of the approximating Galton-Watson process exceeds the threshold N^α is asymptotically equal to $\pi(c,x)$ for any $0 < \alpha < \beta$.

In the case $\beta > \frac{4}{7}$ we can choose the level to be reached as N^α with $\alpha = 1 - \frac{3\beta}{4} + \varepsilon$ for some $\varepsilon > 0$ small enough such that $1 - \frac{3\beta}{4} + \varepsilon < \min\left\{\beta, 1 - \frac{\beta}{2}\right\}$. Once the level $N^{1-\frac{3\beta}{4}+\varepsilon}$ is crossed, whp at most two generations later the total host population gets removed, see Proposition 2.5.1. The final epidemic phase is so quick, since once at least $N^{1-\beta}$ hosts are infected, infections generated by

pairs of parasites attacking a host from different edges take over. Indeed, if $I_n^{(N)} \in \Theta(N^\gamma)$ in some generation n for some $\gamma > 0$ (and $\overline{I}_n^{(N)} \ll N$) $\Theta(v_N N^\gamma)$ offspring parasites are generated. From these parasites $\Theta(N^{2\gamma+\beta})$ pairs of parasites can be formed. The majority of these pairs consists of parasites that have been generated on different vertices. The probability that such a pair of parasites attacks the same vertex is approximately $\frac{1}{N}$. For $\gamma > 1-\beta$ we have $2\gamma+\beta-1 > \gamma$. Hence, when $\Theta(N^\gamma)$ hosts are infected for some $1-\beta < \gamma < 1$, more hosts get infected by parasites attacking a vertex from different edges than by pairs of parasites moving along the same edge. Furthermore, for $1-\frac{3\beta}{4} < \gamma < 1-\frac{\beta}{2}$ after one generation $\Theta(N^{2\gamma+\beta-1})$ hosts get infected and, since $2\gamma+\beta-1 > 1-\beta/2$ and $2(1-\frac{\beta}{2})+\beta-1 = 1$, after another generation on average all hosts get killed.

In the case $\beta \leq \frac{4}{7}$ the argument is slightly more involved, since in this case it is not possible to approximate whp $\mathcal{I}^{(N)}$ from below by the Galton-Watson process $\mathcal{Z}_l^{(N)}$ until $N^{1-\frac{3\beta}{4}+\varepsilon}$ hosts get infected. If the number of infected hosts exceeds the level N^β, then with non-trivial probability an edge is attacked from both ends simultaneously by pairs of parasites or single successful parasites. In this case none of these parasites cause an infection of a host, because the vertices to which these parasites are heading to are already empty. However, we can derive an upper bound on the number of parasites involved in such events and remove the corresponding branches in the lower Galton-Watson process. Since these parasites make up only a vanishing proportion of the total parasite population, the growth of the corresponding truncated Galton-Watson process is asymptotically the same as that of the original Galton-Watson process. Hence, for the truncated Galton-Watson process essentially the same techniques can be applied to finish the proof concerning the probability of invasion in the case $\beta \leq \frac{4}{7}$.

The details of the proof can be found in Sections 2.3 to 2.6. In Section 2.3, we are dealing with an upper bound for the invasion probability. In Section 2.4 we derive a lower bound of the probability that N^α hosts get infected for $0 < \alpha \leq 1-\frac{3}{4}\beta+\varepsilon$. In Section 2.5, we show that when $N^{1-\frac{3}{4}\beta+\varepsilon}$ hosts got infected, then whp the remaining hosts will also die in at most two generations. A detailed proof of Theorem 2.2.2 (ii) can be found in Section 2.6.

In the setting of Theorem 2.2.2(i) the number v_N of offspring parasites generated at an infection is negligible compared to $\sqrt{d_N}$. Parasites are unlikely to cooperate. Hence, invasion could only be achieved by successful single parasites. But since we are considering the parameters regime $v_N \cdot \rho_N \to x \leq 1$, successful single parasites are too rare for invasion. Hence, the parasite population infects only a negligible proportion of the host population before it dies out and so for any $u \in (0,1]$ the invasion probability is $o(1)$.

On the contrary, if the number v_N of offspring parasites is large compared to $\sqrt{d_N}$, then the infection of a single host leads to an asymptotically infinite number of further host infections. At least one of the infected hosts triggers the invasion of the host population whp.

2.2.3 Generalisations

The results of Theorem 2.2.2 can be extended to more general settings. Next we point out some of these and discuss how the proofs would need to be modified. We carry out detailed proofs only in the setting of Theorem 2.2.2 to keep the notation and proofs simple.

1.) Instead of assuming that the number d_N of half-edges per vertex and the number v_N of parasite offspring, as well as the probability ρ_N are deterministic, it would also be possible to draw

these numbers in an iid manner per vertex/host/parasite according to some distributions $\mathcal{D}^{(N)}$, $\mathcal{V}^{(N)}$ and $\mathcal{P}^{(N)}$. Our proofs can be easily adapted, if the distributions are sufficiently concentrated. More precisely, this is for example the case, if one can show that for iid random variables $(Y_i^{(N)})$ distributed as $\mathcal{D}^{(N)}$, $\mathcal{V}^{(N)}$ with corresponding expectation μ_N we have that for some $c_N \in o(\mu_N)$

$$\mathbb{P}\left(\bigcap_{i=1}^{N}\{|Y_i^{(N)} - \mu_N| \leq c_N\}\right) = \left(1 - \mathbb{P}\left(|Y_i^{(N)} - \mu_N| > c_N\right)\right)^N \to 1, \tag{2.2.1}$$

and given the total number of parasites, that can be generated, is M_N if the iid random variables $(Y_i^{(N)})$ are distributed as $\mathcal{P}^{(N)}$ we have

$$\mathbb{P}\left(\bigcap_{i=1}^{M_N}\{|Y_i^{(N)} - \mu_N| \leq c_N\}\right) = \left(1 - \mathbb{P}\left(|Y_i^{(N)} - \mu_N| > c_N\right)\right)^{M_N} \to 1.$$

This is for $\mathcal{D}^{(N)}$ for example fulfilled if $Y_1^{(N)}$ is distributed as a discretized normal distribution with mean $\mu_N \in \Theta(N^\beta)$ and variance $\sigma_N^2 \in o(N^{2\beta-\delta})$ for some $\delta > 0$ or is $Pois(N^\beta)$. If $(Y_i^{(N)})$ has a heavy-tailed distribution with mean $\mu_N = N^\beta$ and $Y_i^{(N)} - \mu_N$ has a Pareto-tail, then Condition (2.2.1) is fulfilled, if the tail is of order $\tau > \frac{1}{\beta}$. Similar distributions can be chosen for $\mathcal{V}^{(N)}$ and $\mathcal{P}^{(N)}$.

2.) While for many viruses our assumption $v_N \to \infty$ might be well justified (since viruses often generate a large number of offspring), for some host-parasite systems it might be more appropriate to assume $v_N \equiv v$. If $d_N \to \infty$, cooperative parasites whp won't invade the host population, as in Theorem 2.2.2 (i). If $d_N \equiv d$ (i.e. in a setting of a sparse graph), $v \geq 2$, $d > 2$ (for the almost sure existence of a giant component) and $\rho_N v \to x \in [0, 1]$, we expect that some (non zero) proportion of the host population can be infected with some non trivial probability (that asymptotically equals the survival probability of an appropriate GWP).
After parasite reproduction the v offspring parasites are distributed uniformly at random over the d edges. At the beginning of invasion the parasites that do not occupy the edge, over which the host that generated the offspring parasites got infected, are whp moving to a susceptible vertex. Hence, a suitable candidate for a GWP, which total size approximates the number of infected and removed hosts, should have an offspring distribution that is close to the distribution of the number of the $d-1$ edges that get occupied by at least two parasites or by single successful parasites. One would start the GWP in generation two with a number of lines that equals the random number of hosts that get infected in the first generation.
The asymptotic probability to invade a non-trivial proportion of the host population should be equal to the asymptotic survival probability of these GWPes. Given invasion a certain proportion $u, u > 0$ of the host population eventually gets infected. The level u should be bounded from below by the survival probability of a suitable approximating backward branching process, see e.g. [68] for a construction of such a backward process in the case of a Reed-Frost model. In contrast to the setting of Theorem 2.2.2 cooperation from different edges is not sufficiently strong to accelerate the order of the speed at which parasites spread at the end of the invasion process. Indeed, from ℓ_N infected hosts by cooperation from different edges of order $(\ell_N)^2/N$ further hosts get infected. This number is of the same order as the number of host that get infected by cooperation over the same edge if $\ell_N \in \Theta(\ell_N^2/N)$, i.e. only when already of order N hosts are infected.
While cooperation from different edges seems not to accelerate the speed of infection, it might lead

to the infection of a non-trivial proportion of the host population, since once of order $\Theta(N)$ hosts are infected cooperation from the same edge and cooperation from different edges contribute to the infection process on the same order.

3.) In our model we implicitly assume that bacteria get blocked only for a single generation after a phage attack. In reality this blocking may last for a longer time. In this case our result on the asymptotic of the invasion probability remains the same. Indeed, recall that as long as $\overline{I}_n^{(N)} < N^\gamma$ for some $\gamma < \frac{1}{2}(1-\beta) < 1-\beta$ the number of vertices attacked from different edges is negligible. Assuming $\overline{I}_n^{(N)} = N^\gamma$ for some $\gamma < \frac{1}{2}(1-\beta)$ we also have whp $I_n^{(N)} = \Theta(N^\gamma)$ and the probability that a blocked vertex (which number is of order $N^\gamma v_N$) is attacked by another parasite in generation n is $\mathcal{O}(N^\gamma v_N \frac{N^\gamma v_N}{N})$. This probability is non-trivial for $\gamma \geq \frac{1}{2}(1-\beta)$ (in the setting of Theorem 2.2.2(ii)). Since invasion of the host population is already decided if the frequency of infected host reaches N^ε for some $\varepsilon > 0$, at this stage of the epidemic invasion of the host population occurs anyway with probability $1 - o(1)$.

4.) In reality the number of offspring parasites generated during an infection could depend on the number of parasites infecting a host. In the scaling of Theorem 2.2.2 (ii) the probability that a host gets infected by k parasites, for $k \geq 3$, from a set of parasites of size v_N located on the same vertex scales as $N^{-\frac{(k-2)\beta}{2}}$. As long as $v_N^{(k)} N^{-\frac{(k-2)\beta}{2}} \in o(v_N)$, where $v_N^{(k)}$ is the number of offspring generated at reproduction of k parasites infecting a host, these kind of reproduction events have only a negligible impact on the initial spread of the parasite population. Hence, in this case the asymptotic of the invasion probability remains the same, since parasites generated on different vertices will start to jointly infect hosts only when the frequency of parasites is so high that whp the parasite population will invade the host population anyway.

5.) Instead of assuming that the graph on which the epidemic spreads is fixed over the whole time period, one may want to consider evolving graphs, for which edges may be rewired over time. We conjecture that for evolving graphs that rewire at most every generation the results of Theorem 2.2.2 remain valid at least if $\beta > \frac{1}{2}$.
Indeed the proof of Theorem 2.2.2 is to a large extent based on couplings with Galton-Watson processes. For these couplings the number of parasites generated at infection of a host as well as the edges, along which offspring parasites move, are assigned to the vertices independent of the generation when a host gets infected. If the graph is changing over time such a construction could lead to failures of the couplings. However as long as the number of infected hosts of the upper and resp. lower Galton-Watson process coincide exactly with the actual number of infected hosts, this construction yields couplings also for evolving graphs.
For the upper bound on the invasion probability we need the coupling to hold until time $\tau_{\ell_N,0}$ at which the GWP dies out or its total size reaches a level ℓ_N, for some sequence ℓ_N converging to ∞ arbitrarily slowly. In the proof of Theorem 2.2.2 (ii) we show, that the upper Galton-Watson process and the actual number of infected hosts coincide exactly whp until time $\tau_{\ell_N,0}$.
For the lower bound on the probability of invasion we need to couple the total number of infected hosts with the total size of the lower Galton-Watson process until it reaches the level $N^{1-\beta+\varepsilon}$ for some $\varepsilon > 0$ or the GWP dies out. When the level $N^{1-\beta+\varepsilon}$ is reached cooperation from different edges already took over and completes the invasion. The actual number of infected hosts and the number of individuals in the lower Galton-Watson process differs, when vertices get attacked from

pairs of parasites originating from different hosts. These events start to play a role when of order $\sqrt{N}$ hosts get infected. If $\beta > \frac{1}{2}$, $N^{1-\beta} \ll \sqrt{N}$, i.e. the lower GWP coincides sufficiently long with $\mathcal{I}^{(N)}$.
Similarly, one can adapt the proofs of Theorem 2.2.2 (i) and (iii) to the setting of evolving graphs. In summary, (at least) for $\beta > \frac{1}{2}$, the statements of Theorem 2.2.2 should also hold for evolving graphs.

6.) Phages that are not able to block CRISPR-resistant bacteria may have a chance to replicate in bacteria that have been blocked by ACR-phages before. However, by a similar reasoning as in item 3.) of this subsection and the sketch of the proof of Theorem 2.2.2(ii) this is only likely when the amount of this type of phages is of order $N^{1-\beta/2}$, that is this type of phages must be much more frequent than ACR-phages initially.

7.) In our model we assume that parasites that hit empty vertices keep moving further and hosts are not reproducing. These parasite have only a negligible impact on the fate of the parasite population. Hence, the statements of Theorem 2.2.2 remain valid, if we assume that parasites die (or die with a certain probability) when hitting an empty vertex.

Similarly, if hosts may reproduce (e.g. on empty nearest-neighbour spots) and the offspring numbers per host are sufficiently bounded (e.g. uniformly bounded in N) our results remain valid. Indeed, the probability that at least $N^{1-\beta+\varepsilon}$ hosts get infected is asymptotically independent on the state of the vertices on which hosts have been killed already, because the probability to re-hit these vertices is small when the overall number of infected hosts is $\ll N$. After reaching the level $N^{1-\beta+\varepsilon}$ the parasite population expands faster with every generation and in only a finite number of generations the host population gets killed whp. Host reproduction cannot curb this strong parasite expansion, when the offspring numbers are uniformly bounded in N.

8.) Instead of considering the above configuration models, we could have also considered random d_N-regular graphs. For these to exist we would need to assume that $d_N = o(\sqrt{N})$. Furthermore, biologically it seems reasonable that parasites can move from one host to another one over different routes. If several parasites move away from the same vertex this may result in multiple edges, which do not exist for random regular graphs, which makes it more difficult to motivate biologically the consideration of these graphs. Nevertheless given $d_N = o(\sqrt{N})$, we suspect the same result to hold when the configuration model is replaced by the random d_N-regular graph model since multiples edges or self loops do not play a role in the infection process.

2.3 Upper bound on the invasion probability

Consider the setting of Theorem 2.2.2 (ii). In this section we prepare all results to show that the invasion probability is asymptotically upper bounded by $\pi(c, x)$. We first introduce the Galton-Watson process $\mathcal{Z}_u^{(N)}$, see Definition 2.3.1. This process is constructed as follows. When the number of infected hosts is sufficiently small and the number of susceptible hosts is still sufficiently large, hosts most likely get infected by pairs of parasites occupying the same half-edge or by successful single parasites. Hence, we estimate the probability that an infected host infects j other hosts, for any j (not too large), by a lower bound on the sum over (k, ℓ) with $k+\ell = j$ of the probabilities that out of v_N parasites, which are originating from the same vertex, $2k$ parasites are distributed as pairs onto k different half-edges, the remaining $v_N - 2k$ parasites are distributed separately on different

half-edges and ℓ of them are successful single parasites. In all other cases we estimate the number of infected hosts by v_N which is the maximal number of hosts that can get infected by v_N parasites. We show in Proposition 2.3.2 that $\overline{\mathcal{I}}^{(N)}$ can whp be estimated from above by the total size of the Galton-Watson process $\mathcal{Z}_u^{(N)}$ until it reaches some level ℓ_N, with $\ell_N \to \infty$ and $\ell_N \in o(N)$. Only after crossing the level ℓ_N it gets likely that two parasites located on different half-edges attack the same host. In this case it could happen that $\overline{\mathcal{I}}^{(N)}$ is no longer dominated by the total size of the Galton-Watson process. However, since the level ℓ_N tends to ∞, the probability that the total size of $\mathcal{Z}_u^{(N)}$ reaches the level ℓ_N is asymptotically equal to its survival probability which is asymptotically equal to $\pi(c,x)$, see Proposition 2.3.3. Consequently, the invasion probability of the host-parasite model is asymptotically bounded from above by $\pi(c,x)$.

Definition 2.3.1. *(Upper Galton-Watson process)*
Let $0 < \delta < \frac{1}{2}$, and $a_N \to \infty$ satisfying $a_N \in o\left(\sqrt{d_N}\right)$. Let $\mathcal{Z}_u^{(N)} = \left(Z_{n,u}^{(N)}\right)_{n\in\mathbb{N}_0}$ be a Galton-Watson process with $Z_{0,u}^{(N)} = 1$ almost surely, and offspring distribution $\left(p_{j,u}^{(N)}\right)_{j\in\mathbb{N}_0}$ with

$$p_{j,u}^{(N)} := \sum_{k+\ell=j} \left(\frac{(v_N - 2a_N)^2}{2d_N}\right)^k \frac{1}{k!} \exp\left(-\frac{v_N^2}{2d_N}\right)\left(1 - \frac{1}{d_N^\delta}\right) \frac{((v_N - 2a_N)\rho_N)^\ell}{\ell!} (1-\rho_N)^{v_N} \tag{2.3.1}$$

for all $1 \le j < a_N$ and

$$p_{v_N,u}^{(N)} := 1 - \sum_{j=0}^{a_N} p_{j,u}^{(N)}. \tag{2.3.2}$$

Denote by $\overline{\mathcal{Z}}_u^{(N)} = \left(\overline{Z}_{n,u}^{(N)}\right)_{n\in\mathbb{N}_0}$ where $\overline{Z}_{n,u}^{(N)} := \sum_{i=0}^n Z_{i,u}^{(N)}$, that is $\overline{Z}_{n,u}^{(N)}$ gives the total size of $\mathcal{Z}_u^{(N)}$ accumulated till generation n.

The main results of this section are stated in the next two propositions.

Proposition 2.3.2. *(Coupling from above)*
Consider a sequence $(\ell_N)_{N\in\mathbb{N}}$ with $\ell_N \to \infty$ and $\ell_N^3 v_N^2 \in o(N)$. Introduce the stopping time

$$\tau_{\ell_N,0}^{(N)} := \inf\left\{n \in \mathbb{N}_0 : \overline{Z}_{n,u}^{(N)} \ge \ell_N \text{ or } Z_{n,u}^{(N)} = 0\right\}.$$

Then it exists a coupling between $\left(\overline{I}_n^{(N)}\right)_{n\in\mathbb{N}_0}$ and $\left(\overline{Z}_{n,u}^{(N)}\right)_{n\in\mathbb{N}_0}$ such that

$$\lim_{N\to\infty} \mathbb{P}\left(\overline{I}_n^{(N)} \le \overline{Z}_{n,u}^{(N)} \ \forall n \le \tau_{\ell_N,0}^{(N)}\right) = 1.$$

Proposition 2.3.3. *(Probability for the total size of the upper GWP to reach a level ℓ_N)*
Consider a sequence $(\ell_N)_{N\in\mathbb{N}}$ with $\ell_N \to \infty$. Then, we have

$$\lim_{N\to\infty} \mathbb{P}\left(\exists n \in \mathbb{N}_0 : \overline{Z}_{n,u}^{(N)} \ge \ell_N\right) = \pi(c,x).$$

In Subsection 2.3.1, we will prove Proposition 2.3.2. In Subsection 2.3.2 we will study (in a quite general setting) the asymptotic survival probability of a sequence of Galton-Watson processes and afterwards give the proof of Proposition 2.3.3.

2.3.1 Proof of Proposition 2.3.2

To prepare the proof of Proposition 2.3.2 we make temporarily two assumptions. First, we ignore infections of hosts by parasites attacking a vertex from different edges. In Proposition 2.3.6 we will show that this assumption is whp fulfilled as long as the number of infected and removed hosts $\overline{I}^{(N)} = R^{(N)} + I^{(N)}$ stays below a certain level ℓ_N. Secondly, we assume that all vertices that get attacked are occupied by hosts and any vertex is connected to exactly d_N different neighbouring vertices. Under the first assumption this second assumption leads to an upper bound on the number of infected hosts.

Consider a vertex that is occupied by v_N parasites. Denote by $L^{(N)}$ the random number of hosts that get removed after movement of the parasites to neighbouring vertices.
The probability distribution of $L^{(N)}$ is given by

$$\mathbb{P}(L^{(N)} = 0) = \frac{d_N!}{d_N^{v_N}(d_N - v_N)!}(1-\rho_N)^{v_N},$$

and for $k \in \mathbb{N}$

$$\begin{aligned}\mathbb{P}(L^{(N)} = k) = &\binom{v_N}{k}\frac{d_N!}{d_N^{v_N}(d_N - v_N)!}\rho_N^k(1-\rho_N)^{v_N-k} \\ &+\sum_{j=1}^{k}\sum_{\substack{k_1,\dots,k_j\geq 2 \\ k_1+\dots+k_j\leq v_N-(k-j)}}\prod_{\ell=1}^{j}\binom{v_N-(k_1+\dots+k_{\ell-1})}{k_\ell}\binom{v_N-(k_1+\dots+k_j)}{k-j} \\ &\cdot\frac{1}{\prod_{s=2}^{v_N-(k+j)+2}|\{i\in\{1,\dots,j\},k_i=s\}|!} \\ &\cdot\frac{d_N!}{(d_N-j-(v_N-(k_1+\dots+k_j)))!}\rho_N^{k-j}(1-\rho_N)^{v_N-(k-j)-(k_1+\dots+k_j)},\end{aligned}$$

because k hosts get infected after movement of v_N parasites if either all parasites move over different edges and exactly k vertices get infected by single successful parasites (and the remaining single parasites are unsuccessful) or if j for $1 \leq j \leq k$ edges get occupied by at least 2 parasites and the remaining parasites move along different edges and exactly $j-k$ of them are successful.
We have $L^{(N)} \leq v_N$ a.s. and, as for the birthday problem, the probability that $L^{(N)}$ is zero is asymptotically 1, if $v_N \in o(\sqrt{d_N})$. In the situation of Theorem 2.2.2 (ii), i.e. for $v_N \sim c\sqrt{d_N}$, with $c > 0$, the probability that $L^{(N)}$ is zero is asymptotically non-trivial.

Denote by $D_{k,\ell}^{(N)}$ the event that (under the just stated two assumptions) after parasite movement exactly $k+\ell$ hosts get infected by k pairs of parasites moving along the same edge and ℓ successful single parasites, and all the remaining parasites die without infecting a host. The next proposition states that the events $\left(D_{k,\ell}^{(N)}\right)_{k,\ell\in\mathbb{N}_0}$ are typical, while all other events occur asymptotically only with negligible probabilities.

Proposition 2.3.4. *Assume the conditions of Theorem 2.2.2 (ii) are fulfilled. Then*

$$\lim_{N\to\infty}\mathbb{P}\left(\bigcup_{k,\ell=0}^{\infty}D_{k,\ell}^{(N)}\right) = 1.$$

Proof. Let $(k,\ell) \in \mathbb{N}_0^2$. Denote by

$$w_k^{(N)} := \frac{\binom{v_N}{2}\binom{v_N-2}{2}\cdots\binom{v_N-2(k-1)}{2}}{k! d_N^k} \cdot \frac{d_N!}{d_N^{v_N-k}(d_N-(v_N-k))!} \tag{2.3.3}$$

the probability to create exactly k pairs of parasites out of v_N parasites when placing the parasites on d_N spots. We have

$$\begin{aligned} \mathbb{P}(D_{k,\ell}^{(N)}) &= w_k^{(N)} \cdot \binom{v_N-2k}{\ell} \rho_N^\ell \left(1-\rho_N\right)^{v_N-(2k+\ell)} \\ &\sim \left(\frac{c^2}{2}\right)^k \frac{1}{k!} \exp\left(-\frac{c^2}{2}\right) \cdot \frac{x^\ell}{\ell!} \exp\left(-x\right) =: p_{k,\ell}, \end{aligned} \tag{2.3.4}$$

and for all $j \in \mathbb{N}_0$

$$\mathbb{P}\left(\bigcup_{k+\ell=j} D_{k,\ell}^{(N)}\right) = \sum_{k+\ell=j} \mathbb{P}(D_{k,\ell}^{(N)}) \sim \sum_{k+\ell=j} p_{k,\ell} = \left(\frac{c^2}{2}+x\right)^j \frac{1}{j!} \exp\left(-\left(\frac{c^2}{2}+x\right)\right) := p_j, \tag{2.3.5}$$

since the sum of two independent Poisson variables is again Poisson. As the $\mathrm{Pois}\left(\frac{c^2}{2}+x\right)$-probability masses $(p_j)_{j\geq 0}$ sum up to 1, we find for all $\varepsilon > 0$ a $\widetilde{J} > 0$, such that for all $J \geq \widetilde{J}$

$$1-\varepsilon \leq \sum_{j=0}^{J} p_j \leq 1.$$

and by (2.3.5) for $\widetilde{J}$, there exists $\widetilde{N}$ such that for all $N \geq \widetilde{N}$

$$\left| \sum_{j=0}^{\widetilde{J}} \mathbb{P}\left(\bigcup_{k+\ell=j} D_{k,\ell}^{(N)}\right) - \sum_{k=0}^{\widetilde{J}} p_j \right| \leq \varepsilon.$$

Consequently

$$1-2\varepsilon \leq \sum_{j=0}^{\widetilde{J}} \mathbb{P}\left(\bigcup_{k+\ell=j} D_{k,\ell}^{(N)}\right) \leq \sum_{j=0}^{\infty} \mathbb{P}\left(\bigcup_{k+\ell=j} D_{k,\ell}^{(N)}\right) \leq 1,$$

which yields the claim since ε was arbitrary. □

We show next that the offspring distribution of the upper Galton-Watson process $\mathcal{Z}_u^{(N)}$ stochastically dominates $L^{(N)}$ for N large enough, which yields that as long as we can and do ignore infections of hosts by parasites attacking hosts from different edges, $\overline{\mathcal{I}}^{(N)}$ can be upper bounded by $\overline{\mathcal{Z}}_u^{(N)}$.

Proposition 2.3.5. *Under the assumptions of Theorem 2.2.2 (ii) the random variables $Z_{1,u}^{(N)}$ and $L^{(N)}$ can be coupled such that for N large enough*

$$\mathbb{P}(L^{(N)} \leq Z_{1,u}^{(N)}) = 1.$$

Proof. Recall that we denoted by $(p_{j,u}^{(N)})_{j\in\mathbb{N}_0}$ the offspring distribution of the GWP $\mathcal{Z}_u^{(N)}$, see (2.3.1) and (2.3.2), and we fixed a level a_N for the definition of $\mathcal{Z}_u^{(N)}$. For the proof of the proposition it suffices to show that for $j \leq a_N$ we have $p_{u,j}^{(N)} \leq \mathbb{P}\left(\bigcup_{k+\ell=j} D_{k,\ell}^{(N)}\right) \leq \mathbb{P}(L^{(N)} = j)$, since by definition $\mathbb{P}(Z_{1,u}^{(N)} = v_N) = 1 - \mathbb{P}(Z_{1,u}^{(N)} \leq a_N)$, and $L^{(N)} \leq v_N$ a.s. We have $\mathbb{P}\left(\bigcup_{k+\ell=j} D_{k,\ell}^{(N)}\right) = \sum_{k+\ell=j} \mathbb{P}\left(D_{k,\ell}^{(N)}\right)$ for all $j \in \mathbb{N}_0$. We use (2.3.4) and (2.3.3) to estimate the sum. The first factor in Equation (2.3.3) can be lower bounded by

$$\frac{\binom{v_N}{2}\binom{v_N-2}{2}\cdots\binom{v_N-2(k-1)}{2}}{k!d_N^k} \geq \left(\frac{(v_N-2a_N)^2}{2d_N}\right)^k \frac{1}{k!},$$

and the second and forth factor of the product in Equation (2.3.4) can be lower bounded by

$$\binom{v_N-2k}{\ell} \geq \frac{(v_N-2a_N)^\ell}{\ell!},$$
$$(1-\rho_N)^{v_N-(2k+\ell)} \geq (1-\rho_N)^{v_N},$$

for $k+\ell \leq a_N$.

It remains to estimate the second factor of Equation (2.3.3), i.e. $\dfrac{d_N!}{d_N^{v_N-k}(d_N-(v_N-k))!}$. Expanding the factorials up to second order we obtain

$$\frac{d_N!}{d_N^{v_N-k}(d_N-(v_N-k))!} = \exp\left(-\frac{(v_N-k)^2}{2d_N}\right)\cdot\left[1+\frac{1}{2}\frac{v_N-k}{d_N}\left(1-\frac{1}{3}\frac{(v_N-k)^2}{d_N}\right)+\mathcal{O}\left(\frac{1}{d_N}\right)\right].$$

Hence, for N large enough and $0<\delta<\frac{1}{2}$

$$\frac{d_N!}{d_N^{v_N-k}(d_N-(v_N-k))!} \geq \exp\left(-\frac{v_N^2}{2d_N}\right)\left(1-\frac{1}{d_N^\delta}\right),$$

which concludes the proof. □

So far we ignored infections of hosts by parasites attacking a vertex from different edges. Next we find a sequence of levels ℓ_N, such that (i) $\ell_N \to \infty$ and (ii) as long as the number $\overline{I}^{(N)}$ of infected and removed hosts is bounded by ℓ_N, these kind of infections are unlikely to occur.

For any $y>0$ denote by

$$\overline{\tau}_y^{(N)} := \inf\{n \in \mathbb{N}_0 : \overline{I}_n^{(N)} \geq y\},$$

the first time at which the number of infected and removed hosts exceeds the level y and by

$$\tau_D^{(N)} := \inf\{n \in \mathbb{N}_0 : \text{ a vertex of } S_n^{(N)} \text{ is hit by parasites from different edges}\}. \tag{2.3.6}$$

In the next proposition it is shown that infections of hosts by parasites attacking a vertex from different edges can be neglected as long as the number of infected and removed hosts $\overline{I}^{(N)}$ is of order $o((N/v_N^2)^{\frac{1}{3}})$.

Proposition 2.3.6. *Choose a sequence $(\ell_N)_{N\in\mathbb{N}}$, such that $\ell_N \to \infty$ and $\ell_N^3 v_N^2 \in o(N)$. Then*

$$\lim_{N\to\infty} \mathbb{P}\left(\tau_D^{(N)} \le \overline{\tau}_{\ell_N}^{(N)}, \overline{\tau}_{\ell_N}^{(N)} < \infty\right) = 0.$$

Proof. Recall that we denoted by $S_n^{(N)}, I_n^{(N)}$ and $R_n^{(N)}$ the sets of susceptible, infected and empty vertices, resp., in generation n. For the proof of the proposition we need to control the probability that a vertex is hit by at least two parasites from different edges simultaneously. We first show that it is unlikely to re-hit an already empty vertex till generation $\overline{\tau}_{\ell_N}^{(N)}$. Hence, only parasites on infected vertices remain as candidates for simultaneous infections of parasites from different edges. However, as we will show below, the number of susceptible vertices till generation $\overline{\tau}_{\ell_N}^{(N)}$ is large and each susceptible vertex has roughly d_N free half-edges. That makes it unlikely to hit a susceptible vertex simultaneously from different edges.
For a rigorous proof denote by $A_n^{(N)}$ the number of parasites on empty vertices in generation n and by

$$\tau_A^{(N)} := \inf\{n \in \mathbb{N}_0 : A_n^{(N)} \ge 1\},$$

the first generation when at least one parasite hits a vertex of $R^{(N)}$.
We show next that

$$\lim_{N\to\infty} \mathbb{P}(\tau_A^{(N)} \le \overline{\tau}_{\ell_N}^{(N)}, \overline{\tau}_{\ell_N}^{(N)} < \infty) = 0. \tag{2.3.7}$$

Let

$$\tau_{\text{no inf}}^{(N)} := \inf\{n \in \mathbb{N}_0 : \ I_n^{(N)} = 0\},$$

be the first generation at which no host gets infected. Note that at generation $\tau_{\text{no inf}}^{(N)}$ the infection process is not necessarily finished, as parasites may remain on empty vertices. However, this is whp not the case if $\tau_{\text{no inf}}^{(N)} < \overline{\tau}_{\ell_N}^{(N)}$. More precisely we claim,

$$\lim_{N\to\infty} \mathbb{P}(\tau_A^{(N)} \le \tau_{\text{no inf}}^{(N)} \wedge \overline{\tau}_{\ell_N}^{(N)}) = 0. \tag{2.3.8}$$

Given we have shown (2.3.8), we also have (2.3.7), since

$$\{\tau^{(N)}_{\text{no inf}} < \tau^{(N)}_A \leq \overline{\tau}^{(N)}_{\ell_N}, \overline{\tau}^{(N)}_{\ell_N} < \infty\} = \emptyset,$$

and hence

$$\lim_{N\to\infty} \mathbb{P}(\tau^{(N)}_A \leq \overline{\tau}^{(N)}_{\ell_N}, \overline{\tau}^{(N)}_{\ell_N} < \infty) = \lim_{N\to\infty} \mathbb{P}(\tau^{(N)}_A \leq \tau^{(N)}_{\text{no inf}} \wedge \overline{\tau}^{(N)}_{\ell_N}, \overline{\tau}^{(N)}_{\ell_N} < \infty).$$

So, lets prove (2.3.8). First of all we have by definition of $\tau^{(N)}_{\text{no inf}}$ that $\tau^{(N)}_{\text{no inf}} \wedge \overline{\tau}^{(N)}_{\ell_N} \leq \ell_N$. Furthermore, the number of parasites generated in some generation n with $n \leq \tau^{(N)}_{\text{no inf}} \wedge \overline{\tau}^{(N)}_{\ell_N}$ is bounded by $\ell_N v_N$ and the total number of half-edges formed for vertices of the set $R^{(N)}_n$ is at most $\ell_N \cdot d_N$. The number of half-edges not yet connected to other half-edges in the graph is at least as large as the number of free half-edges of the vertices in the set $S^{(N)}_n$, which is bounded from below by $(N-\ell_N)d_N - \ell_N v_N \geq (N-2\ell_N)d_N$. (Note that the summand $-\ell_N v_N$ has to be added to account for the potential attacks that do not lead to an infection of a host). Hence, the number of parasites that move to an empty vertex in any generation n with $n \leq \tau^{(N)}_{\text{no inf}} \wedge \overline{\tau}^{(N)}_{\ell_N}$ can be estimated from above by the following iid random variables $(H^{(N)}_n)_{n\in\mathbb{N}}$. Assume for each n (independently of each other), $\ell_N v_N$ numbers are chosen randomly and without replacement from the set $\{1, ..., (N-2\ell_N)d_N\}$. Let $H^{(N)}_n$ count the numbers falling into the set $\{1, ..., \ell_N d_N\}$. Then we have

$$\begin{aligned}
\mathbb{P}\left(\tau^{(N)}_A \leq \overline{\tau}^{(N)}_{\ell_N} \wedge \tau^{(N)}_{\text{no inf}}\right) &\leq \mathbb{P}\left(\exists\, n \leq \ell_N : H^{(N)}_n \geq 1\right) \\
&\leq \ell_N \mathbb{P}\left(H^{(N)}_1 \geq 1\right) \\
&\leq \ell_N \left(1 - \frac{N_1!}{(N_1 - l_N v_N)!} \cdot \frac{1}{((N-2\ell_N)d_N)^{\ell_N v_N}}\right),
\end{aligned}$$

where $N_1 := (N-3\ell_N)d_N$. Using an asymptotic expansion of the factorial, we get

$$\frac{N_1!}{(N_1 - \ell_N v_N)!} \cdot \frac{1}{((N-2l_N)d_N)^{\ell_N v_N}} = 1 + \mathcal{O}\left(\frac{\ell_N^2 v_N^2}{N}\right),$$

so using the assumption $\ell_N^3 v_N^2 = o(N)$, we have proven Equation (2.3.8).

To finish the proof of the proposition it remains to show that susceptible vertices are not hit simultaneously by parasites from different edges before generation $\overline{\tau}^{(N)}_{\ell_N}$. Recall the definition of $\tau^{(N)}_D$ in (2.3.6). If $\overline{\tau}^{(N)}_{\ell_N} > \tau^{(N)}_{\text{no inf}}$, then using (2.3.7) whp $\overline{\tau}^{(N)}_{\ell_N} = \infty$ and hence it suffices to show

$$\lim_{N\to\infty} \mathbb{P}\left(\tau^{(N)}_D \leq \overline{\tau}^{(N)}_{\ell_N} \wedge \tau^{(N)}_{\text{no inf}}, \overline{\tau}^{(N)}_{\ell_N} < \infty\right) = 0.$$

Denote by $S^{(N)}_{n,\text{free}}$ the set of susceptible vertices for which all half-edges are still free. As before the number of parasites in the graph is smaller than $\ell_N v_N$ for any generation n with $n < \overline{\tau}^{(N)}_{\ell_N} \wedge \tau^{(N)}_{\text{no inf}}$ and $|S^{(N)}_{n,\text{free}}| \geq N - \ell_N v_N$. Define this time the following sequence of iid random variables $(G^{(N)}_n)_{n\in\mathbb{N}}$. Consider $N - \ell_N v_N$ boxes each containing d_N balls. Assume (for each n independently) $\ell_N v_N$ balls

are drawn randomly and without replacement out of the boxes (that are refilled for each n). Let $G_n^{(N)}$ be the number of boxes from which at least two balls were drawn. Then we can estimate

$$\begin{aligned}\mathbb{P}\left(\tau_D^{(N)} \leq \overline{\tau}_{\ell_N}^{(N)} \wedge \tau_{\text{no inf}}^{(N)}, \overline{\tau}_{\ell_N}^{(N)} < \infty\right) &\leq \mathbb{P}\left(\exists\, n \leq \ell_N : G_n^{(N)} \geq 1\right) \\ &\leq \ell_N \mathbb{P}\left(G_1^{(N)} \geq 1\right) \\ &= \ell_N \left(1 - d_N^{\ell_N v_N} \cdot \frac{N_2!}{(N_2 - \ell_N v_N)!} \cdot \frac{(N_2 d_N - \ell_N v_N)!}{(N_2 d_N)!}\right) \quad (2.3.9)\end{aligned}$$

where $N_2 := N - \ell_N v_N$. Using an asymptotic expansion of the factorial, we get

$$d_N^{\ell_N v_N} \cdot \frac{N_2!}{(N_2 - \ell_N v_N)!} \cdot \frac{(N_2 d_N - \ell_N v_N)!}{(N_2 d_N)!} = 1 + \mathcal{O}\left(\frac{(\ell_N v_N)^2}{N}\right),$$

which shows that the left hand side of (2.3.9) converges to 0. □

Proof of Proposition 2.3.2. By Proposition 2.3.6 whp no infection of hosts by parasites attacking from different edges occurs till $\overline{\mathcal{I}}^{(N)}$ reaches the level ℓ_N for any sequence $(\ell_N)_{N\in\mathbb{N}}$ with $\ell_N \to \infty$ and $\ell_N^3 v_N^2 \in o(N)$. Hence, it suffices to consider the case that such infections do not occur and Proposition 2.3.5 can be applied. Consequently, as long as $\overline{\mathcal{I}}^{(N)}$ has not reached the level ℓ_N, the number of hosts that get infected from an infected vertex in the next generation can whp be estimated from above by the offspring number of the GWP $\mathcal{Z}_u^{(N)}$, which yields the claim of Proposition 2.3.2. □

2.3.2 Asymptotic survival probabilities of sequences of GWPes and the proof of Proposition 2.3.3

Before we give the proof of Proposition 2.3.3 we establish some general results about the asymptotic survival probability of a sequence of Galton-Watson processes.

Consider a Galton-Watson process $\mathcal{Z} = (Z_n)_{n\in\mathbb{N}_0}$ with offspring distribution $(p_k)_{k\in\mathbb{N}_0}$ and with $Z_0 = 1$ almost surely, and a sequence of Galton-Watson processes $\mathcal{Z}^{(N)} = \left(Z_n^{(N)}\right)_{n\in\mathbb{N}_0}$ with offspring distributions $\left(p_k^{(N)}\right)_{k\in\mathbb{N}_0}$ and with $Z_0^{(N)} = 1$ almost surely, for all $N \in \mathbb{N}$.

Denote by Φ and $\Phi^{(N)}$, resp., the probability generating functions of the offspring distributions $(p_k)_{k\in\mathbb{N}_0}$ and $\left(p_k^{(N)}\right)_{k\in\mathbb{N}_0}$, by π and $\pi^{(N)}$ the corresponding survival probabilities, and by $q := 1 - \pi$ and $q^{(N)} := 1 - \pi^{(N)}$ the corresponding extinction probabilities. Denote also by $\overline{\mathcal{Z}}^{(N)} = \left(\overline{Z}_n^{(N)} := \sum_{i=0}^n Z_i^{(N)}\right)_{n\in\mathbb{N}_0}$ the process that counts the total size of the GWP $\mathcal{Z}^{(N)}$ till generation n.

Recall that $\Phi^{(N)}$ converges uniformly to Φ, if the corresponding offspring distributions converge in total variation distance, in particular, if

$$\sum_{k=0}^{\infty} |q_k^{(N)} - q_k| \to 0,$$

see [71], Proposition 4.2, or as one readily checks, if there exists an $\mathbb{N}-$valued sequence $(K_N)_{N\in\mathbb{N}}$ with $K_N \to \infty$ such that

$$\sum_{k=0}^{K_N} |q_k^{(N)} - q_k| \to 0.$$

Lemma 2.3.7. *Consider the just defined Galton-Watson processes $\mathcal{Z}$ and $(\mathcal{Z}^{(N)})_{N\in\mathbb{N}}$. Furthermore, let $(a_N)_{N\in\mathbb{N}}$ be an $\mathbb{N}$-valued sequence with $a_N \to \infty$. Assume that the generating functions $(\Phi^{(N)})_{N\in\mathbb{N}}$ converge uniformly in $[0,1]$ to Φ. Then the following holds:*
a)

$$|\pi^{(N)} - \pi| \to 0,$$

b)

$$\mathbb{P}\left(Z_{a_N}^{(N)} = 0\right) \to q,$$

c)

$$\mathbb{P}\left(\exists n \in \mathbb{N}_0 : Z_n^{(N)} \geq a_N\right) \to \pi,$$

d)

$$\mathbb{P}\left(\exists\ n \in \mathbb{N}_0 : \overline{Z}_n^{(N)} \geq a_N\right) \to \pi.$$

Proof. We show a detailed proof in the case $\pi > 0$, with analogous arguments one also shows the claim in the case $\pi = 0$. Recall that the extinction probabilities q and $q^{(N)}$ are characterised as the smallest fixed points in $[0,1]$ of the generating functions Φ and $\Phi^{(N)}$ respectively. Consider the function

$$g(s) := \Phi(q+s) - (q+s),$$

for $s \in [-q, 1-q]$. We have $g(s) = 0$, iff $s = 0$ or $s = 1-q$. Furthermore $g > 0$ for $s < 0$ and g is decreasing up to some $s_0 > 0$.
Let $0 < \varepsilon < s_0$, and

$$\eta < \min\{g(-\varepsilon), -g(\varepsilon)\}.$$

Since by assumption $\Phi^{(N)}$ converges uniformly to Φ we find an $N_0 \in \mathbb{N}$ such that for all $N \geq N_0$

$$|\Phi^{(N)}(s) - \Phi(s)| < \eta, \tag{2.3.10}$$

for all $s \in [0,1]$ and hence for all $N \geq N_0$

$$\begin{aligned}\Phi^{(N)}(q-\varepsilon) &\geq \Phi(q-\varepsilon) - \eta = g(-\varepsilon) + q - \varepsilon - \eta > q - \varepsilon,\\ \Phi^{(N)}(q+\varepsilon) &\leq \Phi(q+\varepsilon) + \eta = g(\varepsilon) + q + \varepsilon + \eta < q + \varepsilon.\end{aligned}$$

Since $\Phi^{(N)}$ is monotonically increasing on $[0,1]$ and continuous, the smallest non-negative fixed point of $\Phi^{(N)}$ is contained in the interval $[q-\varepsilon, q+\varepsilon]$ which implies a).
Denote by $(\Phi \pm \eta)(s) := \Phi(s) \pm \eta$, and $(\Phi \pm \eta)_n(s) := (\Phi \pm \eta) \circ \cdots \circ (\Phi \pm \eta)(s)$ the n-fold composition of $(\Phi \pm \eta)$. An iterated application of (2.3.10) yields for all $n \in \mathbb{N}$

$$(\Phi + \eta)_n(0) \geq \Phi_n^{(N)}(0) \geq (\Phi - \eta)_n(0).$$

The sequences $((\Phi - \eta)_n(0))_{n\in\mathbb{N}}$ and $((\Phi + \eta)_n(0))_{n\in\mathbb{N}}$ are increasing and converge for $n \to \infty$ to the smallest non-negative fixed point of $\Phi - \eta$ and $\Phi + \eta$, respectively. While the fixed point of $\Phi - \eta$ is larger than $q - \varepsilon$, by definition of η, the fixed point of $\Phi + \eta$ is smaller than $q + \varepsilon$.
In particular, we have that there exists $\widetilde{n} \in \mathbb{N}$, such that for all $N \geq N_0$ and for all $n \geq \widetilde{n}$

$$q - \varepsilon \leq \Phi_n^{(N)}(0) \leq q + \varepsilon.$$

Since $a_N \to \infty$, there exists $N_1 \in \mathbb{N}$ such that $\forall N \geq N_1, a_N \geq \widetilde{n}$.
Finally we have for all $N \geq N_2 := \max\{N_0, N_1\}$

$$q - \varepsilon \leq \Phi_{a_N}^{(N)}(0) \leq q + \varepsilon,$$

which proves b).
The extinction-explosion principle for Galton-Watson processes yields

$$\mathbb{P}\left(Z_n^{(N)} > 0 \ \forall n \in \mathbb{N}_0\right) \leq \mathbb{P}\left(\exists\, n \in \mathbb{N}_0 : Z_n^{(N)} \geq a_N\right).$$

Hence, by a)

$$\pi + o(1) \leq \mathbb{P}\left(\exists\, n \in \mathbb{N}_0 : Z_n^{(N)} \geq a_N\right). \tag{2.3.11}$$

Furthermore

$$\begin{aligned}
\pi^{(N)} &= \mathbb{P}\left(Z_n^{(N)} > 0 \ \forall n \in \mathbb{N}_0\right) \\
&= \mathbb{P}\left(\left\{\exists\, n \in \mathbb{N}_0 :\ Z_n^{(N)} \geq a_N\right\} \cap \left\{Z_n^{(N)} > 0 \ \forall n \in \mathbb{N}_0\right\}\right) \\
&\geq \mathbb{P}\left(\exists\, n \in \mathbb{N}_0 : Z_n^{(N)} \geq a_N\right) \cdot \left(1 - \left(q^{(N)}\right)^{a_N}\right).
\end{aligned}$$

By a) we have that $\mathcal{Z}^{(N)}$ is supercritical for N large enough, which implies

$$\left(q^{(N)}\right)^{a_N} \to 0.$$

Consequently

$$\mathbb{P}\left(\exists\, n \in \mathbb{N}_0 :\ Z_n^{(N)} \geq a_N\right) \leq \frac{\pi^{(N)}}{1 - \left(q^{(N)}\right)^{a_N}} = \pi^{(N)} \cdot (1 + o(1)) = \pi + o(1),$$

which, together with (2.3.11), concludes the proof of c).
For proving d), it only remains to show that

$$\mathbb{P}\left(\left\{\exists\, n \in \mathbb{N}_0 : \overline{Z}_n^{(N)} \geq a_N\right\} \cap \left\{\exists\, n \in \mathbb{N}_0 : Z_n^{(N)} = 0\right\}\right) = o(1).$$

Let $(c_N)_{N\in\mathbb{N}}$ be a sequence with $c_N \to \infty$ and $\frac{a_N}{c_N} \to \infty$ and consider the subsets

$$A^{(N)} := \left\{\exists\, n \in \mathbb{N}_0 : \overline{Z}_n^{(N)} \geq a_N,\ \exists\, i \leq n : Z_i^{(N)} \geq c_N\right\} \cap \left\{\exists\, n \in \mathbb{N}_0 : Z_n^{(N)} = 0\right\},$$
$$B^{(N)} := \left\{\exists\, n \in \mathbb{N}_0 : \overline{Z}_n^{(N)} \geq a_N,\ Z_i^{(N)} < c_N\ \forall i \leq n\right\} \cap \left\{\exists\, n \in \mathbb{N}_0 : Z_n^{(N)} = 0\right\}.$$

By definition

$$\left\{\exists\, n \in \mathbb{N}_0 : \overline{Z}_n^{(N)} \geq a_N\right\} \cap \left\{\exists\, n \in \mathbb{N}_0 :\ Z_n^{(N)} = 0\right\} = A^{(N)} \sqcup B^{(N)}.$$

According to c) we have

$$\mathbb{P}\left(A^{(N)}\right) \leq \mathbb{P}\left(\left\{\exists i \in \mathbb{N}_0 : Z_i^{(N)} \geq c_N\right\} \cap \left\{\exists n \in \mathbb{N}_0 : Z_n^{(N)} = 0\right\}\right) \to 0,$$

Furthermore

$$B^{(N)} \subset \left\{Z^{(N)}_{\lfloor \frac{a_N}{c_N} \rfloor} > 0\right\} \cap \left\{\exists n \in \mathbb{N}_0 : Z_n^{(N)} = 0\right\},$$

so according to a) and b) applied with the sequence $\left(\lfloor \frac{a_N}{c_N} \rfloor\right)_{N\in\mathbb{N}}$ we get

$$\begin{aligned}
\mathbb{P}\left(B^{(N)^c}\right) \geq& \mathbb{P}\left(\left\{Z^{(N)}_{\lfloor \frac{a_N}{c_N} \rfloor} = 0\right\} \sqcup \left\{Z_n^{(N)} > 0\ \forall n \in \mathbb{N}_0\right\}\right) \\
&= \mathbb{P}\left(\left\{Z^{(N)}_{\lfloor \frac{a_N}{c_N} \rfloor} = 0\right\}\right) + \mathbb{P}\left(\left\{Z_n^{(N)} > 0\ \forall n \in \mathbb{N}_0\right\}\right) \\
&= q + o(1) + \pi + o(1) \\
&= 1 - o(1),
\end{aligned}$$

which yields $\mathbb{P}\left(A^{(N)} \sqcup B^{(N)}\right) \to 0$. □

We are now ready to prove Proposition 2.3.3.

Proof of Proposition 2.3.3. By Lemma 2.3.7 d) it suffices to show that the sequence of generating functions $\Phi_u^{(N)}$ belonging to the offspring distributions $\left(p_{j,u}^{(N)}\right)_{j\in\mathbb{N}_0}$ of $\mathcal{Z}_u^{(N)}$ converges uniformly on $[0,1]$ to the generating function $\Phi^{(c,x)}$ of the Pois $\left(\frac{c^2}{2} + x\right)$-distribution. We will denote by $(p_j)_{j\in\mathbb{N}_0}$ the probability weights of the Pois $\left(\frac{c^2}{2} + x\right)$-distribution. According to the remark just before Lemma 2.3.7 it suffices to find a sequence $(K_N)_{N\in\mathbb{N}}$ with $K_N \to \infty$ for which $\sum_{j=0}^{K_N} |p_{j,u}^{(N)} - p_j| \to 0$. We set $K_N = a_N$ and use in the following calculation the asymptotics

$$\left(\frac{(v_N-2a_N)^2}{2d_N}\right)^k \exp\left(\frac{-v_N^2}{2d_N}\right) = \left(\frac{c^2}{2}\right)^k \exp\left(\frac{-c^2}{2}\right)(1-h_N)^{k+1},$$
$$((v_N-2a_N)\rho_N)^\ell (1-\rho_N)^{v_N} = x^\ell \exp(-x)(1+o(1))^{\ell+1},$$

where $(h_N)_{N\in\mathbb{N}}$ denotes some appropriate sequence of order $\mathcal{O}\left(\frac{\max\{a_N, r_N\}}{\sqrt{d_N}}\right)$ and $r_N := v_N - c\sqrt{d_N}$. For all $j \geq 0$

$$|p_{j,u}^{(N)} - p_j| \leq \sum_{k+\ell=j} \frac{1}{k!}\left(\frac{c^2}{2}\right)^k \frac{x^\ell}{\ell!} \exp\left(-\left(\frac{c^2}{2}+x\right)\right) \Big| (1-h_N)^{k+1}(1+o(1))^{\ell+1} - 1\Big|.$$

The last term can be upper bounded in the following way

$$\begin{aligned}
|(1-h_N)^{k+1}(1+o(1))^{\ell+1} - 1| &\leq h_N \Big|\sum_{i=1}^{k+1}\binom{k+1}{i}(-h_N)^{i-1}\Big| + o(1)\Big|\sum_{i=1}^{\ell+1}\binom{\ell+1}{i}o(1)^{i-1}\Big| \\
&\quad + o(1)h_N\Big|\sum_{i=1}^{k+1}\binom{k+1}{i}(-h_N)^{i-1}\sum_{i=1}^{\ell+1}\binom{\ell+1}{i}o(1)^{i-1}\Big| \\
&\leq 3\max\{h_N, o(1)\}2^{k+\ell+2}.
\end{aligned}$$

It follows that

$$\begin{aligned}
\sum_{j=0}^{a_N}|p_{j,u}^{(N)} - p_j| &\leq \sum_{j=0}^{a_N} 12\max\{h_N, o(1)\}\sum_{k+\ell=j}\frac{1}{k!}\left(\frac{c^2}{2}\right)^k\frac{x^\ell}{\ell!}\exp\left(-\left(\frac{c^2}{2}+x\right)\right)2^{k+\ell} \\
&\leq 12\max\{h_N, o(1)\}\exp\left(\frac{c^2}{2}+x\right) \\
&\to 0,
\end{aligned}$$

which ends the proof. □

2.4 Coupling from below with (truncated) Galton-Watson processes

2.4.1 Establishing invasion

Consider again the setting of Theorem 2.2.2 (*ii*). The next proposition gives a lower bound on the probability that the parasite population infects at least N^α hosts for $0 < \alpha < \beta$.

Proposition 2.4.1. *Consider the setting of Theorem 2.2.2 (ii) and let* $0 < \alpha < \beta$. *Then*

$$\liminf_{N\to\infty} \mathbb{P}\left(\exists n \in \mathbb{N}_0 : \overline{I}_n^{(N)} \geq N^\alpha\right) \geq \pi(c,x).$$

Remark 2.4.2. *Proposition 2.4.1 together with the results from Section 2.3 yield*

$$\lim_{N\to\infty} \mathbb{P}\left(\exists n \in \mathbb{N}_0 : \overline{I}_n^{(N)} \geq N^\alpha\right) = \pi(c,x).$$

The remainder of this subsection is devoted to the proof of Proposition 2.4.1, which is given at the end of this subsection. First we introduce a simpler host-parasite model, see Definition 2.4.3, that lower bounds the number of infected and removed hosts $\overline{I}^{(N)}$ of the original host-parasite model a.s. In this model hosts can get infected only by pairs of parasites moving along the same edge or by successful single parasites. In the following, we will refer to either a pair of parasites moving along the same edge or a successful single parasite as an *infecting unit*. We show then that whp the simpler process can be coupled with a Galton-Watson process from below until N^α hosts get infected, see Proposition 2.4.7. The total size of this lower Galton-Watson process reaches any level ℓ_N where $\ell_N \to \infty$ with asymptotic probability $\pi(c,x)$, see Lemma 2.4.6, in particular the level N^α. This yields the claimed lower bound.

Definition 2.4.3 (A simpler model involving only infecting units)**.** *For a sequence of parameters $(N, d_N, v_N, \rho_N)_{N\in\mathbb{N}}$ introduce the following host-parasite model defined on the same random configuration model (with N vertices and d_N half-edges per vertex) as the original model. Initially on each vertex a single host is placed. We start the infection process by infecting a randomly chosen host. A random number of infecting units is generated according to the following distribution with probability weights $(p_j^{(N)})_{j\in\mathbb{N}_0}$ where for all $1 \leq j \leq v_N$*

$$p_j^{(N)} := \sum_{k+\ell=j, k\leq\lfloor v_N/2\rfloor} w_k^{(N)} \cdot \binom{v_N - 2k}{\ell} \rho_N^\ell \left(1-\rho_N\right)^{v_N-(2k+\ell)}, \tag{2.4.1}$$

and

$$p_0^{(N)} := 1 - \sum_{j=1}^{v_N} p_k^{(N)},$$

where $w_k^{(N)}$ denotes the probability defined in (2.3.3). Afterwards, the host dies and the infection process continues in discrete generations as follows. At the beginning of each generation, infecting units move, independently of each other, to nearest neighbour vertices along different, randomly chosen edges. If a host is attacked by at least one infecting unit, then the host gets infected. In each infected host, independently a random number of infecting units is produced according to the distribution $(p_j^{(N)})_{j\in\mathbb{N}_0}$. Afterwards the infected hosts and all the infecting units that infected the hosts die. If an infecting unit moves to an empty vertex, then it dies.
Denote by $J_n^{(N)}$ the number of hosts that get infected at generation n in this simpler model and the epidemic process by $\mathcal{J}^{(N)} = (J_n^{(N)})_{n\in\mathbb{N}_0}$. Furthermore we denote by $\overline{J}_n^{(N)} = \sum_{i=0}^n J_i^{(N)}$ the total number of hosts infected till generation n in this simpler host-parasite model and by $\overline{\mathcal{J}}^{(N)} = (\overline{J}_n^{(N)})_{n\in\mathbb{N}_0}$ the corresponding process.

Proposition 2.4.4. *For all $N \in \mathbb{N}$ it is possible to couple $\mathcal{J}^{(N)}$ and $\mathcal{I}^{(N)}$ such that almost surely $\forall n \in \mathbb{N}_0$*

$$\overline{J}_n^{(N)} \leq \overline{I}_n^{(N)}.$$

Proof. Consider the same realisation of the configuration model for both host-parasite models and assume that the same host gets initially infected.
Enumerate the d_N half-edges of each vertex and denote by $V_i^{(N)} \in \{0,\dots,v_N\}^{d_N}$ the occupancy vector of the half-edges linked to vertex i (when host i gets infected) by the v_N offspring parasites generated at its infection in the original host-parasite model. By definition, the random variables $(V_i^{(N)})_{1\le i\le N}$ are iid. A coupling of $\overline{\mathcal{I}}^{(N)}$ and $\overline{\mathcal{J}}^{(N)}$ is obtained as follows. Use the same occupancy vector $V_i^{(N)}$ when host i gets infected for the simpler host-parasite model but modify it as follows: Assume that in the original and in the simpler model the same single parasites are chosen to be successful and apply the subsequent rules:

- If exactly k pairs of parasites occupy k different half-edges, the remaining parasites move separately along different half-edges, and if exactly ℓ of them are successful single parasites, for some $0 \le k \le \lfloor v_N/2 \rfloor$ and $0 \le \ell \le v_N$ such that $0 \le 2k+\ell \le v_N$, then in the simpler model all pairs of parasites and successful single parasites are kept and the remaining parasites are removed.
- If according to the occupancy vector $V_i^{(N)}$ at least one half-edge is occupied by at least three parasites, update $V_i^{(N)}$ for the simpler host-parasite model by removing all parasites, i.e. in particular no pairs of parasites or successful single parasite remain.

With this procedure the number of infecting units is distributed according to the distribution given in (2.4.1). Moreover, hosts get either simultaneously infected in both host-parasite models or first in the original model and later possibly also in the simpler model. Hence, the number of infected hosts in the simpler model is bounded from above by $\overline{I}_n^{(N)}$ in any generation n. □

Our next step is to couple $\mathcal{J}^{(N)}$ with the Galton-Watson process $\mathcal{Z}_l^{(N)}$ which is defined next.

Definition 2.4.5 (Lower Galton-Watson Process)**.** *Let $0 < \delta < \frac{1}{2}$ and $(a_N)_{N\in\mathbb{N}}$ be a sequence with $a_N \to \infty$ and $a_N \in o\left(\sqrt{d_N}\right)$. Furthermore assume $(\theta_N)_{N\in\mathbb{N}}$ is a $[0,1]$-valued sequence with $\theta_N \to 0$. Let $\mathcal{Z}_l^{(N)} = \left(Z_{n,l}^{(N)}\right)_{n\in\mathbb{N}_0}$ be a Galton-Watson process with mixed binomial offspring distribution $Bin\left(\widetilde{Z}^{(N)}, 1-\theta_N\right)$, where the probability weights $\left(\widetilde{p}_k^{(N)}\right)_{k\in\mathbb{N}_0}$ of $\widetilde{Z}^{(N)}$ are for all $1 \le j \le a_N$*

$$\widetilde{p}_j^{(N)} := \sum_{k+\ell=j} \left(\frac{(v_N-2a_N)^2}{2d_N}\right)^k \frac{1}{k!} \exp\left(-\frac{v_N^2}{2d_N}\right)\left(1-\frac{1}{d_N^\delta}\right) \frac{((v_N-2a_N)\rho_N)^\ell}{\ell!} (1-\rho_N)^{v_N},$$

and

$$\widetilde{p}_0^{(N)} := 1 - \sum_{j=1}^{a_N} \widetilde{p}_j^{(N)}.$$

Denote by $\Phi_l^{(N)}$ the generating function of the offspring distribution $\left(p_{k,l}^{(N)}\right)_{k\in\mathbb{N}_0}$ of $\mathcal{Z}_l^{(N)}$, and by $\pi_l^{(N)}$ and $q_l^{(N)}$ the survival and extinction probability of $\mathcal{Z}_l^{(N)}$. Furthermore, denote by $\overline{Z}_{n,l}^{(N)} := \sum_{i=0}^n Z_{i,l}^{(N)}$ the total size of the Galton-Watson process until generation n and $\overline{\mathcal{Z}}_l = \left(\overline{Z}_{n,l}^{(N)}\right)_{n\in\mathbb{N}_0}$ the corresponding process.

Lemma 2.4.6. *Let $(\ell_N)_{N\in\mathbb{N}}$ be a sequence with $\ell_N \to \infty$. Assume $Z_{0,l}^{(N)} = 1$ a.s. Then*

$$\lim_{N\to\infty} \mathbb{P}\left(\exists n \in \mathbb{N}_0 : \overline{Z}_{n,l}^{(N)} \geq \ell_N\right) = \pi(c, x).$$

Proof. We proceed as in the proof of Proposition 2.3.3 and show that

$$\sum_{j=1}^{a_N} |\widetilde{p}_j^{(N)} - p_j| \to 0,$$

where $(p_j)_{j\in\mathbb{N}_0}$ denote the probability weights of the Pois$(c^2/2 + x)$-distribution. Using the same asymptotics as in the proof of Proposition 2.3.3, we have for all $j \geq 1$

$$|\widetilde{p}_j^{(N)} - p_j| \leq \sum_{k+\ell=j} \frac{1}{k!}\left(\frac{c^2}{2}\right)^k \frac{x^\ell}{\ell!} \exp\left(-\left(\frac{c^2}{2} + x\right)\right) \left| (1-h_N)^{k+1} (1+o(1))^{\ell+1} - 1\right|,$$

where $(h_N)_N$ is an appropriate sequence with $h_N = \mathcal{O}\left(\frac{\max\{a_N, r_N\}}{\sqrt{d_N}}\right)$. As in the proof of Proposition 2.3.3, the last term can be upper bounded by

$$|(1-h_N)^{k+1}(1+o(1))^{\ell+1} - 1| \leq 3\max\{h_N, o(1)\} 2^{k+\ell+2}.$$

It follows that

$$\begin{aligned}
\sum_{j=1}^{a_N} |\widetilde{p}_j^{(N)} - p_j| &\leq \sum_{j=1}^{a_N} 12\max\{h_N, o(1)\} \sum_{k+\ell=j} \frac{1}{k!}\left(\frac{c^2}{2}\right)^k \frac{x^\ell}{\ell!} \exp\left(-\left(\frac{c^2}{2}+x\right)\right) 2^{k+\ell} \\
&\leq 12\max\{h_N, o(1)\} \exp\left(\frac{c^2}{2} + x\right) \to 0,
\end{aligned}$$

which also implies that $|\widetilde{p}_0^{(N)} - p_0| \to 0$, because $a_N \to \infty$. Furthermore, we can estimate

$$\begin{aligned}
\sum_{i=1}^{a_N} |p_{i,l}^{(N)} - \widetilde{p}_i^{(N)}| &\leq \sum_{i=1}^{a_N} \widetilde{p}_i^{(N)} |1 - (1-\theta_N)^i| + \sum_{i=1}^{\infty} \sum_{j\geq i+1} \widetilde{p}_j^{(N)} \binom{j}{i} (1-\theta_N)^i \theta_N^{j-i} \\
&\leq \theta_N \sum_{i=1}^{\infty} \widetilde{p}_i^{(N)} 2^i + \sum_{j=2}^{\infty} \widetilde{p}_j^{(N)} \sum_{i=1}^{j-1} \binom{j}{i} (1-\theta_N)^i \theta_N^{j-i} \\
&\leq \theta_N \left(\sum_{i=0}^{\infty} \frac{1}{i!} \left(\frac{c^2}{2} + x\right)^i \exp\left(\frac{-c^2}{2} + x\right) 2^i + \sum_{i=1}^{\infty} |\widetilde{p}_i^{(N)} - p_i| 2^i \right) \\
&\quad + \sum_{j=2}^{\infty} \widetilde{p}_j^{(N)} (1 - (1-\theta_N)^j) \\
&\leq \theta_N \left[\exp\left(\frac{c^2}{2} + x\right) + 12\max\{h_N, o(1)\} \exp\left(\frac{c^2}{2} + x\right) \right] + \theta_N \sum_{j=2}^{\infty} \widetilde{p}_j^{(N)} 2^j
\end{aligned}$$

$$\leq 2\theta_N \left[\exp\left(\frac{c^2}{2}+x\right) + 12\max\{h_N, o(1)\}\exp\left(\frac{c^2}{2}+x\right)\right]$$
$$\to 0,$$

which implies $|p_{0,l}^{(N)} - \widetilde{p}_0| \to 0$ as well. An application of the triangle inequality ends the proof. □

Next we show that the process $\overline{\mathcal{Z}}_l^{(N)}$ indeed bounds from below the number of infected hosts $\overline{\mathcal{J}}^{(N)}$ in the simpler host-parasite model. Recall that $d_N \in \Theta(N^\beta)$.

Proposition 2.4.7. *Let* $0 < \alpha < \beta$, $\overline{\sigma}_{N^\alpha}^{(N)} := \inf\{n \in \mathbb{N}_0 : \overline{J}_n^{(N)} \geq N^\alpha\}$ *and consider* $\overline{\mathcal{Z}}_l^{(N)}$ *with* $\theta_N := \frac{2N^\alpha \log(N)}{N - N^\alpha}$. *Then*

$$\lim_{N\to\infty} \mathbb{P}\left(\overline{Z}_{n,l}^{(N)} \leq \overline{J}_n^{(N)} \ \forall n < \overline{\sigma}_{N^\alpha}^{(N)}\right) = 1.$$

To prepare the proof of Proposition 2.4.7 in the next lemma we estimate in the simpler host-parasite model the total number of infecting units $\overline{M}^{(N)}$ that can maximally be generated during the epidemic, and the total number of infecting units $M^{\alpha,(N)}$ that are generated until in total N^α hosts get infected.

Lemma 2.4.8. *Assume the conditions of Theorem 2.2.2 (ii) are fulfilled and* $0 < \alpha < 1$. *Then we have*

$$\lim_{N\to\infty} \mathbb{P}\left(\overline{M}^{(N)} \leq N\log(N)\right) = 1,$$
$$\lim_{N\to\infty} \mathbb{P}\left(M^{\alpha,(N)} \leq N^\alpha \log(N)\right) = 1.$$

Proof. Denote by $M_i^{(N)}$ the number of infecting units generated in host i if it gets infected in the simpler model, i.e. $\overline{M}^{(N)} = \sum_{i=1}^N M_i^{(N)}$ and $M^{\alpha,(N)} \sim \sum_{i=1}^{N^\alpha} M_i^{(N)}$. By construction $M_i^{(N)}$ is distributed according to the probability distribution defined in (2.4.1) and the random variables $(M_i^{(N)})_{1\leq i\leq N}$ are i.i.d. An application of Markov's inequality yields

$$\mathbb{P}\left(\sum_{i=1}^N M_i^{(N)} \geq N\log(N)\right) \leq \frac{\mathbb{E}[M_1^{(N)}]}{\log(N)} \to 0,$$
$$\mathbb{P}\left(\sum_{i=1}^{N^\alpha} M_i^{(N)} \geq N^\alpha\log(N)\right) \leq \frac{\mathbb{E}[M_1^{(N)}]}{\log(N)} \to 0,$$

because the expectations $\left(\mathbb{E}[M_1^{(N)}]\right)_{N\in\mathbb{N}}$ are uniformly bounded. Indeed, recall the definition of the probability $w_k^{(N)}$ in (2.3.3). We have

$$\mathbb{E}[M_i^{(N)}] = \sum_{j=0}^{v_N} j \sum_{k+\ell=j, k\leq\lfloor v_N/2\rfloor} w_k^{(N)} \cdot \binom{v_N - 2k}{\ell} \rho_N^\ell (1-\rho_N)^{v_N-(2k+\ell)}$$

$$\leq \sum_{j=1}^{v_N} j \sum_{k+\ell=j} \left(\frac{v_N^2}{2d_N}\right)^k \frac{1}{k!} \cdot \frac{(v_N\rho_N)^\ell}{l!}$$
$$= \sum_{j=1}^{v_N} j \frac{\left(\frac{v_N^2}{2d_N} + v_N\rho_N\right)^j}{j!}$$
$$\leq \exp\left(\frac{v_N^2}{2d_N} + v_N\rho_N\right) \cdot \left(\frac{v_N^2}{2d_N} + v_N\rho_N\right) < \infty, \tag{2.4.2}$$

because $\frac{v_N^2}{2d_N} \to \frac{c^2}{2}$ and $v_N\rho_N \to x$. □

Proof of Proposition 2.4.7. Using the same kind of calculations as in the proof of Proposition 2.3.5 we can show that for all $1 \leq j \leq a_N$, $\widetilde{p}_j^{(N)} \leq \mathbb{P}(\bigcup_{k+\ell=j} D_{k,\ell}^{(N)})$, see Equations (2.3.4) and (2.3.5). In other words whenever a host gets infected we can estimate the number of infecting units, generated on the corresponding vertex according to the simpler model, from below by the corresponding number of offspring in the Galton-Watson process $\mathcal{Z}_l^{(N)}$, since $\widetilde{p}_0^{(N)} = 1 - \sum_{i=1}^{a_N} \tilde{p}_j^{(N)}$.

However, in the host-parasite model "ghost" infections may occur, when a) an already empty vertex is attacked by an infecting unit over a free half-edge, b) a vertex is attacked by more than one infecting unit or c) two infecting units attack an edge from different ends (and hence both infecting units hit empty vertices).

We will show next that each infecting unit generated before generation $\overline{\sigma}_{N^\alpha}^{(N)}$ is involved in one of the events a) or b) (independently of the other infecting units) with probability at most θ_N. Furthermore, we will show that an event of type c) occurs before generation $\overline{\sigma}_{N^\alpha}^{(N)}$ only with negligible probability $o(1)$. Consequently, by removing infecting units with probability θ_N the number of offspring of infected hosts can whp be bounded from below by the number of offspring drawn according to the distribution with weights $(p_{k,l}^{(N)})_{k\in\mathbb{N}_0}$ from Definition 2.4.5 for any generation $n < \overline{\sigma}_{N^\alpha}^{(N)}$. This yields the claimed coupling of $(\overline{J}_n^{(N)})_{n\in\mathbb{N}_0}$ and $(\overline{Z}_{n,l}^{(N)})_{n\in\mathbb{N}_0}$ before generation $\overline{\sigma}_{N^\alpha}^{(N)}$.

We first control the probabilities of the events a) and b).

a) Before generation $\overline{\sigma}_{N^\alpha}^{(N)}$ the number of free half-edges linked to an empty vertex is bounded by $N^\alpha d_N$. Hence, the probability that an infecting unit on a half-edge gets connected to a half-edge of an empty vertex is bounded from above by $\frac{N^\alpha d_N}{Nd_N - N^\alpha v_N} \sim \frac{1}{N^{1-\alpha}}$, since the total number of free half-edges is at least $Nd_N - N^\alpha v_N$.

b) Before generation $\overline{\sigma}_{N^\alpha}^{(N)}$, the number of empty vertices in the graph is smaller than N^α. Consequently, the probability that two infecting units attack the same vertex can be estimated from above by $\frac{d_N}{Nd_N - N^\alpha d_N} \sim \frac{1}{N}$. By Lemma 2.4.8 the total number of infecting units generated before generation $\overline{\sigma}_{N^\alpha}^{(N)}$ is whp bounded by $N^\alpha \log(N)$. Hence, each infecting unit is involved in an event of type *b*) with probability at most $N^\alpha \log(N) \cdot \frac{d_N}{Nd_N - N^\alpha d_N} \sim \frac{\log(N)}{N^{1-\alpha}}$.

In summary, $\theta_N = 2 \cdot N^\alpha \log(N) \cdot \frac{d_N}{Nd_N - N^\alpha d_N}$ yields an upper bound on the probability that an infecting unit is involved in one of the events of type a) or b). Since $\alpha < 1$ we have $\theta_N \in o(1)$.

It remains to show that whp events of type c) do not occur until generation $\overline{\sigma}_{N^\alpha}^{(N)}$. According to Lemma 2.4.8 whp the number of infecting units that can be generated during the epidemic is at most $N\log(N)$ and before generation $\overline{\sigma}_{N^\alpha}^{(N)}$ the total number of generated infecting units can be

estimated from above by $N^\alpha \log(N)$. Hence, whp we can estimate the probability that before time $\overline{\sigma}^{(N)}_{N^\alpha}$ none of the infecting units moves along an edge, on which end another infecting unit is located on, by

$$\begin{aligned}
&\frac{Nd_N - N\log(N)}{Nd_N - 1} \cdots \frac{Nd_N - N\log(N) - (N^\alpha \log(N) - 1)}{Nd_N - 1 - 2(N^\alpha \log(N) - 1)} \\
=&\frac{1}{\prod_{i=0}^{N^\alpha \log(N)-1}(Nd_N - 1 - 2i)} \cdot \frac{(Nd_N - N\log(N))!}{(Nd_N - N\log(N) - N^\alpha \log(N))!} \\
\geq& \left(\frac{(Nd_N - N\log(N) - N^\alpha \log(N))}{Nd_N}\right)^{N^\alpha \log(N)} \\
=& \left(1 - \frac{(N - N^\alpha)\log(N)}{Nd_N}\right)^{N^\alpha \log(N)} \\
=&1 - o(1),
\end{aligned}$$

where the last equality holds because $\alpha < \beta$. □

We conclude this section with the proof of Proposition 2.4.1.

Proof of Proposition 2.4.1. By Proposition 2.4.4 we can show the claim of the proposition for the event $\{\exists n \in \mathbb{N}_0 : \overline{J}^{(N)}_n \geq N^\alpha\}$ instead of the event $\{\exists n \in \mathbb{N}_0 : \overline{I}^{(N)}_n \geq N^\alpha\}$. According to Proposition 2.4.7 the process $\overline{\mathcal{J}}^{(N)}$ can whp be coupled from below by $\overline{\mathcal{Z}}^{(N)}_l$. By Lemma 2.4.6, the process $\overline{\mathcal{Z}}^{(N)}_l$ reaches at least the level N^α with asymptotic probability $\pi(c, x)$, which concludes the proof. □

2.4.2 Growing further at exponential speed

In Section 2.4.1 we showed that N^α hosts will get infected with asymptotic probability $\pi(c, x)$ for any $0 < \alpha < \beta$. In Section 2.5 we will see that the total host population will go extinct whp in at most 2 generations if at least $N^{1-\frac{3}{4}\beta+2\varepsilon}$ hosts get infected for any $\varepsilon > 0$. If $\beta > \frac{4}{7}$ we have $1 - \frac{3}{4}\beta < \beta$ and hence, with the results of the next section we can prove Theorem 2.2.2 (ii). The aim of this section is to argue that also in the case $\beta \leq \frac{4}{7}$ whp $N^{1-\frac{3}{4}\beta+2\varepsilon}$ hosts will get infected once N^α hosts have been removed for some $0 < \alpha < \beta$. Hence, we assume in the remainder of this subsection that

$$\beta \leq \frac{4}{7}.$$

We will truncate the process $\mathcal{Z}^{(N)}_l$ at certain time points. The resulting process $\mathcal{Z}^{(N)}_t = (Z^{(N)}_{n,t})_{n\in\mathbb{N}_0}$ grows asymptotically at the same speed as $\mathcal{Z}^{(N)}_l$ and can be coupled with $\overline{I}^{(N)}$ until the level $N^{1-\frac{3}{4}\beta+2\varepsilon}$ is reached. The coupling of $\mathcal{Z}^{(N)}_l$ with $\overline{I}^{(N)}$ fails if two infecting units attack an edge from two different ends at the same generation. In this case none of the two infecting units can reproduce because the vertices they are moving to are empty. Since in each generation, the number of infecting units involved in these events is small we can remove from time to time (the ancestors of) these infecting units without changing the asymptotic speed of exponential growth. Define $k_0 \in \mathbb{N}$ through $k_0 - 1 := \sup\{k \in \mathbb{N} : k\beta \leq 1 - \frac{3}{4}\beta\}$, in particular we have

$$(k_0-1)\beta \le 1-\frac{3}{4}\beta < k_0\beta.$$

Definition 2.4.9. *Let $\delta<\beta$ and $\varepsilon>0$ small enough such that $k_0(\beta-\delta)\ge 1-\frac{3}{4}\beta+2\varepsilon$. Define $\mathcal{Z}_t^{(N)}=(Z_{n,t}^{(N)})_{n\in\mathbb{N}_0}$, with $Z_{0,t}^{(N)}:=1$ almost surely, and let $Z_t^{(N)}$ evolve as a GWP with offspring distribution $(p_{k,l}^{(N)})_{k\in\mathbb{N}_0}$ until time $\overline{\sigma}_1^{(N)}:=\inf\{n\in\mathbb{N}_0 : \overline{Z}_{n,t}^{(N)}\ge N^{\beta-\delta}\}$. We set $Z^{(N)}_{\overline{\sigma}_1^{(N)}+1,t}:=\max\{Z^{(N)}_{\overline{\sigma}_1^{(N)},t}-N^{\beta-\frac{3}{2}\delta},0\}$. Assume that the process $Z_t^{(N)}$ is defined until generation $\overline{\sigma}_i^{(N)}+1$ for some $i\le k_0-1$, then let the process grow as a GWP with offspring distribution $(p_{k,l}^{(N)})_{k\in\mathbb{N}_0}$ until generation $\overline{\sigma}_{i+1}^{(N)}:=\inf\{n\in\mathbb{N}_0 : \overline{Z}_{n,t}^{(N)}\ge N^{(i+1)(\beta-\delta)}\}$. Set $Z^{(N)}_{\overline{\sigma}_{i+1}^{(N)}+1,t}:=\max\{Z^{(N)}_{\overline{\sigma}_{i+1}^{(N)},t}-N^{i\beta-\frac{2i+1}{2}\delta},0\}$.*

Proposition 2.4.10. *Let $\overline{\tau}^{(N)}:=\inf\{n\in\mathbb{N}_0 : \overline{I}_n^{(N)}\ge N^{1-\frac{3}{4}\beta+2\varepsilon}\}$. Then*

$$\lim_{N\to\infty}\mathbb{P}\left(\overline{Z}_{n,t}^{(N)}\le\overline{I}_n^{(N)}\ \forall n\le\overline{\tau}^{(N)}\right)=1.$$

Proof. The coupling of $\overline{Z}_l^{(N)}$ and $\overline{I}^{(N)}$ fails if two infecting units attack an edge from both ends, because in this situation the corresponding branches in the Galton-Watson process have offspring but the corresponding infecting units do not infect any host. These infecting units cannot be treated independently and hence we cannot arrive at a coupling by thinning the Galton-Watson process. Instead we will remove the corresponding lines in the Galton-Watson process in pairs.

If at some generation the number of infected hosts is $\mathcal{O}(N^\alpha)$, then in this generation whp $\mathcal{O}(N^\alpha\log(N))$ infecting units are generated, see Lemma 2.4.8. Because whp the total number of infecting units is at most $N\log(N)$, see Lemma 2.4.8 again, an application of Chebyshev's Inequality yields that whp no more than $\mathcal{O}\left(\frac{N^\alpha\log(N)N\log(N)}{d_N N}\right)=\mathcal{O}\left(N^{\alpha-\beta}\log(N)^2\right)$ pairs of infecting units attack an edge from both ends. Within the time intervals $([\overline{\sigma}_i^{(N)}+1,\overline{\sigma}_{i+1}^{(N)}])_i$ in each generation each individual has on average at least $c^2/2+x+o(1)$ offspring. Since within any time interval $[\overline{\sigma}_i^{(N)}+1,\overline{\sigma}_{i+1}^{(N)}]$ the process grows exponentially fast, for $1\le i\le k_0$, whp at most $\mathcal{O}(N^{(i+1)(\beta-\delta)-\beta}\log^3(N))=o(N^{i\beta-\frac{2i+1}{2}\delta})$ pairs of infecting units are placed on two different ends of an edge. If we remove this number of pairs of infecting units at time $\overline{\sigma}_i^{(N)}+1$ and then let the process evolve like a GWP with offspring distribution $(p_{k,l}^{(N)})_{k\in\mathbb{N}_0}$, the total size of the resulting process whp lower bounds $\overline{I}^{(N)}$ until generation $\overline{\sigma}_{i+1}^{(N)}$. Continuing this algorithm till generation $\overline{\sigma}_{k_0}^{(N)}$, we arrive at the desired result. □

Lemma 2.4.11. *Assume the process $\mathcal{Z}_t^{(N)}$ is constructed by means of the probability weights $(p_{k,l}^{(N)})_{k\in\mathbb{N}_0}$ with $\theta_N=\frac{2N^\alpha\log(N)}{N-N^\alpha}$ for some $k_0\beta<\alpha<1$. Assume ε is small enough such that $1-\frac{3}{4}\beta+2\varepsilon<k_0\beta$. Then*

$$\lim_{N\to\infty}\mathbb{P}\left(\exists n\in\mathbb{N}_0 : \overline{Z}_{n,t}^{(N)}\ge N^{1-\frac{3}{4}\beta+2\varepsilon}\right)=\pi(c,x).$$

Proof. Since $\mathcal{Z}_t^{(N)}$ and $\mathcal{Z}_l^{(N)}$ coincide until the level N^γ is reached for any $\gamma\le\beta-\delta$ an application of Lemma 2.4.6 yields that the level $N^{\beta-\delta}$ is reached with asymptotic probability $\pi(c,x)$. If the

level $N^{\beta-\delta}$ has been reached, the level $N^{1-\frac{3}{4}\beta+2\varepsilon} \gg N^{\beta-\delta}$ will be reached whp. Indeed once a level ℓ_N has been reached by a supercritical GWP for some sequence $\ell_N \to \infty$, the GWP will explode whp. Since $Z^{(N)}_{\overline{\sigma}_1^{(N)}+1,t} \sim Z^{(N)}_{\overline{\sigma}_1^{(N)},t} = Z^{(N)}_{\overline{\sigma}_1^{(N)},\ell}$ and between generations $\overline{\sigma}_1^{(N)}+1$ and $\overline{\sigma}_2^{(N)}$, $\mathcal{Z}_t^{(N)}$ evolves as a supercritical GWP, we have $\overline{\sigma}_2^{(N)} < \infty$ whp. Repeating this argument $k_0 - 1$ times, we reach the level $N^{1-\frac{3}{4}\beta+2\varepsilon}$ whp. □

From Proposition 2.4.10 and Lemma 2.4.11 it follows that $\overline{\mathcal{I}}^{(N)}$ reaches the level $N^{1-\frac{3}{4}\beta+2\varepsilon}$ asymptotically with probability $\pi(c,x)$. Hence, for the proof of Theorem 2.2.2 (ii) it remains to show that $\overline{\mathcal{I}}^{(N)}$ reaches the level N after hitting the level $N^{1-\frac{3}{4}\beta+2\varepsilon}$ whp. This is the topic of Section 2.5.

2.5 Final phase of the epidemic

In this section we consider again the setting of Theorem 2.2.2 (ii). We aim to show that once N^α hosts got infected eventually whp also the remaining hosts get infected. Assume in the following that $\varepsilon > 0$ is small enough such that $1 - \frac{3\beta}{4} + 2\varepsilon < 1 - \frac{\beta}{2}$. Recall

$$\overline{\tau}^{(N)} = \inf\{n \in \mathbb{N}_0 : N^{1-\frac{3}{4}\beta+2\varepsilon} \le \overline{I}_n^{(N)}\},$$

and define

$$\tau^{(N)} := \inf\{n \in \mathbb{N}_0 : N^{1-\frac{3}{4}\beta+\varepsilon} \le I_n^{(N)}\}.$$

Proposition 2.5.1. *For ε defined as at the beginning of this section we have*

$$\lim_{N\to\infty} \mathbb{P}\left(\overline{I}^{(N)}_{\tau^{(N)}+2} = N \middle| \overline{\tau}^{(N)} < \infty\right) = 1.$$

The key observation for the proof of Proposition 5.1 is that infection by cooperation of parasites that attack a host from different edges determine the infection dynamics when $I_n^{(N)} \gg N^{1-\beta}$. Our assumptions on ε guarantee that $\frac{(N^{1-\frac{3\beta}{4}+2\epsilon}v_N)^2}{N} \ll N$. In Lemma 2.5.2 we will show that $\tau^{(N)} \le \overline{\tau}^{(N)}$ whp conditioned on $\overline{\tau}^{(N)} < \infty$. Hence, we have $N^{1-\frac{3}{4}\beta+\varepsilon} \le \overline{I}^{(N)}_{\tau^{(N)}} \ll N$ and one generation further we have $N^{1-\beta/2+\varepsilon} \ll \overline{I}^{(N)}_{\tau^{(N)}+1} \le N$ and also $N^{1-\beta/2+\varepsilon} \ll I^{(N)}_{\tau^{(N)}+1}$. Consequently, in the following generation either the remaining hosts get infected, since $\frac{N^{1-\beta/2+\varepsilon}v_N}{N} \gg N$ or (when already all hosts got infected) the number of removed hosts is N.

In the following we first state and prepare for the proof of Lemma 2.5.2, then we give the proof of this lemma and finish the section with the proof of Proposition 2.5.1.

Lemma 2.5.2. *For ε, $\tau^{(N)}$ as well as $\overline{\tau}^{(N)}$ defined as at the beginning of the section*

$$\lim_{N\to\infty} \mathbb{P}\left(\tau^{(N)} \le \overline{\tau}^{(N)}\right) = 1.$$

To prove Lemma 2.5.2 we control the time the approximating processes $(\mathcal{Z}_t^{(N)})_{N\in\mathbb{N}}$ need to reach some level N^α. We start with a rather classical result on branching processes. We give its proof in the Appendix for the sake of completeness.

Lemma 2.5.3. *Let $(Z_n)_{n\in\mathbb{N}_0}$ be a Galton-Watson process with $m := \mathbb{E}[Z_1] > 1$. Introduce $\tau_{N^\alpha} := \inf\{n \in \mathbb{N}_0 : Z_n \geq N^\alpha\}$. Assume $Z_0 = N^\gamma - \varphi(N)$ such that $Z_0 \geq 1$, where $0 \leq \gamma < \alpha$ and $\varphi(N) \in o(N^\gamma)$. Denote by W the almost sure limit of the non-negative martingale $(\frac{Z_n}{Z_0}m^{-n})_{n\in\mathbb{N}_0}$. Conditioning on $\{W > 0\}$*

$$\frac{\tau_{N^\alpha}\log m}{(\alpha-\gamma)\log N} \to 1, \text{ almost surely.}$$

Next we consider a family of Galton-Watson processes $((Z_n^{(\varepsilon)})_{n\in\mathbb{N}_0})_{\varepsilon>0}$, for which mean offspring numbers m_ε are converging to some limit $m > 1$ when $\varepsilon \downarrow 0$. In this case the time to reach the level N^α from a level N^γ is, conditioned on non-extinction, also not larger than $(1+\delta)\frac{(\alpha-\gamma)\log N}{\log m}$ for ε small enough and $\delta > 0$.

Lemma 2.5.4. *Let $\mathcal{Z}^{(\varepsilon)} = (Z_n^{(\varepsilon)})_{n\in\mathbb{N}_0}$ be a Galton-Watson Process. Denote by $m_\varepsilon := \mathbb{E}[Z_1^{(\varepsilon)}] = m - f(\varepsilon)$ the mean number of offspring, where $f(\varepsilon) \underset{\varepsilon\to 0}{\longrightarrow} 0$, and $m > 1$. Introduce $W^{(\varepsilon)}$ the almost sure limit of the non-negative martingale $(Z_n^{(\varepsilon)}m_\varepsilon^{-1})_{n\in\mathbb{N}_0}$, and $\tau_{N^\alpha}^{(\varepsilon)} := \inf\{n \in \mathbb{N}_0 : Z_n^{(\varepsilon)} \geq N^\alpha\}$, the first time at which $\mathcal{Z}^{(\varepsilon)}$ reaches the size N^α.*
If $Z_0^{(\varepsilon)} = N^\gamma - \varphi(N)$ such that $Z_0^{(\varepsilon)} \geq 1$, where $0 \leq \gamma < \alpha$ and $\varphi(N) \in o(N^\gamma)$, then for all $\delta > 0$ and for all $\varepsilon > 0$ small enough

$$\lim_{N\to\infty} \mathbb{P}\left(\tau_{N^\alpha}^{(\varepsilon)} \leq (1+\delta)\frac{(\alpha-\gamma)\log(N)}{\log(m)}\Big| W^{(\varepsilon)} > 0\right) = 1.$$

Proof. Lemma 2.5.3 gives that for all $\delta > 0$

$$\lim_{N\to\infty} \mathbb{P}\left(\tau_{N^\alpha}^{(\varepsilon)} \leq (1+\delta)\frac{(\alpha-\gamma)\log(N)}{\log(m_\varepsilon)}\Big| W^{(\varepsilon)} > 0\right) = 1.$$

And using that $m_\varepsilon \to m$ when $\varepsilon \to 0$, it directly follows the result of this Lemma. □

Finally we consider a sequence of GWPes $((Z_n^{(N)})_{n\in\mathbb{N}_0})_{N\in\mathbb{N}}$, whose offspring distributions depend on N, and the level that we are interested to reach depends on N as well.

Lemma 2.5.5. *Let $\left(\left(Z_n^{(N)}\right)_{n\in\mathbb{N}_0}\right)_{N\in\mathbb{N}}$ be a sequence of GWPes whose offspring distributions are denoted by $(p_k^{(N)})_{k\in\mathbb{N}_0}$. Denote by $\Phi^{(N)}$ the corresponding sequences of generating functions of the offspring distributions. Introduce $\tau_{N^\alpha}^{(N)} := \inf\{n \in \mathbb{N}_0 : Z_n^{(N)} \geq N^\alpha\}$ and $\overline{\tau}_{N^\alpha}^{(N)} := \inf\{n \in \mathbb{N}_0 : \sum_{i=0}^n Z_i^{(N)} \geq N^\alpha\}$. Let $(p_k)_{k\in\mathbb{N}_0}$ be a probability distribution and Φ its generating function, satisfying $1 < m := \Phi'(1) < \infty$. Assume that*

$$\sum_{k=0}^{\infty} |p_k^{(N)} - p_k| \to 0. \tag{2.5.1}$$

If $Z_0^{(N)} = N^\gamma - \varphi(N)$ such that $Z_0^{(N)} \geq 1$, where $0 \leq \gamma < \alpha$ and $\varphi(N) \in o(N^\gamma)$, then for all $\delta > 0$

$$\lim_{N\to\infty} \mathbb{P}\left(\tau_{N^\alpha}^{(N)} \leq (1+\delta)\frac{(\alpha-\gamma)\log(N)}{\log m}\Big| \overline{\tau}_{N^\alpha}^{(N)} < \infty\right) = 1.$$

Proof. Using Assumption (2.5.1) it follows from the remark just before Lemma 2.3.7 that the sequence $(\Phi^{(N)})_{N\in\mathbb{N}}$ converges uniformly to the generating function Φ.
Consider a family of natural numbers $(K_\varepsilon)_{\varepsilon>0}$ satisfying $K_\varepsilon \underset{\varepsilon\to 0}{\longrightarrow} \infty$ and $K_\varepsilon^2\varepsilon^\gamma \underset{\varepsilon\to 0}{\longrightarrow} 0$, where $0 < \gamma < 1$. We introduce the GWP $\left(Z_n^{(\varepsilon)}\right)_{n\in\mathbb{N}_0}$, whose offspring distribution $\left(p_k^{(\varepsilon)}\right)_{k\in\mathbb{N}_0}$ is defined as follows. For all $1 \le k \le K_\varepsilon$

$$p_k^{(\varepsilon)} := \max\{p_k - \varepsilon^\gamma, 0\},$$

and

$$p_0^{(\varepsilon)} := 1 - \sum_{k=1}^{K_\varepsilon} p_k^{(\varepsilon)}.$$

This definition implies that the generating functions $\Phi^{(\varepsilon)}$ converge uniformly in $[0,1]$ to Φ, as well as the mean number of offspring $m_\varepsilon := \mathbb{E}[Z_1^{(\varepsilon)}]$ converges to m, when $\varepsilon \to 0$. Indeed, we have for all $0 \le s \le 1$

$$\begin{aligned}|\Phi(s) - \Phi^{(\varepsilon)}(s)| &\le \sum_{k=1}^{K_\varepsilon} s^k\varepsilon^\gamma + \sum_{k=K_\varepsilon+1}^{\infty} s^k p_k + \left(p_0^{(\varepsilon)} - p_0\right)\\ &\le 2K_\varepsilon\varepsilon^\gamma + 2\sum_{k=K_\varepsilon+1}^{\infty} p_k\\ &\underset{\varepsilon\to 0}{\longrightarrow} 0,\end{aligned}$$

since $K_\varepsilon\varepsilon^\gamma \underset{\varepsilon\to 0}{\longrightarrow} 0$ and $K_\varepsilon \underset{\varepsilon\to 0}{\longrightarrow} \infty$. And also

$$|m - m_\varepsilon| \le \sum_{k=1}^{K_\varepsilon} k\varepsilon^\gamma + \sum_{k=K_\varepsilon+1}^{\infty} kp_k \le K_\varepsilon^2\varepsilon^\gamma + \sum_{k=K_\varepsilon+1}^{\infty} kp_k \underset{\varepsilon\to 0}{\longrightarrow} 0,$$

because $K_\varepsilon^2\varepsilon^\gamma \underset{\varepsilon\to 0}{\longrightarrow} 0$ and $m < \infty$.
Moreover, Assumption (2.5.1) implies that $\sup_{k\in\mathbb{N}_0}|p_k^{(N)} - p_k| \to 0$, so there exists N_ε such that $N \ge N_\varepsilon$ and for all $k \in \mathbb{N}_0$

$$p_k^{(N)} \ge \max\left\{p_k - \frac{\varepsilon^\gamma}{2}, 0\right\}.$$

Consequently, for all $N \ge N_\varepsilon$ we have $p_k^{(\varepsilon)} \le p_k^{(N)}$ for all $k \ge 1$ and $p_0^{(\varepsilon)} \ge p_0^{(N)}$. Hence, we can couple $\left(Z_n^{(\varepsilon)}\right)_{n\in\mathbb{N}_0}$ and $\left(Z_n^{(N)}\right)_{n\in\mathbb{N}_0}$ such that for all $n \in \mathbb{N}$

$$Z_n^{(\varepsilon)} \le Z_n^{(N)},$$

and

$$Z_0^{(N)} = Z_0^{(\varepsilon)}.$$

Lemma 2.5.4 and the convergence $m_\varepsilon \underset{\varepsilon\to 0}{\longrightarrow} m$ gives that for all $\delta > 0$ and for all $\varepsilon > 0$ small enough

$$\lim_{N\to\infty} \mathbb{P}\left(\tau_{N^\alpha}^{(\varepsilon)} \leq (1+\delta)\frac{(\alpha-\gamma)\log(N)}{\log(m)} \Big| W^{(\varepsilon)} > 0\right) = 1,$$

where $\tau_{N^\alpha}^{(\varepsilon)} := \inf\{n \in \mathbb{N}_0 : Z_n^{(\varepsilon)} \geq N^\alpha\}$ and $W^{(\varepsilon)}$ is the almost sure limit of the non-negative martingale $\left(Z_n^{(\varepsilon)}/(Z_0^{(\varepsilon)} m_\varepsilon^n)\right)_{n\in\mathbb{N}_0}$. The coupling yields

$$\begin{aligned} 1 &\geq \mathbb{P}\left(\tau_{N^\alpha}^{(N)} \leq (1+\delta)\frac{(\alpha-\gamma)\log(N)}{\log(m)} \Big| W^{(\varepsilon)} > 0\right) \\ &\geq \mathbb{P}\left(\tau_{N^\alpha}^{(\varepsilon)} \leq (1+\delta)\frac{(\alpha-\gamma)\log(N)}{\log(m)} \Big| W^{(\varepsilon)} > 0\right) \\ &\to 1, \end{aligned}$$

which yields that for all $\delta > 0$ and for all $\varepsilon > 0$ small enough

$$\lim_{N\to\infty} \mathbb{P}\left(\tau_{N^\alpha}^{(N)} \leq (1+\delta)\frac{(\alpha-\gamma)\log(N)}{\log(m)} \Big| W^{(\varepsilon)} > 0\right) = 1.$$

Denote by $E_N := \{\tau_{N^\alpha}^{(N)} \leq (1+\delta)\frac{(\alpha-\gamma)\log(N)}{\log(m)}\}$, by $F_\varepsilon := \{W^{(\varepsilon)} > 0\}$, and by $G_N := \{\overline{\tau}_{N^\alpha}^{(N)} < \infty\}$. The coupling implies that $F_\varepsilon \subset G_N$. Lemma 2.3.7 d) and the uniform convergence of the generating functions $\Phi^{(N)}$ to Φ give that $\lim_{N\to\infty} \mathbb{P}(G_N) = \pi$, where π is the survival probability of the GWP with generating function Φ. Lemma 2.3.7 a) and the uniform convergence of the generating functions $\Phi^{(\varepsilon)}$ give that $\lim_{\varepsilon\to 0} \mathbb{P}(F_\varepsilon) = \pi$.
We have

$$\begin{aligned} \mathbb{P}(E_N|F_\varepsilon) &= \frac{\mathbb{P}(E_N \cap F_\varepsilon)}{\mathbb{P}(F_\varepsilon)} + \frac{\mathbb{P}(E_N \cap (G_N \backslash F_\varepsilon))}{\mathbb{P}(F_\varepsilon)} - \frac{\mathbb{P}(E_N \cap (G_N \backslash F_\varepsilon))}{\mathbb{P}(F_\varepsilon)} \\ &\leq \mathbb{P}(E_N|G_N) \cdot \frac{\mathbb{P}(G_N)}{\mathbb{P}(F_\varepsilon)}, \end{aligned}$$

and taking the $\liminf_{N\to\infty}$ gives that

$$1 \leq \liminf_{N\to\infty} (\mathbb{P}(E_N|G_N)) \cdot \frac{\pi}{\mathbb{P}(F_\varepsilon)},$$

and finally by taking the limit when $\varepsilon \to 0$, we get

$$\liminf_{N\to\infty} \mathbb{P}(E_N|G_N) \geq 1,$$

and since it is a sequence of probability terms, it follows that

$$\lim_{N\to\infty} \mathbb{P}(E_N|G_N) = 1,$$

which is the result of this lemma. □

We apply the last lemma iteratively to the sequence of processes $(\mathcal{Z}_t^{(N)})_{N\in\mathbb{N}}$ introduced in Definition 2.4.9.

Lemma 2.5.6. *Assume the process $\mathcal{Z}_t^{(N)}$ is constructed by means of the offspring probability weights $\left(p_{k,l}^{(N)}\right)_{k\in\mathbb{N}_0}$ with $\theta_N = \frac{2N^\alpha \log(N)}{N-N^\alpha}$ for some $k_0\beta < \alpha < 1$. Then $\forall \delta > 0$*

$$\lim_{N\to\infty} \mathbb{P}\left(\overline{\tau}_{N^\alpha,t}^{(N)} \le (1+\delta)\frac{\alpha\log(N)}{\log\left(\frac{c^2}{2}+x\right)} \Big| \overline{\tau}_{N^\alpha,t}^{(N)} < \infty\right) = 1,$$

where $\overline{\tau}_{N^\alpha,t}^{(N)} := \inf\{n \in \mathbb{N}_0 : \overline{Z}_{n,t}^{(N)} \ge N^\alpha\}$.

Proof. Since $\mathcal{Z}_t^{(N)}$ is, except at the time points $\overline{\sigma}_i^{(N)}$, a GWP, we can apply iteratively Lemma 2.5.5 where Assumption (2.5.1) is obtained in the proof of Lemma 2.4.6. □

Finally we come to the proof of Lemma 2.5.2.

Proof of Lemma 2.5.2. If for every generation n before $\overline{\tau}^{(N)}$, the number of infected hosts at generation n satisfies $I_n^{(N)} \le N^{1-\frac{3\beta}{4}+\varepsilon}$, then $\overline{\tau}^{(N)} \ge N^\varepsilon$.
But the coupling from below works whp at least until generation $\overline{\tau}^{(N)}$, and thanks to Lemma 2.5.6, we know that the total size of the process $\mathcal{Z}_t^{(N)}$ will reach $N^{1-\frac{3\beta}{4}+2\varepsilon}$ within a time of order $\log(N)$. This implies that there exists $n \le \overline{\tau}^{(N)}$ for which $I_n^{(N)} \ge N^{1-\frac{3\beta}{4}+\varepsilon}$. □

Lemma 2.5.7. *For ε defined as at the beginning of this section*

$$\lim_{N\to\infty} \mathbb{P}\left(\overline{I}_{\overline{\tau}^{(N)}}^{(N)} \le N^{1-\frac{\beta}{2}+5\varepsilon} \Big| \overline{\tau}^{(N)} < \infty\right) = 1.$$

Proof. The number of newly infected vertices is the sum of vertices that get attacked by successful single parasites or by several parasites simultaneously. The number of vertices that get infected by single successful parasites or pairs of parasites that move along the same edge denoted by $A^{(N)}$ is whp bounded from above by $N^{1-\frac{3}{4}\beta+3\varepsilon}$. We will show that the number of vertices that get infected by parasites attacking the vertex from different edges is whp bounded above by $N^{1-\frac{\beta}{2}+5\varepsilon}$.

At generation $\overline{\tau}^{(N)}-1$ less than $N^{1-\frac{3\beta}{4}+2\varepsilon}$ vertices are infected, and so there are less than $b_N := v_N N^{1-\frac{3\beta}{4}+2\varepsilon}$ available parasites. Also the number of susceptible hosts is bigger than $N - N^{1-\frac{3\beta}{4}+2\varepsilon}$, and as we will show below whp they all have more than $d_N - \varphi(N)$ free half-edges for some sequence $(\varphi(N))_{N\in\mathbb{N}}$ where $\varphi(N) = O(1)$, see (2.5.3).
Denote by $D_i^{(N)}$ the number of free half-edges of vertex i at generation $\overline{\tau}^{(N)}-1$. Assume we have S_N boxes with box $i \le S_N$ containing $D_i^{(N)}$ positions, and assume b_N balls are distributed uniformly on the positions of the boxes, such that each position gets occupied at most once, and let $G_i^{(N)}$ be the number of balls put into box i. Then we have whp

$$\overline{I}_{\overline{\tau}^{(N)}}^{(N)} \le A^{(N)} + \sum_{i\in S_N} \mathbb{1}_{\{G_i^{(N)}\ge 2\}},$$

because $A^{(N)} + \sum_{i\in S_N} \mathbb{1}_{\{G_i^{(N)}\ge 2\}} \ge N^{1-\frac{3}{4}\beta+2\varepsilon}$ whp. Denote by $G_N := \sum_{i\in S_N} \mathbb{1}_{\{G_i^{(N)}\ge 2\}}$. We will show that

$$\lim_{N\to\infty} \mathbb{P}\left(G_N \leq N^{1-\frac{\beta}{2}+5\varepsilon}\right) = 1.$$

Denote by $T := \sum_{j\in S_N} D_j^{(N)}$, $T_i := T - D_i^{(N)}$ and $T_{i,j} := T - \left(D_i^{(N)} + D_j^{(N)}\right)$. To estimate the expectation and variance of G_N we estimate the probabilities of the events $\{G_i^{(N)} \leq 1\}$ and $\{G_i^{(N)} \leq 1\} \cap \{G_j^{(N)} \leq 1\}$ for $i \neq j$ conditioned on S_N. Since $\mathbb{P}\left(\{N - N^{1-\frac{3\beta}{4}+2\varepsilon} \leq S_N \leq N\}\right) = 1$ and $\mathbb{P}\left(\bigcap_{i=1}^{S_N} \{d_N - \varphi(N) \leq D^{(N)} \leq d_N\} \,\middle|\, S_N\right) \to 1$, Lemma 2.6.1 can be applied whp. Hence, we have whp

$$\mathbb{P}\left(\{G_i^{(N)} \leq 1\}|S_N\right) = 1 - \frac{1}{2}\frac{b_N^2}{S_N^2} + \frac{1}{3}\frac{b_N^3}{S_N^3} - \frac{1}{8}\frac{b_N^4}{S_N^4} + \mathcal{O}\left(\frac{b_N^5}{S_N^5}\right), \tag{2.5.2}$$

and for all $i \neq j$

$$\mathbb{P}\left(\{G_i^{(N)} \leq 1\} \cap \{G_j^{(N)} \leq 1\}|S_N\right) = 1 - \frac{b_N^2}{S_N^2} + \frac{2}{3}\frac{b_N^3}{S_N^3} + \mathcal{O}\left(\frac{b_N^5}{S_N^5}\right).$$

Using (2.5.2) we get

$$\mathbb{E}[G^{(N)}|S_N] = S_N\left(1 - \mathbb{P}\left(\{G_i^{(N)} \leq 1\}|S_N\right)\right) = \frac{b_N^2}{2S_N} + o\left(\frac{b_N^2}{S_N}\right),$$

and because $\mathbb{P}\left(\{N - N^{1-\frac{3\beta}{4}+2\varepsilon} \leq S_N \leq N\}\right) = 1$, it follows that $E[G^{(N)}] = \mathcal{O}\left(N^{1-\frac{\beta}{2}+4\varepsilon}\right)$. The variance of $G^{(N)}$ conditioned on S_N is estimated in Lemma 2.6.1 as

$$\mathbb{V}[G^{(N)}|S_N] = \mathcal{O}\left(\frac{b_N^4}{S_N^2} \cdot \frac{b_N}{S_N}\right),$$

as long as $S_N \sim N$. The law of total variance yields

$$\begin{aligned}\mathbb{V}[G^{(N)}] &= \mathbb{E}[\mathbb{V}[G^{(N)}|S_N]] + \mathbb{V}[\mathbb{E}[G^{(N)}|S_N]] \\ &= \mathcal{O}\left(\frac{b_N^5}{S_N^3}\right) + \mathbb{E}[\mathbb{E}[G^{(N)}|S_N]^2] - \mathbb{E}[G^{(N)}].\end{aligned}$$

The term $\mathbb{E}[\mathbb{E}[G^{(N)}|S_N]^2] = \sum_{i=N-N^{1-\frac{3\beta}{4}+2\varepsilon}}^{N} \mathbb{P}(S_N = i)\left(\frac{b_N^2}{2i} + o\left(\frac{b_N^2}{i}\right)\right)^2 \sim \frac{b_N^4}{N^2}$. This means that $\mathbb{V}[\mathbb{E}[G^{(N)}|S_N]]$ can not exceed $\mathcal{O}\left(\frac{b_N^4}{N^2}\right)$, so an application of Chebyshev's Inequality yields the statement of the lemma.

It remains to show that the number of free half-edges of each susceptible vertex is sufficiently close to d_N. Denote by $H_i^{(N)}$ the number of half-edges that are already formed for vertex i in generation $\overline{\tau}^{(N)} - 1$, for $i \in \{1, \cdots, N\}$. We show that

$$\lim_{N\to\infty} \mathbb{P}\left(H_i^{(N)} \leq \varphi(N), \forall i \in S_{\overline{\tau}^{(N)}-1}^{(N)}\right) = 1, \tag{2.5.3}$$

for any $\varphi(N)$ such that $\liminf_N \varphi(N) \geq 5$. Indeed, consider the following experiment: Assume we have $N - N^{1-\frac{3\beta}{4}+2\varepsilon}$ boxes, each with d_N positions, and we distribute uniformly at random $v_N N^{1-\frac{3\beta}{4}+2\varepsilon}$ balls on the positions, such that each position gets occupied by at most one ball. Denote again by $(G_i^{(N)})_i$ the number of balls in box i. Then we have

$$\mathbb{P}\left(H_i^{(N)} \leq \varphi(N), \forall i \in S^{(N)}_{\overline{\tau}^{(N)}-1}\right) \geq \mathbb{P}\left(G_i^{(N)} \leq \varphi(N), \forall i \leq N - N^{1-\frac{3\beta}{4}+2\varepsilon}\right),$$

and assuming w.l.o.g. $\varphi(N) = 5$, we have

$$\begin{aligned}
\mathbb{P}\left(\exists i : G_i^{(N)} \geq \varphi(N)\right) &= \frac{(N - N^{1-\frac{3\beta}{4}+2\varepsilon})\binom{v_N N^{1-\frac{3\beta}{4}+2\varepsilon}}{\varphi(N)} d_N!(d_N(N - N^{1-\frac{3\beta}{4}+2\varepsilon}) - \varphi(N))!}{(d_N - \varphi(N))!(d_N(N - N^{1-\frac{3\beta}{4}+2\varepsilon}))!} \\
&\leq N \cdot \frac{(v_N N^{1-\frac{3\beta}{4}+2\varepsilon})^{\varphi(N)}}{\varphi(N)!} \cdot \frac{d_N^{\varphi(N)}}{(d_N(N - N^{1-\frac{3\beta}{4}+2\varepsilon}) - \varphi(N))^{\varphi(N)}} \\
&\leq N \exp(\varphi(N)) \left(\frac{v_N N^{1-\frac{3\beta}{4}+2\varepsilon} d_N}{(d_N(N - N^{1-\frac{3\beta}{4}+2\varepsilon}) - \varphi(N))\varphi(N)}\right)^{\varphi(N)} \to 0.
\end{aligned}$$

□

Lemma 2.5.8. *In the setting of Theorem 2.2.2 (ii) there exists a constant $C > 0$ such that*

$$\lim_{N\to\infty} \mathbb{P}\left(I^{(N)}_{\tau^{(N)}+1} \geq C \cdot N^{1-\frac{\beta}{2}+\varepsilon} \middle| \overline{\tau}^{(N)} < \infty\right) = 1.$$

for $\varepsilon > 0$ small enough.

Proof. According to Lemma 2.5.2, $\tau^{(N)} \leq \overline{\tau}^{(N)}$ whp. Thus using Lemma 2.5.7 the number of empty vertices at generation $\tau^{(N)}$ is whp at most $N^{1-\frac{\beta}{2}+5\varepsilon}$. By definition of $\tau^{(N)}$ there are at least $N^{1-\frac{3\beta}{4}+\varepsilon}$ infected individuals, and so at least $\Theta(v_N N^{1-\frac{3\beta}{4}+\varepsilon})$ parasites participate in new infections.
First we are going to show that the number of pairs of parasites present on infected vertices at generation $\tau^{(N)}$ are negligible compared to v_N. Denote by $A^{(N)}_{\tau^{(N)}}$ the number of parasites occupying an edge alone at generation $\tau^{(N)}$. Then for all functions φ_1, satisfying $\varphi_1(N) \to \infty$, we have

$$\lim_{N\to\infty} \mathbb{P}\left(A^{(N)}_{\tau^{(N)}} \geq N^{1-\frac{3\beta}{4}+\varepsilon}(v_N - \varphi_1(N)) \middle| \overline{\tau}^{(N)} < \infty\right) = 1. \tag{2.5.4}$$

Indeed, denote by $(K_i^{(N)})_{i\in\{1,\dots,N\}}$ the iid random variables giving the number of half-edges (connected to the vertices i, for $i \in \{1, ..., N\}$) that are occupied by at least two parasites in the generations at which the vertices get infected. We have for all $0 \leq k \leq \lfloor \frac{v_N}{2} \rfloor$

$$\mathbb{P}(K_1^{(N)} = k) \leq \frac{\binom{v_N}{2}\binom{v_N-2}{2}\dots\binom{v_N-2(k-1)}{2}}{k! \cdot d_N^{v_N}}.$$

Using Markov's inequality, we obtain that

$$\mathbb{P}\left(\sum_{i=1}^{N^{1-\frac{3\beta}{4}+\varepsilon}} K_i^{(N)} \geq N^{1-\frac{3\beta}{4}+\varepsilon}\varphi_1(N)\right) \leq \frac{\mathbb{E}[K_1^{(N)}]}{\varphi_1(N)} \to 0,$$

since $\mathbb{E}(K_1^{(N)})$ is uniformly bounded in N, see for a similar calculation Equation (2.4.2).
Denote by $H_i^{(N)}$ the number of half-edges that have already been formed for vertex i till generation $\tau^{(N)}$. Using Lemma 2.5.2 and a similar computation as the one at the end of the proof of Lemma 2.5.7 we obtain

$$\lim_{N\to\infty} \mathbb{P}\left(H_i^{(N)} \leq \varphi_2(N)\ \forall i \in I_{\tau^{(N)}}^{(N)}\right) = 1, \tag{2.5.5}$$

where $\liminf_N \varphi_2(N) \geq 5$.
Thus, using (2.5.4) and (2.5.5) the number of parasites that may cooperate by infecting a host from different edges is whp bounded from below by $N^{1-\frac{3\beta}{4}+\varepsilon}(v_N - 2\widetilde{\varphi}(N))$, where $\widetilde{\varphi}(N) := \max\{\varphi_1(N), \varphi_2(N)\}$.
In addition it can also happen that a parasite attacks a half-edge on which another parasite is located. In this case, these two parasites cannot infect a host. An upper bound for the probability that a parasite is involved in such kind of event is whp $\frac{Nv_N}{Nd_N - 2Nv_N}$. And so a lower bound on the number of available parasite is $N^{1-\frac{3\beta}{4}+\varepsilon}(v_N - 2\widetilde{\varphi}(N))(1 - \frac{v_N}{d_N - 2v_N})$. With this estimate we derive a whp lower bound on the number of infections occurring in the next generation.

Consider N boxes, assume the $N^{1-\frac{\beta}{2}+5\varepsilon}$ first ones (corresponding to the empty vertices) contain d_N positions, and the remaining ones (corresponding to the susceptible vertices) have $d_N - \varphi(N)$ positions, where $\liminf_N \varphi(N) \geq 5$ and $\varphi(N) = o(d_N)$. Assume that $N^{1-\frac{3\beta}{4}+\varepsilon}(v_N - 2\widetilde{\varphi}(N))(1 - \frac{v_N}{d_N - 2v_N})$ balls are uniformly distributed on the positions. Let $G^{(N)}$ be the number of boxes that contain $d_N - \varphi(N)$ positions and into which at least two balls are thrown. $G^{(N)}$ yields whp an estimate from below for the number of new infections. Using the same kind of computations as in the proof of Lemma 2.5.7 (using Chebyshev's Inequality, estimating expectation and variance of $G^{(N)}$) we arrive at the statement of the lemma. □

Lemma 2.5.9. *Under the conditions of Theorem 2.2.2 (ii) it holds*

$$\lim_{N\to\infty} \mathbb{P}\left(\overline{I}_{\tau^{(N)}+2}^{(N)} = N \middle| \overline{\tau}^{(N)} < \infty\right) = 1.$$

Proof. We aim to show that all hosts that have not been infected so far, get infected whp in generation $\tau^{(N)} + 2$. According to Lemma 2.5.8 we have whp $I_{\tau^{(N)}+1}^{(N)} \geq C \cdot N^{1-\frac{\beta}{2}+2\varepsilon}$. Hence, we have whp at least $C \cdot N^{1-\frac{\beta}{2}+2\varepsilon} v_N$ parasites that may infect the remaining hosts. However, some of these parasites may be placed on already linked half-edges or occupy half-edges together with other parasites. Hosts that got infected in generation $\tau^{(N)} + 1$ have been attacked by at most one parasite in any generation $n \leq \tau^{(N)}$. By Lemma 2.5.6 whp $\tau^{(N)} \leq (1+\delta)\frac{\alpha \log N}{\log\left(\frac{c^2}{2}+x\right)}$ hence the number of formed edges is whp limited by $(1+\delta)\frac{\alpha \log N}{\log\left(\frac{c^2}{2}+x\right)}$ for any of these hosts for any $\delta > 0$.
Furthermore in generation $\tau^{(N)}$ we have according to Lemma 2.5.7 and because $\tau^{(N)} \leq \overline{\tau}^{(N)}$ whp $I_{\tau^{(N)}} \leq N^{1-\frac{\beta}{2}+5\varepsilon}$. So by an application of Chebyshev's Inequality we can estimate that a host gets

attacked in generation $\tau^{(N)}+1$ by at most $\frac{N^{1-\frac{\beta}{2}+6\varepsilon}v_N}{N} \sim N^{6\varepsilon}$ parasites with probability $1-\frac{1}{N^\varepsilon}$. Consequently, at least a proportion $1-\frac{1}{N^\varepsilon}$ of the hosts infected at generation $\tau^{(N)}+1$ occupy whp a vertex with at least

$$e_N := d_N - \left((1+\delta)\frac{\alpha \log N}{\log\left(\frac{c^2}{2}+x\right)}\right) - \frac{N^{1-\frac{\beta}{2}+6\varepsilon}v_N}{N}$$

free half-edges and the probability that the parasites generated in these hosts occupy a half-edge that has been linked before or that is occupied already by another parasite can be estimated from above by

$$\frac{v_N + d_N - e_N}{d_N} \sim \frac{v_N}{d_N},$$

for $\varepsilon > 0$ small enough.
In summary, we have whp at least

$$m_N := C \cdot N^{1-\frac{\beta}{2}+2\varepsilon}\left(1-\frac{1}{N^\varepsilon}\right) v_N \left(1-\frac{v_N+d_N-e_N}{d_N}\right)$$

free half-edges occupied with at least one parasite that may attack so far uninfected hosts.
Similarly an up to generation $\tau^{(N)}+2$ uninfected host has whp at least

$$f_N := d_N - \left((1+\delta)\frac{\alpha \log N}{\log\left(\frac{c^2}{2}+x\right)}\right)$$

free half-edges. So, the probability that an up to generation $\tau^{(N)}+2$ uninfected host gets attacked by at most one of the m_N parasites (and hence with high probability remains uninfected) can be estimated from above by

$$\left(\left(1-\frac{f_N}{d_N N - v_N N}\right)^{m_N} + \left(1-\frac{f_N}{d_N N - v_N N}\right)^{m_N-1} m_N \frac{f_N}{d_N N - v_N N}\right)(1+o(1))$$
$$\sim N^{2\varepsilon}\exp(-N^{2\varepsilon}).$$

The number of uninfected hosts at the beginning of generation $\tau^{(N)}+2$ is at most N. Consequently, the probability that at least one of these hosts remains uninfected till the end of generation $\tau^{(N)}+2$ can be estimated from above by a probability proportional to

$$N\left(\exp(-N^{2\varepsilon})N^{2\varepsilon}\right) = o(1),$$

which yields the claim of Lemma 2.5.9. □

Proof of Proposition 2.5.1. According to Lemma 2.5.2 once $\overline{\mathcal{I}}_n^{(N)}$ has reached the level $N^{1-\frac{3\beta}{4}+2\varepsilon}$ also $\mathcal{I}^{(N)}$ has reached the level $N^{1-\frac{3\beta}{4}+\varepsilon}$. Moreover, according to Lemma 2.5.7 whp $\overline{I}^{(N)}_{\overline{\tau}^{(N)}} \in \mathcal{O}(N^{1-\frac{\beta}{2}+5\varepsilon})$. Consequently, according to Lemma 2.5.8, the size of $\mathcal{I}^{(N)}$ is at generation $\tau^{(N)}+1$ whp at least $C \cdot N^{1-\frac{\beta}{2}+2\varepsilon}$ for some appropriate constant $C > 0$. Finally, we can apply Lemma 2.5.9, which yields the result. □

2.6 Proof of the main Theorem

Proof of Theorem 2.2.2
We start with the proof of Theorem 2.2.2 (ii):
For the upper bound on the invasion probability consider for a given $\ell_N > 0$ the event

$$F^{(N)}_{\ell_N} := \{\exists n \in \mathbb{N}_0 : \overline{I}^{(N)}_n \geq \ell_N\}.$$

Then given $0 < u \leq 1$ we have for any ℓ_N with $\ell_N \leq uN$

$$\mathbb{P}(E^{(N)}_u) \leq \mathbb{P}(F^{(N)}_{\ell_N}).$$

For any sequence $(\ell_N)_{N\in\mathbb{N}}$ with $\ell_N \to \infty$ and $\ell_N^3 v_N^2 \in o(N)$ we have by Proposition 2.3.2

$$\mathbb{P}(F^{(N)}_{\ell_N}) \leq \mathbb{P}\left(\exists n \in \mathbb{N}_0 : \overline{Z}^{(N)}_{n,u} \geq \ell_N\right),$$

and by Proposition 2.3.3

$$\lim_{N\to\infty} \mathbb{P}\left(\exists n \in \mathbb{N}_0 : \overline{Z}^{(N)}_{n,u} \geq \ell_N\right) = \pi(c,x).$$

Since for any given $0 < u \leq 1$ and any sequence (ℓ_N) with $\ell_N^3 v_N^2 \in o(N)$ we have for N large enough $\ell_N \leq uN$. Hence, in summary

$$\limsup_{N\to\infty} \mathbb{P}(E^{(N)}_u) \leq \pi(c,x),$$

which yields the claimed upper bound on the invasion probability.

For the lower bound we first apply Lemma 2.4.11, which yields the lower bound $\pi(c,x) + o(1)$ on the probability that N^α hosts eventually get infected with $\alpha = 1 - \frac{3}{4}\beta + 2\varepsilon$. Furthermore we can choose $\varepsilon > 0$ small enough such that $\alpha < 1 - \frac{\beta}{2}$. Then the assumptions of Proposition 2.5.1 are fulfilled and we obtain the claimed upper bound on the invasion probability, since once the level $N^{1-\frac{3\beta}{4}+2\varepsilon}$ is reached with probability 1+ o(1) the remaining hosts get infected as well, in particular any proportion u of the host population for $0 < u \leq 1$.

Proof of Theorem 2.2.2 (i):
The proof of Theorem 2.2.2 (i) relies on the same arguments as the proof of Theorem 2.2.2 (ii).

Indeed, since $v_N = o(\sqrt{d_N})$ we have for any $c > 0$ that the upper Galton-Watson process from Definition 2.3.1, where v_N in this Definition is replaced by $c\sqrt{d_N}$, can be coupled with $\mathcal{I}^{(N)}$, such that $\overline{\mathcal{I}}^{(N)}$ is bounded from above by $\overline{\mathcal{Z}}_u^{(N)}$ until $\overline{\mathcal{I}}^{(N)}$ is not further increasing or is exceeding the threshold ℓ_N for an appropriate sequence $(\ell_N)_{N\in\mathbb{N}}$ fulfilling the conditions of Proposition 2.3.2. Consequently, by Proposition 2.3.3 for all $0 < u \leq 1$, the invasion probability satisfies $\mathbb{P}(E_u^{(N)}) \leq \pi(c, x) + o(1)$. But since $x \leq 1$, we have $\lim_{c\downarrow 0} \pi(c, x) = 0$ and so the statement follows, since $c > 0$ was arbitrary.

Proof of Theorem 2.2.2 (iii):
Trivially the invasion probability is upper bounded by 1. For the lower bound we can again rely on results of the proof of Theorem 2.2.2 (ii). We consider, alongside the host-parasite model with the parameters (d_N, v_N, ρ_N) fulfilling the conditions from Theorem 2.2.2 (iii), a host-parasite model with parameters $(d_N, v_N^{(c)}, \rho_N)$, where we set $v_N^{(c)} = c\sqrt{d_N}$, i.e. the parameters $(d_N, v_N^{(c)}, \rho_N)$ fulfill the conditions from Theorem 2.2.2 (ii). We couple these two host-parasite models by following, in the second host-parasite model, at each host infection instead of all v_N parasite offspring only the first $v_N^{(c)}$ parasites. In this manner the process $\overline{\mathcal{I}}^{(N)}$ can be estimated from below by the corresponding process $\overline{\mathcal{I}}_c^{(N)}$ of the host-parasite model with parameters $(d_N, v_N^{(c)}, \rho_N)$. According to Theorem 2.2.2 (ii) a lower bound on the invasion probability of this model is $\pi(c, x) + o(1)$. Since for $c \to \infty$ we have $\pi(c, x) \to 1$ and c can be chosen arbitrarily large this yields the claim of Theorem 2.2.2 (iii).

Acknowledgements We thank Anton Wakolbinger for valuable hints and comments.

Funding CP acknowledges support from the German Research Foundation through grant PO-2590/1-1.

Appendix

Proof of Lemma 2.5.3. Using the almost sure convergence of $\left(\frac{Z_n}{Z_0} m^{-n}\right)_{n\in\mathbb{N}_0}$ to W, it follows that for all $\omega \in \{W > 0\}$, for all $\varepsilon > 0$ there exists $\tilde{n} \in \mathbb{N}_0$, such that for all $n \geq \tilde{n}$

$$(W - \varepsilon)m^n \leq \frac{Z_n}{Z_0} \leq (W + \varepsilon)m^n.$$

Introduce

$$\tau_{\underline{N}^\alpha} := \inf\left\{n \in \mathbb{N}_0 : (W + \varepsilon)m^n \geq \frac{N^\alpha}{Z_0}\right\},$$
$$\tau_{\overline{N}^\alpha} := \inf\left\{n \in \mathbb{N}_0 : (W - \varepsilon)m^n \geq \frac{N^\alpha}{Z_0}\right\}.$$

We have $\tau_{\underline{N}^\alpha} \leq \tau_{N^\alpha} \leq \tau_{\overline{N}^\alpha}$, for N large enough, and the following lower and upper bounds for $\tau_{\underline{N}^\alpha}$ and $\tau_{\overline{N}^\alpha}$ respectively hold for ε small enough

Notation	Meaning	Defined in
d_N	number of edges per vertex scaling: $\Theta(N^\beta), 0 < \beta < 1$	Section 2.2.1
v_N	number of offspring parasites, scaling in Theorem 2.2.2 (ii): $v_N \sim c\sqrt{d_N}$	—— " ——
ρ_N	infection probability of a single parasite, scaling: $\rho_N v_N \to x \in [0,1]$	—— " ——
$\pi(c,x)$	survival probability of a GWP with $\text{Pois}(\frac{c^2}{2}+x)$ offspring distribution	
$\mathcal{I}^{(\mathcal{N})} = \left(I_n^{(N)}\right)_{n\in\mathbb{N}}$	process counting the number of infected hosts	—— " ——
$\overline{\mathcal{I}}^{(N)} = \left(\overline{I}_n^{(N)}\right)_{n\in\mathbb{N}}$	process counting the total number of hosts infected before generation n	—— " ——
$\mathcal{Z}_a^{(N)} = \left(Z_{n,a}^{(N)}\right)_{n\in\mathbb{N}}, a = u, l$	GWP used for approximating $\overline{I}^{(N)}$ from above (a=u) and from below (a=l)	Def. 2.3.1 and Def. 2.4.5, resp.
$\overline{\mathcal{Z}}_a^{(N)} = \left(\overline{Z}_{n,a}^{(N)}\right)_{n\in\mathbb{N}}, a = u, l$	total size of the process $\mathcal{Z}_a^{(N)}$ until generation n	—— " ——
$\left(p_{k,a}^{(N)}\right)_{k\in\mathbb{N}_0}, a = u, l$	probability weights of the offspring distribution of $\mathcal{Z}_a^{(N)}$	—— " ——
$\mathcal{J}^{(N)} = \left(J_n^{(N)}\right)_{n\in\mathbb{N}}$	process counting the number of infected hosts in the model from Definition 2.4.3	Def. 2.4.3
$\overline{\mathcal{J}}^{(N)} = \left(\overline{J}_n^{(N)}\right)_{n\in\mathbb{N}}$	process counting the total number of hosts infected before generation n in the model from Definition 2.4.3	—— " ——

Table 2.1: Table of frequently used notation

$$\tau_{\underline{N}^\alpha} \geq \frac{(\alpha-\gamma)\log N}{\log m} - \frac{\log(W+\varepsilon)}{\log m} - \frac{\log\left(1-\frac{\varphi(N)}{N^\gamma}\right)}{\log(m)},$$
$$\tau_{\overline{N}^\alpha} \leq \frac{(\alpha-\gamma)\log N}{\log m} - \frac{\log(W-\varepsilon)}{\log m} - \frac{\log\left(1-\frac{\varphi(N)}{N^\gamma}\right)}{\log(m)} + 1,$$

which finally yields the following inequality

$$1 - \frac{\log(W+\varepsilon)+\log\left(1-\frac{\varphi(N)}{N^\gamma}\right)}{(\alpha-\gamma)\log N} \leq \frac{\tau_{N^\alpha}\log(m)}{(\alpha-\gamma)\log N} \leq 1 - \frac{\log(W-\varepsilon)+\log\left(1-\frac{\varphi(N)}{N^\gamma}\right)-\log(m)}{(\alpha-\gamma)\log N}.$$

Taking the limit $N\to\infty$ concludes the proof. □

For the proof of Lemma 2.5.7 we need in addition to Lemma 2.5.3 estimates on the number of vertices that get attacked by at least two parasites. For this purpose we consider the following experiment.

Let $(S_N)_{N\in\mathbb{N}}$, $(D_i^{(N)})_{1\leq i\leq S_N, N\in\mathbb{N}}$ be deterministic sequences of integers with $S_N \sim N$ and $D_i^{(N)} = d_N + O(1)$. Assume we have S_N boxes with box number i having $D_i^{(N)}$ many positions, and assume $b_N := v_N N^{1-\frac{3\beta}{4}+2\varepsilon} \in \Theta(N^{1-\frac{\beta}{4}+2\varepsilon})$ balls are uniformly distributed on the positions of the boxes, such that each position gets occupied at most once, for some $\varepsilon>0$ small enough that $1-\frac{\beta}{4}+2\varepsilon<1$. Denote by $G_i^{(N)}$ the number of balls in box number i, and by $G^{(N)} := \sum_{i=1}^{S_N} \mathbb{1}_{\{G_i^{(N)}\geq 2\}}$ the number of boxes containing at least 2 balls. The following statements on the random variables $G_i^{(N)}$ and $G^{(N)}$ we apply in the proof of Lemma 2.5.7.

Lemma 2.6.1.

$$\mathbb{P}\left(\{G_i^{(N)}\leq 1\}\right) = 1 - \frac{b_N^2}{2S_N^2} + \frac{b_N^3}{3S_N^3} - \frac{b_N^4}{8S_N^4} + \mathcal{O}\left(\frac{b_N^5}{S_N^5}\right),$$
$$\mathbb{P}\left(\{G_i^{(N)}\leq 1\}\cap\{G_j^{(N)}\leq 1\}\right) = 1 - \frac{b_N^2}{S_N^2} + \frac{2}{3}\frac{b_N^3}{S_N^3} + \mathcal{O}\left(\frac{b_N^5}{S_N^5}\right),$$
$$\mathbb{V}[G^{(N)}] = \mathcal{O}\left(\frac{b_N^4}{S_N^2}\cdot\frac{b_N}{S_N}\right).$$

Proof. During the computation, we are using the following asymptotic estimates $\frac{1}{d_N} = o\left(\frac{b_N^4}{S_N^4}\right)$, $\frac{1}{S_N} = o\left(\frac{b_N^4}{S_N^4}\right)$, $\frac{1}{S_N d_N} = o\left(\frac{b_N^5}{S_N^5}\right)$, $\frac{b_N}{S_N d_N} = o\left(\frac{b_N^5}{S_N^5}\right)$, $\frac{b_N}{S_N^2} = o\left(\frac{b_N^5}{S_N^5}\right)$, $\frac{b_N^2}{S_N^2 d_N} = o\left(\frac{b_N^5}{S_N^5}\right)$.

To prepare the proof of the three estimates we first expand a few typical factors that will arise in the calculations of the two probability terms. Denote by $T := \sum_{i=1}^{S_N} D_i^{(N)}$ the total number of

positions. We expand

$$\begin{aligned}\left(\frac{T-D_i^{(N)}-b_N}{T-b_N}\right)^{b_N} &= \exp\left\{b_N \log\left[1-\frac{1}{S_N}\cdot\left(1+o\left(\frac{b_N^4}{S_N^4}\right)\right)\right]\right\}\\ &= \exp\left(-\frac{b_N}{S_N}\right)\left(1+o\left(\frac{b_N^5}{S_N^5}\right)\right)\\ &= 1-\frac{b_N}{S_N}+\frac{b_N^2}{2S_N^2}-\frac{b_N^3}{6S_N^3}+\frac{b_N^4}{24S_N^4}+\mathcal{O}\left(\frac{b_N^5}{S_N^5}\right).\end{aligned}$$

and similarly we have for $k\in\mathbb{N}$ such that $\frac{b_N^k}{T^{k-1}}=o\left(\frac{b_N^5}{S_N^5}\right)$

$$\begin{aligned}&\frac{\left(1-\frac{b_N}{T}\right)^T}{\left(1-\frac{b_N}{T-D_i^{(N)}}\right)^{T-D_i^{(N)}}}\\ =&\exp\left\{T\left[\log(1-\tfrac{b_N}{T})-\log\left(1-\frac{b_N}{T}\left[1+\tfrac{1}{S_N}+o\left(\tfrac{1}{S_N d_N}\right)\right]\right)\right]\right\}\\ &\qquad\cdot\exp\left(D_i^{(N)}\log\left(1-\tfrac{b_N}{T}\left[1+\mathcal{O}\left(\tfrac{1}{S_N}\right)\right]\right)\right)\\ =&\exp\left(T\left(-\tfrac{b_N}{T}-\cdots-\tfrac{b_N^k}{kT^k}\right)\right)\exp\left(-\tfrac{b_N}{S_N}+o\left(\tfrac{b_N^5}{S_N^5}\right)\right)\\ &\qquad\cdot\exp\left\{T\left[\tfrac{b_N}{T}\left(1+\tfrac{1}{S_N}+o\left(\tfrac{1}{S_N d_N}\right)\right)+\cdots+\tfrac{b_N^k}{kT^k}\left(1+\tfrac{1}{S_N}+o\left(\tfrac{1}{S_N d_N}\right)\right)+o\left(\tfrac{b_N^k}{T^k}\right)\right]\right\}\\ =&\exp\left(-\tfrac{b_N}{S_N}+o\left(\tfrac{b_N^5}{S_N^5}\right)\right)\exp\left(\tfrac{b_N}{S_N}+o\left(\tfrac{b_N^5}{S_N^5}\right)\right)\\ =&1+o\left(\tfrac{b_N^5}{S_N^5}\right).\end{aligned}$$

Using the asymptotic expansion of the factorial and the two previous estimates we get

$$\begin{aligned}&\frac{(T-D_i^{(N)})!}{(T-D_i^{(N)}-b_N)!}\cdot\frac{(T-b_N)!}{T!}\\ =&\frac{(T-D_i^{(N)})^{T-D_i^{(N)}}}{(T-D_i^{(N)}-b_N)^{T-D_i^{(N)}-b_N}}\cdot\left(\frac{T-D_i^{(N)}}{T-D_i^{(N)}-b_N}\right)^{\frac{1}{2}}\\ &\qquad\cdot\frac{(T-b_N)^{T-b_N}}{T^T}\cdot\left(\frac{T-b_N}{T}\right)^{\frac{1}{2}}\left(1+o\left(\frac{b_N^5}{S_N^5}\right)\right)\\ =&\frac{\left(1-\frac{b_N}{T}\right)^T}{\left(1-\frac{b_N}{T-D_i^{(N)}}\right)^{T-D_i^{(N)}}}\cdot\left(\frac{T-D_i^{(N)}-b_N}{T-b_N}\right)^{b_N}\left(1-\frac{b_N}{T-D_i^{(N)}}\right)^{-\frac{1}{2}}\\ &\qquad\cdot\left(1-\frac{b_N}{T}\right)^{\frac{1}{2}}\left(1+o\left(\frac{b_N^5}{S_N^5}\right)\right)\\ =&1-\frac{b_N}{S_N}+\frac{b_N^2}{2S_N^2}-\frac{b_N^3}{6S_N^3}+\frac{b_N^4}{24S_N^4}+\mathcal{O}\left(\frac{b_N^5}{S_N^5}\right),\end{aligned}$$

and with similar calculations

$$\frac{(T-D_i^{(N)}-D_j^{(N)})!}{(T-D_i^{(N)}-D_j^{(N)}-b_N)!}\cdot\frac{(T-b_N)!}{T!}=1-\frac{2b_N}{S_N}+\frac{2b_N^2}{S_N^2}-\frac{4b_N^3}{3S_N^3}+\frac{2b_N^4}{3S_N^4}+\mathcal{O}\left(\frac{b_N^5}{S_N^5}\right).$$

Now we are ready to estimate the two probabilities

$$\begin{aligned}
&\mathbb{P}(\{G_i^{(N)}\le 1\})\\
=&\frac{(T-D_i)!}{((T-D_i-b_N)!}\cdot\frac{(T-b_N)!}{T!}+b_ND_i^{(N)}\cdot\frac{(T-D_i)!}{(T-D_i-(b_N-1))!}\cdot\frac{(T-b_N)!}{T!}\\
=&\frac{(T-D_i)!}{(T-D_i-b_N)!}\cdot\frac{(T-b_N)!}{T!}\left(1+\frac{b_ND_i^{(N)}}{T-D_i-(b_N-1)}\right)\\
=&\frac{(T-D_i)!}{(T-D_i-b_N)!}\cdot\frac{(T-b_N)!}{T!}\left(1+\frac{b_N}{S_N}+o\left(\frac{b_N^5}{S_N^5}\right)\right)\\
=&\left(1-\frac{b_N}{S_N}+\frac{b_N^2}{2S_N^2}-\frac{b_N^3}{6S_N^3}+\frac{b_N^4}{24S_N^4}+\mathcal{O}\left(\frac{b_N^5}{S_N^5}\right)\right)\left(1+\frac{b_N}{S_N}+o\left(\frac{b_N^5}{S_N^5}\right)\right)\\
=&1-\frac{b_N^2}{2S_N^2}+\frac{b_N^3}{3S_N^3}-\frac{b_N^4}{8S_N^4}+\mathcal{O}\left(\frac{b_N^5}{S_N^5}\right).
\end{aligned}$$

$$\begin{aligned}
&\mathbb{P}(\{G_i^{(N)}\le 1\}\cap\{G_j^{(N)}\le 1\})\\
=&\frac{(T-b_N)!}{T!}\cdot\frac{(T-D_i^{(N)}-D_j^{(N)})!}{(T-D_i^{(N)}-D_j^{(N)}-b_N)!}\\
&\quad+b_ND_i^{(N)}\frac{(T-b_N)!}{T!}\cdot\frac{(T-D_i^{(N)}-D_j^{(N)})!}{(T-D_i^{(N)}-D_j^{(N)}-(b_N-1))!}\\
&\quad+b_ND_j^{(N)}\frac{(T-b_N)!}{T!}\cdot\frac{(T-D_i^{(N)}-D_j^{(N)})!}{(T-D_i^{(N)}-D_j^{(N)}-(b_N-1))!}\\
&\quad+b_N(b_N-1)D_i^{(N)}D_j^{(N)}\frac{(T-b_N)!}{T!}\cdot\frac{(T-D_i^{(N)}-D_j^{(N)})!}{(T-D_i^{(N)}-D_j^{(N)}-(b_N-2))!}\\
=&\frac{(T-b_N)!}{T!}\cdot\frac{(T-D_i^{(N)}-D_j^{(N)})!}{(T-D_i^{(N)}-D_j^{(N)}-b_N)!}\cdot\left[1+\frac{b_ND_i^{(N)}}{T-D_i^{(N)}-D_j^{(N)}-(b_N-1)}\right.\\
&\qquad\left.+\frac{b_ND_j^{(N)}}{T-D_i^{(N)}-D_j^{(N)}-(b_N-1)}\right.
\end{aligned}$$

$$
\begin{aligned}
&\qquad + \frac{b_N(b_N-1)D_i^{(N)}D_j^{(N)}}{(T-D_i^{(N)}-D_j^{(N)}-(b_N-1))(T-D_i^{(N)}-D_j^{(N)}-(b_N-2))}\Bigg] \\
&= \frac{(T-b_N)!}{T!}\cdot\frac{(T-D_i^{(N)}-D_j^{(N)})!}{(T-D_i^{(N)}-D_j^{(N)}-b_N)!} \\
&\qquad \cdot\left(1+\frac{2b_N}{S_N}\left(1+o\left(\frac{b_N^4}{S_N^4}\right)\right)+\frac{b_N^2}{S_N^2}\left(1+o\left(\frac{b_N^3}{N^3}\right)\right)\right) \\
&= \left(1-\frac{2b_N}{S_N}+\frac{2b_N^2}{S_N^2}-\frac{4}{3}\frac{b_N^3}{S_N^3}+\frac{2}{3}\frac{b_N^4}{S_N^4}+\mathcal{O}\left(\frac{b_N^5}{S_N^5}\right)\right)\cdot\left(1+\frac{2b_N}{S_N}+\frac{b_N^2}{S_N^2}+o\left(\frac{b_N^5}{S_N^5}\right)\right) \\
&= 1-\frac{b_N^2}{S_N^2}+\frac{2}{3}\frac{b_N^3}{S_N^3}+\mathcal{O}\left(\frac{b_N^5}{S_N^5}\right).
\end{aligned}
$$

The estimate on the variance is obtained using the two previous computations

$$
\begin{aligned}
&\mathbb{V}[G^{(N)}] = \mathbb{V}\left[\sum_{i=1}^{S_N}\mathbb{1}_{\{i\geq 2\}}\right] = \mathbb{E}\left[\left(\sum_{i=1}^{S_N}\mathbb{1}_{\{G_i^{(N)}\geq 2\}}\right)^2\right]-\left(\mathbb{E}\left[\sum_{i=1}^{S_N}\mathbb{1}_{\{G_i^{(N)}\geq 2\}}\right]\right)^2 \\
&= \sum_{i\neq j}\mathbb{E}\left[\mathbb{1}_{\{G_i^{(N)}\geq 2\}}\mathbb{1}_{\{G_j^{(N)}\geq 2\}}\right]+\mathbb{E}\left[\sum_{i=1}^{S_N}\mathbb{1}_{\{G_i^{(N)}\geq 2\}}\right]-\left(\mathbb{E}\left[\sum_{i=1}^{S_N}\mathbb{1}_{\{G_i^{(N)}\geq 2\}}\right]\right)^2 \\
&= S_N(S_N-1)\mathbb{P}\left(\{G_i^{(N)}\geq 2\}\cap\{G_j^{(N)}\geq 2\}\right)+S_N\mathbb{P}(\{G_i^{(N)}\geq 2\})-S_N^2\mathbb{P}(\{G_i^{(N)}\geq 2\}^2 \\
&= S_N^2\left(\mathbb{P}(\{G_i^{(N)}\geq 2\}\cap\{G_j^{(N)}\geq 2\})-\mathbb{P}(\{G_i^{(N)}\geq 2\})^2\right) \\
&\qquad + S_N\left(\mathbb{P}(\{G_i^{(N)}\geq 2\})-\mathbb{P}(\{G_i^{(N)}\geq 2\}\cap\{G_j^{(N)}\geq 2\})\right) \\
&= S_N^2(\mathbb{P}(\{G_i^{(N)}\leq 1\}\cap\{G_j^{(N)}\leq 1\})-\mathbb{P}(\{G_i^{(N)}\leq 1\})^2) \\
&\qquad + S_N(\mathbb{P}(\{G_i^{(N)}\leq 1\})-\mathbb{P}(\{G_i^{(N)}\leq 1\}\cap\{G_j^{(N)}\leq 1\})) \\
&= S_N^2\left[1-\frac{b_N^2}{S_N^2}+\frac{2}{3}\frac{b_N^3}{S_N^3}+\mathcal{O}\left(\frac{b_N^5}{S_N^5}\right)-\left(1-\frac{b_N^2}{2S_N^2}+\frac{b_N^3}{3S_N^3}-\frac{b_N^4}{8S_N^4}+\mathcal{O}\left(\frac{b_N^5}{S_N^5}\right)\right)^2\right] \\
&+ S_N\left[1+\mathcal{O}\left(\frac{b_N^2}{S_N^2}\right)-\left(1+\mathcal{O}\left(\frac{b_N^2}{S_N^2}\right)\right)\right] \\
&= S_N^2\left[1-\frac{b_N^2}{S_N^2}+\frac{2}{3}\frac{b_N^3}{S_N^3}+\mathcal{O}\left(\frac{b_N^5}{S_N^5}\right)-\left(1-\frac{b_N^2}{S_N^2}+\frac{2}{3}\frac{b_N^3}{S_N^3}+\mathcal{O}\left(\frac{b_N^5}{S_N^5}\right)\right)\right]+\mathcal{O}\left(\frac{b_N^2}{S_N}\right) \\
&= \mathcal{O}\left(\frac{b_N^5}{S_N^3}\right)+\mathcal{O}\left(\frac{b_N^2}{S_N}\right) \\
&= \mathcal{O}\left(\frac{b_N^4}{S_N^2}\cdot\frac{b_N}{S_N}\right).
\end{aligned}
$$

□

Chapter 3

Spatial Invasion of Cooperative Parasites

This Chapter corresponds to the publication [2] "Spatial Invasion of Cooperative Parasites", written in collaboration with Cornelia Pokalyuk, Marco Seiler and Hung Tran for the simulations. It is published in Theoretical Population Biology for the special issue for the 60th birthday of Alison Etheridge.

Abstract:

In this paper we study invasion probabilities and invasion times of cooperative parasites spreading in spatially structured host populations. The spatial structure of the host population is given by a random geometric graph on $[0,1]^n$, $n \in \mathbb{N}$, with a Poisson(N)-distributed number of vertices and in which vertices are connected over an edge when they have a distance of at most r_N with r_N of order $N^{(\beta-1)/n}$ for some $0 < \beta < 1$. At a host infection many parasites are generated and parasites move along edges to neighbouring hosts. We assume that parasites have to cooperate to infect hosts, in the sense that at least two parasites need to attack a host simultaneously. We find lower and upper bounds on the invasion probability of the parasites in terms of survival probabilities of branching processes with cooperation. Furthermore, we characterise the asymptotic invasion time.

An important ingredient of the proofs is a comparison with infection dynamics of cooperative parasites in host populations structured according to a complete graph, i.e. in well-mixed host populations. For these infection processes we can show that invasion probabilities are asymptotically equal to survival probabilities of branching processes with cooperation.

Furthermore, we build on proof techniques developed in [1], where an analogous invasion process has been studied for host populations structured according to a configuration model.

We substantiate our results with simulations.

Keywords: cooperation, host-parasite population dynamics, invasion probability, invasion time, spatial host population structure, random geometric graph.

3.1 Introduction

Understanding the dynamics of infection processes is a highly relevant and active research field. In this study we are interested in the spread of cooperative parasites in spatially structured host

populations. Cooperative behaviour is observed in many biological systems, see [72]. The main biological motivation for our model stems from observations made on phages, that is viruses infecting bacteria. Bacteria own various mechanisms to defend against phages. Defense on the basis of CRISPR-Cas system is widespread in bacteria. Certain phages, called anti-CRISPR phages, can overcome these defense mechanism by cooperation. Only when anti-CRISPR-phages infect simultaneously or subsequently a CRISPR-resistant bacterium the infection gets likely to be successful, see [73, 74].

Besides the motivation stemming from application, models which incorporate cooperative mechanisms are also highly interesting from a mathematical point. For example González Casanova, Pardo and Perez [75] show that for a branching process with cooperation the survival probability is positive as long as the probability to generate offspring for pairs of individuals is non-zero. In case of survival it explodes in finite time. In the papers [76], [77] and more recently [78] mean-field limits of systems with cooperative reproduction are studied. Mach et al. find in [78] that the mean-field equation corresponding to certain interacting particle systems with cooperation can have more fixed points than the corresponding mean-field equations of classical infection models such as the contact process. This can be seen as evidence that in the microscopic model there could exist more extremal invariant laws as compared to the non-cooperative infection models. Sturm and Swart studied in [33] such a cooperative microscopic model. To be precise they considered a nearest neighbour cooperative branching-coalescing random walk on $\mathbb{Z}$. In comparison with the classical branching-coalescing random walk a subcritical phase exists, where the system ends up with only one particle. Superficially, this cooperative branching-coalescing system seems to be similar to a contact process, but a closer look reveals some apparent differences. For example [33] show that the decay rates in the subcritical regime are polynomial and not exponential as for the contact process.

In [1] the invasion of cooperative parasites in host populations structured according to a configuration model was studied. In this paper a parameter regime was considered, in which parasites have many offspring and a parasite can reach many, but not all hosts. In the critical scale this resulted in an invasion process which could be approximated initially by a Galton Watson process with roughly Poisson offspring numbers.

In this manuscript we consider the spread of cooperative parasites in host populations that have a finite-dimensional spatial structure. More precisely, we assume that the (immobile) hosts are distributed on an n-dimensional cube $[0,1]^n$ according to a Poisson point process. Parasites can move in every generation up to some fixed distance in space and attack the hosts located in this region. As in [1] we consider a parameter regime in which parasites have many offspring and can reach many, but not all hosts, as well as hosts need to be attacked jointly by parasites for successful parasite reproduction.

However, in contrast to the case considered in [1] the invasion process is already in the initial phase poorly approximated by a Galton-Watson process. The reason is that parasites generated in different hosts have in the spatial setting often a good chance to cooperate, because infected hosts are located close to each other. To arrive at lower and upper bounds on the invasion probability we compare the spread with infection dynamics caused by cooperative parasites spreading on complete graphs. The number of vertices of these complete graphs yield upper and lower bounds on the number of hosts, with which parasites generated on different hosts can cooperate. We show that the invasion probabilities of these infection processes on complete graphs are asymptotically equal to survival probabilities of certain branching processes with cooperation, a result that is of interest on its own. Furthermore, we show that the spatial infection processes can be coupled from above and below with these branching processes with cooperation until either the parasite population dies

or a sufficiently large amount of hosts are infected so that afterwards with high probability (i.e. asymptotically with probability 1, abbreviated as whp in the following) the parasite population will spread through the whole host population.

Once a sufficiently large number of hosts is infected, we show that the parasite population spreads whp at linear, almost maximal speed. As in the considered scaling the initial phase, which is decisive for survival of the parasite population, takes place only on a negligible amount of space, invasion time is basically determined by the time frame in which the parasite population spreads linearly fast. This yields our asymptotic result on the invasion time. Here again a clear difference to the model in [1] occurs, in which the final phase of invasion is finished after a constant number of steps.

By means of simulations we study the fit of the upper and lower bound on the invasion probability and our prediction for the invasion time. Interestingly, we find that the upper bound on the invasion probability matches very well with simulated invasion probabilities.

3.2 Main results

3.2.1 Model definition and main results

Consider the n-dimensional cube $[0,1]^n$, which we will denote by $M = M_n$ in the following. Measure distances on M according to the maximum metric denoted by ρ, i.e. for $x = (x_1, ..., x_n), y = (y_1, ..., y_n) \in [0,1]^n$ we have $\rho(x,y) = \max_{i=1,...,n}\{|x_i - y_i|\}$. Consider a homogeneous Poisson point process with intensity measure $N\lambda^n(\,\cdot\,)$ on $[0,1]^n$, where λ^n denotes the Lebesgue measure on $\mathbb{R}^n$ and $N \in \mathbb{N}$. In particular the number of Poisson points contained in a set $S \subset [0,1]^n$ of volume s is Poisson distributed with parameter sN. Denote the set of the Poisson points by $\mathcal{V} = \mathcal{V}^{(N)}$. Build a random geometric graph on $[0,1]^n$ by connecting all points in $\mathcal{V}^{(N)}$ over an edge which have a distance of at most $r_N > 0$ with respect to the metric ρ. Denote the set of edges by $\mathcal{E} = \mathcal{E}^{(N)}$ and the random geometric graph by $\mathcal{G} = \mathcal{G}^{(N)} = \mathcal{G}^{(r_N)} = (\mathcal{V}^{(N)}, \mathcal{E}^{(N)})$.

On $\mathcal{G}$ we consider the following infection process. At the beginning place on each vertex a host. These hosts can get infected with parasites. Choose the vertex $x_0 = x_0^{(N)} \in \mathcal{V}$ closest to the center of the cube $[0,1]^n$, denoted by x_c. We assume that the host on this vertex gets infected in the first generation $g = 0$. This means that the host dies and $v_N \in \mathbb{N}$ offspring parasites are generated on x_0. Then the infection process continues in discrete generations. At the beginning of each generation each parasite chooses uniformly at random and independently of all other parasites an edge, that is adjacent to the vertex on which the parasite is located. Along this edge the parasite moves to the neighbouring vertex and attacks the host on this vertex, if a host is still available. After movement of parasites, offspring parasites are generated and hosts die according to the following rules. If a vertex is occupied by a host and at least two parasites attack the host, the host on the vertex gets infected, dies and v_N parasites are generated. If only a single parasite attacks a host, it dies and the host stays alive. If a parasite arrives at an unoccupied vertex, it dies.

If a vertex is occupied/not occupied with a host, in the following we will call this vertex *occupied/unoccupied vertex*. Sometimes we also speak of *a susceptible/so far uninfected vertex*, if a host on a vertex did not yet get infected. Similarly, we say that a *vertex* is *infected* in some generation g, if the host on the vertex is in this generation infected.

Denote by $\mathcal{S}_g = \mathcal{S}_g^{(N)} \in \mathcal{V}$, $\mathcal{I}_g = \mathcal{I}_g^{(N)}$ and by $\mathcal{R}_g = \mathcal{R}_g^{(N)}$, resp., the set of occupied and uninfected, the infected and the unoccupied, resp., vertices in generation g. We set $S_g := |\mathcal{S}_g|$, $I_g := |\mathcal{I}_g|$ and $R_g := |\mathcal{R}_g|$. Furthermore $\overline{I}_g{}^{(N)} = \sum_{i=0}^{g} I_i{}^{(N)}$ is the number of hosts that got infected

till generation g. Let $\mathbf{I} = \mathbf{I}^{(N)} = (I_g{}^{(N)})_{g\geq 0}$ and $\overline{\mathbf{I}} = \overline{\mathbf{I}}^{(N)} = (\overline{I}_g{}^{(N)})_{g\geq 0}$ be the corresponding processes.

To state our main results about the infection process we need the definition of *branching processes with cooperation in discrete time.*

Definition 3.2.1 (Branching process with cooperation in discrete time)**.** *Let $\mathcal{L}_o$ and $\mathcal{L}_c$ be two probability distributions on $\mathbb{N}_0$. A discrete-time branching process with cooperation (DBPC, for short) $\mathbf{Z} = (Z_g)_{g\geq 0}$ with offspring distribution $\mathcal{L}_o$ and cooperation distribution $\mathcal{L}_c$ is recursively defined as follows. Assume $Z_0 = k$ a.s. for some $k \in \mathbb{N}$, then for any $g \geq 1$, Z_g is defined as*

$$Z_g := \sum_{i=1}^{Z_{g-1}} X_{g,i} + \sum_{i,j=1, i>j}^{Z_{g-1}} Y_{g,i,j},$$

where $(X_{g,i})_{g,i}$ and $(Y_{g,i,j})_{g,i,j}$ are sequences of independent and identically distributed random variables with $X_{g,i} \sim \mathcal{L}_o$ and $Y_{g,i,j} \sim \mathcal{L}_c$. We denote by $\overline{\mathbf{Z}} = (\overline{Z}_g)_{g\geq 0}$ the total size process, i.e.

$$\overline{Z}_g = \sum_{i=0}^{g} Z_i.$$

In the following we will denote the probability weights of the distributions $\mathcal{L}_o$ and $\mathcal{L}_c$, resp. by $(p_{k,o})_{k\in\mathbb{N}_0}$ and $(p_{k,c})_{k\in\mathbb{N}_0}$.

Remark 3.2.2. *Branching processes with cooperation have been mostly studied in continuous time in more general settings, like branching process with (pairwise) interactions, see e.g. [79], [80], [81], [75], [82] and [83]. In particular, in [80] formulas for extinction probabilities for the case of branching processes with cooperation have been determined.*

A central object for our results is a DBPC with Poisson offspring and cooperation distributions or rather its survival probability. Therefore, we fix in the next definition some notation for these processes.

Definition 3.2.3 (DBPC with Poisson offspring and cooperation distribution)**.** *Let $a > 0$. Denote by $\mathbf{P}^{(a)}$ a DBPC with offspring distribution $\mathcal{L}_o \sim Poi(a^2/2)$ and cooperation distribution $\mathcal{L}_a \sim Poi(a^2)$. Furthermore, we denote by $\pi(a)$ the survival probability of $\mathbf{P}^{(a)}$.*

Denote by

$$d_N := (2r_N)^n N,$$

which is the expected number of direct neighbours a vertex of $\mathcal{G}^{(N)}$ (with an asymptotically non-vanishing distance to the boundary of M) is connected to in dimension n. Furthermore, denote for all $u \in (0, 1]$

$$E_u^{(N)} := \left\{\exists g \in \mathbb{N}_0 : \overline{I}_g^{(N)} \geq u \cdot \#\mathcal{V}^{(N)}\right\}$$

the event that at least a proportion u of the host population dies during the infection process.

Our main result is the following theorem.

Theorem 3.2.4. *Consider the above defined sequence of infection processes $(\mathbf{I}^{(N)})_{N\in\mathbb{N}}$ on $[0, 1]^n$ for some $n \in \mathbb{N}$. Assume $r_N = \frac{1}{2}N^{(\beta-1)/n}$ for some $0 < \beta < 1$ and let $0 < u \leq 1$. Then it holds:*

1) Invasion probability

(i) If $v_N \in o\big(\sqrt{d_N}\big)$, *then* $\lim_{N\to\infty} \mathbb{P}\big(E_u^{(N)}\big) = 0$.

(ii) If $v_N \sim a\sqrt{d_N}$ *for some* $0 < a < \infty$, *then*

$$\pi\big(\tfrac{a}{\sqrt{2^n}}\big) \le \liminf_{N\to\infty} \mathbb{P}\big(E_u^{(N)}\big) \le \limsup_{N\to\infty} \mathbb{P}\big(E_u^{(N)}\big) \le \pi(a).$$

(iii) If $\sqrt{d_N} \in o(v_N)$, *then* $\lim_{N\to\infty} \mathbb{P}\big(E_u^{(N)}\big) = 1$.

2) Invasion time
Assume $v_N \sim a\sqrt{d_N}$ *for some* $0 < a < \infty$. *Denote by*

$$T^{(N)} := \inf\left\{ g \in \mathbb{N} \mid \bar{I}_g^{(N)} = \#\mathcal{V}^{(N)} \right\}.$$

Then

$$\mathbb{P}\left(\left\lfloor \frac{1}{2r_N} \right\rfloor \le T^{(N)} \le \left\lceil \frac{1}{2r_N} \right\rceil + \mathcal{O}\left(\kappa_N\right) \middle| T^{(N)} < \infty \right) \to 1 \text{ as } N \to \infty$$

with $\kappa_N = \max\left\{ \log(\log(N)), N^{\frac{(\beta/2)-1}{n}+\delta} r_N^{-2} \right\}$, *for any* $\delta > 0$.

Remark 3.2.5. • *In Theorem 3.2.4 1) (ii) we obtain bounds for both* $\liminf \mathbb{P}(E_u^{(N)})$ *and* $\limsup \mathbb{P}(E_u^{(N)})$. *We believe that the limit of* $\mathbb{P}(E_u^{(N)})$ *exists.*

Simulations suggest that the upper bound provides a good approximation of the actual invasion probability, see Section 3.2.4. An analysis of the initial phase of the epidemic, when infected parasites start to spread around the initially infected vertex, would be helpful to understand how close the upper bound is to the asymptotic invasion probability (that we believe to exist).

- *In the critical scaling* $v_N \sim a\sqrt{d_N}$ *with* $0 < a < \infty$, *one can show that the invasion probability is also lower bounded by the survival probability* $\varphi(a)$ *of a Galton-Watson process with a* $Poi(a^2/2)$ *offspring distribution. To arrive at this lower bound, it suffices to neglect cooperation of parasites generated by different infected hosts at the beginning of the infection process. In particular such a result allows for a comparison of the invasion probabilities of the infection processes defined on the random geometric graph and on the configuration model with* d_N *half-edges per vertex and* v_N *offspring parasites (which analysis is derived in [1]). It shows that having a finite-dimensional spatial structure of the population increases the invasion probability compared to scenario where the host population is structured by a configuration model leading to an infection graph with a locally tree like structure.*

- *In Lemma 3.3.2 below, we prove, that for any* $a > 0$ *the survival probability* $\pi(a)$ *of a DBPC* $\mathbf{P}^{(a)}$ *as defined in Definition 3.2.3 is strictly positive. Therefore, the invasion probability is in Case 1) (ii) of Theorem 3.2.4 for any* $a > 0$ *strictly positive.*

  This contrasts the situation studied in [1] where for $a^2/2 \le 1$ *the invasion probability is asymptotically 0 (for a host population structured according to a configuration model instead of a random geometric graph on* $[0,1]^n$.

- *We assume that parasites die, if they move to an empty vertex. Our results should remain valid, if e.g. parasites survive and move forward in the next generation.*

The proof of Theorem 3.2.4 is formally given in Section 3.5, at the end of this section we will give a sketch of the proof. The proof of the lower bound on the invasion probability is based on an asymptotic result on the invasion probability of an analogously defined infection process when the host population is not structured according to a random geometric graph on the cube, but according to a complete graph. This model mimics the spread of cooperative parasites in well-mixed host populations and is neither covered by the parameter regime considered in [1] nor by Theorem 3.2.4. Therefore the result is of interest on its own. We state it next.

We consider the same host-parasite dynamics as previously described except for the fact that we consider it on a complete graph with D_N vertices. To avoid confusion with the other model we denote the number of newly generated parasites in case of a successful host infection by V_N. Furthermore the corresponding infection process is denoted by $\mathcal{J}$. Analogously as before the total number of infected hosts in generation g is denoted by $J_g^{(N)}$ and the total number of infected hosts up until generation g is denoted by $\overline{J}_g^{(N)}$, where similarly as before $\mathbf{J} = \mathbf{J}^{(N)} = \left(J_g^{(N)}\right)_{g\geq 0}$ and $\overline{\mathbf{J}} = \overline{\mathbf{J}}^{(N)} = \left(\overline{J}_g^{(N)}\right)_{g\geq 0}$ are the corresponding processes. We are interested in the event $F_u^{(N)}$ that eventually a proportion u, for $u \in (0,1]$, of the host population gets infected, i.e.

$$F_u^{(N)} = \left\{\exists g \in \mathbb{N}_0 : \overline{J}_g^{(N)} \geq u \cdot D_N\right\}.$$

We show that the invasion probability is in the critical regime asymptotically equal to the survival probability of a branching process with cooperation.

Before we state the result we introduce the following notation. We write that $f \in \Theta(g)$ if f grows asymptotically as fast as g, i.e.

$$0 < \liminf_{x\to\infty} \frac{f(x)}{g(x)} \leq \limsup_{x\to\infty} \frac{f(x)}{g(x)} < \infty.$$

Theorem 3.2.6. *Assume $D_N \in \Theta\left(N^\beta\right)$ for some $0 < \beta < 1$ and let $0 < u \leq 1$. The following invasion regimes hold:*

(i) Assume $V_N \in o\left(\sqrt{D_N}\right)$. Then

$$\lim_{N\to\infty} \mathbb{P}\left(F_u^{(N)}\right) = 0.$$

(ii) Assume $V_N \sim a\sqrt{D_N}$ for some $0 < a < \infty$. Then the invasion probability of parasites satisfies

$$\lim_{N\to\infty} \mathbb{P}\left(F_u^{(N)}\right) = \pi(a) > 0.$$

(iii) Assume $\sqrt{D_N} \in o(V_N)$, then $\lim_{N\to\infty} \mathbb{P}\left(F_u^{(N)}\right) = 1$.

The proof of Theorem 3.2.6 is given in Section 3.4. Next we sketch the proofs of Theorem 3.2.4 and 3.2.6.

Hereinafter we will often use the following terminology. We call an infection a *CoSame* infection (for cooperation from the same edge), if a host gets infected by two parasites (originating from the same vertex) that moved along the same edge to the vertex on which the infected host is located on, and we call an infection a *CoDiff* infection (for cooperation from different edges), if a host gets infected by two parasites that moved along different edges to the vertex the infected host is located on.

3.2.2 Sketches of the proofs of the main results

Sketch of the proof of Theorem 3.2.6: Case (ii): To arrive at an upper bound on the invasion probability we couple whp the total number of currently infected and currently empty vertices $\overline{\mathbf{J}}$ from above with the total size of a DBPC $\overline{\mathbf{Z}}_u^{(N)}$ until $\overline{\mathbf{Z}}_u^{(N)}$ remains constant or reaches at least the level ℓ_N for a sequence ℓ_N with $\ell_N \to \infty$ sufficiently slowly, see Proposition 3.4.3. The probability that $\overline{\mathbf{Z}}_u^{(N)}$ reaches the level ℓ_N is asymptotically equal to $\pi(a)$, see Proposition 3.4.4, as the appproximating DBPC has asymptotically the survival probability $\pi(a)$. In case the level ℓ_N is reached we upper bound the probability, that afterwards also the remaining hosts get infected, by 1.

For the lower bound we couple whp $\overline{\mathbf{J}}$ from below with the total size of a DBPC $\overline{\mathbf{Z}}_\ell^{(N)}$ that has asymptotically the survival probability $\pi(a)$ of a DBPC $\mathbf{P}^{(a)}$ until $\overline{\mathbf{Z}}_\ell^{(N)}$ remains constant or reaches the level ℓ_N for some sequence ℓ_N with $\ell_N \in \Theta(N^\varepsilon)$ and $\varepsilon > 0$ small enough, see Proposition 3.4.6 and Proposition 3.4.7. As for the upper bound the probability to reach the level ℓ_N for $\overline{\mathbf{Z}}_\ell^{(N)}$ is asymptotically equal to $\pi(a)$. When $\overline{\mathbf{J}}$ reaches the level ℓ_N, we show that the total number of empty vertices grows in a constant number of generations to a level $N^{\frac{\beta}{2}+\delta}$ for some small $\delta > 0$ whp, see Lemma 3.4.12. After reaching this particular level the remaining hosts get infected whp in a single generation.

Case (i): We show that asymptotically with probability 1 the parasite population does not survive the first generation.

Case (iii): We show that we can whp couple from below $\overline{\mathbf{J}}$ with the total size process of a Galton-Watson process with approximately Poi$(a^2/2)$ offspring distribution until N^α hosts get infected or the parasite population dies out for any $0 < \alpha < \beta$ and any $a > 0$. This can be done by ignoring infection due to cooperation. By choosing $\alpha > \beta/2$ we can show that once the level N^α is reached by $\overline{\mathbf{J}}$ whp after one more generation the remaining hosts get infected. Since the probability to reach the level N^α for the considered Galton-Watson process is asymptotically equal to the survival probability φ_a of a Galton-Watson process with Poi$(a^2/2)$ offspring distribution and φ_a tends to 1 for $a \to \infty$ the result follows.

We proceed with a *sketch of the proof of Theorem 3.2.4*:

Before we start recall that d_N denotes the expected number of neighbours and v_N the offspring number of parasites generated at a host infection.

Claim 1) (ii) and Claim 2): For our upper bound on the survival probability we couple (as in the case of the complete graph) $\overline{\mathbf{I}}$ with the total size of a DBPC with offspring and cooperation distributions that are approximately Poisson distributions until a certain number ℓ_N of hosts get infected or the parasite population dies out, for a sequence $\ell_N \to \infty$ sufficiently slowly, see Proposition 3.5.7. The parameter of the approximating Poisson distribution for the offspring distribution is roughly $\frac{a^2}{2}$, since if all vertices are occupied with hosts the number of CoSame infections is on average approximately $\binom{v_N}{2}\frac{1}{d_N}$. The Poisson parameter of the cooperation distribution is roughly a^2, since cooperation is maximal if two balls centered around vertices, on which parasites have been generated in the same generation, are completely overlapping. In case of a complete overlap the number of cooperation events is roughly $v_N^2 \frac{1}{d_N} \sim a^2$ on average . Then we show that the probability to reach with the total size of the upper DBPC the level ℓ_N is asymptotically equal to the survival probability of the DBPC. This yields the upper bound, since again we upper bound the probability to infect the remaining hosts afterwards by 1.

For the lower bound we consider the spread of the parasites restricted to a certain *complete*

neighbourhood $\mathcal{C}(x_0)$ of the vertex x_0, that gets initially infected. The set $\mathcal{C}(x_0)$ contains all Poisson points with a distance $r_N/2$ to x_0. Since any two points in $\mathcal{C}(x_0)$ have a distance of at most r_N any two points are connected over an edge, in other words the restriction of $\mathcal{G}^{(N)}$ to points in $\mathcal{C}(x_0)$ is a complete graph. Consequently, also the infection process restricted to $\mathcal{C}(x_0)$ is an infection process on a complete graph. In particular, we can apply Theorem 3.2.6 to show that the probability to infected at least N^ε vertices can be asymptotically lower bounded by the survival probability $\pi\left(\frac{a}{\sqrt{2^n}}\right)$ of a DBPC $\mathbf{P}^{(a/\sqrt{2^n})}$. The parameter of the offspring distribution is roughly $\frac{a^2}{2^{n+1}}$, since $\binom{v_N}{2}$ pairs of parasites can be generated per infected host and the probability that both parasites of a pair of parasites hit the same vertex and that the vertex lies in $\mathcal{C}(x_0)$ is roughly $\frac{1}{2^n}\frac{1}{d_N}$, because the first parasite has to attack a vertex in the complete neighbourhood, which happens with probability approximately $\left(\left(2\frac{r_N}{2}\right)^n N\right)/\left((2r_N)^n N\right) = 1/2^n$, and the second parasite has to attack the same vertex. So the number of CoSame infections per host is roughly $\binom{v_N}{2}\frac{1}{2^n}\frac{1}{d_N} \to \frac{a^2}{2^{n+1}}$. Similarly the parameter for the cooperation distribution is $v_N^2\frac{1}{2^n}\frac{1}{d_N} \sim \frac{a^2}{2^n}$. If the infection process is successful N^ε many hosts are infected within at most $\mathcal{O}(\log(\log(N)))$ generations. Afterwards in a finite number of generations in the complete neighbourhood the remaining hosts get infected and the infection process expands at a distance $r_N(1-o(1))$ per generation, see also Figure 3.6 and Section 3.5.2 for more details. On the other hand the invasion time is lower bounded by $1/(2r_N)$, since parasites can move in any generation at most at a distance r_N and the infection starts in the center of the cube. This explains our Claim 2) on the invasion time.
Case (i): As in the case of the complete graph we show that asymptotically with probability 1 the parasite population does not survive the first generation.
Case (iii): Again as in the case of the complete graph we show that we can whp couple from below $\bar{\mathbf{I}}$ with the total size process of a Galton-Watson process with an approximately $\mathrm{Poi}(a^2/2)$ offspring distribution until $N^{\beta'}$ hosts get infected or the parasite population dies out for any $0 < \beta' < \beta$ and any $a > 0$. In addition we can show that when the level $N^{\beta'}$ is reached there exists a ball of radius r_N which contains at least $N^{\beta'}/\log(N)$ infected hosts. By choosing $\beta' > \beta/2$ we can show that once the level $N^{\beta'}$ is reached whp after one more generation the remaining hosts in this ball get infected. Afterwards the infection expands by a distance $r_N(1-o(1))$ in every generation whp (similar as in Case (ii)) until the remaining hosts are all infected. Since the probability to reach the level $N^{\beta'}$ for the total size of the considered Galton-Watson process is asymptotically equal to the survival probability φ_a of a Galton-Watson process with $\mathrm{Poi}(a^2/2)$ offspring distribution and φ_a tends to 1 for $a \to \infty$ the result follows.

3.2.3 Invasion on Riemannian manifolds

In this section we give a brief heuristic justification why our results should also carry over to a setting where the unit cube $[0,1]^n$ is replaced by a Riemannian manifold. Instead of considering the spread of the parasite population in host populations structured according to a random geometric graph on an n-dimensional cube it is natural to assume that the host population is located on a manifold. We can generalize our model to this setting as follows. Let (M', g) be a compact, connected, orientable, n-dimensional Riemannian manifold with Riemannian metric g. Assume without loss of generality that $vol(M') = 1$, where $vol(M')$ denotes the volume of M' calculated according to the volume induced by g. Denote furthermore by ρ' the metric on M' induced by g. Consider a homogeneous Poisson point process with intensity measure $N \cdot vol(\,\cdot\,)$ on M' (for this point process the number of vertices contained in a set $S \subset M'$ with volume $vol(S) = s$ is Poisson distributed

with parameter sN). We denote the set of the Poisson points by $\mathcal{V}' = \mathcal{V}'^{(N)}$ and build a random geometric graph on M' by connecting all points in $\mathcal{V}'$ over an edge which have a distance of at most r_N with respect to the metric ρ'. Denote the set of edges by $\mathcal{E}' = \mathcal{E}'^{(N)}$ and the random geometric graph by $\mathcal{G}' = \mathcal{G}'^{(N)} = \mathcal{G}'^{(r_N)} = (\mathcal{V}'^{(N)}, \mathcal{E}'^{(N)})$.

Given $\mathcal{G}'$ we can consider an infection process (with the components $(\mathcal{S}'_g, \mathcal{I}'_g, \mathcal{R}'_g)$) in the same way in which we defined it on the random geometric graph on the cube.

Denote by

$$d'_N := \frac{\pi^{n/2}}{\Gamma(n/2+1)} (r_N)^n N,$$

which is the expected number of vertices a vertex of $\mathcal{G}^{(N)}$ is connected to in dimension n (if the distance of the vertex to the boundary of M' is asymptotically non-vanishing, in case M' has a boundary) and let $p \in M'$. Denote by

$$\Delta(p) := \max_{q \in M} \{\rho'(q,p)\}$$

the maximal distance between p and any other point $q \in M$. Furthermore, denote as before by

$$E'^{(N)}_u := \left\{ \exists g \in \mathbb{N}_0 : \overline{I}'^{(N)}_g \geq u \cdot \#\mathcal{V}'^{(N)} \right\}.$$

Then we believe that the following statements hold at least for $n \in \{1, 2\}$.

Assume $r_N = \left(\frac{\Gamma(n/2+1)}{\pi^{n/2}}\right)^{1/n} N^{\frac{\beta-1}{n}}$ for some $0 < \beta < 1$, let $0 < u \leq 1$. Assume the infection process is started in a vertex $x_0^{(N)} \in \mathcal{V}'^{(N)}$ that has asymptotically a positive distance to the boundary of M' (if M' has a boundary).

1) Invasion probability

   (i) If $v_N \in o\left(\sqrt{d'_N}\right)$, then
   $$\lim_{N\to\infty} \mathbb{P}\left(E'^{(N)}_u\right) = 0.$$

   (ii) If $v_N \sim a\sqrt{d'_N}$ for some $0 < a < \infty$, then
   $$\pi\left(\tfrac{a}{\sqrt{2^n}}\right) \leq \liminf_{N\to\infty} \mathbb{P}\left(E'^{(N)}_u\right) \leq \limsup_{N\to\infty} \mathbb{P}\left(E'^{(N)}_u\right) \leq \pi(a).$$

   (iii) If $\sqrt{d'_N} \in o(v_N)$, then $\lim_{N\to\infty} \mathbb{P}\left(E'^{(N)}_u\right) = 1$.

2) Invasion time: Assume $v_N \sim a\sqrt{d'_N}$ for some $0 < a < \infty$. Denote by
   $$T'^{(N)} := \inf\left\{ g \in \mathbb{N} | \overline{I}^{(N)}_g = \#\mathcal{V}'^{(N)} \right\}.$$
   Then
   $$\mathbb{P}\left(\left\lfloor \Delta\left(x_0^{(N)}\right) r_N^{-1} \right\rfloor \leq T'^{(N)} \leq \left\lceil \Delta\left(x_0^{(N)}\right) r_N^{-1} \right\rceil + \mathcal{O}(\kappa_N) \,\middle|\, T'^{(N)} < \infty \right) \underset{N\to\infty}{\to} 1,$$
   with $\kappa_N = \max\left\{ \log(\log(N)), N^{\frac{\frac{\beta}{2}-1}{n}+\delta} r_N^{-2} \right\}$ for any $\delta > 0$.

The main reason why these results should hold is that the decision, if eventually invasion takes place, is made in a neighbourhood of x_0 that is asymptotically of negligible volume. Indeed, only N^ε many hosts need to be infected such that whp subsequently the whole host population gets infected. Since N^ε many hosts are directly connected to x_0 for $\varepsilon > 0$ small enough, we need to consider only a neighbourhood of x_0 of negligible volume to decide on invasion. As a consequence at the beginning the invasion process is essentially the same as a corresponding process on $[0,1]^n$ with distances measured according to the Euclidean distance due to the following relationship between the volume of an Euclidean ball in $[0,1]^n$ and the volume of a ball of the manifold. For any sequence $h_N \to 0$ it holds

$$\frac{vol(B_{h_N}(x))}{vol(\tilde{B}_{h_N}(0))} = 1 - \frac{S}{6(n+2)} h_N^2 + o(h_N^2),$$

where $vol(B_{h_N}(x))$ denotes the volume of a (geodesic) ball of radius h_N centered in $x \in M'$ and $vol(\tilde{B}_{h_N}(0))$ denotes the volume of an n-dimensional Euclidean ball of radius h_N centered in 0 and S the scalar curvature in x, see [84], Section XII.8. Since M' is compact and the scalar curvature is a continuous function on M', the scalar curvature of M' is bounded from above and below. In particular, for $h_N = r_N$ the number of points connected to $x \in \mathcal{V}'$ is Poisson distributed with parameter $N\frac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)} r_N^n + O(N r_N^{n+2})$, since

$$vol(\tilde{B}_{r_N}(0)) = \frac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)} r_N^n.$$

In Theorem 3.2.4 we consider the maximum metric to measure distances between two points. With this metric we easily can cover M with balls (that are cubes as well) to control the spread of parasites across M. A similar construction is also possible with Euclidean balls (at least in the case $n \in \{1,2\}$), the notation is just a bit more complicated. Therefore, considering the Euclidean metric or maximum metric should not influence the invasion probability as long as the ratio of the expected number of vertices contained in a ball and the number of offspring parasites generated at infection is asymptotically the same. The invasion time in general differs for two different metrics, because the function $\Delta(p)$ depends on the metric.

3.2.4 Simulating spatial invasion of cooperative parasites

We supplement our findings with simulation results for moderately sized, finite N. We simulated invasion of parasites in host populations structured according to random geometric graphs on (i) the interval $[0,1]$ with the euclidean metric (which agrees with the maximum metric, since $n = 1$), (ii) the square $[0,1]^2$ using the maximum metric and (iii) the unit 2-sphere S^2 using spherical distances (to substantiate our conjecture given in Remark 3.2.3 at least by means of simulations).

To ease computations in the case of the 2-sphere, we generate points on the unit 2-sphere S^2, instead of the sphere with radius $1/\sqrt{4\pi}$ which would have as required in Remark 3.2.3 a surface area (aka volume) of 1. This simplification benefits both generation and evaluation of point distances in our implementation of the process and only requires appropriate rescaling. The distance between two points x and y is then simply given by $\arccos(x \cdot y)$ as the radius is of length 1. Uniform points on S^2 can be generated by a two-step scheme in which first the polar angles (θ, ϕ) are sampled using inverse transform sampling. To this end, let $U_1, U_2 \sim U(0,1)$. We compute $\theta = 2\pi \cdot U_1$ as well as $\phi = \arccos(1 - 2 \cdot U_2)$ and obtain Cartesian coordinates by a standard transformation.

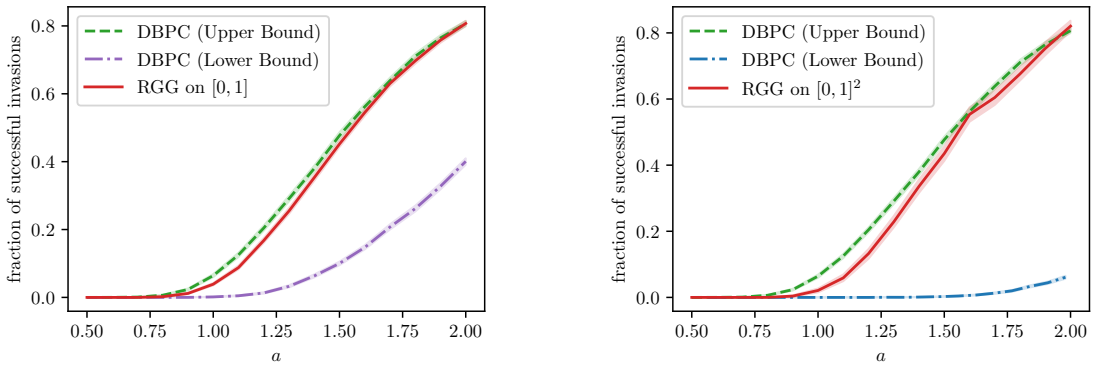

Figure 3.1: Simulated invasion probabilities with a host population structured by a random geometric graph (RGG, for short) on $[0,1]$ (top) and $[0,1]^2$ (bottom) for $N = 10^6$ and $\beta = 0.7$ as well as simulated survival probabilities $\pi\left(\frac{a}{\sqrt{2^n}}\right)$ and $\pi(a)$.

In general, storing and operating on an explicit representation of $\mathcal{G}$ takes space in the order of $|\mathcal{E}^{(N)}|$. In the case when $Nd_N \sim N^{1+\beta}$ gets prohibitively large rendering parameter combinations of N and $d_N = N^\beta$ is infeasible for general-purpose compute architectures. Optimizations, however, are possible by implicit representations of $\mathcal{G}$ using the coordinates of $\mathcal{V}^{(N)}$. Realizations of this are straight-forward for the interval $[0,1]$ and can be adapted using Quadtrees for 2-dimensional spaces [85].

Invasion probabilities

In Theorem Theorem 3.2.4 we claim that for $0 < a < \infty$ and $v_N \sim a\sqrt{d_N}$

$$\pi\left(\tfrac{a}{\sqrt{2^n}}\right) \leq \liminf_{N\to\infty} \mathbb{P}\left(E_u^{(N)}\right) \leq \limsup_{N\to\infty} \mathbb{P}\left(E_u^{(N)}\right) \leq \pi(a).$$

In Figure 3.1 and Figure 3.2 simulation results are depicted that show the fraction of cases, in which the host population got completely infected, for parasites spreading in host populations structured according to random geometric graphs on the interval $[0,1]$, the square $[0,1]^2$ and the sphere S^2. The simulated survival probabilities $\pi\left(a/\sqrt{2^n}\right)$ and $\pi(a)$ (where in the simulations we assume that survival took place if the DBPC attains size N) appear to be appropriate upper and lower bounds. The upper bound gives a particularly good approximation to the observed actual invasion probability. For the upper bound one assumes that the chance for two parasites, which have been generated on different vertices, to cooperate is roughly $1/d_N$, which is actually only true if the distance of the two vertices is 0. Therefore it might be surprising that the upper bound gives such a good fit. However, since parasites perform symmetric random walks a large part of parasites stays in a neighbourhood of x_0 and parasites that are close together have due to CoDiff infections a higher chance to generate offspring, which implies that parasites located in densely populated regions have in general more offspring parasites than parasites located in sparsely populated regions. This effect remains until a significant proportion of the host population in a r_N-ball gets infected, but at this time point invasion is essentially already decided. Consequently, the probability that a typical pair

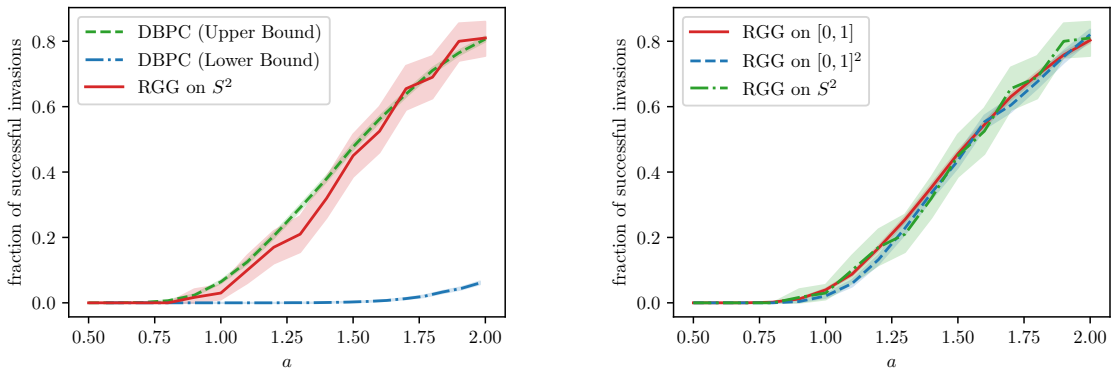

Figure 3.2: Fraction of successful invasions for $N = 10^6$ and $\beta = 0.7$ of infection processes spreading on host populations structured by a RGG on S^2 (top) and for comparison on $[0,1]$, $[0,1]^2$ and S^2 (bottom).

of parasites produces CoDiff infections could be in the initial phase pretty close to $1/d_N$.

Conversely, our lower bound is relatively far away from the simulated invasion probabilities. To arrive at the lower bound we consider only parasites spreading within the complete neighbourhood of the initially infected host. The larger n the smaller this complete neighbourhood is in comparison to the r_N- neighbourhood of a host. Therefore the larger n the worse is the lower bound.

Our asymptotic upper bound of the invasion probability does only depend on the ratio of the number v_N of parasites generated on a vertex and the (asymptotic) number of direct neighbours of a typical vertex, but neither depends on the dimension nor (in the setting considered in Remark 3.2.3) on the curvature of the manifold. We suppose that this is also the case for the invasion probabilities. In Figure 3.2 we present a direct comparison of simulated invasion probabilities for infection processes on $[0,1]$, $[0,1]^2$ and S^2 and see that the probabilities are very close to each other (even for finite N).

Finally we simulated invasion probabilities of the infection processes on the complete graphs that we use for a coupling from below. In Figure 3.3 one can observe that the simulated invasion probabilities match very well with the probabilities $\pi\left(a/\sqrt{2^n}\right)$ and $\pi(a)$ of the corresponding DBPCs.

Invasion time

In Figure 3.4 and Figure 3.5 we present the invasion time of simulated infection processes on the interval $[0,1]$, square $[0,1]^2$ and sphere S^2, respectively. For reference we plot also the asymptotic order of the invasion times derived in Theorem 3.2.4 and Remark 3.2.3. In Figure 3.4 we observe a matching overlap that improves for increasing N for all considered values of a in the 1-dimensional case.

For large β values the simulations showcase a higher invasion time than predicted. This can be explained as follows: We show in Theorem 3.2.4 that the invasion time is asymptotically proportional to $N^{1-\beta}$. In particular, the larger β is the shorter is the invasion time. For $N \to \infty$

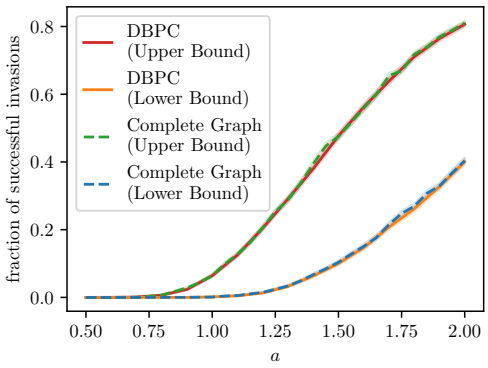

Figure 3.3: Simulated invasion probabilities with a host population structured by a complete graph for $N = 10^6$ and $\beta = 0.7$ as well as simulated survival probabilities $\pi\left(\frac{a}{\sqrt{2}}\right)$ and $\pi(a)$.

invasion is dominated by the time necessary to reach from an infinitesimally small neighbourhood of x_0 points close to the boundary of $[0,1]^n$ or in the setting of Remark 3.2.3 the point that has the largest distance to the host that got initially infected. The initial phase until for the first time all direct neighbours of a vertex get infected is only of order $\log\log(N)$. For β close to 1 and finite N however both time frames are of approximately the same length, which explains the deviation from the theoretical prediction where the initial phase is ignored. In Figure 3.5 we plotted the invasion time when the initial phase is removed. One observes that for intermediate and larger values of β the gap between the predicted and simulated invasion time disappears. For larger values of β the simulated invasion times lie slightly below the predicted invasion times. Probably this is caused by parasites spreading the infection further before the initial phase is over.

Also for small β values we observe that simulated invasion times are generally higher than the predicted ones, even when the initial phases are removed. This deviation is particularly pronounced for invasion on $[0,1]^2$, where the maximum metric is used. This can be explained as follows. As we pointed out in the sketch of the proof of Theorem 3.2.4 the parasite population expands furthest due to parasites born at the boundary of an r_N neighbourhood. When on the square the maximum distance is used r_N-squares on the diagonal can get infected fastest by parasites at the corners of neighbouring r_N-squares. However, when N is not large, the number of parasites located in the corners is pretty small, so that they might not be able to move the front forward as quickly as predicted for $N \to \infty$.

This behaviour is further studied in Figure 3.6 where the progress of the infection process is tracked along balls of radius $r_N/2$ on the unit-square $[0,1]^2$ for different values of β. The smaller β the less vertices are located at the corners leading to a decreased speed of expansion. As shown in Figure 3.6, for $\beta = 0.7$ and $\beta = 0.5$ one observes that after the initial phase the parasite population expands linearly (almost) by a factor 1 (as predicted), while for $\beta = 0.3$ (when in each box with edge length r_N only $\approx N^{0.3} \approx 63$ vertices are contained) the population expands also linearly, but only by a factor of (almost) 2.

In the following the manuscript is structured as follows. In Section 3.3 we show several properties (of sequences) of DBPCs that we will need in the subsequent section. Afterwards in Section 3.4 we

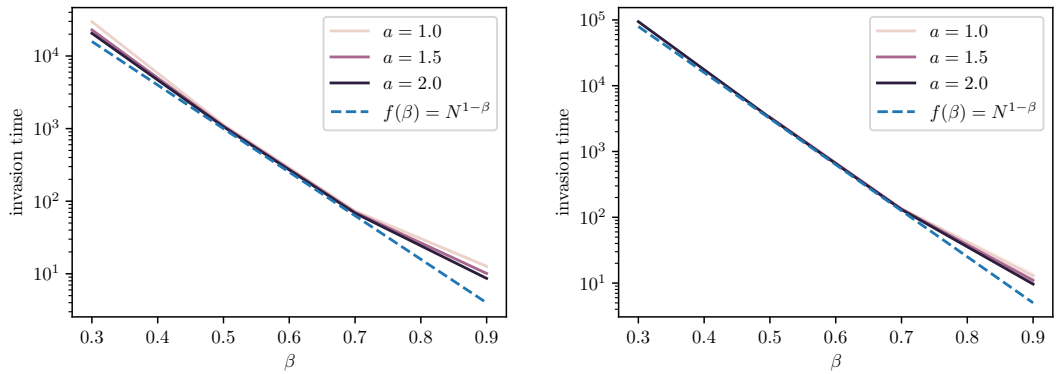

Figure 3.4: Invasion time on random geometric graphs on $[0,1]$ for $N = 10^6$ (top) and $N = 10^7$ (bottom) with varying a and β.

will prove Theorem 3.2.6. Finally in the last section we will prepare and give the proof of Theorem 3.2.4.

3.3 Discrete branching processes with cooperation

In this section we collect properties of (sequences of) DBPCs that we need in the following. For Galton-Watson processes some of these statements are well-known or have been proven in [1]. Since the proof techniques are similar, we do not give all of the proofs in the main text, but provide some of them in the supplementary material.

We start with the extinction-explosion principle, which is well-known for branching processes and also holds for DBPCs.

Lemma 3.3.1. *(Extinction-explosion principle for DBPCs) Let* $\mathbf{Z}$ *be a DBPC satisfying* $p_{1,o} \neq 1$ *and* $(p_{0,o}, p_{1,c}) \neq (1,1)$*. Then there exists a nullset* $\mathcal{N}$ *such that*

$$\{\forall g \in \mathbb{N}_0 : Z_g > 0\} \subseteq \{\forall i \in \mathbb{N}, \exists g_0 \in \mathbb{N}_0, \forall g \geq g_0 : Z_g \geq i\} \cup \mathcal{N}.$$

For the proof one shows that all states but 0 are transient states. Details of the proof can be found in the supplementary material.

Note that the conditions $p_{1,o} \neq 1$ and $(p_{0,o}, p_{1,c}) \neq (1,1)$ are necessary to exclude two pathological cases where the statement does obviously not hold. The first case is where the process stays constant at 1 individual in each generation, and the second condition ensures that the DBPC with 3 and less individuals can further increase and is not stuck below 3.

Before we proceed we introduce some useful notation. We denote the expectation and the variance of the offspring distribution by μ_o and ν_o and for the cooperation distribution by μ_c and ν_c.

In contrast to a Galton-Watson process we aim to show that except for pathological cases a DBPC always has a positive survival probability.

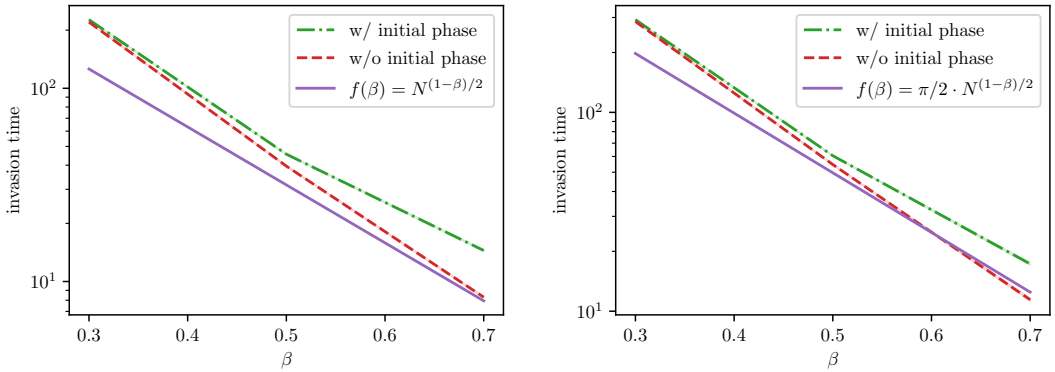

Figure 3.5: Invasion time on random geometric graphs on $[0,1]^2$ (top) and S^2 (bottom) for $N = 10^6$ with $a = 2.0$.

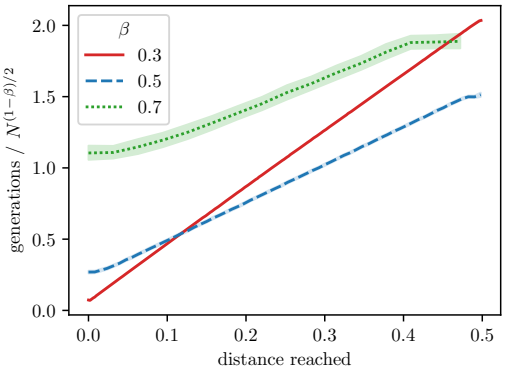

Figure 3.6: Distance reached for successful invasions on $[0,1]^2$ with $N = 10^6$.

Lemma 3.3.2. *Let* $\mathbf{Z}$ *be a DBPC with* $\mu_o, \mu_c, \nu_o, \nu_c \in (0,\infty)$ *and* $p_{0,o} + p_{1,o} < 1$. *Suppose* $Z_0 = x \in \mathbb{N}$, *then* $\mathbf{Z}$ *has a positive survival probability, i.e.*

$$\mathbb{P}_x(Z_g > 0 \,\forall\, g > 0) > 0.$$

The proof of Lemma 3.3.2 is based on the next lemma, which we immediately formulate in a more general setting to be able to apply it also later in another context and which basically states that if a DBPC attains a certain level, then in subsequent generations the size will up to a constant factor (that does not depend on the size) be squared in the next generations due to cooperation with a certain non-vanishing probability.

In the subsequent sections we will deal often with sequences of DBPCs rather than a single process. We will often need the following assumption to be fulfilled.

Assumption 3.3.3. *Let* $\left(\mathbf{Z}^{(N)}\right)_{N\in\mathbb{N}_0}$ *be a sequence of DBPCs for which* ${\mu_o}^{(N)}, {\mu_c}^{(N)}, {\nu_o}^{(N)}, {\nu_c}^{(N)}$ *denote the expectations and the variances of the offspring and cooperation distributions. We assume*

that

$$\mu_o^{(N)}, \mu_c^{(N)}, \nu_o^{(N)}, \nu_c^{(N)} \underset{N\to\infty}{\longrightarrow} \mu_o, \mu_c, \nu_o, \nu_c \in (0,\infty),$$

and we introduce

$$\mu := \min(\mu_o, \mu_c) > 0, \text{ and } \nu := \max(\nu_o, \nu_c) > 0.$$

Lemma 3.3.4. *Let $(\mathbf{Z}^{(N)})_{N\in\mathbb{N}_0}$ be a sequence of DBPCs which satisfies Assumption 3.3.3. Furthermore, set $f_i(k) := \frac{k^{2^i}\mu^{2^i-1}}{8^{2^i-1}}$ for $i \geq 1$ and $k \in \mathbb{N}$. Then there exists $N_0 \in \mathbb{N}$ such that for any $N \geq N_0$, $g \in \mathbb{N}$, $M \in \mathbb{N}$ and $k \geq L := \lceil \mu^{-1}(8+\nu)^2 \rceil$ it holds that*

$$\mathbb{P}\left(\bigcap_{i=1}^{M}\left\{Z_{g+i}^{(N)} > f_i(k)\right\} \;\middle|\; Z_g^{(N)} = k\right) \geq \prod_{i=1}^{M}\left(1 - \frac{6\nu}{f_i(k)\mu}\right) \geq \prod_{i=1}^{M}\left(1 - \frac{3}{4(8+\nu)^{2^i-1}}\right).$$

The claim can be shown by several applications of Tchebychev's inequality. A detailed proof can be found in the supplementary material.

Now we can prove Lemma 3.3.2.

Proof of Lemma 3.3.2. The assumption $p_{0,o} + p_{1,o} < 1$ implies that there is a strictly positive probability that $Z_{g+1} > Z_g$ for all initial values $x > 0$. Let us consider $x \in \{1, ..., L-1\}$ first. Since $L = \lceil \mu^{-1}(8+\nu))^2 \rceil$ is finite the previous observation implies that there exists a $p_1 > 0$ such that for any x it exists a $g_0 \in \mathbb{N}$ such that

$$\mathbb{P}_x\big(Z_{g_0} \geq L\big) \geq p_1.$$

If $x \geq L$ obviously we can set $g_0 = 0$. Applying Lemma 3.3.4 for $\mathbf{Z}^{(N)} \equiv \mathbf{Z}$ one obtains

$$\mathbb{P}\left(\bigcap_{i=1}^{M}\{Z_{g_0+i} > f_i(L)\} \;\middle|\; Z_{g_0} = L\right) \geq \prod_{i=1}^{M}\left(1 - \frac{3}{4(8+\nu)^{2^i-1}}\right).$$

Noticing that the events appearing on the left-hand side are monotone-decreasing for growing M it follows from upper σ-continuity of the measure $\mathbb{P}$ that we have

$$\mathbb{P}\left(\bigcap_{i=1}^{\infty}\{Z_{g_0+i} > f_i(L)\} \;\middle|\; Z_{g_0} = L\right) \geq \prod_{i=1}^{\infty}\left(1 - \frac{3}{4(8+\nu)^{2^i-1}}\right) > 0,$$

where it follows that the right-hand side is strictly positive by comparison with a geometric sum. We have that

$$\bigcap_{i=1}^{\infty}\{Z_{g_0+i} > f_i(L)\} \cap \{Z_{g_0} \geq L\} \subset \{Z_g > 0 \;\forall g \geq 0\},$$

then by Markov property and monotonicity we get

$$\mathbb{P}_x\left(Z_g > 0 \,\forall g \geq 0\right) \geq \mathbb{P}(Z_{g_0+i} > 0 \,\forall i \geq 0 \mid Z_{g_0} = L)\mathbb{P}_x(Z_{g_0} \geq L) > 0.$$

□

The next lemma claims that a DBPC also survives whp when a level b_N, that tends to ∞ (arbitrarily slowly) with N, has been reached.

Lemma 3.3.5. *Let $(\mathbf{Z}^{(N)})_{N\in\mathbb{N}_0}$ be a sequence of DBPCs which satisfy Assumption 3.3.3 and let $(b_N)_{N\in\mathbb{N}}$ be an $\mathbb{N}$-valued sequence with $b_N \to \infty$ as $N \to \infty$. Then for all $g \in \mathbb{N}$ it holds that*

$$\mathbb{P}\left(Z_{g+i}^{(N)} > 0 \ \forall i \geq 0 \mid Z_g^{(N)} \geq b_N\right) \underset{N\to\infty}{\longrightarrow} 1.$$

The proof relies on Lemma 3.3.4, we provide a detailed proof in the supplementary material.

Finally we are able to derive results on the growth speed of the population size in case of survival. The next lemma shows that for any sequence $b_N \to \infty$, reaching the level b_N or dying out is at most of order $\log\log(b_N)$.

Proposition 3.3.6. *Let $(\mathbf{Z}^{(N)})_{N\in\mathbb{N}_0}$ be a sequence of DBPCs which satisfy Assumption 3.3.3. Assume ${Z_0}^{(N)} = x$ for some $x \in \mathbb{N}$. Furthermore, assume that there exists an $N_0 > 0$ such that*

$$\inf_{N\geq N_0} \mathbb{P}\left(Z_1^{(N)} > L | Z_0^{(N)} = x\right) > 0, \tag{3.3.1}$$

where $L =: \lceil \mu^{-1}(8+\nu)^2 \rceil$. Let $(b_N)_{N\in\mathbb{N}}$ be a $\mathbb{N}$-valued sequence with $b_N \to \infty$ as $N \to \infty$ and denote by

$$\tau_{b_N,0}^{\mathbf{Z}^{(N)}} := \inf\left\{g \in \mathbb{N} : Z_g^{(N)} \geq b_N \text{ or } Z_g^{(N)} = 0\right\}.$$

Then there exists a constant $C > 0$ such that

$$\mathbb{P}_x\left(\tau_{b_N,0}^{\mathbf{Z}^{(N)}} \leq C\log\log(b_N)\right) \to 1 \quad \text{as } N \to \infty.$$

Proof. Set $e_N := \log(b_N)$. Note that we consider N large enough such that $e_N \geq L$. We will first show that there exist $p_0 > 0, C, N_1 \in \mathbb{N}$ independent of N, such that for all $N \geq N_1$ and all $x \in \mathbb{N}$

$$\mathbb{P}_x\left(\tau_{e_N,0}^{\mathbf{Z}^{(N)}} \leq C\log\log(e_N)\right) \geq p_0. \tag{3.3.2}$$

From this follows that for all $m \in \mathbb{N}$

$$\mathbb{P}_x\left(\tau_{e_N,0}^{\mathbf{Z}^{(N)}} \leq mC\log\log(e_N)\right) \geq \sum_{i=1}^{m}(1-p_0)^{i-1}p_0.$$

In particular for $m = g_N := \frac{\log\log(b_N)}{\log\log(e_N)}$ it follows that

$$\mathbb{P}_x\left(\tau_{e_N,0}^{\mathbf{Z}^{(N)}} \leq Cg_N\log\log(e_N)\right) \geq 1-(1-p_0)^{\lfloor g_N\rfloor} \to 1, \tag{3.3.3}$$

where we used that by choice of e_N it follows that $g_N \to \infty$ as $N \to \infty$. In fact we will see that $C := \log(2)^{-1}$ is a suitable choice. Now we will show (3.3.2). By assumption (3.3.1) and the fact that L does not depend on N there exists a $p_1 > 0$ such that for any $x \in \{1, ..., L-1\}$

$$\mathbb{P}\left(Z_g^{(N)} = 0 \text{ or } Z_g^{(N)} > L \mid Z_{g-1}^{(N)} = x\right) \geq p_1.$$

Consequently, for any $x \in \{1, ..., L-1\}$ we can lower bound the time to reach the state 0 or a state $> L$ by a geometrically distributed random variable with success probability p_1. Reasoning as above shows that the waiting time to hit 0 or a state $> L$ is with probability $1 - o(1)$ bounded by $\log\log(e_N)$.

Before we continue we briefly state two facts we need. Recall that $f_i(k) = \frac{k^{2^i}\mu^{2^i-1}}{8^{2^i-1}}$ for $i \geq 0$, where $f_0(k) = k$. First we see that we have the relation

$$\frac{f_{i-1}^2(k)\mu}{8} = \left(\frac{k^{2^{i-1}}\mu^{2^{i-1}-1}}{8^{2^{i-1}-1}}\right)^2 \frac{\mu}{8} = \frac{k^{2^i}\mu^{2^i-1}}{8^{2^i-1}} = f_i(k). \tag{3.3.4}$$

Furthermore, we assumed that $k > \mu^{-1}(8+\nu)^2$, and therefore

$$f_i(k) > \frac{(8+\nu)^{2^{i+1}}\mu^{2^i-1}}{\mu^{2^i}8^{2^i-1}} = \frac{(8+\nu)^{2^i}(8+\nu)^{2^i}}{\mu 8^{2^i-1}} > \frac{8(8+\nu)^{2^i}}{\mu}. \tag{3.3.5}$$

We show next, that if $Z_g^{(N)} \geq L$ after $C\log\log(e_N)$ further generations the level e_N will be reached with some probability $p_2 > 0$ for any N large enough. Lemma 3.3.4 implies that for $k \geq L$ and $M = C\log\log(e_N)$ it follows that

$$\mathbb{P}\left(\bigcap_{i=1}^{C\log\log(e_N)} \left\{Z_{g+i}^{(N)} > f_i(k)\right\} \,\middle|\, Z_g^{(N)} = k\right) \geq \prod_{i=1}^{C\log\log(e_N)} \left(1 - \frac{3}{4(8+\nu)^{2^i-1}}\right) > p_2 > 0.$$

Now by (3.3.5) we know that $f_i(k) > \mu^{-1}8(8+\nu)^{2^i}$. Thus, by our choice $C = (\log(2))^{-1}$ we have that

$$f_{C\log\log(e_N)}(k) \geq \frac{8}{\mu}(e_N)^{\log(8)}.$$

Since $\log(8) > 1$ we get that $f_{C\log\log(e_N)}(k) \geq e_N$ for N large enough, i.e. this means that we in fact reach a level higher than e_N. Thus, we can conclude (3.3.2).

We again apply Lemma 3.3.4 such that we get that

$$\mathbb{P}\left(\bigcap_{i=1}^{C\log\log(b_N)} \left\{Z_{g+i}^{(N)} > f_i(e_N)\right\} \,\middle|\, Z_g^{(N)} = e_N\right) \geq \prod_{i=1}^{C\log\log(b_N)} \left(1 - \frac{6\nu}{f_i(e_N)\mu^2}\right).$$

By an analogous calculation as before we see that

$$f_{C\log\log(b_N)}(e_N) = \frac{8}{\mu} b_N^{\log\left(\frac{e_N\mu}{8}\right)}.$$

We know that $\log\left(\frac{e_N\mu}{8}\right) > 1$ since we choose N large enough such that $e_N \geq L$, and thus we have that

$$f_{C\log(\log(b_N))}(e_N) \geq b_N.$$

Now we use that $1 - x > \exp(-2x)$ for $x < 1$ and that $f_i(e_N) \geq e_N$ for all i, which again follows by the fact that $e_N \geq L$. These two observations allow us to conclude for N large enough that

$$\prod_{i=1}^{C\log\log(b_N)} \left(1 - \frac{6\nu}{f_{i-1}^2(e_N)\mu^2}\right) \geq \exp\left(-12\nu C\frac{\log\log(b_N)}{\log(b_N)\mu}\right) \underset{N\to\infty}{\to} 1. \tag{3.3.6}$$

Finally using the strong Markov property at the stopping time $\tau^{\mathbf{Z}^{(N)}}_{e_N,0}$ gives that

$$\mathbb{P}\big(\tau^{\mathbf{Z}^{(N)}}_{b_N,0} \le 2C\log\log(b_N)\big) \ge \mathbb{P}\big(\tau^{\mathbf{Z}^{(N)}}_{e_N,0} \le C\log\log(b_N), Z^{(N)}_{\tau^{\mathbf{Z}^{(N)}}_{e_N,0}} = 0\big)$$
$$+ \mathbb{P}\left(\tau^{\mathbf{Z}^{(N)}}_{e_N,0} \le C\log\log(b_N), Z^{(N)}_{\tau^{\mathbf{Z}^{(N)}}_{e_N,0}} \ge e_N\right) \cdot \mathbb{P}\left(\tau^{\mathbf{Z}^{(N)}}_{b_N,0} \le C\log\log(b_N) \;\middle|\; Z^{(N)}_0 = e_N\right).$$

Now the right-hand side converges to 1 as $N \to \infty$ according to (3.3.3) and (3.3.6). This concludes the proof of the claim. □

The following lemma states that the probability of reaching an arbitrary high level, that tends to ∞ as $N \to \infty$, at some generation or up to some generation is asymptotically equal to the survival probability for a sequence of DBPCs.

Proposition 3.3.7. *Consider a sequence of DBPCs $(\mathbf{Z}^{(N)})_{N\in\mathbb{N}}$ with offspring and cooperation distributions $({p_{k,o}}^{(N)})_{k\in\mathbb{N}_0}$ and $({p_{k,c}}^{(N)})_{k\in\mathbb{N}_0}$ respectively, which satisfies Assumption 3.3.3. Furthermore, let $\mathbf{Z}^\infty$ be a DBPC with offspring and cooperation distribution $(p_{k,o})_{k\in\mathbb{N}_0}$ and $(p_{k,c})_{k\in\mathbb{N}_0}$. Assume that ${p_{k,o}}^{(N)} \to p_{k,o}$ and ${p_{k,c}}^{(N)} \to p_{k,c}$ as $N \to \infty$ for all $k \ge 0$. Then for any $\mathbb{N}$-valued sequence $(b_N)_{N\in\mathbb{N}}$ with $b_N \to \infty$ it holds that*

$$\begin{aligned}\lim_{N\to\infty} \mathbb{P}\left(\forall g \in \mathbb{N}_0 : Z^{(N)}_g > 0\right) &= \lim_{N\to\infty} \mathbb{P}\left(\exists g \in \mathbb{N}_0 : Z^{(N)}_g \ge b_N\right)\\ &= \lim_{N\to\infty} \mathbb{P}\left(\exists g \in \mathbb{N}_0 : \overline{Z}^{(N)}_g \ge b_N\right)\\ &= \pi,\end{aligned}$$

where π denotes the survival probability of $\mathbf{Z}^\infty$.

We provide the proof in the supplementary material. The two first equalities are shown using a similar method as in the proof of Lemma 3.7 of [1]. The last equality is shown by using an exact coupling argument between the processes $\mathbf{Z}^{(N)}$ and $\mathbf{Z}^\infty$ until the total size of both processes reaches a sufficiently high level to ensure that the processes survive whp, or they both die out. In the supplementary materials we introduce Lemma 3.6.1 which is used for this proof.

3.4 Invasion of cooperative parasites in host populations structured by a complete graph

In this section we prepare and give the proof of Theorem 3.2.6. Therefore, let us briefly recall the setting. We consider a host-parasite dynamics as introduced in Section 2 on a complete graph with D_N vertices, where we assume that $D_N \in \Theta(N^\beta)$ with $0 < \beta < 1$. In case of an host infection V_N parasites are generated. The number of infected hosts in generation g is denoted by $J_g = J^{(N)}_g$ and the total number until generation g is denoted by $\overline{J}_g = \overline{J}^{(N)}_g$. In Subsection 3.4.1 and 3.4.2 we do the preparatory work to show Theorem 3.2.6 (ii), and thus we consider the critical scaling $V_N \sim a\sqrt{D_N}$ in these subsections, where $a > 0$.

In this and the next section we often will make use of the inequalities $\exp(-x) \ge 1 - x \ge \exp(-x)\exp(-x^2)$, $\exp(-x) \le 1 - x/2$ for $x \in [0, 1/2]$ and Bernoulli's inequality $(1 + x)^i \ge 1 + ix$ for $i \in \mathbb{N}$ and $x \ge -1$.

Furthermore, we will compare the infection dynamics happening within one generation often with balls-into-boxes experiments. The following lemma gives control about certain events happening in these experiments.

Lemma 3.4.1. *Let $(m'_N), (V'_N), (h'_N), (\ell'_N), (D'_N)$ be non-negative sequences with $1 \leq m'_N \leq h'_N \leq \ell'_N$, assume $\frac{{\ell'_N}^4 {V'_N}^3}{{D'_N}^2} \in o(1)$, $\ell'_N \to \infty$ for $N \to \infty$, as well as $V'_N \sim a\sqrt{D'_N}$ for some $a > 0$. Consider $D'_N - m'_N$ boxes (for N large enough such that $D'_N - h'_N > 0$) and $m'_N V'_N$ balls. Assume the balls are put independently and uniformly at random into the boxes. Consider the event $C_k^{(h'_N)}$, that k many of the first $D'_N - h'_N$ boxes contain exactly two balls and the remaining boxes contain at most one ball. We have for all $k \leq \ell'_N$ and for N large enough*

$$\mathbb{P}(W'_N = k) \exp\left(-\frac{{\ell'_N}^5 {V'_N}^3}{{D'_N}^2}\right) \leq \mathbb{P}\left(C_k^{(h'_N)}\right) \leq \left(\frac{(m'_N V'_N)^2}{2D'_N}\right)^k \frac{1}{k!} \exp\left(-\frac{(m'_N V'_N - 2\ell'_N)^2}{2D'_N}\right) \cdot \exp\left(\frac{{\ell'_N}^2 V'_N}{D'_N}\right),$$

where W'_N is Poisson distributed with parameter $\frac{\left(m'_N V'_N - 2\ell'_N\right)^2}{2D'_N}$.

A proof can be found in the supplementary material.

Let $(\ell_N)_N$ be a sequence such that $\ell_N \in o(V_N)$ and $\ell_N \to \infty$. Then, in the following we will denote by $W_{o,m}^{(N)}$ a Poisson distributed random variable with parameter $\frac{m(V_N - 2\ell_N)^2}{2D_N} = mu_N$ with

$$u_N := \frac{(V_N - 2\ell_N)^2}{2D_N}$$

for any $m \in \mathbb{N}$ and similarly by $W_{c,m}^{(N)}$ a Poisson distributed random variable with parameter $\frac{m(V_N - 2\ell_N)^2}{D_N} = 2mu_N$.

3.4.1 Upper bound on the invasion probability on a complete graph

To derive an upper bound on the invasion probability we estimate from above the total number of infected hosts by the total size of a branching process with cooperation with offspring and cooperation distributions that are approximately Poisson distributed.

Definition 3.4.2. *(Upper DBPC)*
Let $\ell_N \underset{N\to\infty}{\to} \infty$ satisfying $\frac{\ell_N^7 V_N^3}{D_N^2} \in o(1)$. Let $\mathbf{Z}_u^{(N)} = \left(Z_{g,u}^{(N)}\right)_{g\in\mathbb{N}_0}$ be a branching process with cooperation with $Z_{0,u}^{(N)} = 1$ almost surely, and offspring and cooperation distribution with probability weights $p_{u,o}^{(N)} = \left(p_{j,u,o}^{(N)}\right)_{j\in\mathbb{N}_0}$ and $p_{u,c}^{(N)} = \left(p_{j,u,c}^{(N)}\right)_{j\in\mathbb{N}_0}$, respectively with

$$p_{j,u,o}^{(N)} := \mathbb{P}\left(W_{o,1}^{(N)} = j\right) \exp\left(-\frac{\ell_N^5 V_N^3}{D_N^2}\right),$$

for all $0 \le j \le \ell_N$ *and*

$$p^{(N)}_{\ell_N+1,u,o} := 1 - \sum_{j=0}^{\ell_N} p^{(N)}_{j,u,o},$$

as well as

$$p^{(N)}_{j,u,c} := \mathbb{P}(W^{(N)}_{c,1} = j) \exp\left(-2\frac{\ell_N^5 V_N^3}{D_N^2}\right),$$

for all $0 \le j \le \ell_N$ *and*

$$p^{(N)}_{\ell_N+1,u,c} := 1 - \sum_{j=0}^{\ell_N} p^{(N)}_{j,u,c}.$$

Denote by $\overline{\mathbf{Z}}^{(N)}_u := \left(\overline{Z}^{(N)}_{g,u}\right)_{g\in\mathbb{N}_0}$, *where* $\overline{Z}^{(N)}_{g,u} := \sum_{i=0}^{g} Z^{(N)}_{i,u}$, *that is* $\overline{Z}^{(N)}_{g,u}$ *gives the total size of* $\mathbf{Z}^{(N)}_u$ *accumulated till generation* g.

In the next proposition, we show that the total size of the infection process $\overline{\mathbf{J}}^{(N)}$ can be coupled whp from above with the total size of the DBPC of Definition 3.4.2 $\overline{\mathbf{Z}}^{(N)}_u$ up to the first random generation at which $\overline{\mathbf{Z}}^{(N)}_u$ reaches the size ℓ_N (for $\ell_N \to \infty$ not too fast as considered in the definition of the DBPC) or the process $\mathbf{Z}^{(N)}_u$ dies out.

Proposition 3.4.3. *For any sequence* $(b_N)_{N\in\mathbb{N}}$ *introduce the stopping time*

$$\tau^{\overline{\mathbf{Z}}_u}_{b_N,0} := \inf\left\{g \in \mathbb{N}_0 : \overline{Z}^{(N)}_{g,u} \ge b_N \text{ or } Z^{(N)}_{g,u} = 0\right\}.$$

Let $(\ell_N)_{N\in\mathbb{N}}$ *be the same sequence as used in Definition 3.4.2. Then*

$$\lim_{N\to\infty} \mathbb{P}\left(\overline{J}^{(N)}_g \le \overline{Z}^{(N)}_{g,u}, \forall n < \tau^{\overline{\mathbf{Z}}_u}_{\ell_N,0}\right) = 1,$$

and

$$\lim_{N\to\infty} \mathbb{P}\left(J^{(N)}_{\tau^{\overline{\mathbf{Z}}_u}_{\ell_N,0}} = 0 \;\middle|\; Z^{(N)}_{\tau^{\overline{\mathbf{Z}}_u}_{\ell_N,0},u} = 0\right) = 1$$

and

$$\lim_{N\to\infty} \mathbb{P}\left(\overline{J}^{(N)}_{\tau^{\overline{\mathbf{Z}}_u}_{\ell_N,0}} \ge \ell_N \;\middle|\; \overline{J}^{(N)}_{\tau^{\overline{\mathbf{Z}}_u}_{\ell_N,0}} \ge \overline{Z}^{(N)}_{\tau^{\overline{\mathbf{Z}}_u}_{\ell_N,0},u}\right) = 1.$$

Proof. In order to show that $\mathcal{J}^{(N)}$ and $\mathbf{Z}^{(N)}_u$ can be coupled such that $\overline{J}^{(N)}_g \le \overline{Z}^{(N)}_{g,u}$ for all $g < \tau^{\overline{\mathbf{Z}}_u}_{\ell_N,0}$ whp we show that

$$\mathbb{P}\Big(\overline{J}^{(N)}_{g+1} = k \mid \overline{J}^{(N)}_g = m\Big) \le \mathbb{P}\Big(\overline{Z}^{(N)}_{g+1,u} = k \mid \overline{Z}^{(N)}_{g,u} = m\Big)$$

for all $k, m < \ell_N$. Then, one can use the Markov property to show that

$$\mathbb{P}\big(\{\overline{J}_g^{(N)} \le k_g, \dots, \overline{J}_1^{(N)} \le k_1\} \cap \{g < \tau_{\ell_N,0}^{\overline{\mathbf{Z}}_u}\}\big) \le \mathbb{P}\big(\{\overline{Z}_{g,u}^{(N)} \le k_g, \dots, \overline{Z}_{1,u}^{(N)} \le k_1\} \cap \{g < \tau_{\ell_N,0}^{\overline{\mathbf{Z}}_u}\}\big)$$

where $k_1, \dots, k_g \in \mathbb{N}$ i.e. that $\overline{\mathbf{Z}}_u^{(N)}$ stochastically dominates $\overline{\mathbf{J}}^{(N)}$ until $\overline{\mathbf{Z}}_u^{(N)}$ reaches the level ℓ_N. Having this one can deduce by Strassen's theorem that $\overline{\mathbf{J}}^{(N)}$ and $\overline{\mathbf{Z}}_u^{(N)}$ can be coupled as claimed.

Up to generation

$$\tau_{\ell_N,0}^{\overline{\mathbf{J}}} := \inf\left\{g \in \mathbb{N}_0 : \overline{J}_g^{(N)} \ge \ell_N \text{ or } J_g^{(N)} = 0\right\}$$

the total number of parasites that are moving in the graph is upper bounded by $\ell_N V_N$. Consider (for N large enough such that $D_N - \ell_N > 0$) the following experiment with $D_N - \ell_N$ boxes, $\ell_N V_N$ balls. Assume that balls are thrown uniformly at random in the boxes. The probability that there exists a box with at least 3 balls in it can be upper bounded as follows

$$\mathbb{P}\left(\exists \text{ 1 box with more than 3 balls}\right) \le D_N \left(\ell_N V_N\right)^3 \frac{1}{\left(D_N - \ell_N\right)^3} \sim \frac{\ell_N^3 V_N^3}{D_N^2} \to 0.$$

This means that with the assumed scaling of ℓ_N it is unlikely that such an event occurs before generation $\tau_{\ell_N,0}^{\overline{\mathbf{Z}}_u}$. Consequently for whp couplings, we can only focus on infections generated by pairs of parasites.

Now consider a complete graph with $1 \le m_N < \ell_N$ infected vertices and at most $\ell_N - 1$ empty or infected vertices. The probability that k infections are generated for $0 \le k \le \ell_N$ can be estimated from above by the probability that k boxes are filled with at least two balls and the remaining boxes are filled with at most one ball in the following balls-into-boxes experiment: Consider $D_N - m_N$ boxes and $m_N V_N$ balls. Place the balls uniformly at random into the boxes. Denote by $C_k^{(N)}$ the event that k boxes contain exactly two balls and all other boxes contain at most one ball.

By Lemma 3.4.1 with $D_N' = D_N, m_N' = m_N, V_N' = V_N$, $h_N' = m_N$ and $\ell_N' = \ell_N$ we can estimate for $m_N \ge 2$

$$\begin{aligned}\mathbb{P}\left(C_k^{(N)}\right) &\ge \left(\frac{\left(m_N V_N - 2\ell_N\right)^2}{2D_N}\right)^k \frac{1}{k!} \exp\left(-\frac{\left(m_N V_N - 2\ell_N\right)^2}{2D_N}\right) \exp\left(-\frac{\ell_N^5 V_N^3}{D_N^2}\right) \\ &\ge \left(m_N^2 u_N\right)^k \frac{1}{k!} \exp\left(-m_N^2 \frac{V_N^2}{2D_N}\right) \exp\left(-\frac{\ell_N^5 V_N^3}{D_N^2}\right), \end{aligned} \tag{3.4.1}$$

again for N large enough, and for $m_N = 1$

$$\mathbb{P}\big(C_k^{(N)}\big) \ge \left(u_N\right)^k \frac{1}{k!} \exp\left(-u_N\right) \exp\left(-\frac{\ell_N^5 V_N^3}{D_N^2}\right). \tag{3.4.2}$$

In order to prove that $\overline{\mathbf{J}}^{(N)}$ can be coupled with $\overline{\mathbf{Z}}_u^{(N)}$ such that $\overline{\mathbf{Z}}_u^{(N)}$ dominates $\overline{\mathbf{J}}^{(N)}$, we show that the probability $\mathbb{P}\big(Z_{n+1,u}^{(N)} = k | Z_{n,u}^{(N)} = m_N\big)$ can be estimated from above by the lower bound derived in (3.4.1) for $m_N \ge 2$ and by (3.4.2) for $m_N = 1$.

Consider independent random variables $\left(X_i^{(N)}\right)_{i\in\mathbb{N}}$ and $\left(Y_{(i,j)}^{(N)}\right)_{(i,j)\in\mathbb{N}^2}$ with probability weights $p_{u,o}^{(N)}$ and $p_{u,c}^{(N)}$ respectively. We have that

$$\begin{aligned}
\mathbb{P}\left(Z_{g+1,u}^{(N)}=k|Z_{g,u}^{(N)}=m_N\right) &= \mathbb{P}\left(\sum_{i=1}^{m_N}X_i^{(N)}+\sum_{i,j=1,i>j}^{m_N}Y_{(i,j)}^{(N)}=k\right) \qquad (3.4.3)\\
&= \sum_{\substack{k_o,k_c:\\ k_o+k_c=k,k_o,k_c\geq 0}} \mathbb{P}\left(\sum_{i=1}^{m_N}X_i^{(N)}=k_o\right)\mathbb{P}\left(\sum_{i,j=1,i>j}^{m_N}Y_{(i,j)}^{(N)}=k_c\right)\\
&= \sum_{\substack{k_o,k_c:\\ k_o+k_c=k,k_o,k_c\geq 0}} \mathbb{P}(W_{o,m_N}^{(N)}=k_o)\exp\left(-m_N\frac{\ell_N^5V_N^3}{D_N^2}\right)\mathbb{P}(W_{c,\binom{m_N}{2}}^{(N)}=k_c)\exp\left(-2\binom{m_N}{2}\frac{\ell_N^5V_N^3}{D_N^2}\right)\\
&= \exp\left(-m_N^2\frac{\ell_N^5V_N^3}{D_N^2}\right)\sum_{\substack{k_o,k_c:\\ k_o+k_c=k,\\ k_o,k_c\geq 0}}\mathbb{P}(W_{o,m_N}^{(N)}=k_o)\mathbb{P}(W_{o,m_N(m_N-1)}^{(N)}=k_c)\\
&= \exp\left(-m_N^2\frac{\ell_N^5V_N^3}{D_N^2}\right)\mathbb{P}(W_{o,m_N^2}^{(N)}=k),
\end{aligned}$$

where we have used that

$$\begin{aligned}
\mathbb{P}\left(\sum_{i=1}^{m_N}X_i^{(N)}=k\right) &= \sum_{\substack{(k_1,\dots,k_{m_N}):\\ k_1+\cdots+k_{m_N}=k}}\left[\prod_{i=1}^{m_N}\mathbb{P}(W_{o,1}^{(N)}=k_i)\exp\left(-\frac{\ell_N^5V_N^3}{D_N^2}\right)\right]\\
&= \exp\left(-m_N\frac{\ell_N^5V_N^3}{D_N^2}\right)\mathbb{P}(W_{o,m_N}^{(N)}=k)
\end{aligned}$$

and a similar reasoning for $Y_{(i,j)}^{(N)}$, as well as

$$\mathbb{P}\left(W_{c,\binom{m_N}{2}}^{(N)}=k_c\right)=\mathbb{P}\left(W_{o,m_N(m_N-1)}^{(N)}=k_c\right).$$

Since

$$\exp\left(-\frac{m_N^2V_N^2}{2D_N}\right)\exp\left(-\frac{\ell_N^5V_N^3}{D_N^2}\right)\geq \quad \exp\left(-m_N^2\frac{\ell_N^5V_N^3}{D_N^2}\right)\exp\left(-m_N^2u_N\right)$$

for $1\leq m_N\leq \ell_N$ and N large enough, we have that (3.4.1) $\geq$ (3.4.3) for $m_N\geq 2$ and (3.4.2) $\geq$ (3.4.3) for $m_N=1$ for N large enough. Thus, because of the Markov property we can successively couple the two processes until $\overline{\mathbf{Z}}_u^{(N)}$ reaches the level ℓ_N. □

Using the results of the previous Section 3.3 we will show that for the upper DBPC defined in Definition 3.4.2 the probability of reaching an arbitrary high number of individuals a t some generation is asymptotically the same as the survival probability of a DBPC whose offspring and cooperation distributions are respectively $\mathrm{Poi}\left(a^2/2\right)$ and $\mathrm{Poi}(a^2)$ distributed. Let $X\sim\mathrm{Poi}(\lambda)$ with $\lambda>0$, then we denote the probability weight of X in j by $p_j(\lambda)=\mathbb{P}(X=j)$.

Proposition 3.4.4. *(Probability for the total size of the upper DBPC to reach a level b_N).* *Consider a sequence $(b_N)_{N\in\mathbb{N}}$ with $b_N \underset{N\to\infty}{\to} \infty$ and assume that $V_N \sim a\sqrt{D_N}$ for $0 < a < \infty$. Then, we have*

$$\lim_{N\to\infty} \mathbb{P}\left(\exists g \in \mathbb{N}_0 : \overline{Z}^{(N)}_{g,u} \geq b_N\right) = \pi(a).$$

Proof. The claim follows as an application of Proposition 3.3.7. Thus, we check that the sequence $\left(\mathbf{Z}^{(N)}_u\right)_{N\in\mathbb{N}}$ satisfies the assumptions of Proposition 3.3.7. Let us first consider the convergence of $p^{(N)}_{j,u,o} \to p_j\left(\frac{a^2}{2}\right)$ for every $j \in \mathbb{N}_0$. Note that for a given j, we can choose N large enough, such that $j \leq \ell_N$ and hence

$$p^{(N)}_{j,u,o} = \mathbb{P}\left(W^{(N)}_{o,1} = j\right) \exp\left(-\frac{\ell_N^5 V_N^3}{D_N^2}\right).$$

By the choice of ℓ_N and as we assumed that $V_N \sim a\sqrt{D_N}$ we have $u_N \to \frac{a^2}{2}$. Thus, by continuity it follows that

$$(u_N)^j \frac{1}{j!} \exp(-u_N) \exp\left(-\frac{\ell_N^5 V_N^3}{D_N^2}\right) \underset{N\to\infty}{\longrightarrow} \left(\frac{a^2}{2}\right)^j \frac{1}{j!} \exp\left(-\frac{a^2}{2}\right) = p_j\left(\frac{a^2}{2}\right).$$

Thus, we showed that $p^{(N)}_{j,u,o} \to p_j\left(\frac{a^2}{2}\right)$ as $N \to \infty$ for every $j \in \mathbb{N}$. Analogously one can show that $p^{(N)}_{j,u,c} \to p_j(a^2)$ as $N \to \infty$ for every $j \in \mathbb{N}$. Next we need to check that the first and second moment of the offspring and cooperation distribution converge.

Let $X^{(N)}$ be distributed according to the offspring distributions $\left(p^{(N)}_{j,u,o}\right)_{j\geq 0}$ of the upper DBPC $\mathbf{Z}^{(N)}_u$. Then

$$\begin{aligned}\mathbb{E}\left[X^{(N)}\right] &= \sum_{j=1}^{\ell_N+1} j p^{(N)}_{j,u,o} \\ &= \sum_{j=1}^{\ell_N} (u_N)^j \frac{1}{(j-1)!} e^{-u_N} \exp\left(-\frac{\ell_N^5 V_N^3}{D_N^2}\right) + (\ell_N+1)\left(1 - \sum_{j=0}^{\ell_N} p^{(N)}_{j,u,o}\right).\end{aligned}$$

Since $\ell_N \to \infty$ and $u_N \underset{N\to\infty}{\to} \frac{a^2}{2}$, we have

$$\lim_{N\to\infty} e^{-u_N} \sum_{j=0}^{\ell_N-1} \frac{(u_N)^j}{j!} = 1.$$

It follows that

$$\exp\left(-\frac{\ell_N^5 V_N^3}{D_N^2}\right) \sum_{j=1}^{\ell_N} (u_N)^j \frac{1}{(j-1)!} e^{-u_N} = u_N \exp\left(-\frac{\ell_N^5 V_N^3}{D_N^2}\right) \left(\sum_{j=0}^{\ell_N-1} \frac{(u_N)^j}{j!} e^{-u_N}\right) \to \frac{a^2}{2}$$

and

$$(\ell_N+1)\left(1 - \sum_{j=0}^{\ell_N} p^{(N)}_{j,u,o}\right) = (\ell_N+1)\left(1 - \exp\left(-\frac{\ell_N^5 V_N^3}{D_N^2}\right) \mathbb{P}\left(W^{(N)}_{o,1} \leq \ell_N\right)\right).$$

Now by Markov's inequality follows that

$$\mathbb{P}\left(W_{o,1}^{(N)} \leq \ell_N\right) = \mathbb{P}\left((W_{o,1}^{(N)})^2 \leq \ell_N^2\right) \geq 1 - \frac{\mathbb{E}[(W_{o,1}^{(N)})^2]}{\ell_N^2} = 1 - \frac{u_N + u_N^2}{\ell_N^2}.$$

Hence

$$\begin{aligned}(\ell_N + 1)\left(1 - \sum_{j=0}^{\ell_N} p_{j,u,o}^{(N)}\right) &\leq (\ell_N + 1)\left(1 - \exp\left(-\frac{\ell_N^5 V_N^3}{D_N^2}\right)\right) + (u_N + u_N^2)\frac{\ell_N + 1}{\ell_N^2} \\ &\leq \frac{(\ell_N + 1)\ell_N^5 V_N^3}{D_N^2} + (u_N + u_N^2)\frac{\ell_N + 1}{\ell_N^2} \to 0.\end{aligned}$$

Consequently

$$\mathbb{E}\left[X^{(N)}\right] \to \frac{a^2}{2}$$

as $N \to \infty$. Similarly, we have for the second moment

$$\mathbb{E}\Big[\big(X^{(N)}\big)^2\Big] = \sum_{j=1}^{\ell_N} j^2 \,(u_N)^j \,\frac{1}{j!} e^{-u_N} \exp\left(-\frac{\ell_N^5 V_N^3}{D_N^2}\right) + (\ell_N + 1)^2\Big(1 - \sum_{j=0}^{\ell_N} p_{j,u,o}^{(N)}\Big).$$

The second term again vanishes in the limit by the same argument as before just that we use the Markov inequality for the third moment such that

$$\mathbb{P}\left(W_{o,1}^{(N)} \leq \ell_N\right) \geq 1 - \frac{u_N + 3u_N^2 + u_N^3}{\ell_N^3},$$

which yields that

$$(\ell_N + 1)^2\left(1 - \sum_{j=0}^{\ell_N} p_{j,u,o}^{(N)}\right) \leq \frac{(\ell_N + 1)^2 \ell_N^5 V_N^3}{D_N^2} + (u_N + 3u_N^2 + u_N^3)\frac{(\ell_N + 1)^2}{\ell_N^3} \to 0,$$

as $N \to \infty$ and

$$\exp\left(-\frac{\ell_N^5 V_N^3}{D_N^2}\right) \sum_{j=1}^{\ell_N} j^2 \frac{(u_N)^j}{j!} e^{-u_N} = \exp\left(-\frac{\ell_N^5 V_N^3}{D_N^2}\right) u_N \left(u_N \sum_{j=0}^{\ell_N - 2} \frac{(u_N)^j}{j!} e^{-u_N} + \sum_{j=0}^{\ell_N - 1} \frac{(u_N)^j}{j!} e^{-u_N}\right).$$

Now one can show analogously as before that

$$\mathbb{E}\Big[\big(X^{(N)}\big)^2\Big] \to \frac{a^4}{4} + \frac{a^2}{2}.$$

For the expectations and the second moments of the cooperation distributions one argues analogously, except that one shows convergence to a^2 and $a^4 + a^2$, respectively. □

3.4.2 Lower bound on the invasion probability on a complete graph

Lower bound on the probability to infect at least N^ε many hosts

We first aim to show that the total number of infected hosts until the parasite population dies out or N^ε hosts are infected for $\varepsilon > 0$ small enough can be lower bounded whp by the total size process of a DBPC. This DBPC we introduce next.

Definition 3.4.5. *(Lower DBPC)*
Let $\ell_N \underset{N\to\infty}{\to} \infty$ satisfying $\frac{\ell_N^7 V_N \log(\log(N))}{D_N} \in o(1)$. Let $\mathbf{Z}_\ell^{(N)} = \left(Z_{g,\ell}^{(N)}\right)_{g\in\mathbb{N}_0}$ be a branching process with cooperation with $Z_{0,\ell}^{(N)} = 1$ almost surely, and offspring distribution $p_{\ell,o}^{(N)} = \left(p_{j,\ell,o}^{(N)}\right)_{j\in\mathbb{N}_0}$ as well as cooperation distribution $p_{\ell,c}^{(N)} = \left(p_{j,\ell,c}^{(N)}\right)_{j\in\mathbb{N}_0}$ with

$$p_{j,\ell,o}^{(N)} := \mathbb{P}(W_{o,1}^{(N)} = j)\exp\left(-\frac{\ell_N V_N}{D_N}\right),$$

for all $0 < j \le \ell_N$ and

$$p_{0,\ell,o}^{(N)} := 1 - \sum_{j=1}^{\ell_N} p_{j,\ell,o}^{(N)},$$

as well as

$$p_{j,\ell,c}^{(N)} := \mathbb{P}(W_{c,1}^{(N)} = j)\exp\left(-\frac{\ell_N V_N}{D_N}\right),$$

for all $0 < j \le \ell_N$ and

$$p_{0,\ell,c}^{(N)} := 1 - \sum_{j=1}^{\ell_N} p_{j,\ell,c}^{(N)}.$$

Denote by $\overline{\mathbf{Z}}_\ell^{(N)} := \left(\overline{Z}_{g,\ell}^{(N)}\right)_{g\in\mathbb{N}_0}$ with $\overline{Z}_{g,\ell}^{(N)} := \sum_{i=0}^{g} Z_{i,\ell}^{(N)}$, that is $\overline{Z}_{g,\ell}^{(N)}$ gives the total size of $\mathbf{Z}_\ell^{(N)}$ accumulated till generation g.

Proposition 3.4.6. *For any sequence $(b_N)_{N\in\mathbb{N}}$ introduce the stopping time*

$$\tau_{b_N,0}^{\overline{\mathbf{Z}}_\ell} := \inf\left\{g \in \mathbb{N}_0 : \overline{Z}_{g,\ell}^{(N)} \ge b_N \text{ or } Z_{g,\ell}^{(N)} = 0\right\}.$$

Consider a sequence $(\ell_N)_{N\in\mathbb{N}}$ with $\ell_N = N^\varepsilon$ for $\varepsilon > 0$ such that $\frac{\ell_N^7 V_N \log(\log(N))}{D_N} \in o(1)$. Then

$$\lim_{N\to\infty} \mathbb{P}\left(\overline{J}_g^{(N)} \ge \overline{Z}_{g,\ell}^{(N)}, \forall g \le \tau_{\ell_N,0}^{\overline{\mathbf{Z}}_\ell}\right) = 1. \tag{3.4.4}$$

Proof. As in the proof of Proposition 3.4.3 in the scaling $\ell_N^7 V_N^3 / D_N^2 = o(1)$ it is unlikely, that up to the first generation, at which the total infection process reaches size ℓ_N, an infection occurs due to more than a pair of parasites. Consequently for a coupling whp we can only focus on infections generated by pairs of parasites, and do not need to treat infections generated by at least 3 parasites.

Now consider a complete graph with exactly $m_N < \ell_N$ infected vertices and with at most $\ell_N - 1$ empty or infected vertices. The number of new infections generated on such a graph in the next generation can be lower bounded by the number of infections arising in the following experiment: Consider D_N boxes and $m_N V_N$ balls. Distribute the balls uniformly at random into the boxes. Assume a new infection is created for each of the first $D_N - \ell_N$ boxes that contains at least two balls. Let A_{m_N} be the number of infections generated in this experiment and let $B_k{}^{(N)}$ be the event that exactly k of the $D_N - \ell_N$ first boxes contain exactly two balls and all other boxes have at most one ball. We have

$$\mathbb{P}\big(J_g^{(N)} \geq A_{m_N} \mid J_{g-1}^{(N)} = m_N\big) = 1$$

and

$$\mathbb{P}\big(A_{m_N} = k\big) \geq \mathbb{P}\big(B_k^{(N)}\big).$$

We will show below that there exists a constant $C_1 > 0$ such that for N large enough

$$\left|\mathbb{P}\left(B_k^{(N)}\right) - \mathbb{P}\left(Z_{g+1,\ell}^{(N)} = k \mid Z_{g,\ell}^{(N)} = m_N\right)\right| \leq \left(\frac{m_N^2 V_N^2}{2D_N}\right)^k \frac{1}{k!} \exp\left(-\frac{m_N^2 V_N^2}{2D_N}\right)\left(\frac{C_1 V_N \ell_N^4}{D_N}\right) + \frac{C_1 \ell_N^6 V_N}{D_N} \tag{3.4.5}$$

for any k with $0 \leq k \leq \ell_N$. Since

$$\sum_{k=0}^{\ell_N} \left(\frac{m_N^2 V_N^2}{2D_N}\right)^k \frac{1}{k!} \exp\left(-\frac{m_N^2 V_N^2}{2D_N}\right)\left(\frac{C_1 V_N \ell_N^4}{D_N}\right) + \frac{C_1 \ell_N^7 V_N}{D_N} \leq \frac{2C_1 \ell_N^7 V_N}{D_N} \to 0$$

we can couple (by means of Lemma 3.6.3) the balls into boxes experiment with the lower DBPC $\mathbf{Z}_\ell^{(N)}$, such that given that $\{Z_{g,\ell}^{(N)} = m_N\}$ the event $B_k^{(N)}$ occurs whp, if $\{Z_{g+1,\ell}^{(N)} = k\}$ for any $k \in \{0, \ldots \ell_N\}$ and vice-versa. We can repeat this argument till $\tau_{\ell_N,0}^{\overline{\mathbf{Z}}_\ell}$ whp. Indeed, by Proposition 3.3.6 it exists $C_2 > 0$ such that $\mathbb{P}\big(\tau_{\ell_N,0}^{\overline{\mathbf{Z}}_\ell} \leq C_2 \log(\log(\ell_N))\big) \to 1$, as by analogous arguments as in the proof of Proposition 3.4.4 it can be shown that the first and second moments of $Z_{1,\ell}^{(N)}$ are uniformly bounded in N. Since

$$\left(1 - \frac{2C_1 \ell_N^7 V_N}{D_N}\right)^{C_2 \log\log(\ell_N)} \to 1$$

it follows that we can couple whp subsequently performed balls into boxes problems and $\mathbf{Z}_\ell^{(N)}$ for any generation g with $g \leq \tau_{\ell_N,0}^{\overline{\mathbf{Z}}_\ell}$, which implies (3.4.4).

So to finish the proof it remains to show (3.4.5).

We start by controlling the probabilities of the events $B_k^{(N)}$. By Lemma 3.4.1 with $D'_N = D_N$, $m'_N = m_N$, $\ell'_N = \ell_N$, $V'_N = V_N$ and $h'_N = \ell_N$ we can estimate

$$\mathbb{P}\big(B_k^{(N)}\big) \leq \left(\frac{(m_N V_N)^2}{2D_N}\right)^k \frac{1}{k!} \exp\left(-\frac{(m_N V_N - 2\ell_N)^2}{2D_N}\right) \exp\left(\frac{\ell_N^2 V_N}{D_N}\right), \tag{3.4.6}$$

and

$$\mathbb{P}\left(B_k\right) \geq \mathbb{P}(W_B^{(N)} = k) \exp\left(-\frac{\ell_N^5 V_N^3}{D_N^2}\right) \tag{3.4.7}$$

for a Poisson distributed random variable $W_B^{(N)}$ with parameter $\frac{(m_N V_N - 2\ell_N)^2}{2D_N}$ and N large enough.

Next we control the transition probabilities of $\mathbf{Z}_\ell^{(N)}$. Consider independent random variables $(X_i{}^{(N)})_{i\in\mathbb{N}}$ and $(Y_{(i,j)}^{(N)})_{i<j}$ with probability weights $p_{\ell,o}^{(N)}$ and $p_{\ell,c}^{(N)}$, respectively. We have

$$\begin{aligned}
\mathbb{P}\left(X_1^{(N)} = 0\right) &= 1 - \sum_{j=1}^{\ell_N} p_{j,\ell,o}^{(N)} \\
&= 1 - \sum_{j=1}^{\ell_N} \exp\left(-\frac{\ell_N V_N}{D_N}\right) \mathbb{P}\left(W_{o,1}^{(N)} = j\right) \\
&= 1 + \exp\left(-\frac{\ell_N V_N}{D_N}\right) \exp(-u_N) - \sum_{j=0}^{\ell_N} \exp\left(-\frac{\ell_N V_N}{D_N}\right) \mathbb{P}\left(W_{o,1}^{(N)} = j\right) \\
&= \exp\left(-\frac{\ell_N V_N}{D_N}\right) \exp(-u_N) + 1 - \exp\left(-\frac{\ell_N V_N}{D_N}\right) + \exp\left(-\frac{\ell_N V_N}{D_N}\right) \mathbb{P}\left(W_{o,1}^{(N)} \geq \ell_N + 1\right).
\end{aligned}$$

We define the constant

$$c_o^{(N)} := 1 + \exp\big(-\frac{\ell_N V_N}{D_N}\big)\big(\mathbb{P}\big(W_{o,1}^{(N)} \geq \ell_N + 1\big) - 1\big).$$

We have $\mathbb{P}\big(W_{o,1}^{(N)} \geq \ell_N + 1\big) \in \mathcal{O}\Big(\frac{u_N^{\ell_N}}{(\ell_N)!}\Big)$, in particular $c_o^{(N)} \in \Theta\Big(\frac{\ell_N V_N}{D_N}\Big)$, where we used that $\ell_N \sim N^\varepsilon$, and thus $\frac{u_N^{\ell_N}}{(\ell_N)!}$ decays exponentially fast in N. Let us recall that by definition

$$\mathbb{P}\left(X_i^{(N)} = k\right) = \tfrac{(u_N)^k}{k!} \exp(-u_N) \exp\left(-\tfrac{\ell_N V_N}{D_N}\right) + c_o^{(N)} \mathbb{1}_{\{k=0\}},$$

where $0 \leq k \leq \ell_N$. We see that for $0 \leq k_o \leq \ell_N$ it holds that

$$\mathbb{P}\left(\sum_{i=1}^{m_N} X_i^{(N)} = k_o\right) = \sum_{\substack{k_1,\ldots,k_m: \\ k_1+\ldots+k_m=k_o}} \prod_{i=1}^{m_N} \mathbb{P}\left(X_i^{(N)} = k_i\right).$$

This allows us to derive the lower and upper bound

$$\mathbb{P}\left(\sum_{i=1}^{m_N} X_i^{(N)} = k_o\right) \geq \exp\left(-m_N \frac{\ell_N V_N}{D_N}\right) \mathbb{P}\left(W_{o,m_N}^{(N)} = k_o\right), \tag{3.4.8}$$

and

$$\mathbb{P}\left(\sum_{i=1}^{m_N} X_i^{(N)} = k_o\right) \leq \exp\left(-m_N \frac{\ell_N V_N}{D_N}\right) \mathbb{P}\left(W_{o,m_N}^{(N)} = k_o\right) + m_N c_o^{(N)},$$

where $W^{(N)}_{o,m_N} \sim \mathrm{Poi}(m_N u_N)$ and if $k_1 + \ldots + k_{m_N} = k_o$, then the number of k_i with $k_i = 0$ is at most m_N.

Now we obtain analogously as before that

$$\mathbb{P}\left(Y^{(N)}_{1,2} = 0\right) = \exp\left(-\frac{\ell_N V_N}{D_N}\right)\exp(-u_N) + c^{(N)}_c$$

where the constant is defined as

$$c^{(N)}_c := 1 + \exp\left(-\frac{\ell_N V_N}{D_N}\right)\left(\mathbb{P}\left(\tilde{Y}^{(N)} \geq \ell_N + 1\right) - 1\right) \in \Theta\left(\frac{\ell_N V_N}{D_N}\right),$$

with $\tilde{Y}^{(N)} \sim \mathrm{Poi}(2u_N)$. Similarly as before we arrive at the lower and upper bounds

$$\begin{aligned}
\mathbb{P}\left(\sum_{i,j=1,i<j}^{m_N} Y^{(N)}_{(i,j)} = k_c\right) &\geq \mathbb{P}\left(W^{(N)}_{c,\binom{m_N}{2}} = k_c\right)\exp\left(-\binom{m_N}{2}\frac{\ell_N V_N}{D_N}\right) \\
\mathbb{P}\left(\sum_{i,j=1,i<j}^{m_N} Y^{(N)}_{(i,j)} = k_c\right) &\leq \mathbb{P}\left(W^{(N)}_{c,\binom{m_N}{2}} = k_c\right)\exp\left(-\binom{m_N}{2}\frac{\ell_N V_N}{D_N}\right) + \binom{m_N}{2} c^{(N)}_c.
\end{aligned} \tag{3.4.9}$$

So in summary we have

$$\begin{aligned}
\mathbb{P}\left(Z^{(N)}_{g+1,\ell} = k | Z^{(N)}_{g,\ell} = m_N\right) &= \mathbb{P}\left(\sum_{i=1}^{m_N} X^{(N)}_i + \sum_{i,j=1,i<j}^{m_N} Y^{(N)}_{(i,j)} = k\right) \\
&= \sum_{k_o,k_c:k_o+k_c=k,k_o,k_c\geq 0} \mathbb{P}\left(\sum_{i=1}^{m_N} X^{(N)}_i = k_o\right)\mathbb{P}\left(\sum_{i,j=1,i<j}^{m_N} Y^{(N)}_{(i,j)} = k_c\right)
\end{aligned}$$

and hence using (3.4.8) and (3.4.9) for any $0 \leq k \leq \ell_N$ we have for an appropriate constant $C > 0$

$$\begin{aligned}
\mathbb{P}(W^{(N)}_{o,m_N^2} = k)\exp\left(-m_N^2\frac{\ell_N V_N}{D_N}\right) &\leq \mathbb{P}\left(Z^{(N)}_{g+1,\ell} = k | Z^{(N)}_{g,\ell} = m_N\right) \\
&\leq \mathbb{P}(W^{(N)}_{o,m_N^2} = k)\exp\left(-m_N^2\frac{\ell_N V_N}{D_N}\right) + \ell_N^3 c^{(N)}_c + \ell_N^2 c^{(N)}_o + \ell_N^4 c^{(N)}_c c^{(N)}_o \\
&\leq \mathbb{P}(W^{(N)}_{o,m_N^2} = k)\exp\left(-m_N^2\frac{\ell_N V_N}{D_N}\right) + \frac{C\ell_N^6 V_N}{D_N}.
\end{aligned} \tag{3.4.10}$$

Subtracting upper and lower, resp., bounds of the transition probabilities of $\mathbf{Z}^{(N)}_\ell$ from the lower and upper, resp. bounds of $\mathbb{P}(B_k)$ and taking the modulus yields (3.4.5). Indeed, by (3.4.7) and

(3.4.10) and as we can estimate

$$\begin{aligned}\mathbb{P}(W_B^{(N)} = k) &= \left(\frac{(m_n V_n - 2\ell_N)^2}{k}\right)^k \frac{1}{k!} \exp\left(-\frac{(m_N V_N - 2\ell_N)^2}{2D_N}\right) \\ &= \left(\frac{m_N^2 (V_n - 2\ell_N)^2}{2D_n}\right)^k \frac{1}{k!} \exp\left(-\frac{m_N^2 (V_n - 2\ell_N)^2}{2D_N}\right)\left(1 + O\left(\frac{\ell_N^3 V_N}{D_N}\right)\right) \\ &= \mathbf{P}(W^{(N)}_{o,m_N^2} = k)\left(1 + \mathcal{O}\left(\frac{\ell_N^3 V_N}{D_N}\right)\right),\end{aligned}$$

we have for a constant $C > 0$ that may change from line to line

$$\begin{aligned}&\mathbb{P}(Z^{(N)}_{g+1,\ell} = k | Z^{(N)}_{g,\ell} = m_N) - \mathbb{P}(B_k) \\ &\qquad \le \mathbb{P}(W^{(N)}_{o,m_N^2} = k)\left(\exp\left(-m_N^2 \frac{\ell_N V_N}{D_N}\right) - \exp\left(-\frac{\ell_N^5 V_N^3}{D_N^2}\right)\right) + \frac{C\ell_N^6 V_N}{D_N} \\ &\qquad \le \left(\frac{(m_N V_N)^2}{2D_N}\right)^k \frac{1}{k!} \exp\left(-\frac{(m_N V_N)^2}{2D_N}\right) \frac{\ell_N^4 V_N}{D_N} + \frac{C\ell_N^6 V_N}{D_N}\end{aligned}$$

since

$$\begin{aligned}\exp\left(-\frac{m_N^2 (V_N - 2\ell_N)^2}{2D_N}\right) &\le \exp\left(-\frac{m_N^2 V_N^2}{2D_N}\right) \exp\left(\frac{2m_N^2 \ell_N V_N}{D_N}\right) \\ &\le \exp\left(-\frac{m_N^2 V_N^2}{2D_N}\right)\left(1 + \frac{\ell_N^4 V_N}{D_N}\right)\end{aligned}$$

for N large enough. Furthermore, we have that

$$\begin{aligned}\mathbb{P}(W^{(N)}_{o,m_N^2} = k) &= \left(\frac{m_N^2 (V_N - 2\ell_N)^2}{2D_N}\right)^k \frac{1}{k!} \exp\left(-\frac{m_N^2 (V_N - 2\ell_N)^2}{2D_N}\right) \\ &\ge \left(\frac{m_N^2 V_N^2}{2D_N}\right)^k \left(1 - \frac{2\ell_N}{V_N}\right)^{2k} \frac{1}{k!} \exp\left(-\frac{m_N^2 (V_N - 2\ell_N)^2}{2D_N}\right) \\ &\ge \left(\frac{m_N^2 V_N^2}{2D_N}\right)^k \frac{1}{k!} \exp\left(-\frac{m_N^2 (V_N - 2\ell_N)^2}{2D_N}\right) - \frac{C\ell_N^4 V_N}{D_N},\end{aligned}$$

where we used again Bernoulli's inequality in the second inequality. Now we have for N large enough by (3.4.6) and (3.4.10)

$$\begin{aligned}&\mathbb{P}(B_k) - \mathbb{P}(Z^{(N)}_{g+1,\ell} = k | Z^{(N)}_{g,\ell} = m_N) \\ &\qquad \le \left(\frac{m_N^2 V_N^2}{2D_N}\right)^k \frac{1}{k!} \exp\left(-\frac{m_N^2 (V_N - 2\ell_N)^2}{2D_N}\right)\left(\frac{2\ell_N^2 V_N}{D_N} + \frac{6\ell_N^3 V_N}{D_N}\right) + \frac{C\ell_N^6 V_N}{D_N} \\ &\qquad \le \left(\frac{m_N^2 V_N^2}{2D_N}\right)^k \frac{1}{k!} \exp\left(-\frac{m_N^2 V_N^2}{2D_N}\right) \frac{7\ell_N^3 V_N}{D_N} + \frac{C\ell_N^6 V_N}{D_N}.\end{aligned}$$

This yields the claim. □

As a counterpart of Proposition 3.4.4, we show that the total size of the lower DBPC of Definition 3.4.5 reaches a level tending to infinity with asymptotically the survival probability of a DBPC whose offspring and cooperation distributions are respectively $\mathrm{Poi}\left(\frac{a^2}{2}\right)$ and $\mathrm{Poi}(a^2)$ distributed.

Proposition 3.4.7. *(Probability for the total size of the lower DBPC to reach a level b_N). Consider a sequence $(b_N)_{N\in\mathbb{N}}$ with $b_N \underset{N\to\infty}{\to} \infty$ and the scaling $V_N \sim a\sqrt{D_N}$. Then, we have*

$$\lim_{N\to\infty} \mathbb{P}\left(\exists g \in \mathbb{N}_0 : \overline{Z}_{g,\ell}^{(N)} \geq b_N\right) = \pi(a).$$

Proof. This Proposition is shown by the same line of argument as Proposition 3.4.4, i.e. basically one applies Lemma 3.3.7. □

Final phase of an epidemic on a complete graph

In this subsection we are going to show that if the total size of the infection process reaches the level N^ε for $\varepsilon > 0$, then in a finite number of generations, all the hosts are killed. For that we will intensively use the following Lemma. Recall that we consider $D_N \in \Theta(N^\beta)$ and $V_N \sim a\sqrt{D_N}$.

Lemma 3.4.8. *Let $\varphi_1(N), \varphi_2(N)$ and $\varphi_3(N)$ such that it holds $\varphi_i(N) = o(D_N)$ for $i \in \{1,2\}$ and $\varphi_3(N) = o(V_N)$. Let $0 < \varepsilon < \beta/2$ and $H(N)$ such that $N^\varepsilon = o(H(N))$ and $H(N) = o(\sqrt{D_N})$. Consider the following experiment: Assume $H(N)(V_N - \varphi_3(N))$ balls are distributed uniformly into $D_N - \varphi_1(N)$ boxes. Denote by $G^{(N)}$ the number of boxes among the first $D_N - \varphi_1(N) - \varphi_2(N)$ boxes that contain at least 2 balls. Then it holds:*

(i) Define $\ell := \inf\left\{i \geq 2 : H(N)^{i+1} = o\left(\sqrt{D_N}^{i-1}\right)\right\} < \infty$. Let $f_1(N)$ such that $\log(H(N)V_N) = o(f_1(N))$. Then we have

$$\mathbb{P}\left(\frac{H^2(N)}{f_1(N)} \leq G^{(N)}\right) \geq 1 - \Theta\left(\frac{H(N)^{\ell+1}}{\sqrt{D_N}^{\ell-1}}\right). \tag{3.4.11}$$

(ii) Let $f_2(N) \to \infty$. Then we have

$$\mathbb{P}\left(G^{(N)} \leq H^2(N) f_2(N)\right) \geq 1 - \Theta\left(\tfrac{1}{f_2(N)}\right). \tag{3.4.12}$$

A proof can be found in the supplementary material.
Now, we introduce for an arbitrary sequence $(b_N)_N$ with $b_N \to \infty$ the stopping times

$$\tau_{b_N}^{\overline{\mathbf{J}}} := \inf\left\{g \in \mathbb{N}_0 : \overline{J}_g^{(N)} \geq b_N\right\}, \tag{3.4.13}$$

and

$$\tau_{b_N}^{\mathbf{J}} := \inf\left\{g \in \mathbb{N}_0 : J_g^{(N)} \geq b_N\right\}. \tag{3.4.14}$$

Proposition 3.4.9. *It holds*

$$\mathbb{P}\left(\tau_{N^\varepsilon \log^{-1}(N)}^{\mathbf{J}} \leq \tau_{N^\varepsilon}^{\overline{\mathbf{J}}} \;\middle|\; \right) \underset{N\to\infty}{\to} 1.$$

Proof. If for any generation before $\tau^{\overline{\mathbf{J}}}_{N^\varepsilon}$ the number of infected vertices is strictly smaller than $\frac{N^\varepsilon}{\log(N)}$ then this would mean that the number of generations until the total size of the infection process reaches the level N^ε is at least $\log(N)$. But this contradicts the fact that it exists a constant $C > 0$ such that

$$\mathbb{P}\left(\tau^{\overline{\mathbf{J}}}_{N^\varepsilon} \leq C\log(\log(N)) \;\middle|\; \tau^{\overline{\mathbf{J}}}_{N^\varepsilon} < \infty\right) \to 1,$$

which follows from coupling from below with the DBPC of Definition 3.4.5 and Proposition 3.3.6. □

Lemma 3.4.10. *We have*

$$\mathbb{P}\Big(\overline{J}^{(N)}_{\tau^{\mathbf{J}}_{N^\varepsilon \log^{-1}(N)}} \leq \frac{N^{2\varepsilon}}{\log(N)} \;\Big|\; \tau^{\overline{\mathbf{J}}}_{N^\varepsilon} < \infty\Big) \underset{N\to\infty}{\longrightarrow} 1.$$

Proof. By definition at generation $\tau^{\mathbf{J}}_{N^\varepsilon \log^{-1}(N)}$ the number of infected vertices is at least $\frac{N^\varepsilon}{\log(N)}$. At generation $\tau^{\mathbf{J}}_{N^\varepsilon \log^{-1}(N)} - 1$, whp we have $\overline{J}^{(N)}_{\tau^{\mathbf{J}}_{N^\varepsilon \log^{-1}(N)}-1} \leq N^\varepsilon$ because otherwise we have a contradiction to Proposition 3.4.9. Then to bound from above the total number of infected vertices up to generation $\tau^{\mathbf{J}}_{N^\varepsilon \log^{-1}(N)}$, it suffices to estimate the number of infections generated within one generation, when at the beginning of this generation at most N^ε hosts are infected. This estimate is obtained by an application of Lemma 3.4.8 with $H(N) = \frac{N^\varepsilon}{\log(N)}$, $\varphi_1(N) = N^\varepsilon$, $\varphi_2(N) = 0$, $\varphi_3(N) = 0$, $f_2(N) = \frac{1}{2}\log(N)$ and an arbitrary function f_1 satisfying the condition of Lemma 3.4.8. Indeed, since before generation $\tau^{(N)}$ the total number of parasites on the graph is at most $\frac{N^\varepsilon}{\log(N)} v_N$, the number of new infections generated is controlled from above using the previous experiment. □

Next choose $0 < \varepsilon < \frac{\beta}{2}$ such that for all $k \in \mathbb{N}$, $2^k\varepsilon \neq \frac{\beta}{2}$. Then define $\overline{k}$ as the largest $k \in \mathbb{N}$ satisfying $2^{k+1}\varepsilon < \beta$. In particular it holds $2^{\overline{k}+1}\varepsilon > \frac{\beta}{2}$ because otherwise $2^{\overline{k}+2}\varepsilon < \beta$ which contradicts the definition of $\overline{k}$.

Lemma 3.4.11. *Let $k \in \{0, \dots, \overline{k}\}$. We have*

$$\mathbb{P}\Big(J^{(N)}_{\tau^{\mathbf{J}}_{N^\varepsilon \log^{-1}(N)}+k} \geq \frac{N^{2^k\varepsilon}}{\log^{\alpha_k}(N)}, \overline{J}^{(N)}_{\tau^{\mathbf{J}}_{N^\varepsilon \log^{-1}(N)}+k} \leq \frac{N^{2^{k+1}\varepsilon}}{\log(N)} \Big| \tau^{\overline{\mathbf{J}}}_{N^\varepsilon} < \infty\Big) \to 1,$$

where we set $\alpha_0 := 1$ and $\alpha_k := 2\alpha_{k-1} + 2$ for all $k \geq 1$.

Proof. We prove the claim via induction over k. For $k = 0$ the claim follows by Lemma 3.4.10. Next we prove the claim for $k+1$ assuming the claim holds for all $0 \leq j \leq k$.

For the lower bound on the number of infected vertices at generation $\tau^{\mathbf{J}}_{N^\varepsilon \log^{-1}(N)}+k+1$, Lemma 3.4.8 can be applied with $H(N) = \frac{N^{2^k\varepsilon}}{\log^{\alpha_k}(N)}$, $\varphi_1(N) = 0$, $\varphi_2(N) = \frac{N^{2^{k+1}\varepsilon}}{\log(N)}$, $\varphi_3(N) = 0$, $f_1(N) = \log^2(N)$ and an arbitrary function f_2 with $f_2(N) \to \infty$, which yields that the number of infected vertices at generation $\tau^{\mathbf{J}}_{N^\varepsilon \log^{-1}(N)} + k + 1$ is whp at least of order $\frac{1}{\log^2(N)}\left(\frac{N^{2^k\varepsilon}}{\log^{\alpha_k}(N)}\right)^2 = \frac{N^{2^{k+1}\varepsilon}}{\log^{\alpha_{k+1}}(N)}$.

Indeed by considering D_N boxes we lower bound the probability for a parasite to attack an occupied vertex, which is $\frac{1}{D_N - 1}$ in the case of the complete graph. According to the induction hypothesis we have considered whp by Lemma 3.4.8 the minimal number of parasites which is

$\frac{N^{2^k\varepsilon}}{\log^{\alpha_k}(N)}V_N$. In the balls into boxes experiment new infections are (only) counted when reaching one of the $D_N - \frac{N^{2^{k+1}\varepsilon}}{\log(N)}$ first boxes whereas in the original process there are at least this number of occupied vertices.

To arrive at the upper bound on the number of empty vertices, apply Lemma 3.4.8 with $H(N) = \frac{N^{2^{k+1}\varepsilon}}{\log(N)}$, $\varphi_1(N) = H(N)$, $\varphi_2(N) = 0$, $\varphi_3(N) = 0$, $f_2(N) = \log(N)$ and an arbitrary function f_1 that satisfies the conditions of Lemma 3.4.8, since in the previous upper bound the number of empty vertices is bounded by $\frac{N^{2^{k+1}\varepsilon}}{\log(N)}$. So according to Lemma 3.4.8 the number of empty vertices at generation $\tau^{\mathbf{J}}_{N^\varepsilon \log^{-1}(N)} + k + 1$ is whp at most of order $\log(N)\left(\frac{N^{2^{k+1}\varepsilon}}{\log(N)}\right)^2 = \frac{N^{2^{k+2}\varepsilon}}{\log(N)}$. □

Applying Lemma 3.4.11 with $k = \overline{k}$ we obtain

$$\mathbb{P}\Big(J^{(N)}_{\tau^{\mathbf{J}}_{N^\varepsilon \log^{-1}(N)}+\overline{k}} \geq \frac{N^{2^{\overline{k}}\varepsilon}}{\log^{\alpha_{\overline{k}}}(N)}, \overline{J}^{(N)}_{\tau^{\mathbf{J}}_{N^\varepsilon \log^{-1}(N)}+\overline{k}} \leq \frac{N^{2^{\overline{k}+1}\varepsilon}}{\log(N)} \Big| \tau^{\overline{\mathbf{J}}}_{N^\varepsilon} < \infty\Big) \to 1.$$

Define $\delta = \frac{1}{2}\left(2^{\overline{k}+1}\varepsilon - \frac{\beta}{2}\right) > 0$. In the next Lemma we show that at generation $\overline{k}+1$ the number of infected vertices is at least of order $N^{\frac{\beta}{2}+\delta}$.

Lemma 3.4.12. *It holds*

$$\mathbb{P}\left(J^{(N)}_{\tau^{\mathbf{J}}_{N^\varepsilon \log^{-1}(N)}+\overline{k}+1} \geq N^{\frac{\beta}{2}+\delta} \mid \tau^{\overline{\mathbf{J}}}_{N^\varepsilon} < \infty\right) \to 1.$$

Proof. Here we apply again Lemma 3.4.8 to obtain this lower bound. More precisely with the following set of parameters: $H(N) = \frac{N^{2^{\overline{k}}\varepsilon}}{\log^{\alpha_{\overline{k}}}(N)}$, $\varphi_1(N) = 0$, $\varphi_2(N) = \frac{N^{2^{\overline{k}+1}\varepsilon}}{\log(N)}$, $\varphi_3(N) = 0$, $f_1(N) = \log^2(N)$ and an arbitrary function f_2. We obtain that whp $J_{\tau^{\mathbf{J}}_{N^\varepsilon \log^{-1}(N)}+\overline{k}+1} \geq \frac{N^{2^{\overline{k}+1}\varepsilon}}{\log^{\alpha_{\overline{k}+1}}(N)} \geq N^{\frac{\beta}{2}+\delta}$, by definition of δ. □

In the next lemma we show that in one more generation whp any vertex will be reached by at least 2 parasites, in other words each of the remaining hosts gets infected whp.

Lemma 3.4.13. *We have*

$$\mathbb{P}\left(\overline{J}^{(N)}_{\tau^{\mathbf{J}}_{N^\varepsilon \log^{-1}(N)}+\overline{k}+2} = D_N \mid \tau^{\overline{\mathbf{J}}}_{N^\varepsilon} < \infty\right) \to 1.$$

Proof. We aim to show that all hosts that have not been infected so far, get infected whp in generation $\tau^{\mathbf{J}}_{N^\varepsilon \log^{-1}(N)} + \overline{k} + 2$. According to Lemma 3.4.12 we have whp $J^{(N)}_{\tau^{\mathbf{J}}_{N^\varepsilon \log^{-1}(N)}+\overline{k}+1} \geq N^{\frac{\beta}{2}+\delta}$. Hence we have whp at least $m_N := N^{\frac{\beta}{2}+\delta}V_N$ parasites that may infect the remaining hosts. So, the probability that an up to generation $\tau^{\mathbf{J}}_{N^\varepsilon \log^{-1}(N)} + \overline{k} + 2$ uninfected host gets attacked by at most one of the m_N parasites (and hence with high probability remains uninfected) can be estimated from above by

$$\left(1 - \frac{1}{D_N - 1}\right)^{m_N} + \left(1 - \frac{1}{D_N - 1}\right)^{m_N - 1} \frac{m_N}{D_N - 1} = \Theta\left(N^\delta \exp\left(-aN^\delta\right)\right),$$

because

$$\frac{m_N}{D_N} = \frac{N^{\beta/2+\delta}V_N}{D_N} \sim \frac{N^{\beta/2}N^{\delta}a\sqrt{D_N}}{D_N} = aN^{\delta}\frac{N^{\beta/2}}{\sqrt{D_N}} = \Theta(N^{\delta}).$$

The number of uninfected hosts at the beginning of generation $\tau^{\mathbf{J}}_{N^{\varepsilon}\log^{-1}(N)} + \overline{k} + 2$ is at most D_N. Consequently, the probability that at least one of these hosts remains uninfected till the end of generation $\tau^{\mathbf{J}}_{N^{\varepsilon}\log^{-1}(N)} + \overline{k} + 2$ can be estimated from above by a probability proportional to

$$D_N N^{\delta}\exp\left(-aN^{\delta}\right) = o(1),$$

which yields the claim of Lemma 3.4.13. □

3.4.3 Proof of Theorem 3.2.6 (ii)

Now we have all necessary ingredients to prove Theorem 3.2.6 (ii).
The first step is to show

$$\limsup_{N\to\infty}\mathbb{P}\big(F_u^{(N)}\big) \le \pi(a). \tag{3.4.15}$$

For a sequence $(\ell_N)_{N\in\mathbb{N}}$ introduce the event

$$A_{\ell_N} := \left\{\exists g \in \mathbb{N}_0 : \overline{J}_g^{(N)} \ge \ell_N\right\}.$$

Then it follows that for all $0 < u \le 1$ and any sequence $\ell_N \le uD_N$ we have

$$\mathbb{P}\big(F_u^{(N)}\big) \le \mathbb{P}\big(A_{\ell_N}\big). \tag{3.4.16}$$

Taking a sequence ℓ_N satisfying both $\ell_N \to \infty$ and $\frac{\ell_N^7 V_N^3}{D_N^2} \in o(1)$ we have by Proposition 3.4.3 that

$$\mathbb{P}\big(\mathcal{A}_{\ell_N}\big) \le \mathbb{P}\left(\exists g \in \mathbb{N}_0 : \overline{Z}_{g,u}^{(N)} \ge \ell_N\right) + o(1). \tag{3.4.17}$$

Proposition 3.4.4 gives that

$$\lim_{N\to\infty}\mathbb{P}\left(\exists g \in \mathbb{N}_0 : \overline{Z}_{g,u}^{(N)} \ge \ell_N\right) = \pi(a). \tag{3.4.18}$$

In summary combining (3.4.16), (3.4.17) and (3.4.18) gives exactly (3.4.15).

The second step is to show

$$\liminf_{N\to\infty}\mathbb{P}\big(F_u^{(N)}\big) \ge \pi(a). \tag{3.4.19}$$

Proposition 3.4.6 combined with Proposition 3.4.7 gives that

$$\liminf_{N\to\infty}\mathbb{P}\left(\exists g \in \mathbb{N}, \overline{J}_g^{(N)} \ge N^{\varepsilon}\right) \ge \pi(a), \tag{3.4.20}$$

for $\varepsilon > 0$ small enough. Then Lemma 3.4.13 yields that conditioned on the event $\left\{\exists g \in \mathbb{N}, \overline{J}_g^{(N)} \ge N^{\varepsilon}\right\}$ whp all the vertices on the graph finally get infected. Combined with (3.4.20) the claim of (3.4.19) follows.

3.4.4 Proof of Theorem 3.2.6(i)

In this subsection we prove Theorem 3.2.6 (i). Recall that in this case $V_N \in o(\sqrt{D_N})$.

We initially start with one individual, i.e. $J_0^{(N)} = 1$. We determine the probability that the parasite population gets extinct after one generation. For that we consider the following experiment, where we distribute uniformly at random and independently V_N balls into $D_N - 1$ boxes. The probability of extinction after one generation is the same as the probability of the event that all boxes contain at most one ball. Thus, we get that

$$\mathbb{P}\big(J_1^{(N)} = 0\big) = \prod_{i=0}^{V_N-1} \left(1 - \frac{i}{D_N - 1}\right) \geq \exp\left(-\frac{1}{D_N - 1} \sum_{i=0}^{V_N-1} i\right) \geq \exp\left(-\frac{V_N^2}{2(D_N - 1)}\right).$$

We assumed that $V_N \in o(\sqrt{D_N})$ which implies that $\frac{V_N^2}{2(D_N-1)} \to 0$ as $N \to \infty$, and thus the right hand side converges to 1. On the other hand for any $u \in (0, 1]$ an obvious upper bound for the invasion probability is $\mathbb{P}\big(F_u^{(N)}\big) \leq 1 - \mathbb{P}\big(J_1^{(N)} = 0\big)$. This implies that

$$\lim_{N\to\infty} \mathbb{P}\big(F_u^{(N)}\big) \leq 1 - \lim_{N\to\infty} \mathbb{P}\big(J_1^{(N)} = 0\big) = 0.$$

3.4.5 Proof of Theorem 3.2.6 (iii)

In this subsection we are going to prove Theorem 3.2.6 (iii). In this case $\sqrt{D_N} \in o(V_N)$. The proof is based on a coupling from below of the total size of the infection process with the total size of a Galton-Watson process whose offspring distribution is close to a $\mathrm{Poi}\big(\frac{a^2}{2}\big)$ distribution until a level N^α is reached, where $0 < \alpha < \beta$, or until the process dies out. This coupling is possible for any $a > 0$ which yields that the total size of the infection process reaches the level N^α with asymptotically probability 1. Then by choosing $\alpha > \beta/2$ one shows that there exists a generation in which there are at least $N^{\widetilde{\alpha}}$ infected individuals, for some $\widetilde{\alpha} > \beta/2$. In the subsequent generation, all remaining hosts are infected, in the same manner as in Subsection 3.4.2.

We will show

$$\lim_{N\to\infty} \mathbb{P}\left(\exists g \in \mathbb{N}_0 : \overline{J}_g^{(N)} = D_N\right) = 1,$$

which yields the claim of Theorem 3.2.6 (iii).

Recall that we denoted in the setting of the complete graph the infected hosts by $\mathcal{J}^{(N)}$. Analogously as in the setting of the random geometric graph we denote by $\mathcal{S}^{(N)}$ the uninfected hosts and by $\mathcal{R}^{(N)}$ the dead hosts. The first step is to couple $\big(\mathcal{S}^{(N)}, \mathcal{J}^{(N)}, \mathcal{R}^{(N)}\big)$ to a process $\big(\widetilde{\mathcal{S}}^{(N)}, \widetilde{\mathcal{J}}^{(N)}, \widetilde{\mathcal{R}}^{(N)}\big)$ in which infections are only generated by pairs of parasites originating from the same vertex, but not if a host gets infected only by parasites stemming from different vertices.

For every vertex x we only need to determine once to which neighbours the V_N offspring parasites move, since afterwards the vertex cannot be used anymore. We denote by $\mathcal{H}_x^{(N)} \subset \{1, \ldots, D_N\} \backslash \{x\}$ the set of all vertices which are occupied by at least two or more of the V_N offspring parasites generated on x after their movement. With this we can build the coupling of the two processes step by step. We consider for both processes the initial configuration where only vertex 1 is currently infected and all other vertices are susceptible, i.e.

$$\big(\mathcal{S}_0^{(N)}, \mathcal{J}_0^{(N)}, \mathcal{R}_0^{(N)}\big) = \big(\widetilde{\mathcal{S}}_0^{(N)}, \widetilde{\mathcal{J}}_0^{(N)}, \widetilde{\mathcal{R}}_0^{(N)}\big) = \big(\{2, \ldots, D_N\}, \{1\}, \emptyset\big).$$

Then assume that we constructed the process until generation $g \geq 0$. Then from g to $g+1$ the dynamics are as follows

$$\widetilde{\mathcal{J}}_{g+1}^{(N)} = \bigcup_{x \in \widetilde{\mathcal{J}}_g^{(N)}} \mathcal{H}_x^{(N)} \setminus (\widetilde{\mathcal{I}}_g^{(N)} \cup \widetilde{\mathcal{R}}_g^{(N)}),$$
$$\widetilde{\mathcal{S}}_{g+1}^{(N)} = \widetilde{\mathcal{S}}_g^{(N)} \setminus \widetilde{\mathcal{J}}_{g+1}^{(N)} \quad \text{and} \quad \widetilde{\mathcal{R}}_{g+1}^{(N)} = \widetilde{\mathcal{R}}_g^{(N)} \cup \widetilde{\mathcal{J}}_g^{(N)}.$$

In words every vertex $y \in \mathcal{H}_x^{(N)}$ which is attacked by at least two parasites that are originating from a single vertex $x \in \widetilde{\mathcal{J}}_g^{(N)}$ is added to $\widetilde{\mathcal{J}}_{g+1}^{(N)}$, except for vertices which were already attacked at a previous generation, i.e. $y \in \widetilde{\mathcal{J}}_g^{(N)} \cup \widetilde{\mathcal{R}}_g^{(N)}$. Furthermore, all previously infected hosts $\widetilde{\mathcal{J}}_g^{(N)}$ are declared as removed and all vertices which were infected in this generation $\widetilde{\mathcal{J}}_{g+1}^{(N)}$ are removed from the set of susceptible vertices.

In the process $\left(\mathcal{S}^{(N)}, \mathcal{J}^{(N)}, \mathcal{R}^{(N)}\right)$ cooperation from different infected vertices for the spread of the epidemic is allowed. Since we defined movement of parasites independent from the generation at which vertices get infected, we have by construction that

$$\widetilde{\mathcal{J}}_g^{(N)} \cup \widetilde{\mathcal{R}}_g^{(N)} \subset \mathcal{J}_g^{(N)} \cup \mathcal{R}_g^{(N)}, \forall g \in \mathbb{N}_0, \tag{3.4.21}$$

almost surely. As by cooperation only more infections are generated, it is not possible that a vertex x which is susceptible for both processes at a generation g gets infected at generation $g+1$ for the process $\left(\widetilde{\mathcal{S}}^{(N)}, \widetilde{\mathcal{J}}^{(N)}, \widetilde{\mathcal{R}}^{(N)}\right)$ but not for the process $\left(\mathcal{S}^{(N)}, \mathcal{J}^{(N)}, \mathcal{R}^{(N)}\right)$.

The infection process $\left(\widetilde{\mathcal{S}}^{(N)}, \widetilde{\mathcal{J}}^{(N)}, \widetilde{\mathcal{R}}^{(N)}\right)$ is monotone with respect to the parameter V_N, in contrast to the original process $\left(\mathcal{S}^{(N)}, \mathcal{J}^{(N)}, \mathcal{R}^{(N)}\right)$. Now let $a > 0$ and consider $V_N^{(a)} = a\sqrt{D_N}$ as well as $\left(\widetilde{\mathcal{S}}^{(N,a)}, \widetilde{\mathcal{J}}^{(N,a)}, \widetilde{\mathcal{R}}^{(N,a)}\right)$ to be the analogously defined infection process. Infections are only generated by pairs of parasites originating from the same vertex as well as the number of parasites generated at an infection event is $V_N^{(a)}$. Since we assume that $D_N \in o(V_N^2)$ it follows for N large enough that $V_N^{(a)} \leq V_N$. Thus, by monotonicity it follows that we can couple $\left(\widetilde{\mathcal{S}}^{(N,a)}, \widetilde{\mathcal{J}}^{(N,a)}, \widetilde{\mathcal{R}}^{(N,a)}\right)$ and $\left(\widetilde{\mathcal{S}}^{(N)}, \widetilde{\mathcal{J}}^{(N)}, \widetilde{\mathcal{R}}^{(N)}\right)$, such that

$$\left|\widetilde{\mathcal{R}}_g^{(N,a)} \cup \widetilde{\mathcal{J}}_g^{(N,a)}\right| \leq \left|\widetilde{\mathcal{R}}_g^{(N)} \cup \widetilde{\mathcal{J}}_g^{(N)}\right|. \tag{3.4.22}$$

For the sequence of processes $\left(\widetilde{\mathcal{S}}^{(N,a)}, \widetilde{\mathcal{J}}^{(N,a)}, \widetilde{\mathcal{R}}^{(N,a)}\right)$ we can show (by a coupling with Galton-Watson processes) that the probability to infect eventually N^α host is asymptotically lower bounded by the survival probability φ_a of a Galton-Watson process with $\mathrm{Poi}\left(\frac{a^2}{2}\right)$ offspring distribution. The proof of this statement can be found in the proof of Lemma 3.6.4, where this statement is formulated, in the supplementary material (since it can be shown by very similar arguments that have been used to show Proposition 4.7 combined with Lemma 3.7) in [1].

Because this result is true for any $a > 0$, taking the limit when a goes to ∞ gives, together with (3.4.22) and (3.4.21),

$$\lim_{N \to \infty} \mathbb{P}\left(\exists g \in \mathbb{N}_0 : \overline{J}_g^{(N)} \geq N^\alpha\right) = 1. \tag{3.4.23}$$

Now let $\frac{\beta}{2} < \alpha < \beta$ and recall the definition of the stopping times defined in (3.4.13) and (3.4.14). Then one can show as in Proposition 3.4.9 that

$$\mathbb{P}\left(\tau_{N^\alpha \log^{-2}(N)}^{\mathbf{J}} \leq \tau_{N^\alpha}^{\overline{\mathbf{J}}} \mid \tau_{N^\alpha}^{\overline{\mathbf{J}}} < \infty\right) \to 1.$$

Indeed, if for any generation before $\tau^{\overline{\mathbf{J}}}_{N^\alpha}$ the number of infected vertices is strictly smaller than $\frac{N^\alpha}{\log^2(N)}$ then this means that the number of generations to reach the level N^α for the total size of the infection process is at least $\log^2(N)$. But this is in contradiction with the couplings of (3.4.21) and (3.6.13) and Lemma 5.5 from [1].
Then using a similar approach as in the proof of Lemma 3.4.13, one shows that

$$\mathbb{P}\left(\overline{J}^{(N)}_{\tau^{\mathbf{J}}_{N^\alpha \log^{-2}(N)+1}} = D_N \mid \tau^{\overline{\mathbf{J}}}_{N^\alpha} < \infty\right) \to 1. \tag{3.4.24}$$

Finally combining (3.4.23) and (3.4.24) it follows that

$$\mathbb{P}\left(\exists g \in \mathbb{N}_0 : \overline{J}^{(N)}_g = D_N\right) \to 1,$$

which completes the proof.

3.5 Invasion on a random geometric graph

Let us recall the setting of Theorem 3.2.4. Recall that we denote by $\mathcal{G}^{(N)} = (\mathcal{V}^{(N)}, \mathcal{E}^{(N)})$ a random geometric graph, where the vertices are given by a homogeneous Poisson point process on $[0,1]^n$ with intensity measure $N\lambda^n(\,\cdot\,)$ and vertices share an edge if they are at a distance less than $r_N = \frac{1}{2}N^{\frac{\beta-1}{n}}$ apart. Thus, every vertex (sufficiently far from the boundary) has on average $d_N = N^\beta$ direct neighbours. In case of an host infection v_N parasites are generated. We denote the hosts which were infected in generation g by $\mathcal{I}^{(N)}_g \subset \mathcal{V}^{(N)}$, the number of infected hosts in generation g by $I_g = I^{(N)}_g$ and the total number until generation g by $\overline{I}_g = \overline{I}^{(N)}_g$.

Before we start to study the behaviour of the host-parasite infection, we show some properties of $\mathcal{G}^{(N)}$, which we will need in the subsequent section. We show that $\mathcal{G}^{(N)}$ is whp connected and is fairly dense in the sense that the number of vertices contained in every ball of radius r_N is of order N^β. We denote by $B_r(x)$ the ball of radius r around $x \in [0,1]^n$ with respect to the maximum norm ρ.

Lemma 3.5.1. *Let $0 < \alpha \le \beta$, $\gamma \in \left(\frac{2}{n+2}\alpha, \alpha\right)$ and $\delta \in (0,1]$.*

1. *The graph $\mathcal{G}^{(N)} = (\mathcal{V}^{(N)}, \mathcal{E}^{(N)})$ is connected whp as $N \to \infty$.*
2. *Set $\varepsilon_N = \varepsilon_N(\alpha) := \frac{1}{2}N^{\frac{\alpha-1}{n}}$ and $\boldsymbol{Vol}_x(\delta\varepsilon_N) := vol(B_{\delta\varepsilon_N}(x) \cap [0,1]^n) \in \Theta(N^{\alpha-1})$ and*

$$\mathcal{U}_N(\alpha) := \left\{\left||\mathcal{V}^{(N)} \cap B_{\delta\varepsilon_N}(x)| - N\,\boldsymbol{Vol}_x(\delta\varepsilon_N)\right| \le \delta^{n-1}(n+1)N^{\frac{(n-1)\alpha+\gamma}{n}}\ \forall\, x \in [0,1]\right\}. \tag{3.5.1}$$

It holds that $\lim_{N\to\infty}\mathbb{P}\left(\mathcal{U}_N(\alpha)\right) = 1$.

Proof of Lemma 3.5.1. Choose $0 < \gamma < \beta$ and $0 < \varepsilon < \gamma/2$. The idea of the proof is to define disjoint balls $K(l)$ for $l = (l_1, \dots, l_n) \in \mathbb{N}^n$ with side length $N^{(\gamma-1)/n}$ which cover the whole unit ball, i.e. $[0,1]^n \subset \bigcup K(l)$. In the second step we gain control on the asymptotic number of Poisson points contained in every ball simultaneously, i.e. we will show with the help of Lemma 3.6.6 that every ball contains $N^\gamma \pm N^{\gamma/2+\varepsilon}$ many points with high probability. A technical problem is that we defined our Poisson point set $\mathcal{V}^{(N)}$ only on $[0,1]^n$. Not for every N are we able to perfectly cover

the unit ball with our balls $K(l)$ such that $[0,1]^n = \bigcup K(l)$. Thus, we need to extend our Poisson point set. This can be easily done by sampling independent Poisson points with intensity measure $N\mathrm{d}x$ on $[0,2]^n \backslash [0,1]^n$. We denote this Poisson point set by $\mathcal{V}'_N$. Now we set $\mathcal{V}''_N := \mathcal{V}^{(N)} \cup \mathcal{V}'_N$, so $\mathcal{V}''_N$ is a Poisson point set on $[0,2]^n$ with intensity measure $N\mathrm{d}x$.

Let us set $\overline{M} := \left\{0, \dots, \lceil N^{\frac{1-\gamma}{n}} \rceil\right\}^n$ and $\underline{M} := \left\{0, \dots, \lfloor N^{\frac{1-\gamma}{n}} \rfloor\right\}^n$. Define balls of edge length $N^{\frac{\gamma-1}{n}}$ by setting $K(l) := [0, N^{\frac{\gamma-1}{n}})^n + N^{\frac{\gamma-1}{n}} l$, where $l \in \overline{M}$. Set $\overline{k} := |\overline{M}|$ and $\underline{k} := |\underline{M}|$. For these balls we have

$$\bigcup_{l \in \underline{M}} K(l) \subset [0,1]^n \subset \bigcup_{l \in \overline{M}} K(l).$$

Set $X_N := |K(\mathbf{0}) \cap \mathcal{V}''_N|$, where $\mathbf{0} = (0, \dots, 0) \in \mathbb{R}^n$ then

$$\mathbb{P}\Big(\bigcap_{l \in \overline{M}} \{N^\gamma - N^{\frac{\gamma}{2}+\varepsilon} \le |B(l) \cap \mathcal{V}''_N| \le N^\gamma + N^{\frac{\gamma}{2}+\varepsilon}\}\Big) = \mathbb{P}\Big(N^{-\frac{\gamma}{2}-\varepsilon} |X_N - N^\gamma| \le 1\Big)^{\overline{k}}.$$

According to Lemma 3.6.6, where we control the size of Poisson random variables via moderate deviations,

$$\lim_{n\to\infty} -\frac{2}{N^{2\varepsilon}} \log\big(\mathbb{P}\big(N^{-\frac{\gamma}{2}-\varepsilon}|X_N - N^\gamma| > 1\big)\big) = 1.$$

This implies that

$$\log\big(\mathbb{P}\big(N^{-\frac{\gamma}{2}-\varepsilon}|X_N - N^\gamma| > 1\big)\big) = -\frac{N^{2\varepsilon}}{2}(1 + h(N^{2\varepsilon}))$$

where $h(x) \in o(1)$ as $x \to \infty$. Since $\overline{k} = \lceil N^{\frac{1-\gamma}{n}} \rceil^n$ with Bernoulli's inequality

$$\Big(1 - \mathbb{P}\big(N^{-\frac{\gamma}{2}-\varepsilon}|X_N - N^\gamma| > 1\big)\Big)^{\overline{k}} \ge 1 - \lceil N^{\frac{1-\gamma}{n}} \rceil^n \exp\big(-\frac{N^{2\varepsilon}}{2}(1 + h(N^{2\varepsilon}))\big) \to 1$$

as $N \to \infty$. Thus, we have shown that all balls $(K(l))_{l \in \overline{M}}$ simultaneously contain with high probability $N^\gamma \pm N^{\frac{\gamma}{2}+\varepsilon}$ many Poisson points as $N \to \infty$.

1. The first claim is a direct consequence of what we just showed. Let $l \in \underline{M}$, and consider a ball $K(l) \subset [0,1]^n$, then it follows that every vertex $x \in \mathcal{V}^{(N)}$ contained in $K(l)$ is connected to every other vertex contained in the same ball $K(l)$ since $\gamma < \beta$. This means that the vertices in a ball $K(l)$ form a complete graph for every $l \in \underline{M}$. Furthermore, for N large enough it holds that $2N^{\frac{\gamma-1}{n}} < N^{\frac{\beta-1}{n}}$, and thus every vertex contained in a ball $K(l)$ is connected to every vertex contained in all adjacent balls $K(l')$. Thus, we have shown that the random geometric graph with vertex set $\mathcal{V}^{(N)} \cap \bigcup_{l \in \underline{M}} K(l)$ forms a connected graph with high probability.

   It remains to argue that every vertex $x \in \mathcal{V}^{(N)} \cap [0,1]^n \backslash \bigcup_{l \in \underline{M}} K(l)$ is connected to its neighbouring ball. Since $\gamma < \beta$ it holds $B_{r_N}(x) \cap \bigcup_{l \in M} K(l) \neq \emptyset$ for a vertex $x \in \mathcal{V}^{(N)} \cap [0,1]^n \backslash \bigcup_{l \in \underline{M}} K(l)$. Hence, for N large enough these vertices are connected to its closest ball $K(l)$ with high probability, since with high probability every ball $K(l)$ is non-empty for $N \to \infty$.

2. Now we consider the hyperrectangle $B_{\delta\varepsilon_N}(x)\cap[0,1]$, where $x\in[0,1]^n$, and we denote its edge lengths by $(\kappa_i)_{1\le i\le n}$. Note that $\frac{\delta}{2}N^{\frac{1-\alpha}{n}}\le\kappa_i\le\delta N^{\frac{1-\alpha}{n}}$, and thus it holds for the volume

$$(\tfrac{\delta}{2})^n N^{1-\alpha}\le\mathbf{Vol}_x(\delta\varepsilon_N)\le\delta^n N^{1-\alpha}.$$

Furthermore, the set $B_{\delta\varepsilon_N}(x)\cap[0,1]$, contains at least $\prod_{i=1}^n\lfloor\kappa_i N^{\frac{1-\gamma}{n}}\rfloor^n$ many balls with edge length $N^{\frac{1-\gamma}{n}}$. This means that whp the set $B_{\delta\varepsilon_N}(x)\cap[0,1]$ contains at least $\big(\prod_{i=1}^n\lfloor\kappa_i N^{\frac{1-\gamma}{n}}\rfloor^n\big)(N^\gamma-N^{\frac{\gamma}{2}+\varepsilon})$ many vertices.

$$\begin{aligned}\prod_{i=1}^n\lfloor\kappa_i N^{\frac{1-\gamma}{n}}\rfloor^n(N^\gamma-N^{\frac{\gamma}{2}+\varepsilon})&\ge\prod_{i=1}^n\big(\kappa_i N^{\frac{1-\gamma}{n}}-1\big)(N^\gamma-N^{\frac{\gamma}{2}+\varepsilon})\\&\ge N\mathbf{Vol}_x(\delta\varepsilon_N)-\delta^{n-1}nN^{\frac{(n-1)\alpha+\gamma}{n}}+R_-(N),\end{aligned}$$

where $R_-(N)=-\delta^nN^{\alpha-\frac{\gamma}{2}+\varepsilon}+o(N^{\alpha-\frac{\gamma}{2}+\varepsilon})$. Note that we used that $\kappa_i\le\delta N^{\frac{1-\alpha}{n}}$ and that $\mathbf{Vol}_x(\delta\varepsilon_N)$ is of order $N^{1-\alpha}$. Since it holds

$$\alpha-\frac{\gamma}{2}<\frac{(n-1)\alpha+\gamma}{n}\iff\frac{2}{(n+2)}\alpha\le\gamma$$

for all $n\ge1$ we can choose ε small enough such that $R_-(N)$ consists only of lower order terms with the leading order term having a negative sign. This means that for N large enough it follows that

$$\prod_{i=1}^n\lfloor\kappa_i N^{\frac{1-\gamma}{n}}\rfloor^n(N^\gamma-N^{\frac{\gamma}{2}+\varepsilon})\ge N\mathbf{Vol}_x(\delta\varepsilon_N)-\delta^{n-1}(n+1)N^{\alpha-\frac{\gamma}{2}+\varepsilon}.\tag{3.5.2}$$

On the other hand $B_{\delta\varepsilon_N}(x)\cap[0,1]$ can be covered by $\prod_{i=1}^n\big(\lfloor\kappa_i N^{\frac{1-\gamma}{n}}\rfloor+1\big)$ many balls. Thus, we obtain similarly as before that

$$\prod_{i=1}^n\big(\lfloor\kappa_i N^{\frac{1-\gamma}{n}}\rfloor+1\big)(N^\gamma+N^{\frac{\gamma}{2}+\varepsilon})\le N\mathbf{Vol}_x(\delta\varepsilon_N)+\delta^{n-1}nN^{\frac{(n-1)\alpha+\gamma}{n}}+R_+(N),$$

where $R_+(N)=\delta^nN^{\alpha-\frac{\gamma}{2}+\varepsilon}+o(N^{\alpha-\frac{\gamma}{2}+\varepsilon})$ and we used one more time that $\kappa_i\le\delta N^{\frac{1-\alpha}{n}}$ and that $\mathbf{Vol}_x(\delta\varepsilon_N)$ is of order $N^{1-\alpha}$. Again for N large enough we get that

$$\prod_{i=1}^n\big(\lfloor\kappa_i N^{\frac{1-\gamma}{n}}\rfloor+1\big)(N^\gamma+N^{\frac{\gamma}{2}+\varepsilon})\le\delta^nN\mathbf{Vol}_x(\delta\varepsilon_N)+\delta^{n-1}(n+1)N^{\frac{(n-1)\alpha+\gamma}{n}}.\tag{3.5.3}$$

Now (3.5.2) and (3.5.3) imply that

$$\big||\mathcal{V}^{(N)}\cap B_{\delta\varepsilon_N}(x)|-N\mathbf{Vol}_x(\delta\varepsilon_N)\big|\le\delta^{n-1}(n+1)N^{\frac{(n-1)\alpha+\gamma}{n}}.$$ □

Remark 3.5.2. *Under the event $\mathcal{U}_N(\alpha)$ the graph $\mathcal{G}^{(N)}$ is connected.*

Remark 3.5.3. *The optimal choice of γ to minimize the order of the error term is to choose γ close to $\frac{2}{2+n}\alpha$, which leads to an order close to*

$$\frac{(n-1)\alpha+\gamma}{n} = \Big(\frac{n+1}{n+2}\Big)\alpha.$$

But the result of Lemma 3.5.1 does not allow for this choice. Thus, one reasonable choice would for example be

$$\gamma = \frac{3}{3+n}\alpha > \frac{2}{2+n}\alpha,$$

then we get that the order of the error term is

$$\frac{(n-1)\alpha+\gamma}{n} = \Big(\frac{n+2}{n+3}\Big)\alpha,$$

which yields for $n=1$ the value $\frac{3}{4}\alpha$.

For our approach we need a quite subtle control of the number of hosts on two different spatial scales. To be precise we need to have control on how many hosts an arbitrary ball of radius $\varepsilon_N(\frac{\beta}{2}+\xi) = \frac{1}{2}N^{\frac{\beta/2+\xi-1}{n}}$ has, where $\xi \in (0, \frac{\beta}{4})$, as well as on how many hosts a ball of radius $\frac{1}{2}r_N$ contains. The need for this will only become apparent in Lemma 3.5.15.

Thus, we only consider realisations of the graph $\mathcal{G}^{(N)}$ which are contained in the event $\mathcal{U}_N(\beta/2+\xi)$ which was defined in Lemma 3.5.1. Note that we will use the choice of γ discussed in Remark 3.5.3, that is $\gamma = \frac{3}{3+n}\alpha$. Since Lemma 3.5.1 yields that $\mathbb{P}(\mathcal{U}_N(\beta/2+\xi)) \to 1$ as $N \to \infty$ this is justified. For these realisations we know that for every $x \in [\varepsilon_N, 1-\varepsilon_N]^n$ we have that

$$\big||\mathcal{V}^{(N)} \cap B_{\varepsilon_N}(x)| - N^{\frac{\beta}{2}+\xi}\big| \le (n+1)N^{\frac{n+2}{n+3}\left(\frac{\beta}{2}+\xi\right)}.$$

Thus, we have control on the number of vertices on fairly fine spatial scale. This allows us to gain control on a courser scale as well. By calculations similar as (3.5.2) and (3.5.3) in the proof of Lemma 3.5.1 one can show that there exist $C(n) > 0$ and $c(n) \in (0,1)$ such that for all $x \in [\frac{r_N}{2}, 1-\frac{r_N}{2}]^n$ it holds that

$$\big||\mathcal{V}^{(N)} \cap B_{2^{-1}r_N}(x)| - 2^{-n}N^{\beta}\big| \le C(n)N^{c(n)\beta}.$$

Remark 3.5.4. *A precise calculation would yield that the two constants are explicitly given by*

$$C(n) = 2^{-(n-1)}(n+2) \quad \text{and} \quad c(n) = \frac{(2n+5)}{2(n+3)}.$$

One can see that the error for the courser scale would be better by a direct application of Lemma 3.5.1, but this would not give us the control on both spatial scales simultaneously.

In Subsections 3.5.1 and 3.5.2 we lay the groundwork to show Theorem 3.2.4 1) (ii) and 2), and thus we consider the critical scaling $v_N \sim a\sqrt{d_N}$ in these subsections.

3.5.1 Upper bound on the invasion probability

To derive an upper bound on the invasion probability we couple whp the total number of (currently and previously) infected hosts from above with the total size of a DBPC whose offspring and cooperation laws are approximately Poisson distributed until the DBPC dies out or reaches at least the level ℓ_N, for a well chosen ℓ_N.

Let

$$\delta_{N,\ell} := N^{\beta} - N^{\beta-\varepsilon'}, \tag{3.5.4}$$

$$\delta_{N,u} := N^{\beta} + N^{\beta-\varepsilon'} \tag{3.5.5}$$

for some $0 < \varepsilon' < \beta$. According to Lemma 3.5.1 whp every ball of radius r_N fully contained in $[0,1]^n$ contains at least $\delta_{N,\ell}$ and at most $\delta_{N,u}$ vertices $x \in \mathcal{V}^{(N)}$.

Definition 3.5.5. *(Upper DBPC)*
Let $\ell_N \underset{N\to\infty}{\to} \infty$ satisfying $\ell_N \in o(\log\log N)$. Let $\mathbf{Z}_u^{(N)} = \left(Z_{g,u}^{(N)}\right)_{g\in\mathbb{N}_0}$ be a branching process with cooperation with $Z_{0,u}^{(N)} = 1$ almost surely, and offspring and cooperation distributions with probability weights $p_{u,o}^{(N)} = \left(p_{j,u,o}^{(N)}\right)_{j\in\mathbb{N}_0}$ and $p_{u,c}^{(N)} = \left(p_{j,u,c}^{(N)}\right)_{j\in\mathbb{N}_0}$, respectively with

$$p_{j,u,o}^{(N)} := \left(\frac{(v_N - \ell_N^2)^2}{2\delta_{N,u}}\left(1 - 3N^{-\varepsilon'} - \frac{2^{\ell_N}\ell_N^2}{\delta_{N,\ell}}\right)\right)^j \frac{1}{j!}\exp\left(-\frac{v_N^2}{2\delta_{N,u}}\right)$$

for all $0 \le j \le \ell_N$ and

$$p_{\ell_N+1,u,o}^{(N)} := 1 - \sum_{j=0}^{\ell_N} p_{j,u,o}^{(N)},$$

as well as

$$p_{j,u,c}^{(N)} := \left(\frac{(v_N - \ell_N^2)^2}{\delta_{N,u}}\left(1 - 3N^{-\varepsilon'} - \frac{2^{\ell_N}\ell_N^2}{\delta_{N,\ell}}\right)\right)^j \frac{1}{j!}\exp\left(-\frac{v_N^2}{\delta_{N,u}}\right)$$

for all $0 \le j \le \ell_N$ and

$$p_{\ell_N+1,u,c}^{(N)} := 1 - \sum_{j=0}^{\ell_N} p_{j,u,c}^{(N)}.$$

Denote by $\overline{\mathbf{Z}}_u^{(N)} := \left(\overline{Z}_{g,u}^{(N)}\right)_{g\in\mathbb{N}_0}$ where $\overline{Z}_{g,u}^{(N)} := \sum_{i=0}^{g} Z_{i,u}^{(N)}$, that is $\overline{Z}_{g,u}^{(N)}$ gives the total size of $\mathbf{Z}_u^{(N)}$ accumulated till generation g.

Proposition 3.5.6. *(Probability that the total size of the upper DBPC reaches a level b_N).*
Consider a sequence $(b_N)_{N\in\mathbb{N}}$ with $b_N \underset{N\to\infty}{\to} \infty$ and assume that $v_N \sim a\sqrt{d_N}$ for $0 < a < \infty$. Then, we have

$$\lim_{N\to\infty} \mathbb{P}\left(\exists g \in \mathbb{N}_0 : \overline{Z}_{g,u}^{(N)} \ge b_N\right) = \pi(a).$$

Proof. This proposition is shown by the same line of argument as Proposition 3.4.4, i.e. basically one applies Lemma 3.3.7. □

Proposition 3.5.7. *Consider a sequence* $(\ell_N)_{N\in\mathbb{N}}$ *fulfilling* $\ell_N \underset{N\to\infty}{\to} \infty$ *and* $\ell_N \in o(\log\log(N))$. *Introduce the stopping time*

$$\tau^{\overline{\mathbf{Z}}_u}_{\ell_N,0} := \inf\left\{g \in \mathbb{N}_0 : \overline{Z}^{(N)}_{g,u} \geq \ell_N \text{ or } Z^{(N)}_{g,u} = 0\right\}.$$

Then

$$\lim_{N\to\infty} \mathbb{P}\Big(\overline{I}^{(N)}_g \leq \overline{Z}^{(N)}_{g,u}, \forall g < \tau^{\overline{\mathbf{Z}}_u}_{\ell_N,0}\Big) = 1$$

and

$$\lim_{N\to\infty} \mathbb{P}\left(I^{(N)}_{\tau^{\overline{\mathbf{Z}}_u}_{\ell_N,0}} = 0 \;\middle|\; Z^{(N)}_{\tau^{\overline{\mathbf{Z}}_u}_{\ell_N,0}} = 0\right) = 1.$$

Proof. For the proof we couple the infection process with another model, that uses the same infection rules but assumes that every generation empty vertices are reoccupied by an host. This increases only the number of infections when one assumes (as we will do here) that the movement of the parasites is independent of the generation in which they are generated. Denote by $\tilde{I}^{(N)} = (\tilde{I}^{(N)}_g)_{g\in\mathbb{N}_0}$ the corresponding process that counts the number of infections generated in this modified model. We have $I^{(N)}_g \leq \tilde{I}^{(N)}_g$ for all $g \in \mathbb{N}_0$ a.s. Next we show that $\tilde{I}^{(N)}_g \leq Z^{(N)}_{g,u}$ whp for all $g < \tau^{\overline{\mathbf{Z}}_u}_{\ell_N,0}$. We say that in generation g we have k $\tilde{I}^{(N)}$ infections, if $\tilde{I}^{(N)}_g = k$ and we say that in generation g we have k $\mathbf{Z}^{(N)}_u$ infections, if $Z^{(N)}_{g,u} = k$. Start with generation $g = 0$. Since initially only a single vertex is infected, in the first generation only CoSame infections are possible. As in [1] we can couple $\tilde{I}^{(N)}_1$ with $Z^{(N)}_{1,u}$, such that $\mathbb{P}\left(\tilde{I}^{(N)}_1 \leq Z^{(N)}_{1,u}\right) = 1$ for N large enough, see Proposition 3.5 in [1]. Next we proceed iteratively. Assume in generation g $m = m_N < \ell_N$ vertices are $\tilde{I}^{(N)}$-infected. If $m = 1$ we can use the coupling as in generation 0 and add independently additional CoSame and CoDiff infections according to the DBPC distribution in $\mathbf{Z}^{(N)}_u$, if $Z^{(N)}_{1,u} > 1$.

If $m > 1$, let $w_1, ..., w_m$ be the infected vertices and denote by $\overline{\mathcal{D}}_i$ the set and by $\overline{D}_i$ the number of vertices in the ball of radius r_N around vertex w_i for $i = 1, ..., m$. For $y \in \{0,1\}^m$ denote by $\mathcal{D}_y$ the set and by D_y the number of vertices that are contained in the balls that are centered around vertices w_j, $j \in \{1, ..., m\}$ which have a 1 at the j-th position of the vector y and are not contained in the other balls. So for example for $m = 3$ D_{001} gives the number of vertices that are contained only in the ball around vertex w_3, but not in the balls centered around vertices w_1 and w_2.

For a vector $x = (x_1, ..., x_m, x_{1,2}, x_{1,3}, ..., x_{m-1,m}) \in \mathbb{N}^{m+\binom{m}{2}}$ denote by E_x the event that in the next generation x_i CoSame infections occur caused by exactly two parasites generated on vertex w_i for $i = 1, ..., m$, $x_{i,j}$ Codiff infections occur caused by exactly two parasites being generated on vertex w_i and vertex w_j for $i, j \in \{1, ..., m\}$ with $i < j$ and all other vertices get attacked by at most one parasite.

To determine the probability of the event E_x we distinguish different cases. Let for $y \in \{0,1\}^m$ denote by $x^o_{y,i}$ the number of CoSame infections caused by parasites generated on vertex w_i attacking vertices in $\mathcal{D}_y$ as well as by $x^c_{y,i,j}$ the number of CoDiff infections generated by parasites from vertices w_i and w_j that are attacking vertices in $\mathcal{D}_y$ as well as by $x^r_{i,j}$ the number of parasites originating from vertex w_j and attacking a vertex without any other parasite in $\overline{\mathcal{D}}_i$. The probability of E_x is given by the sum of the probabilities of infection patterns corresponding to vectors $(x^o_{y,i})_y$, $(x^c_{y,i,j})_y$ and $(x^r_{i,j})_{i>j}$ with $\sum_y x^o_{y,i} = x_i$ where $x^o_{y,i} = 0$ if the ith coordinate of y is 0, and $\sum_y x^c_{y,i,j} = x_{i,j}$ where $x^c_{y,i,j} = 0$, if the ith or jth coordinate of y is 0, such that $D_y > \underline{x}_y$ with $\underline{x}_y = \sum_i x^o_{y,i} +$

$\sum_{i=1}^{m-1}\sum_{j=i+1}^{m} x^c_{y,i,j}$ for all $y \in \{0,1\}^m$. The probability of an infection pattern according to the vectors $x^o = (x^o_{y,i})_{y,i}$ and $x^c = (x^c_{y,i,j})_{y,i,j}$ is given by the product of the three factors $p_o = p_o(x^o, x^c)$, $p_c = p_c(x^o, x^c)$ and $p_r = p_r(x^o, x^c)$ representing the CoSame, CoDiff and single infections with

$$p_o = \prod_{i=1}^{m} \binom{v_N}{2} \cdots \binom{v_N - 2\,(x_i - 1)}{2} \frac{1}{D_i^{2x_i}} \prod_y \frac{(D_y - \overline{x}_{y,i})!}{(D_y - \overline{x}_{y,i} - x^o_{y,i})!} \frac{1}{x^o_{y,i}!}$$

where $\overline{x}_{y,i} = \sum_{j=1}^{i-1} x^o_{y,j}$. The factor $\binom{v_N}{2} \cdots \binom{v_N - 2(x_i-1)}{2}$ gives the number of possibilities to choose x_i pairs of parasites from the v_N parasites generated on vertex w_i, for $i = 1, ..., m$, $\frac{(D_y - \overline{x}_{y,i})!}{(D_y - \overline{x}_{y,i} - x^o_{y,i})!} \frac{1}{x^o_{y,i}!}$ gives the number of possibilities to choose for $x^o_{y,i}$ pairs of parasites a location in D_y, when we already distributed the pairs of parasites generated on vertices w_j for $j = 1, ..., i-1$ on D_y. $\frac{1}{D_i^{2x_i}}$ is the probability to place the pairs of parasites exactly on these locations in D_y.

$$p_c = \prod_{i=1}^{m} \prod_{j=i+1}^{m} \frac{(v_N - \underline{x}_{1,i,j})!}{(v_N - \underline{x}_{1,i,j} - x_{i,j})!} \frac{(v_N - \underline{x}_{2,i,j})!}{(v_N - \underline{x}_{2,i,j} - x_{i,j})!} \cdot \prod_y \frac{(D_y - \underline{x}_{y,i,j})!}{(D_y - \underline{x}_{y,i,j} - x^c_{y,i,j})!} \left(\frac{1}{D_i D_j}\right)^{x^c_{y,i,j}} \frac{1}{(x^c_{y,i,j})!}$$

with $\underline{x}_{1,i,j} = 2x_i + \sum_{\ell=i+1}^{j-1} x_{i,\ell}$ and $\underline{x}_{2,i,j} = 2x_j + \sum_{\ell=1}^{i-1} x_{\ell,j}$, $\underline{x}_{y,i,j} = \sum_{k=1}^{m} x^o_{y,k} + \sum_{k=1}^{i-1}\sum_{\ell=k+1}^{m} x^c_{y,k,\ell} + \sum_{k=i+1}^{j-1} x^c_{y,i,k}$. The factor $\frac{(v_N - \underline{x}_{1,i,j})!}{(v_N - \underline{x}_{1,i,j} - x_{i,j})!}$ gives the number of possibilities to choose $x_{i,j}$ parasites from the parasites generated on vertex w_i, when the parasites for the CoSame infections as well as the parasites for the CoDiff infections of the vertex pairs $(w_i, w_{i+1}), \cdots, (w_i, w_{j-1})$ have already been determined. The factor $\frac{(D_y - \underline{x}_{y,i,j})!}{(D_y - \underline{x}_{y,i,j} - x^c_{y,i,j})!} \frac{1}{(x^c_{y,i,j})!}$ gives the number of possibilities to choose in D_y the $x^c_{y,i,j}$ locations for the pairs of parasites generating a CoDiff infection from vertex w_i and w_j, when the locations for the CoSame infections as well as for the CoDiff infections of vertex pairs $(w_1, w_2), \cdots, (w_i, w_{j-1})$ have already been determined. Finally, the factor $\left(\frac{1}{D_i}\right)^{x^c_{y,i,j}} \left(\frac{1}{D_j}\right)^{x^c_{y,i,j}}$ is the probability to place the pairs of parasites generating the CoDiff infections on exactly these locations.

$$p_r = \prod_{i=1}^{m} \frac{D_i - x_i - \sum_{j=1}^{i-1} x_{j,i} - \sum_{j=i+1}^{m} x_{i,j} - \sum_{j=1}^{i-1} x^r_{i,j}}{D_i} \cdots \frac{D_i + x_i - \sum_{j=1}^{i-1} x^r_{i,j} - v_N + 1}{D_i}.$$

p_r is the probability to place the remaining parasites all onto different vertices.

To analyse the above probabilities, consider only configurations $(x^o_{y,i})_{y,i}, (x^c_{y,i,j})_{y,i,j}$ with positive entries for vectors y for which $1/D_y \in o(1/(d_N)^{1-\varepsilon'})$ and only values $x_i, x_{i,j} \le \ell_N$, because the sum of the remaining probabilities is $O(d_N^{-\varepsilon'})$. Under this assumption we can estimate

$$p_o \ge \prod_{i=1}^{m} \prod_y \left(\frac{(v_N - \ell_N^2)^2}{2\delta_{N,u}} \frac{D_y - \ell_N^2}{\delta_{N,u}}\right)^{x^o_{y,i}} \frac{1}{x^o_{y,i}!}. \tag{3.5.6}$$

Then by setting $t_N^y = \frac{(v_N-\ell_N^2)^2}{2\delta_{N,u}} \frac{D_y-\ell_N^2}{\delta_{N,u}}$ we can write

$$\begin{aligned}
(3.5.6) &= \prod_{i=1}^m \prod_y \left(\frac{(v_N-\ell_N^2)^2}{2\delta_{N,u}} \frac{D_y-\ell_N^2}{\delta_{N,u}} \right)^{x_{y,i}^o} \frac{1}{x_{y,i}^o!} \\
&= \prod_{i=1}^m \prod_y (t_N^y)^{x_{y,i}^o} \frac{1}{x_{y,i}^o!} \exp(-t_N^y) \exp(t_N^y) \\
&= \prod_{i=1}^m \left(\prod_{y:x_{y,i}^o \neq 0} (t_N^y)^{x_{y,i}^o} \frac{1}{x_{y,i}^o!} \exp(-t_N^y) \right) \left(\prod_{y:x_{y,i}^o \neq 0} \exp(t_N^y) \right) \\
&\geq \prod_{i=1}^m \mathbb{P}(Y_i = x_i) \exp(a_{N,i}),
\end{aligned}$$

for $\mathrm{Poi}(a_{N,i})$ distributed random variables Y_i with

$$a_{N,i} = \frac{(v_N-\ell_N^2)^2}{2\delta_{N,u}} \left(\frac{\sum_{y:x_{y,i}^o \neq 0} (D_y - \ell_N^2)}{\delta_{N,u}} \right).$$

By assumption $\sum_{y:x_{y,i}^o \neq 0} D_y \in [N^\beta - 2N^{\beta-\varepsilon'}, N^\beta + N^{\beta-\varepsilon'}]$ and so $a_{N,i} \geq \frac{(v_N-\ell_N^2)^2}{2\delta_{N,u}} \left(1 - 3N^{-\varepsilon'} - \frac{2^{\ell_N}\ell_N^2}{\delta_{N,u}}\right) =: a_N$. Consequently, we have $(3.5.6) \geq \prod_{i=1}^m \mathbb{P}(Y_i = x_i) \exp(a_N)$.

Similarly, we have

$$\begin{aligned}
p_c &\geq \prod_{i=1}^m \prod_{j=i+1}^m (v_N - \ell_N^2)^{2x_{i,j}} \prod_{y:x_{y,i,j}^c \neq 0} \left(\frac{D_y - \ell_N^2}{\delta_{N,u}} \right)^{x_{y,i,j}^c} \left(\frac{1}{\delta_{N,u}} \right)^{x_{y,i,j}^c} \frac{1}{x_{y,i,j}^c!} \\
&\geq \prod_{i=1}^m \prod_{j=i+1}^m \mathbb{P}(Y_{i,j} = x_{i,j}) \exp(2a_N)
\end{aligned}$$

with $Y_{i,j} \sim \mathrm{Poi}(2a_{N,i,j})$ and

$$a_{N,i,j} = \frac{(v_N-\ell_N^2)^2}{2\delta_{N,u}} \left(\frac{\sum_{y:x_{y,i,j}^c \neq 0} (D_y - \ell_N^2)}{\delta_{N,u}} \right).$$

Furthermore $p_r \geq \exp\left(-m^2 a_N\right)$ for N large enough, since $m \geq 2$. Consequently, we have

$$\mathbb{P}(E_x) \geq \prod_{i=1}^m \mathbb{P}(Y_i = x_i) \prod_{k=1}^m \prod_{\ell=k+1}^m \mathbb{P}(Y_{k,\ell} = x_{k,\ell}).$$

Since the random variables Y_i and $Y_{k,\ell}$ have finite moments, we can control (e.g. with Markov's inequality) the probability $\mathbb{P}\left(\bigcup_{x,x\leq \ell_N} E_x\right)$ by

$$\mathbb{P}\left(\bigcup_{x,x\leq \ell_N} E_x \right) = 1 - o\left(\frac{1}{\ell_N} \right),$$

where we write $x \leq \ell_N$, if $x_i \leq \ell_N$ and $x_{i,j} \leq \ell_N$ for all $i, j \in \{1, ..., m\}$ with $j > i$.

Since

$$\exp\left(-\frac{v_N^2}{2\delta_{N,u}}\right) \leq \min_{i,k,\ell}\{\exp(-a_{N,i}), \exp(-a_{N,k,\ell})\},$$

we have for any $0 \leq k \leq \ell_N$

$$\mathbb{P}\left(Z_{g,u}^{(N)} = k \mid Z_{g-1,u}^{(N)} = m\right) \leq \sum_{\substack{x \in \mathbb{N}^{m+\binom{m}{2}}: \\ \sum_{i=1}^m x_i + \sum_{i,j=1, i<j}^m x_{i,j} = k}} \mathbb{P}(E_x)$$

and hence, we can couple the processes $(\tilde{I}_g)$ and $(Z_{g,u})_g$ such that $\tilde{I}_g \leq Z_{g,u}$ whp for any $g \leq \tau_{\ell_N,0}^{\overline{\mathbf{Z}}_u}$, since whp $\tau_{\ell_N,0}^{\overline{\mathbf{Z}}_u} \leq C \log(\log(\ell_N))$ by Proposition 3.3.6. □

3.5.2 Lower bound on the invasion probability

Establishing invasion

In this section we show that in the random geometric graph the level N^ε is reached with at least the probability with which a well chosen lower DBPC reaches this level for some $0 < \varepsilon < \beta/4$.

Let us recall the complete neighbourhood $\mathcal{C}(x)$ of a vertex $x \in [0,1]^n$, which we already introduced in words in Subsection 3.2.2. To be precise

$$\mathcal{C}(x) = \mathcal{C}^{(N)}(x) := \left\{y \in \mathcal{V}^{(N)} : d(x,y) \leq \tfrac{r_N}{2}\right\} \subset \mathcal{V}^{(N)},$$

i.e. the set $\mathcal{C}(x)$ contains all vertices of the graph $\mathcal{G}^{(N)}$ lying in $B_{r_N/2}(x) \cap [0,1]^n$. We call this subset a complete neighbourhood, since it forms a complete graph as a subgraph of $\mathcal{G}^{(N)}$. Furthermore, recall that we denote by x_c the center of $[0,1]^n$ and by $x_0 \in \mathcal{V}^{(N)}$ the vertex with the smallest distance to x_c. We assumed that x_0 is the vertex that gets infected first, i.e. $\mathcal{I}_0 = \{x_0\}$.

As already mentioned before we will only consider realisations of the underlying graph $\mathcal{G}^{(N)}$ contained in the event $\mathcal{U}_N(\beta/2 + \xi)$ (see Equation (3.5.1) for a definition of $\mathcal{U}_N$). Note these realisations are in particular connected. Furthermore every complete neighbourhood $\mathcal{C}(x)$ contains whp at least $\frac{N^\beta}{2^n} - C(n)N^{c(n)\beta}$ and at most $\frac{N^\beta}{2^n} + C(n)N^{c(n)\beta}$ vertices for $x \in \left[\frac{r_N}{2}, 1 - \frac{r_N}{2}\right]^n$. Since $\mathbb{P}(\mathcal{U}_N(\beta/2 + \xi)) \to 1$ as $N \to \infty$ we will condition on this event and introduce the notation

$$\widetilde{\mathbb{P}}_N(\,\cdot\,) := \mathbb{P}(\,\cdot\,|\,\mathcal{U}_N(\beta/2 + \xi)).$$

Now we analyse the probability of infecting at least N^ε hosts by studying the infection process in the complete neighbourhood $\mathcal{C}(x_c)$ around the center point. Since we condition on $\mathcal{U}_N(\beta/2 + \xi)$ the set $\mathcal{C}(x_c)$ contains in particular the initially infected host x_0.

Definition 3.5.8. *(Subinfection process) The subinfection process $\mathcal{H}^{(N)} = (\mathcal{H}_g^{(N)})_{g \geq 0}$ on the complete neighbourhood $\mathcal{C}(x_c)$ of x_c is defined as follows.*

We set $\mathcal{H}_0^{(N)} = \{x_0\} \cap \mathcal{C}(x_c) \subset \mathcal{I}_0^{(N)}$. Assume the process is defined up to generation $g \geq 0$, then conditional on $\sigma(\mathcal{S}_m^{(N)}, \mathcal{I}_m^{(N)}, \mathcal{R}_m^{(N)}, \mathcal{H}_m^{(N)} : m \leq g)$ set $\mathcal{H}_{g+1}^{(N)} \subset \mathcal{I}_{g+1}^{(N)}$ to be the set of all infected hosts contained in $\mathcal{C}(x_c)$ generated by previously infected hosts $x \in \mathcal{H}_g^{(N)}$. We set $H_g^{(N)} := |\mathcal{H}_g^{(N)}|$ for all $g \geq 0$.

For any sequence $(b_N)_{N\in\mathbb{N}}$ define

$$\tau^{\mathbf{I}}_{b_N} := \inf\left\{g \in \mathbb{N} : I^{(N)}_g \geq b_N\right\},$$
$$\tau^{\bar{\mathbf{I}}}_{b_N} := \inf\left\{g \in \mathbb{N} : \bar{I}^{(N)}_g \geq b_N\right\}.$$

Lemma 3.5.9. *(Coupling between $\mathcal{H}^{(N)}$ and an infection process on a complete graph)* *There exists a coupling between $\mathcal{H}^{(N)}$ and the infection process $\mathcal{J}^{(N)} = (\mathcal{J}^{(N)}_g)_{g\geq 0}$ on a complete graph as defined in Section 3.4 with $\widetilde{d}_N$ vertices, where $\widetilde{d}_N := \left|\mathcal{C}^{(N)}(x_c)\right|$, and $\widetilde{v}_N$ offspring parasites, where $\widetilde{v}_N$ is defined in the proof of this lemma and is asymptotically equal to $\frac{v_N}{2^n}$, such that $(H^{(N)}_g)_g$ dominates $(J^{(N)}_g)_g$ for all $g \leq \tau^{\bar{\mathbf{I}}}_{N^\varepsilon}$ whp, i.e.*

$$\widetilde{\mathbb{P}}_N\left(\forall g \leq \tau^{\bar{\mathbf{I}}}_{N^\varepsilon}, H^{(N)}_g \geq J^{(N)}_g\right) \underset{N\to\infty}{\longrightarrow} 1.$$

Proof. Since we condition on the event $\mathcal{U}_N(\beta/2+\xi)$ the complete neighbourhood $\mathcal{C}(x_c)$ forms a complete graph as a subgraph of $\mathcal{G}$ and $\left||\mathcal{C}(x_c)| - \frac{N^\beta}{2^n}\right| \leq C(n)N^{c(n)\beta}$. In particular $\mathcal{J}^{(N)}$ is defined on a complete graph of the same size as $\mathcal{C}(x_c)$.

Now assume that some host x is currently infected according to the subinfection process $\mathcal{H}^{(N)}$. Then out of the v_N parasites which are generated in host x only those parasites which are moving to vertices in $\mathcal{C}(x_c)$ are counted in $\mathcal{H}^{(N)}$. The number of parasites which originate from x and move to a host also contained in $\mathcal{C}(x_c)$ is bounded from below by $\widetilde{v}_N = \frac{v_N}{2^n} - v_N \cdot p_N$ whp, where p_N is determined by the upper and lower bound on the number of vertices in $\mathcal{C}(x_c)$ and any ball of radius r_N conditioned on the event $\mathcal{U}_N$. In particular, $p_N \in \mathcal{O}(N^{-u})$ for some sufficiently small $u > 0$. Indeed, a parasite chooses uniformly at random the neighbour it is moving toward, and thus the number of parasites moving to $\mathcal{C}(x_c)$ is binomially distributed with parameter v_N and $\frac{1}{2^n}(1+\mathcal{O}(p_N))$. So the claimed whp lower bound follows by an application of Markov's inequality. Repeating this argument one can show that for each of at most N^ε infections the number of generated parasites for every single infection is bounded from below by $\frac{v_N}{2^n} - v_N \cdot p_N$ whp. Consequently, since we choose the number of generated parasites for $\mathcal{J}^{(N)}$ to be $\widetilde{v}_N$ only less infections can be created with respect to the infection process $\mathcal{J}^{(N)}$ in comparison to $\mathcal{H}^{(N)}$.

The claimed coupling could fail when two parasites generated during the process $\mathcal{I}^{(N)}$ move from outside of $\mathcal{C}(x_c)$ to an empty vertex contained in $\mathcal{C}(x_c)$. In this case no infection happens at this vertex with respect to the subinfection process $\mathcal{H}^{(N)}$, and thus with respect to the infection process $\mathcal{J}^{(N)}$ potentially more infections could have been generated. However such an event is only possible, if with respect to the infection process $\mathcal{I}^{(N)}$ two pairs of parasites are attacking the same vertex in $\mathcal{C}(x_c)$ (at the same or different generations). So this particular event is contained in the event that at least one vertex of $\mathcal{C}(x_c)$ is attacked by at least four parasites up to generation $\tau^{\bar{\mathbf{I}}}_{N^\varepsilon}$.

Until generation $\tau^{\bar{\mathbf{I}}}_{N^\varepsilon}$ less than N^ε vertices get infected cumulatively over all generations. So it is possible to estimate from above the probability that such an event happens before generation $\tau^{\bar{\mathbf{I}}}_{N^\varepsilon}$ by estimating the probability of the event A in the following experiment: Assume $N^\varepsilon v_N$ balls placed uniformly at random into $\widetilde{d}_N$ boxes and we are interested in the event A that it exists (at least) one box containing at least four balls. Indeed, the probability of the event A gives an upper bound, all balls (corresponding to parasites in the original process) are put into $\widetilde{d}_N$ boxes (parasites have a larger choice of vertices where they can move to in the original process). This increases the probability for one box to contain four balls.

We upper bounded the probability of A as follows

$$\begin{aligned}\widetilde{\mathbb{P}}_N(A) &\leq \left(\tfrac{d_N}{2^n} + C(n)N^{c(n)\beta}\right)\left(\frac{N^{\varepsilon}v_N}{\frac{d_N}{2^n} - C(n)N^{c(n)\beta}}\right)^4 \\ &= \mathcal{O}\left(\frac{N^{4\varepsilon}}{d_N}\right) \underset{N\to\infty}{\longrightarrow} 0,\end{aligned}$$

because $\varepsilon < \frac{\beta}{4}$.

So $H_g^{(N)}$ can be coupled with $J_g^{(N)}$ such that whp $H_g^{(N)} \geq J_g^{(N)}$ for all $g \leq \tau^{\bar{\mathbf{I}}}_{N^\varepsilon}$. □

Lemma 3.5.10. *It holds that*

$$\liminf_{N\to\infty} \widetilde{\mathbb{P}}_N\left(\exists g \in \mathbb{N}, \bar{I}_g^{(N)} \geq N^{\varepsilon}\right) \geq \pi\left(\frac{a}{\sqrt{2^n}}\right),$$

where $\pi\left(\frac{a}{\sqrt{2^n}}\right)$ *is the survival probability of a DBPC with offspring and cooperation distribution* $Poi\left(\frac{a^2}{2^{n+1}}\right)$ *and* $Poi\left(\frac{a^2}{2^n}\right)$.

Proof. The claim follows by Lemma 3.5.9 and the results of Section 3.4. □

Increasing from a total number of N^ε infections to $N^{\frac{\beta}{2}+\delta}$ infections within a single ball

In this subsection we will show that the total number of infected hosts increases whp from N^ε to $N^{\beta/2+\delta}$ within a single ball, with δ defined in (3.5.9).

Cover the space with non-overlapping balls, such that all balls have an edge length of at most r_N and such that all balls except of those having a non-empty intersection with the boundary of $[0,1]^n$ have an edge length of precisely r_N. Furthermore, assume that one of the balls is centered around x_c, that is the vertex set of this ball is equal to $\mathcal{C}(x_c)$. Label the balls and denote by $\mathbf{K}$ the set of labels and by $\mathcal{V}_k^{(N)}$ the set of vertices in ball k. Furthermore, denote by $I_g^{(N)}(k)$ the number of infected vertices in ball k in generation g.

Lemma 3.5.11. *It holds*

$$\widetilde{\mathbb{P}}_N\left(\exists g \leq \tau^{\bar{\mathbf{I}}}_{N^\varepsilon}, \exists k \in \mathbf{K}, I_g^{(N)}(k) \geq \frac{N^\varepsilon}{\log(N)} \middle| \tau^{\bar{\mathbf{I}}}_{N^\varepsilon} < \infty\right) \to 1.$$

Proof. For a sequence $(g_N)_{N\in\mathbb{N}}$ let

$$\Gamma_{g_N}^{(N)} := \left\{k \in \mathbf{K} : \exists g \leq g_N, I_g^{(N)}(k) \geq 1\right\}$$

be the set of labels of balls, in which at least one host gets infected up to generation g_N. At each generation a parasite may move a distance of at most r_N. So in dimension n, in g_N generations, the number of balls of diameter r_N that can be reached is $(2g_N+1)^n$, so $|\Gamma_{g_N}^{(N)}| \leq (2g_N+1)^n$.

Using Lemma 3.5.9 and the coupling from below with the DBPC $\mathbf{Z}_\ell^{(N)}$ until generation $\tau^{\bar{\mathbf{I}}}_{N^\varepsilon}$ and applying Proposition 3.3.6 to the DBPC $\mathbf{Z}_\ell^{(N)}$ we obtain that it exists a $C > 0$ such that

$$\widetilde{\mathbb{P}}_N\left(\tau^{\bar{\mathbf{I}}}_{N^\varepsilon} \leq C\log(\log(N)) \mid \tau^{\bar{\mathbf{I}}}_{N^\varepsilon} < \infty\right) \to 1. \tag{3.5.7}$$

Combining these two results we obtain that whp

$$\left|\Gamma^{(N)}_{\tau^{\overline{\mathbf{I}}}_{N^\varepsilon}}\right| \leq (2C\log(\log(N))+1)^n.$$

If in any generation before generation $\tau^{\overline{\mathbf{I}}}_{N^\varepsilon}$ the number of infected individuals in any ball of $\Gamma^{(N)}_{\tau^{\overline{\mathbf{I}}}_{N^\varepsilon}}$ is smaller than $\frac{N^\varepsilon}{\log(N)}$, the total number of infected individuals up to generation $\tau^{\overline{\mathbf{I}}}_{N^\varepsilon}$ would be upper bounded by

$$\frac{N^\varepsilon}{\log(N)}\left(2C\log(\log(N))+1\right)^n C\log(\log(N)) = o(N^\varepsilon),$$

which gives a contradiction. □

Next let $\sigma^{(N)}$ be the stopping time, at which for the first time in one of the balls at least $\frac{N^\varepsilon}{\log(N)}$ hosts get infected

$$\sigma^{(N)} := \inf\left\{g \in \mathbb{N} : \exists k \in \mathbf{K}, I^{(N)}_g(k) \geq \frac{N^\varepsilon}{\log(N)}\right\}.$$

The last lemma exactly states that

$$\widetilde{\mathbb{P}}_N\left(\sigma^{(N)} \leq \tau^{\overline{I}}_{N^\varepsilon} \;\middle|\; \tau^{\overline{I}}_{N^\varepsilon} < \infty\right) \underset{N\to\infty}{\longrightarrow} 1.$$

Now we will show that after a finite number of generations after generation $\sigma^{(N)}$, there is whp one ball $k \in \mathbf{K}$ with at least $N^{\frac{\beta}{2}+\delta}$ infected vertices for some $\delta > 0$ which is sufficiently small.

To achieve this goal, we will argue in the same manner as we have done in Subsection 3.4.2. Choose $\varepsilon > 0$ such that for all $g \in \mathbb{N}$, $2^g\varepsilon \neq \frac{\beta}{2}$. Then define $\overline{g}$ as the largest $g \in \mathbb{N}$ satisfying $2^{g+1}\varepsilon < \beta$. In particular it is $2^{\overline{g}+1}\varepsilon > \frac{\beta}{2}$ fulfilled because otherwise we would have $2^{\overline{g}+2}\varepsilon < \beta$ which contradicts the definition of $\overline{g}$.

Denote by

$$S^{(N)} := \left\{k \in \mathbf{K} : I^{(N)}_{\sigma^{(N)}}(k) \geq \frac{N^\varepsilon}{\log(N)}\right\} \tag{3.5.8}$$

the set of balls that contain at least $\frac{N^\varepsilon}{\log(N)}$ infected vertices in generation $\sigma^{(N)}$. By definition of $\sigma^{(N)}$ the set $S^{(N)}$ is not empty almost surely, if $\sigma^{(N)} < \infty$.

Lemma 3.5.12. *We have*

$$\widetilde{\mathbb{P}}_N\left(\overline{I}^{(N)}_{\sigma^{(N)}} \leq \frac{N^{2\varepsilon}}{\log(N)} \;\middle|\; \tau^{\overline{I}}_{N^\varepsilon} < \infty\right) \underset{N\to\infty}{\longrightarrow} 1.$$

Proof. By definition, at generation $\sigma^{(N)}-1$ the number of infected vertices in each ball i is at most $\frac{N^\varepsilon}{\log(N)}$ and the total number of balls that have been infected is whp at most $(2C\log(\log(N))+1)^n$. To show that $\overline{I}^{(N)}_{\sigma^{(N)}} \leq \frac{N^{2\varepsilon}}{\log(N)}$, we will control the number of infections in each ball by applying a similar argument as in Lemma 3.4.10 in the context of the complete graph.
At generation $\sigma^{(N)}-1$ whp we have $\overline{I}^{(N)}_{\sigma^{(N)}-1} \leq N^\varepsilon$ because otherwise we would have a contradiction to Lemma 3.5.11. Then to bound from above at generation $\sigma^{(N)}$ the total number of infected vertices up to this generation, it suffices to add to N^ε an upper bound on the number of new infections generated in generation $\sigma^{(N)}$.

In each ball, there are at most $\frac{N^\varepsilon}{\log(N)}v_N$ parasites that will move. Because of the sizes of the balls, each ball can receive infections from outside only due to its 3^n-1 neighbouring balls. To arrive at an upper bound on the number of new infections generated in a ball, one can compare the situation with the following balls-into-boxes experiment: Consider $d_N := \frac{N^\beta}{2^n} - C(n)N^{c(n)\beta} - N^\varepsilon$ boxes. Put $(3^n-1)\frac{N^\varepsilon}{\log(N)}v_N$ balls into the boxes uniformly at random and count the number of boxes that contain at least two balls. Applying Lemma 3.4.8 Equation (3.4.12) with $H(N) = (3^n-1)\frac{N^\varepsilon}{\log(N)}$, $\varphi_1(N) = N^\varepsilon$, $\varphi_2(N) = 0$, $\varphi_3(N) = 0$ $f_2(N) = \frac{1}{(3^n-1)^2}\frac{\log(N)}{(2C\log(\log(N))+1)^n}$, we can estimate from above the number of infections generated in a ball by $\frac{N^{2\varepsilon}}{(\log N)(2C\log^n(\log(N)))}$ with probability $1-\Theta\left(\frac{\log^n(\log(N))}{\log(N)}\right)$. Such result holds for each ball. Since in generation $\sigma^{(N)}-1$ whp at most $(2C\log(\log(N))+1)^n$ balls have been infected we can apply this argument for all ball and obtain that whp the total number of infected hosts does not exceed $\frac{N^{2\varepsilon}}{\log(N)}$. □

Lemma 3.5.13. *Let $g \in \{0,\dots,\overline{g}\}$ we have*

$$\widetilde{\mathbb{P}}_N\left(\forall k \in S^{(N)}, I^{(N)}_{\sigma^{(N)}+g}(k) \geq \frac{N^{2^g\varepsilon}}{\log^{\alpha_g}(N)}, \overline{I}^{(N)}_{\sigma^{(N)}+g} \leq \frac{N^{2^{g+1}\varepsilon}}{\log(N)} \middle| \tau^{\overline{I}}_{N^\varepsilon} < \infty\right) \to 1,$$

where $\alpha_0 = 1$ and for all $g \geq 1, \alpha_g = 2\alpha_{g-1}+2$.

Proof. The proof is obtained by induction. First for $g=0$ the result is given by Lemma 3.5.12. Then let $g \leq \overline{g}-1$, assume the result is obtained for $0 \leq j \leq g$. Now we will show the result for $g+1$.

To derive the lower bound on the number of infected vertices in a ball $k \in S^{(N)}$ at generation $\sigma^{(N)}+g+1$, one can consider only the infections generated due to infected vertices inside this ball. According to the induction hypothesis there are at least $\frac{N^{2^g\varepsilon}}{\log^{\alpha_g}(N)}$ infected vertices in the ball. Among the parasites generated on these vertices, at least $\frac{N^{2^g\varepsilon}}{\log^{\alpha_g}(N)}\left(\frac{1}{2^n}v_N - v_n\cdot p_N\right)$ (where p_N is defined in the proof of Lemma 3.5.9) of them will move to vertices in the ball. Then it suffices to apply Lemma 3.4.8 Equation (3.4.11) where d_N in this Lemma is equal to $\frac{1}{2^n}N^\beta + C(n)N^{c(n)\beta}$, with $H(N) := \frac{1}{2^n}\frac{N^{2^g\varepsilon}}{\log(N)}$, $\varphi_1(N) = 0$, $\varphi_2(N) = \frac{N^{2^{g+1}\varepsilon}}{\log(N)}$, $\varphi_3(N) = 2^n v_N\cdot p_N$, $f_1(N) = \frac{\log^2(N)}{2^{n+1}}$, which gives that the number of infected vertices at generation $\sigma^{(N)}+g+1$ is whp at least of order $\frac{2^{n+1}}{\log^2(N)}\left(\frac{1}{2^n}\frac{N^{2^g\varepsilon}}{\log^{\alpha_g}(N)}\right)^2 = \frac{N^{2^{g+1}\varepsilon}}{\log^{\alpha_{g+1}}(N)}$. Because there are whp at most $\Theta(\log(\log(N)))$ ball in $S^{(N)}$ and by Equation (3.4.11) of Lemma 3.4.8, the statement holds whp for all balls in $S^{(N)}$.

Indeed considering $\frac{1}{2^n}N^\beta + C(n)N^{c(n)\beta}$ boxes (for the balls-into-boxes experiment) lower bounds the probability for a parasite to move to an occupied vertex, because whp there are at most $\frac{1}{2^n}N^\beta + C(n)N^{c(n)\beta}$ many vertices in the ball. Furthermore, according to the induction assumption we have considered the minimal number of parasites which is $\frac{N^{2^g\varepsilon}}{\log^{\alpha_g}(N)}\left(\frac{1}{2^n}v_N - v_N\cdot p_N\right)$ and new infections are counted when reaching one of the $\frac{1}{2^n}d_N - \frac{N^{2^{g+1}\varepsilon}}{\log(N)}$ first boxes whereas in the original process there are at least this number of occupied vertices.

To derive the upper bound on the number of empty vertices, we control for each ball the number of new infections generated in generation $\sigma^{(N)}+g+1$. Since by induction the number of empty vertices in generation $\sigma^{(N)}+g$ is $\frac{N^{2^{g+1}\varepsilon}}{\log(N)}$ whp, we apply Lemma 3.4.8 Equation (3.4.12) with $H(N) =$

$(3^n-1)\frac{N^{2^{g+1}\varepsilon}}{\log(N)}$, $\varphi_1(N)=\frac{N^{2^{g+1}\varepsilon}}{\log(N)}$, $\varphi_2(N)=0$, $\varphi_3(N)=0$, $f_2(N)=\frac{1}{(3^n-1)^2}\frac{\log(N)}{\left(2\log^{\frac{1}{n+2}}(N)+g+2\right)^n}$ to estimate the number of new infections in generation $\sigma^{(N)}+g+1$ in each ball in $S^{(N)}$. The lemma yields that in each ball there are at most $\frac{N^{2^{g+2}\varepsilon}}{\log^2(N)}\frac{\log(N)}{\left(2\log^{\frac{1}{n+2}}(N)+g+2\right)^n}$ new infections whp. Since there are whp at most $(2C\log(\log(N))+g+2)^n$ balls and since $f_2(N)=\Theta\left(\log(N)^{\frac{2}{n+2}}\right)$ whp for all balls the number of new infections is bounded from above by $\frac{N^{2^{g+2}\varepsilon}}{\log^2(N)}\frac{\log(N)}{\left(2\log^{\frac{1}{n+2}}(N)+g+2\right)^n}$, see Equation (3.4.12) of Lemma 3.4.8. Consequently, the total number of empty vertices at generation $\sigma^{(N)}+g+1$ is whp at most

$$\frac{N^{2^{g+2}\varepsilon}}{\log^2(N)}\frac{\log(N)}{\left(2\log^{\frac{1}{n+2}}(N)+g+2\right)^n}\left(2C\log(\log(N))+g+2\right)^n\leq\frac{N^{2^{g+2}\varepsilon}}{\log(N)}.$$

□

Applying Lemma 3.5.13 for $g=\overline{g}$ gives that

$$\widetilde{\mathbb{P}}_N\left(\forall k\in S^{(N)},I^{(N)}_{\sigma^{(N)}+\overline{g}}(k)\geq\frac{N^{2^{\overline{g}}\varepsilon}}{\log^{\alpha_{\overline{g}}}(N)},\overline{I}^{(N)}_{\sigma^{(N)}+\overline{g}}\leq\frac{N^{2^{\overline{g}+1}\varepsilon}}{\log(N)}\middle|\tau^{\overline{I}}_{N^\varepsilon}<\infty\right)\to 1.$$

Define

$$\delta=\frac{1}{2}\left(2^{\overline{g}+1}\varepsilon-\frac{\beta}{2}\right)>0. \tag{3.5.9}$$

In the next lemma we show that at generation $\sigma^{(N)}+\overline{g}+1$ the number of infected vertices in each ball of $S^{(N)}$ is at least of order $N^{\frac{\beta}{2}+\delta}$.

Lemma 3.5.14. *It holds*

$$\widetilde{\mathbb{P}}_N\left(\forall k\in S^{(N)},I^{(N)}_{\sigma^{(N)}+\overline{g}+1}(k)\geq N^{\frac{\beta}{2}+\delta}|\tau^{\overline{I}}_{N^\varepsilon}<\infty\right)\to 1.$$

Proof. Here we apply again Lemma 3.4.8 to obtain this lower bound. More precisely with the following set of parameters: $H(N)=\frac{1}{2^n}\frac{N^{2^{\overline{g}}\varepsilon}}{\log^{\alpha_{\overline{g}}}(N)}$, $\varphi_1(N)=0$, $\varphi_2(N)=\frac{N^{2^{\overline{g}+1}\varepsilon}}{\log(N)}$, $\varphi_3(N)=2^n v_N\cdot p_N$, $f_1(N)=\frac{\log^2(N)}{2^{n+1}}$. We obtain that whp $I_{\tau^I_{N^\varepsilon}+\overline{g}+1}\geq\frac{N^{2^{\overline{g}+1}\varepsilon}}{\log^{\alpha_{\overline{g}+1}}(N)}\geq N^{\frac{\beta}{2}+\delta}$, by definition of δ. □

"Pulled travelling wave" epidemic spread

We start with a general lemma that we will use multiple times in this subsection. It says that when a ball of diameter ε_N is fully infected, then in the next generation, all the vertices in the neighboring area of diameter $2r_N-\varepsilon_N$ are visited by at least two parasites whp.

Lemma 3.5.15. *Consider a ball of diameter ε_N centered around a point $x\in[0,1]^n$, denoted by B_2, where $\varepsilon_N\in\Theta\left(N^{\frac{\beta/2+\xi-1}{n}}\right)$ for $\xi>0$ small enough. Assume that the proportion of currently infected vertices in this ball is asymptotically 1. Then in the next generation all the vertices in the ball centered around x with diameter $2r_N-\varepsilon_N$, denoted by B_1, are attacked by at least 2 parasites with probability $1-\mathcal{O}\left(\varepsilon_N^n N v_N\exp\left(-\frac{\varepsilon_N^n}{(2r_N)^n}v_N\right)\right)$ conditioned on $\mathcal{U}_N(\beta/2+\xi)$.*

Proof. We estimate from above the probability that at least one vertex is visited by at most 1 parasite in the following generation. Denote by K_1 the number of vertices in B_1. Due to the conditioning on $\mathcal{U}_N(\beta/2+\xi)$ it holds

$$(2r_N-\varepsilon_N)^n N - C(n)N^{c(n)\beta} \leq K_1 \leq (2r_N-\varepsilon_N)^n N + C(n)N^{c(n)\beta} \in \Theta(d_N).$$

Denote by $\underline{K}_1 = (2r_N-\varepsilon_N)^n N - C(n)N^{c(n)\beta}$. Furthermore, within B_2 due to our conditioning at least $\varepsilon_N^n N v_N/2$ and at most $2\varepsilon_N^n N v_N$ parasites are generated. Denote the number of parasites generated in B_2 by K_2. Since B_2 is contained in B_1 and B_1 has diameter $2r_N-\varepsilon_N$ every vertex in B_2 is connected over an edge to any vertex in B_1. Hence, the probability p that a particular vertex in B_1 gets attacked by at most one parasite generated in B_2 can be estimated from above by

$$\begin{aligned} p &\leq \left(1-\frac{1}{\underline{K}_1}\right)^{K_2} + K_2\left(1-\frac{1}{\underline{K}_1}\right)^{K_2-1}\frac{1}{\underline{K}_1} \\ &\leq \left(1-\frac{1}{\underline{K}_1}\right)^{\varepsilon_N^n N v_N/2} + 2\varepsilon_N^n N v_N\left(1-\frac{1}{\underline{K}_1}\right)^{\varepsilon_N^n N v_N/2-1}\frac{1}{\underline{K}_1}. \end{aligned}$$

Consequently, we can estimate from above the probability that at least one of the vertices gets attacked by at most one parasite by

$$K_1\left(1-\frac{1}{\underline{K}_1}\right)^{\varepsilon_N^n N v_N/2} + 2\varepsilon_N^n N v_N\left(1-\frac{1}{\underline{K}_1}\right)^{\varepsilon_N^n N v_N/2-1}\frac{K_1}{\underline{K}_1} \in \mathcal{O}\left(\varepsilon_N^n N v_N \exp\left(-\frac{\varepsilon_N^n}{(2r_N)^n}v_N\right)\right),$$

from which follows the claim. □

This lemma implies that within one generation all vertices in the ball B_1 get attacked by at least two parasites whp from parasites generated on vertices in B_2, since ε_N is chosen such that $\varepsilon_N^n v_N / r_N^n \sim N^\xi$. In particular this means that if in B_2 almost all the vertices that are not contained in B_1 (which is asymptotically completely infected) are still occupied by a host, then all these hosts in B_2 get infected in the next generation. This allows us to repeat the same argument subsequently. Due to the exponential decay of the error term $\mathcal{O}\left(\varepsilon_N^n N v_N \exp\left(-\frac{\varepsilon_N^n}{(2r_N)^n}v_N\right)\right)$ we can apply this argument for many balls, in particular for the $2^n N^{1-\beta}$ many balls of diameter r_N. In particular, this implies that we can show that a pulled traveling wave in any direction is created by repeating the argument, as long as the invasion is not stopped by a region in which no susceptible hosts are available anymore and which cannot be crossed by parasites.
However, such a region with a non-trivial proportion of hosts killed and with a diameter of at least $r_N-\varepsilon_N$ (such that it cannot be crossed by a travelling wave whp) cannot arise, because if in a ball of size $r_N-\varepsilon_N$ at least k_N, with $k_N\to\infty$ arbitrarily slowly, hosts get infected by couplings with DBPCs (which have a positive survival probability) we can show that in this region either a new infection wave is started or a travelling wave is hitting the ball. Consequently balls cannot be slowly depleted and we will reach the boundary of $[0,1]^n$ whp after at most

$$\frac{1}{2(r_N-\varepsilon_N)} = \frac{1}{2r_N}\left(1+\mathcal{O}\left(\frac{\varepsilon_N}{r_N}\right)\right) \tag{3.5.10}$$

many generations.

3.5.3 Proof of Theorem 3.2.4 1) (ii):

Now we have all necessary materials to prove Theorem 3.2.4 1) (ii). The first step is to show

$$\limsup_{N\to\infty} \mathbb{P}\big(E_u^{(N)}\big) \le \pi(a). \tag{3.5.11}$$

For a sequence $(\ell_N)_{N\in\mathbb{N}}$ introduce the event

$$A_{\ell_N} := \Big\{\exists g \in \mathbb{N}_0 : \overline{I}_g^{(N)} \ge \ell_N\Big\}.$$

Then for all $0 < u \le 1$ and any sequence $\ell_N \le u|\mathcal{V}^{(N)}|$ we have

$$\widetilde{\mathbb{P}}_N\big(E_u^{(N)}\big) \le \widetilde{\mathbb{P}}_N\big(A_{\ell_N}\big). \tag{3.5.12}$$

Let (ℓ_N) be a sequence with $\ell_N \to \infty$ as well as with $\ell_N \in o(\log(\log(N)))$. Then Proposition 3.5.7 yields

$$\widetilde{\mathbb{P}}_N\big(A_{\ell_N}\big) \le \widetilde{\mathbb{P}}_N\big(\exists g \in \mathbb{N}_0 : \overline{Z}_{g,u}^{(N)} \ge \ell_N\big) + o(1). \tag{3.5.13}$$

Proposition 3.5.6 gives that

$$\lim_{N\to\infty} \mathbb{P}\big(\exists g \in \mathbb{N}_0 : \overline{Z}_{g,u}^{(N)} \ge \ell_N\big) = \pi(a). \tag{3.5.14}$$

In summary combining (3.5.12), (3.5.13), (3.5.14) and that $\mathcal{U}_N(\beta/2+\xi)$ occurs whp we obtain (3.5.11).

The second step is to show

$$\liminf_{N\to\infty} \mathbb{P}\big(E_u^{(N)}\big) \ge \pi\left(\frac{a}{\sqrt{2^n}}\right). \tag{3.5.15}$$

Lemma 3.5.10 gives that

$$\liminf_{N\to\infty} \widetilde{\mathbb{P}}_N\big(\exists g \in \mathbb{N}, \overline{I}_g^{(N)} \ge N^{\varepsilon}\big) \ge \pi\left(\frac{a}{\sqrt{2^n}}\right). \tag{3.5.16}$$

Then Lemma 3.5.14 yields that in all balls of the set $S^{(N)}$, see (3.5.8) for its definition, the infection level is at least of order $N^{\beta/2+\delta}$ in the random generation $\sigma^{(N)} + \overline{g} + 1$, i.e.

$$\widetilde{\mathbb{P}}_N\big(\forall k \in S^{(N)}, I^{(N)}_{\sigma^{(N)}+\overline{g}+1}(k) \ge N^{\frac{\beta}{2}+\delta} | \tau^{\overline{I}}_{N^{\varepsilon}} < \infty\big) \to 1.$$

Then arguing as in the proof of Lemma 3.4.13 one can show that whp all hosts on vertices contained in the balls of the set $S^{(N)}$ get killed in one more generation, that is

$$\widetilde{\mathbb{P}}_N\big(\forall k \in S^{(N)}, \overline{I}^{(N)}_{\sigma^{(N)}+\overline{g}+2}(k) = |\mathcal{V}_k^{(N)}|\big) \to 1.$$

And finally using the results from Subsection 3.5.2 one can show that whp every host eventually gets infected conditioned on the event $\{\tau^{\overline{I}}_{N^{\varepsilon}} < \infty\}$, which combined with (3.5.16) and that $\mathcal{U}_N(\beta/2+\xi)$ occurs whp, gives (3.5.15).

3.5.4 Proof of Theorem 3.2.4 1) (i)

Assume $v_N \in o(\sqrt{d_N})$. Then using a similar approach as in Subsection 3.4.4 which is to show that whp there are no infected individuals at generation 1, one obtains the result.

3.5.5 Proof of Theorem 3.2.4 1) (iii)

In this section we assume $\sqrt{d_N} \in o(v_N)$. We will prove that

$$\lim_{N\to\infty} \mathbb{P}\big(\exists g \in \mathbb{N}_0 : \overline{I}_g^{(N)} = |\mathcal{V}^{(N)}|\big) = 1.$$

Proof. The proof is split into two parts. First we argue that we can reach with high probability a level N^α for any $\alpha < \beta$ in a time of order $\log^2(N)$. In the second part we show that similar as in the critical scaling the host population is killed by a traveling wave.

We closely follow the proof of Lemma 3.6.4 and the proof strategy in Subsection 3.4.5 for the first part. We build an infection process $\big(\widehat{\mathcal{S}}^{(N)}, \widehat{\mathcal{I}}^{(N)}, \widehat{\mathcal{R}}^{(N)}\big)$, in which infections are only transmitted due to parasites originating from the same vertex and $v_N^{(a)} = a\sqrt{d_N} = aN^{\frac{\beta}{2}}$ many parasites are generated. This means a host is only infected if at least two parasites which originate from the same vertex attack the host simultaneously. Note that $d_N \in o(v_N)$, which means that for every $a > 0$ there exists an N large enough such that $v_N^{(a)} \leq v_N$. Thus, analogously as we showed in Subsection 3.4.5, for every $a > 0$ we can couple this process to the original infection process $\big(\mathcal{S}^{(N)}, \mathcal{I}^{(N)}, \mathcal{R}^{(N)}\big)$ such that

$$\widehat{\mathcal{I}}_g^{(N)} \cup \widehat{\mathcal{R}}_g^{(N)} \subset \mathcal{I}_g^{(N)} \cup \mathcal{R}_g^{(N)}, \forall g \in \mathbb{N}_0,$$

for N large enough. Denote by $H_x^{(N)}$ the number of vertices which get attacked by at least two parasites originating from x. Denote by $\deg(x)$ the degree of vertex $x \in \mathcal{V}^{(N)}$. Then

$$\mathbb{P}\left(H_x^{(N)} = k\right) \geq \frac{\prod_{i=1}^k \binom{v_N - 2(i-1)}{2}}{k!\deg(x)^k} \frac{\deg(x)!}{(\deg(x) - k - v_N)!\deg(x)^{v_N - k}},$$

where we only consider infections resulting from cooperation from the same edge and ignore infections generated by groups of 3 or more parasites, since these events happen with a negligible probability. Recall the definition of $\delta_{N,\ell}$ and $\delta_{N,u}$ in (3.5.4) and (3.5.5). Set

$$A^{(N)} := \big\{\delta_{N,\ell} \leq \deg(x) \leq \delta_{N,u} \,\forall\, x \in \mathcal{V}^{(N)}\big\}.$$

By Lemma 3.5.1 it follows that $\mathbb{P}\big(A^{(N)}\big) \to 1$ as $N \to \infty$. Thus, $\delta_{N,\ell}$ and $\delta_{N,u}$ act as a uniform lower and upper bound on $\deg(x)$ for all $x \in \mathcal{V}^{(N)}$ with high probability and we can again conclude analogously as in Proposition 3.5 in [1] that

$$\begin{aligned}
&\frac{\prod_{i=1}^k \binom{v_N - 2(i-1)}{2}}{k!\deg(x)^k} \frac{\deg(x)!}{(\deg(x) - k - v_N)!\deg(x)^{v_N - k}} \\
&\qquad \geq \left(\frac{(v_N - 2a_N)^2}{2\deg(x)}\right)^k \frac{1}{k!} \exp\left(-\frac{v_N^2}{2\deg(x)}\right)\left(1 - \frac{1}{\deg(x)^\delta}\right) \\
&\qquad \geq \left(\frac{(v_N - 2a_N)^2}{2\delta_{N,u}}\right)^k \frac{1}{k!} \exp\left(-\frac{v_N^2}{2\delta_{N,\ell}}\right)\left(1 - \frac{1}{\delta_{N,\ell}^\delta}\right)
\end{aligned}$$

for $0 \leq k \leq a_N$. This suggests that we can couple the process $\big(|\widehat{\mathcal{I}}_g^{(N)} \cup \widehat{\mathcal{R}}^{(N)}|_g\big)_{g\geq 0}$ with an appropriately chosen branching process until we reach a level N^α with $\alpha < \beta$.

Definition 3.5.16. *(Modified lower Galton-Watson Process) Let $0 < \delta < \frac{1}{2}$ and $(a_N)_{N\in\mathbb{N}}$ be a sequence with $a_N \to \infty$ and $a_N \in o\big(N^{\frac{\beta}{2}}\big)$. Furthermore, assume $(\vartheta_N)_{N\in\mathbb{N}}$ is a $[0,1]$-valued sequence with $\vartheta_N \to 0$ as $N \to \infty$. Let $\mathbf{Y}_l^{(N)} = \big(Y_{g,l}^{(N)}\big)_{g\in\mathbb{N}_0}$ be a Galton-Watson process with mixed binomial offspring distribution $Bin\big(\widehat{Y}^{(N)}, 1-\vartheta_N\big)$, where the probability weights $\big(\hat{p}_k^{(N)}\big)_{k\in\mathbb{N}_0}$ of $\widehat{Y}^{(N)}$ are for all $1 \leq j \leq a_N$*

$$\hat{p}_j^{(N)} := \left(\frac{(v_N - 2a_N)^2}{2\delta_{N,u}}\right)^k \frac{1}{k!} \exp\left(-\frac{v_N^2}{2\delta_{N,\ell}}\right)\left(1 - \frac{1}{\delta_{N,\ell}^\delta}\right),$$

and

$$\hat{p}_0^{(N)} := 1 - \sum_{j=1}^{a_N} \hat{p}_j^{(N)}.$$

Denote by $\overline{Y}_{g,l}^{(N)} := \sum_{i=0}^g Y_{i,l}^{(N)}$ the total size of the Galton-Watson process until generation g and by $\overline{\mathbf{Y}_l} = \big(\overline{Y}_{g,l}^{(N)}\big)_{g\in\mathbb{N}_0}$ the corresponding process.

Now let $0 < \alpha < \beta$ and define

$$\overline{\sigma}_{N^\alpha}^{(N)} = \inf\left\{g \in \mathbb{N}_0 : |\widehat{\mathcal{I}}_g^{(N)} \cup \widehat{\mathcal{R}}_g^{(N)}| \geq N^\alpha\right\}.$$

One can show similarly as in proof of Lemma 3.6.4 Equation (3.6.13) that

$$\mathbb{P}\left(|\widetilde{\mathcal{I}}_g^{(N)} \cup \widetilde{\mathcal{R}}_g^{(N)}| \geq \overline{Y}_{g,l}^{(N)} \,\forall\, g \leq \overline{\sigma}_{N^\alpha}^{(N)}\right) \to 1$$

as $N \to \infty$. Indeed, as in the proof of Lemma 3.6.4 essentially we need to control the probability that a) an already empty vertex is re-attacked by at least two parasites moving along the same edge or b) a vertex gets simultaneously attacked by several pairs of parasites moving along the same edge.

In the following we will call pairs of parasites moving along the same edge packs of parasites. Similar as before we need to determine that each pack of parasites generated by an infected vertex before generation $\overline{\sigma}_{N^\alpha}^{(N)}$ is involved in one of the events a) or b) (independently of the other packs of parasites) with probability at most ϑ_N. In this case we can remove packs of parasites with probability ϑ_N such that the number of new infections generated by an infected host can with high probability be bounded from below by the number of offspring drawn according to the distribution with weights $(p_{k,l}^{(N)})_{k\in\mathbb{N}_0}$ from Definition 3.5.16 for any generation $n \leq \overline{\sigma}_{N^\alpha}^{(N)}$.

Next we determine an upper bound on the probabilities of the events a) and b).

a) Before generation $\overline{\sigma}_{N^\alpha}^{(N)}$ the probability that a pack of parasites originating from a vertex x attacks an already empty vertex is bounded from above by

$$\frac{N^\alpha}{\deg(x)} \leq \frac{N^\alpha}{\delta_{N,\ell}^\delta} = \frac{N^\alpha}{N^\beta - (d+1)N^{\frac{d+2}{d+3}\beta}} \in \Theta\Big(\frac{1}{N^{\beta-\alpha}}\Big).$$

b) Before generation $\overline{\sigma}^{(N)}_{N^\alpha}$, the number of empty vertices in the graph is smaller than N^α. The probability that two packs of parasites coming from 2 different vertices x and y attack the same vertex is bound by

$$\frac{|(\mathcal{N}_x \cap \mathcal{N}_y)\backslash\{x,y\}|}{\deg(x)\deg(y)} \leq \frac{1}{\deg(x)\wedge\deg(y)} \leq \frac{1}{\delta_{N,l}} \in \Theta\Big(\frac{1}{N^\beta}\Big),$$

where $\mathcal{N}_x = \{y \in \mathcal{V}^{(N)} : x \in \mathcal{E}^{(N)}\}$ denotes the neighbourhood of $x \in \mathcal{V}^{(N)}$. An application of Markov's inequality yields that the total number of packs of parasites generated before generation $\overline{\sigma}^{(N)}_{N^\alpha}$ is with high probability bounded by $N^\alpha \log(N)$, as in Lemma 4.8 in [1]. Hence, each pack of parasites is involved in an event of type b) with probability at most $N^\alpha \log(N) \cdot (\delta_{N,l})^{-1} = \Theta\left(\frac{\log(N)}{N^{\beta-\alpha}}\right)$.

Set $\vartheta_N := 2\frac{N^\alpha \log(N)}{\delta_{N,l}} \in \Theta\left(\frac{\log(N)}{N^{\beta-\alpha}}\right)$, then ϑ is an upper bound on the probability that a pack of parasites is involved in one of the events of type a) or b). For $\alpha < \beta$ we have $\vartheta_N \in o(1)$. By the exact same line of arguments as in Lemma 3.6.4 one can conclude that

$$\lim_{N\to\infty} \mathbb{P}\left(\exists g \in \mathbb{N}_0 : \overline{I}^{(N)}_g \geq N^\alpha\right) = 1 \tag{3.5.17}$$

for any $\alpha < \beta$. Using the same approach as in the proof of Lemma 3.5.11, where the only difference is a coupling from below with Galton-Watson processes instead of DBPC, one can show that under the event $\left\{\exists g \in \mathbb{N}_0 : \overline{I}^{(N)}_g \geq N^\alpha\right\}$ it exists a box of diameter r_N, in which at least $\frac{N^\alpha}{\log^2(N)}$ hosts got infected.
Taking $\frac{\beta}{2} < \alpha < \beta$ and using a similar approach as in the proof of Lemma 3.4.13, one shows that it exists a box of diameter r_N where all the hosts are killed and the number of infected individuals is of the order N^β. Then arguing as in Subsection 3.5.2 one shows that whp every vertices are killed by the infection process. Combined with (3.5.17) the result follows. □

3.5.6 Proof of Theorem 3.2.4 2)

Proof of Theorem 3.2.4 2). At each generation a parasite can move at most to a distance r_N meaning that the minimal number of generations the infection process $\overline{\mathbf{I}}^{(N)}$ needs for killing every host is at least the number of boxes of diameter r_N that separates the initial vertex to the boundaries of the domain. In dimension n, using the max norm, this number is exactly $\frac{1}{2r_N}$, giving the result $\mathbb{P}\big(\lfloor\frac{1}{2r_N}\rfloor \leq T^{(N)}\big) = 1$.
According to (3.5.7), Lemma 3.5.14 and applying a similar reasoning as in Lemma 3.4.13, and using that $\mathcal{U}_N(\beta/2+\xi)$ occurs whp, one shows that it exists $C > 0$ such that

$$\mathbb{P}\left(\widetilde{\sigma}^{(N)} \leq C\log(\log(N)) \;\middle|\; \widetilde{\sigma}^{(N)} < \infty\right) \underset{N\to\infty}{\longrightarrow} 1,$$

where $\widetilde{\sigma}^{(N)} := \inf\left\{g \in \mathbb{N} : \exists k \in \mathbf{K}, \overline{I}^{(N)}_g(k) = |\mathcal{V}^{(N)}_k|\right\}$. Moreover Equation (3.5.10) gives that after time $\sigma^{(N)}$, under the event $\mathcal{U}_N(\beta/2+\xi)$ which in particular guarantees that the graph $\mathcal{G}^{(N)}$ is connected, the remaining time up to total infection is upper bounded by $\frac{1}{2r_N}\left(1+\mathcal{O}\left(\frac{\varepsilon_N}{r_N}\right)\right)$ with

$\varepsilon_N = N^{\frac{\beta/2-1}{n}+\delta}$ for some $\delta > 0$ (where $\delta = \frac{\xi}{n}$ in Lemma 3.5.15). Combining these two facts gives that

$$\mathbb{P}\left(T^{(N)} \leq \left\lceil \frac{1}{2r_N} \right\rceil + \mathcal{O}(\kappa_N) \middle| T^{(N)} < \infty\right) \to 1,$$

with $\kappa_N := \max\left(\log(\log(N)), \frac{\varepsilon_N}{r_N^2}\right)$. □

Acknowledgements We thank Alison for her inspiring contributions to the theory of population genetics and wish her that her life stays as rich as it is! Furthermore, we thank two anonymous referees for careful reading of our manuscript and helpful comments. MS and HT were supported by the LOEWE programme of the state of Hessen (CMMS). CP acknowledges support from the German Research Foundation through grant PO-2590/1-1. HT acknowledges support from the German Research Foundation through grant ME-2088/5-1. VB and CP acknowledge support during the JTP 2022 "Stochastic Modelling in the Life Science" funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2047/1 – 390685813. VB would like to thank the Chair "Modélisation Mathématique et Biodiversité" of VEOLIA-Ecole Polytechnique-MNHN-F.X.

CRediT authorship contribution statement **Vianney Brouard:** Conceptualization, Methodology, Formal analysis, Investigation, Writing – original draft, Writing – review & editing **Cornelia Pokalyuk:** Conceptualization, Methodology, Formal analysis, Investigation, Writing – original draft, Writing – review & editing, Supervision, Funding acquisition **Marco Seiler:** Methodology, Formal analysis, Investigation, Writing – original draft, Writing – review & editing **Hung Tran:** Software, Validation, Investigation, Data Curation, Writing – original draft, Visualization.

3.6 Supplementary material

Proof of Lemma 3.3.1. If $p_{0,o} = 0$ the result follows by applying the extinction-explosion principle to the super-critical Galton-Watson process formed by the offspring generated by the initial individual. Due to assumption $p_{1,o} \neq 1$ this process is super-critical.

For the remaining cases we will show that all states except 0 are transient states, which yields the result.

Assume that $p_{0,o} > 0$ and that $p_{0,c} = 0$. Note that this means that we get at least one offspring from every possible cooperation of parents. Thus if we have $Z_g = k$ parents, we get at least $\frac{k(k-1)}{2}$ many offspring due to cooperation. But it holds that $\frac{k(k-1)}{2} > k$ for $k \geq 4$. Thus, if at some generation n we have that $Z_g \geq 4$, then we know that $Z_{g+1} > Z_g$ almost surely. This implies that $Z_g \to \infty$ as $n \to \infty$ almost surely, if $Z_{g_0} \geq 4$ for some $g_0 \in \mathbb{N}$. On the other hand since we exclude that $p_{0,o} = 1$ and $p_{1,c} = 1$ we have

$$\mathbb{P}(0 < Z_g \leq 3 \mid 0 < Z_{g-1} \leq 3) \leq c$$

for some $c < 1$. Consequently, the event $\{0 < Z_g \leq 3 \, \forall g \geq 0\}$ is a null-set and so all states but 0 are transient.

Assume that $p_{0,o} > 0$ and $p_{0,c} > 0$. Let us assume that in some generation g_0 we have $Z_{g_0} = k$ for some $k \geq 1$. If the process dies out in the next generation it enters the trap 0 such that it can

never return to k. Thus, an obvious lower bound for the probability to never hit k again is

$$\mathbb{P}\left(Z_g \neq k \forall g \geq g_0 | Z_{g_0} = k\right) \geq p_{0,o}^{k} p_{0,c}^{\binom{k}{2}} > 0$$

for $k \geq 1$. But this already implies that

$$\mathbb{P}\left(Z_g = k \text{ for some } g > g_0 \;\middle|\; Z_{g_0} = k\right) < 1,$$

i.e. the state k is transient.

□

Proof of Lemma 3.3.4. We set

$$E^{(N)}(k) := \mathbb{E}\left[Z_g^{(N)} \mid Z_{g-1}^{(N)} = k\right] = \left(k\mu_o^{(N)} + \binom{k}{2}\mu_c^{(N)}\right) \geq \binom{k+1}{2}\mu^{(N)},$$
$$V^{(N)}(k) := \mathbb{V}\left(Z_g^{(N)} | Z_{g-1}^{(N)} = k\right) = \left(k\nu_o^{(N)} + \binom{k}{2}\nu_c^{(N)}\right) \leq \binom{k+1}{2}\nu^{(N)},$$

where we define $\mu^{(N)} := \min\{\mu_o^{(N)}, \mu_c^{(N)}\}$ and $\nu^{(N)} := \max\{\nu_o^{(N)}, \nu_c^{(N)}\}$.

Since we assumed that the first and second moments of the offspring and cooperation distributions converge, it exists a N_0 such that

$$\frac{1}{2}\left(\frac{k(k+1)}{2}\mu\right) \leq E^{(N)}(k) \text{ and } V^{(N)}(k) \leq \frac{3}{2}\left(\frac{k(k+1)}{2}\nu\right),$$

for $N \geq N_0$. By Tchebychev's inequality for any $k \geq L$ we have

$$\begin{aligned}
\mathbb{P}\left(Z_g^{(N)} \geq \frac{k^2\mu}{8} \middle| Z_{g-1}^{(N)} = k\right) &\geq \mathbb{P}\left(Z_g^{(N)} \geq \frac{E^{(N)}(k)}{2} \middle| Z_{g-1}^{(N)} = k\right) \\
&\geq \mathbb{P}\left(\left|Z_g^{(N)} - E^{(N)}(k)\right| \leq \frac{E^{(N)}(k)}{2} \middle| Z_{g-1}^{(N)} = k\right) \\
&\geq 1 - \frac{4V^{(N)}(k)}{\left(E^{(N)}(k)\right)^2} \geq 1 - \frac{48(k(k+1)\nu)}{(k(k+1)\mu)^2} \geq 1 - \frac{48\nu}{k^2\mu^2}.
\end{aligned} \tag{3.6.1}$$

We choose $f_i(k) = \frac{k^{2^i}\mu^{2^i-1}}{8^{2^i-1}}$, where $f_0(k) = k$. Recall from (3.3.4) and (3.3.5) that we have

$$\frac{f_{i-1}^2(k)\mu}{8} = f_i(k) \quad \text{and} \quad f_i(k) > \frac{8(8+\nu)^{2^i}}{\mu}. \tag{3.6.2}$$

Now applying (3.6.1) and the first part of (3.6.2) recursively implies that

$$\mathbb{P}\left(\bigcap_{i=1}^{M}\left\{Z_{g+i}^{(N)} > f_i(k)\right\} \;\middle|\; Z_g^{(N)} = k\right) \geq \prod_{i=1}^{M}\left(1 - \frac{48\nu}{f_{i-1}^2(k)\mu^2}\right) = \prod_{i=1}^{M}\left(1 - \frac{6\nu}{f_i(k)\mu}\right)$$

and by the second part of (3.6.2) it follows $f_i(k)\mu > 8(8+\nu)^{2^i}$, which yields that

$$\mathbb{P}\left(\bigcap_{i=1}^{M}\left\{Z_{g+i}^{(N)} > f_i(k)\right\} \;\middle|\; Z_g^{(N)} = k\right) \geq \prod_{i=1}^{M}\left(1 - \frac{3}{4(8+\nu)^{2^i-1}}\right),$$

where we used that $\nu(8+\nu)^{-1} < 1$. □

Proof of Lemma 3.3.5. By Lemma 3.3.4 follows that if N_0 is large enough such that $b_N > L = \lceil \mu^{-1}(8+\nu)^2 \rceil$ for all $N \geq N_0$,

$$\mathbb{P}\left(\bigcap_{i=1}^{\infty} \left\{ Z_{g+i}^{(N)} > f_i(b_N) \right\} \,\middle|\, Z_g^{(N)} = b_N \right) \geq \prod_{i=1}^{\infty} \left(1 - \frac{6\nu}{f_i(b_N)\mu} \right),$$

where $f_i(b_N) = \frac{b_N^{2^i} \mu^{2^i-1}}{8^{2^i-1}}$ and $f_0(b_N) = b_N$. Without loss of generality we can assume that b_N is monotonically increasing in N, which implies that $\log\left(1 - \frac{6\nu}{f_i(b_N)\mu}\right)$ is monotonically increasing to 0 as $N \to \infty$. Furthermore, note that for N large enough $\inf_{i \geq 0} \frac{6\nu}{f_i(b_N)\mu} \leq \frac{1}{2}$, and thus

$$0 \geq \sum_{i=1}^{\infty} \log\left(1 - \frac{6\nu}{f_i(b_N)\mu}\right) \geq -\sum_{i=1}^{\infty} \frac{12\nu}{f_i(b_N)\mu} > -\infty,$$

for all N large enough, where we used that $1 - x \geq e^{-2x}$ for $x \in [0, \frac{1}{2}]$. Now by using continuity of $\exp(\cdot)$ and $\log(\cdot)$ and applying the monotone convergence theorem we obtain that

$$\lim_{N\to\infty} \exp\left(\log\left(\prod_{i=1}^{\infty} \left(1 - \frac{6\nu}{f_i(b_N)\mu}\right)\right)\right) = \exp\left(\sum_{i=1}^{\infty} \log\left(\lim_{N\to\infty} \left(1 - \frac{6\nu}{f_i(b_N)\mu}\right)\right)\right) = 1,$$

since $1 - \frac{6\nu}{f_i(b_N)\mu} \to 1$ as $N \to \infty$ for all $i \geq 1$. □

A consequence of the extinction-explosion principle is the following lemma, which states that for a DBPC the probability of reaching an arbitrary high level, that tends to ∞, at some generation or up to some generation is asymptotically the same as surviving. It is a special case of Proposition 3.3.7 when $\mathbf{Z}^{(N)} \equiv \mathbf{Z}$. We need it to prove Proposition 3.3.7 and other statements.

Lemma 3.6.1. *Let $\mathbf{Z}$ be a DBPC with survival probability $\pi > 0$ and satisfying $p_{1,o} \neq 1$ and $(p_{0,o}, p_{1,c}) \neq (1,1)$. Then for any sequence $(b_N)_{N\in\mathbb{N}}$ satisfying $b_N \to \infty$ we have*

$$\begin{aligned} \lim_{N\to\infty} \mathbb{P}\left(\exists g \in \mathbb{N}_0 : \overline{Z}_g \geq b_N\right) &= \lim_{N\to\infty} \mathbb{P}\left(\exists g \in \mathbb{N}_0 : Z_g \geq b_N\right) \\ &= \pi. \end{aligned}$$

The proof follows basically along the same arguments as the corresponding Lemma 3.7 in [1].

Proof of Lemma 3.6.1. First we will show that

$$\mathbb{P}\left(\exists g \in \mathbb{N}_0 : Z_g \geq b_N\right) \to \pi.$$

By the extinction-explosion principle for DBPC, proven in Lemma 3.3.1, we have that $\pi \leq \mathbb{P}\left(\exists g \in \mathbb{N}_0 : Z_g \geq b_N\right)$. Then

$$\begin{aligned} \pi &= \mathbb{P}\left(Z_g > 0, \forall g \in \mathbb{N}_0\right) \\ &= \mathbb{P}\left(\{\exists g \in \mathbb{N}_0 : Z_g \geq b_N\} \cap \{Z_g > 0, \forall g \in \mathbb{N}_0\}\right) \\ &= \mathbb{P}\left(\{\exists g \in \mathbb{N}_0 : Z_g \geq b_N\}\right) \cdot \mathbb{P}\left(\{Z_g > 0, \forall g \in \mathbb{N}_0\} | \{\exists g \in \mathbb{N}_0 : Z_g \geq b_N\}\right). \end{aligned}$$

Using the strong Markov property, one can show that

$$\mathbb{P}\left(\{Z_g > 0, \forall g \in \mathbb{N}_0\} \,\middle|\, \{\exists g \in \mathbb{N}_0 : Z_g \geq b_N\}\right) \geq \mathbb{P}\left(Z_g > 0, \forall g \in \mathbb{N}_0 \mid Z_0 = b_N\right).$$

Then because the interaction is a cooperative one, a DBPC starting in b_N can be coupled with b_N independent DBPCs starting in 1 such that we get

$$\mathbb{P}\left(\exists g \in \mathbb{N}_0 : Z_g = 0 \mid Z_0 = b_N\right) \leq \left(\mathbb{P}\left(\exists g \in \mathbb{N}_0 : Z_g = 0 | Z_0 = 1\right)\right)^{b_N}.$$

Introducing $q < 1$ as the extinction probability of the DBPC starting in 1, we finally obtain

$$\pi \geq \mathbb{P}\left(\{\exists g \in \mathbb{N}_0 : Z_g \geq b_N\}\right)\left(1 - q^{b_N}\right).$$

It follows that

$$\mathbb{P}\left(\{\exists g \in \mathbb{N}_0 : Z_g \geq b_N\}\right) \leq \frac{\pi}{1 - q^{b_N}} \to \pi.$$

Hence we have shown that

$$\mathbb{P}\left(\exists g \in \mathbb{N}_0 : Z_g \geq b_N\right) \underset{N\to\infty}{\longrightarrow} \pi. \tag{3.6.3}$$

For proving the remaining equality it remains to show that

$$\mathbb{P}\left(\left\{\exists\ g \in \mathbb{N}_0 : \overline{Z}_g \geq b_N\right\} \cap \left\{\exists\ g \in \mathbb{N}_0 :\ Z_g = 0\right\}\right) = o(1)$$

due to the extinction-explosion principle for DBPCs. Let $(c_N)_{N\in\mathbb{N}}$ be a sequence with $c_N \underset{N\to\infty}{\to} \infty$ and $\frac{b_N}{c_N} \underset{N\to\infty}{\to} \infty$ and consider the subsets

$$A^{(N)} := \left\{\exists\ g \in \mathbb{N}_0 : \overline{Z}_g \geq b_N,\ \exists\ i \leq g,\ Z_i \geq c_N\right\} \cap \left\{\exists\ g \in \mathbb{N}_0 :\ Z_g = 0\right\},$$
$$B^{(N)} := \left\{\exists\ g \in \mathbb{N}_0 : \overline{Z}_g \geq b_N,\ \forall\ i \leq g,\ Z_i < c_N\right\} \cap \left\{\exists\ g \in \mathbb{N}_0 :\ Z_g = 0\right\}.$$

By definition

$$A^{(N)} \sqcup B^{(N)} := \left\{\exists\ g \in \mathbb{N}_0 : \overline{Z}_g \geq b_N\right\} \cap \left\{\exists\ g \in \mathbb{N}_0 :\ Z_g = 0\right\}.$$

The extinction-explosion principle together with (3.6.3) yields that

$$\mathbb{P}\left(A^{(N)}\right) \leq \mathbb{P}\left(\left\{\exists\ g \in \mathbb{N}_0,\ Z_g \geq c_N\right\} \cap \left\{\exists\ g \in \mathbb{N}_0,\ Z_g = 0\right\}\right) \underset{N\to\infty}{\to} 0.$$

Furthermore

$$B^{(N)} \subset \left\{Z_{\lfloor \frac{b_N}{c_N} \rfloor} > 0\right\} \cap \left\{\exists\ g \in \mathbb{N}_0,\ Z_g = 0\right\},$$

which gives

$$\begin{aligned}\mathbb{P}\left(B^{(N)^c}\right) &\geq \mathbb{P}\left(\left\{Z_{\lfloor \frac{b_N}{c_N} \rfloor} = 0\right\} \sqcup \left\{\forall\ g \in \mathbb{N}_0,\ Z_g > 0\right\}\right) \\ &= \mathbb{P}\left(\left\{Z_{\lfloor \frac{b_N}{c_N} \rfloor} = 0\right\}\right) + \mathbb{P}\left(\left\{\forall\ g \in \mathbb{N}_0,\ Z_g > 0\right\}\right) \to 1,\end{aligned}$$

because for any sequence u_N tending to infinity we have $\mathbb{P}(Z_{u_N} = 0) \to 1 - \pi$, which follows by monotonicity of the events since $\{Z_{g+1} = 0\} \subset \{Z_g = 0\}$ for all $g \geq 0$. Hence, we have

$$\mathbb{P}\left(A^{(N)} \sqcup B^{(N)}\right) \underset{N\to\infty}{\to} 0.$$

□

Proof of Proposition 3.3.7. Due to Assumption 3.3.3 we have that neither $p_{1,o}^{(N)} \neq 1$ nor $(p_{0,o}^{(N)}, p_{1,c}^{(N)}) \neq (1,1)$ for N large enough. Due to Lemma 3.3.1, for an arbitrary $(b_N)_{N\in\mathbb{N}}$ such that $b_N \to \infty$ we have that it exists A_N such that $\mathbb{P}(A_N) = 0$ and

$$\{\forall g \in \mathbb{N}_0 : Z_g^{(N)} > 0\} \backslash A_N \subset \{\exists g \in \mathbb{N}_0 : Z_g^{(N)} \geq b_N\}.$$

Then using Lemma 3.3.5 we obtain that

$$\mathbb{P}\left(\{\exists g \in \mathbb{N}_0 : Z_g^{(N)} \geq b_N\} \cap \{\exists g \in \mathbb{N}_0 : Z_g^{(N)} = 0\}\right) \to 0.$$

Consequently using that $\mathbb{P}(\bigcup_{N\in\mathbb{N}} A_N) = 0$ it follows that if the limit exists it satisfies

$$\lim_{N\to\infty} \mathbb{P}\left(\forall g \in \mathbb{N}_0 : Z_g^{(N)} > 0\right) = \lim_{N\to\infty} \mathbb{P}\left(\exists g \in \mathbb{N}_0 : Z_g^{(N)} \geq b_N\right).$$

Let $(c_N)_{N\in\mathbb{N}}$ be a sequence with $c_N \to \infty$ and $\frac{b_N}{c_N} \to \infty$. In order to show, if the limit exists, that

$$\lim_{N\to\infty} \mathbb{P}\left(\exists g \in \mathbb{N}_0 : Z_g^{(N)} \geq b_N\right) = \lim_{N\to\infty} \mathbb{P}\left(\exists g \in \mathbb{N}_0 : \overline{Z}_g^{(N)} \geq b_N\right),$$

it remains to show that

$$\mathbb{P}\left(\exists g \in \mathbb{N}_0 : \overline{Z}_g^{(N)} \geq b_N \text{ and } \forall i \leq g, Z_i^{(N)} \leq c_N\right) \to 0. \tag{3.6.4}$$

In particular for $\tau_{b_N}^{\overline{\mathbf{Z}}^{(N)}} := \inf\{g \in \mathbb{N}_0 : \overline{Z}_g^{(N)} \geq b_N\}$ we have

$$\mathbb{P}\left(\exists g \in \mathbb{N}_0 : \overline{Z}_g^{(N)} \geq b_N \text{ and } \forall i \leq g, Z_i^{(N)} \leq c_N\right) \leq \mathbb{P}\left(\left\lceil \frac{b_N}{c_N} \right\rceil \leq \tau_{b_N}^{\overline{\mathbf{Z}}^{(N)}} < \infty\right).$$

But according to Proposition 3.3.6 it follows that

$$\mathbb{P}\left(\tau_{b_N}^{\overline{\mathbf{Z}}^{(N)}} \leq C \log\left(\log\left(b_N\right)\right) \mid \tau_{b_N}^{\overline{\mathbf{Z}}^{(N)}} < \infty\right) \to 1.$$

In particular taking $c_N \to \infty$ such that $\log\log(b_N) = o\left(\frac{b_N}{c_N}\right)$ implies that $\mathbb{P}\left(\lceil \frac{b_N}{c_N} \rceil \leq \tau_{b_N}^{\overline{\mathbf{Z}}^{(N)}} < \infty\right) \to 0$, which gives (3.6.4).

To conclude the proof it only remains to show that

$$\lim_{N\to\infty} \mathbb{P}\left(\exists g \in \mathbb{N}_0 : \overline{Z}_g^{(N)} \geq b_N\right) = \pi.$$

Note that $\sum_{k=0}^{\infty} |p_{k,o}^{(N)} - p_{k,o}| \leq 2$, and thus by dominated convergence it follows that

$$\lim_{N\to\infty} \sum_{k=0}^{\infty} \left|p_{k,o}^{(N)} - p_{k,o}\right| = 0.$$

Analogously follows that $\lim_{N\to\infty}\sum_{k=0}^{\infty}|p_{k,c}^{(N)}-p_{k,c}|=0$. From this follows that for a given sequence $(K_N)_{N\in\mathbb{N}}\subset\mathbb{N}$ such that $K_N\to\infty$ as $N\to\infty$ we find a sequence $(\varepsilon_N)_{N\in\mathbb{N}}$ with $\varepsilon_N\to 0$ as $N\to\infty$ and with

$$\max\left(\sum_{k=0}^{K_N}|p_{k,o}^{(N)}-p_{k,o}|,\sum_{k=0}^{K_N}|p_{k,c}^{(N)}-p_{k,c}|\right)\leq\varepsilon_N, \tag{3.6.5}$$
$$\max\left(\sum_{k=K_N+1}^{\infty}p_{k,o},\sum_{k=K_N+1}^{\infty}p_{k,c}\right)\leq\varepsilon_N.$$

Note that this implies

$$\begin{aligned}\sum_{k=K_N+1}^{\infty}p_{k,o}^{(N)}&\leq\left|\sum_{k=K_N+1}^{\infty}(p_{k,o}^{(N)}-p_{k,o})\right|+\sum_{k=K_N+1}^{\infty}p_{k,o}\\&\leq\left|\sum_{k=0}^{K_N}(p_{k,o}^{(N)}-p_{k,o})\right|+\varepsilon_N\leq 2\varepsilon_N.\end{aligned}$$

By the exact same calculation we get the same bound for the sum of $p_{k,c}^{(N)}$ from K_N+1 to ∞ and this yields that

$$\max\left(\sum_{k=K_N+1}^{\infty}p_{k,o}^{(N)},\sum_{k=K_N+1}^{\infty}p_{k,c}^{(N)}\right)\leq 2\varepsilon_N. \tag{3.6.6}$$

We know by assumption that $\varepsilon_N\to 0$. Consider now a sequence $(e_N)_{N\in\mathbb{N}}$ such that $e_N\to\infty$ and $\varepsilon_N e_N^2\to 0$. The first step is to prove that

$$\lim_{N\to\infty}\mathbb{P}\left(\exists g\in\mathbb{N}_0:\overline{Z}_g^{(N)}\geq e_N\right)=\pi. \tag{3.6.7}$$

We start by showing that whp. the sequence of DBPC $\left(\mathbf{Z}^{(N)}\right)_{N\in\mathbb{N}}$ and the limiting DBPC $\mathbf{Z}^{\infty}$ can be exactly coupled until their total size reaches the level e_N or they both die out. Introduce the stopping time of the first generation that the total size of $\mathbf{Z}^{\infty}$ reaches the level e_N or that it dies out as

$$\tau_{e_N,0}^{\overline{\mathbf{Z}}^{\infty}}=\inf\left\{g\in\mathbb{N}_0:\overline{Z}_g^{\infty}\geq e_N\text{ or }Z_g^{\infty}=0\right\}.$$

By definition we have that almost surely $\overline{Z}^{\infty}_{\tau^{\overline{\mathbf{Z}}^{\infty}}_{e_N,0}-1}<e_N$ which means that in order to make an exact coupling between $\mathbf{Z}^{(N)}$ and $\mathbf{Z}^{\infty}$ until generation $\tau_{e_N,0}^{\overline{\mathbf{Z}}^{\infty}}$, there are at most e_N^2 offspring and cooperation independent random variables to couple. Till this point we have not specified the joint distribution of the offspring and cooperation random variables $(X_i,Y_{j,k})_{i\in\mathbb{N},j<k}$ and $(X_i^{(N)},Y_{j,k}^{(N)})_{i\in\mathbb{N},j<k}$. We couple them in such a way that $\mathbb{P}\big(X_i\neq X_i^{(N)}\big)$ and $\mathbb{P}\big(Y_{j,k}\neq Y_{j,k}^{(N)}\big)$ are minimized. For a single random variable this can be done for each pair recursively via the maximal coupling, see Theorem 2.9 in [86], such that

$$\mathbb{P}\big(X_i\neq X_i^{(N)}\big)=\frac{1}{2}\sum_{k=0}^{\infty}|p_{k,o}^{(N)}-p_{k,o}|\leq 2\varepsilon_N$$

and

$$\mathbb{P}\big(Y_{j,k} \neq Y_{j,k}^{(N)}\big) = \frac{1}{2}\sum_{k=0}^{K_N} |p_{k,o}^{(N)} - p_{k,o}| \leq 2\varepsilon_N,$$

where we used the bounds from (3.6.5) and (3.6.6). Since these are families of independent random variables and also the offspring and cooperation random variables across different generations are independent, the probability that the e_N^2 relevant offspring and cooperation independent random variables are equal is lower bounded by $(1-2\varepsilon_N)^{e_N^2} \to 1$ by the choice of (e_N). In summary we have

$$\mathbb{P}\big(\overline{Z}_g^{(N)} = \overline{Z}_g^{\infty}, \forall g \leq \tau_{e_N,0}^{\overline{\mathbf{Z}}^{\infty}}\big) \geq (1-2\varepsilon_N)^{e_N^2} \to 1 \tag{3.6.8}$$

as $N \to \infty$. Let us now define the event

$$C_N := \left\{\overline{Z}_g^{(N)} = \overline{Z}_g^{\infty}, \forall g \leq \tau_{e_N,0}^{\overline{\mathbf{Z}}^{\infty}}\right\}.$$

We see that by monotonicity and Lemma 3.6.1 that

$$\mathbb{P}\left(\left\{\exists g \in \mathbb{N}_0 : \overline{Z}_g^{\infty} \geq e_N\right\} \cap C_N\right) \leq \mathbb{P}\left(\exists g \in \mathbb{N}_0 : \overline{Z}_g^{\infty} \geq e_N\right) \to \pi$$

as $N \to \infty$. On the other hand, by monotonicity and Equation (3.6.8) we see that

$$\mathbb{P}\left(\left\{\exists g \in \mathbb{N}_0 : \overline{Z}_g^{\infty} \geq e_N\right\} \cap C_N^c\right) \leq \mathbb{P}\left(\exists g \leq \tau_{e_N,0}^{\overline{\mathbf{Z}}^{\infty}} : \overline{Z}_g^{(N)} \neq \overline{Z}_g^{\infty}\right) \to 0$$

as $N \to \infty$. This yields that

$$\lim_{N\to\infty} \mathbb{P}\left(\left\{\exists g \in \mathbb{N}_0 : \overline{Z}_g^{\infty} \geq e_N\right\} \cap C_N\right) = \lim_{N\to\infty} \mathbb{P}\left(\exists g \in \mathbb{N}_0 : \overline{Z}_g^{\infty} \geq e_N\right) = \pi.$$

But C_N states that the Z^{∞} and $Z^{(N)}$ are coupled until $\tau_{e_N,0}^{\overline{\mathbf{Z}}^{\infty}}$, and therefore we also know that

$$\mathbb{P}\left(\left\{\exists g \in \mathbb{N}_0 : \overline{Z}_g^{\infty} \geq e_N\right\} \cap C_N\right) = \mathbb{P}\left(\left\{\exists g \in \mathbb{N}_0 : \overline{Z}_g^{(N)} \geq e_N\right\} \cap C_N\right)$$

for all $N > 0$. But this equality already implies Equation (3.6.7), i.e.

$$\lim_{N\to\infty} \mathbb{P}\left(\left\{\exists g \in \mathbb{N}_0 : \overline{Z}_g^{(N)} \geq e_N\right\}\right) = \pi.$$

This concludes the proof since we have shown it for $b_N = e_N$ and since we have shown it for one specific choice it follows also for an arbitrary sequence $(b_N)_{N\in\mathbb{N}}$ because of the extinction-explosion principle shown in Lemma 3.3.1. □

Lemma 3.6.2. *Consider sequences (D_N), (V_N), (m_N), (f_N), (g_N), (h_N), (k_N) such that $V_N \sim a\sqrt{D_N}$ for some $a > 0$, $f_N \geq h_N \geq 0$, $g_N, m_N \geq 0$ and assume $k_N \geq \max\{m_N, f_N, g_N\}$ as well as $k_N \to \infty$ for $N \to \infty$ and $\frac{k_N^4 V_N^3}{D_N^2} \in o(1)$. Then*

$$\begin{aligned}\frac{(D_N - f_N)!}{(D_N - f_N - (m_N V_N - g_N))!(D_N - h_N)^{m_N V_N - g_N}} &\geq \exp\left(-\frac{(m_N V_N - g_N)^2}{2D_N}\right)\exp\left(-\frac{k_N^4 V_N^3}{D_N^2}\right)\\ &\geq \exp\left(-\frac{(m_N V_N)^2}{2D_N}\right)\exp\left(-\frac{k_N^4 V_N^3}{D_N^2}\right).\end{aligned}$$

On the other hand we have

$$\frac{(D_N - f_N)!}{(D_N - f_N - (m_N V_N - g_N))!(D_N - h_N)^{m_N V_N - g_N}} \leq \exp\left(-\frac{(m_N V_N - g_N)^2}{2D_N}\right) \exp\left(\frac{m_N V_N}{D_N}\right).$$

Proof of Lemma 3.6.2. For completeness we show the inequality

$$1 - x \geq \exp(-x)\exp(-x^2) \tag{3.6.9}$$

for $x \in [0, \frac{1}{2}]$ first. We have

$$1 - x + x^2/2 \geq \exp(-x),$$

so

$$(1-x)\left(1 + \frac{x^2}{2(1-x)}\right) \geq \exp(-x)$$

which is equivalent to

$$1 - x \geq \exp(-x)\frac{1}{\left(1 + \frac{x^2}{2(1-x)}\right)}.$$

We have

$$1 + \frac{x^2}{2(1-x)} \leq 1 + x^2 < \exp(x^2)$$

which yields

$$1 - x \geq \exp(-x)\exp(-x^2).$$

We have

$$\begin{aligned} &\frac{(D_N - f_N)!}{(D_N - f_N - (m_N V_N - g_N))!(D_N - h_N)^{m_N V_N - g_N}} \\ &\qquad = \left(1 - \frac{f_N - h_N}{D_N - h_N}\right) \cdots \left(1 - \frac{f_N - h_N}{D_N - h_N} - \frac{m_N V_N - g_N - 1}{D_N - h_N}\right) \end{aligned} \tag{3.6.10}$$

Since $f_N - h_N \geq 0$ and $D_N - h_N \to \infty$ as $N \to \infty$ by assumption, we can estimate for N large enough by inequality (3.6.9)

$$\begin{aligned} (3.6.10) \geq \exp\left(\frac{-(f_N - h_N)(m_N V_N - g_N)}{D_N - h_N}\right) \cdot \exp\left(-\sum_{i=1}^{m_N V_N - g_N - 1} \frac{i}{D_N - h_N}\right) \\ \cdot \exp\left(-\left(\frac{f_N - h_N + m_N V_N}{D_N - h_N}\right)^2 m_N V_N\right) \end{aligned} \tag{3.6.11}$$

Finally, since $\frac{1}{D_N - h_N} \leq \frac{1}{D_N}(1 + 2\frac{h_N}{D_N})$ and

$$\left(\frac{(f_N + m_N V_N)^2 m_N V_N}{D_N^2} + \frac{f_N(m_N V_N)}{D_N}\right)\left(1 + \frac{5h_N}{D_N}\right) + \frac{h_N(m_N V_N - g_N)^2}{D_N} \leq \frac{k_N^4 V_N^3}{D_N^2}$$

for N large enough, we can estimate

$$(3.6.11) \geq \exp\left(-\frac{(m_N V_N - g_N)^2}{2D_N}\right)\exp\left(-\frac{k_N^4 V_N^3}{D_N^2}\right).$$

Furthermore, for the upper bound we estimate

$$\begin{aligned}
&\frac{(D_N - f_N)!}{(D_N - f_N - (m_N V_N - g_N))!(D_N - h_N)^{m_N V_N - g_N}} \\
&\qquad \leq \exp\left(\frac{-(f_N - h_N)(m_N V_N - g_N)}{D_N - h_N}\right)\cdot \exp\left(-\sum_{i=1}^{m_N V_N - g_N - 1}\frac{i}{D_N - h_N}\right) \\
&\qquad \leq \exp\left(-\frac{(m_N V_N - g_N - 1)^2}{2D_N}\right) \\
&\qquad \leq \exp\left(-\frac{(m_N V_N - g_N)^2}{2D_N}\right)\exp\left(\frac{m_N V_N}{D_N}\right).
\end{aligned}$$

□

Proof of Lemma 3.4.1. We have that

$$\begin{aligned}
\mathbb{P}\left(C_k^{(h'_N)}\right) = \frac{\prod_{i=1}^k \binom{m'_N V'_N - 2(i-1)}{2}}{k!(D'_N - m'_N)^k}\cdot \frac{(D'_N - h'_N)!}{(D'_N - h'_N - k)!(D'_N - m'_N)^k} \\
\cdot \frac{(D'_N - m'_N - k)!}{[(D'_N - m'_N - k) - (m'_N V'_N - 2k)]!(D'_N - m'_N)^{m'_N V'_N - 2k}}.
\end{aligned}$$

In particular applying Lemma 3.6.2 with $f_N = m'_N + k$, $g_N = 2k$, $h_N = m'_N$ and $k_N = \ell'_N$, one obtains that

$$\begin{aligned}
\mathbb{P}\left(C_k^{(h'_N)}\right) &\leq \left(\frac{m'_N V'_N}{2D'_N}\right)^k \frac{1}{k!}\exp\left(\ell'_N \frac{m'_N}{D'_N}\right)\cdot \exp\left(-\frac{(m'_N V'_N - 2\ell'_N)^2}{2D'_N}\right)\exp\left(\frac{m'_N V'_N}{D'_N}\right) \\
&\leq \left(\frac{m'_N V'_N}{2D'_N}\right)^k \frac{1}{k!}\exp\left(-\frac{(m'_N V'_N - 2\ell'_N)^2}{2D'_N}\right)\cdot \exp\left(\frac{{\ell'_N}^2 V'_N}{D'_N}\right).
\end{aligned}$$

Applying again Lemma 3.6.2 with $f_N = m'_N + k$, $g_N = 2k$, $h_N = m'_N$ and $k_N = \ell'_N$ we obtain

$$\begin{aligned}
\mathbb{P}\left(C_k^{(h'_N)}\right) &\geq \left(\frac{(m'_N V'_N - 2\ell'_N)^2}{2D'_N}\right)^k \frac{1}{k!}\exp\left(-4\frac{{\ell'_N}^2}{D'_N - m'_N}\right)\exp\left(-\frac{(m'_N V'_N)^2}{2D'_N}\right)\exp\left(-\frac{{\ell'_N}^4 {V'_N}^3}{{D'_N}^2}\right) \\
&\geq \left(\frac{(m'_N V'_N - 2\ell'_N)^2}{2D'_N}\right)^k \frac{1}{k!}\exp\left(-\frac{(m'_N V'_N - 2\ell'_N)^2}{2D'_N}\right)\exp\left(-\frac{{\ell'_N}^5 {V'_N}^3}{{D'_N}^2}\right).
\end{aligned}$$

□

Lemma 3.6.3. *Let X and Y be two random variables with state spaces $S = \{s_1, ..., s_n\}$ and $\tilde{S} = \{\tilde{s}_1, ..., \tilde{s}_n\}$ of size n for some $n \in \mathbb{N}$. Assume there exists an $\varepsilon > 0$ such that for all $i \in \{1, ..., n\}$ we have*

$$|\mathbb{P}(X = s_i) - \mathbb{P}(Y = \tilde{s}_i)| \leq \varepsilon.$$

Then there exist a probability space $(\Omega, \mathcal{F}, \tilde{\mathbb{P}})$ and random variables $\tilde{X}$ and $\tilde{Y}$ defined on $(\Omega, \mathcal{F}, \tilde{\mathbb{P}})$, such that $\tilde{X} \sim X$, $\tilde{Y} \sim Y$ and $\tilde{\mathbb{P}}\left((\tilde{X}, \tilde{Y}) \in \bigcup_{i=1}^{n} \{(s_i, \tilde{s}_i)\}\right) \geq 1 - n\varepsilon$. By a (common) slight abuse of notation, we will write (here and in the main text) that we can couple X and Y such that

$$\mathbb{P}\left((X, Y) \in \bigcup_{i=1}^{n} \{(s_i, \tilde{s}_i)\}\right) \geq 1 - n\varepsilon.$$

Proof. Denote by $p_i := \min\{\mathbb{P}(X = s_i), \mathbb{P}(Y = \tilde{s}_i)\}$. Let U be a random variable defined on some probability space $(\Omega, \mathcal{F}, \tilde{\mathbb{P}})$ and uniformly distributed on $[0, 1]$. Set $\tilde{X}(\omega) := s_k$ and $\tilde{Y}(\omega) := \tilde{s}_k$, if $\sum_{i=1}^{k-1} p_i \leq U(\omega) \leq \sum_{i=1}^{k} p_i$, where we set $\sum_{\emptyset} := 0$. Furthermore, let $p_i^X := \mathbb{P}(X = s_i) - p_i$ and set $\tilde{X}(\omega) := s_j$, if $\sum_{i=1}^{n} p_i + \sum_{i=1}^{j-1} p_i^X \leq U(\omega) \leq \sum_{i=1}^{n} p_i + \sum_{i=1}^{j} p_i^X$ and analogously define $\tilde{Y}(\omega)$ for $U(\omega) > \sum_{j=1}^{n} p_j$. Then X and Y have the same distribution as $\tilde{X}$ and $\tilde{Y}$, because e.g. $\tilde{\mathbb{P}}(\tilde{X} = s_i) = p_i + p_i^X = \mathbb{P}(X = s_i)$. Furthermore, $\tilde{\mathbb{P}}\big((\tilde{X}, \tilde{Y}) \in \cup_{i=1}^{n}(s_i, \tilde{s}_i)\big) \geq 1 - n\varepsilon$, since $\sum_{i=1}^{n} p_i \geq 1 - \sum_{i=1}^{n} |\mathbb{P}(X = s_i) - \mathbb{P}(Y = \tilde{s}_i)| \geq 1 - n\varepsilon$. □

Proof of Lemma 3.4.8. Denote by $S_N = D_N - \varphi_1(N) - \varphi_2(N)$, $b_N = D_N - \varphi_1(N)$, and $h(N) = H(N)(v_N - \varphi_3(N))$. Introduce for all $i \leq b_N$ the random variable $G_i^{(N)}$ which counts the number of balls in box i. We have

$$\begin{aligned}
&\mathbb{P}\left(G_1^{(N)} \leq 1\right) = \mathbb{P}\left(G_1^{(N)} = 0\right) + \mathbb{P}\left(G_1^{(N)} = 1\right) \\
&= \left(1 - \frac{1}{b_N}\right)^{h(N)} + \frac{h(N)}{b_N}\left(1 - \frac{1}{b_N}\right)^{h(N)-1} \\
&= \exp\left(h(N)\log\left(1 - \frac{1}{b_N}\right)\right)\left[1 + \frac{h(N)}{b_N}\frac{1}{1 - \frac{1}{b_N}}\right] \\
&= \left[1 - \frac{h(N)}{b_N} + \frac{1}{2}\left(\frac{h(N)}{b_N}\right)^2 + \mathcal{O}\left(\left(\frac{h(N)}{b_N}\right)^3\right)\right]\left(1 - \mathcal{O}\left(\frac{h(N)}{b_N^2}\right)\right)\left[1 + \frac{h(N)}{b_N} + \mathcal{O}\left(\frac{h(N)}{b_N^2}\right)\right] \\
&= 1 - \frac{1}{2}\left(\frac{h(N)}{b_N}\right)^2 + \mathcal{O}\left(\left(\frac{h(N)}{b_N}\right)^3\right).
\end{aligned}$$

Combining this previous computation with the fact that $S_N \frac{h^2(N)}{b_N^2} = \Theta(H^2(N))$ gives that

$$\mathbb{E}\left[G^{(N)}\right] = \mathbb{E}\left[\sum_{i=1}^{S_N} \mathbb{1}_{\left\{G_i^{(N)} \geq 2\right\}}\right] = S_N\left(1 - \mathbb{P}\left(G_1^{(N)} \leq 1\right)\right) = \Theta\left(H^2(N)\right).$$

By using Markov Inequality we obtain that

$$\mathbb{P}\left(G^{(N)} \geq H^2(N) f_2(N)\right) \leq \frac{\mathbb{E}[G^{(N)}]}{H^2(N) f_2(N)} = \Theta\left(\frac{1}{f_2(N)}\right),$$

which gives exactly (3.4.12).
Due to the scaling of $H(N)$ it could happen that some boxes contain more than 3 balls. In order to deal with such a situation introduce

$$\ell := \inf\left\{i \geq 2 : H(N)^{i+1} = o\left(\sqrt{D_N}^{i-1}\right)\right\}.$$

The scaling $N^\varepsilon = o(H(N))$ ensures that $\ell < \infty$. Now we have

$$\mathbb{P}\left(\exists i \leq b_N, G_i^{(N)} \geq \ell + 1\right) \leq b_N \frac{h(N)^{\ell+1}}{b_N^{\ell+1}} = \Theta\left(\frac{H(N)^{\ell+1} V_N^{\ell+1}}{D_N^\ell}\right) = \Theta\left(\frac{H(N)^{\ell+1}}{\sqrt{D_N}^{\ell-1}}\right) \to 0,$$

by definition of ℓ. Then under the event

$$\left\{\forall i \leq b_N, G_i^{(N)} \leq \ell\right\},$$

getting $G^{(N)} = k$ is obtained for any composition of balls, in which there exist exactly k groups of balls that consist of 2 to ℓ many balls and in which all other balls are grouped alone. The number of different kind of compositions is upper bounded by k^ℓ, because for all $2 \leq j \leq \ell$ the number of boxes getting exactly j balls is upper bounded by k. For each composition there are at most $(H(N)V_N)^{k\ell}$ different possibilities to fill the k groups with (at most ℓ) balls. An upper bound on the number of possibilities for the boxes, where the k groups of balls are placed to, is b_N^k. The number of remaining balls to be placed on different boxes is lower bounded by $h(N) - k\ell$. We can estimate in summary

$$\mathbb{P}\left(G^{(N)} \leq \frac{H^2(N)}{f_1(N)} \middle| \{\forall i \leq b_N, G_i \leq \ell\}\right) \leq \sum_{k=0}^{H^2(N) f_1^{-1}(N)} k^\ell H^{k\ell}(N) V_N^{k\ell} \prod_{i=0}^{h(N)-k\ell} \left(1 - \frac{i}{b_N}\right). \quad (3.6.12)$$

Also using that

$$\prod_{i=0}^{h(N)-k\ell} \left(1 - \frac{i}{b_N}\right) = \exp\left(\sum_{i=0}^{h(N)-k\ell} \log\left(1 - \frac{i}{b_N}\right)\right)$$

and that $\log(1-x) \leq -x$ for all $x \in [0,1)$, one can show, using that $h(N) = o(b_N)$, that for all $k \leq H^2(N) f_1^{-1}(N)$ we have

$$\prod_{i=0}^{h(N)-k\ell} \left(1 - \frac{i}{b_N}\right) \leq \exp\left(-\sum_{i=0}^{h(N)-k\ell} \frac{i}{b_N}\right) \leq \exp\left(-\sum_{i=0}^{h(N)-H^2(N)\ell} \frac{i}{b_N}\right).$$

Because $\sum_{i=0}^{h(N)-H^2(N)\ell} \frac{i}{b_N} \sim \frac{H^2(N) v_N^2}{2 b_N} \sim H^2(N) \frac{a^2}{2}$, it follows that for N large enough we have

$$\exp\left(-\sum_{i=0}^{h(N)-H^2(N)\ell} \frac{i}{b_N}\right) \leq \exp\left(-H^2(N)\frac{a^2}{4}\right).$$

Then using the natural bound

$$k^\ell \left(H(N)V_N\right)^{k\ell} \leq H^{2\ell}(N) \left(H(N)V_N\right)^{H^2(N) f_1^{-1}(N) \ell}$$

for all $k \leq H^2(N) f_1^{-1}(N)$, we get that for N large enough

$$\begin{aligned}(3.6.12) &\leq \exp\left(-H^2(N)\frac{a^2}{4}\right) H^{2(\ell+1)}(N)\left(H(N)V_N\right)^{H^2(N)f_1^{-1}(N)\ell} \\ &= \exp\left(-\frac{H^2(N)}{4}\left[a^2 - 4\ell f_1^{-1}(N)\log\left(H(N)V_N\right)\right]\right) H^{2(\ell+1)}(N) \\ &\leq \exp\left(-\frac{H^2(N)a^2}{8}\right),\end{aligned}$$

where for the last inequality we use that $\log(H(N)V_N) = o(f_1(N))$.
Finally to conclude the proof we have

$$\begin{aligned}\mathbb{P}\left(G^{(N)} \leq \frac{H^2(N)}{f_1(N)}\right) &\leq \mathbb{P}\left(\exists i \leq b_N, G_i^{(N)} \geq \ell+1\right) + \mathbb{P}\left(G^{(N)} \leq \frac{H^2(N)}{f_1(N)} \,\middle|\, \{\forall i \leq b_N, G_i \leq \ell\}\right) \\ &\leq \Theta\left(\frac{H(N)^{\ell+1}}{\sqrt{D_N}^{\ell-1}}\right).\end{aligned}$$

□

Lemma 3.6.4. *Let $0 < \alpha < \beta$, $a > 0$ and consider the sequence of processes $\left(\widetilde{\mathcal{S}}^{(N,a)}, \widetilde{\mathcal{J}}^{(N,a)}, \widetilde{\mathcal{R}}^{(N,a)}\right)_{N\in\mathbb{N}}$ defined in Section 3.4.5. It holds*

$$\lim_{N\to\infty} \mathbb{P}\left(\left\{\exists g \geq 0 : \left|\widetilde{\mathcal{R}}_g^{(N,a)}\right| \geq N^\alpha\right\}\right) \geq \varphi_a,$$

where φ_a denotes the survival probability of a Galton-Watson process with $Poi\left(\frac{a^2}{2}\right)$ offspring distribution.

In the proof of Lemma 3.6.4 we will couple

$$\left(\widetilde{\mathcal{S}}^{(N,a)}, \widetilde{\mathcal{J}}^{(N,a)}, \widetilde{\mathcal{R}}^{(N,a)}\right)_N$$

with the Galton-Watson process defined next.

Definition 3.6.5. *(Lower Galton-Watson Process)*
Let $0 < \delta < \frac{1}{2}$ and $(b_N)_{N\in\mathbb{N}}$ be a sequence which fulfills $b_N \to \infty$ and $b_N \in o\left(\sqrt{D_N}\right)$. Furthermore assume $(\theta_N)_{N\in\mathbb{N}}$ is a $[0,1]$-valued sequence with $\theta_N \to 0$. Let $\mathbf{X}_l^{(N)} = \left(X_{g,l}^{(N)}\right)_{g\in\mathbb{N}_0}$ be a Galton-Watson process with mixed binomial offspring distribution $Bin\left(\widetilde{X}^{(N)}, 1-\theta_N\right)$, where the probability weights $\left(\widetilde{p}_k^{(N)}\right)_{k\in\mathbb{N}_0}$ of $\widetilde{X}^{(N)}$ are for all $1 \leq j \leq b_N$

$$\widetilde{p}_j^{(N)} := \left(\frac{\left(V_N^{(a)} - 2b_N\right)^2}{2D_N}\right)^j \frac{1}{j!} \exp\left(-\frac{\left(V_N^{(a)}\right)^2}{2D_N}\right)\left(1 - \frac{1}{D_N^\delta}\right),$$

and

$$\widetilde{p}_0^{(N)} := 1 - \sum_{j=1}^{b_N} \widetilde{p}_j^{(N)}.$$

Denote by $\Phi_l^{(N)}$ *the generating function of the offspring distribution* $\left(p_{k,l}^{(N)}\right)_{k\in\mathbb{N}_0}$ *of* $\mathbf{X}_l^{(N)}$*, and by* $\overline{X}_{g,l}^{(N)} := \sum_{i=0}^{g} X_{i,l}^{(N)}$ *the total size of the Galton-Watson process until generation* g *and* $\overline{\mathbf{X}_l} = \left(\overline{X}_{g,l}^{(N)}\right)_{g\in\mathbb{N}_0}$ *the corresponding process.*

Proof of Lemma 3.6.4. Introduce

$$\overline{\eta}_{N^\alpha}^{(N)} = \inf\left\{g \in \mathbb{N}_0 : \left|\widetilde{\mathcal{J}}_g^{(N,a)} \cup \widetilde{\mathcal{R}}_g^{(N,a)}\right| \geq N^\alpha\right\}.$$

We are going to show that

$$\mathbb{P}\left(|\widetilde{\mathcal{J}}_g^{(N,a)} \cup \widetilde{\mathcal{R}}_g^{(N,a)}| \geq \overline{X}_{g,l}^{(N)} \,\forall\, g \leq \overline{\eta}_{N^\alpha}^{(N)}\right) \to 1 \tag{3.6.13}$$

for the process $\overline{\mathbf{X}}_l^{(N)}$ defined in Definition 3.6.5.

Consider $D_N - 1$ boxes, assume V_N many balls are put uniformly at random into the boxes. Denote by $C_j^{(N)}$ the event that exactly j boxes contain at least 2 balls. One can show using similar calculations as in the proof of Proposition 3.5 of [1] that

$$\widetilde{p}_j^{(N)} \leq \mathbb{P}\left(C_j^{(N)}\right), \forall 1 \leq j \leq b_N,$$

This means that whenever a vertex x gets infected one can estimate from below how many of its neighbors are visited by at least 2 of its V_N parasites, which we call a *pack of parasites*, by the corresponding number of offspring in the Galton-Watson process $\mathbf{X}_l^{(N)}$, since $\widetilde{p}_0^{(N)} = 1 - \sum_{i=1}^{b_N} \tilde{p}_j^{(N)}$.

However, in the process $\left(\widetilde{\mathcal{S}}^{(N,a)}, \widetilde{\mathcal{J}}^{(N,a)}, \widetilde{\mathcal{R}}^{(N,a)}\right)$ "ghost" infections may occur, when a) an already empty vertex is attacked by at least 2 parasites coming from the same infected vertex, or when b) a vertex is attacked by at least two packs of parasites coming from different vertices.

We show next that each pack of parasites of size at least 2 generated by an infected vertex before generation $\overline{\eta}_{N^\alpha}^{(N)}$ is involved in one of the events a) or b) (independently of the other packs of parasites) with probability at most θ_N. Consequently, by removing packs of parasites with probability θ_N the number of new infection generated by an infected hosts can with high probability be bounded from below by the number of offspring drawn according to the distribution with weights $(p_{k,l}^{(N)})_{k\in\mathbb{N}_0}$ from Definition 3.6.5 for any generation $n \leq \overline{\eta}_{N^\alpha}^{(N)}$.

Now we upper bound the probabilities of the events a) and b).

a) Before generation $\overline{\eta}_{N^\alpha}^{(N)}$ the probability that a pack of parasites goes to an empty vertex is bounded from above by $\frac{N^\alpha}{D_N - 1} = \Theta\left(\frac{1}{N^{\beta-\alpha}}\right)$.

b) Before generation $\overline{\eta}_{N^\alpha}^{(N)}$, the number of empty vertices in the graph is smaller than N^α. The probability that two packs of parasites coming from 2 different vertices attack the same vertex is $\frac{1}{D_N - 2} = \Theta\left(\frac{1}{N^\beta}\right)$. Using Markov inequality one can show that the total number of packs of parasites generated before generation $\overline{\eta}_{N^\alpha}^{(N)}$ is with high probability bounded by $N^\alpha \log(N)$, as in Lemma 4.8 in [1]. Hence, each pack of parasites is involved in an event of type b) with probability at most $N^\alpha \log(N) \cdot \frac{1}{D_N - 2} = \Theta\left(\frac{\log(N)}{N^{\beta-\alpha}}\right)$.

In summary, $\theta_N = 2\frac{N^\alpha \log(N)}{D_N - 2} = \Theta\left(\frac{\log(N)}{N^{\beta-\alpha}}\right)$ yields an upper bound on the probability that a pack of parasites is involved in one of the events of type a) or b). Since $\alpha < \beta$ we have $\theta_N \in o(1)$. Then

taking the Galton Watson process $\mathbf{X}_l^{(N)}$ of Definition 3.6.5 with $\theta = 2\frac{N^\alpha \log(N)}{D_N - 2}$ we have exactly proven (3.6.13).

As in the proof of Proposition 3.3 of [1] one can show the uniform convergence of $\Phi_l^{(N)}$ to the generating function of a $\mathrm{Poi}\big(\frac{a^2}{2}\big)$ distribution, such that applying Lemma 3.7 of [1] gives that

$$\lim_{N\to\infty} \mathbb{P}\left(\exists g \in \mathbb{N}_0 : \overline{X}_{g,l}^{(N)} \geq N^\alpha\right) = \varphi_a. \tag{3.6.14}$$

Combining (3.6.13) and (3.6.14) we get that

$$\liminf_{N\to\infty} \mathbb{P}\left(\exists g \in \mathbb{N}_0 : |\widetilde{\mathcal{R}}_g^{(N,a)}| \geq N^\alpha\right) \geq \varphi_a.$$

□

For the proof of Lemma 3.6.6 we have to control moderate deviations of Poisson distributed random variables.

Lemma 3.6.6. *Let $0 < \gamma < 1$, $N \in \mathbb{N}$ and $X_N \sim Poi(N^\gamma)$. Then for any $0 < \varepsilon < \gamma/2$ it holds that*

$$\lim_{N\to\infty} \frac{-2}{N^{2\varepsilon}} \log\left(\mathbb{P}\left(|X_N - N^\gamma| > N^{\frac{\gamma}{2}+\varepsilon}\right)\right) = 1.$$

The proof of Lemma 3.5.1 is based on the following lemma.

Lemma 3.6.7. *Let $X = (X_t)$ be a Poisson process with intensity 1 on $[0,\infty)$ and let $a(t)$ be a function such that $\frac{t}{a(t)} \to \infty$ and $\frac{a(t)}{\sqrt{t}} \to \infty$. Then for every Borel-set $B \subset \mathbb{R}$ it holds that*

$$\limsup_{t\to\infty} \frac{t}{a(t)^2} \log\left(\mathbb{P}\Big(\frac{1}{a(t)}(X_t - t) \in B\Big)\right) \leq -\inf_{x\in\overline{B}} \frac{x^2}{2}$$

and

$$\liminf_{t\to\infty} \frac{t}{a(t)^2} \log\left(\mathbb{P}\Big(\frac{1}{a(t)}(X_t - t) \in B\Big)\right) \geq -\inf_{x\in\mathring{B}} \frac{x^2}{2},$$

where $\mathring{B}$ denote the interior and $\overline{B}$ the closure of B.

Proof of Lemma 3.6.7. This is a direct consequence of Theorem 1.1' found in [87]. We will now check the conditions of this theorem. First denote by $Y_t = X_t - t$ the compensated Poisson process, note that $(Y_t)_{t\geq 0}$ is a martingale. Let $\delta > 0$, then it holds that

$$\mathbb{E}\Big[\exp\big(\delta \sup_{s\leq t\leq s+1} |Y_t - Y_s|\big) \mid \sigma(Y_u : u \leq s)\Big] = \mathbb{E}\Big[\exp\big(\delta \sup_{0\leq t\leq 1} |Y_t|\big)\Big],$$

where we used that the Poisson process has independent and stationary increments and that $Y_0 = 0$. Furthermore it holds that

$$\mathbb{E}\Big[\exp\big(\delta \sup_{0\leq t\leq 1} |Y_t|\big)\Big] \leq \mathbb{E}\big[\exp\big(\delta X_1\big)\big] e^1 = \exp(\exp(\delta)),$$

where we used that X_t is a monotone process and that the moment generating function of a Poisson distribution with parameter 1 is given through $t \mapsto \exp(e^t - 1)$. This provides $(A1)'$. The second condition $(A2)'$ follows from the fact that

$$\mathbb{E}\Big[\frac{1}{t}(Y_{s+t} - Y_s)^2 - 1 \;\Big|\; \sigma(Y_u : u \leq s)\Big] = \mathbb{E}\Big[\frac{1}{t}Y_t^2\Big] - 1 = 0,$$

where we used again that the process has stationary and independent increments. Now the claimed moderate deviation principle follows from Theorem 1.1' in [87]. □

Proof of Lemma 3.6.6. This follows from Lemma 3.6.7 by choosing $a(t) = t^{\frac{1}{2}+\frac{\varepsilon}{\gamma}}$, $B = [0,1]^c$ and consider the subsequent $(X_{N^\gamma} - N^\gamma)_{N\geq 0}$ instead of $(X_t - t)_{t\geq 0}$. Then plugging the choices in we get that

$$\lim_{N\to\infty} \frac{1}{N^{2\varepsilon}} \log\left(\mathbb{P}\left(|X_{N^\gamma} - N^\gamma| > N^{\frac{\gamma}{2}+\varepsilon}\right)\right) = -\frac{1}{2}.$$

which provides the claim. □

Part II

Cancer Evolution

Chapter 4

Neutral (and Deleterious) Cancer Evolution

This Chapter corresponds to a major revision of the preprint [21] ≪Genetic Composition of Supercritical Branching Populations under Power Law Mutation Rates≫ for the Annals of Applied Probability .

Abstract:
We aim to understand the evolution of the genetic composition of cancer cell populations. To achieve this, we consider an individual-based model representing a cell population where cells divide, die and mutate along the edges of a finite directed graph (V, E). The process starts with only one wild-type cell. Following typical parameter values in cancer cell populations we study the model under *power law mutation rates*, in the sense that the mutation probabilities are parameterized by negative powers of a scaling parameter n and the typical sizes of the population of interest are positive powers of n. Under a *non-increasing growth rate condition*, we describe the time evolution of the first-order asymptotics of the size of each subpopulation in the $\log(n)$ time scale, as well as in the random time scale at which the wild-type population, resp. the total population, reaches the size n^t. In particular, such results allow for the perfect characterization of evolutionary pathways. Without imposing any conditions on the growth rates, we describe the time evolution of the order of magnitude of each subpopulation, whose asymptotic limits are positive non-decreasing piecewise linear continuous functions.
Keywords: cancer evolution, multitype branching processes, finite graph, long time behaviour, power law mutation rates, population genetics.
MSC2020 subject classifications: 60J80, 60J27, 60F99, 92D15, 92D25.

4.1 Introduction and presentation of the model

Consider a population of cells characterised by a phenotypic trait, where the trait space V is finite. For all $v \in V$ denote by $(Z_v(t))_{t\in\mathbb{R}^+}$ the number of cells of trait v at time t in the population, and $\big(\mathcal{Z}(t) := (Z_v(t))_{v\in V}\big)_{t\in\mathbb{R}^+}$ the global process. Assume that $0 \in V$ and

$$\forall v \in V, Z_v(0) = \mathbb{1}_{\{v=0\}}, \text{ almost surely.}$$

Cells with trait 0 are called *wild-type cells*, and all cells with trait $v \in V\backslash\{0\}$ are called *mutant cells*. The population dynamics will follow a continuous-time branching process on $\mathbb{N}_0^V$. More precisely, cells divide (giving birth to two daughter cells) and die with rates depending only on their phenotypic trait. The birth, death and growth rate functions are respectively

$$\alpha : V \longrightarrow \mathbb{R}^+, \beta : V \longrightarrow \mathbb{R}^+ \text{ and } \lambda := \alpha - \beta.$$

We use the words "division" and "birth" synonymously. During a division event of a cell of trait $v \in V$, the two daughter cells may independently mutate. The mutation landscape across traits is encoded via a directed graph structure (V, E) on the trait space, where $E \subset \{(v, u), \forall v, u \in V^2\}$ is a set of ordered pairs over V such that for all $v \in V$, $(v, v) \cap E = \emptyset$, and there exists a path from 0 to v within E. In other words, (V, E) represents a finite directed graph without self-loops, with each vertex belonging to the connected component of 0. Mutation directly from trait v to trait u is possible if and only if $(v, u) \in E$. Let $\mu : E \longrightarrow [0, 1]$ be a mutation kernel satisfying

$$\forall v \in V, \overline{\mu}(v) := \sum_{u \in V:(v,u)\in E} \mu(v, u) \leq 1.$$

A daughter cell mutates from its mother trait v to trait u with probability $\mu(v, u)$, meaning that $\overline{\mu}(v)$ is its total mutation probability. Notice that backward mutations are permitted in this model.

Finally the exact transition rates from a state $z = (z_v)_{v\in V} \in \mathbb{N}_0^V$ of the process $\mathcal{Z}$ are

$$z \mapsto \begin{cases} z - \delta_v, \text{ at rate } z_v\beta(v), \\ z - \delta_v + \delta_u + \delta_w, \text{ at rate } 2z_v\alpha(v)\mu(v, u)\mu(v, w)\mathbb{1}_{\{(v,u)\in E\}}\mathbb{1}_{\{(v,w)\in E\}}\mathbb{1}_{\{u\neq w\}}, \\ z - \delta_v + 2\delta_u, \text{ at rate } z_v\alpha(v)\mu(v, u)^2\mathbb{1}_{\{(v,u)\in E\}}, \\ z + \delta_v, \text{ at rate } z_v\alpha(v)\left(1 - \overline{\mu}(v)\right)^2 + 2\sum\limits_{u\in V:(u,v)\in E} z_u\alpha(u)\mu(u, v)\left(1 - \overline{\mu}(u)\right), \end{cases}$$

where $\forall v \in V, \delta_v = \left(\mathbb{1}_{\{u=v\}}\right)_{u\in V}$. Throughout the paper, the growth rate of the wild-type subpopulation $\lambda(0)$ is assumed to be strictly positive, to ensure that the wild-type subpopulation survives with positive probability.

The biological motivation for this model is to capture the time dynamics of the genetic composition of a cell population during carcinogenesis. Tumors are typically detected when they reach a large size, around 10^9 cells. The mutation rates per base pair per cell division are generally estimated to be of order 10^{-9}, see [4, 5]. Thus, the framework of a *power law mutation rates limit* naturally arises. A parameter $n \in \mathbb{N}$ is used to quantify both the decrease of the mutation probabilities, expressed as a negative power of n, and the typical population size, expressed as a positive power of n, at which we are interested in understanding the genetic composition. The aim is to obtain asymptotic results on the sizes of all the mutant subpopulations when n goes to infinity. This is a classical stochastic regime studied in particular in [6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17]. Such a regime is referred to in [8, 10] as the *large population rare mutations limit*. However, we have chosen the more precise term *power law mutation rates* to distinguish this regime from the classical *rare mutations limit*, which is generally used in the context of adaptive dynamics to separate evolutionary and ecological scales, where the mutation probabilities $\mu^{(n)}$ typically scale as $e^{-Cn} \ll \mu^{(n)} \ll \frac{1}{n\log(n)}$. Indeed, under the power law mutation rates limit, the mutation probabilities are of a higher order compared to those under the rare mutations limit if for instance $\mu^{(n)} \propto n^{-\alpha}$ with $\alpha \in (0, 1]$.

To be more precise, let $L := \{\ell(v,u) \in \mathbb{R}_+^*, \forall (v,u) \in E\}$ be a set of strictly positive labels on the edges of the graph, where $\mathbb{R}_+^* := \{x \in \mathbb{R}, x > 0\}$. Introduce a sequence of models $\left(\mathcal{Z}^{(n)}\right)_{n\in\mathbb{N}}$, where for each $n \in \mathbb{N}$, $\mathcal{Z}^{(n)}$ corresponds to the process described above with the mutation kernel $\mu^{(n)} : E \longrightarrow [0,1]$ satisfying

$$\forall (v,u) \in E, n^{\ell(v,u)}\mu^{(n)}(v,u) \underset{n\to\infty}{\longrightarrow} \mu(v,u) \in \mathbb{R}^+. \tag{4.1.1}$$

For all $t \in \mathbb{R}_+^*$, the stopping times corresponding to the first time that the wild-type subpopulation $Z_0^{(n)}$, respectively the total population $Z_{tot}^{(n)} := \sum_{v\in V} Z_v^{(n)}$, reaches the level n^t, are defined as

$$\eta_t^{(n)} := \inf\left\{u \in \mathbb{R}^+ : Z_0^{(n)}(u) \geq n^t\right\} \text{ and } \sigma_t^{(n)} := \inf\left\{u \in \mathbb{R}^+ : Z_{tot}^{(n)}(u) \geq n^t\right\}.$$

These are motivated by two different biological interpretations in different scenarios. For instance, when considering metastasis the wild-type subpopulation $Z_0^{(n)}$ may represent the primary tumor, and the mutant subpopulations $Z_v^{(n)}$, for all $v \in V\backslash\{0\}$, may correspond to secondary tumors. As clinicians typically have access to the size rather than the age of a tumor, it is biologically relevant to estimate the genetic composition of the secondary tumors when the primary one has reached a given size. This is mathematically encoded by examining the first-order asymptotics of $Z_v^{(n)}\left(\eta_t^{(n)}\right)$ for all $v \in V\backslash\{0\}$. Another biological scenario involves the total population $Z_{tot}^{(n)}$ representing a single tumor. It is appropriate to obtain theoretical results about the size of the mutant subpopulations $Z_v^{(n)}$ for all $v \in V\backslash\{0\}$ when the tumor has reached a given size. This corresponds exactly to looking at the first-order asymptotics of $Z_v^{(n)}\left(\sigma_t^{(n)}\right)$. Every time that results can be stated either with $\eta_t^{(n)}$ or $\sigma_t^{(n)}$, the following notation will be used

$$\rho_t^{(n)} := \eta_t^{(n)} \text{ or } \sigma_t^{(n)}. \tag{4.1.2}$$

In the present work the cell population will be studied on different time scales: the random time scale

$$\left(\rho_t^{(n)} + s\right)_{(t,s)\in\mathbb{R}^+\times\mathbb{R}}; \tag{4.1.3}$$

and the following deterministic approximation

$$\left(\mathfrak{t}_t^{(n)} + s\right)_{(t,s)\in\mathbb{R}^+\times\mathbb{R}}, \text{ with } \mathfrak{t}_t^{(n)} := t\frac{\log(n)}{\lambda(0)}. \tag{4.1.4}$$

Intuitively, the lineage of wild-type cells generated from the cancer-initiating cell constitutes the first subpopulation that will generate mutations. Understanding its growth, therefore, provides the natural time scale to consider for observing mutations. The birth and death rates of this lineage are $\alpha(0)\left(1-\overline{\mu}^{(n)}(0)\right)^2$ and $\beta(0)+\alpha(0)\left(\overline{\mu}^{(n)}(0)\right)^2$, respectively. Due to the power law mutation rates regime specified in Equation (4.1.1), these rates converge to $\alpha(0)$ and $\beta(0)$ when n grows to ∞. Consequently, this lineage should therefore behave asymptotically as a birth and death process with rates $\alpha(0)$ and $\beta(0)$. Indeed, such a result emerges from the natural martingale associated to a birth and death process, see Lemma 4.3.1. In particular the growth rate of this lineage is close to $\lambda(0)$, thus this population reaches a size of order n^t approximately at the deterministic time $\mathfrak{t}_t^{(n)}$, see Lemma 4.3.2.

For any finite directed labelled graph (V, E, L), under the following *non-increasing growth rate condition*

$$\forall v \in V, \lambda(v) \leq \lambda(0), \tag{4.1.5}$$

the first-order asymptotics of the mutant subpopulation sizes $Z_v^{(n)}$ are obtained on both random and deterministic time scales (4.1.3) and (4.1.4), see Theorem 4.2.7. Assumption (4.1.5) can be biologically motivated. Historically, tumour dynamics has been seen under the prism of clonal expansion of selective mutations, i.e. $\lambda(v) > \lambda(0)$. Nevertheless, the paradigm of neutral cancer evolution has recently been considered, see [40, 41, 3, 42, 43]. This means that all the selective mutations are already present in the cancer-initiating cell, and any mutations that occur subsequently are neutral (i.e. $\lambda(v) = \lambda(0)$). With Assumption (4.1.5), deleterious mutations (i.e. $\lambda(v) < \lambda(0)$) are also permitted. This paradigm has been introduced because the genetic heterogeneity inside a tumour could be explained by considering neutral mutations only. Various statistical methods have been developed to infer the evolutionary history of tumours, including test of neutral evolution, see [44, 45, 46] for details.

Without any assumption on the growth rate function λ, we study the system on the deterministic time scale of Equation (4.1.4). As in [6, 9, 11, 12, 13, 14, 15, 17], we obtain the asymptotic behaviour of the *stochastic exponent* processes

$$\forall v \in V, X_v^{(n)}(t) := \frac{\log^+\left(Z_v^{(n)}\left(\mathfrak{t}_t^{(n)}\right)\right)}{\log(n)/\lambda(0)}. \tag{4.1.6}$$

These results are presented in Theorem 4.2.9. Here we are tracking the exponent of n for each subpopulation, whereas Theorem 4.2.7 is a more refined result that gives the size directly in terms of n. To our knowledge, it is the first model capturing this level of refinement on the asymptotic behaviours under the power law mutation rates regime (4.1.1). Two significant new conclusions emerge.

First, Theorem 4.2.7 shows the remarkable result that under Assumption (4.1.5) the randomness in the first-order asymptotics of the size of any mutant subpopulation is fully described by the stochasticity of only one random variable W, which encodes the long-time randomness of the lineage of wild-type cells issued from the cancer-initiating cell. More precisely, the stochasticity for any mutant subpopulation size is fully driven, at least to first order, by the randomness in the growth of the wild-type subpopulation and not by the dynamics of any lineage of a mutant cell nor by the stochasticity generating the mutations.

Second, Theorem 4.2.7 characterises the exact effective evolutionary pathways, in the sense of the pathways that asymptotically contribute to the growth of the mutant subpopulations. More precisely, if the length of a pathway is defined as the sum of the labels of its edges, asymptotic results on the stochastic exponent give that for any trait v, among the pathways from 0 to v, only those of minimal length can asymptotically contribute to the growth of trait v. However, having results on the first-order asymptotics of the size of the mutant subpopulations allows us to see which of those minimal length pathways actually contribute to the dynamics of trait v. More specifically, among the minimal length pathways only those with the maximal number of neutral mutations on their edges asymptotically contribute to the growth of trait v. Indeed, for each neutral mutation in a pathway, an additional multiplicative factor of order $\log(n)$ appears in the first-order asymptotics. Such a theoretical result opens the door for developing new statistical methods to infer the underlying graph structure from data, i.e. to infer the evolutionary history of tumours, as

well as for designing new statistical estimators for biologically relevant parameters, alongside new neutral (and deleterious) cancer evolution tests.

Moreover it is, to our knowledge, the first time that this power law mutation rates limit has been studied on the random time scale of Equation (4.1.3). From a biological point of view, it is more interesting to obtain results on such a random time scale rather than a deterministic one. We find that the randomness in the first-order asymptotics of any mutant subpopulation size is fully described by the stochasticity in the survival of the lineage of wild-type cells issued from the cancer-initiating cell.

In [8, 10], Cheek and Antal study a model that can be seen as an application of the model of the present work via a specific finite directed labelled graph (V, E, L), the finite-dimensional hypercube. Among their results, they fully characterise, in distribution, the asymptotic sizes of all the mutant subpopulations around the random time at which the wild-type subpopulation reaches the typical size allowing mutations to occur. In their setting, it corresponds to $\left(\eta_1^{(n)} + s\right)_{s\in\mathbb{R}}$. In particular, they obtain that the asymptotic sizes of all the mutant subpopulations around this random time $\eta_1^{(n)}$ are finite almost surely, following generalised Luria-Delbrück distributions, see [10, Theorem 5.1]. The original Luria and Delbrück model, introduced in [88], has generated many subsequent works, see in particular [89, 90, 91, 92, 93, 8, 10]. Two major features explain the latter result. The first one is that asymptotically only a finite number of mutant cells are generated from the wild-type subpopulation until time $\eta_1^{(n)}$, following a Poisson distribution. The second one is that all the lineages of the mutant cells generated from mutational events of the wild-type subpopulation have, up to time $\eta_1^{(n)}$, only an asymptotically finite random time to grow, which is exponentially distributed. We extend their results to larger times, typically when the total mutation rate from the subpopulation of a trait v to the subpopulation of a trait u is growing as a positive power of n, instead of remaining finite.

In [6], Durrett and Mayberry study the exponentially growing Moran model. They consider the same mutation regime; their total population size grows exponentially fast at a fixed rate, and new individuals in the population choose their trait via a selective frequency-dependent process. In Theorem 4.2.9, a similar result is obtained for the case of a multitype branching population. In particular, for this setting, the exponential speed of the total population (and of the dominant subpopulations) growth evolves over time. More specifically, we show that the speed is a non-decreasing piecewise constant function going from $\lambda(0)$ to $\max_{v\in V}\lambda(v)$, and taking values only from the set $\{\lambda(v), \forall v \in V\}$, see Theorem 4.2.9.

In [8, 7, 9, 11, 12, 13, 14, 15, 17], the authors consider the power law mutation rates limit of Equation (4.1.1) in the special case where all different traits mutate with the same scaling of a fixed order of a negative power of n. In contrast, in the present work, the power law mutation rates are more general by allowing traits to mutate with different scalings, as in [10, 16].

As in [8, 10], compared to the different models in [6, 7, 11, 12, 13, 14, 16, 17], the initial population $\mathcal{Z}^{(n)}(0)$ is not assumed to have a macroscopic size. This introduces an additional randomness in how the wild-type subpopulation stochastically grows to reach a macroscopic size. However, contrary to [8, 10], we condition neither on the survival of the wild-type subpopulation nor on the finiteness of the stopping times of Equation (4.1.2).

In [23], Nicholson and Antal study a similar model under a slightly less general non-increasing growth rate condition. More precisely, in their case, all the growth rates of the mutant populations are strictly smaller than the growth rate of the wild-type population: $\forall v \in V\backslash\{0\}, \lambda(v) < \lambda(0)$. However, the main difference remains the mutation regime. In their case, only the last mutation is in

the power law mutation rates regime, while all other mutations have a fixed probability independent of n. In Theorem 4.2.7 the case where all mutations are in the power law mutation rates regime is analysed. Additionally, Nicholson and Antal were interested in obtaining the distribution of the first time that a mutant subpopulation gets a mutant cell, whereas in the present work, the first-order asymptotics of the sizes of the mutant subpopulations are studied over time.

In [24], Nicholson, Cheek and Antal study the case of a mono-directional graph where time tends to infinity with fixed mutation probabilities. In particular, they obtain the almost sure first-order asymptotics of the mutant subpopulation sizes. Under a non-increasing growth rate condition, they are able to characterise the distribution of the random variables they obtain in the limit. Without any condition on the growth rates, they study the distribution of the random limit under the small mutation probabilities limit, using the hypothesis of an approximating model with less stochasticity. Note that the mutation regime they study is not the power law mutation rates limit of Equation (4.1.1) as considered in the present work. Under the latter regime, both the size of the population goes to infinity and the mutation probabilities to 0, through the parameter n.

In [62], Gunnarson, Leder and Zhang study a similar model to the one in the present work and are also interested in capturing the time-evolution of the genetic diversity of a cell population, using in their case the well-known summary statistic called the *site frequency spectrum* (SFS for short). The main difference lies in the considered mutation regime which is not the power law mutation rates limit. In their case, the mutation probabilities are fixed. Additionally, they restrict the study to the neutral cancer evolution case. In particular, as in the present work, they capture the first-order asymptotics of the SFS at a fixed time and at the random time at which the population first reaches a certain size. Two noticeable similarities in the results are that the first-order asymptotics of the SFS converge to a random limit when evaluated at a fixed time and to a deterministic limit when evaluated at the previous stochastic time. One could argue that in the present work the correct convergence in the latter case is actually a stochastic limit. But the randomness is fully given by the survival of the wild-type lineage of the cancer-initiating cell, so conditioned on such an event, in the end, the limit is a deterministic one. In particular the results of Gunnarson, Leder and Zhang are all conditioned on the non extinction of the population.

In [16], Gamblin, Gandon, Blanquart and Lambert study a model of an exponentially growing asexual population that undergoes cyclic bottlenecks under the power law mutation rates limit. Their trait space is composed of 4 subpopulations $00, 10, 01$ and 11, where two pathways of mutations are possible: $00 \mapsto 10 \mapsto 11$ and $00 \mapsto 01 \mapsto 11$. They study the special case where one mutation (10) has a high rate but is a weakly beneficial mutation whereas the other mutation (01) has a low rate but is a strongly beneficial mutation. In particular they show the notable result that due to cyclic bottlenecks only a unique evolutionary pathway unfolds, but modifying their intensity and period implies that all pathways can be explored. Their work relies on a deterministic approximation of the wild-type subpopulation 00 and some parts of the analysis of the model's behaviour are obtained only through heuristics. The present work, and more specifically Theorem 4.2.9, because it considers selective mutations, can be used and adapted to consider the case of cyclic bottlenecks in order to prove rigorously their results, both in the specific trait space that they consider and in a general finite directed trait space.

The rest of the paper is organised as follows. In Section 4.2, we give the results and their biological interpretations. Sections 4.3 and 4.4 are dedicated to proving Theorem 4.2.7, which assumes Equation (4.1.5). In Section 4.3, we provide the mathematical construction of the model for an infinite mono-directional graph using Poisson point measures, as well as the proof in this particular case. The generalisation of the proof from an infinite mono-directional graph to a general

finite directed graph is given in Section 4.4.

4.2 Main results and biological interpretation

In Subsection 4.2.1 the first-order asymptotics of the size of each mutant subpopulation on the time scales (4.1.3) and (4.1.4) are provided under the non-increasing growth rate condition (4.1.5). In Subsection 4.2.2, the asymptotic result on the stochastic exponent of each mutant subpopulation is presented without any assumption on the growth rate function λ. In each subsection, biological interpretations of the results are provided.

4.2.1 First-order asymptotics of the mutant subpopulation sizes under non-increasing growth rate condition

In this subsection, we assume that the graph (V, E, L) satisfies the non-increasing growth rate condition given by Equation (4.1.5).

Heuristics for a general finite graph

The next definitions, notations and results are initially motivated by heuristics for the simplest possible graph: a wild-type and a mutant population where only mutations from wild-type to mutant cells are considered. Specifically, we consider the graph $(V, E, L) = (\{0,1\}, \{(0,1)\}, \{\ell(0,1)\})$, as illustrated in Figure 4.1. Under the power law mutation rates regime, the intrinsic birth and

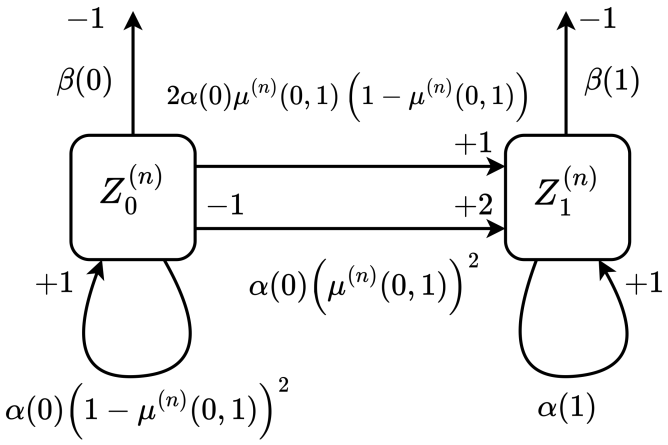

Figure 4.1: Graphical representation of the model with two traits and without backward mutation

death rates of the wild-type subpopulation, $\alpha(0)\left(1-\mu^{(n)}(0,1)\right)^2$ and $\beta(0)+\alpha(0)\left(\mu^{(n)}(0,1)\right)^2$, respectively, are so close to $\alpha(0)$ and $\beta(0)$ that its natural martingale asymptotically behaves like that of a birth and death process with rates $\alpha(0)$ and $\beta(0)$ (see Lemma 4.3.1). This allows us to approximate the growth of the wild-type subpopulation as an exponential growth with parameter $\lambda(0)$. Then, if it survives, at time $\mathfrak{t}_t^{(n)}$ (see (4.1.4)), its size is of order $\Theta\left(n^t\right)$ (see Lemma 4.3.2),

where we use the standard Landau notation for Θ. Given this, we understand why it is necessary to wait until time $\mathfrak{t}^{(n)}_{\ell(0,1)}$ before observing any mutations. Indeed, with a mutation probability scaling as $n^{-\ell(0,1)}$, the total mutation probability up to time $\mathfrak{t}^{(n)}_t$ scales as

$$\int_0^t n^u n^{-\ell(0,1)} d\big(u\frac{\log(n)}{\lambda(0)}\big) = \frac{n^{-\ell(0,1)}}{\lambda(0)}\left(n^t - 1\right),$$

which starts to be of order 1 for $t \geq \ell(0,1)$. This is formalised by D. Cheek and T. Antal in [8, 10]. An illustration is provided in Figure 4.2. Some heuristics for the size of the mutant subpopulation

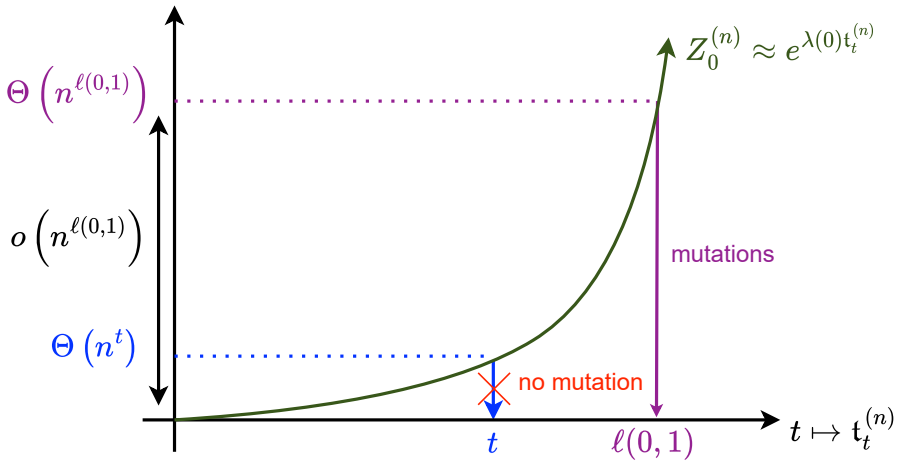

Figure 4.2: Heuristics for the first-occurrence time of mutant cells

at time $\mathfrak{t}^{(n)}_t$, for $t \geq \ell(0,1)$, can also be derived. For $\ell(0,1) \leq u \leq t$, the number of new mutations generated at time $\mathfrak{t}^{(n)}_u$ scales as $\exp\left(\lambda(0)(u - \ell(0,1))\frac{\log(n)}{\lambda(0)}\right)$. The remaining time for these new mutant cells to grow exponentially at rate $\lambda(1)$ until time $\mathfrak{t}^{(n)}_t$ is $\mathfrak{t}^{(n)}_{t-u}$. This implies that their lineages reach a size at time $\mathfrak{t}^{(n)}_t$ of order

$$\Theta\Big(\exp\Big(\left[\lambda(1)t + (\lambda(0) - \lambda(1))u - \lambda(0)\ell(0,1)\right]\frac{\log(n)}{\lambda(0)}\Big)\Big). \tag{4.2.1}$$

Two scenarios are then possible:

- If $\lambda(1) < \lambda(0)$: Equation (4.2.1) is maximised for $u = t$ and equals $n^{t-\ell(0,1)}$. This means that the dynamics of the mutant subpopulation is driven by mutations from the wild-type subpopulation rather than by its intrinsic growth. More precisely, its size order at time $\mathfrak{t}^{(n)}_t$ is determined entirely by the mutations generated at that time -and so is of order $n^{t-\ell(0,1)}$- and not by the lineages arising from mutations at earlier times. Biologically, these mutations are termed *deleterious*.
- If $\lambda(1) = \lambda(0)$: Equation (4.2.1) is independent of u and equals $\Theta\left(n^{t-\ell(0,1)}\right)$ for any $\ell(0,1) \leq u \leq t$. This indicates that lineages of mutant cells generated from mutations at any time

between $\mathrm{t}^{(n)}_{\ell(0,1)}$ and $\mathrm{t}^{(n)}_{t}$ have the same order of size at time $\mathrm{t}^{(n)}_{t}$. In other words, there is a balance in the dynamics of the mutant subpopulation between the contributions of mutations and its intrinsic growth. This is a consequence of assuming $\lambda(1) = \lambda(0)$. These mutations are referred to as *neutral mutation*, even though biologically speaking, this would more precisely mean the restrictive condition $\alpha(1) = \alpha(0)$ and $\beta(1) = \beta(0)$. Therefore, to capture the total size of the mutant subpopulation at time $\mathrm{t}^{(n)}_{t}$, one must integrate all lineages resulting from mutational events over the time $\mathrm{t}^{(n)}_{u}$, for $\ell(0,1) \leq u \leq t$. This gives exactly the order $\Theta\left((t-\ell(0,1))\log(n)n^{t-\ell(0,1)}\right)$.

To summarise, for this simple graph, the size of the mutant subpopulation after time $\mathrm{t}^{(n)}_{\ell(0,1)}$ scales as

$$\Theta\Big(n^{t-\ell(0,1)}\left[\mathbb{1}_{\{\lambda(0)>\lambda(1)\}} + \mathbb{1}_{\{\lambda(0)=\lambda(1)\}}(t-\ell(0,1))\log(n)\right]\Big). \tag{4.2.2}$$

Notice, in particular, that in any case, the mutant subpopulation exhibits exponential growth at rate $\lambda(0)$ after time $\mathrm{t}^{(n)}_{\ell(0,1)}$, as indicated by the factor $n^{t-\ell(0,1)}$. An illustration of this heuristic can be found in Figure 4.3, which visually represents the growth dynamics of the mutant subpopulation over time.

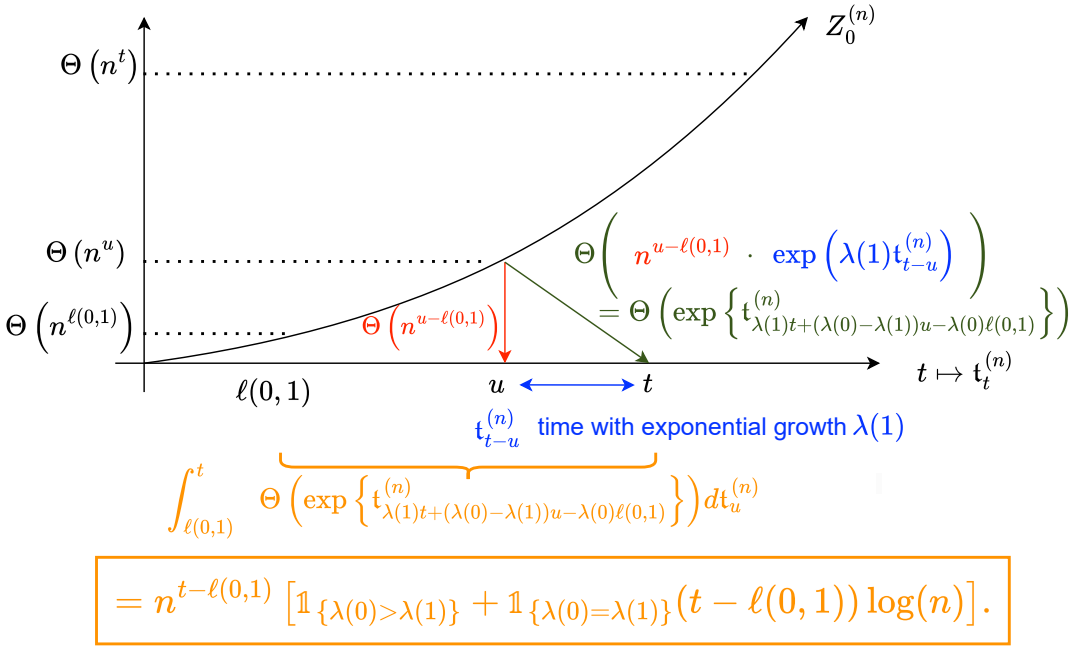

Figure 4.3: Heuristics for the size of the mutant subpopulation after time $\mathrm{t}^{(n)}_{\ell(0,1)}$

These heuristics on this simple graph can be used as an elementary brick for developing heuristics on a general finite graph. Considering a vertex $v \in V\backslash\{0\}$, there may be multiple mutational

pathways from the initial vertex 0 to v. It is important to understand which pathways actually contribute to the size order of the mutant subpopulation of trait v. Using both the previous heuristics on the time required for mutations to occur and the fact that after this time, the mutant subpopulation grows exponentially at rate $\lambda(0)$, along with an additional $\log(n)$ factor if the mutation is neutral, it seems natural to iteratively apply this reasoning to a mutational pathway, encoded via a mono-directional graph. In the following, we will use the term *'walk'* instead of *'pathway'*, favouring the nomenclature of graph theory over the biological terminology. For any given walk from 0 to v, the needed time u, in the time scale $\mathrm{t}_u^{(n)}$, to observe a cell of trait v generated via this specific walk is the sum of the labels of the edges along this walk, which is referred to as the *length* of the walk. After this time, this subpopulation of cells of trait v grows exponentially at rate $\lambda(0)$. Moreover, as observed in (4.2.2), for each neutral mutation along the walk, an additional multiplicative factor of order $\log(n)$ is included in the size order. This leads to three key observations about the total mutant subpopulation of trait v:

- First occurrence of cells: Cells of trait v first appear after a time equal to the minimum of the lengths of all walks from 0 to v.
- Effective evolutionary pathways: After this time, only walks whose lengths equal this minimum might contribute to the size order of the mutant subpopulation of trait v. This is because any time delay creates an exponential delay in the size order. This fact is captured asymptotically in Theorem 4.2.9.
- Neutral mutation factor: The additional multiplicative factor of $\log(n)$ due to neutral mutations implies that, among the walks from 0 to v with lengths equal to the aforementioned minimum, only those with the maximal number of neutral mutations actually contribute to the size order of the mutant subpopulation of trait v. Specifically, these walks contribute with a factor of $\log(n)$ raised to the power given by this maximal number of neutral mutations. This fact is asymptotically captured in Theorem 4.2.7. Additionally, for each of these admissible walks, an additional time integral is obtained at each neutral mutation, as observed in (4.2.2).

An illustration of this reasoning is provided with an example in Figure 4.4.

Notations and definitions:

Now, the natural definitions derived from these heuristics are formally established before presenting the results.

Definition 4.2.1 (Deleterious and neutral vertices)**.** *A vertex $v \in V$ is called a neutral vertex if $\lambda(v) = \lambda(0)$, and a deleterious vertex if $\lambda(v) < \lambda(0)$.*

Remark 4.2.2. *In the previous definition, the terms "neutral" or "deleterious" for a mutation are based on comparing its growth rate to that of the wild-type subpopulation. However, it is possible to have a mutation from a vertex v to a vertex u where $\lambda(v) < \lambda(u) \leq \lambda(0)$. Although such a mutation could theoretically be considered selective, since $\lambda(u) > \lambda(v)$, the previous definition categorises it as either neutral or deleterious, depending on the value of $\lambda(u)$ relative to $\lambda(0)$. This nomenclature emerges from the fact that, under Assumption (4.1.5), any mutant subpopulation grows exponentially at rate $\lambda(0)$, as developed in the earlier heuristics. Thus, this legitimates the previous definition, assuming (4.1.5) holds.*

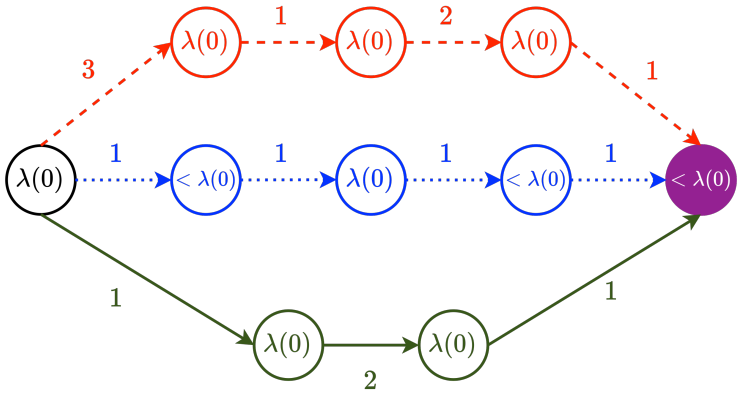

Figure 4.4: Heuristics for the contribution of walks to the size order of the plain purple mutant subpopulation: in this example, the dashed red walk has a length of 7, while the dotted blue and plain green walks have a length of 4. Therefore, only the two latter walks may contribute to the size order of the plain purple mutant subpopulation, making them sub-admissible walks. However, the dotted blue walk has only one neutral mutation, whereas the plain green walk has two neutral mutations. As a result, only the plain green walk will ultimately contribute to the size order of the purple mutant subpopulation. For $t \geq 4$, at time $\mathfrak{t}_t^{(n)}$, it will grow as $\log^2(n)e^{\lambda(0)\mathfrak{t}_{t-4}^{(n)}}$. Notice, in particular, that the dashed red walk has the maximal number of neutral mutations, which is 3. However, since it is not a sub-admissible walk, the multiplicative factor of $\log(n)$ remains 2 instead of 3.

The following definition provides a structured framework to analyse the contribution of evolutionary pathways to the growth of mutant subpopulations. It does so by introducing the adapted vocabulary, for the neutral and deleterious evolutionary context of the model, associated with walks in labelled graphs. We use the term *'walk'* here according to the standard nomenclature of graph theory.

Definition 4.2.3 (Walk in the graph)**.** *A walk $\gamma = (v(0), \cdots, v(k))$ in the graph (V, E) is defined as a sequence of vertices linking $v(0)$ to $v(k)$ such that for all $0 \leq i \leq k, v(i) \in V$, and for all $0 \leq i \leq k-1, (v(i), v(i+1)) \in E$. We will sometimes use the term 'path' to refer to a walk that visits only distinct vertices. Given a walk $\gamma = (v(0), v(1), \cdots v(k))$ in the labelled graph (V, E, L), we define:*

- *The sum of the labels of the edges and the sum over the first i edges of the walk γ, respectively:*

$$t(\gamma) := \sum_{i=0}^{k-1} \ell(v(i), v(i+1)) \text{ and for all } i \leq k, t_\gamma(i) := \sum_{j=0}^{i-1} \ell(v(j), v(j+1)).$$

- *The subset of neutral heads of the edges of the walk γ and its cardinality:*

$$\gamma_{neut} = \{v(i), 1 \leq i \leq k : \lambda(v(i)) = \lambda(0)\} \text{ and } \theta(\gamma) := |\gamma_{neut}|.$$

- *The weights $w_{neut}(\gamma)$ and $w_{del}(\gamma)$ associated with the neutral and deleterious vertices of the walk γ, respectively:*

$$w_{neut}(\gamma) := \prod_{1\leq i\leq k, \lambda(v(i))=\lambda(0)} \frac{2\alpha(v(i-1))\mu(v(i-1),v(i))}{\lambda(0)},$$
$$w_{del}(\gamma) := \prod_{1\leq i\leq k, \lambda(v(i))<\lambda(0)} \frac{2\alpha(v(i-1))\mu(v(i-1),v(i))}{\lambda(0)-\lambda(v(i))}.$$

  Along a walk, the constant of the asymptotic contribution of a vertex- depending on its parameters and those of the upstream vertex- takes a distinct form based on whether the vertex is neutral or deleterious. This distinction motivates the use of the separate weights $w_{neut}(\gamma)$ and $w_{del}(\gamma)$.

- *The time dependence associated with the neutral vertices: Let σ be an increasing function from $\{1,\cdots,\theta(\gamma)\}$ to $\{1,\cdots,k\}$, such that $v(\sigma_i)$ is the i-th neutral vertex of the walk γ. For all $t>0$, define the multiple integral $I_\gamma(t)$ as*

$$I_\gamma(t) := \int_{t_\gamma(\sigma_{\theta(\gamma)})}^{t\vee t_\gamma(\sigma_{\theta(\gamma)})} \int_{t_\gamma(\sigma_{\theta(\gamma)-1})}^{u_1} \cdots \int_{t_\gamma(\sigma_{\theta(\gamma)-k})}^{u_k} \cdots \int_{t_\gamma(\sigma_1)}^{u_{\theta(\gamma)-1}} du_{\theta(\gamma)}\cdots du_1.$$

  Along a walk, for each neutral vertex that is visited, an additional integral over the time parameter appears in the asymptotic limit, as described in the heuristics. This motivates the definition of $I_\gamma(t)$.

- *The weight of the walk γ at time t:*

$$w_\gamma(t) := w_{del}(\gamma)w_{neut}(\gamma)I_\gamma(t). \tag{4.2.3}$$

  This expression captures the total weight of a walk γ at time t, accounting for both the deleterious and neutral visited vertices, and the integrals over the time parameters associated with these neutral vertices.

The next remark provides a recursive formula for computing the weight of a walk γ at a given time t.

Remark 4.2.4. *The weight $w_\gamma(t)$ of the walk $\gamma=(v(0),\cdots,v(k))$ at time t can be recursively expressed in terms of the weight $w_{\overleftarrow{\gamma}}(t)$ associated with the walk $\overleftarrow{\gamma}:=(v(0),\cdots,v(k-1))$, which is the same walk as γ up to the second-to-last vertex (i.e. without the final vertex $v(k)$). The recursive equation, which considers whether the last vertex $v(k)$ is deleterious or neutral, is given by*

$$\begin{aligned} w_\gamma(t) =& 2\alpha(v(k-1))\mu(v(k-1),v(k)) \\ &\cdot\Big(\mathbb{1}_{\{\lambda(k)<\lambda(0)\}}\frac{1}{\lambda(0)-\lambda(v(k))}w_{\overleftarrow{\gamma}}(t) + \mathbb{1}_{\{\lambda(k)=\lambda(0)\}}\frac{1}{\lambda(0)}\int_{t(\overleftarrow{\gamma})}^{t\vee t(\overleftarrow{\gamma})} w_{\overleftarrow{\gamma}}(s)ds\Big). \end{aligned}$$

Definition 4.2.5 (Admissible walks)**.** *For all $v\in V$, let $P(v)$ denote the set of all walks γ in the graph (V,E) that link the vertex 0 to the vertex v. We define the:*

- *The minimum total label sum among all walks from vertex* 0 *to vertex* v*:*

$$t(v) := \min_{\gamma \in P(v)} t(\gamma).$$

- *The maximum number of neutral vertices among the shortest walks from vertex* 0 *to vertex* v*:*

$$\theta(v) := \max_{\gamma \in P(v), t(\gamma) = t(v)} \theta(\gamma).$$

- *The set of admissible walks from vertex* 0 *to vertex* v*:*

$$A(v) := \{\gamma \in P(v) : t(\gamma) = t(v) \text{ and } \theta(\gamma) = \theta(v)\}.$$

Remark 4.2.6. *In the previous definition, the set* $A(v)$ *is referred to as the set of admissible walks because, as indicated by the heuristics, only walks belonging to* $A(v)$ *contribute to the growth dynamics of the mutant subpopulation of trait* v*. This is formally established in Theorem 4.2.7.*

First-order asymptotic results

Under Assumption (4.1.5), the more refined result can now be formally stated. The model is mathematically constructed in Section 4.4 (see (4.4.1), (4.4.2), (4.4.3), (4.4.4) and (4.4.5)) using independent Poisson Point Measures. The following theorem provides the asymptotic results for this specific mathematical construction of the model. The convergences are, in particular, obtained in probability. For any mathematical construction of the model other than the one given in Section 4.4, the convergences hold at least in distribution in the appropriate Skorokhod space, see Remark 4.2.8. A motivation for the normalising term $d_v^{(n)}(t,s)$, introduced in the the following theorem, is provided below in Remark 4.2.8.

Theorem 4.2.7. *Assume that the general finite directed labelled graph* (V, E, L) *satisfies both the power law mutation rates regime described in* (4.1.1) *and the non-increasing growth rate graph condition given in* (4.1.5)*. Let* $h_n = \frac{\log(n)}{\log(\log(n))\theta_{\max}+\varphi_n}$*, where* $\varphi_n \underset{n\to\infty}{\to} \infty$ *such that* $h_n \underset{n\to\infty}{\to} \infty$ *and where* $\theta_{max} := \max_{v \in V\setminus\{0\}} \theta(v)$*. Let also* ψ_n *such that* $\sqrt{\log(n)} = o(\psi_n)$*. Define for all* $(t,s) \in \mathbb{R}^+ \times \mathbb{R}$*,*

$$\begin{aligned} d_v^{(n)}(t,s) :=& \mathbb{1}_{\left\{t \in [0, t(v) - h_n^{-1})\right\}} + \mathbb{1}_{\left\{t \in [t(v) - h_n^{-1}, t(v))\right\}} \psi_n \log^{\theta(v)-1}(n) \\ &+ \mathbb{1}_{\{t \in [t(v), \infty)\}} n^{t - t(v)} \log^{\theta(v)}(n) e^{\lambda(0)s}. \end{aligned} \tag{4.2.4}$$

Let $(T, M) \in \left(\mathbb{R}_+^*\right)^2$ *and* $0 < T_1 < T_2$*. Using the mathematical definition of the model given in Section 4.4 (see* (4.4.1)*,* (4.4.2)*,* (4.4.3)*,* (4.4.4) *and* (4.4.5))*, there exists a random variable* W*, properly defined in* (4.4.6)*, satisfying*

$$W \overset{law}{:=} Ber\Big(\frac{\lambda(0)}{\alpha(0)}\Big) \otimes Exp\Big(\frac{\lambda(0)}{\alpha(0)}\Big),$$

such that for all $v \in V\setminus\{0\}$*, we obtain the convergence results in probability in* $L^\infty([0,T]\times[-M,M])$ *for Equation* (4.2.5) *and in* $L^\infty([T_1,T_2]\times[-M,M])$ *for Equations* (4.2.6)*,* (4.2.7) *and* (4.2.8)*:*

- ***Deterministic time scale*** (4.1.4)***:***
  If $\lambda(v) = \lambda(0)$*, then*

$$\frac{Z_v^{(n)}\left(\mathfrak{t}_t^{(n)} + s\right)}{d_v^{(n)}(t,s)} \underset{n\to\infty}{\longrightarrow} W \sum_{\gamma \in A(v)} w_\gamma(t). \tag{4.2.5}$$

  If $\lambda(v) < \lambda(0)$*, then*

$$\frac{Z_v^{(n)}\left(\mathfrak{t}_{t(v)+t}^{(n)} + s\right)}{n^t \log^{\theta(v)}(n) e^{\lambda(0)s}} \underset{n\to\infty}{\longrightarrow} W \sum_{\gamma \in A(v)} w_\gamma(t(v)+t). \tag{4.2.6}$$

- ***Random time scale*** (4.1.3)***:*** *Consider* $\left(\rho_t^{(n)}\right)_{t\in\mathbb{R}^+}$ *as defined in* (4.1.2).
  If $\lambda(v) = \lambda(0)$*, then*

$$\frac{Z_v^{(n)}\left(\rho_t^{(n)} + s\right)}{d_v^{(n)}(t,s)} \underset{n\to\infty}{\longrightarrow} \mathbb{1}_{\{W>0\}} \sum_{\gamma \in A(v)} w_\gamma(t). \tag{4.2.7}$$

  If $\lambda(v) < \lambda(0)$*, then*

$$\frac{Z_v^{(n)}\left(\rho_{t(v)+t}^{(n)} + s\right)}{n^t \log^{\theta(v)}(n) e^{\lambda(0)s}} \underset{n\to\infty}{\longrightarrow} \mathbb{1}_{\{W>0\}} \sum_{\gamma \in A(v)} w_\gamma(t(v)+t). \tag{4.2.8}$$

The proof of Theorem 4.2.7 relies on a martingale approach using Doob's and Maximal Inequalities. The initial step involves controlling the growth of the lineage of wild-type cells originated from the initial cell, for both the deterministic and random time scales (4.1.4) and (4.1.3) (see Lemma 4.4.3 and 4.4.4). For any vertex $v \in V\backslash\{0\}$, there may be several mutational walks in the graph (V, E) that start from 0 and lead to v. Understanding the contribution of each of these walks to the first-order asymptotics of the size of the mutant subpopulation of trait v is essential. The proof proceeds in 2 steps:

(i) Consider an infinite mono-directional graph under Assumption (4.1.5) and establish the result for this specific graph, see Section 4.3. Performing this step for an infinite graph is particularly helpful in handling cycles (such as backward mutations) in a general finite directed graph.

(ii) Identify and exclude walks from the initial vertex 0 to v that do not contribute to the first-order asymptotics of the size of the mutant subpopulation of trait v, see Section 4.4.

Remark 4.2.8. *1.* ***Mathematical construction:*** *For any mathematical construction other than the one given in Section 4.4, the convergences hold at least in distribution in* $\mathbb{D}([0,T]\times[-M,M])$ *for Equation* (4.2.5) *and in* $\mathbb{D}([T_1,T_2]\times[-M,M])$ *for Equations* (4.2.6), (4.2.7) *and* (4.2.8).

2. ***An additional*** $\log(n)$ ***factor:*** *Notice that a multiplicative factor of* $\log^{\theta(v)}(n)$ *is captured after time* $\mathfrak{t}_{t(v)}^{(n)}$*, see Equations* (4.2.4), (4.2.5), (4.2.6), (4.2.7) *and* (4.2.8). *Obtaining a result*

on the stochastic exponents (see (4.1.6)) does not capture such a factor. For instance, with the model of Figure 4.1, if $\lambda(1) = \lambda(0)$, Theorem 4.2.7 gives that after time $\ell(0,1)$, $Z_1^{(n)}(\mathfrak{t}_t^{(n)})$ behaves approximately as $\log(n)e^{\lambda(0)\mathfrak{t}^{(n)}_{t-\ell(0,1)}}$. However, what is captured with $X_1^{(n)}(t)$ after time $\ell(0,1)$ is asymptotically $\lambda(0)(t-\ell(0,1))$, see Theorem 4.2.9.

3. ***Stochasticity of the limits:*** *The random variable W is explicitly defined as the almost sure limit of the natural positive martingale associated to a specific birth and death branching process with rates $\alpha(0)$ and $\beta(0)$; see (4.4.6). The martingale associated to the lineage of wild-type cells issued from the initial cell behaves similarly to the one associated to the aforementioned birth and death branching process (see Lemma 4.4.3). Thus, W quantifies the randomness of this lineage over the long time. Due to the power law mutation rates regime, mutations arise after a long time, so the stochasticity of this lineage is already captured by W. Notice that under Assumption (4.1.5), the randomness in the first-order asymptotics of any mutant subpopulation size is described completely by W. This means that the stochasticity of these subpopulations is driven primarily by the randomness in the growth of the wild-type subpopulation rather than by the one of the mutational process or of any lineage of mutant cells. In particular, if the process starts with a large number of wild-type cells instead of just one, the first-order asymptotics of the size of the mutant subpopulations would be entirely deterministic.*

4. ***Selective cancer evolution:*** *It seems quite natural not to obtain such a result when considering selective mutation ($\lambda(v) > \lambda(0)$). Indeed, a selective mutation imply that any time advantage translates directly into a growth advantage. Thus, the stochasticity of the mutational process, as well as the randomness in the lineages of the mutant cells, cannot be ignored. Therefore, expecting to control the stochasticity of the mutant subpopulation solely by controlling the randomness in the wild-type subpopulation, without also accounting for the randomness in the mutational process and the mutant lineages, is vain. More precisely, using a martingale approach to derive the first-order asymptotics cannot be successful for a selective mutation. Technically, this is because the expected size of the selective mutant subpopulation is of a higher order than its typical asymptotic size. Indeed, the rare event of the initial cell mutating to the selective trait extremely quickly, an event that asymptotically vanishes, is responsible for this discrepancy between the expected value and the typical asymptotic size of the selective mutant subpopulation. Nevertheless, when examining the stochastic exponent (4.1.6), the martingale approach allows us to obtain convergence results as given in Theorem 4.2.9. This is because the aforementioned rare event contributes only a factor proportional to its probability to the expected value of the stochastic exponent, meaning it actually asymptotically neither contributes to the typical size nor to the expected value of the stochastic exponent of the selective mutant subpopulation. Generalisation to derive the first-order asymptotics when considering selective mutations is a work in progress.*

5. ***Definition of neutral mutation:*** *In view of Theorem 4.2.7, the mathematical definition of neutral mutation, $\lambda(v) = \lambda(0)$, is well-understood, as opposed to the more restrictive but biologically meaningful condition of having both $\alpha(v) = \alpha(0)$ and $\beta(v) = \beta(0)$. Indeed, maintaining the same growth rate $\lambda(v) = \lambda(0)$, while changing the birth and death rates $\alpha(v)$ and $\beta(v)$ alters the distribution of any lineage of mutant cells. Consequently, one might naturally expect that this would alter the stochasticity of the mutant subpopulation size. However, this is not the case. The randomness in the first-order asymptotics is fully summed up by the random*

variable W. Thus, it is entirely consistent that, under the neutral assumption, the condition pertains only to the growth rate function rather than to the birth and death rate functions.

6. ***Motivation of*** $d_v^{(n)}(t,s)$***:*** *Considering the time scale $\mathfrak{t}_t^{(n)}$, the result slightly differs depending on whether the vertex is neutral or deleterious. Indeed, when looking at the asymptotic behaviour for a deleterious vertex v, the result holds strictly after time $t(v)$, whereas, in the case of a neutral vertex, the entire trajectory from the initial time can be analysed. Mathematically, this difference arises from the additional multiplicative factor of $\log(n)$ in the first-order asymptotics when considering a neutral mutation. This factor allows us to control the quadratic variation at time $t(v)$ for the martingale associated to the mutant subpopulation. Three distinct regimes are obtained, as indicated by (4.2.4) and (4.2.5) :*

   (i) Up to time $t(v) - h_n^{-1}$: with high probability, no mutational pathway from 0 to v has generated a mutant cell of trait v. Since $h_n \to \infty$ and satisfies $h_n = o(\log(n))$, $t(v)$ can be interpreted as the first time -when considering the time scale accelerated by $\log(n)$- at which it becomes asymptotically possible to observe the first occurrence of a mutant cell of trait v. This result is also true for deleterious mutations, see Lemma 4.3.6.

   (ii) For $t \in \big[t(v) - h_n^{-1}, t(v)\big)$: in this time interval, some mutant cells of trait v are produced, but the interval's length is insufficient to achieve any power of n for the size of the mutant subpopulation of trait v. We succeed to dominate its growth by $\psi_n \log^{\theta(v)-1}(n)$, with a well-chosen ψ_n. Heuristically, the total number of mutant cells of trait v resulting from a mutational event up to time t is of order $\Theta\big(\log^{\theta(v)-1}(n)\big)$. With the remaining time for these mutant cells' lineages to grow, we manage to control the size of the mutant subpopulation of trait v by at most $\sqrt{\log(n)}\log^{\theta(v)-1}(n)$. Consequently, dividing by any function ψ_n satisfying $\sqrt{\log(n)} = o(\psi_n)$ results in an asymptotic limits of 0. This result also holds for deleterious mutations, see Lemma 4.3.7. The $\sqrt{\log(n)}$ factor in the growth control comes from a mathematical analysis using a martingale approach, particularly considering the time scale accelerated by $\log(n)$. With further refinement, we conjecture that the actual size of the mutant subpopulation at time $t(v)$ is of order $\Theta\big(\big(\mathbb{1}_{\{\lambda(0)=\lambda(v)\}}\log(\log(n)) + \mathbb{1}_{\{\lambda(0)>\lambda(v)\}}\big)\log^{\theta(v)-1}(n)\big)$.

   (iii) For $t \in [t(v), \infty)$: with high probability, the number of mutant cells of trait v grows exponentially at rate $\lambda(0)$. A supplementary multiplicative factor $\log^{\theta(v)}(n)$ is present due to the neutral mutations on the walks in $A(v)$. Thus, the growth scales globally as $n^{(t-t(v))}\log^{\theta(v)}(n)w_v(t)$.

7. ***Differences between the time scales:*** *When comparing point (i) and (ii) of Theorem 4.2.7, notice that the result transitions from the deterministic time scale $\mathfrak{t}_t^{(n)}$ to the random time scale $\rho_t^{(n)}$ merely by switching W to $\mathbb{1}_{\{W>0\}}$. This seemingly surprising fact can be explained by the essential role of W. As mentioned in Remark 4.2.8 3., W encodes the long-term stochasticity of the lineage of wild-type cells originating from the initial cell. By showing that the time scale $\mathfrak{t}_t^{(n)}$ serves as the correct deterministic approximation of $\rho_t^{(n)}$ (see Lemma 4.4.4), it follows that obtaining an asymptotic result on time scale $\mathfrak{t}_t^{(n)}$ also yields a result for the time scale $\rho_t^{(n)}$. This idea is formalised using a technique similar to that in [56, Lemma 3]. The switch from W to $\mathbb{1}_{\{W>0\}}$ in the result occurs because the time scale $\rho_t^{(n)}$ inherently carries the stochasticity of the random variable W. Consequently, the only remaining randomness that needs to be*

considered is the survival of the lineage from the initial cell, which is asymptotically given by $\mathbb{1}_{\{W>0\}}$.

4.2.2 Result for a general finite directed labelled graph

This subsection does not require the non-increasing growth rate condition of Equation (4.1.5). Without this assumption, a martingale approach fails to obtain the first-order asymptotics of the mutant subpopulation sizes. However, the stochastic exponents of the mutant subpopulations, as defined in (4.1.6), can be uniformly tracked over time. In particular, we show that, under the event $\{W > 0\}$, the limits are positive deterministic non-decreasing piecewise linear continuous functions. Such limits are defined via a recursive algorithm tracking their slopes over time. More precisely, we show that the slopes can only increase and take values from the growth rate function.

In the tracking algorithm, two different kinds of updates can be made:

- **Birth of a new trait:** The first update is the birth of a new trait which takes as its slope the maximum between its inner growth rate and the slope of the subpopulation that gave birth to it. In fact, it could also happen that many subpopulations give birth to it at the same time; in this case it is the maximum of their slopes that is compared to the inner growth rate of the born trait. Such a comparison on the growth rates indicates which mechanism is driving the subpopulation growth: either its inner growth if this subpopulation is selective compared to the subpopulation(s) that is/are giving birth to it, or conversely the mutational process if it is deleterious. The neutral case corresponds to a balance of these two mechanisms, as previously mentioned in Theorem 4.2.7.

- **Growth driven by another trait:** The second kind of update is when a live trait v increases its slope because another live trait u among its incoming neighbours, with a higher slope, has reached its typical size so that the mutational contribution from trait u now drives the growth of trait v. Consequently trait v now takes the slope of trait u. Again potentially many traits u among the incoming neighbours of trait v can reach at the same time the typical size for the mutational contribution to drive the growth of trait v; in this case the growth of trait v is driven by the trait u with the maximal slope. This kind of update encodes the possibility in the evolutionary process that the driving mechanism of a subpopulation can change over time, always triggering an increase in the actual growth of the subpopulation.

How these two different kinds of updates happen in the tracking algorithm is made formal in the following theorem. Moreover, they can happen at the same time for different vertices. The complexity of such an algorithm comes mostly from the generality both on the growth rate function and on the trait structure. Under the non-increasing growth rate condition (4.1.5), the limiting functions $(x_v)_{x\in V}$ have an explicit form, see Corollary 4.2.12; this is also true when the graph structure is mono-directional, see Corollary 4.2.10.

Theorem 4.2.9. *Let* $0 < T_1 < T_2$. *The stochastic exponents defined in* (4.1.6) *satisfy*

$$\Big(\big(X_v^{(n)}(t)\big)_{v\in V}\Big)_{t\in[T_1,T_2]} \underset{n\to\infty}{\longrightarrow} \mathbb{1}_{\{W>0\}}\Big(\big(x_v(t)\big)_{v\in V}\Big)_{t\in[T_1,T_2]}$$

in probability in $L^\infty[T_1, T_2]$. *For each* $v \in V$, x_v *is a positive deterministic non-decreasing piecewise linear continuous function obtained via a recursive approach tracking its slope over time. In particular there exist* $k^* \in \mathbb{N}$ *and* $0 = \Delta_0 < \Delta_1 < \cdots < \Delta_{k^*} < \infty$ *such that the slopes of* $(x_v)_{v\in V}$ *change*

only at the times $(\Delta_j)_{j\in\{0,\cdots,k^*\}}$. *For* $j \in \{0,\cdots,k^*\}$, *at time* Δ_j *two kinds of updates in the slopes can occur: (i) either a new trait starts to grow or (ii) an already growing trait increases its slope due to a growth driven now by another more selective trait. The algorithm tracks the following quantities for all* $j \in \{0,\cdots,k^*\}$ *at time* Δ_j*:*

- *the set of alive traits,* A_j,
- *the set of not-yet-born traits,* U_j,
- *the slope of* x_v, $\lambda_j(v)$,
- *and the set of traits whose growth is driven by trait* v, $C_j(v)$.

Initialisation: *Set* $A_0 = \{0\}$, $U_0 = V\setminus\{0\}$ *and for all* $v \in V$

$$x_v(0) = 0, \lambda_0(v) = \lambda(0)\mathbb{1}_{\{v=0\}}, \text{ and } C_0(v) = \emptyset.$$

Induction: *Let* $j \in \{0,\cdots,k^*-1\}$. *Assume that there exist times* $0 = \Delta_0 < \Delta_1 < \cdots < \Delta_j < \infty$ *such that* $(x_v)_{v\in V}$ *are positive deterministic non-decreasing piecewise linear continuous functions defined on* $[0,\Delta_j]$, *where changes of slopes occur only on the discrete set* $\{\Delta_1,\cdots,\Delta_j\}$. *Also assume that there exist* $\lambda_j(v)$, A_j, U_j, *and* $C_j(v)$, *respectively the slope of* x_v, *the set of alive vertices and not-yet-born vertices, and the set of vertices whose growth is driven by* v, *everything at time* Δ_j.

Then there exists $\Delta_{j+1} \in (\Delta_j,\infty)$ *such that* $(x_v)_{v\in V}$ *are constructed during the time period* $[\Delta_j,\Delta_{j+1}]$ *according to the following. For all* $v \in V$ *and for all* $t \geq \Delta_j$ *let*

$$y_v(t) = (t-\Delta_j)\lambda_j(v) + x_v(\Delta_j).$$

For all $v \in U_j$ *define*

$$\begin{aligned}
&\forall u \in A_j \text{ such that } (u,v) \in E, \delta_{u,v} := \inf\{t \geq \Delta_j : y_u(t) \geq \lambda(0)\ell(u,v)\},\\
&\delta_v := \inf_{u\in A_j:(u,v)\in E} \delta_{u,v},\\
&\nu(v) := \{u \in A_j : (u,v) \in E \text{ and } \delta_{u,v} = \delta_v\}.
\end{aligned}$$

For all $v \in A_j$ *define*

$$\begin{aligned}
&B_j(v) := \{u \in A_j : (v,u) \in E \text{ and } \lambda_j(v) > \lambda_j(u)\},\\
&\forall u \in B_j(v), \delta_{v,u} := \inf\{t \geq \Delta_j : y_v(t) \geq y_u(t) + \lambda(0)\ell(v,u)\},\\
&\delta_v := \inf_{u\in B_j(v)} \delta_{v,u},\\
&\nu(v) := \{u \in B_j(v) : \delta_{v,u} = \delta_v\}.
\end{aligned}$$

Then define $\Delta_{j+1} := \inf_{v\in V}\delta_v$ *and* $\nu_{j+1} := \{v \in V : \delta_v = \Delta_{j+1}\}$. *Then proceed to the following updates:*

- *Let* $A_{j+1} := A_j \cup (\nu_{j+1}\cap U_j)$ *and* $U_{j+1} = U_j\setminus(\nu_{j+1}\cap U_j)$. *Also let* $\forall v \in U_{j+1}$, $\lambda_{j+1}(v) = \lambda_j(v) = 0$, $C_{j+1}(v) = C_j(v) = \emptyset$.

- *For all* $v \in \nu_{j+1} \cap A_j$, *introduce the set* $\nu^{(-)}(v) := \{u \in \nu(v) : \exists w \in \nu_{j+1} \cap A_j, \lambda_j(w) > \lambda_j(v), \text{ and } u \in \nu(w)\}$.

  Then let $C_{j+1}(v) := C_j(v) \cup \bigcup_{u \in \nu(v) \setminus \nu^{(-)}(v)} (\{u\} \cup C_j(u))$. *For all* $u \in \nu(v) \setminus \nu^{(-)}(v)$ *and* $w \in C_j(u)$, $\lambda_{j+1}(u) = \lambda_{j+1}(w) = \lambda_j(v)$.

- *For all* $v \in A_j$ *whose slope has not been updated yet, let* $\lambda_{j+1}(v) = \lambda_j(v)$. *And for all* $v \in A_j$ *whose set* $C_j(v)$ *has not been updated yet, let* $C_{j+1}(v) := C_j(v)$.

- *For all* $v \in \nu_{j+1} \cap U_j$, *let* $\lambda_{j+1}(v) := \max\left(\lambda(v), \max_{u \in \nu(v)} \lambda_{j+1}(u)\right)$, *and* $C_{j+1}(v) = C_j(v) = \emptyset$. *If* $\lambda_{j+1}(v) \geq \lambda(v)$, *introduce the following set* $\nu^+(v) := \{u \in \nu(v) : \lambda_{j+1}(u) = \max_{w \in \nu(v)} \lambda_{j+1}(w)\}$, *and for all* $u \in \nu^+(v)$, $C_{j+1}(u) := C_{j+1}(u) \cup \{v\}$.

For any mathematical construction other than the one given in Section 4.4 (see (4.4.1), (4.4.2), (4.4.3), (4.4.4) and (4.4.5)), the convergences are at least in distribution in $\mathbb{D}([T_1, T_2])$.

The proof of Theorem 4.2.9 is a minor adaptation of the proofs found in [6]. Specifically, by adapting the arguments from [6, Propositions 2 and 4] to the context of the present model, Theorem 4.2.9 follows. For this reason, and in the interest of brevity, we do not provide an explicit proof of Theorem 4.2.9.

The only notable difference is that the process does not start from a macroscopic state. However, it can be easily shown that, conditioned on $\{W = 0\}$, no mutant cells are generated asymptotically, since with high probability the wild-type subpopulation can't survive in the $\log(n)$-accelerated time scale. Additionally, conditioned on $\{W > 0\}$, the first phase, corresponding to the growth of the wild-type subpopulation leading to the macroscopic state that allows for the generation of the first mutant cell, is straightforward to capture.

When considering a(n) (infinite) mono-directional graph, the structure of such a graph is sufficiently simple to allow for an explicit form of the limiting functions $(x_v)_{v \in V}$, see the next corollary. In particular, there is only one possible slope change that can happen at a time. More specifically, when a not-yet-born trait becomes alive due to the previous trait reaching the typical size allowing for mutations. When this happens, the new born trait takes the slope the maximum between its inner growth rate or the current slope of the previous trait (as mentioned in point (i) of the heuristics preceding Theorem 4.2.9). Any alive trait cannot update its slope because no backward mutation is permitted with this graph structure. Moreover, only a single trait becomes alive at each time, due to the scaling labels $\ell(i, i+1)$ being positive.

Corollary 4.2.10 (Theorem 4.2.9 applied to a mono-directional graph)**.** *Assume the graph is infinite and mono-directional, i.e.* $(V, E) = (\mathbb{N}_0, \{(i, i+1), i \in \mathbb{N}_0\})$ *and that* $\ell^* := \inf\{\ell(i, i+1), i \in \mathbb{N}_0\} > 0$. *Then the limiting functions* $(x_i)_{i \in \mathbb{N}_0}$ *of Theorem 4.2.9 have the following simplified form:*

$$\forall t \in \mathbb{R}^+, x_i(t) := \lambda_{\max}(i)(t - \widetilde{t}(i))_+,$$

where $\lambda_{\max}(i) = \max_{j \in \{0, \cdots, i\}} \lambda(j)$ *and* $\widetilde{t}(i) := \sum_{j=0}^{i-1} \frac{\ell(j, j+1)\lambda(0)}{\lambda_{\max}(j)}$.

Remark 4.2.11. *Using the previous corollary, the limits* $(x_v)_{v \in V}$ *defined in Theorem 4.2.9 can be rewritten by using the decomposition via walks. More specifically, let* $v \in V$, *then for any walk* $\gamma \in P(v)$ *define* x_γ *as the limit obtained by applying the previous corollary to the mono-directional graph indexed by this walk* γ. *Then we have* $x_v = \max_{\gamma \in P(v)} x_\gamma$. *The maximum is well-defined because for all* $t \in \mathbb{R}^+$ *the set* $\{\gamma \in P(v) : x_\gamma(t) > 0\}$ *is finite.*

Theorem 4.2.9 is more general than Theorem 4.2.7 in the sense that there is no assumption on the growth rate function, but it is a less refined result. In Remark 4.2.8 1. we made explicit one contribution of Theorem 4.2.7 about capturing a multiplicative factor of $\log(n)$ using the example of Figure 4.1. Next we are going to do a full comparison of Theorem 4.2.7 and 4.2.9 on the example of Figure 4.4.

Comparison between Theorems 4.2.7 and 4.2.9:

The asymptotic function x obtained through Theorem 4.2.9 for the plain purple trait is $x(t) = \mathbb{1}_{\{t\geq 4\}}\lambda(0)(t-4)$. In the caption of Figure 4.4, it is already made explicit that only the plain green walk will contribute to the size order of the plain purple mutant subpopulation. If one denotes respectively by 1, 2 and 3 the vertices on the plain green walk such that this walk is exactly $(0,1,2,3)$, where 3 is the plain purple vertex, the asymptotic limits for vertex 3, captured by Theorem 4.2.7, is for all $t \geq 4$,

$$\begin{aligned}&\frac{2\alpha(0)\mu(0,1)}{\lambda(0)}\cdot\frac{2\alpha(1)\mu(1,2)}{\lambda(0)}\cdot\frac{2\alpha(2)\mu(2,3)}{\lambda(0)-\lambda(3)}W\int_3^t\left(\int_1^u ds\right)du\cdot n^{t-4}\log^2(n)\\ &\quad=\left(\frac{t^2}{2}-t-\frac{3}{2}\right)\frac{8\alpha(0)\alpha(1)\alpha(2)\mu(0,1)\mu(1,2)\mu(2,3)}{\lambda^2(0)\left(\lambda(0)-\lambda(3)\right)}Wn^{t-4}\log^2(n).\end{aligned}$$

In particular, Theorem 4.2.9 captures only the power of n which is $t-4$ whereas Theorem 4.2.7 captures the stochasticity W, a supplementary scaling factor $\log^2(n)$, a time polynomial $\frac{t^2}{2}-t-\frac{3}{2}$ and also a constant depending only on the parameters of the visited vertices $\frac{8\alpha(0)\alpha(1)\alpha(2)\mu(0,1)\mu(1,2)\mu(2,3)}{\lambda^2(0)(\lambda(0)-\lambda(3))}$. To our knowledge, this is the first time that this level of refinement has been captured under the power law mutation rates limit.

Now we make explicit the form of the limiting functions $(x_v)_{v\in V}$ in the special case where we assume the non-increasing growth rate condition. Under this condition, the limiting functions $(x_v)_{v\in V}$ take a very simple form. The only quantity one has to consider is the time $t(v)$ at which trait v becomes alive, where $t(v)$ is defined in Definition 4.2.5. Then after this time, trait v grows at speed $\lambda(0)$ due to the non-increasing growth rate condition. This is made formal in the next corollary.

Corollary 4.2.12 (Theorem 4.2.9 applied with the non-increasing growth rate condition of (4.1.5))**.** *Assume the non-increasing growth rate condition of* (4.1.5)*. Then the limiting functions* $(x_v)_{v\in V}$ *of Theorem 4.2.9 have the following simplified form:*

$$\forall t\in\mathbb{R}^+, x_v(t) = \lambda(0)\left(t-t(v)\right)_+,$$

where $\forall x\in\mathbb{R}, x_+ := x\mathbb{1}_{\{x\in\mathbb{R}^+\}}$.

4.3 First-order asymptotics of the size of the mutant subpopulations for an infinite mono-directional graph

In this section, we consider the model described in Section 4.1 within a specific infinite mono-directional graph setting:

$$(V,E) = (\mathbb{N}_0, \{(i,i+1), i\in\mathbb{N}_0\}).$$

Studying this special case will enable us to address cycles, particularly those generated by backward mutations, in the more general finite graph scenario. We assume the non-increasing growth rate condition given in (4.1.5). For simplicity of notation, we introduce the following new notations for all $i \in \mathbb{N}_0$

$$\mu_i^{(n)} := \mu^{(n)}(i, i+1) \text{ and } \ell(i) := \ell(i, i+1).$$

In other words, the mutation regime is

$$\forall i \in \mathbb{N}_0, n^{\ell(i)} \mu_i^{(n)} \underset{n\to\infty}{\longrightarrow} \mu_i. \tag{4.3.1}$$

Assume that $\ell^* := \inf\{\ell(i) : i \in \mathbb{N}_0\} > 0$. For all $i \in \mathbb{N}_0$, denote by α_i, β_i and λ_i the division, death and growth rates associated to trait i instead of $\alpha(i), \beta(i)$ and $\lambda(i)$. With this setting, three different scenarios can occur during a division event of a cell of trait $i \in \mathbb{N}_0$:

- both daughter cells keep the trait i of their mother cell, with probability $\left(1-\mu_i^{(n)}\right)^2$,
- exactly one daughter cell mutates to the next trait $i+1$ when the second daughter cell keeps the trait i of its mother cell, with probability $2\mu_i^{(n)}\left(1-\mu_i^{(n)}\right)$,
- both daughter cells mutate to the next trait $i+1$, with probability $\left(\mu_i^{(n)}\right)^2$.

This model is graphically represented in Figure 4.5.

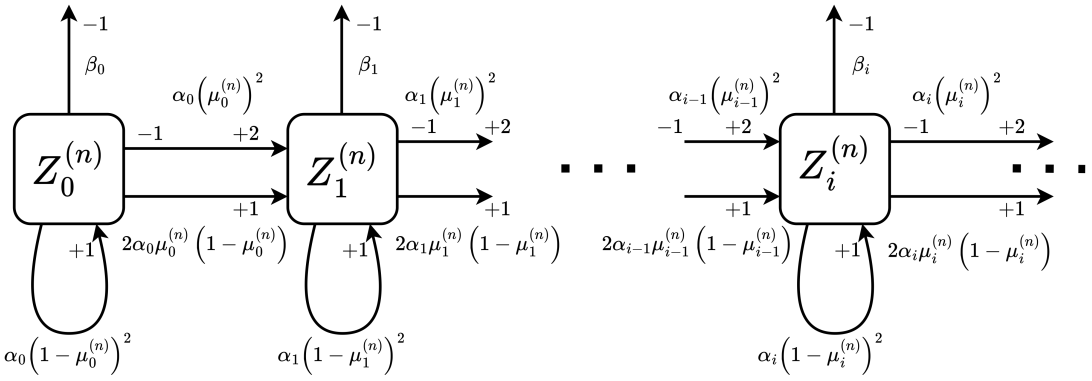

Figure 4.5: Dynamical representation of the infinite mono-directional graph

In particular, any lineage of a cell of trait i follows a birth-death branching process with birth rate $\alpha_i\left(1-\mu_i^{(n)}\right)^2$ and death rate $\beta_i + \alpha_i^{(n)}\left(\mu_i^{(n)}\right)^2$. Thus, we introduce the birth, death and growth rates of any lineage of a cell with trait i as follows

$$\alpha_i^{(n)} := \alpha_i\left(1-\mu_i^{(n)}\right)^2, \beta_i^{(n)} := \beta_i + \alpha_i^{(n)}\left(\mu_i^{(n)}\right)^2 \text{ and } \lambda_i^{(n)} := \alpha_i^{(n)} - \beta_i^{(n)} = \lambda_i - 2\alpha_i\mu_i^{(n)}.$$

Compared to the general finite graph, for this mono-directional graph, there is only one path from trait 0 to any trait $i \in \mathbb{N}$, implying that

$$t(i) = \sum_{i=0}^{i-1} \ell(i) \text{ and } \theta(i) = |\{j \in \{1, \cdots, i\} : \lambda_j = \lambda_0\}|.$$

Let $w_i := w_{(0,1,\cdots,i)}$ denote the weight on the path $(0,\cdots,i)$. The sequence $\left(\left(Z_i^{(n)}\right)_{i\in\mathbb{N}_0}\right)_{n\in\mathbb{N}}$ is mathematically constructed using independent Poisson Point Measures (PPMs). Let $Q_0^b(ds,d\theta)$, $Q_0^d(ds,d\theta)$, $(Q_i(ds,d\theta))_{i\in\mathbb{N}}$, $(N_i(ds,d\theta))_{i\in\mathbb{N}_0}$, and $(Q_i^m(ds,d\theta))_{i\in\mathbb{N}_0}$ be independent PPMs with intensity $dsd\theta$. The subpopulation of wild-type cells is

$$\begin{aligned} Z_0^{(n)}(t) := 1 &+ \int_0^t \int_{\mathbb{R}^+} \mathbb{1}_{\left\{\theta\leq\alpha_0^{(n)}Z_0^{(n)}(s^-)\right\}} Q_0^b(ds,d\theta) \\ &- \int_0^t \int_{\mathbb{R}^+} \mathbb{1}_{\left\{\theta\leq\beta_0 Z_0^{(n)}(s^-)\right\}} Q_0^d(ds,d\theta) - H_0^{(n)}(t) \end{aligned} \tag{4.3.2}$$

and for all $i\in\mathbb{N}$ the mutant subpopulation of trait i is

$$\begin{aligned} Z_i^{(n)}(t) := &\int_0^t \int_{\mathbb{R}^+} \left(\mathbb{1}_{\left\{\theta\leq\alpha_i^{(n)}Z_i^{(n)}(s^-)\right\}} - \mathbb{1}_{\left\{\alpha_i^{(n)}Z_i^{(n)}(s^-)\leq\theta\leq\left(\alpha_i^{(n)}+\beta_i\right)Z_i^{(n)}(s^-)\right\}}\right) Q_i(ds,d\theta) \\ &+ K_{i-1}^{(n)}(t) + 2H_{i-1}^{(n)}(t) - H_i^{(n)}(t), \end{aligned}$$

where for all $i\in\mathbb{N}_0$

$$K_i^{(n)}(t) := \int_0^t \int_{\mathbb{R}^+} \mathbb{1}_{\left\{\theta\leq 2\alpha_i\mu_i^{(n)}\left(1-\mu_i^{(n)}\right)Z_i^{(n)}(s^-)\right\}} N_i(ds,d\theta)$$

and

$$H_i^{(n)}(t) := \int_0^t \int_{\mathbb{R}^+} \mathbb{1}_{\left\{\theta\leq\alpha_i\left(\mu_i^{(n)}\right)^2 Z_i^{(n)}(s^-)\right\}} Q_i^m(ds,d\theta).$$

The processes $\left(K_i^{(n)}(t)\right)_{t\in\mathbb{R}^+}$ and $\left(H_i^{(n)}(t)\right)_{t\in\mathbb{R}^+}$ count the number of mutations up to time t in the subpopulation of trait i that result in exactly one and exactly two mutated daughter cells of trait $i+1$.

Let $(Z_0(t))_{t\in\mathbb{R}^+}$ be the birth-death branching process with birth and death rates α_0 and β_0 constructed in the following way

$$Z_0(t) := 1 + \int_0^t \int_{\mathbb{R}^+} \mathbb{1}_{\{\theta\leq\alpha_0 Z_0(s^-)\}} Q_0^b(ds,d\theta) - \int_0^t \int_{\mathbb{R}^+} \mathbb{1}_{\{\theta\leq\beta_0 Z_0(s^-)\}} Q_0^d(ds,d\theta). \tag{4.3.3}$$

Notice that with this construction, the following monotone coupling immediately holds:

$$\forall t\geq 0, Z_0^{(n)}(t) \leq Z_0(t) \ a.s. \tag{4.3.4}$$

Denote by

$$W := \lim_{t\to\infty} e^{-\lambda_0 t} Z_0(t) \tag{4.3.5}$$

the almost sure limit of the positive martingale $\left(e^{-\lambda_0 t}Z_0(t)\right)_{t\in\mathbb{R}^+}$, whose law is

$$W \overset{law}{=} Ber\left(\frac{\lambda_0}{\alpha_0}\right) \otimes Exp\left(\frac{\lambda_0}{\alpha_0}\right), \tag{4.3.6}$$

see [22, Section 1.1] or [94, Theorem 1].

4.3.1 The wild-type subpopulation dynamics

Using the same PPMs Q_0^b and Q_0^d in the construction of $\left(Z_0^{(n)}\right)_{n\in\mathbb{N}}$ and Z_0 (see Equations (4.3.2) and (4.3.3)) allows us to control the size dynamics of the previous sequence over time by comparing it with the size of Z_0. More precisely, we show that the natural martingale associated with $Z_0^{(n)}$ can be compared to the natural one of Z_0. This comparison follows from the fact that $\left(\alpha_0^{(n)},\beta_0^{(n)}\right)\to(\alpha_0,\beta_0)$ as $n\to\infty$. The control is obtained along the entire trajectory and in probability. The rate of convergence is quantified to be at most of order $\mathcal{O}\left(\mu_0^{(n)}\right)$.

Lemma 4.3.1. *There exist $C(\alpha_0,\lambda_0)>0$ and $N\in\mathbb{N}$ such that for all $\varepsilon>0$ and $n\geq N$,*

$$\mathbb{P}\Big(\sup_{t\in\mathbb{R}^+}\left|e^{-\lambda_0 t}Z_0(t)-e^{-\lambda_0^{(n)}t}Z_0^{(n)}(t)\right|\geq\varepsilon\Big)\leq\frac{C(\alpha_0,\lambda_0)}{\varepsilon^2}\mu_0^{(n)}\underset{n\to\infty}{\longrightarrow}0.$$

Proof of Lemma 4.3.1. Let the filtration $(\mathcal{F}_t)_{t\geq0}$ defined for all $t\geq0$ as

$$\mathcal{F}_t:=\sigma(Q_0^b((0,s]\times A)),Q_0^d((0,s]\times A),s\leq t,A\in\mathcal{B}(\mathbb{R}^+)).$$

Notice that $\left(e^{-\lambda_0 t}Z_0(t)-e^{-\lambda_0^{(n)}t}Z_0^{(n)}(t)\right)_{t\in\mathbb{R}^+}$ is a martingale, with respect to $(\mathcal{F}_t)_{t\in\mathbb{R}^+}$, as the difference between the two martingales $\left(e^{-\lambda_0 t}Z_0(t)\right)_{t\in\mathbb{R}^+}$ and $\left(e^{-\lambda_0^{(n)}t}Z_0^{(n)}(t)\right)_{t\in\mathbb{R}^+}$. Let $(f(m))_{m\in\mathbb{N}}$ be a non decreasing sequence satisfying $f(m)\underset{m\to\infty}{\to}\infty$. Using Doob's Inequality in L^2 (see [95, Proposition 3.15]) we derive

$$\begin{aligned}\mathbb{P}\Big(\sup_{t\in[0,f(m)]}\left|e^{-\lambda_0 t}Z_0(t)-e^{-\lambda_0^{(n)}t}Z_0^{(n)}(t)\right|\geq\varepsilon\Big)\leq\frac{4}{\varepsilon^2}\mathbb{E}\Big[e^{-2\lambda_0 f(m)}Z_0(f(m))^2\\+e^{-2\lambda_0^{(n)}f(m)}Z_0^{(n)}(f(m))^2-2e^{-(\lambda_0+\lambda_0^{(n)})f(m)}Z_0(f(m))Z_0^{(n)}(f(m))\Big].\end{aligned}\tag{4.3.7}$$

Ito's formula and Equation (4.3.4) give

$$\mathbb{E}\left[Z_0(t)Z_0^{(n)}(t)\right]=1+\int_0^t\left(\lambda_0+\lambda_0^{(n)}\right)\mathbb{E}\left[Z_0(s)Z_0^{(n)}(s)\right]ds+\int_0^t\left(\alpha_0^{(n)}+\beta_0\right)\mathbb{E}\left[Z_0^{(n)}(s)\right]ds.$$

By solving this equation, we obtain for all $t\geq0$

$$\mathbb{E}\left[Z_0(t)Z_0^{(n)}(t)\right]=\frac{\alpha_0+\alpha_0^{(n)}}{\lambda_0}e^{\left(\lambda_0+\lambda_0^{(n)}\right)t}-\frac{\alpha_0^{(n)}+\beta_0}{\lambda_0}e^{\lambda_0^{(n)}t}.\tag{4.3.8}$$

Similarly we have

$$\begin{aligned}\mathbb{E}\left[\left(Z_0(t)\right)^2\right]&=\frac{2\alpha_0}{\lambda_0}e^{2\lambda_0 t}-\frac{\alpha_0+\beta_0}{\lambda_0}e^{\lambda_0 t}\leq\frac{2\alpha_0}{\lambda_0}e^{2\lambda_0 t},\\\mathbb{E}\left[\left(Z_0^{(n)}(t)\right)^2\right]&=\frac{2\alpha_0^{(n)}}{\lambda_0^{(n)}}e^{2\lambda_0^{(n)}t}-\frac{\alpha_0^{(n)}+\beta_0^{(n)}}{\lambda_0^{(n)}}e^{\lambda_0^{(n)}t}\leq\frac{2\alpha_0^{(n)}}{\lambda_0^{(n)}}e^{2\lambda_0^{(n)}t}.\end{aligned}\tag{4.3.9}$$

Consequently, combining (4.3.7), (4.3.8) and (4.3.9) yields

$$\begin{aligned}&\mathbb{P}\Big(\sup_{t\in[0,f(m)]}\left|e^{-\lambda_0 t}Z_0(t)-e^{-\lambda_0^{(n)}t}Z_0^{(n)}(t)\right|\geq\varepsilon\Big)\\&\qquad\leq\frac{4}{\varepsilon^2}\Big(\frac{2\alpha_0}{\lambda_0}+\frac{2\alpha_0^{(n)}}{\lambda_0^{(n)}}-2\frac{\alpha_0+\alpha_0^{(n)}}{\lambda_0}+2\frac{\alpha_0^{(n)}+\beta_0}{\lambda_0}e^{-\lambda_0 f(m)}\Big).\end{aligned}$$

The event $\left\{ \sup_{t\in[0,f(m)]} \left|e^{-\lambda_0 t}Z_0(t) - e^{-\lambda_0^{(n)} t}Z_0^{(n)}(t)\right| \geq \varepsilon\right\}$ is increasing with respect to the parameter m. Therefore, taking the limit as $m \to \infty$ and applying the monotonicity of the measure, it follows that

$$\mathbb{P}\Big(\sup_{t\in\mathbb{R}^+} \left|e^{-\lambda_0 t}Z_0(t) - e^{-\lambda_0^{(n)} t}Z_0^{(n)}(t)\right| \geq \varepsilon\Big) \leq \frac{4}{\varepsilon^2}\Big(\frac{2\alpha_0}{\lambda_0} + \frac{2\alpha_0^{(n)}}{\lambda_0^{(n)}} - 2\frac{\alpha_0+\alpha_0^{(n)}}{\lambda_0}\Big).$$

Recalling that $\lambda_0^{(n)} = \lambda_0 - 2\alpha_0\mu_0^{(n)}$ it easily follows that $\frac{2\alpha_0^{(n)}}{\lambda_0^{(n)}} = \frac{2\alpha_0}{\lambda_0} + \frac{4\beta_0\alpha_0}{\lambda_0^2}\mu_0^{(n)} + \mathcal{O}\big(\big(\mu_0^{(n)}\big)^2\big)$ as well as $2\frac{\alpha_0+\alpha_0^{(n)}}{\lambda_0} = \frac{4\alpha_0}{\lambda_0} - \frac{4\alpha_0}{\lambda_0}\mu_0^{(n)} + \mathcal{O}\big(\big(\mu_0^{(n)}\big)^2\big)$. Finally we have

$$\begin{aligned}\mathbb{P}\Big(\sup_{t\in\mathbb{R}^+} \left|e^{-\lambda_0 t}Z_0(t) - e^{-\lambda_0^{(n)} t}Z_0^{(n)}(t)\right| \geq \varepsilon\Big) &\leq \frac{4}{\varepsilon^2}\Big(\frac{4\beta_0\alpha_0}{\lambda_0^2} + \frac{4\alpha_0}{\lambda_0}\Big)\mu_0^{(n)} + \mathcal{O}\big(\big(\mu_0^{(n)}\big)^2\big)\\ &= \frac{16\alpha_0^2}{\varepsilon^2\lambda_0^2}\mu_0^{(n)} + \mathcal{O}\big(\big(\mu_0^{(n)}\big)^2\big),\end{aligned}$$

which concludes the proof. □

The next lemma provides an asymptotic comparison between the stopping times $\eta_t^{(n)}$, at which the wild-type subpopulation reaches the size n^t, and the deterministic times $\mathfrak{t}_t^{(n)}$. This asymptotic comparison is given in probability and is conditioned on $\{W > 0\}$. It clarifies why these deterministic times are the natural candidates for studying the asymptotic behaviour of the mutant subpopulations at the corresponding stopping times. The result is obtained uniformly over time intervals whose lengths tend to infinity, but not too quickly.

Lemma 4.3.2. *For all $\varepsilon > 0$, $(T_1, T_2) \in \mathbb{R}^+$ and φ_n such that $\log(n) = o(\varphi_n)$ and $\varphi_n = o\left(n^{\ell(0)}\right)$, we have*

$$\mathbb{P}\Big(\sup_{t\in\left[T_1, T_2\frac{\varphi_n}{\log(n)}\right]} \left|\eta_t^{(n)} - \Big(\mathfrak{t}_t^{(n)} - \frac{\log(W)}{\lambda_0}\Big)\right| \geq \varepsilon \Big| W > 0\Big) \underset{n\to\infty}{\longrightarrow} 0.$$

Proof of Lemma 4.3.2. Let $\varepsilon > 0$ and for all $n \in \mathbb{N}$ introduce the event

$$A^{(n)} := \Big\{ \sup_{t\in\left[T_1, T_2\frac{\varphi_n}{\log(n)}\right]} \left|\eta_t^{(n)} - \Big(\mathfrak{t}_t^{(n)} - \frac{\log(W)}{\lambda_0}\Big)\right| \geq \varepsilon\Big\}.$$

Step 1: We begin by showing that for all $0 < \delta_1 < \delta_2$

$$\mathbb{P}\left(A^{(n)} \cap \{\delta_1 < W < \delta_2\}\right) \underset{n\to\infty}{\longrightarrow} 0. \tag{4.3.10}$$

Let $\nu > 0$ and $\tilde{\varepsilon} < \frac{\delta_1}{2}$. Firstly, since $e^{-\lambda_0 t}Z_0(t) \underset{t\to\infty}{\to} W$ almost surely, it immediately follows that $Y(t) := \sup_{s\in[t,\infty]} \left|e^{-\lambda_0 s}Z_0(s) - W\right| \underset{t\to\infty}{\longrightarrow} 0$ almost surely and as a consequence in probability. Thus, introducing the event $B_t := \{Y(t) \leq \tilde{\varepsilon}\}$ for all $t > 0$, there exists $t_1 > 0$ such that for all $t \geq t_1$, $\mathbb{P}(B_t) \geq 1 - \frac{\nu}{3}$. Secondly, using Lemma 4.3.1, there exists $n_1 \in \mathbb{N}$ such that for all $n \geq n_1$, $\mathbb{P}\left(C^{(n)}\right) \geq 1 - \frac{\nu}{3}$ where $C^{(n)} := \left\{ \sup_{t\in\mathbb{R}^+} \left|e^{-\lambda_0 t}Z_0(t) - e^{-\lambda_0^{(n)} t}Z_0^{(n)}(t)\right| \leq \tilde{\varepsilon}\right\}$. Combining these two facts, we obtain the following inequality for all $n \geq n_1$

$$\mathbb{P}\left(A^{(n)} \cap \{\delta_1 < W < \delta_2\}\right) \leq \mathbb{P}\left(A^{(n)} \cap \{\delta_1 < W < \delta_2\} \cap B_{t_1} \cap C^{(n)}\right) + \frac{2\nu}{3}. \tag{4.3.11}$$

It remains to show that $\mathbb{P}\left(A^{(n)} \cap \{\delta_1 < W < \delta_2\} \cap B_{t_1} \cap C^{(n)}\right) \leq \frac{\nu}{3}$ for sufficiently large n. Under the event B_{t_1} we have

$$\forall s \geq t_1, (W - \widetilde{\varepsilon})\, e^{\lambda_0 s} \leq Z_0(s) \leq (W + \widetilde{\varepsilon})\, e^{\lambda_0 s}.$$

Given that $\lambda_0^{(n)} \leq \lambda_0$, we obtain that under the event $C^{(n)}$, for all $n \geq n_1$

$$\forall s \in \mathbb{R}^+, \left(e^{-\lambda_0 s} Z_0(s) - \widetilde{\varepsilon}\right) e^{\lambda_0^{(n)} s} \leq Z_0^{(n)}(s) \leq \left(e^{-\lambda_0 s} Z_0(s) + \widetilde{\varepsilon}\right) e^{\lambda_0^{(n)} s} \leq Z_0(s) + \widetilde{\varepsilon} e^{\lambda_0 s}.$$

Combining the two previous inequalities, it follows that under $\{\delta_1 < W < \delta_2\} \cap B_{t_1} \cap C^{(n)}$, we have

$$\forall s \geq t_1, \forall n \geq n_1, (W - 2\widetilde{\varepsilon})\, e^{\lambda_0^{(n)} s} \leq Z_0^{(n)}(s) \leq (W + 2\widetilde{\varepsilon})\, e^{\lambda_0 s} \leq (\delta_2 + 2\widetilde{\varepsilon})\, e^{\lambda_0 s}.$$

Notice that, by definition of $\widetilde{\varepsilon}$, we have $W - 2\widetilde{\varepsilon} > 0$ under the event $\{\delta_1 < W\}$. Now, we introduce the following quantities, which almost surely increase with time:

$$\begin{aligned}
\underline{T}^{(n)}_{\delta_2,t} &:= \inf\{s > 0 : (\delta_2 + 2\widetilde{\varepsilon}) e^{\lambda_0 s} \geq n^t\} = \frac{1}{\lambda_0}\left(t\log(n) - \log(\delta_2 + 2\widetilde{\varepsilon})\right), \\
\underline{T}^{(n)}_{t} &:= \inf\{s > 0 : (W + 2\widetilde{\varepsilon}) e^{\lambda_0 s} \geq n^t\} = \frac{1}{\lambda_0}\left(t\log(n) - \log(W + 2\widetilde{\varepsilon})\right), \\
\overline{T}^{(n)}_{t} &:= \inf\{s > 0 : (W - 2\widetilde{\varepsilon}) e^{\lambda_0^{(n)} s} \geq n^t\} = \frac{1}{\lambda_0^{(n)}}\left(t\log(n) - \log(W - 2\widetilde{\varepsilon})\right).
\end{aligned}$$

We have that there exists $n_2 \in \mathbb{N}$ such that for all $n \geq n_2$, $t_1 \leq \underline{T}^{(n)}_{\delta_2,T_1}$. Moreover, under the event $\{\delta_1 < W < \delta_2\} \cap B_{t_1} \cap C^{(n)}$, we have for all $n \geq \max(n_1, n_2)$ and for all $t \in \left[T_1, T_2 \frac{\varphi_n}{\log(n)}\right]$,

$$\underline{T}^{(n)}_{\delta_2,T_1} \leq \underline{T}^{(n)}_{\delta_2,t} \leq \underline{T}^{(n)}_{t} \leq \eta^{(n)}_t \leq \overline{T}^{(n)}_{t}.$$

Using that $\lambda_0/\lambda_0^{(n)} = 1/\left(1 - 2\alpha_0\mu_0^{(n)}/\lambda_0\right)$, and from the previous equation, we derive that for all $t \in \left[T_1, T_2 \frac{\varphi_n}{\log(n)}\right]$ and for all $n \geq \max(n_1, n_2)$,

$$\begin{aligned}
&\frac{t\log(n)}{\lambda_0} - \frac{\log(W)}{\lambda_0} - \frac{\log\left(1 + 2\tilde{\varepsilon}/W\right)}{\lambda_0} \leq \eta^{(n)}_t \\
&\qquad \leq \Big(\frac{t\log(n)}{\lambda_0} - \frac{\log(W)}{\lambda_0} - \frac{\log(1 - 2\tilde{\varepsilon}/W)}{\lambda_0}\Big)\Big(1 - 2\alpha_0\mu_0^{(n)}/\lambda_0\Big)^{-1}.
\end{aligned}$$

From this we obtain

$$\begin{aligned}
-\frac{\log(1 + 2\tilde{\varepsilon}/W)}{\lambda_0} &\leq \eta^{(n)}_t - \Big(\frac{t\log(n)}{\lambda_0} - \frac{\log(W)}{\lambda_0}\Big) \\
&\leq \Big(1 - 2\alpha_0\mu_0^{(n)}/\lambda_0\Big)^{-1}\Big(\Big(\frac{t\log(n)}{\lambda_0} - \frac{\log(W)}{\lambda_0}\Big)\frac{2\alpha_0\mu_0^{(n)}}{\lambda_0} - \frac{\log(1 - 2\tilde{\varepsilon}/W)}{\lambda_0}\Big).
\end{aligned}$$

In particular, this implies that for all $n \geq \max(n_1, n_2)$,

$$\begin{aligned}
\sup_{t \in \left[T_1, T_2 \frac{\varphi_n}{\log(n)}\right]} \Big|\eta^{(n)}_t - \Big(\frac{t\log(n)}{\lambda_0} - \frac{\log(W)}{\lambda_0}\Big)\Big| &\leq \max\Big\{\frac{\log(1 + 2\tilde{\varepsilon}/W)}{\lambda_0} \\
&; \Big(1 - 2\alpha_0\mu_0^{(n)}/\lambda_0\Big)^{-1}\Big(\Big(\frac{T_2\varphi_n}{\lambda_0} - \frac{\log(W)}{\lambda_0}\Big)\frac{2\alpha_0\mu_0^{(n)}}{\lambda_0} - \frac{\log(1 - 2\tilde{\varepsilon}/W)}{\lambda_0}\Big)\Big\}.
\end{aligned}$$

Denote by $D^{(n)}$ the right-hand side of the last inequality. Then it directly follows that

$$\mathbb{P}\left(A^{(n)} \cap \{\delta_1 < W < \delta_2\} \cap B_{t_1} \cap C^{(n)}\right) \leq \mathbb{P}\left(\{D^{(n)} \geq \varepsilon\} \cap \{\delta_1 < W < \delta_2\}\right). \tag{4.3.12}$$

Because φ_n was defined such that $\varphi_n \mu_0^{(n)} \underset{n\to\infty}{\to} 0$, it is possible to find an adequate $\widetilde{\varepsilon} > 0$ and $n_3 \in \mathbb{N}$ such that for all $n \geq n_3$, $\mathbb{P}\left(\{D^{(n)} \geq \varepsilon\} \cap \{\delta_1 < W < \delta_2\}\right) \leq \frac{\nu}{3}$. Together with (4.3.11) and (4.3.12), we deduce (4.3.10).

Step 2: We are going to prove that $\mathbb{P}\left(A^{(n)} \cap \{W > 0\}\right) \underset{n\to\infty}{\longrightarrow} 0$. We have

$$\mathbb{P}\left(A^{(n)} \cap \{W > 0\}\right) \leq \mathbb{P}\left(A^{(n)} \cap \{\delta_1 < W < \delta_2\}\right) + \mathbb{P}\left(0 < W < \delta_1\right) + \mathbb{P}\left(W > \delta_2\right).$$

Using Equation (4.3.10), we obtain

$$\limsup_{n\to\infty} \mathbb{P}\left(A^{(n)} \cap \{W > 0\}\right) \leq \mathbb{P}\left(0 < W < \delta_1\right) + \mathbb{P}\left(\delta_2 < W\right).$$

Taking the limit as $(\delta_1, \delta_2) \underset{n\to\infty}{\longrightarrow} (0, \infty)$, and noting that W is finite almost surely (see (4.3.6)), we conclude the proof. □

Remark 4.3.3. *From Lemma 4.3.2, the useful results*

$$\mathbb{P}\Big(\sup_{t\in\left[T_1, T_2 \frac{\varphi_n}{\log(n)}\right]} \left|\frac{\eta_t^{(n)}}{\log(n)}\lambda_0 - t\right| \geq \varepsilon \Big| W > 0\Big) \underset{n\to\infty}{\longrightarrow} 0$$

and

$$\mathbb{P}\Big(\sup_{t\in\left[T_1, T_2 \frac{\varphi_n}{\log(n)}\right]} \left|e^{-\lambda_0\left(\eta_t^{(n)} - \mathfrak{t}_t^{(n)}\right)} - W\right| \geq \varepsilon \Big| W > 0\Big) \underset{n\to\infty}{\longrightarrow} 0 \tag{4.3.13}$$

follow.

4.3.2 The mutant subpopulations dynamics in the deterministic time scale (Theorem 4.2.7 (i))

In this subsection, Equations (4.2.5) and (4.2.6) are proven for the mono-directional graph. The proof will be carried out in two steps. First, we will show the result for a fixed $s \in \mathbb{R}$, uniformly in the parameter t. Then, in the second step, we will establish the result uniformly in the parameter s by adapting a method developed in [56, Lemma 3].

Uniform control on the time parameter t

In this subsection, we will prove the following proposition, which is a less refined result than (4.2.5) and (4.2.6), as it is not uniform in the parameter s.

Proposition 4.3.4. *Let $i \in \mathbb{N}$, $(\psi_n(i), h_n(i)) \underset{n\to\infty}{\to} \infty$ such that there exist $\varphi_n(i) \underset{n\to\infty}{\to} \infty$ such that $h_n(i) = \frac{\log(n)}{\log(\log(n))\theta(i-1)+\varphi_n(i)}$ and $\sqrt{\log(n)} = o(\psi_n(i))$. For all $(t,s) \in \mathbb{R}^+ \times \mathbb{R}$ define*

$$d_i^{(n)}(t,s) := \mathbb{1}_{\{t\in[0,t(i)-h_n^{-1}(i))\}} + \mathbb{1}_{\{t\in[t(i)-h_n^{-1}(i),t(i))\}}\psi_n \log^{\theta(i)-1}(n) + \mathbb{1}_{\{t\in[t(i),\infty)\}} n^{t-t(i)} \log^{\theta(i)}(n) e^{\lambda(0)s}.$$

Let $T > 0$, $0 < T_1 < T_2$, and $s \in \mathbb{R}$. We have

- *If $\lambda_i = \lambda_0$ then $t \mapsto Z_i^{(n)}\left(\mathfrak{t}_t^{(n)} + s\right)/d_i^{(n)}(t,s)$ converges in probability in $L^\infty([0,T])$ to $Ww_i(t)$.*
- *If $\lambda_i < \lambda_0$ then $t \mapsto Z_i^{(n)}\left(\mathfrak{t}_{t(i)+t}^{(n)} + s\right)/n^t \log^{\theta(i)}(n) e^{\lambda_0 s}$ converges in probability in $L^\infty([T_1,T_2])$ to $Ww_i(t(i)+t)$.*

The proof is carried out by induction on $i \in \mathbb{N}$. For $i \geq 2$, we assume that Proposition 4.3.4 holds for $i-1$. In the base case, $i = 1$, Proposition 4.3.4 is proved without any assumptions. As long as the proof is similar for the initialisation and the inductive step, the specific step under consideration will not be indicated. To make the proof as clear as possible, it is divided into several lemmas. All results are derived using a martingale approach. In the next lemma, we introduce the martingales considered for all mutant subpopulations and compute their quadratic variations.

Lemma 4.3.5. *For all $i \in \mathbb{N}$ define*

$$M_i^{(n)}(t) := Z_i^{(n)}(t) e^{-\lambda_i^{(n)} t} - \int_0^t 2\alpha_{i-1}\mu_{i-1}^{(n)} e^{-\lambda_i^{(n)} s} Z_{i-1}^{(n)}(s) ds. \tag{4.3.14}$$

$\left(M_i^{(n)}(t)\right)_{t\geq 0}$ is a martingale, with quadratic variation

$$\left\langle M_i^{(n)} \right\rangle_t = \int_0^t 2\alpha_{i-1}\mu_{i-1}^{(n)} e^{-2\lambda_i^{(n)} s} Z_{i-1}^{(n)}(s) ds + \left(\alpha_i^{(n)} + \beta_i^{(n)}\right) \int_0^t e^{-2\lambda_i^{(n)} s} Z_i^{(n)}(s) ds. \tag{4.3.15}$$

Since the proof of this lemma is fairly standard, it can be found in the Appendix. We can now proceed to prove Proposition 4.3.4. This proof is structures as follows:

1. **Neutral case ($\lambda_i = \lambda_0$):** The proof begins by addressing the neutral case. This part is divided into three major steps, each corresponding to a different time regime for the normalising term $d_i^{(n)}(t,s)$:
   (i) **First time regime ($t \in [0, t(i) - h_n^{-1}(i)]$):** Lemma 4.3.6 establishes the asymptotic result for this interval.
   (ii) **Second time regime ($t \in [t(i) - h_n^{-1}(i)]$):** Lemma 4.3.7 covers the convergence within this interval. The proof begins with a first step where the result is established under a more restrictive condition on $\psi_n(i)$. This step is further divided using: Lemma 4.3.8, which handles the finite variation process associated with the mutant subpopulation, Lemma 4.3.9, which controls the expected value of the size of the mutant subpopulation and Lemma 4.3.10, which controls the expected value of the quadratic variation of the martingale associated with the mutant subpopulation. The second step of the proof proceeds by relaxing the aforementioned restrictive condition on $\psi(n)$ from step 1.

(iii) **Third time regime ($t \in [t(i), T]$):** Lemma 4.3.11 controls the finite variation process associated with the mutant subpopulation in this regime.

2. **Deleterious case ($\lambda_i < \lambda_0$):** After completing the neutral case, the proof proceeds to the deleterious case, using some of the previously proven lemmas.

Proof of Proposition 4.3.4. Let $i \in \mathbb{N}$. For $i \geq 2$ assume that Proposition 4.3.4 holds for $i-1$. We begin by proving the result when i is a neutral trait; specifically, we aim to establish Proposition 4.3.4 (i). All the lemmas mentioned in the proof are not restricted to this neutral assumption and also apply to deleterious mutant traits.

1. Neutral case: Assume that $\lambda_i = \lambda_0$. Let $(\psi_n(i), h_n(i))$ as in Proposition 4.3.4 and $\varepsilon > 0$. Notice that

$$\mathbb{P}\Big(\sup_{t\in[0,T]} \Big| \frac{Z_i^{(n)}\left(\mathfrak{t}_t^{(n)} + s\right)}{d_i^{(n)}(t,s)} - W w_i(t) \Big| \geq 3\varepsilon \Big)$$

$$\leq \mathbb{P}\Big(\sup_{t\in\left[0, t(i) - h_n^{-1}(i)\right)} Z_i^{(n)}\left(\mathfrak{t}_t^{(n)} + s\right) \geq \varepsilon \Big) \tag{4.3.16}$$

$$+ \mathbb{P}\Big(\sup_{t\in\left[t(i) - h_n^{-1}(i), t(i)\right)} \frac{Z_i^{(n)}\left(\mathfrak{t}_t^{(n)} + s\right)}{\psi_n(i)\log^{\theta(i-1)}(n)} \geq \varepsilon \Big) \tag{4.3.17}$$

$$+ \mathbb{P}\Big(\sup_{t\in[t(i),T]} \Big| \frac{Z_i^{(n)}\left(\mathfrak{t}_t^{(n)} + s\right)}{n^{t-t(i)}\log^{\theta(i)}(n) e^{\lambda_0 s}} - W w_i(t) \Big| \geq \varepsilon \Big), \tag{4.3.18}$$

where we used that $\omega_i(t) = 0$ for all $t \leq t(i)$. We will show that (4.3.16), (4.3.17) and (4.3.18) converge to 0 as n goes to infinity.

(i) First time regime ($t \in [0, t(i) - h_n^{-1}(i)]$), convergence to 0 of (4.3.16): The characterisation of $t(i)$ as the first time mutant cells of trait i appear on the time scale $t \mapsto \mathfrak{t}_t^{(n)}$ is made explicit in the next lemma. More precisely, we exactly show that up until time $t(i) - h_n^{-1}(i)$, asymptotically no mutant cells of trait i are generated. In particular, the convergence to 0 of (4.3.16) is deduced from the next lemma.

Lemma 4.3.6. *Let $i \in \mathbb{N}$, and $h_n(i) = \frac{\log(n)}{\log(\log(n))\theta(i-1) + \varphi_n(i)}$, where $\varphi_n(i) \underset{n\to\infty}{\to} \infty$ such that $h_n(i) \underset{n\to\infty}{\to} \infty$, and $s \in \mathbb{R}$. For $i \geq 2$, we prove that if Proposition 4.3.4 holds for $i-1$ then*

$$\mathbb{P}\Big(\sup_{t\in\left[0, t(i) - h_n^{-1}(i)\right]} Z_i^{(n)}\left(\mathfrak{t}_t^{(n)} + s\right) = 0 \Big) \underset{n\to\infty}{\longrightarrow} 1. \tag{4.3.19}$$

For $i = 1$, we prove (4.3.19) without any conditions.

Proof of Lemma 4.3.6. Notice first that

$$\Big\{ \sup_{t\in\left[0, t(i) - h_n^{-1}(i)\right]} Z_i^{(n)}\left(\mathfrak{t}_t^{(n)} + s\right) = 0 \Big\} = A^{(n)} \cap B^{(n)}, \tag{4.3.20}$$

where $A^{(n)} := \big\{ K_{i-1}^{(n)}\big(\mathfrak{t}^{(n)}_{t(i) - h_n^{-1}(i)} + s\big) = 0 \big\}$ and $B^{(n)} := \big\{ H_{i-1}^{(n)}\big(\mathfrak{t}^{(n)}_{t(i) - h_n^{-1}(i)} + s\big) = 0 \big\}$. Indeed, the event on the left-hand side of (4.3.20) is satisfied if and only if no mutant cell of the subpopulation

$Z_i^{(n)}$ is generated from the subpopulation $Z_{i-1}^{(n)}$ up until time $\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)} + s$. This corresponds almost surely to $A^{(n)} \cap B^{(n)}$. In what follows, we will provide the proof that $\mathbb{P}\left(A^{(n)}\right) \underset{n\to\infty}{\longrightarrow} 1$. The proof that $\mathbb{P}\left(B^{(n)}\right) \underset{n\to\infty}{\longrightarrow} 1$ can be established using a similar method, so it will not be detailed here. This will conclude the proof of Lemma 4.3.6. Therefore, we now address the proof that $\mathbb{P}\left(A^{(n)}\right) \underset{n\to\infty}{\longrightarrow} 1$, which will vary slightly depending on whether $i = 1$ or $i \geq 2$.

Case $i = 1$: For all $\widetilde{t} \in \mathbb{R}^+$ and $\varepsilon \in \mathbb{R}^+$, let $C_{\varepsilon,\widetilde{t}} := \left\{ \sup_{s\in[\widetilde{t},\infty]} \left| e^{-\lambda_0 s} Z_0(s) - W \right| \leq \varepsilon \right\}$. Using the almost sure inequality (4.3.4), under the event $C_{\varepsilon,\widetilde{t}}$, we have

$$K_0^{(n)}\left(\mathfrak{t}^{(n)}_{t(1)-h_n^{-1}(1)} + s\right) \leq \int_0^{\widetilde{t}} \int_{\mathbb{R}^+} \mathbb{1}_{\left\{\theta \leq 2\alpha_0 \mu_0^{(n)} \sup_{v\in[0,\widetilde{t}]} Z_0(v)\right\}} N_0(du, d\theta) \tag{4.3.21}$$
$$+ \int_{\widetilde{t}}^{\mathfrak{t}^{(n)}_{t(1)-h_n^{-1}(1)}+s} \int_{\mathbb{R}^+} \mathbb{1}_{\left\{\theta \leq 2\alpha_0 \mu_0^{(n)} e^{\lambda_0 u}(\varepsilon + W)\right\}} N_0(du, d\theta).$$

Let us set the following notations

$$D_{\widetilde{t}}^{(n)} := \left\{ \int_0^{\widetilde{t}} \int_{\mathbb{R}^+} \mathbb{1}_{\left\{\theta \leq 2\alpha_0 \mu_0^{(n)} \sup_{v\in[0,\widetilde{t}]} Z_0(v)\right\}} N_0(du, d\theta) = 0 \right\},$$
$$E_{\varepsilon,\widetilde{t}}^{(n)} := \left\{ \int_{\widetilde{t}}^{\mathfrak{t}^{(n)}_{t(1)-h_n^{-1}(1)}+s} \int_{\mathbb{R}^+} \mathbb{1}_{\left\{\theta \leq 2\alpha_0 \mu_0^{(n)} e^{\lambda_0 u}(\varepsilon + W)\right\}} N_0(du, d\theta) = 0 \right\}.$$

Using Equation (4.3.21) we have that

$$\mathbb{P}\Big(A^{(n)}\Big) \geq \mathbb{P}\Big(A^{(n)} \cap C_{\varepsilon,\widetilde{t}}\Big) \geq \mathbb{P}\Big(C_{\varepsilon,\widetilde{t}} \cap D_{\widetilde{t}}^{(n)} \cap E_{\varepsilon,\widetilde{t}}^{(n)}\Big).$$

It remains to show that the right-hand side converges to 1. By the definition of W as the almost sure time limit of the positive martingale $e^{-\lambda_0 t} Z_0(t)$, it follows that $\mathbb{P}\big(C_{\varepsilon,\widetilde{t}}\big) \underset{\widetilde{t}\to\infty}{\longrightarrow} 1$. We also have that $\sup_{v\in[0,\widetilde{t}]} Z_0(v)$ is finite almost surely. Combined with the fact that Z_0 and N_0 are independent, we deduce that $\mathbb{P}\big(D_{\widetilde{t}}^{(n)}\big) \underset{n\to\infty}{\longrightarrow} 1$, because $\mu_0^{(n)} \underset{n\to\infty}{\longrightarrow} 0$. Recall the distribution of W, given in (4.3.6). Since W and N_0 are independent, we have

$$\mathbb{P}\Big(E_{\varepsilon,\widetilde{t}}^{(n)}\Big) = \frac{\beta_0}{\alpha_0} \mathbb{P}\Big(\int_{\widetilde{t}}^{\mathfrak{t}^{(n)}_{t(1)-h_n^{-1}(1)}+s} \int_{\mathbb{R}^+} \mathbb{1}_{\left\{\theta \leq 2\alpha_0 \mu_0^{(n)} e^{\lambda_0 u}\varepsilon\right\}} N_0(du, d\theta) = 0 \Big) + \frac{\lambda_0}{\alpha_0} \tag{4.3.22}$$
$$\cdot \int_0^{\infty} \frac{\lambda_0}{\alpha_0} e^{-\frac{\lambda_0}{\alpha_0} w} \mathbb{P}\Big(\int_{\widetilde{t}}^{\mathfrak{t}^{(n)}_{t(1)-h_n^{-1}(1)}+s} \int_{\mathbb{R}^+} \mathbb{1}_{\left\{\theta \leq 2\alpha_0 \mu_0^{(n)} e^{\lambda_0 u}(\varepsilon + w)\right\}} N_0(du, d\theta) = 0 \Big) dw$$

$$
\begin{aligned}
&= \frac{\beta_0}{\alpha_0} \exp\Big(- \int_{\widetilde{t}}^{\mathfrak{t}^{(n)}_{t(1)-h_n^{-1}(1)}+s} 2\alpha_0 \mu_0^{(n)} e^{\lambda_0 u} \varepsilon du \Big) \\
&\qquad + \frac{\lambda_0}{\alpha_0} \int_0^\infty \frac{\lambda_0}{\alpha_0} e^{-\frac{\lambda_0}{\alpha_0} w} \exp\Big(- \int_{\widetilde{t}}^{\mathfrak{t}^{(n)}_{t(1)-h_n^{-1}(1)}+s} 2\alpha_0 \mu_0^{(n)} e^{\lambda_0 u} (\varepsilon + w) du \Big) dw \\
&\geq \frac{\beta_0}{\alpha_0} \exp\Big(- \frac{2\alpha_0 \Big(n^{t(1)} \mu_0^{(n)} \Big)}{\lambda_0} \varepsilon e^{\lambda_0 s} e^{-h_n^{-1}(1)\log(n)} \Big) \\
&\qquad + \frac{\lambda_0}{\alpha_0} \int_0^\infty \frac{\lambda_0}{\alpha_0} e^{-\frac{\lambda_0}{\alpha_0} w} \exp\Big(- \frac{2\alpha_0 \Big(n^{t(1)} \mu_0^{(n)} \Big)}{\lambda_0} (\varepsilon + w) e^{\lambda_0 s} e^{-h_n^{-1}(1)\log(n)} \Big) dw \\
&\underset{n\to\infty}{\longrightarrow} 1,
\end{aligned}
$$

where we first use that for all $w \geq 0$,

$$
\frac{2\alpha_0 \big(n^{t(1)} \mu_0^{(n)} \big)}{\lambda_0} (\varepsilon + w) e^{\lambda_0 s} e^{-h_n^{-1}(1)\log(n)} \underset{n\to\infty}{\longrightarrow} 0.
$$

This follows from the choice of $h_n(1)$, which ensures that $h_n^{-1}(1)\log(n) \underset{n\to\infty}{\longrightarrow} \infty$. Then, we apply the dominated convergence theorem to obtain

$$
\int_0^\infty \frac{\lambda_0}{\alpha_0} e^{-\frac{\lambda_0}{\alpha_0} w} e^{\Big(- \frac{2\alpha_0 \left(n^{t(1)} \mu_0^{(n)} \right)}{\lambda_0} (\varepsilon + w) e^{\lambda_0 s} e^{-h_n^{-1}(1)\log(n)} \Big)} dw \underset{n\to\infty}{\to} \int_0^\infty \frac{\lambda_0}{\alpha_0} e^{-\frac{\lambda_0}{\alpha_0} w} dw = 1.
$$

Finally, we have shown that $\mathbb{P}\left(A^{(n)} \right) \underset{n\to\infty}{\longrightarrow} 1$, which concludes the proof for the case $i = 1$.

Case $i \geq 2$: Let $\widetilde{t}(i) := \frac{t(i)+t(i-1)}{2}$ and $\Psi_n \underset{n\to\infty}{\longrightarrow} \infty$. For $\varepsilon > 0$, define

$$
C_\varepsilon^{(n)} := \Big\{ \sup_{t\in[0,t(i)]} \Big| \frac{Z_{i-1}^{(n)}\Big(\mathfrak{t}_t^{(n)} \Big)}{d^{(n)}(t)} - W \mathbb{1}_{\{t \geq \widetilde{t}(i)\}} w_{i-1}(t) \Big| \leq \varepsilon \Big\},
$$

where

$$
d^{(n)}(t) = \mathbb{1}_{\left\{t\in[0,\widetilde{t}(i))\right\}} n^{\widetilde{t}(i)-t(i-1)} \log^{\theta(i-1)}(n) \Psi_n + \mathbb{1}_{\left\{t\in\left[\widetilde{t}(i),t(i)\right]\right\}} n^{t-t(i-1)} \log^{\theta(i-1)}(n).
$$

Under the event $C_\varepsilon^{(n)}$, we have

$$
\begin{aligned}
K_{i-1}^{(n)}\Big(\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)} + s \Big) &\leq \int_0^{\mathfrak{t}^{(n)}_{t(i)}} \int_{\mathbb{R}^+} \mathbb{1}_{\left\{\theta \leq 2\alpha_{i-1}\mu_{i-1}^{(n)} \varepsilon n^{\widetilde{t}(i)-t(i-1)} \log^{\theta(i-1)}(n) \Psi_n\right\}} N_{i-1}(du, d\theta) \quad (4.3.23) \\
&+ \int_{\mathfrak{t}^{(n)}_{t(i)}}^{\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)}+s} \int_{\mathbb{R}^+} \mathbb{1}_{\left\{\theta \leq 2\alpha_{i-1}\mu_{i-1}^{(n)} (\varepsilon + W w_{i-1}(t(i))) e^{\lambda_0 u} n^{-t(i-1)} \log^{\theta(i-1)}(n)\right\}} N_{i-1}(du, d\theta).
\end{aligned}
$$

Let us introduce the events

$$D_\varepsilon^{(n)} := \Big\{ \int_0^{\mathrm{t}_{\widetilde{t}(i)}^{(n)}} \int_{\mathbb{R}^+} \mathbb{1}_{\left\{\theta \le 2\alpha_{i-1}\mu_{i-1}^{(n)} \varepsilon n^{\widetilde{t}(i)-t(i-1)} \log^{\theta(i-1)}(n)\Psi_n\right\}} N_{i-1}(du, d\theta) = 0 \Big\},$$

$$E_\varepsilon^{(n)} := \Big\{ \int_{\mathrm{t}_{\widetilde{t}(i)}^{(n)}}^{\mathrm{t}_{t(i)-h_n^{-1}(i)}^{(n)}+s} \int_{\mathbb{R}^+} \mathbb{1}_{\left\{\theta \le 2\alpha_{i-1}\mu_{i-1}^{(n)}(\varepsilon + W w_{i-1}(t(i)))e^{\lambda_0 u} n^{-t(i-1)} \log^{\theta(i-1)}(n)\right\}} \cdot N_{i-1}(du, d\theta) = 0 \Big\}.$$

From (4.3.23) we obtain $\mathbb{P}\big(A^{(n)}\big) \ge \mathbb{P}\big(A^{(n)} \cap C_\varepsilon^{(n)}\big) \ge \mathbb{P}\big(C_\varepsilon^{(n)} \cap D_\varepsilon^{(n)} \cap E_\varepsilon^{(n)}\big)$. It remains to show that the right-hand side converges to 1. By assuming Proposition 4.3.4 holds for trait $i-1$, it follows that $\mathbb{P}\big(C_\varepsilon^{(n)}\big) \underset{n\to\infty}{\longrightarrow} 1$. Secondly, we have

$$\mathbb{P}\Big(D_\varepsilon^{(n)}\Big) = \exp\Big(- \widetilde{t}(i) \frac{\log^{\theta(i-1)+1}(n)}{\lambda_0} 2\alpha_{i-1}\mu_{i-1}^{(n)} \varepsilon \sqrt{n^{\ell(i-1)}} \Psi_n \Big) \underset{n\to\infty}{\longrightarrow} 1,$$

because $\widetilde{t}(i) - t(i-1) = \ell(i-1)/2$, and Ψ_n can be chosen to satisfy both $\Psi_n \underset{n\to\infty}{\longrightarrow} \infty$ and $\log^{\theta(i-1)+1}(n)\Psi_n \sqrt{n^{\ell(i-1)}} \mu_{i-1}^{(n)} \underset{n\to\infty}{\longrightarrow} 0$. Using a similar approach as in the computation of (4.3.22), we get

$$\begin{aligned}
\mathbb{P}\Big(E_\varepsilon^{(n)}\Big) &\ge \frac{\beta_0}{\alpha_0} \exp\Big[- \frac{2\alpha_{i-1}\Big(n^{\ell(i-1)}\mu_{i-1}^{(n)}\Big)}{\lambda_0} \varepsilon \log^{\theta(i-1)}(n) e^{\lambda_0 s} e^{-h_n^{-1}(i)\log(n)} \Big] \\
&+ \frac{\lambda_0}{\alpha_0} \int_0^\infty \frac{\lambda_0}{\alpha_0} e^{-\frac{\lambda_0}{\alpha_0} w} \\
&\cdot \exp\Big(- \frac{2\alpha_{i-1}\Big(n^{\ell(i-1)}\mu_{i-1}^{(n)}\Big)}{\lambda_0} (\varepsilon + w w_{i-1}(t(i))) \log^{\theta(i-1)}(n) e^{\lambda_0 s} e^{-h_n^{-1}(i)\log(n)} \Big) dw \\
&\underset{n\to\infty}{\longrightarrow} 1,
\end{aligned}$$

where, $\forall w \ge 0$,

$$\frac{2\alpha_{i-1}\Big(n^{\ell(i-1)}\mu_{i-1}^{(n)}\Big)}{\lambda_0} (\varepsilon + w w_{i-1}(t(i))) \log^{\theta(i-1)}(n) e^{\lambda_0 s} e^{-h_n^{-1}(i)\log(n)} \underset{n\to\infty}{\longrightarrow} 0,$$

because

$$\log^{\theta(i-1)}(n) e^{-h_n^{-1}(i)\log(n)} = \exp\big(\theta(i-1)\log(\log(n)) - \log(n) h_n^{-1}(i)\big) \underset{n\to\infty}{\longrightarrow} 0$$

by hypothesis on $h_n(i)$. Then, we apply the dominated convergence theorem to get

$$\begin{aligned}
\int_0^\infty \frac{\lambda_0}{\alpha_0} e^{-\frac{\lambda_0}{\alpha_0} w} \exp\Big(- \frac{2\alpha_{i-1}\Big(n^{\ell(i-1)}\mu_{i-1}^{(n)}\Big)}{\lambda_0} \\
\cdot (\varepsilon + w w_{i-1}(t(i))) \log^{\theta(i-1)}(n) e^{\lambda_0 s} e^{-h_n^{-1}(i)\log(n)} \Big) dw \\
\underset{n\to\infty}{\longrightarrow} \int_0^\infty \frac{\lambda_0}{\alpha_0} e^{-\frac{\lambda_0}{\alpha_0} w} dw = 1.
\end{aligned}$$

Finally, we have shown that $\mathbb{P}\left(A^{(n)}\right) \underset{n\to\infty}{\to} 1$, which concludes the proof. □

(ii) Second time regime $t \in [t(i) - h_n^{-1}(i), t(i)]$, convergence to 0 of (4.3.17): In the next lemma we show that in the time interval $[t(i)-h_n^{-1}(i), t(i)]$, the size of the mutant subpopulation of trait i does not achieve any power of n. We control its growth by the factor $\psi_n(i)\log^{\theta(i-1)}(n)$, with a well-chosen function $\psi_n(i)$. Heuristically, the total number of mutant cells of trait i generated from mutational events up to time $t(i)$ is of order $\mathcal{O}\big(\log^{\theta(i-1)}(n)\big)$. Moreover, with the remaining time for the lineages of these mutant cells to grow, we are able to control the size of the mutant subpopulation of trait i by at most $\sqrt{\log(n)}\log^{\theta(i-1)}(n)$. Consequently, by dividing by any function $\psi_n(i)$ that satisfies $\sqrt{\log(n)} = o(\psi_n(i))$, the asymptotic limit is 0.

Lemma 4.3.7. *Let $i \in \mathbb{N}$, $h_n(i) = \frac{\log(n)}{\log(\log(n))\theta(i-1)+\varphi_n(i)}$, where $\varphi_n(i) \underset{n\to\infty}{\to} \infty$ such that $h_n(i) \underset{n\to\infty}{\to} \infty$, $\psi_n(i) \to \infty$ such that $\sqrt{\log(n)} = o(\psi_n(i))$, $s \in \mathbb{R}$ and $\varepsilon > 0$. For $i \geq 2$, we prove that if Proposition 4.3.4 holds for $i-1$ then*

$$\mathbb{P}\Big(\sup_{t\in\left[t(i)-h_n^{-1}(i),t(i)\right]} \frac{Z_i^{(n)}\left(\mathfrak{t}_t^{(n)}+s\right)}{\psi_n(i)\log^{\theta(i-1)}(n)} \geq \varepsilon\Big) \underset{n\to\infty}{\longrightarrow} 0. \tag{4.3.24}$$

For $i = 1$, we prove (4.3.24) without any conditions.

Proof. We begin by proving the same result under the more restrictive condition $\log(n)e^{\varphi_n(i)} = o(\psi_n^2(i))$.

Step 1: Let $\psi_n(i)$ satisfying the previous equation. For all $t \in \left[t(i) - h_n^{-1}(i), t(i)\right]$, we have

$$\begin{aligned}
&\frac{Z_i^{(n)}\left(\mathfrak{t}_t^{(n)}+s\right)}{\psi_n(i)\log^{\theta(i-1)}(n)} \\
&= \frac{Z_i^{(n)}\left(\mathfrak{t}_t^{(n)}+s\right)e^{-\lambda_i^{(n)}\left(\mathfrak{t}_t^{(n)}+s\right)} - Z_i^{(n)}\left(\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)}+s\right)e^{-\lambda_i^{(n)}\left(\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)}+s\right)}}{\psi_n(i)\log^{\theta(i-1)}(n)e^{-\lambda_i^{(n)}\left(\mathfrak{t}_t^{(n)}+s\right)}} \\
&\qquad + \frac{Z_i^{(n)}\left(\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)}+s\right)e^{-\lambda_i^{(n)}\left(\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)}+s\right)}}{\psi_n(i)\log^{\theta(i-1)}(n)e^{-\lambda_i^{(n)}\left(\mathfrak{t}_t^{(n)}+s\right)}} \\
&= \frac{M_i^{(n)}\left(\mathfrak{t}_t^{(n)}+s\right) - M_i^{(n)}\left(\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)}+s\right)}{\psi_n(i)\log^{\theta(i-1)}(n)e^{-\lambda_i^{(n)}\left(\mathfrak{t}_t^{(n)}+s\right)}} + \frac{\int_{\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)}+s}^{\mathfrak{t}_t^{(n)}+s} 2\alpha_{i-1}\mu_{i-1}^{(n)}e^{-\lambda_i^{(n)}u}Z_{i-1}^{(n)}(u)du}{\psi_n(i)\log^{\theta(i-1)}(n)e^{-\lambda_i^{(n)}\left(\mathfrak{t}_t^{(n)}+s\right)}} \\
&\qquad + \frac{Z_i^{(n)}\left(\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)}+s\right)}{\psi_n(i)\log^{\theta(i-1)}(n)e^{-\lambda_i^{(n)}\mathfrak{t}^{(n)}_{t-t(i)+h_n^{-1}(i)}}}.
\end{aligned} \tag{4.3.25}$$

Then, we have

$$\mathbb{P}\Big(\sup_{t\in\left[t(i)-h_n^{-1}(i),t(i)\right]}\frac{Z_i^{(n)}\big(\mathfrak{t}_t^{(n)}+s\big)}{\psi_n(i)\log^{\theta(i-1)}(n)}\geq 3\varepsilon\Big)$$

$$\leq \mathbb{P}\Big(\sup_{t\in\left[t(i)-h_n^{-1}(i),t(i)\right]}\Big|\frac{M_i^{(n)}\big(\mathfrak{t}_t^{(n)}+s\big)-M_i^{(n)}\big(\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)}+s\big)}{\psi_n(i)\log^{\theta(i-1)}(n)e^{-\lambda_i^{(n)}\left(\mathfrak{t}_t^{(n)}+s\right)}}\Big|\geq\varepsilon\Big) \tag{4.3.26}$$

$$+\mathbb{P}\Big(\sup_{t\in\left[t(i)-h_n^{-1}(i),t(i)\right]}\frac{\int_{\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)}+s}^{\mathfrak{t}_t^{(n)}+s}2\alpha_{i-1}\mu_{i-1}^{(n)}e^{-\lambda_i^{(n)}u}Z_{i-1}^{(n)}(u)du}{\psi_n(i)\log^{\theta(i-1)}(n)e^{-\lambda_i^{(n)}\left(\mathfrak{t}_t^{(n)}+s\right)}}\geq\varepsilon\Big) \tag{4.3.27}$$

$$+\mathbb{P}\Big(\sup_{t\in\left[t(i)-h_n^{-1}(i),t(i)\right]}\frac{Z_i^{(n)}\big(\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)}+s\big)}{\psi_n(i)\log^{\theta(i-1)}(n)e^{-\lambda_i^{(n)}\mathfrak{t}^{(n)}_{t-t(i)+h_n^{-1}(i)}}}\geq\varepsilon\Big). \tag{4.3.28}$$

We have (4.3.28) $\leq \mathbb{P}\Big(Z_i^{(n)}\big(\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)}+s\big)\geq 1\Big)$, because a necessary condition for fulfilling the condition of interest is that there is at least one mutant cell of trait i at time $\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)}+s$. Then, applying Lemma 4.3.6 shows that (4.3.28) converges to 0. The convergence to 0 of the term (4.3.27) follows by applying the subsequent lemma. Note, in particular, that $(\psi_n(i),h_n(i))$ satisfies the condition of this lemma.

Lemma 4.3.8. *Let $i\in\mathbb{N}$, $h_n(i)=\frac{\log(n)}{\log(\log(n))\theta(i-1)+\varphi_n(i)}$, where $\varphi_n(i)\underset{n\to\infty}{\to}\infty$ such that $h_n(i)\underset{n\to\infty}{\to}\infty$, $\psi_n(i)\underset{n\to\infty}{\to}\infty$ such that $\log(n)=o(\psi_n(i)h_n(i))$, $s\in\mathbb{R}$ and $\varepsilon>0$. For $i\geq 2$, we prove that if Proposition 4.3.4 holds for $i-1$ then*

$$\mathbb{P}\Big(\sup_{t\in[t(i)-h_n^{-1}(i),t(i)]}\frac{\int_{\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)}+s}^{\mathfrak{t}_t^{(n)}+s}2\alpha_{i-1}\mu_{i-1}^{(n)}e^{-\lambda_i^{(n)}u}Z_{i-1}^{(n)}(u)du}{\psi_n(i)\log^{\theta(i-1)}(n)e^{-\lambda_i^{(n)}\left(\mathfrak{t}_t^{(n)}+s\right)}}\geq\varepsilon\Big)\underset{n\to\infty}{\longrightarrow}0. \tag{4.3.29}$$

For $i=1$, we prove (4.3.29) without any conditions.

Proof of Lemma 4.3.8. Let

$$a_t^{(n)}:=\frac{\int_{\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)}+s}^{\mathfrak{t}_t^{(n)}+s}2\alpha_{i-1}\mu_{i-1}^{(n)}e^{-\lambda_i^{(n)}u}Z_{i-1}^{(n)}(u)du}{\psi_n(i)\log^{\theta(i-1)}(n)e^{-\lambda_i^{(n)}\left(\mathfrak{t}_t^{(n)}+s\right)}}.$$

Our aim is to prove that for all $\varepsilon>0$,

$$\mathbb{P}\Big(\sup_{t\in\left[t(i)-h_n^{-1}(i),t(i)\right]}a_t^{(n)}\leq\varepsilon\Big)\underset{n\to\infty}{\to}1. \tag{4.3.30}$$

Case $i = 1$: We have

$$a_t^{(n)} = \frac{e^{\lambda_1^{(n)}\left(\mathfrak{t}_t^{(n)}+s\right)}}{\psi_n(1)} \int_{\mathfrak{t}^{(n)}_{t(1)-h_n^{-1}(1)}+s}^{\mathfrak{t}_t^{(n)}+s} 2\alpha_0\mu_0^{(n)}\Big[W + \left(e^{-\lambda_0 u}Z_0(u) - W\right) + \left(e^{-\lambda_0^{(n)}u}Z_0^{(n)}(u) - e^{-\lambda_0 u}Z_0(u)\right)\Big] e^{\left(\lambda_0^{(n)}-\lambda_1^{(n)}\right)u} du.$$

Let us define

$$B_{\widetilde{\varepsilon}}^{(n)} := \Big\{ \sup_{u\in\mathbb{R}^+} \left|e^{-\lambda_0 u}Z_0(u) - e^{-\lambda_0^{(n)}u}Z_0^{(n)}(u)\right| \le \widetilde{\varepsilon}\Big\},$$
$$C_{x,\widetilde{\varepsilon}} := \Big\{ \sup_{u\in[x,\infty]} |e^{-\lambda_0 u}Z_0(u) - W| \le \widetilde{\varepsilon}\Big\}.$$

According to Lemma 4.3.1 and the definition of W (see (4.3.5)) we have both that $\mathbb{P}\big(B_{\widetilde{\varepsilon}}^{(n)}\big) \underset{n\to\infty}{\to} 1$ and $\mathbb{P}\big(C_{\sqrt{\log(n)},\widetilde{\varepsilon}}\big) \underset{n\to\infty}{\to} 1$. Then, for sufficiently large n, under the event $B_{\widetilde{\varepsilon}}^{(n)} \cap C_{\sqrt{\log(n)},\widetilde{\varepsilon}}$, we have

$$a_t^{(n)} \le 2\alpha_0\left(n^{t(1)}\mu_0^{(n)}\right)(W + 2\widetilde{\varepsilon})I_n,$$

where $I_n := \frac{e^{\lambda_1^{(n)}\left(\mathfrak{t}_t^{(n)}+s\right)}}{\psi_n(1)n^{t(1)}} \int_{\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(1)}+s}^{\mathfrak{t}_t^{(n)}+s} e^{\left(\lambda_0-\lambda_1^{(n)}\right)u} du$. In the case where $\lambda_1 < \lambda_0$, we have that

$$I_n \le \frac{e^{\lambda_1^{(n)}\left(\mathfrak{t}_t^{(n)}+s\right)}}{\psi_n(1)n^{t(1)}} \frac{e^{\left(\lambda_0-\lambda_1^{(n)}\right)\left(\mathfrak{t}_t^{(n)}+s\right)}}{\lambda_0-\lambda_1} = \frac{e^{-\lambda_0\mathfrak{t}^{(n)}_{t(1)-t}}e^{\lambda_0 s}}{\psi_n(1)(\lambda_0-\lambda_1)} \le \frac{e^{\lambda_0 s}}{\psi_n(1)(\lambda_0-\lambda_1)}. \tag{4.3.31}$$

In the case where $\lambda_1 = \lambda_0$, recalling that $\lambda_1^{(n)} = \lambda_0 - 2\alpha_1\mu_1^{(n)}$, we obtain

$$\begin{aligned} I_n &\le \frac{e^{\lambda_0 s}e^{-2\alpha_1\mu_1^{(n)}\left(\mathfrak{t}_t^{(n)}+s\right)}}{\psi_n(1)} \frac{e^{2\alpha_1\mu_1^{(n)}\left(\mathfrak{t}_t^{(n)}+s\right)} - e^{2\alpha_1\mu_1^{(n)}\left(\mathfrak{t}^{(n)}_{t(1)-h_n^{-1}(1)}+s\right)}}{2\alpha_1\mu_1^{(n)}} \\ &= \frac{e^{\lambda_0 s}}{\psi_n(1)} \frac{1 - e^{-2\alpha_1\mu_1^{(n)}\mathfrak{t}^{(n)}_{t-t(1)+h_n^{-1}(1)}}}{2\alpha_1\mu_1^{(n)}} \\ &\le \frac{e^{\lambda_0 s}\log(n)}{\psi_n(1)h_n(1)\lambda_0}, \end{aligned} \tag{4.3.32}$$

where for the last inequality, we use the fact that $\mathfrak{t}^{(n)}_{t-t(1)+h_n^{-1}(1)} \le \log(n)/(h_n(1)\lambda_0)$ and apply the following equation with $a = 2\alpha_1\mu_1^{(n)} > 0$ and $x = \mathfrak{t}^{(n)}_{t-t(1)+h_n^{-1}(1)}$

$$\forall x \ge 0, \forall a > 0, \frac{1 - e^{-ax}}{a} \le x. \tag{4.3.33}$$

In any case, since W is a finite random variable (see (4.3.6)), we find (4.3.30).

Case $i \geq 2$: Assume Proposition 4.3.4 holds for $i-1$. We have $\mathbb{P}\left(B_{\widetilde{\varepsilon}}^{(n)}\right) \underset{n\to\infty}{\to} 1$ with

$$B_{\widetilde{\varepsilon}}^{(n)} := \Big\{ \sup_{v\in\left[t(i)-h_n^{-1}(i),t(i)\right]} \Big| \frac{Z_{i-1}^{(n)}\left(\mathfrak{t}_v^{(n)}+s\right)}{n^{v-t(i-1)}e^{\lambda_0 s}\log^{\theta(i-1)}(n)} - Ww_{i-1}(v)\Big| \leq \widetilde{\varepsilon}\Big\}.$$

Using the change of variable $u = \mathfrak{t}_v^{(n)} + s$ and the fact that $t(i-1) = t(i) - \ell(i-1)$, notice that

$$a_t^{(n)} = \frac{e^{\lambda_i^{(n)}\left(\mathfrak{t}_t^{(n)}+s\right)}}{\psi_n(i)n^{t(i)}} \int_{t(i)-h_n^{-1}(i)}^{t} 2\alpha_{i-1}\left(n^{\ell(i-1)}\mu_{i-1}^{(n)}\right) \cdot \frac{Z_{i-1}^{(n)}\left(\mathfrak{t}_v^{(n)}+s\right)}{n^{v-t(i-1)}e^{\lambda_0 s}\log^{\theta(i-1)}(n)} e^{\left(\lambda_0-\lambda_i^{(n)}\right)\left(\mathfrak{t}_v^{(n)}+s\right)}\frac{\log(n)}{\lambda_0}dv.$$

Since w_{i-1} is a non-decreasing function, it follows that, under the event $B_{\widetilde{\varepsilon}}^{(n)}$,

$$a_t^{(n)} \leq 2\alpha_{i-1}\left(n^{\ell(i-1)}\mu_{i-1}^{(n)}\right)\left(Ww_{i-1}(t(i)) + \widetilde{\varepsilon}\right)\frac{e^{\lambda_i^{(n)}\left(\mathfrak{t}_t^{(n)}+s\right)}}{\psi_n(i)n^{t(i)}} \int_{\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)}+s}^{\mathfrak{t}_t^{(n)}+s} e^{\left(\lambda_0-\lambda_i^{(n)}\right)u}du.$$

By similar computations as in (4.3.31) and (4.3.32), (4.3.30) follows. □

Now, we will prove that (4.3.26) converges to 0. We begin by introducing two lemmas, whose proofs are provided in the Appendix, which allow us to control both the expected size of any mutant subpopulation and the quadratic variation associated to the martingale $M_i^{(n)}$. First, a natural upper bound on the expected growth of each mutant subpopulation can be easily obtained, as stated in the next lemma.

Lemma 4.3.9. *For all $i \in \mathbb{N}_0$ and $u \geq 0$,*

$$\mathbb{E}\left[Z_i^{(n)}(u)\right] \leq C_i\mu_{\otimes,i}^{(n)}u^{\theta(i)}e^{\lambda_0 u},$$

where $\mu_{\otimes,i}^{(n)} := \prod_{j=1}^{i}\mu_{j-1}^{(n)}$ and $C_i := \prod_{j=1}^{i} 2\alpha_{j-1}\left(\mathbb{1}_{\{\lambda_j=\lambda_0\}} + \mathbb{1}_{\{\lambda_j<\lambda_0\}}\frac{1}{\lambda_0-\lambda_j}\right)$.

Notice that there are three key components. The first is the mutational cost to produce such mutant cells, represented by the term $\mu_{\otimes,i}^{(n)}$. The second component is the contribution over time of all neutral mutations along the path leading to the mutant subpopulation in question. The third component is the exponential growth at rate λ_0 exhibited by the wild-type subpopulation. Additionally, using the expression for the quadratic variation of the martingale associated to a mutant subpopulation, given in Equation (4.3.15), and the previous Lemma 4.3.9, a natural upper bound on its mean is derived and summarized in the next lemma.

Lemma 4.3.10. *Let $0 < t_1^{(n)} < t_2$ and $s \in \mathbb{R}$. There exist $N \in \mathbb{N}$ and $C(i) > 0$ such that, for all $n \geq N$, we have*

$$\mathbb{E}\left[\left\langle M_i^{(n)}\right\rangle_{\mathfrak{t}_{t_2}^{(n)}+s} - \left\langle M_i^{(n)}\right\rangle_{\mathfrak{t}_{t_1}^{(n)}+s}\right] \leq C(i)\mu_{\otimes,i}^{(n)}\Big[\mathbb{1}_{\{\lambda_i=\lambda_0\}}\frac{e^{-\lambda_0 s}\left(\mathfrak{t}_{t_1^{(n)}}^{(n)}+s\right)^{\theta(i)}}{n^{t_1^{(n)}}}\left(\mathfrak{t}_{t_2}^{(n)}+s\right)^{\theta(i)}$$
$$\cdot\left(\mathbb{1}_{\{\lambda_0>2\lambda_i\}}e^{(\lambda_0-2\lambda_i)\left(\mathfrak{t}_{t_2}^{(n)}+s\right)} + \mathbb{1}_{\{\lambda_0=2\lambda_i\}}\left(\mathfrak{t}_{t_2}^{(n)}+s\right) + \mathbb{1}_{\{\lambda_i<\lambda_0<2\lambda_i\}}e^{-(2\lambda_i-\lambda_0)\left(\mathfrak{t}_{t_1^{(n)}}^{(n)}+s\right)}\right)\Big].$$

Now we can prove that (4.3.26) converges to 0. Using the fact that $\log^{\theta(i-1)}(n) = e^{\lambda_0 \mathfrak{t}^{(n)}_{h_n^{-1}(i)}} e^{-\varphi_n(i)}$, we can rewrite for all $t \in \big[t(i) - h_n^{-1}(i), t(i)\big]$

$$\begin{aligned}
&\frac{\Big|M_i^{(n)}\big(\mathfrak{t}_t^{(n)} + s\big) - M_i^{(n)}\big(\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)} + s\big)\Big|}{\psi_n(i)\log^{\theta(i-1)}(n) e^{-\lambda_i^{(n)}\left(\mathfrak{t}_t^{(n)}+s\right)}} \\
&\qquad = \frac{\Big|M_i^{(n)}\big(\mathfrak{t}_t^{(n)} + s\big) - M_i^{(n)}\big(\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)} + s\big)\Big|}{\psi_n(i) e^{-\varphi_n(i)} e^{\left(\lambda_0-\lambda_i^{(n)}\right)\mathfrak{t}^{(n)}_{t-t(i)+h_n^{-1}(i)}} e^{-\lambda_i^{(n)}\left(\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)}+s\right)} e^{\lambda_0 \mathfrak{t}^{(n)}_{t(i)-t}}} \\
&\qquad \leq \frac{\Big|M_i^{(n)}\big(\mathfrak{t}_t^{(n)} + s\big) - M_i^{(n)}\big(\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)} + s\big)\Big|}{\psi_n(i) e^{-\varphi_n(i)} e^{\left(\lambda_0-\lambda_i^{(n)}\right)\mathfrak{t}^{(n)}_{t-t(i)+h_n^{-1}(i)}} e^{-\lambda_i^{(n)}\left(\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)}+s\right)}}.
\end{aligned}$$

In the case where $\lambda_i = \lambda_0$, we simplify the denominator using that $e^{\left(\lambda_0-\lambda_i^{(n)}\right)\mathfrak{t}^{(n)}_{t-t(i)+h_n^{-1}(i)}} \geq 1$. Then, we apply Doob's inequality to the martingale $\big(M_i^{(n)}\big(\mathfrak{t}_t^{(n)} + s\big) - M_i^{(n)}\big(\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)} + s\big)\big)_{t\geq t(i)}$, and use the property that if M is a square integrable martingale, then $\mathbb{E}[(M(t) - M(s))^2] = \mathbb{E}[M^2(t) - M^2(s)] = \mathbb{E}[\langle M\rangle_t - \langle M\rangle_s]$. It follows that

$$\begin{aligned}
(4.3.26) &\leq \mathbb{P}\Big(\sup_{t\in\left[t(i)-h_n^{-1}(i),t(i)\right]} \Big|\frac{M_i^{(n)}\big(\mathfrak{t}_t^{(n)} + s\big) - M_i^{(n)}\big(\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)} + s\big)}{\psi_n(i) e^{-\varphi_n(i)} e^{-\lambda_i^{(n)}\left(\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)}+s\right)}}\Big| \geq \varepsilon\Big) \\
&\leq \frac{4e^{2\lambda_i \mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)}} e^{2\lambda_i^{(n)} s}}{\varepsilon^2 \psi_n^2(i) e^{-2\varphi_n(i)}} \mathbb{E}\Big[\big\langle M_i^{(n)}\big\rangle_{\mathfrak{t}^{(n)}_{t(i)}+s} - \big\langle M_i^{(n)}\big\rangle_{\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)}+s}\Big].
\end{aligned}$$

Applying Lemma 4.3.10 at times $t_1^{(n)} = t(i) - h_n^{-1}(i)$ and $t_2 = t(i)$, there exists a constant $C = C(s, i, \varepsilon)$ (which may change from line to line) such that

$$(4.3.26) \leq \frac{Ce^{2\varphi_n(i)}}{\psi_n^2(i)}\big(n^{t(i)}\mu^{(n)}_{\otimes,i}\big)\big(\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)} + s\big)^{\theta(i)} n^{-h_n^{-1}(i)}.$$

Note that

$$n^{t(i)}\mu^{(n)}_{\otimes,i} = \prod_{j=1}^{i} n^{\ell(j-1)}\mu^{(n)}_{j-1} \underset{n\to\infty}{\longrightarrow} \prod_{j=1}^{i}\mu_{j-1} < \infty. \tag{4.3.34}$$

Then, for n large enough, and recalling that $\theta(i) = \theta(i-1) + 1$, we have

$$(4.3.26) \leq C\frac{\log(n)e^{\varphi_n(i)}}{\psi_n^2(i)} \underset{n\to\infty}{\longrightarrow} 0,$$

according to the scaling of $\psi_n(i)$. In the case where $\lambda_i < \lambda_0$, using the Maximal inequality (see [96, Chapter VI, page 72]) applied to the supermartingale

$$\Big[\frac{M_i^{(n)}\big(\mathfrak{t}_t^{(n)} + s\big) - M_i^{(n)}\big(\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)} + s\big)}{\psi_n(i) e^{-\varphi_n(i)} e^{\left(\lambda_0-\lambda_i^{(n)}\right)\mathfrak{t}^{(n)}_{t-t(i)+h_n^{-1}(i)}} e^{-\lambda_i^{(n)}\left(\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)}+s\right)}}\Big]_{t\geq t(i)-h_n^{-1}(i)},$$

it follows that

$$(4.3.26) \leq \frac{3e^{\lambda_i^{(n)}\left(\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)}+s\right)}}{\varepsilon\psi_n(i)e^{-\varphi_n(i)}} \sup_{t\in\left[t(i)-h_n^{-1}(i),t(i)\right]} f^{(n)}(t), \tag{4.3.35}$$

where $f^{(n)}(t) := e^{-\left(\lambda_0-\lambda_i^{(n)}\right)\mathfrak{t}^{(n)}_{t-t(i)+h_n^{-1}(i)}}\mathbb{E}\left[\left\langle M_i^{(n)}\right\rangle_{\mathfrak{t}^{(n)}_t+s}-\left\langle M_i^{(n)}\right\rangle_{\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)}+s}\right]^{\frac{1}{2}}$. According to Lemma 4.3.10 applied with $t_1^{(n)} = t(i) - h_n^{-1}(i)$ and $t_2 = t(i)$, we have that

$$\begin{aligned}\sup_{t\in\left[t(i)-h_n^{-1}(i),t(i)\right]} f^{(n)}(t) &\leq C\left(\mu^{(n)}_{\otimes,i}\right)^{\frac{1}{2}}\left(\mathfrak{t}^{(n)}_{t(i)}+s\right)^{\frac{\theta(i)}{2}}e^{\left(\lambda_0-\lambda_i^{(n)}\right)\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)}} \\ &\cdot\Big(\mathbb{1}_{\{\lambda_0>2\lambda_i\}}e^{-\frac{\lambda_0}{2}\mathfrak{t}^{(n)}_{t(i)}}e^{\frac{\lambda_0-2\lambda_i}{2}s}+\mathbb{1}_{\{\lambda_0=2\lambda_i\}}\left(\mathfrak{t}^{(n)}_{t(i)}+s\right)^{\frac{1}{2}}e^{-\left(\lambda_0-\lambda_i^{(n)}\right)\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)}} \\ &+\mathbb{1}_{\{\lambda_i<\lambda_0<2\lambda_i\}}e^{-\frac{\lambda_0}{2}\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)}}e^{-\frac{2\lambda_i-\lambda_0}{2}s}\Big).\end{aligned} \tag{4.3.36}$$

Combining (4.3.35), (4.3.36), (4.3.34), and using the facts that $e^{\lambda_0\mathfrak{t}^{(n)}_{t(i)}} = n^{t(i)}$, $e^{\lambda_0\mathfrak{t}^{(n)}_{h_n^{-1}(i)}} = \log^{\theta(i-1)}(n)e^{\varphi_n(i)}$ and $\theta(i-1) = \theta(i)$, it follows that

$$\begin{aligned}(4.3.26) &\leq \frac{Ce^{\varphi_n(i)}}{\varepsilon\psi_n(i)}e^{\lambda_i^{(n)}s}\left(n^{t(i)}\mu^{(n)}_{\otimes,i}\right)^{\frac{1}{2}}\left(\mathfrak{t}^{(n)}_{t(i)}+s\right)^{\frac{\theta_i}{2}}e^{\frac{\lambda_0}{2}\mathfrak{t}^{(n)}_{t(i)}}e^{-\lambda_0\mathfrak{t}^{(n)}_{h_n^{-1}(i)}} \\ &\cdot\Big(\mathbb{1}_{\{\lambda_0>2\lambda_i\}}e^{-\frac{\lambda_0}{2}\mathfrak{t}^{(n)}_{t(i)}}e^{\frac{\lambda_0-2\lambda_i}{2}s}+\mathbb{1}_{\{\lambda_0=2\lambda_i\}}\left(\mathfrak{t}^{(n)}_{t(i)}+s\right)^{\frac{1}{2}}e^{-(\lambda_0-\lambda_i)\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)}} \\ &\qquad+\mathbb{1}_{\{\lambda_i<\lambda_0<2\lambda_i\}}e^{-\frac{\lambda_0}{2}\mathfrak{t}^{(n)}_{t(i)-h_n^{-1}(i)}}e^{-\frac{2\lambda_i-\lambda_0}{2}s}\Big) \\ &\leq \frac{Ce^{\varphi_n(i)}}{\varepsilon\psi_n(i)}\log^{\frac{\theta(i)}{2}}(n)\Big(\mathbb{1}_{\{\lambda_0>2\lambda_i\}}\frac{e^{-\varphi_n(i)}}{\log^{\theta(i-1)}(n)}+\mathbb{1}_{\{\lambda_0=2\lambda_i\}}\log^{\frac{1}{2}}(n)e^{-\frac{\lambda_0}{2}\mathfrak{t}^{(n)}_{h_n^{-1}(i)}} \\ &\qquad+\mathbb{1}_{\{\lambda_i<\lambda_0<2\lambda_i\}}e^{-\frac{\lambda_0}{2}\mathfrak{t}^{(n)}_{h_n^{-1}(i)}}\Big) \\ &\leq \frac{C}{\varepsilon\psi_n(i)}\Big(\mathbb{1}_{\{\lambda_0>2\lambda_i\}}\frac{1}{\log^{\frac{\theta(i)}{2}}(n)}+\mathbb{1}_{\{\lambda_0=2\lambda_i\}}\sqrt{\log(n)}e^{\frac{\varphi_n(i)}{2}}+\mathbb{1}_{\{\lambda_1<\lambda_0<2\lambda_i\}}e^{\frac{\varphi_n(i)}{2}}\Big) \\ &\underset{n\to\infty}{\longrightarrow} 0,\end{aligned}$$

according to the scaling of $\psi_n(i)$.

Step 2: Let $\psi_n(i)$ satisfy $\sqrt{\log(n)} = o(\psi_n(i))$, but such that $\log(n)e^{\varphi_n(i)} \neq o(\psi_n^2(i))$. Let $\widetilde{\varphi}_n(i)$ be such that $\log(n)e^{\widetilde{\varphi}_n(i)} = o(\psi_n^2(i))$, and define $\widetilde{h}_n(i) := \frac{\log(n)}{\log(\log(n))\theta(i-1)+\widetilde{\varphi}_n(i)}$. Notice, in

particular, that $\widetilde{h}_n(i) \geq h_n(i)$. We have

$$\begin{aligned}\mathbb{P}\Big(\sup_{t\in\left[t(i)-h_n^{-1}(i),t(i)\right]} \frac{Z_i^{(n)}\big(\mathfrak{t}_t^{(n)}+s\big)}{\psi_n(i)\log^{\theta(i-1)}(n)} \geq \varepsilon\Big) \\ \leq \mathbb{P}\Big(\sup_{t\in\left[t(i)-h_n^{-1}(i),t(i)-\widetilde{h}_n^{-1}(i)\right]} Z_i^{(n)}\big(\mathfrak{t}_t^{(n)}+s\big) \geq \varepsilon\Big) \\ + \mathbb{P}\Big(\sup_{t\in\left[t(i)-\widetilde{h}_n^{-1}(i),t(i)\right]} \frac{Z_i^{(n)}\big(\mathfrak{t}_t^{(n)}+s\big)}{\psi_n(i)\log^{\theta(i-1)}(n)} \geq \varepsilon\Big),\end{aligned}$$

where the first term on the right-hand side converges to 0 according to Lemma 4.3.6, and the second term converges from Step 1 of this proof. □

(iii) Third time regime ($t \in [t(i), T]$)**, convergence to 0 of** (4.3.18)**:** Applying similar computations as in (4.3.25), notice that for all $t \geq t(i)$

$$\begin{aligned}\frac{Z_i^{(n)}\big(\mathfrak{t}_t^{(n)}+s\big)}{n^{t-t(i)}\log^{\theta(i)}(n)e^{\lambda_0 s}} &= \frac{M_i^{(n)}\big(\mathfrak{t}_t^{(n)}+s\big) - M_i^{(n)}\big(\mathfrak{t}_{t(i)}^{(n)}+s\big)}{n^{-t(i)}\log^{\theta(i)}(n)e^{\left(\lambda_0-\lambda_i^{(n)}\right)\left(\mathfrak{t}_t^{(n)}+s\right)}} \\ &+ \frac{\int_{\mathfrak{t}_{t(i)}^{(n)}+s}^{\mathfrak{t}_t^{(n)}+s} 2\alpha_{i-1}\mu_{i-1}^{(n)}Z_{i-1}^{(n)}(u)e^{-\lambda_i^{(n)}u}du}{n^{-t(i)}\log^{\theta(i)}(n)e^{\left(\lambda_0-\lambda_i^{(n)}\right)\left(\mathfrak{t}_t^{(n)}+s\right)}} + \frac{Z_i^{(n)}\big(\mathfrak{t}_{t(i)}^{(n)}+s\big)}{\log^{\theta(i)}(n)e^{\left(\lambda_0-\lambda_i^{(n)}\right)\mathfrak{t}_{t-t(i)}^{(n)}}e^{\lambda_0 s}}.\end{aligned}$$

Then this allows to write

$$\begin{aligned}\mathbb{P}\Big(\sup_{t\in[t(i),T]} \Big|\frac{Z_i^{(n)}\big(\mathfrak{t}_t^{(n)}+s\big)}{n^{t-t(i)}\log^{\theta(i)}(n)e^{\lambda_0 s}} - Ww_i(t)\Big| \geq 3\varepsilon\Big) \\ \leq \mathbb{P}\Big(\sup_{t\in[t(i),T]} \Big|\frac{M_i^{(n)}\big(\mathfrak{t}_t^{(n)}+s\big) - M_i^{(n)}\big(\mathfrak{t}_{t(i)}^{(n)}+s\big)}{n^{-t(i)}\log^{\theta(i)}(n)e^{\left(\lambda_0-\lambda_i^{(n)}\right)\left(\mathfrak{t}_t^{(n)}+s\right)}}\Big| \geq \varepsilon\Big) \qquad (4.3.37) \\ + \mathbb{P}\Big(\sup_{t\in[t(i),T]} \Big|\frac{\int_{\mathfrak{t}_{t(i)}^{(n)}+s}^{\mathfrak{t}_t^{(n)}+s} 2\alpha_{i-1}\mu_{i-1}^{(n)}Z_{i-1}^{(n)}(u)e^{-\lambda_i^{(n)}u}du}{n^{-t(i)}\log^{\theta(i)}(n)e^{\left(\lambda_0-\lambda_i^{(n)}\right)\left(\mathfrak{t}_t^{(n)}+s\right)}} - Ww_i(t)\Big| \geq \varepsilon\Big) \qquad (4.3.38) \\ + \mathbb{P}\Big(\sup_{t\in[t(i),T]} \frac{Z_i^{(n)}\big(\mathfrak{t}_{t(i)}^{(n)}+s\big)}{\log^{\theta(i)}(n)e^{\left(\lambda_0-\lambda_i^{(n)}\right)\mathfrak{t}_{t-t(i)}^{(n)}}e^{\lambda_0 s}} \geq \varepsilon\Big). \qquad (4.3.39)\end{aligned}$$

We will show that (4.3.37), (4.3.38) and (4.3.39) converge to 0 as n goes to infinity. For the term (4.3.37), we start by using the fact that $\lambda_0 \geq \lambda_i^{(n)}$ to simplify the denominator. Then, we apply Doob's inequality to the martingale $\big(M_i^{(n)}\big(\mathfrak{t}_t^{(n)}+s\big) - M_i^{(n)}\big(\mathfrak{t}_{t(i)}^{(n)}+s\big)\big)\big)_{t\geq t(i)}$ to obtain

$$\begin{aligned}(4.3.37) &\leq \mathbb{P}\Big(\sup_{t\in[t(i),T]} \Big|\frac{M_i^{(n)}\big(\mathfrak{t}_t^{(n)}+s\big) - M_i^{(n)}\big(\mathfrak{t}_{t(i)}^{(n)}+s\big)}{n^{-t(i)}\log^{\theta(i)}(n)}\Big| \geq \varepsilon\Big) \qquad (4.3.40) \\ &\leq \frac{4n^{2t(i)}}{\varepsilon^2\log^{2\theta(i)}(n)}\mathbb{E}\Big[\big\langle M_i^{(n)}\big\rangle_{\mathfrak{t}_T^{(n)}+s} - \big\langle M_i^{(n)}\big\rangle_{\mathfrak{t}_{t(i)}^{(n)}+s}\Big].\end{aligned}$$

By applying Lemma 4.3.10 at times $t_1^{(n)} = t(i)$ and $t_2 = T$, we obtain

$$\mathbb{E}\Big[\big\langle M_i^{(n)}\big\rangle_{\mathfrak{t}_T^{(n)}+s} - \big\langle M_i^{(n)}\big\rangle_{\mathfrak{t}_{t(i)}^{(n)}+s}\Big] \leq C\frac{e^{-\lambda_0 s}\big(\mathfrak{t}_{t(i)}^{(n)}+s\big)^{\theta(i)}\mu_{\otimes,i}^{(n)}}{n^{t(i)}}. \tag{4.3.41}$$

Then, combining (4.3.40) and (4.3.41), we get

$$\begin{aligned}\mathbb{P}\Big(\sup_{t\in[t(i),T]}\Big|\frac{M_{(i}^{(n)}\big(\mathfrak{t}_t^{(n)}+s\big) - M_i^{(n)}\big(\mathfrak{t}_{t(i)}^{(n)}+s\big)}{n^{-t(i)}\log^{\theta(i)}(n)e^{\left(\lambda_0-\lambda_i^{(n)}\right)\left(\mathfrak{t}_t^{(n)}+s\right)}}\Big| \geq \varepsilon\Big)&\\ \leq \frac{4Ce^{-\lambda_0 s}}{\varepsilon^2\log^{\theta(i)}(n)}\Big(\frac{\mathfrak{t}_{t(i)}^{(n)}+s}{\log(n)}\Big)^{\theta(i)}n^{t(i)}\mu_{\otimes,i}^{(n)}&\\ \underset{n\to\infty}{\longrightarrow} 0,&\end{aligned}$$

as $\theta(i) \geq 1$, since the vertex i is assumed to be neutral. This concludes the proof of the convergence to 0 of (4.3.37). The term (4.3.38) also converges to 0, as shown in the following lemma.

Lemma 4.3.11. *Let $i \in \mathbb{N}$, $T \geq t(i)$, $s \in \mathbb{R}$ and $\varepsilon > 0$. For $i \geq 2$, we prove that if Proposition 4.3.4 holds for $i-1$, then*

$$\mathbb{P}\Big(\sup_{t\in[t(i),T]}\Big|\frac{\int_{\mathfrak{t}_{t(i)}^{(n)}+s}^{\mathfrak{t}_t^{(n)}+s} 2\alpha_{i-1}\mu_{i-1}^{(n)}e^{-\lambda_i^{(n)}u}Z_{i-1}^{(n)}(u)du}{n^{-t(i)}\log^{\theta(i)}(n)e^{\left(\lambda_0-\lambda_i^{(n)}\right)\left(\mathfrak{t}_t^{(n)}+s\right)}} - Ww_i(t)\Big| \geq \varepsilon\Big) \underset{n\to\infty}{\longrightarrow} 0. \tag{4.3.42}$$

For $i = 1$, we prove (4.3.42) without any conditions.

Proof of Lemma 4.3.11. Let $c_n(t,s) := e^{\left(\lambda_0-\lambda_i^{(n)}\right)\left(\mathfrak{t}_t^{(n)}+s\right)}$ and

$$a_t^{(n)} := \frac{\int_{\mathfrak{t}_{t(i)}^{(n)}+s}^{\mathfrak{t}_t^{(n)}+s} 2\alpha_{i-1}\mu_{i-1}^{(n)}e^{-\lambda_i^{(n)}u}Z_{i-1}^{(n)}(u)du}{n^{-t(i)}\log^{\theta(i)}(n)c_n(t,s)}.$$

Our aim is to prove that, for all $\varepsilon > 0$, $\mathbb{P}\big(\sup_{t\in[t(i),T]}\big|a_t^{(n)} - Ww_i(t)\big| \leq \varepsilon\big) \underset{n\to\infty}{\to} 1$.

Case $i = 1$: We have

$$\begin{aligned}a_t^{(n)} = \frac{n^{t(1)}}{\log^{\theta(1)}(n)c_n(t,s)}\int_{\mathfrak{t}_{t(1)}^{(n)}+s}^{\mathfrak{t}_t^{(n)}+s} 2\alpha_0\mu_0^{(n)}\Big[W + \big(e^{-\lambda_0 u}Z_0(u) - W\big)&\\ + \Big(e^{-\lambda_0^{(n)}u}Z_0^{(n)}(u) - e^{-\lambda_0 u}Z_0(u)\Big)\Big]e^{\left(\lambda_0^{(n)}-\lambda_1^{(n)}\right)u}du.&\end{aligned}$$

For all $\varepsilon > 0$, introduce the events

$$\begin{aligned}B_{\widetilde{\varepsilon}}^{(n)} &:= \Big\{\sup_{u\in\mathbb{R}^+}\Big|e^{-\lambda_0 u}Z_0(u) - e^{-\lambda_0^{(n)}u}Z_0^{(n)}(u)\Big| \leq \widetilde{\varepsilon}\Big\},\\ C_{x,\widetilde{\varepsilon}} &:= \Big\{\sup_{u\in[x,\infty]}|e^{-\lambda_0 u}Z_0(u) - W| \leq \widetilde{\varepsilon}\Big\}.\end{aligned}$$

According to Lemma 4.3.1 and the definition of W (see (4.3.5)), we have both that $\mathbb{P}\big(B^{(n)}_{\widetilde{\varepsilon}}\big) \underset{n\to\infty}{\to} 1$ and $\mathbb{P}\big(C_{\sqrt{\log(n)},\widetilde{\varepsilon}}\big) \underset{n\to\infty}{\to} 1$. Notice that when $\lambda_1 < \lambda_0$, we have the following bound

$$\frac{1}{c_n(t,s)} \int_{\mathfrak{t}^{(n)}_{t(1)}+s}^{\mathfrak{t}^{(n)}_t+s} e^{\left(\lambda_0-\lambda_1^{(n)}\right)u} du = \frac{1}{\lambda_0-\lambda_1^{(n)}} \frac{c_n(t,s)-c_n(t(1),s)}{c_n(t,s)} \le \frac{1}{\lambda_0-\lambda_1}, \tag{4.3.43}$$

and when $\lambda_1 = \lambda_0$, we have

$$\frac{1}{c_n(t,s)} \int_{\mathfrak{t}^{(n)}_{t(1)}+s}^{\mathfrak{t}^{(n)}_t+s} e^{\left(\lambda_0-\lambda_1^{(n)}\right)u} du = \frac{1-e^{-\left(\lambda_0-\lambda_1^{(n)}\right)\mathfrak{t}^{(n)}_{t-t(1)}}}{\lambda_0-\lambda_1^{(n)}} \le \mathfrak{t}^{(n)}_{t-t(1)}, \tag{4.3.44}$$

where for the last inequality, we use (4.3.33) applied with $a = \lambda_0 - \lambda_1^{(n)} = 2\alpha_1\mu_1^{(n)} > 0$ and $x = \mathfrak{t}^{(n)}_{t-t(1)}$. It follows that, for sufficiently large n (such that $\mathfrak{t}^{(n)}_{t(1)} + s \ge \sqrt{\log(n)}$), under the event $B^{(n)}_{\widetilde{\varepsilon}} \cap C_{\sqrt{\log(n)},\widetilde{\varepsilon}}$, we have that

$$\begin{aligned} a_t^{(n)} &\le \frac{n^{t(1)}}{\log^{\theta(1)}(n) c_n(t)} \int_{\mathfrak{t}^{(n)}_{t(1)}+s}^{\mathfrak{t}^{(n)}_t+s} 2\alpha_0\mu_0^{(n)}\big(W+2\widetilde{\varepsilon}\big) e^{\left(\lambda_0-\lambda_1^{(n)}\right)u} du \\ &\le 2\alpha_0\big(n^{t(1)}\mu_0^{(n)}\big)(W+2\widetilde{\varepsilon})\Big(\mathbb{1}_{\{\lambda_1<\lambda_0\}}\frac{1}{\lambda_0-\lambda_1} + \mathbb{1}_{\{\lambda_1=\lambda_0\}}\frac{1}{\lambda_0}(t-t(1))\Big), \end{aligned}$$

since $\theta(1) = \mathbb{1}_{\{\lambda_1=\lambda_0\}}$. By definition, we have

$$w_1(t) = 2\alpha_0\mu_0\big(\mathbb{1}_{\{\lambda_1<\lambda_0\}}\frac{1}{\lambda_0-\lambda_1} + \mathbb{1}_{\{\lambda_1=\lambda_0\}}\frac{1}{\lambda_0}(t-t(1))\big).$$

This implies that

$$a_t^{(n)} - Ww_1(t) \le \frac{w_1(t)}{\mu_0} W\big(n^{t(1)}\mu_0^{(n)} - \mu_0\big) + C\widetilde{\varepsilon},$$

where $C > 0$ is a sufficiently large constant.

Introduce the event

$$D^{(n)}_{\widetilde{\varepsilon}} := \big\{ \sup_{t\in[t(1),T]} \big| \frac{w_1(t)}{\mu_0} W\big(n^{t(1)}\mu_0^{(n)} - \mu_0\big)\big| \le \widetilde{\varepsilon}\big\}.$$

This event satisfies $\mathbb{P}\big(D^{(n)}_{\widetilde{\varepsilon}}\big) \underset{n\to\infty}{\to} 1$ because W is finite almost surely, $n^{t(1)}\mu_0^{(n)} \underset{n\to\infty}{\to} \mu_0$ and $w_1(t)$ is bounded from above on $[t(1),T]$. Under $B^{(n)}_{\widetilde{\varepsilon}} \cap C_{\sqrt{\log(n)},\widetilde{\varepsilon}} \cap D^{(n)}_{\widetilde{\varepsilon}}$, we have for all $t \in [t(1),T]$,

$$a_t^{(n)} - Ww_1(t) \le (C+1)\widetilde{\varepsilon}.$$

Similarly, it follows that under $B^{(n)}_{\widetilde{\varepsilon}} \cap C_{\sqrt{\log(n)},\widetilde{\varepsilon}} \cap D^{(n)}_{\widetilde{\varepsilon}}$

$$\sup_{t\in[t(1),T]} |a_t^{(n)} - Ww_1(t)| \le (C+1)\widetilde{\varepsilon}.$$

By choosing $\widetilde{\varepsilon} > 0$ such that $(C+1)\widetilde{\varepsilon} \le \varepsilon$, we deduce that under the event $B^{(n)}_{\widetilde{\varepsilon}} \cap C_{\sqrt{\log(n)},\widetilde{\varepsilon}} \cap D^{(n)}_{\widetilde{\varepsilon}}$,

$$\sup_{t\in[t(1),T]} |a^{(n)}_t - W w_1(t)| \le \varepsilon.$$

This concludes the proof for the case $i = 1$, since $\mathbb{P}\big(B^{(n)}_{\widetilde{\varepsilon}} \cap C_{\sqrt{\log(n)},\widetilde{\varepsilon}} \cap D^{(n)}_{\widetilde{\varepsilon}}\big) \underset{n\to\infty}{\longrightarrow} 1$.

Case $i \ge 2$: Assume Proposition 4.3.4 holds for $i-1$. In particular, we have $\mathbb{P}\big(B^{(n)}_{\widetilde{\varepsilon}}\big) \underset{n\to\infty}{\longrightarrow} 1$, where

$$B^{(n)}_{\widetilde{\varepsilon}} := \Big\{ \sup_{v\in[t(i),T]} \Big| \frac{Z^{(n)}_{i-1}\big(\mathfrak{t}^{(n)}_v + s\big)}{n^{v-t(i-1)} e^{\lambda_0 s} \log^{\theta(i-1)}(n)} - W w_{i-1}(v) \Big| \le \widetilde{\varepsilon} \Big\}.$$

Using the change of variable $u = \mathfrak{t}^{(n)}_v + s$ and the fact that $t(i-1) = t(i) - \ell(i-1)$ we obtain

$$a^{(n)}_t = \int_{t(i)}^t 2\alpha_{i-1}\big(n^{\ell(i-1)}\mu^{(n)}_{i-1}\big) \frac{Z^{(n)}_{i-1}\big(\mathfrak{t}^{(n)}_v + s\big)}{n^{v-t(i-1)} e^{\lambda_0 s} \log^{\theta(i)}(n)} \frac{c_n(v,s)}{c_n(t,s)} \frac{\log(n)}{\lambda_0} dv.$$

Notice that when $\lambda_i < \lambda_0$, we have $\theta(i-1) = \theta(i)$, and when $\lambda_i = \lambda_0$, we have $\theta(i-1) = \theta(i) - 1$. Additionally, we use that $v \mapsto c_n(v,s)$ and w_{i-1} are non-decreasing functions, and we apply similar computations as in (4.3.43) and (4.3.44), replacing the index 1 with i to find, under $B^{(n)}_{\widetilde{\varepsilon}}$, that

$$\begin{aligned} a^{(n)}_t \le{}& 2\alpha_{i-1}\big(n^{\ell(i-1)}\mu^{(n)}_{i-1}\big) \\ &\cdot \Big[\mathbb{1}_{\{\lambda_i<\lambda_0\}} \frac{W w_{i-1}(t) + \widetilde{\varepsilon}}{\lambda_0 - \lambda_i} + \mathbb{1}_{\{\lambda_i=\lambda_0\}} W \frac{1}{\lambda_0} \int_{t(i)}^t \big(w_{i-1}(v) + \widetilde{\varepsilon}\big)\, dv \Big]. \end{aligned}$$

By definition (see (4.2.3) and Remark 4.2.4), we have

$$w_i(t) = 2\alpha_{i-1}\mu_{i-1}\Big(\mathbb{1}_{\{\lambda_i<\lambda_0\}} \frac{w_{i-1}(t)}{\lambda_0 - \lambda_i} + \mathbb{1}_{\{\lambda_i=\lambda_0\}} \frac{1}{\lambda_0} \int_{t(i)}^t w_{i-1}(u) du \Big).$$

Thus, under the event $C^{(n)}_{\widetilde{\varepsilon}} := \Big\{ W|n^{\ell(i-1)}\mu^{(n)}_{i-1} - \mu_{i-1}| \le \widetilde{\varepsilon} \Big\}$, we find that for all $t \le T$

$$\begin{aligned} a^{(n)}_t - W w_i(t) &\le 2\alpha_{i-1} \Big[\mathbb{1}_{\{\lambda_i<\lambda_0\}} \frac{1}{\lambda_0 - \lambda_1} \big(w_{i-1}(T) + \big(n^{\ell(i-1)}\mu^{(n)}_{i-1}\big)\big) \\ &\qquad + \mathbb{1}_{\{\lambda_i=\lambda_0\}} \frac{1}{\lambda_0} \Big(\int_{t(i)}^T w_{i-1}(u) du + T\big(n^{\ell(i-1)}\mu^{(n)}_{i-1}\big)\Big)\Big] \widetilde{\varepsilon} \\ &\le C\widetilde{\varepsilon}, \end{aligned}$$

where C is a positive constant that depends only on the parameters and on T, but is independent of n. Recalling that $n^{\ell(i-1)}\mu^{(n)}_{i-1}$ converges and that W is finite almost surely (see (4.3.6)) we obtain that the event $C^{(n)}_{\widetilde{\varepsilon}}$ satisfies $\mathbb{P}\big(C^{(n)}_{\widetilde{\varepsilon}}\big) \underset{n\to\infty}{\longrightarrow} 1$.

Then, by choosing $\widetilde{\varepsilon} > 0$ such that $C\widetilde{\varepsilon} \le \varepsilon$, we have shown that under $B^{(n)}_{\widetilde{\varepsilon}} \cap C^{(n)}_{\widetilde{\varepsilon}}$,

$$\sup_{t\in[t(i),T]} a^{(n)}_t - W w_i(t) \le \varepsilon.$$

With similar computations, it can also be shown that under $B_{\widetilde{\varepsilon}}^{(n)} \cap C_{\widetilde{\varepsilon}}^{(n)}$,

$$\sup_{t\in[t(i),T]} |a_t^{(n)} - Ww_i(t)| \leq \varepsilon.$$

We conclude the proof by noting that $\mathbb{P}\big(C_{\widetilde{\varepsilon}}^{(n)}\big) \underset{n\to\infty}{\to} 1$ and $\mathbb{P}\big(B_{\widetilde{\varepsilon}}^{(n)}\big) \underset{n\to\infty}{\to} 1$, as established by the induction assumption. □

Given that $\lambda_0 \geq \lambda_i^{(n)}$, the term (4.3.39) satisfies

$$\mathbb{P}\Big(\sup_{t\in[t(i),T]} \frac{Z_i^{(n)}\big(\mathfrak{t}_{t(i)}^{(n)}+s\big)}{\log^{\theta(i)}(n)e^{\left(\lambda_0-\lambda_i^{(n)}\right)\mathfrak{t}_{t-t(i)}^{(n)}}e^{\lambda_0 s}} \geq \varepsilon\Big) = \mathbb{P}\Big(\frac{Z_i^{(n)}\big(\mathfrak{t}_{t(i)}^{(n)}+s\big)}{\log^{\theta(i)}(n)e^{\lambda_0 s}} \geq \varepsilon\Big) \underset{n\to\infty}{\longrightarrow} 0,$$

where the convergence is obtained by applying Lemma 4.3.7 with $\psi_n(i) = \log(n)e^{\lambda_0 s}$. This is valid because assuming the vertex i is neutral implies that $\theta(i) = \theta(i-1)+1$.

This completes the proof of Proposition 4.3.4 (i).We now turn to Proposition 4.3.4 (ii).

2. Deleterious case: Assume that $\lambda_i < \lambda_0$. Let $0 < T_1 < T_2$. Using similar computations as in (4.3.25), for all $t \in [T_1, T_2]$, we have

$$\begin{aligned}\frac{Z_i^{(n)}\big(\mathfrak{t}_{t(i)+t}^{(n)}+s\big)}{n^t\log^{\theta(i)}(n)e^{\lambda_0 s}} &= \frac{M_i^{(n)}\big(\mathfrak{t}_{t(i)+t}^{(n)}+s\big) - M_i^{(n)}\big(\mathfrak{t}_{t(i)}^{(n)}+s\big)}{n^{-\frac{\lambda_i^{(n)}}{\lambda_0}t(i)}e^{\left(\lambda_0-\lambda_i^{(n)}\right)\left(\mathfrak{t}_t^{(n)}+s\right)}\log^{\theta(i)}(n)} \\ &\quad + \frac{\int_{\mathfrak{t}_{t(i)}^{(n)}+s}^{\mathfrak{t}_{t(i)+t}^{(n)}+s} 2\alpha_{i-1}\mu_{i-1}^{(n)}e^{-\lambda_i^{(n)}s}Z_{i-1}^{(n)}(s)ds}{n^{-\frac{\lambda_i^{(n)}}{\lambda_0}t(i)}e^{\left(\lambda_0-\lambda_i^{(n)}\right)\left(\mathfrak{t}_t^{(n)}+s\right)}\log^{\theta(i)}(n)} + \frac{Z_i^{(n)}\big(\mathfrak{t}_{t(i)}^{(n)}+s\big)}{n^{t\frac{\lambda_0-\lambda_i^{(n)}}{\lambda_0}}e^{\lambda_0 s}\log^{\theta(i)}(n)}.\end{aligned}$$

Then, this allows to write

$$\begin{aligned}&\mathbb{P}\Big(\sup_{t\in[T_1,T_2]}\Big|\frac{Z_i^{(n)}\big(\mathfrak{t}_{t(i)+t}^{(n)}+s\big)}{n^t\log^{\theta(i)}(n)e^{\lambda_0 s}} - Ww_i(t(i)+t)\Big| \geq 3\varepsilon\Big) \\ &\quad\leq \mathbb{P}\Big(\sup_{t\in[T_1,T_2]}\Big|\frac{M_i^{(n)}\big(\mathfrak{t}_{t(i)+t}^{(n)}+s\big) - M_i^{(n)}\big(\mathfrak{t}_{t(i)}^{(n)}+s\big)}{n^{-\frac{\lambda_i^{(n)}}{\lambda_0}t(i)}e^{\left(\lambda_0-\lambda_i^{(n)}\right)\left(\mathfrak{t}_t^{(n)}+s\right)}\log^{\theta(i)}(n)}\Big| \geq \varepsilon\Big) \qquad (4.3.45) \\ &\quad+ \mathbb{P}\Big(\sup_{t\in[T_1,T_2]}\Big|\frac{\int_{\mathfrak{t}_{t(i)}^{(n)}+s}^{\mathfrak{t}_{t(i)+t}^{(n)}+s} 2\alpha_{i-1}\mu_{i-1}^{(n)}e^{-\lambda_i^{(n)}s}Z_{i-1}^{(n)}(s)ds}{n^{-\frac{\lambda_i^{(n)}}{\lambda_0}t(i)}n^{t\frac{\lambda_0-\lambda_i^{(n)}}{\lambda_0}}e^{\left(\lambda_0-\lambda_i^{(n)}\right)s}\log^{\theta(i)}(n)} - Ww_i(t(i)+t)\Big| \geq \varepsilon\Big) \qquad (4.3.46) \\ &\quad+ \mathbb{P}\Big(\frac{Z_i^{(n)}\big(\mathfrak{t}_{t(i)}^{(n)}+s\big)}{n^{T_1\frac{\lambda_0-\lambda_i^{(n)}}{\lambda_0}}e^{\lambda_0 s}\log^{\theta(i)}(n)} \geq \varepsilon\Big). \qquad (4.3.47)\end{aligned}$$

For the convergence to 0 of the term (4.3.45), we first use the fact that $\lambda_i^{(n)} \leq \lambda_i < \lambda_0$, to simplify the denominator. Then, we apply the Maximal inequality to the supermartingale

$$\Big(\frac{M_i^{(n)}\big(\mathfrak{t}_{t(i)+t}^{(n)}+s\big)-M_i^{(n)}\big(\mathfrak{t}_{t(i)}^{(n)}+s\big)}{n^{-\frac{\lambda_i}{\lambda_0}t(i)}e^{(\lambda_0-\lambda_i)\left(\mathfrak{t}_t^{(n)}+s\right)}\log^{\theta(i)}(n)}\Big)_{t\geq 0}$$

to obtain

$$\begin{aligned}
&\mathbb{P}\Big(\sup_{t\in[T_1,T_2]}\Big|\frac{M_i^{(n)}\big(\mathfrak{t}_{t(i)+t}^{(n)}+s\big)-M_i^{(n)}\big(\mathfrak{t}_{t(i)}^{(n)}+s\big)}{n^{-\frac{\lambda_i^{(n)}}{\lambda_0}t(i)}e^{\left(\lambda_0-\lambda_i^{(n)}\right)\left(\mathfrak{t}_t^{(n)}+s\right)}\log^{\theta(i)}(n)}\Big|\geq\varepsilon\Big) \qquad (4.3.48)\\
&\quad\leq\mathbb{P}\Big(\sup_{t\in[T_1,T_2]}\Big|\frac{M_i^{(n)}\big(\mathfrak{t}_{t(i)+t}^{(n)}+s\big)-M_i^{(n)}\big(\mathfrak{t}_{t(i)}^{(n)}+s\big)}{n^{-\frac{\lambda_i}{\lambda_0}t(i)}e^{(\lambda_0-\lambda_i)\left(\mathfrak{t}_t^{(n)}+s\right)}\log^{\theta(i)}(n)}\Big|\geq\varepsilon\Big)\\
&\quad\leq\frac{3e^{-(\lambda_0-\lambda_i)s}}{\varepsilon\log^{\theta(i)}(n)}n^{\frac{\lambda_i}{\lambda_0}t(i)}\sup_{t\in[T_1,T_2]}n^{-t\frac{\lambda_0-\lambda_i}{\lambda_0}}\sqrt{\mathbb{E}\Big[\big\langle M_i^{(n)}\big\rangle_{\mathfrak{t}_{t(i)+t}^{(n)}+s}-\big\langle M_i^{(n)}\big\rangle_{\mathfrak{t}_{t(i)}^{(n)}+s}\Big]}.
\end{aligned}$$

Applying Lemma 4.3.10 at the times $t_1^{(n)}=t(i)$ and $t_2=t(i)+t$, we obtain

$$\begin{aligned}
&\sqrt{\mathbb{E}\Big[\big\langle M_i^{(n)}\big\rangle_{\mathfrak{t}_{t(i)+t}^{(n)}+s}-\big\langle M_i^{(n)}\big\rangle_{\mathfrak{t}_{t(i)}^{(n)}+s}\Big]} \qquad (4.3.49)\\
&\quad\leq C\big(\mathfrak{t}_{t(i)+t}^{(n)}+s\big)^{\frac{\theta(i)}{2}}\sqrt{\mu_{\otimes,i}^{(n)}}\cdot\Big[\mathbb{1}_{\{\lambda_0>2\lambda_i\}}n^{\frac{\lambda_0-2\lambda_i}{2\lambda_0}(t(i)+t)}e^{\frac{\lambda_0-2\lambda_i}{2}s}\\
&\qquad+\mathbb{1}_{\{\lambda_0=2\lambda_i\}}\sqrt{\mathfrak{t}_{t(i)+t}^{(n)}+s}+\mathbb{1}_{\{\lambda_i<\lambda_0<2\lambda_i\}}n^{-\frac{2\lambda_i-\lambda_0}{2\lambda_0}t(i)}e^{-\frac{2\lambda_i-\lambda_0}{2}s}\Big].
\end{aligned}$$

For all $t\in[T_1,T_2]$, we perform the following auxiliary computations, which will be used to obtain the result

$$\begin{aligned}
n^{\frac{\lambda_i}{\lambda_0}t(i)}\sqrt{\mu_{\otimes,i}^{(n)}}n^{-t\frac{\lambda_0-\lambda_i}{\lambda_0}}n^{\frac{\lambda_0-2\lambda_i}{2\lambda_0}(t(i)+t)}&\leq\frac{\sqrt{n^{t(i)}\mu_{\otimes,i}^{(n)}}}{n^{\frac{T_1}{2}}}, \qquad (4.3.50)\\
\mathbb{1}_{\{\lambda_0=2\lambda_i\}}n^{\frac{\lambda_i}{\lambda_0}t(i)}\sqrt{\mu_{\otimes,i}^{(n)}}n^{-t\frac{\lambda_0-\lambda_i}{\lambda_0}}&\leq\frac{\sqrt{n^{t(i)}\mu_{\otimes,i}^{(n)}}}{n^{\frac{T_1}{2}}},\\
n^{\frac{\lambda_i}{\lambda_0}t(i)}\sqrt{\mu_{\otimes,i}^{(n)}}n^{-t\frac{\lambda_0-\lambda_i}{\lambda_0}}n^{-\frac{2\lambda_i-\lambda_0}{2\lambda_0}t(i)}&\leq\frac{\sqrt{n^{t(i)}\mu_{\otimes,i}^{(n)}}}{n^{T_1\frac{\lambda_0-\lambda_i}{\lambda_0}}}.
\end{aligned}$$

Then, combining (4.3.48), (4.3.49) and (4.3.50), we obtain

$$\begin{aligned}
&\mathbb{P}\Big(\sup_{t\in[T_1,T_2]}\Big|\frac{M_i^{(n)}\big(\mathfrak{t}_{t(i)+t}^{(n)}+s\big)-M_i^{(n)}\big(\mathfrak{t}_{t(i)}^{(n)}+s\big)}{n^{-\frac{\lambda_i^{(n)}}{\lambda_0}t(i)}e^{\left(\lambda_0-\lambda_i^{(n)}\right)\left(\mathfrak{t}_t^{(n)}+s\right)}\log^{\theta(i)}(n)}\Big|\geq\varepsilon\Big)\\
&\leq\frac{3Ce^{-(\lambda_0-\lambda_i)s}}{\varepsilon\log^{\theta(i)/2}(n)}\Big(\frac{\mathfrak{t}_{t(i)+T_2}^{(n)}+s}{\log(n)}\Big)^{\frac{\theta(i)}{2}}\sup_{t\in[T_1,T_2]}n^{\frac{\lambda_i}{\lambda_0}t(i)}n^{-t\frac{\lambda_0-\lambda_i}{\lambda_0}}\sqrt{\mu_{\otimes,i}^{(n)}}\Big[\mathbb{1}_{\{\lambda_0>2\lambda_i\}}n^{\frac{\lambda_0-2\lambda_i}{2\lambda_0}(t(i)+t)}
\end{aligned}$$

$$\cdot e^{\frac{\lambda_0-2\lambda_i}{2}s} + \mathbb{1}_{\{\lambda_0=2\lambda_i\}}\sqrt{\mathfrak{t}^{(n)}_{t(i)+t}+s} + \mathbb{1}_{\{\lambda_i<\lambda_0<2\lambda_i\}} n^{-\frac{2\lambda_i-\lambda_0}{2\lambda_0}} e^{-\frac{2\lambda_i-\lambda_0}{2}s}\Big]$$

$$= \frac{3Ce^{-(\lambda_0-\lambda_i)s}}{\varepsilon \log^{\theta(i)/2}(n)}\Big(\frac{\mathfrak{t}^{(n)}_{t(i)+T_2}+s}{\log(n)}\Big)^{\frac{\theta(i)}{2}}\sqrt{n^{t(i)}\mu^{(n)}_{\otimes,i}}$$

$$\cdot\Big(\mathbb{1}_{\{\lambda_0>2\lambda_i\}}\frac{e^{\frac{\lambda_0-2\lambda_i}{2}s}}{n^{\frac{T_1}{2}}} + \mathbb{1}_{\{\lambda_0=2\lambda_i\}}\frac{\sqrt{\mathfrak{t}^{(n)}_{t(i)+T_2}+s}}{n^{\frac{T_1}{2}}} + \mathbb{1}_{\{\lambda_i<\lambda_0<2\lambda_i\}}\frac{e^{-\frac{2\lambda_i-\lambda_0}{2}s}}{n^{T_1\frac{\lambda_0-\lambda_i}{\lambda_0}}}\Big)$$

$$\underset{n\to\infty}{\longrightarrow} 0.$$

The term (4.3.46) converges to 0 by Lemma 4.3.11. The convergence to 0 for the term (4.3.47) is obtained by applying Lemma 4.3.7 with $\psi_n(i) = n^{T_1\frac{\lambda_0-\lambda_i}{\lambda_0}}e^{\lambda_0 s}$. This completes the proof of Proposition 4.3.4. □

Uniform control on the parameter s

In this subsection, we will prove (4.2.5) and (4.2.6) for the mono-directional graph, as stated in Proposition 4.3.4, using an approach inspired by [56, Lemma 3]. Define $u^{(n)}_s := t + \frac{s-M}{\log(n)}\lambda_0$ such that $\mathfrak{t}^{(n)}_t + s = \mathfrak{t}^{(n)}_{u^{(n)}_s} + M$. Notice that

$$0 \le t - u^{(n)}_s \le \frac{2M}{\log(n)}\lambda_0. \tag{4.3.51}$$

1. **Deleterious case:** We begin by showing (4.2.6). We will use that

$$n^t \log^{\theta(i)}(n)e^{\lambda_0 s} = n^{u^{(n)}_s}\log^{\theta(i)}(n)e^{\lambda_0 M}.$$

This gives that

$$\Big|\frac{Z^{(n)}_i\big(\mathfrak{t}^{(n)}_{t(i)+t}+s\big)}{n^t\log^{\theta(i)}(n)e^{\lambda_0 s}} - Ww_i(t(i)+t)\Big| \le \Big|\frac{Z^{(n)}_i\big(\mathfrak{t}^{(n)}_{t(i)+u^{(n)}_s}+M\big)}{n^{u^{(n)}_s}\log^{\theta(i)}(n)e^{\lambda_0 M}} - Ww_i\big(t(i)+u^{(n)}_s\big)\Big| + W\Big|w_i(t(i)+t) - w_i\big(t(i)+u^{(n)}_s\big)\Big|.$$

Since $w_i(t(i)+\cdot)$ is a polynomial function, there exists a constant $C_i > 0$ such that, for all $t \le T_2$ and $s \in [-M, M]$, we have

$$\Big|w_i(t(i)+t) - w_i\big(t(i)+u^{(n)}_s\big)\Big| \le \frac{C_i}{\log(n)}, \tag{4.3.52}$$

due to (4.3.51). Let $0 < \widetilde{T}_1 < T_1$. For n sufficiently large such that $u^{(n)}_s \ge \widetilde{T}_1$ for all $(t,s) \in [T_1, T_2]\times[-M, M]$, we have

$$\Big|\frac{Z^{(n)}_i\big(\mathfrak{t}^{(n)}_{t(i)+t}+s\big)}{n^t\log^{\theta(i)}(n)e^{\lambda_0 s}} - Ww_i(t(i)+t)\Big| \le \sup_{x\in[\widetilde{T}_1,T_2]}\Big|\frac{Z^{(n)}_i\big(\mathfrak{t}^{(n)}_{t(i)+x}+M\big)}{n^x\log^{\theta(i)}(n)e^{\lambda_0 M}} - Ww_i(t(i)+x)\Big| + W\frac{C_i}{\log(n)}.$$

Thus, for sufficiently large n, we have

$$\mathbb{P}\Big(\sup_{s\in[-M,M]}\sup_{t\in[T_1,T_2]}\Big|\frac{Z_i^{(n)}\big(\mathfrak{t}^{(n)}_{t(i)+t}+s\big)}{n^t\log^{\theta(i)}(n)e^{\lambda_0 s}}-Ww_i(t(i)+t)\Big|\geq 2\varepsilon\Big)$$
$$\leq \mathbb{P}\Big(\sup_{x\in[\widetilde{T}_1,T_2]}\Big|\frac{Z_i^{(n)}\big(\mathfrak{t}^{(n)}_{t(i)+x}+M\big)}{n^x\log^{\theta(i)}(n)e^{\lambda_0 M}}-Ww_i(t(i)+x)\Big|\geq\varepsilon\Big)+\mathbb{P}\Big(W\geq\frac{\varepsilon\log(n)}{C_i}\Big),$$

from which (4.2.6) is obtained. Indeed, the first term on the right-hand side converges to 0 according to Proposition 4.3.4 (ii) and the second term converges to 0 since W is finite almost surely (see (4.3.6)).

2. Neutral case: Now, we show (4.2.5). We have

$$\begin{aligned}\Big|\frac{Z_i^{(n)}\big(\mathfrak{t}^{(n)}_t+s\big)}{d_i^{(n)}(t,s)}-Ww_i(t)\Big|&\leq \mathbb{1}_{\{t\in[0,t(i)-h_n^{-1}(i))\}}Z_i^{(n)}\big(\mathfrak{t}^{(n)}_{u_s^{(n)}}+M\big)+\mathbb{1}_{\{t\in[t(i)-h_n^{-1}(i),t(i))\}}\\ &\cdot\frac{Z_i^{(n)}\big(\mathfrak{t}^{(n)}_{u_s^{(n)}}+M\big)}{\psi_n(i)\log^{\theta(i-1)}(n)}+\mathbb{1}_{\{t\geq t(i)\}}\mathbb{1}_{\{u_s^{(n)}<t(i)\}}\Big|\frac{Z_i^{(n)}\big(\mathfrak{t}^{(n)}_{u_s^{(n)}}+M\big)}{\log^{\theta(i)}(n)e^{\lambda_0 M}}e^{\left(t(i)-u_s^{(n)}\right)\log(n)}-Ww_i(t)\Big|\\ &+\mathbb{1}_{\{u_s^{(n)}\geq t(i)\}}\Big|\frac{Z_i^{(n)}\big(\mathfrak{t}^{(n)}_{u_s^{(n)}}+M\big)}{n^{u_s^{(n)}-t(i)}\log^{\theta(i)}(n)e^{\lambda_0 M}}-Ww_i\big(u_s^{(n)}\big)\Big|\\ &+\mathbb{1}_{\{u_s^{(n)}\geq t(i)\}}W\Big|w_i(t)-w_i\big(u_s^{(n)}\big)\Big|.\end{aligned}\tag{4.3.53}$$

As in (4.3.52), there exists a constant C_i such that, for all $(t,s)\in[0,T]\times[-M,M]$, we have

$$\Big|w_i(t)-w_i\big(u_s^{(n)}\big)\Big|\leq\frac{C_i}{\log(n)}.$$

In the case where $t\geq t(i)$ and $u_s^{(n)}<t(i)$, we have that $t(i)-u_s^{(n)}\leq\frac{2M}{\log(n)}\lambda_0$, which, in particular, implies that $e^{\left(t(i)-u_s^{(n)}\right)\log(n)}\leq e^{2M\lambda_0}$. Moreover, since $w_i\big(u_s^{(n)}\big)=0$ (because $w_i(s)=0$ for all $s\in[0,t(i)]$), it follows from the previous inequality that $w_i(t)\leq\frac{C_i}{\log(n)}$. Combining these arguments, we obtain

$$\begin{aligned}\Big|\frac{Z_i^{(n)}\big(\mathfrak{t}^{(n)}_{u_s^{(n)}}+M\big)}{\log^{\theta(i)}(n)e^{\lambda_0 M}}e^{\left(t(i)-u_s^{(n)}\right)\log(n)}-Ww_i(t)\Big|&\leq\frac{Z_i^{(n)}\big(\mathfrak{t}^{(n)}_{u_s^{(n)}}+M\big)}{\log^{\theta(i)}(n)}e^{\lambda_0 M}\\ &+W\frac{C_i}{\log(n)}.\end{aligned}\tag{4.3.54}$$

Finally, using (4.3.53) and (4.3.54), we obtain for all $(t,s) \in [0,T] \times [-M,M]$

$$\begin{aligned}\left|\frac{Z_i^{(n)}\left(\mathfrak{t}_t^{(n)}+s\right)}{d_i^{(n)}(t,s)} - W w_i(t)\right| \leq & \sup_{x\in\left[0,t(i)-h_n^{-1}(i)\right)} Z_i^{(n)}\left(\mathfrak{t}_x^{(n)}+M\right) \\ & + \sup_{x\in[0,t(i)]} \frac{Z_i^{(n)}\left(\mathfrak{t}_x^{(n)}+M\right)}{\psi_n(i)\log^{\theta(i-1)}(n)} \\ & + \sup_{x\in[0,t(i)]} \frac{Z_i^{(n)}\left(\mathfrak{t}_x^{(n)}+M\right)}{\log^{\theta(i)}(n)} e^{\lambda_0 M} + W\frac{2C_i}{\log(n)} \\ & + \sup_{x\in[t(i),T]} \left|\frac{Z_i^{(n)}\left(\mathfrak{t}_x^{(n)}+M\right)}{n^{x-t(i)}\log^{\theta(i)}(n)e^{\lambda_0 M}} - W w_i(x)\right|.\end{aligned}$$

Then we have

$$\begin{aligned}&\mathbb{P}\Big(\sup_{s\in[-M,M]}\sup_{t\in[0,T]}\left|\frac{Z_i^{(n)}\left(\mathfrak{t}_t^{(n)}+s\right)}{d_i^{(n)}(t,s)} - W w_i(t)\right| \geq 5\varepsilon\Big) \\ &\quad\leq \mathbb{P}\Big(\sup_{x\in\left[0,t(i)-\gamma_n^{-1}(i)\right)} Z_i^{(n)}\left(\mathfrak{t}_x^{(n)}+M\right) \geq \varepsilon\Big) + \mathbb{P}\Big(\sup_{x\in[0,t(i)]}\frac{Z_i^{(n)}\left(\mathfrak{t}_x^{(n)}+M\right)}{\psi_n(i)\log^{\theta(i-1)}(n)} \geq \varepsilon\Big) \quad (4.3.55) \\ &\quad+ \mathbb{P}\Big(\sup_{x\in[0,t(i)]}\left|\frac{Z_i^{(n)}\left(\mathfrak{t}_x^{(n)}+M\right)}{e^{-\lambda_0 M}\log^{\theta(i)}(n)}\right| \geq \varepsilon\Big) + \mathbb{P}\Big(W \geq \frac{\varepsilon\log(n)}{2C_i}\Big) \quad (4.3.56) \\ &\quad+ \mathbb{P}\Big(\sup_{x\in[t(i),T]}\left|\frac{Z_i^{(n)}\left(\mathfrak{t}_x^{(n)}+M\right)}{n^{x-t(i)}\log^{\theta(i)}(n)e^{\lambda_0 M}} - W w_i(x)\right| \geq \varepsilon\Big) \quad (4.3.57) \\ &\underset{n\to\infty}{\longrightarrow} 0,\end{aligned}$$

where the different convergences to 0 are obtained as follows:

- Lemma 4.3.6 gives the convergence of the first term in (4.3.55),
- Lemma 4.3.7 gives the convergence of the second term in (4.3.55) and the first term in (4.3.56); for the latter, we apply Lemma 4.3.7 with $\psi_n(i) = e^{-\lambda_0 M}\log(n)$, which is valid because $\theta(i) = \theta(i-1)+1$,
- for the second term in (4.3.56), we use the fact that W is finite almost surely, see (4.3.6),
- Step 3 of the neutral case in the proof of Proposition 4.3.4 directly establishes the convergence of (4.3.57).

Finally, we have proven Equations (4.2.5) and (4.2.6) in the specific case of the infinite mono-directional graph.

4.3.3 First-order asymptotics of the size of the mutant subpopulations on the random time scale (Theorem 4.2.7 (ii))

In this subsection, we will first show that the random time at which the total population reaches the size n^t behaves asymptotically as the random time at which the wild-type subpopulation reaches

the size n^t. This result is obtained uniformly on the time parameter t, conditioned on $\{W > 0\}$, and in probability. Intuitively, for any mutant trait $i \in \mathbb{N}$, the corresponding subpopulation grows exponentially at rate λ_0 after time $t(i)$, see Proposition 4.3.4. Due to these time delays (on the $\log(n)$-accelerated time scale), the total mutant subpopulation remains consistently negligible compared to the wild-type subpopulation. Consequently the difference between $\eta_t^{(n)}$ and $\sigma_t^{(n)}$ converges to 0.

Proposition 4.3.12. *Assume Equation (4.3.1) holds. Then, for all $\varepsilon > 0$ and $0 < T_1 < T_2$, we have*

$$\mathbb{P}\Big(\sup_{t\in[T_1,T_2]} \big(\eta_t^{(n)} - \sigma_t^{(n)}\big) \leq \varepsilon \Big| W > 0 \Big) \underset{n\to\infty}{\longrightarrow} 1.$$

Proof. The proof will be carried out in two steps. We begin by establishing the result under a stronger condition.

Step 1: In this step, we will show that for all $0 < \delta_1 < \delta_2$ and $\varepsilon > 0$ we have

$$\mathbb{P}\Big(\sup_{t\in[T_1,T_2]} \big(\eta_t^{(n)} - \sigma_t^{(n)}\big) \geq \varepsilon \Big| \delta_1 < W < \delta_2 \Big) \underset{n\to\infty}{\longrightarrow} 0. \tag{4.3.58}$$

Let $0 < \delta_1 < \delta_2$. Then there exists $M \in \mathbb{R}^+$ such that

$$\mathbb{P}\Big(\Big| \frac{\log(W)}{\lambda_0} \Big| \leq M \Big| \delta_1 < W < \delta_2 \Big) = 1. \tag{4.3.59}$$

For all $\varepsilon > 0$ introduce the event $A_\varepsilon^{(n)} := \big\{ \sup_{t\in[T_1,T_2]} \big(\eta_t^{(n)} - \sigma_t^{(n)}\big) \geq \varepsilon \big\}$. Assume that there exists $\varepsilon > 0$ such that the sequence $\big(\mathbb{P}\big(A_\varepsilon^{(n)} \big| \delta_1 < W < \delta_2\big)\big)_{n\in\mathbb{N}}$ does not converge to 0. This means that there exists $\eta > 0$ for which there is an infinite subset $N \in \mathbb{N}$ such that for all $n \in N$, $\mathbb{P}\big(A_\varepsilon^{(n)} \big| \delta_1 < W < \delta_2\big) \geq \eta$. For all $\widetilde{\varepsilon} > 0$ introduce the event

$$B_{\widetilde{\varepsilon}}^{(n)} := \Big\{ \sup_{t\in[T_1,T_2]} \Big| \eta_t^{(n)} - \Big(\mathfrak{t}_t^{(n)} - \frac{\log(W)}{\lambda_0} \Big) \Big| \leq \widetilde{\varepsilon} \Big\},$$

which satisfies $\mathbb{P}\big(B_{\widetilde{\varepsilon}}^{(n)} \big| \delta_1 < W < \delta_2\big) \underset{n\to\infty}{\to} 1$, according to Lemma 4.3.2. From this fact, and since $\sigma_t^{(n)} \leq \eta_t^{(n)}$ for all $t > 0$ almost surely, it follows that under $B_{\widetilde{\varepsilon}}^{(n)}$, we have $\sigma_t^{(n)} < \infty$ for all $t \in [T_1, T_2]$. Moreover, it also follows that under $B_{\widetilde{\varepsilon}}^{(n)}$, we have $Z_0^{(n)}\big(\eta_t^{(n)}\big) = n^t$ for all $t \in [T_1, T_2]$. In particular, under $A_\varepsilon^{(n)}$, there exists $t_n \in [T_1, T_2]$ such that $\eta_{t_n}^{(n)} - \sigma_{t_n}^{(n)} \geq \varepsilon$, which implies that $Z_0^{(n)}\big(\sigma_{t_n}^{(n)}\big) \leq n^{t_n} e^{-\lambda_0 \frac{\varepsilon}{2}}$. Otherwise, by applying the strong Markov property, it would lead to a contradiction with $A_\varepsilon^{(n)}$. Combining these reasonings, it follows that under $A_\varepsilon^{(n)} \cap B_{\widetilde{\varepsilon}}^{(n)}$, we have that

$$\sum_{i\geq 1} Z_i^{(n)}\big(\sigma_{t_n}^{(n)}\big) = Z_{tot}^{(n)}\big(\sigma_{t_n}^{(n)}\big) - Z_0^{(n)}\big(\sigma_{t_n}^{(n)}\big) \geq n^{t_n}\big(1 - e^{-\lambda_0 \frac{\varepsilon}{2}}\big) = \Omega\left(n^{t_n}\right), \tag{4.3.60}$$

where we use the standard Landau notation for Ω. However, the result regarding the mutant subpopulations indicates that, due to the power law mutation rates regime, the mutant subpopulations

have a negligible size compared to the wild-type subpopulation. More precisely, under the event $A_\varepsilon^{(n)} \cap B_{\widetilde{\varepsilon}}^{(n)}$, using (4.3.59) and Proposition 4.3.4, we have

$$\begin{aligned}\sum_{i\geq 1} Z_i^{(n)}\big(\sigma_{t_n}^{(n)}\big) &\leq \sup_{u\in[T_1,t_n]} \sum_{i\geq 1} Z_i^{(n)}\big(\eta_u^{(n)}\big) \\ &\leq \sup_{u\in[T_1,t_n]} \sup_{s\in[-(M+\widetilde{\varepsilon}),M+\widetilde{\varepsilon}]} \sum_{i\geq 1} Z_i^{(n)}\big(\mathfrak{t}_u^{(n)}+s\big) \\ &= o(n^{t_n}).\end{aligned} \tag{4.3.61}$$

There is a contradiction between (4.3.60) and (4.3.61), so we have proven (4.3.58) for all $\varepsilon > 0$ and $0 < \delta_1 < \delta_2$.

Step 2: Using a similar method as in Step 2 of the proof of Lemma 4.3.2, one can show that for all $\varepsilon > 0$

$$\mathbb{P}\Big(A_\varepsilon^{(n)}\Big|W>0\Big) \underset{n\to\infty}{\longrightarrow} 0,$$

which concludes the proof. □

In the remainder of this subsection, we will prove the following proposition.

Proposition 4.3.13. *Assume Equation (4.3.1) holds. Let $0 < T_1 < T_2$, $M > 0$ and $\varepsilon > 0$. Consider $\big(\rho_t^{(n)}\big)_{t\in\mathbb{R}^+}$ as defined in (4.1.2). Then, we have*

- *If $\lambda_i = \lambda_0$*

$$\mathbb{P}\Big(\sup_{s\in[-M,M]}\sup_{t\in[T_1,T_2]}\Big|\frac{Z_i^{(n)}\big(\rho_t^{(n)}+s\big)}{d_i^{(n)}(t,s)} - \mathbb{1}_{\{W>0\}}w_i(t)\Big| \geq \varepsilon\Big) \underset{n\to\infty}{\longrightarrow} 0.$$

- *If $\lambda_i < \lambda_0$*

$$\mathbb{P}\Big(\sup_{s\in[-M,M]}\sup_{t\in[T_1,T_2]}\Big|\frac{Z_i^{(n)}\big(\rho_{t(i)+t}^{(n)}+s\big)}{n^t\log^{\theta(i)}(n)e^{\lambda_0 s}} - \mathbb{1}_{\{W>0\}}w_i(t(i)+t)\Big| \geq \varepsilon\Big) \underset{n\to\infty}{\longrightarrow} 0.$$

These results correspond to (4.2.7) and (4.2.8) for the mono-directional graph. The proof will be carried out under the assumption that $\lambda_i = \lambda_0$. The case where $\lambda_i < \lambda_0$ can be addressed using similar reasoning and is left to the reader.

Proof of Proposition 4.3.13. Estimate the quantity of interest from above as

$$\begin{aligned}&\mathbb{P}\Big(\sup_{s\in[-M,M]}\sup_{t\in[T_1,T_2]}\Big|\frac{Z_i^{(n)}\big(\rho_t^{(n)}+s\big)}{d_i^{(n)}(t,s)} - w_i(t)\mathbb{1}_{\{W>0\}}\Big| \geq \varepsilon\Big) \\ &\quad\leq \mathbb{P}\Big(\{W>0\}\cap\Big\{\sup_{s\in[-M,M]}\sup_{t\in[T_1,T_2]}\Big|\frac{Z_i^{(n)}\big(\rho_t^{(n)}+s\big)}{d_i^{(n)}(t,s)} - w_i(t)\Big| \geq \varepsilon\Big\}\Big) \qquad (4.3.62) \\ &\quad+ \mathbb{P}\Big(\{W=0\}\cap\Big\{K_0^{(n)}\big(\rho_{T_2}^{(n)}+M\big)\geq 1\Big\}\cup\Big\{H_0^{(n)}\big(\rho_{T_2}^{(n)}+M\big)\geq 1\Big\}\Big), \qquad (4.3.63)\end{aligned}$$

where, for the term in (4.3.63), we use the fact that a necessary condition for the mutant subpopulation of trait i to be strictly positive is that at least one mutational event from the wild-type subpopulation must have occurred before.

Step 1: The convergence to 0 of (4.3.63) follows from proving that

$$\mathbb{P}\Big(\Big\{\sup_{t\in\mathbb{R}^+} K_0^{(n)}(t)=0\Big\}\cap\Big\{\sup_{t\in\mathbb{R}^+} H_0^{(n)}(t)=0\Big\}\Big|W=0\Big)\underset{n\to\infty}{\longrightarrow}1.$$

Let us first show that $\mathbb{P}\big(\sup_{t\in\mathbb{R}^+} K_0^{(n)}(t)\geq 1\big|W=0\big)\underset{n\to\infty}{\to}0$. Notice that, almost surely, for all $t\in\mathbb{R}^+$

$$K_0^{(n)}(t)\leq \widetilde{K}^{(n)}(t):=\int_0^t\int_{\mathbb{R}^+}\mathbb{1}_{\left\{\theta\leq 2\alpha_0\mu_0^{(n)}Z_0(s^-)\right\}}N_0(ds,d\theta),$$

because, almost surely, for all $t\in\mathbb{R}^+$, we have $Z_0^{(n)}(t)\leq Z_0(t)$. Then it follows that

$$\mathbb{P}\Big(\sup_{t\in\mathbb{R}^+} K_0^{(n)}(t)\geq 1\Big|W=0\Big)\leq\mathbb{E}\Big[\sup_{t\in\mathbb{R}^+}\widetilde{K}^{(n)}(t)\wedge 1\Big|W=0\Big]\underset{n\to\infty}{\longrightarrow}0$$

by dominated convergence. Indeed, for all $\omega\in\{W=0\}$, there exists $T(\omega)\in\mathbb{R}^+$ such that for all $t\geq T(\omega), Z_0(t)=0$. Combined with $\mu_0^{(n)}\underset{n\to\infty}{\to}0$, it follows that there exists $N(\omega)\in\mathbb{N}$ such that for all $n\geq N(\omega)$, we have $\sup_{t\in\mathbb{R}^+}\widetilde{K}^{(n)}(t)=0$. We conclude the proof of Step 1 by showing that $\mathbb{P}\big(\sup_{t\in\mathbb{R}^+} H_0^{(n)}(t)\geq 1\big|W=0\big)\underset{n\to\infty}{\longrightarrow}0$ using similar reasoning.

Step 2: We will show that (4.3.62) converges to 0 in three steps.

Step 2) (i): We begin by showing that for all $\varepsilon>0$ and $\eta>0$ we have

$$\mathbb{P}\Big(\sup_{s\in[-M,M]}\sup_{t\in[T_1,T_2]}\Big|\frac{Z_i^{(n)}\big(\rho_t^{(n)}+s\big)}{d_i^{(n)}(t,s)}e^{-\lambda_0\left[\rho_t^{(n)}-\mathfrak{t}_t^{(n)}\right]}-Ww_i(t)\Big|\geq\varepsilon\Big|W>\eta\Big)\underset{n\to\infty}{\longrightarrow}0.$$

We have

$$\begin{aligned}
&\mathbb{P}\Big(\sup_{t\in[T_1,T_2]}\Big|\rho_t^{(n)}-\mathfrak{t}_t^{(n)}\Big|\geq M\Big|W>\eta\Big)\\
&\quad\leq\mathbb{P}\Big(\sup_{t\in[T_1,T_2]}\Big|\eta_t^{(n)}-\Big(\mathfrak{t}_t^{(n)}-\frac{\log(W)}{\lambda_0}\Big)\Big|\geq\frac{M}{3}\Big|W>\eta\Big)\\
&\quad+\mathbb{P}\Big(\sup_{t\in[T_1,T_2]}\Big|\rho_t^{(n)}-\eta_t^{(n)}\Big|\geq\frac{M}{3}\Big)+\mathbb{P}\Big(\frac{|\log(W)|}{\lambda_0}\geq\frac{M}{3}\Big|W>\eta\Big).
\end{aligned}$$

Let $\delta>0$. Using Lemma 4.3.2, Proposition 4.3.12 and the distribution of W given in (4.3.6), there exist $M>0$ and $N_1\in\mathbb{N}$ such that for all $n\geq N_1$,

$$\mathbb{P}\Big(\sup_{t\in[T_1,T_2]}\Big|\rho_t^{(n)}-\mathfrak{t}_t^{(n)}\Big|\geq M\Big|W>\eta\Big)\leq\frac{\delta}{2}.\tag{4.3.64}$$

Now, we can apply Theorem 4.2.7 (i) Eq. (4.2.5) to get that there exists $N_2\in\mathbb{N}$ such that for all $n\geq N_2$,

$$\mathbb{P}\Big(\sup_{s\in[-M,M]}\sup_{s_1\in[-M,M]}\sup_{t\in[T_1,T_2]}\Big|\frac{Z_i^{(n)}\big(\mathfrak{t}_t^{(n)}+s+s_1\big)}{d_i^{(n)}(t,s+s_1)}-Ww_i(t)\Big|\geq\varepsilon\Big)\leq\frac{\delta}{2}.\tag{4.3.65}$$

Consequently, using Equations (4.3.64) and (4.3.65), we have shown that for all $\delta > 0$, there exists $N := \max(N_1, N_2) \in \mathbb{N}$ such that for all $n \geq N$,

$$\mathbb{P}\Big(\sup_{s\in[-M,M]} \sup_{t\in[T_1,T_2]} \Big| \frac{Z_i^{(n)}\big(\rho_t^{(n)}+s\big)}{d_i^{(n)}(t,s)} e^{-\lambda_0\left[\rho_t^{(n)}-\mathfrak{t}_t^{(n)}\right]} - W w_i(t) \Big| \geq \varepsilon \Big| W > \eta \Big) \leq \delta,$$

which concludes Step 2) (i).

Step 2) (ii): Now, we are going to prove that

$$\mathbb{P}\Big(\sup_{s\in[-M,M]} \sup_{t\in[T_1,T_2]} \Big| \frac{Z_i^{(n)}\big(\rho_t^{(n)}+s\big)}{d_i^{(n)}(t,s)} - w_i(t) \Big| \geq \varepsilon \Big| W > \eta \Big) \underset{n\to\infty}{\longrightarrow} 0.$$

Let $\delta > 0$ and $0 < \widetilde{\varepsilon} < \eta$. According to Remark 4.3.3 Equation (4.3.13) and Proposition 4.3.12, we have

$$\mathbb{P}\left(A_{\widetilde{\varepsilon}}^{(n)} \middle| W > \eta \right) \geq 1 - \frac{\delta}{2}, \text{ where } A_{\widetilde{\varepsilon}}^{(n)} := \Big\{ \sup_{t\in[T_1,T_2]} \Big| e^{-\lambda_0\left(\rho_t^{(n)}-\mathfrak{t}_t^{(n)}\right)} - W \Big| \leq \widetilde{\varepsilon} \Big\}.$$

Combined with Step 2) (i), there exists $N \in \mathbb{N}$ such that for all $n \geq N$, we have $\mathbb{P}\big(A_{\widetilde{\varepsilon}}^{(n)} \cap B_{\widetilde{\varepsilon}}^{(n)} | W > \eta\big) \geq 1 - \delta$, where

$$B_{\widetilde{\varepsilon}}^{(n)} := \Big\{ \sup_{s\in[-M,M]} \sup_{t\in[T_1,T_2]} \Big| \frac{Z_i^{(n)}\big(\rho_t^{(n)}+s\big)}{d_i^{(n)}(t,s)} e^{-\lambda_0\left[\rho_t^{(n)}-\mathfrak{t}_t^{(n)}\right]} - W w_i(t) \Big| \leq \widetilde{\varepsilon} \Big\}.$$

In particular, conditioned on $\{W > \eta\}$, under the event $A_{\widetilde{\varepsilon}}^{(n)} \cap B_{\widetilde{\varepsilon}}^{(n)}$, we have that for all $t \in [T_1, T_2]$ and for all $s \in [-M, M]$,

$$\begin{aligned}
\frac{Z_i^{(n)}\big(\rho_t^{(n)}+s\big)}{d_i^{(n)}(t,s)} - w_i(t) &\leq \left(\widetilde{\varepsilon} + w_i(t) W\right) e^{\lambda_0\left(\rho_t^{(n)}-\mathfrak{t}_t^{(n)}\right)} - w_i(t) \\
&\leq \frac{\widetilde{\varepsilon}}{W - \widetilde{\varepsilon}} + w_i(t)\Big(\frac{W}{W-\widetilde{\varepsilon}} - 1\Big) \\
&\leq \left(1 + w_i(T_2)\right) \frac{\widetilde{\varepsilon}}{\eta - \widetilde{\varepsilon}} \\
&\underset{\widetilde{\varepsilon}\to 0}{\longrightarrow} 0,
\end{aligned}$$

so that we can choose $\widetilde{\varepsilon}$ arbitrarily small such that this upper bound is smaller than ε. By applying a similar approach for the lower bound, we find that, conditioned on $\{W > \eta\}$, under the event $A_{\widetilde{\varepsilon}}^{(n)} \cap B_{\widetilde{\varepsilon}}^{(n)}$, we have that for all $t \in [T_1, T_2]$ and for all $s \in [-M, M]$,

$$\frac{Z_i^{(n)}\big(\rho_t^{(n)}+s\big)}{d_i^{(n)}(t,s)} - w_i(t) \geq -\left(1 + w_i(T_2)\right) \frac{\tilde{\varepsilon}}{\eta - \tilde{\varepsilon}} \underset{\tilde{\varepsilon}\to 0}{\longrightarrow} 0.$$

Consequently, by choosing an appropriate $\widetilde{\varepsilon} > 0$, we have shown that there exists $N \in \mathbb{N}$ such that for all $n \geq N$,

$$\mathbb{P}\Big(\sup_{s\in[-M,M]} \sup_{t\in[T_1,T_2]} \Big| \frac{Z_i^{(n)}\big(\rho_t^{(n)}+s\big)}{d_i^{(n)}(t,s)} - w_i(t) \Big| \leq \varepsilon \Big| W > \eta \Big) \geq 1 - \delta.$$

Step 2) (iii): Introduce the notation $C_\varepsilon^{(n)} := \Big\{ \sup_{s\in[-M,M]} \sup_{t\in[T_1,T_2]} \Big| \frac{Z_i^{(n)}\left(\rho_t^{(n)}+s\right)}{d_i^{(n)}(t,s)} - w_i(t) \Big| \geq \varepsilon \Big\}$. To complete the proof of Step 2, we will show that $\mathbb{P}\big(C_\varepsilon^{(n)} \cap \{W>0\}\big) \underset{n\to\infty}{\to} 0$. We have

$$\mathbb{P}\left(C_\varepsilon^{(n)} \cap \{W>0\}\right) \leq \mathbb{P}\left(C_\varepsilon^{(n)} \cap \{W>\eta\}\right) + \mathbb{P}\left(0<W<\eta\right).$$

Using Step 2) (ii), we obtain

$$\limsup_{n\to\infty} \mathbb{P}\left(C_\varepsilon^{(n)} \cap \{W>0\}\right) \leq \mathbb{P}\left(0<W<\eta\right).$$

Taking the limit as $\eta \underset{n\to\infty}{\to} 0$ completes the proof. □

4.4 First-order asymptotics of the size of the mutant subpopulations for a general finite trait space (Theorem 4.2.7)

As in Section 4.3 the sequence $\left(Z_v^{(n)}, v\in V\right)_{n\in\mathbb{N}}$ is mathematically constructed using independent PPMs. In this construction, each population of trait v is decomposed as the sum of subpopulations indexed by the walks in the graph that start from trait 0 and lead to trait v. An exact definition will be given below. The idea is to apply the reasoning of Section 4.3 to each walk γ from trait 0 to trait v, which will provide the first-order asymptotics for the subpopulation of cells of trait v indexed by γ. By comparing the order of distinct walks, we can then conclude the first-order asymptotics of the size of the mutant subpopulation of trait v. However, this reasoning holds only if there are finitely many walks from trait 0 to trait v. In particular, notice that due to cycles, there may be countably infinitely many walks from trait 0 to trait v. Consequently, the proof requires additional steps, introducing an equivalence relation on the walks. We argue that there are only finitely many equivalent classes, and for each equivalent class, the result follows by adapting the reasoning from Section 4.3. For the equivalent class with infinitely many walks, we show that, with high probability, most of these walks do not asymptotically contribute.

Among wild-type individuals, we define the *primary cell population*, denoted by $\left(Z_{(0)}^{(n)}(t)\right)_{t\geq 0}$, as the set of all cells that have no mutants in their ancestry, tracing back to the initial cell. This corresponds to $Z_0^{(n)}$ in the case of the mono-directional graph.

Definition 4.4.1 (Walks and neighbours)**.** *Define the set of all walks in the graph V starting from trait 0 as $\Gamma(V)$. For a trait $v\in V$, the set of traits to which a cell of trait v may mutate is defined as $N(v) := \{u\in V : (v,u)\in E\}$. For a walk $\gamma = (0,\cdots,\gamma(k)) \in \Gamma(V)$, denote the last trait $\gamma(k)$ visited by γ as $\gamma_{end} := \gamma(k)$, and the sub-walk that does not include this last trait as $\overleftarrow{\gamma} := (0,\cdots,\gamma(k-1))$. Introduce the sets of tuples of the walks in V starting from trait 0, associated with one or two neighbours of the last trait of γ, as*

$$N_\Gamma := \{(\gamma,v) : \gamma\in\Gamma(V), v\in N(\gamma_{end})\},$$

and

$$M_\Gamma := \{(\gamma,(v,u)) : \gamma\in\Gamma(V), (v,u)\in N(\gamma_{end})\times N(\gamma_{end})\}.$$

We then introduce the birth, death and growth rates of any lineage of a cell of trait v as

$$\alpha^{(n)}(v) = \alpha(v)\Big(1-\overline{\mu}^{(n)}(v)\Big)^2 \text{ with } \overline{\mu}^{(n)}(v) := \sum_{u\in V:(v,u)\in E} \mu^{(n)}(v,u),$$
$$\beta^{(n)}(v) = \beta(v) + \alpha(v) \sum_{(u,w)\in N(v)\times N(v)} \mu^{(n)}(v,u)\mu^{(n)}(v,w),$$
$$\lambda^{(n)}(v) = \alpha^{(n)}(v) - \beta^{(n)}(v) = \lambda(v) - 2\alpha(v)\overline{\mu}^{(n)}(v).$$

Let $Q^b_{(0)}(ds,d\theta)$, $Q^d_{(0)}(ds,d\theta)$, $(Q_\gamma(ds,d\theta))_{\gamma\in\Gamma(V)}$, $(Q_{\gamma,v}(ds,d\theta))_{(\gamma,v)\in N_\Gamma}$ and $\big(Q_{\gamma,(v,u)}(ds,d\theta)\big)_{(\gamma,(v,u))\in M_\Gamma}$ be independent PPMs with intensity $dsd\theta$. The subpopulation of primary cells is

$$Z^{(n)}_{(0)}(t) := 1 + \int_0^t \int_{\mathbb{R}^+} \mathbb{1}_{\left\{\theta\le\alpha^{(n)}(0)Z^{(n)}_{(0)}(s^-)\right\}} Q^b_{(0)}(ds,d\theta) \tag{4.4.1}$$
$$- \int_0^t \int_{\mathbb{R}^+} \mathbb{1}_{\left\{\theta\le\beta(0)Z^{(n)}_{(0)}(s^-)\right\}} Q^d_{(0)}(ds,d\theta) - \sum_{(v,u)\in N(0)\times N(0)} H^{(n)}_{(0),(v,u)}(t),$$

and for all $\gamma\in\Gamma(V)$, the subpopulation among the cells of trait γ_{end} whose ancestry traces back to trait 0 with mutations occurring exactly along the edges of γ is

$$Z^{(n)}_\gamma(t) := \int_0^t \int_{\mathbb{R}^+} \Bigg(\mathbb{1}_{\left\{\theta\le\alpha^{(n)}(\gamma_{end})Z^{(n)}_\gamma(s^-)\right\}} \tag{4.4.2}$$
$$- \mathbb{1}_{\left\{\alpha^{(n)}(\gamma_{end})Z^{(n)}_\gamma(s^-)\le\theta\le\left(\alpha^{(n)}(\gamma_{end})+\beta(\gamma_{end})\right)Z^{(n)}_\gamma(s^-)\right\}} \Bigg) Q_\gamma(ds,d\theta)$$
$$+ K^{(n)}_{\overleftarrow{\gamma},\gamma_{end}}(t) + 2H^{(n)}_{\overleftarrow{\gamma},(\gamma_{end},\gamma_{end})} + \sum_{v\in N(\overline{\gamma}_{end}),v\neq\gamma_{end}} \left(H^{(n)}_{\overleftarrow{\gamma},(\gamma_{end},v)} + H^{(n)}_{\overleftarrow{\gamma},(v,\gamma_{end})}\right)(t)$$
$$- \sum_{(v,u)\in N(\gamma_{end})\times N(\gamma_{end})} H^{(n)}_{\gamma,(v,u)}(t),$$

where for all $(\gamma,v)\in N_\Gamma$,

$$K^{(n)}_{\gamma,v}(t) := \int_0^t \int_{\mathbb{R}^+} \mathbb{1}_{\{\theta\le 2\alpha(\gamma_{end})\mu^{(n)}(\gamma_{end},v)\left(1-\overline{\mu}^{(n)}(\gamma_{end})\right)Z^{(n)}_\gamma(s^-)\}} Q_{\gamma,v}(ds,d\theta), \tag{4.4.3}$$

and for all $(\gamma,(v,u))\in M_\Gamma$,

$$H^{(n)}_{\gamma,(v,u)}(t) := \int_0^t \int_{\mathbb{R}^+} \mathbb{1}_{\left\{\theta\le\alpha(\gamma_{end})\mu^{(n)}(\gamma_{end},v)\mu^{(n)}(\gamma_{end},u)Z^{(n)}_\gamma(s^-)\right\}} Q_{\gamma,(v,u)}(ds,d\theta). \tag{4.4.4}$$

The process $\big(K^{(n)}_{\gamma,v}(t)\big)_{t\in\mathbb{R}^+}$, resp. $\big(H^{(n)}_{\gamma,\{v,u\}}(t) := H^{(n)}_{\gamma,(v,u)}(t) + H^{(n)}_{\gamma,(u,v)}(t)\big)_{t\in\mathbb{R}^+}$, counts the number of mutations up to time t from the subpopulation indexed by γ that result in exactly one mutant daughter cell of trait v, resp. two mutant daughter cells of traits $\{v,u\}$. Hence the subpopulation of trait $v\in V$ is

$$Z^{(n)}_v(t) := Z^{(n)}_{(0)}(t)\mathbb{1}_{\{v=0\}} + \sum_{\gamma\in P(v)} Z^{(n)}_\gamma(t), \tag{4.4.5}$$

where $P(v)$, defined in Definition 4.2.5, is the set of all walks from trait 0 to trait v.

Definition 4.4.2 (Limiting birth-death branching process for the primary cell population)**.** *Let $(Z_{(0)}(t))_{t\in\mathbb{R}^+}$ be the birth-death branching process with rates $\alpha(0)$ and $\beta(0)$ respectively, constructed as follows*

$$Z_{(0)}(t) = 1 + \int_0^t \int_{\mathbb{R}^+} \mathbb{1}_{\{\theta\leq\alpha(0)Z_{(0)}(s^-)\}} Q^b_{(0)}(ds, d\theta) - \int_0^t \int_{\mathbb{R}^+} \mathbb{1}_{\{\theta\leq\beta(0)Z_{(0)}(s^-)\}} Q^d_{(0)}(ds, d\theta).$$

Notice that with this construction, the monotone coupling

$$\forall t \geq 0, Z^{(n)}_{(0)}(t) \leq Z_{(0)}(t), a.s.$$

immediately follows.
Introduce the almost sure limit of the positive martingale $\left(e^{-\lambda(0)t}Z_{(0)}(t)\right)_{t\in\mathbb{R}^+}$ as

$$W := \lim_{t\to\infty} e^{-\lambda(0)t} Z_{(0)}(t), \tag{4.4.6}$$

whose law is $W \overset{law}{=} Ber\left(\frac{\lambda(0)}{\alpha(0)}\right) \otimes Exp\left(\frac{\lambda(0)}{\alpha(0)}\right)$, see [22, Section 1.1], or [94, Theorem 1].

Lemma 4.4.3. *There exist $C(\alpha(0), \lambda(0)) > 0$ and $N \in \mathbb{N}$ such that for all $\varepsilon > 0$ and $n \geq N$,*

$$\mathbb{P}\Big(\sup_{t\in\mathbb{R}^+} \left|e^{-\lambda(0)t}Z_{(0)}(t) - e^{-\lambda^{(n)}(0)t}Z^{(n)}_{(0)}(t)\right| \geq \varepsilon\Big) \leq \frac{C(\alpha(0),\lambda(0))}{\varepsilon^2}\overline{\mu}^{(n)}(0) \underset{n\to\infty}{\longrightarrow} 0.$$

Proof. The result is derived by adapting the proof of Lemma 4.3.1 with $\mu^{(n)}_0$ replaced by $\overline{\mu}^{(n)}(0)$. □

Introduce the stopping time of the first time that the primary cell population reaches the size n^t as

$$\tau^{(n)}_t := \inf\left\{u \in \mathbb{R}^+ : Z^{(n)}_{(0)}(u) \geq n^t\right\}.$$

Lemma 4.4.4. *For all $\varepsilon > 0$, $(T_1, T_2) \in \mathbb{R}^+$ and φ_n such that $\log(n) = o(\varphi_n)$ and $\varphi_n = o\left(n^{\min\limits_{v\in N(0)} \ell(0,v)}\right)$, we have*

$$\mathbb{P}\Big(\sup_{t\in\left[T_1, T_2\frac{\varphi_n}{\log(n)}\right]} \left|\tau^{(n)}_t - \left(\mathfrak{t}^{(n)}_t - \frac{\log(W)}{\lambda(0)}\right)\right| \geq \varepsilon \Big| W > 0\Big) \underset{n\to\infty}{\longrightarrow} 0.$$

Proof. By following the proof of Lemma 4.3.2, with $Z^{(n)}_0$ and $\eta^{(n)}_t$ replaced by $Z^{(n)}_{(0)}$ and $\tau^{(n)}_t$, respectively, we obtain the result. □

In the next definition, we introduce an equivalence relation on $\Gamma(V)$. Two walks are considered equivalent if they are identical up to cycles (including cycles formed by backward mutations). More precisely, two walks are equivalent if there exists a minimal walk such that both walks use all the edges of this minimal walk, possible along with additional edges that form cycles. The purpose of this equivalence relation is to establish that, within a class of equivalence, only the walk with the minimal length contributes to the asymptotic size of the mutant subpopulation. In particular, this minimal walk is actually a path, since only distinct vertices are visited.

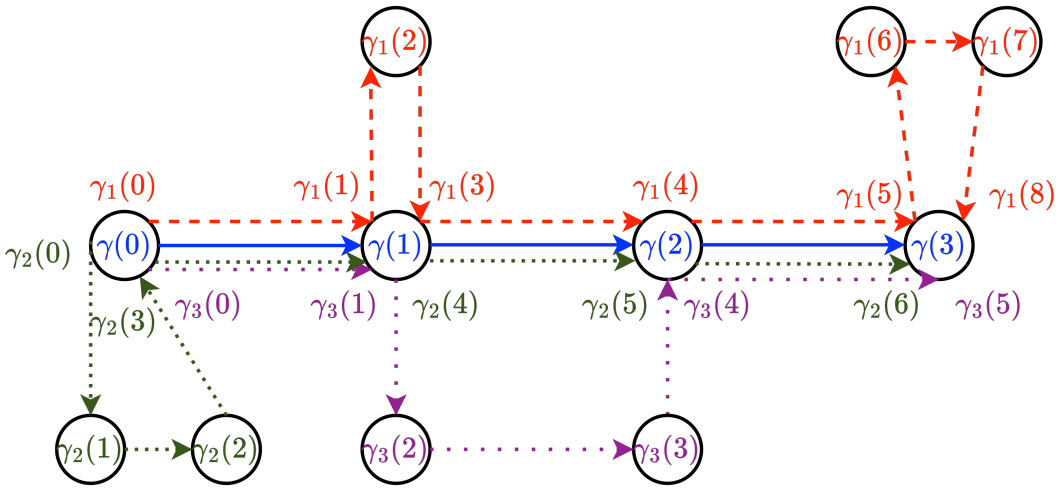

Figure 4.6: Example of Definition 4.4.5: here the walks γ, γ_1 and γ_2, represented respectively in plain blue, dashed red and dense dotted green, respectively, are equivalent. However, the walk γ_3, represented in sparse dotted purple, is not equivalent to any of the other walks. In particular, it is not possible to construct a function σ satisfying condition (ii) of Definition 4.4.5 for the walk γ_3. We have $|\gamma| = 4$, $|\gamma_1| = 9$, $|\gamma_2| = 7$, and $\sigma_1(0) = (0,0)$, $\sigma_1(1) = (1,3)$, $\sigma_1(2) = (4,4)$, $\sigma_1(3) = (5,8)$, $\sigma_2(0) = (0,3)$, $\sigma_2(1) = (4,4)$, $\sigma_2(2) = (5,5)$ and $\sigma_2(3) = (6,6)$.

Definition 4.4.5 (Equivalence relation on $\Gamma(V)$)**.** *We say that two walks γ_1 and γ_2 in $\Gamma(V) \times \Gamma(V)$ are equivalent, denoted by $\gamma_1 \sim \gamma_2$, if and only if there exists $\gamma \in \Gamma(V)$, and for all $j \in \{1,2\}$ there exists*

$$\begin{aligned} \sigma_j : \{0, \cdots |\gamma| - 1\} &\to \{0, \cdots, |\gamma_j| - 1\}^2 \\ i &\mapsto (\underline{\sigma}_j(i), \overline{\sigma}_j(i)) \end{aligned}$$

satisfying :

(i) $\forall j \in \{1,2\}, \underline{\sigma}_j(0) = 0$, *and* $\overline{\sigma}_j(|\gamma| - 1) = |\gamma_j| - 1$,

(ii) $\forall i \in \{0, \cdots |\gamma| - 1\}, \forall j \in \{1,2\}, \underline{\sigma}_j(i) \leq \overline{\sigma}_j(i)$ *and* $\overline{\sigma}_j(i) + 1 = \underline{\sigma}_j(i+1)$,

(iii) $\forall i \in \{0, \cdots, |\gamma| - 1\}, \forall j \in \{1,2\}, \gamma(i) = \gamma_j(\underline{\sigma}_j(i)) = \gamma_j(\overline{\sigma}_j(i))$.

Since the graph is finite, there are only a finite number of equivalence classes. For each walk $\gamma \in \Gamma(V)$, denote by $[\gamma]$ its equivalence class. In each class of equivalence, there is a natural representative candidate which is the walk with the minimum length; we will denote this walk by $\widetilde{\gamma}$. For each $v \in V$, denote by $C(v)$ the set of representative candidates for the walks in $P(v)$. Note that $|C(v)| < \infty$. An illustration of this definition can be found in Figure 4.6..

We introduce the notion of the mono-directional graph associated to a walk γ in the following definition.

Definition 4.4.6. *The mono-directional graph associated to a walk* $\gamma = (0, \gamma(1), \cdots, \gamma(k))$ *is the graph* (V_γ, E_γ) *where*

$$V_\gamma := \{0, \gamma(1), \cdots, \gamma(k)\},$$
$$E_\gamma := \{(0, \gamma(1)), (\gamma(1), \gamma(2)), \cdots, (\gamma(k-1), \gamma(k))\}.$$

In other words, it is the graph composed of the successive subpopulations

$$\big(Z^{(n)}_{(0)}, Z^{(n)}_{(0,\gamma(1))}, \cdots, Z^{(n)}_{\gamma}\big).$$

Now we have all the preliminary results and definitions necessary to prove Theorem 4.2.7.

Proof of Theorem 4.2.7. We prove Equations (4.2.5) and (4.2.7). The proofs of Equations (4.2.6) and (4.2.8) are similar and are left to the reader.

Step 1: Let $\widetilde{\gamma}$ be a representative candidate of an equivalence class. Our first step is to prove, using the results of Section 4.3, that for all $\varepsilon > 0$

$$\mathbb{P}\Big(\sup_{s\in[-M,M]}\sup_{t\in[0,T]}\Big|\sum_{\gamma\in[\widetilde{\gamma}]}\frac{Z^{(n)}_\gamma\big(\mathfrak{t}^{(n)}_t+s\big)}{d^{(n)}_{\widetilde{\gamma}}(t,s)} - W w_{\widetilde{\gamma}}(t)\Big| \geq \varepsilon\Big) \underset{n\to\infty}{\longrightarrow} 0, \tag{4.4.7}$$

where for all $\gamma \in \Gamma(V)$,

$$\begin{aligned} d^{(n)}_\gamma(t,s) :=& \mathbb{1}_{\{t\in[0,t(\gamma)-h_n^{-1}\}} + \mathbb{1}_{\{t\in[t(\gamma)-h_n^{-1},t(\gamma))\}}\psi_n \log^{\theta(\gamma)-1}(n) \\ &+ \mathbb{1}_{\{t\in[t(\gamma),\infty)\}} n^{t-t(\gamma)} \log^{\theta(\gamma)}(n) e^{\lambda(0)s}, \end{aligned}$$

and w_γ is defined in (4.2.3). Notice that

$$\begin{aligned} &\mathbb{P}\Big(\sup_{s\in[-M,M]}\sup_{t\in[0,T]}\Big|\sum_{\gamma\in[\widetilde{\gamma}]}\frac{Z^{(n)}_\gamma\big(\mathfrak{t}^{(n)}_t+s\big)}{d^{(n)}_{\widetilde{\gamma}}(t,s)} - W w_{\widetilde{\gamma}}(t)\Big| \geq \varepsilon\Big) \\ &\quad\leq \mathbb{P}\Big(\sup_{s\in[-M,M]}\sup_{t\in[0,T]}\Big|\frac{Z^{(n)}_{\widetilde{\gamma}}\big(\mathfrak{t}^{(n)}_t+s\big)}{d^{(n)}_{\widetilde{\gamma}}(t,s)} - W w_{\widetilde{\gamma}}(t)\Big| \geq \varepsilon\Big) \qquad (4.4.8) \\ &\quad+ \sum_{\gamma\in[\widetilde{\gamma}]\setminus\{\widetilde{\gamma}\}:t(\gamma)\leq T}\mathbb{P}\Big(\sup_{s\in[-M,M]}\sup_{t\in[0,T]}\Big|\frac{Z^{(n)}_\gamma\big(\mathfrak{t}^{(n)}_t+s\big)}{d^{(n)}_{\widetilde{\gamma}}(t,s)}\Big| \geq \varepsilon\Big) \qquad (4.4.9) \\ &\quad+ \mathbb{P}\Big(\sup_{s\in[-M,M]}\sup_{t\in[0,T]}\sum_{\gamma\in[\widetilde{\gamma}]\setminus\{\widetilde{\gamma}\}:t(\gamma)>T}\Big|\frac{Z^{(n)}_\gamma\big(\mathfrak{t}^{(n)}_t+s\big)}{d^{(n)}_{\widetilde{\gamma}}(t,s)}\Big| \geq \varepsilon\Big). \qquad (4.4.10) \end{aligned}$$

The term in (4.4.8) converges to 0 by applying Equation (4.2.5) to the mono-directional graph defined by the walk $\widetilde{\gamma}$, as proven in Section 4.3. The term in (4.4.9) also converges to 0 since:

- the sum is over a finite set, as we are considering a finite graph with positive labels on the edges ,

- for each $\gamma \in [\widetilde{\gamma}]\backslash\{\widetilde{\gamma}\}$, we have $t(\gamma) > t(\widetilde{\gamma})$ by definition of the representative (see Definition 4.4.5). This implies, by applying Equation (4.2.5) to the mono-directional graph defined by γ, that

$$\mathbb{P}\Big(\sup_{s\in[-M,M]} \sup_{t\in[0,T]} \Big| \frac{Z_{\gamma}^{(n)}\big(\mathfrak{t}_t^{(n)}+s\big)}{d_{\widetilde{\gamma}}^{(n)}(t,s)} \Big| \geq \varepsilon \Big) \underset{n\to\infty}{\longrightarrow} 0.$$

The term in (4.4.10) converges to 0 because

$$\mathbb{P}\Big(\sup_{s\in[-M,M]} \sup_{t\in[0,T]} \sum_{\gamma\in[\widetilde{\gamma}]\backslash\{\widetilde{\gamma}\}:t(\gamma)>T} Z_{\gamma}^{(n)}\big(\mathfrak{t}_t^{(n)}+s\big) = 0 \Big) \underset{n\to\infty}{\to} 1. \tag{4.4.11}$$

Indeed, for each $\gamma \in [\widetilde{\gamma}]\backslash\{\widetilde{\gamma}\}$ satisfying $t(\gamma) > T$, we have

$$\mathbb{P}\Big(\sup_{s\in[-M,M]} \sup_{t\in[0,T]} Z_{\gamma}^{(n)}\big(\mathfrak{t}_t^{(n)}+s\big) = 0 \Big) \underset{n\to\infty}{\longrightarrow} 1, \tag{4.4.12}$$

by applying Lemma 4.3.6 to the mono-directional graph given by γ. It remains to handle the sum over the set $A_{\widetilde{\gamma}}(T) := \{\gamma \in [\widetilde{\gamma}]\backslash\{\widetilde{\gamma}\} : t(\gamma) > T\}$. The easiest situation occurs when $|A_{\widetilde{\gamma}}(T)| < \infty$, as the result follows directly in this case. This situation corresponds exactly to the case where there is no cycle in the graph structure (V, E) for the vertices of $\widetilde{\gamma}$. Now, consider the case $|A_{\widetilde{\gamma}}(T)| = \infty$. In this case, even though Equation (4.4.12) holds for all $\gamma \in A_{\widetilde{\gamma}}(T)$, it does not necessary imply that Equation (4.4.11) is automatically satisfied. The result follows if one can show that there exists a finite subset $B_{\widetilde{\gamma}}(T) \subset A_{\widetilde{\gamma}}(T)$ such that

$$\mathbb{P}\Big(\sup_{s\in[-M,M]} \sup_{t\in[0,T]} \sum_{\gamma\in A_{\widetilde{\gamma}}(T)\backslash B_{\widetilde{\gamma}}(T)} Z_{\gamma}^{(n)}\big(\mathfrak{t}_t^{(n)}+s\big) = 0 \Big| E_{\widetilde{\gamma}}^{(n)} \Big) = 1, \tag{4.4.13}$$

where $E_{\widetilde{\gamma}}^{(n)} := \big\{ \sup_{s\in[-M,M]} \sup_{t\in[0,T]} \sum_{\gamma\in B_{\widetilde{\gamma}}(T)} Z_{\gamma}^{(n)}\big(\mathfrak{t}_t^{(n)}+s\big) = 0 \big\}$. We will now show that $B_{\widetilde{\gamma}}(T)$ exists. The set $[\widetilde{\gamma}]$ consists of walks where, for each vertex v visited by $\widetilde{\gamma}$, there may be a cycle going back to v. Since there are only a finite number of vertices visited by $\widetilde{\gamma}$, and the labels on the vertices are positive, it follows that the number of walks $\gamma \in A_{\widetilde{\gamma}}(T)$ for which we need to control the event that they do not have any cells up to time $\mathfrak{t}_T^{(n)} + M$ is actually finite, and we denote this set by $B_{\widetilde{\gamma}}(T)$. Indeed, for all walks $\gamma \in A_{\widetilde{\gamma}}(T)\backslash B_{\widetilde{\gamma}}(T)$, there exists a walk $\gamma_1 \in B_{\widetilde{\gamma}}(T)$ such that cells in the subpopulation $Z_{\gamma}^{(n)}$ result from (potentially many) mutations of cells in the subpopulation $Z_{\gamma_1}^{(n)}$. Therefore, if one controls with high probability that no cells are generated up to time $\mathfrak{t}_T^{(n)} + M$ for the subpopulations indexed by $\gamma \in B_{\widetilde{\gamma}}^{(n)}$-which is feasible since $B_{\widetilde{\gamma}}^{(n)}$ is finite-it automatically implies by the mechanistic construction of the process that, under such an event, there are almost surely no cells in the subpopulations indexed by $\gamma \in A_{\widetilde{\gamma}}^{(n)}\backslash B_{\widetilde{\gamma}}^{(n)}$. This is precisely the statement of Equation (4.4.13).

Step 2: In this step, Equation (4.2.5) is proven. Notice that for $\gamma \in A(v)$, where $A(v)$ is defined in Definition 4.2.5, we have $d_{\gamma}^{(n)}(t,s) = d_{v}^{(n)}(t,s)$, and also that γ is the representative candidate $\widetilde{\gamma}$ of its equivalence class. In particular, this means that $\sum_{\gamma\in A(v)} w_{\gamma}(t) = \sum_{\widetilde{\gamma}\in C(v):\widetilde{\gamma}\in A(v)} w_{\widetilde{\gamma}}(t)$,

where $C(v)$ is defined in Definition 4.4.5. The proof is obtained by noting that

$$\mathbb{P}\Big(\sup_{s\in[-M,M]}\sup_{t\in[0,T]}\Big|\frac{Z_v^{(n)}\big(\mathfrak{t}_t^{(n)}+s\big)}{d_v^{(n)}(t,s)}-W\sum_{\gamma\in A(v)}w_\gamma(t)\Big|\geq\varepsilon\Big)$$

$$\leq\sum_{\widetilde{\gamma}\in C(v):\widetilde{\gamma}\in A(v)}\mathbb{P}\Big(\sup_{s\in[-M,M]}\sup_{t\in[0,T]}\Big|\sum_{\gamma\in[\widetilde{\gamma}]}\frac{Z_\gamma^{(n)}\big(\mathfrak{t}_t^{(n)}+s\big)}{d_{\widetilde{\gamma}}^{(n)}(t,s)}-Ww_{\widetilde{\gamma}}(t)\Big|\geq\varepsilon\Big)\tag{4.4.14}$$

$$+\sum_{\widetilde{\gamma}\in C(v):\widetilde{\gamma}\notin A(v)}\mathbb{P}\Big(\sup_{s\in[-M,M]}\sup_{t\in[0,T]}\Big|\sum_{\gamma\in[\widetilde{\gamma}]}\frac{Z_\gamma^{(n)}\big(\mathfrak{t}_t^{(n)}+s\big)}{d_v^{(n)}(t,s)}\Big|\geq\varepsilon\Big).\tag{4.4.15}$$

Indeed, (4.4.14) converges to 0 by applying Equation (4.4.7) and because the sum is finite. Similarly, (4.4.15) converges to 0 because the sum is finite and, for all $\widetilde{\gamma}\in C(v),\widetilde{\gamma}\notin A(v)$, we have either $t(\widetilde{\gamma})>t(v)$ or $\theta(\widetilde{\gamma})<\theta(v)$.

Step 3: In this step, we are going to prove Equation (4.2.7). By following the proof of Proposition 4.3.12, replacing $\eta_t^{(n)}$ with $\tau_t^{(n)}$, and defining W as in (4.4.6) instead of (4.3.5), we obtain that for all $0<T_1<T_2$ and for all $\varepsilon>0$,

$$\mathbb{P}\Big(\sup_{t\in[T_1,T_2]}\big(\tau_t^{(n)}-\sigma_t^{(n)}\big)\leq\varepsilon\Big|W>0\Big)\underset{n\to\infty}{\longrightarrow}1.$$

Indeed, because the number of vertices in the graph is finite, and due to Step 2, we have shown that the total number of mutant cells $\sum_{v\in V\setminus\{0\}}Z_v^{(n)}\big(\mathfrak{t}_t^{(n)}+s\big)$ is negligible compared to the number of wild-type cells $Z_{(0)}^{(n)}\big(\mathfrak{t}_t^{(n)}+s\big)$ for any time interval $[T_1,T_2]$. This allows us to apply the reasoning from (4.3.60) and (4.3.61), leading to a straightforward adaptation of the proof of Proposition 4.3.12. By adapting the different proofs from Subsection 4.3.3, we obtain that for all $0<T_1<T_2$, $M>0$ and $\varepsilon>0$,

$$\mathbb{P}\Big(\sup_{s\in[-M,M]}\sup_{t\in[T_1,T_2]}\Big|\frac{Z_v^{(n)}\big(\rho_t^{(n)}+s\big)}{d_v^{(n)}(t,s)}-\mathbb{1}_{\{W>0\}}w_v(t)\Big|\geq\varepsilon\Big)\underset{n\to\infty}{\longrightarrow}0.$$

□

Appendix

Proof of Lemma 4.3.5. For all $t\geq0$ let $\mathcal{F}_{i,t}^{(n)}$ the σ-field generated by $Z_j^{(n)}(s)$ for all $0\leq j\leq i$ and $0\leq s\leq t$. For all $h\geq0$, we have

$$\mathbb{E}\Big[M_i^{(n)}(t+h)-M_i^{(n)}(t)|\mathcal{F}_{i,t}^{(n)}\Big]=\mathbb{E}\Big[Z_i^{(n)}(t+h)\Big|\mathcal{F}_{i,t}^{(n)}\Big]e^{-\lambda_i^{(n)}(t+h)}\tag{4.4.16}$$

$$-Z_i^{(n)}(t)e^{-\lambda_i^{(n)}t}-\int_t^{t+h}2\alpha_{i-1}\mu_{i-1}^{(n)}e^{-\lambda_i^{(n)}s}\mathbb{E}\Big[Z_{i-1}^{(n)}(s)\Big|\mathcal{F}_{i,t}^{(n)}\Big]ds.$$

The forward Chapman-Kolmogorov equation gives the time-differential equation

$$\frac{d\mathbb{E}\Big[Z_i^{(n)}(t)\Big]}{dt}=\lambda_i^{(n)}\mathbb{E}\Big[Z_i^{(n)}(t)\Big]+2\alpha_{i-1}\mu_{i-1}^{(n)}\mathbb{E}\Big[Z_{i-1}^{(n)}(t)\Big],$$

which leads to

$$\mathbb{E}_{Z_i^{(n)}(0)}\Big[Z_i^{(n)}(t)\Big] = Z_i^{(n)}(0)e^{\lambda_i^{(n)}t} + \int_0^t 2\alpha_{i-1}\mu_{i-1}^{(n)}\mathbb{E}_{Z_i^{(n)}(0)}\Big[Z_{i-1}^{(n)}(s)\Big]e^{\lambda_i^{(n)}(t-s)}ds.$$

In particular, by using the Markov property we obtain that

$$\mathbb{E}\Big[Z_i^{(n)}(t+h)\Big|\mathcal{F}_{i,t}^{(n)}\Big] = Z_i^{(n)}(t)e^{\lambda_i^{(n)}h} + \int_t^{t+h} 2\alpha_{i-1}\mu_{i-1}^{(n)}\mathbb{E}\Big[Z_{i-1}^{(n)}(s)|\mathcal{F}_{i,t}^{(n)}\Big]e^{\lambda_i^{(n)}(t+h-s)}ds. \quad (4.4.17)$$

Combining (4.4.16) and (4.4.17), it follows that $\big(M_i^{(n)}(t)\big)_{t\in\mathbb{R}^+}$ is a martingale. Let

$$F^{(n)}(t,x,y) := (e^{-\lambda_i^{(n)}t}x - y)^2.$$

Then, we have

$$\frac{\partial F^{(n)}}{\partial t}(t,x,y) = -2\lambda_i^{(n)}xe^{-\lambda_i^{(n)}t}\sqrt{F^{(n)}(t,x,y)} \text{ and } \frac{\partial F^{(n)}}{\partial y}(t,x,y) = -2\sqrt{F^{(n)}}.$$

Applying Itô's formula with $x = Z_i^{(n)}(t)$ and $y = \int_0^t 2\alpha_{i-1}\mu_{i-1}^{(n)}e^{-\lambda_i^{(n)}s}Z_{i-1}^{(n)}(s)ds$, we obtain

$$\begin{aligned}
\Big(M_i^{(n)}(t)\Big)^2 &= F^{(n)}\Big(t, Z_i^{(n)}(t), \int_0^t 2\alpha_{i-1}\mu_{i-1}^{(n)}e^{-\lambda_i^{(n)}s}Z_{i-1}^{(n)}(s)ds\Big)\\
&= F^{(n)}(0,0,0) - 2\int_0^t 2\alpha_{i-1}\mu_{i-1}^{(n)}e^{-\lambda_i^{(n)}s}Z_{i-1}^{(n)}(s)M_i^{(n)}(s)ds\\
&- 2\lambda_i^{(n)}\int_0^t e^{-\lambda_i^{(n)}s}Z_i^{(n)}(s)M_i^{(n)}(s)ds\\
&+ \int_0^t\int_{\mathbb{R}^+}\Big[\Big(M_i^{(n)}(s^-)\\
&\qquad + e^{-\lambda_i^{(n)}s}\Big\{\mathbb{1}_{\left\{\theta\le\alpha_i^{(n)}Z_i^{(n)}(s^-)\right\}} - \mathbb{1}_{\left\{\alpha_i^{(n)}Z_i^{(n)}(s^-)\le\theta\le\left(\alpha_i^{(n)}+\beta_i\right)Z_i^{(n)}(s^-)\right\}}\Big\}\Big)^2\\
&\qquad\qquad - \Big(M_i^{(n)}(s^-)\Big)^2\Big]Q_i(ds,d\theta)\\
&+ \int_0^t\int_{\mathbb{R}^+}\Big[\Big(M_i^{(n)}(s^-) + e^{-\lambda_i^{(n)}s}\mathbb{1}_{\left\{\theta\le 2\alpha_{i-1}\mu_{i-1}^{(n)}\left(1-\mu_{i-1}^{(n)}\right)Z_{i-1}^{(n)}(s^-)\right\}}\Big)^2\\
&\qquad\qquad - \Big(M_i^{(n)}(s^-)\Big)^2\Big]N_{i-1}(ds,d\theta)
\end{aligned}$$

$$
\begin{aligned}
&+\int_0^t\int_{\mathbb{R}^+}\Big[\Big(M_i^{(n)}(s^-)+e^{-\lambda_i^{(n)}s}2\mathbb{1}_{\left\{\theta\le\alpha_{i-1}\left(\mu_{i-1}^{(n)}\right)^2Z_{i-1}^{(n)}(s^-)\right\}}\Big)^2\\
&\qquad\qquad-\Big(M_i^{(n)}(s^-)\Big)^2\Big]Q_{i-1}^m(ds,d\theta)\\
&+\int_0^t\int_{\mathbb{R}^+}\Big[\Big(M_i^{(n)}(s^-)-e^{-\lambda_i^{(n)}s}\mathbb{1}_{\left\{\theta\le\alpha_i\left(\mu_i^{(n)}\right)^2Z_i^{(n)}(s^-)\right\}}\Big)^2-\Big(M_i^{(n)}\Big)^2\Big]Q_i^m(ds,d\theta)\\
&=-2\int_0^t\Big(2\alpha_{i-1}\mu_{i-1}^{(n)}Z_{i-1}^{(n)}(s)+\lambda_i^{(n)}Z_i^{(n)}(s)\Big)e^{-\lambda_i^{(n)}s}M_i^{(n)}(s)ds\\
&+2\int_0^t\Big(2\alpha_{i-1}\mu_{i-1}^{(n)}Z_{i-1}^{(n)}(s)+\lambda_i^{(n)}Z_i^{(n)}(s)\Big)e^{-\lambda_i^{(n)}s}M_i^{(n)}ds\\
&+\int_0^t\Big[2\alpha_{i-1}\mu_{i-1}^{(n)}Z_{i-1}^{(n)}(s)+\left(\alpha_i^{(n)}+\beta_i^{(n)}\right)Z_i^{(n)}(s)\Big]e^{-2\lambda_i^{(n)}s}ds+\widetilde{M}_i^{(n)}(t)\\
&=\widetilde{M}_i^{(n)}(t)+\int_0^t2\alpha_{i-1}\mu_{i-1}^{(n)}e^{-2\lambda_i^{(n)}s}Z_{i-1}^{(n)}(s)ds+\left(\alpha_i^{(n)}+\beta_i^{(n)}\right)\int_0^te^{-2\lambda_i^{(n)}s}Z_i^{(n)}(s)ds,
\end{aligned}
$$

where $(\widetilde{M}_i^{(n)}(t))_{t\ge0}$ is a martingale. Finally, we obtain

$$
\left\langle M_i^{(n)}\right\rangle_t=\int_0^t2\alpha_{i-1}\mu_{i-1}^{(n)}e^{-2\lambda_i^{(n)}s}Z_i^{(n)}(s)ds+\left(\alpha_i^{(n)}+\beta_i^{(n)}\right)\int_0^te^{-2\lambda_i^{(n)}s}Z_i^{(n)}(s)ds.
$$

□

Proof of Lemma 4.3.9. First, we have that $\mathbb{E}\big[Z_0^{(n)}(u)\big]=e^{\lambda_0^{(n)}u}\le e^{\lambda_0 u}$, which is exactly the result for $i=0$. Then, for $i\in\mathbb{N}$ assume that the result is true for $i-1$. Taking the expected value of the martingale $M_i^{(n)}$ defined in (4.3.14) at time u and using the induction assumption we obtain the following

$$
\begin{aligned}
\mathbb{E}\Big[Z_i^{(n)}(u)\Big]&\le e^{\lambda_i u}\int_0^u2\alpha_{i-1}\mu_{i-1}^{(n)}e^{-\lambda_i s}\mathbb{E}\Big[Z_{i-1}^{(n)}(s)\Big]ds\\
&\le C_{i-1}\mu_{\otimes,i}^{(n)}2\alpha_{i-1}\int_0^ue^{(\lambda_0-\lambda_i)s}dsu^{\theta(i-1)}e^{\lambda_i u}\\
&\le C_{i-1}\mu_{\otimes,i}^{(n)}2\alpha_{i-1}\Big(\mathbb{1}_{\{\lambda_i=\lambda_0\}}u+\mathbb{1}_{\{\lambda_i<\lambda_0\}}\frac{1}{\lambda_0-\lambda_i}e^{(\lambda_0-\lambda_i)u}\Big)u^{\theta(i-1)}e^{\lambda_i u}\\
&=C_{i-1}\mu_{\otimes,i}^{(n)}2\alpha_{i-1}\Big(\mathbb{1}_{\{\lambda_i=\lambda_0\}}+\mathbb{1}_{\{\lambda_i<\lambda_0\}}\frac{1}{\lambda_0-\lambda_i}\Big)u^{\theta(i)}e^{\lambda_0 u},
\end{aligned}
$$

which concludes the proof by induction. □

Proof of Lemma 4.3.10. In the proof, C represents a positive constant that may change from line to line.

Neutral case: Assume that $\lambda_i=\lambda_0$. Applying Lemma 4.3.9, recalling that $\lambda_i^{(n)}=\lambda_0-2\alpha_i\mu_i^{(n)}$, and noting that there exists $N_1\in\mathbb{N}$ such that for all $n\ge N_1$, we have that $e^{4\alpha_i\mu_i^{(n)}\left(\mathfrak{t}_{t_2}^{(n)}+s\right)}\le2$, we obtain

$$
\int_{\mathfrak{t}_{t_1}^{(n)}+s}^{\mathfrak{t}_{t_2}^{(n)}+s}e^{-2\lambda_i^{(n)}u}\mathbb{E}\Big[Z_i^{(n)}(u)\Big]du\le C\mu_{\otimes,i}^{(n)}\int_{\mathfrak{t}_{t_1}^{(n)}+s}^{\mathfrak{t}_{t_2}^{(n)}+s}u^{\theta(i)}e^{-\lambda_0 u}du.
$$

Using integration by parts, we obtain

$$\int_{\mathfrak{t}_{t_1^{(n)}}^{(n)}+s}^{\mathfrak{t}_{t_2}^{(n)}+s} u^{\theta(i)} e^{-\lambda_0 u} du \leq \frac{1}{\lambda(0)} \big(\mathfrak{t}_{t_1^{(n)}}^{(n)} + s\big)^{\theta(i)} e^{-\lambda_0 \left(\mathfrak{t}_{t_1^{(n)}}^{(n)} + s\right)} + \frac{\theta(i)}{\lambda_0} \int_{\mathfrak{t}_{t_1^{(n)}}^{(n)}+s}^{\mathfrak{t}_{t_2}^{(n)}+s} u^{\theta(i)-1} e^{-\lambda_0 u} du.$$

Then, using $\theta(i)$ integrations by parts, there exists $N_2 \in \mathbb{N}$ such that for $n \geq N_2$, we have

$$\int_{\mathfrak{t}_{t_1^{(n)}}^{(n)}+s}^{\mathfrak{t}_{t_2}^{(n)}+s} u^{\theta(i)} e^{-\lambda_0 u} du \leq C \frac{e^{-\lambda_0 s}}{n^{t_1^{(n)}}} \big(\mathfrak{t}_{t_1^{(n)}}^{(n)} + s\big)^{\theta(i)}.$$

It follows that for $n \geq \max(N_1, N_2)$,

$$\int_{\mathfrak{t}_{t_1^{(n)}}^{(n)}+s}^{\mathfrak{t}_{t_2}^{(n)}+s} e^{-2\lambda_i^{(n)} u} \mathbb{E}\Big[Z_i^{(n)}(u)\Big] du \leq C \frac{e^{-\lambda_0 s}}{n^{t_1^{(n)}}} \mu_{\otimes,i}^{(n)} \big(\mathfrak{t}_{t_1^{(n)}}^{(n)} + s\big)^{\theta(i)}.$$

Since vertex i is assumed to be neutral, we have $\theta(i-1) = \theta(i) - 1$. Using similar computation as above, there exists $N_3 \in \mathbb{N}$ such that for $n \geq N_3$, we have

$$\int_{\mathfrak{t}_{t_1^{(n)}}^{(n)}+s}^{\mathfrak{t}_{t_2}^{(n)}+s} \mu_{i-1}^{(n)} e^{-2\lambda_i^{(n)} u} \mathbb{E}\Big[Z_{i-1}^{(n)}(u)\Big] du \leq C \frac{e^{-\lambda_0 s}}{n^{t_1^{(n)}}} \mu_{\otimes,i}^{(n)} \big(\mathfrak{t}_{t_1^{(n)}}^{(n)} + s\big)^{\theta(i)-1}.$$

It follows that for all $n \geq \max(N_1, N_2, N_3)$, we have

$$\mathbb{E}\Big[\big\langle M_i^{(n)}\big\rangle_{\mathfrak{t}_{t_2}^{(n)}+s} - \big\langle M_i^{(n)}\big\rangle_{\mathfrak{t}_{t_1^{(n)}}^{(n)}+s}\Big] \leq C \frac{e^{-\lambda_0 s}}{n^{t_1^{(n)}}} \mu_{\otimes,i}^{(n)} \big(\mathfrak{t}_{t_1^{(n)}}^{(n)} + s\big)^{\theta(i)}.$$

Deleterious case: We now address the case $\lambda_i < \lambda_0$ by applying the same strategy. We obtain

$$\begin{aligned}
\int_{\mathfrak{t}_{t_1^{(n)}}^{(n)}+s}^{\mathfrak{t}_{t_2}^{(n)}+s} e^{-2\lambda_i^{(n)} u} \mathbb{E}\Big[Z_i^{(n)}(u)\Big] du &\leq C \mu_{\otimes,i}^{(n)} \int_{\mathfrak{t}_{t_1^{(n)}}^{(n)}+s}^{\mathfrak{t}_{t_2}^{(n)}+s} u^{\theta(i)} e^{(\lambda_0 - 2\lambda_i)u} du \\
&\leq C \mu_{\otimes,i}^{(n)} \left(\mathfrak{t}_{t_2}^{(n)} + s\right)^{\theta(i)} \Big[\mathbb{1}_{\{\lambda_0 > 2\lambda_i\}} e^{(\lambda_0 - 2\lambda_i)\left(\mathfrak{t}_{t_2}^{(n)} + s\right)} \\
&\qquad + \mathbb{1}_{\{\lambda_0 = 2\lambda_i\}} \big(\mathfrak{t}_{t_2}^{(n)} + s\big) + \mathbb{1}_{\{\lambda_i < \lambda_0 < 2\lambda_i\}} e^{-(2\lambda_i - \lambda_0)\left(\mathfrak{t}_{t_1^{(n)}}^{(n)} + s\right)}\Big].
\end{aligned}$$

Recalling that $\theta(i-1) = \theta(i)$, we find

$$\begin{aligned}
\int_{\mathfrak{t}_{t_1^{(n)}}^{(n)}+s}^{\mathfrak{t}_{t_2}^{(n)}+s} \mu_{i-1}^{(n)} e^{-2\lambda_i^{(n)} u} \mathbb{E}\Big[Z_{i-1}^{(n)}(u)\Big] du &\leq C \mu_{\otimes,i}^{(n)} \big(\mathfrak{t}_{t_2}^{(n)} + s\big)^{\theta(i)} \Big[\mathbb{1}_{\{\lambda_0 > 2\lambda_i\}} e^{(\lambda_0 - 2\lambda_i)(\mathfrak{t}_{t_2}^{(n)} + s)} \\
&\qquad + \mathbb{1}_{\{\lambda_0 = 2\lambda_i\}} \big(\mathfrak{t}_{t_2}^{(n)} + s\big) + \mathbb{1}_{\{\lambda_i < \lambda_0 < 2\lambda_i\}} e^{-(2\lambda_i - \lambda_0)\left(\mathfrak{t}_{t_1^{(n)}}^{(n)} + s\right)}\Big].
\end{aligned}$$

Finally, we have

$$\mathbb{E}\Big[\big\langle M_i^{(n)}\big\rangle_{\mathfrak{t}^{(n)}_{t_2}+s} - \big\langle M_i^{(n)}\big\rangle_{\substack{\mathfrak{t}^{(n)}_{t_1}+s}}\Big] \leq C\mu^{(n)}_{\otimes,i}\big(\mathfrak{t}^{(n)}_{t_2}+s\big)^{\theta(i)} \cdot \Big[\mathbb{1}_{\{\lambda_0>2\lambda_i\}} e^{(\lambda_0-2\lambda_i)(\mathfrak{t}^{(n)}_{t_2}+s)} + \mathbb{1}_{\{\lambda_0=2\lambda_i\}}\big(\mathfrak{t}^{(n)}_{t_2}+s\big) + \mathbb{1}_{\{\lambda_i<\lambda_0<2\lambda_i\}} e^{-(2\lambda_i-\lambda_0)\big(\mathfrak{t}^{(n)}_{t_1}+s\big)}\Big].$$

□

Acknowledgements The author would like to thank Hélène Leman for inspiring and helpful discussions, as well as feedback. The author would like to thank the Chair "Modélisation Mathématique et Biodiversité" of VEOLIA-Ecole Polytechnique-MNHN-F.X.

Funding This research was led with financial support from ITMO Cancer of AVIESAN (Alliance Nationale pour les Sciences de la Vie et de la Santé, National Alliance for Life Sciences Health) within the framework of the Cancer Plan.

Chapter 5

Selective Cancer Evolution

This chapter, a collaborative work in progress with Hélène Leman, aims to generalise Theorem 4.2.7 to the case where no conditions are imposed on the growth rate function. This generalization is particularly relevant in the context of understanding cancer evolution with selective mutations, where refined results are sought. Chapter 4 focused solely (when considering selective cancer evolution) on the stochastic exponents of the mutant subpopulations, whereas in the present chapter, we delve into capturing the first-order asymptotics of their sizes. This endeavor enables the characterization of evolutionary pathways and provides a deeper understanding of the asymptotic stochastic contributions, as already mentioned in Chapter 4.

Within this chapter, we present results on the first-order asymptotics for the case of a finite mono-directional graph, where the last mutation corresponds to the first selective one. We compare the first-order asymptotics of the first-selective mutant subpopulation size with the ones obtain in Chapter 4 for neutral and deleterious mutations. More specifically on three different points: the powers of n and of $\log(n)$ asymptotically captured, as well as the stochasticity of the limit. We also emphasize on the difference of the proof method developed specifically to deal with such selective mutation. Additionally, we discuss the increasing difficulty of generalizing these results to arbitrary monodirectional graphs. A new proof method have to be developed to tackle such difficulty for first obtaining results on any monodirectional graph and then deduce results on any general finite graph using an effective evolutionary pathways approach, as for neutral and deleterious cancer evolution, see Chapter 4, Section 4.4.

5.1 Introduction and microscopic model

In Chapter 4 we studied a toy model of carcinogenesis, primarily focusing on neutral (and deleterious) cancer evolution, where results are stated in Theorem 4.2.7. These results concerning the first-order asymptotics of mutant subpopulation sizes compared to those in current literature regarding stochastic exponents offer new insights into the effective evolutionary pathways within the trait space, as well as into the stochasticity of genetic composition when observing a tumor, but limited to neutral and deleterious mutations only.

In Theorem 4.2.9, selective mutations are considered, enabling discrimination among evolutionary pathways, some of which may contribute asymptotically negligibly. However, this theorem does not precisely characterise the non-negligible pathways, nor does it provide information on the

stochasticity of the genetic composition. The scope of this chapter is to extend the results of Theorem 4.2.7 to a more general cancer evolution framework that allows for selective mutant traits. A novel approach is introduced making progress in obtaining first-order asymptotic results on mutant subpopulation sizes when selective mutations are considered in the trait space. Up to this point, we have achieved the case of a finite labelled mono-directional graph, where there is a unique selective mutation corresponding to the last one. The result, stated in Theorem 5.2.3, reveals several biologically and mathematically intriguing results, which contrast with our previous findings in the context where only neutral and deleterious mutations are considered. In essence, the current approach developed in this chapter successfully addresses scenarios involving only the first selective mutation, but proves inadequate for handling subsequent mutations after the initial selective one. This fact is primarily due to an infinite expectation of the limiting process. Additional work to get the first order asymptotics of the mutant subpopulation sizes for any general finite graph allowing for selective mutations remains to be done in the future.

The model corresponds to the one described in Chapter 4, Section 4.1, using the notations (V, E, L) for the finite directed labelled graph structure representing the evolutionary trait space and $\alpha_v, \beta_v, \lambda_v$ for the birth, death, and growth rates of a trait $v \in V$ respectively. We keep the same initial condition $Z_v(0) = \mathbb{1}_{\{v=0\}}$ for all $v \in V$ almost surely. We are considering the following specific case of a finite labelled mono-directional graph with $k+1$ vertices, for some $k \geq 1$. More precisely, we have

$$(V, E, L) = \Big(\big\{0, \cdots, k\big\}, \big\{(i, i+1), i \in \{0, \cdots, k-1\}\big\}, \big\{\ell(i), i \in \{0, \cdots, k-1\}\big\}\Big). \quad (5.1.1)$$

We use the simplified notations $\mu_i^{(n)} := \mu^{(n)}(i, i+1)$ giving the following power law mutation rates regime

$$\forall 0 \leq i \leq k-1, n^{\ell(i)}\mu_i^{(n)} \underset{n\to\infty}{\longrightarrow} \mu_i < \infty.$$

Assume that the first selective mutant trait is the last one, which mathematically means

$$\forall 0 \leq i \leq k-1, \lambda_i \leq \lambda_0 \text{ and } \lambda_k > \lambda_0. \quad (5.1.2)$$

With this setting three different scenari can happen during a division event of a cell of trait $i \in \{0, \cdots, k-1\}$:

- with probability $\left(1-\mu_i^{(n)}\right)^2$ each daughter cell keeps the trait i of its mother cell,
- with probability $2\mu_i^{(n)}\left(1-\mu_i^{(n)}\right)$ exactly one of the daughter cell mutates to the next trait $i+1$ when the second daughter cell keeps the trait i of its mother cell,
- with probability $\left(\mu_i^{(n)}\right)^2$ both of the daughter cells mutate to the next trait $i+1$.

For $i = k$, during a division event of a cell of trait k, both daughter cells keeps the trait k. A graphical representation of the model can be found in Figure 5.1. Instead of employing the deterministic $\log(n)/\lambda_0$-accelerated time scale of (4.1.4), which is appropriate for neutral (and deleterious) cancer evolution since all subpopulations grow exponentially fast at rate λ_0 according to Theorem 4.2.7, we opt for the $\log(n)$-accelerated time scale $t \mapsto t\log(n)$ in this context. To be fully consistent, we define again, for this new time scale, $t(i)$ and $\theta(i)$ for all $i \in V$.

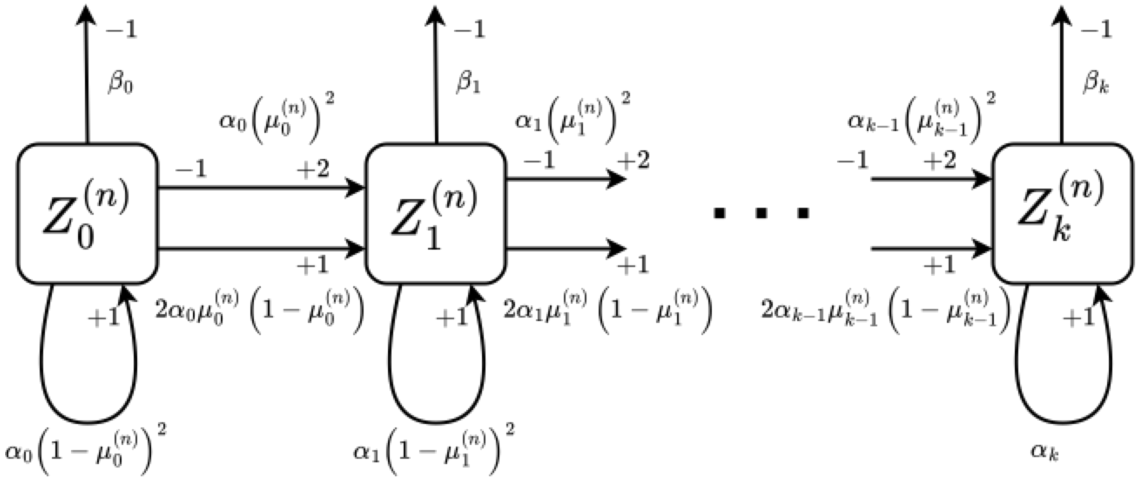

Figure 5.1: Dynamical representation of the finite mono-directional graph

Definition 5.1.1. *Define for all* $1 \leq i \leq k$

$$t(i) := \frac{1}{\lambda_0} \sum_{j=0}^{i-1} \ell(j),$$
$$\theta(i) := \left|\{1 \leq j \leq i : \lambda_j = \lambda_0\}\right|,$$

as respectively the sum of the labels on the edges from trait 0 *to trait* i *renormalized by* λ_0 *and the number of neutral mutations on the path from* 0 *to* i*.*

We rewrite in the following proposition the asymptotic result from Chapter 4, Theorem 4.2.7 in the $\log(n)$-accelerated time scale instead of in the $\log(n)/\lambda_0$-accelerated time scale, for the subpopulations of traits $i \in \{1, \cdots, k-1\}$.

Proposition 5.1.2. *Let* (V, E, L) *be a finite directed labelled mono-directional graph with* $k+1$ *vertices, for some* $k \geq 1$ *as defined in* (5.1.1)*. Assume that the first selective mutation on the graph is the last one as in* (5.1.2)*. Let* $M > 0$ *and* $0 < T_1 < T_2$*. Using the mathematical definition of the model given in Section 5.3, see* (5.3.2)*,* (5.3.3) *and* (5.3.4)*, there exists a random variable* W *properly defined in* (5.3.10) *such that for all* $i \in \{1, \cdots, k-1\}$

$$\left((t,s) \mapsto \frac{Z_i^{(n)}\left((t(i)+t)\log(n)+s\right)}{n^{t\lambda_0}\log^{\theta(i)}(n)e^{\lambda_0 s}}\right) \underset{n\to\infty}{\longrightarrow} W\widetilde{w}_{(0,\cdots,i)}(t(i)+t),$$

where $\widetilde{w}_{(0,\cdots,i)}(t) := w_{(0,\cdots,i)}(t\lambda_0)$ *in probability in* $L^\infty([T_1,T_2]\times[-M,M])$*. For any other mathematical description, the convergence is at least in distribution in* $\mathbb{D}([T_1,T_2]\times[-M,M])$*.*

Remark 5.1.3. *For the trait* $k-1$ *only, the mathematical definition of* $Z_{k-1}^{(n)}$ *is slightly different in this chapter compared to the one in Chapter 4, Section 4.3, due to different techniques used*

in the proof of the first-order asymptotic of trait k (see Subsection 5.3 for more details on the construction). The difference does not hamper however from getting the same first-order asymptotics result for the subpopulation size $Z_{k-1}^{(n)}$. Indeed, only the computation of the quadratic variation of the martingale $M_{k-1}^{(n)}$, defined in Lemma 4.3.5, is modified, but this subpopulation is fully constructed using independent Poison Point Measures, allowing to perfectly adapt the computation done in the proof of this lemma to get the same result at the end.

5.2 Main results and biological interpretation

In this Section, the result regarding the first-order asymptotics of the first selective mutant subpopulation size is stated. Moreover, some mathematical and biological remarks about this result, as well as a sketch of the proof, based on heuristics, are provided.

When considering a selective mutation, the first intuitive idea that arises from the term "selective" is that only the lineages of the first generated mutant cells contribute asymptotically. Indeed, any time delay for a lineage results in an exponential growth loss for its size, providing some qualitative weight to this intuition. However, the challenge lies precisely in quantifying this intuition. Specifically, it is necessary to determine how each lineage of the generated mutant cells contributes asymptotically to the subpopulation size. This raises natural questions for understanding, such as whether only a finite number of surviving lineages contribute asymptotically, and if so, how many, as well as whether it is the first lineage to reach a certain level that contributes, among potentially many other questions. Answering them is not possible solely with results on the stochastic exponents. However, Theorem 5.2.3 quantitatively answers these natural questions on the contribution of all the lineages on the first-order asymptotics of the first selective mutant subpopulation size. For doing that, we are going to separate the process $Z_k^{(n)}$ into 3 different processes having different mathematical analysis, asymptotic contribution and biological meaning:

(i) The first one, denoted by $Z_{k,1}^{(n)}$, corresponds to the process composed of the mutant cells and their lineages generated up to time

$$t_-^{(n)}(k) := t(k)\log(n) - \frac{\theta(k-1)}{\lambda_0}\log(\log(n)).$$

This time corresponds exactly to the first deterministic time at which it becomes likely to observe mutational events from trait $k-1$ to trait k. Indeed the typical order of the subpopulation $Z_{k-1}^{(n)}$ at time $(t(k-1)+t)\log(n)$, that is $n^{t\lambda_0}\log^{\theta(k-1)}(n)$ according to Proposition 5.1.2, reaches $n^{\ell(k-1)}$, the typical size allowing mutations, at time $t_-^{(n)}(k)$. We will show that well renormalized, the subpopulation $Z_{k,1}^{(n)}$ at time $(t(k)+t)\log(n)$ asymptotically follows a compound Luria-Delbrück distribution, in the vein of the works of Cheek and Antal [8, 10].

(ii) The second one, denoted by $Z_{k,2}^{(n)}$, corresponds to the process composed of the mutant cells and their lineages generated between times $\left[t_-^{(n)}(k), t(k)\log(n)\right]$, that is during a time scale of order $\log(\log(n))$. We will show that well renormalized, the subpopulation $Z_{k,2}^{(n)}$ at time $(t(k)+t)\log(n)$ asymptotically follows the large time distribution of an approximate model with less stochasticity, that is independent from n. This approach is inspired by the works of Durrett and Moseley [22], Nicholson and Antal [23] and Nicholson, Cheek and Antal [24].

(iii) The last one, denoted by $Z_{k,3}^{(n)}$, corresponds to the process composed of the mutant cells and their lineages generated after time $t(k)\log(n)$. We will show, using a similar martingale approach as the one of Chapter 4, that asymptotically this process becomes negligible with respect to the total mutant population $Z_k^{(n)}$.

A graphical representation of these three processes is given in Figure 5.2.

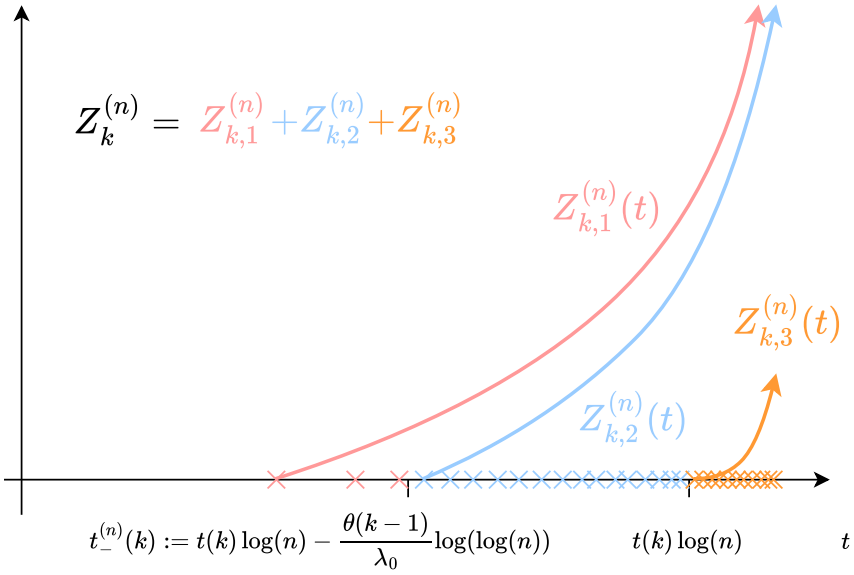

Figure 5.2: Graphical representation of the 3 subprocesses composing $Z_k^{(n)}$

As aforementioned, for stating the first-order asymptotics of the mutant subpopulation $Z_{k,1}^{(n)}$, we introduce a proper definition of the Luria-Delbrück distribution as well as the compound Luria-Delbrück distribution.

Definition 5.2.1 (Luria-Delbrück distribution)**.** *Let $(\xi_i)_{i\in\mathbb{N}}$ be an i.i.d. sequence of exponentially distributed random variables with parameter λ. Let $(Y_i)_{i\in\mathbb{N}}$ be an i.i.d. sequence of birth and death branching processes with rates α and β respectively, satisfying almost surely the initial condition $Y_i(0) = 1$. Let K be a Poisson random variable with parameter ω. The random variables and processes $(\xi_i)_{i\in\mathbb{N}}, (Y_i)_{i\in\mathbb{N}}, K$ are mutually independent. The Luria-Delbrück distribution with parameters $(\lambda, \alpha, \beta, \omega)$, that may be chosen randomly, is defined as the distribution of*

$$B = \sum_{i=1}^{K} Y_i(\xi_i).$$

The compound Luria-Delbrück distribution with parameters $(\lambda, \alpha, \beta, \omega)$ and associated with the distribution $\mathcal{U}$ is defined as the distribution of

$$Z = \sum_{i=1}^{B} U_i,$$

with $(U_i)_{i\in\mathbb{N}}$ *a sequence of i.i.d. random variables following* $\mathcal{U}$ *that is independent from* B*.*

In order to deal with the mathematical analysis of the process $Z_{k,2}^{(n)}$, we introduce the following approximate model with less stochasticity for which asymptotic result concerning the large time limit of the population size is easily derived.

Definition 5.2.2 (Approximate Model)**.** *The approximate model with parameters* $(\lambda, \mu, \alpha, \beta)$ *is defined as the distribution of*

$$Z(t) = \sum_{i=1}^{K(t)} Y_i(t - T_i),$$

where $K(t) = N\left(\int_0^t \mu e^{\lambda s} ds\right)$*, with* N *a Poisson process with intensity 1, independent from the i.i.d. sequence* $(Y_i)_{i\in\mathbb{N}}$ *of birth and death branching processes with rates* α *and* β*, and* $T_i = \inf\{t \geq 0 : K(t) \geq i\}$*. Notice that for all* $t > 0$ *and for all* $i \leq K(t)$*, we have* $t - T_i \geq 0$*. Define* $\widetilde{\lambda} := \alpha - \beta > \lambda$*, then we have*

$$e^{-\widetilde{\lambda}t} Z(t) \underset{t\to\infty}{\longrightarrow} Z^\infty, \tag{5.2.1}$$

almost surely where $Z^\infty \in L^1$ *satisfies* $\mathbb{E}[Z^\infty] = \mu/\left(\widetilde{\lambda} - \lambda\right)$ *and its Laplace transform follows the equation*

$$L(\xi, t) := \mathbb{E}\left[e^{-\xi e^{-\widetilde{\lambda}t} Z(t)}\right] \underset{t\to\infty}{\longrightarrow} \mathbb{E}\left[\exp\left(-\xi \frac{\mu}{\widetilde{\lambda}} \Phi\left(-\xi \frac{\alpha}{\widetilde{\lambda}}, 1, \frac{\widetilde{\lambda} - \lambda}{\widetilde{\lambda}}\right)\right)\right], \tag{5.2.2}$$

where Φ *is the Lerch transcendent defined as* $\forall\ \Re(s) > 0,\ \Re(a) > 0$

$$\Phi(z, s, a) = \frac{1}{\Gamma(s)} \int_0^\infty \frac{t^{s-1} e^{-at}}{1 - z e^{-t}} dt.$$

Equation (5.2.1) is obtained in [24], Proposition 2 for the exact same context, but similar reasoning can be found in [22, 8]. For the sake of completeness, the proof of Equation (5.2.2) is given in Appendix, inspired by the proof of [24], Corollary 2. Now we have all the material to state the result of the first order asymptotics for the first selective mutant subpopulation size.

Theorem 5.2.3 (First-order asymptotics for the first selective mutant subpopulation size)**.** *Let* (V, E, L) *be a finite labelled mono-directional graph with* $k + 1$ *vertices, for some* $k \geq 1$*, as defined in* (5.1.1)*. Assume that the first selective mutation on the graph is the last one as in* (5.1.2)*. Let* $0 < T_1 < T_2$*. Using the mathematical definition of the model given in Section 5.3, see* (5.3.2)*,* (5.3.3)*,* (5.3.4)*,* (5.3.5)*,* (5.3.6)*,* (5.3.7)*,* (5.3.8)*, and* (5.3.9)*, there exists two random variables* $Z_{k,1}^\infty$ *and* $Z_{k,2}^\infty$ *properly defined in Propositions 5.4.3 and 5.4.4 respectively, that are independent conditioning on* W *(properly defined in* (5.3.10)*) such that*

$$\left(\frac{Z_k^{(n)}\left((t(k) + t)\log(n)\right)}{n^{t\lambda_k} \log^{\theta(k-1)\frac{\lambda_k}{\lambda_0}}(n)}\right)_{t\in[T_1,T_2]} \underset{n\to\infty}{\longrightarrow} Z_{k,1}^\infty + Z_{k,2}^\infty,$$

in probability in $L^\infty([T_1,T_2])$. *The random variable* $Z_{k,1}^\infty$ *follows a compound Luria-Delbrück distribution with parameters*

$$\left(\lambda_0, \alpha_k, \beta_k, \frac{2\alpha_{k-1}\mu_{k-1}}{\lambda_0} w_{k-1} W\right)$$

where $w_{k-1} := \widetilde{w}_{(0,\cdots,k-1)}(t(k))$, *and associated with* $\mathcal{U}_k$, *the distribution of the almost sure large limit of the natural martingale associated to a birth and death branching process with rates* α_k *and* β_k *respectively. More precisely,*

$$\mathcal{U}_k := Ber\left(\frac{\lambda_k}{\alpha_k}\right) \otimes Exp\left(\frac{\lambda_k}{\alpha_k}\right).$$

The random variable $Z_{k,2}^\infty$ *follows the distribution of the asymptotic large time limit of the population defined by the approximate model from Definition 5.2.2 with parameters*

$$(\lambda_0, 2\alpha_{k-1}\mu_{k-1}w_{k-1}W, \alpha_k, \beta_k).$$

For any other mathematical description, the convergence is at least in distribution in $\mathbb{D}([T_1,T_2])$.

The proof of this theorem involves employing distinct proof techniques for the three subpopulations $Z_{k,1}^{(n)}$, $Z_{k,2}^{(n)}$ and $Z_{k,3}^{(n)}$ that are summed up here, alongside some heuristics and biological interpretations:

- **Heuristics concerning the convergence of $Z_{k,1}^{(n)}$:** The random variable $Z_{k,1}^\infty$ is the asymptotic limit of the well-renormalized subpopulation $Z_{k,1}^{(n)}$ composed of the mutant cell lineages issued from mutational events before time $t_-^{(n)}(k)$. The analysis is made in two steps. First we adapt a proof from [8] or [10] to demonstrate that at time $t_-^{(n)}(k)$, the number of mutant cells of trait k asymptotically follows a Luria-Delbrück distribution with parameters $\left(\lambda_0, \alpha_k, \beta_k, \frac{2\alpha_{k-1}\mu_{k-1}}{\lambda_0} w_{k-1} W\right)$, which has a infinite expectation. Such proof uses the Poissonian structure of the process of mutation from trait $k-1$ to trait k as well as that conditioning on the number of mutant cells generating due to mutational events up to time $t_-^{(n)}(k)$, the vector of the remaining time for each of the mutant clone to grow up to time $t_-^{(n)}(k)$ is asymptotically the order statistic of a vector of i.i.d. exponential random variables with parameter λ_0. Subsequently, we establish control over the size of all mutant cell lineages present at time $t_-^{(n)}(k)$ by time $(t(k)+t)\log(n)$ using a martingale approach. These lineages experience exponential growth at rate λ_k over a duration of time $(t(k)+t)\log(n) - t_-^{(n)}(k) = t\log(n) + \frac{\theta(k-1)}{\lambda_0}\log(\log(n))$, resulting in a typical size of order $\mathcal{O}\left(n^{t\lambda_k}\log^{\theta(k-1)\frac{\lambda_k}{\lambda_0}}(n)\right)$. This subpopulation corresponds to the lineages of the possible mutant cells generated due to random mutational events appeared before the deterministic time $t_-^{(n)}(k)$. Due to the selectiveness of the mutation, their lineages have at time $t_-^{(n)}(k)$ asymptotically infinite expectancies. Moreover, the fact that they asymptotically contribute to the size order of trait k agrees with the natural intuition behind considering a selective mutation.

- **Heuristics concerning the convergence of $Z_{k,2}^{(n)}$:** Dealing with the appropriate normalization of $Z_{k,2}^{(n)}$ is more complex. For the mutant cells generated between times $t_-^{(n)}(k)$ and $t(k)\log(n)$, and their lineages, we approximate this subpopulation with the simpler model

of Definition 5.2.2 with the adequate parameters that are $(\lambda_0, 2\alpha_{k-1}\mu_{k-1}w_{k-1}W, \alpha_k, \beta_k)$, to reduce stochasticity and facilitate obtaining its asymptotic limit at time $t(k)\log(n)$. We rigorously establish by a martingale argument using an L^1 convergence that at the first-order, this subpopulation behaves asymptotically akin to the considered approximate model up to time $t(k)\log(n)$. In particular, such approximate model grows exponentially fast at rate λ_k during a time $t(k)\log(n) - t_-^{(n)}(k) = \frac{\theta(k-1)}{\lambda_0}\log(\log(n))$, implying that at time $t(k)\log(n)$ the subpopulation $Z_{k,2}^{(n)}$ is of order $\mathcal{O}\left(\log^{\theta(k-1)\frac{\lambda_k}{\lambda_0}}(n)\right)$. Then we obtain its asymptotic limit at time $(t(k)+t)\log(n)$ using a law of large numbers argument. More precisely, after time $t(k)\log(n)$ there are no longer mutational event from trait $k-1$ to trait k that are counted in this subpopulation anymore. Because at this time we have an order of $\mathcal{O}\left(\log^{\theta(k-1)\frac{\lambda_k}{\lambda_0}}(n)\right)$ mutants cells, and that their lineages grow exponentially fast at rate λ_k during a time $t\log(n)$ it resulted in a subpopulation of order $\mathcal{O}\left(\log^{\theta(k-1)\frac{\lambda_k}{\lambda_0}}(n)n^{t\lambda_k}\right)$ at time $(t(k)+t)\log(n)$. This subpopulation encapsulates the idea that an asymptotically infinite number of lineages (but not all of them) significantly contribute to the size of the selective mutant trait subpopulation, despite being characterised by a finite mean random variable. However, these infinite number of lineages are all produced on the slower $\log(\log(n))-$accelerated time scale, meaning that they are asymptotically all condensed at the same time point for the $\log(n)-$accelerated time scale. Moreover, understanding the equation governing the Laplace transform of this limiting random variable, see Definition 5.2.2, provides valuable insights into the actual asymptotic contributions of these lineages, depending on their generation timing.

- **Heuristics concerning the convergence of $Z_{k,3}^{(n)}$:** Dealing with $Z_{k,3}^{(n)}$, the subpopulation containing solely mutant cells generated after time $t(k)\log(n)$ along with their lineages, is done using a similar martingale method as the one used in Chapter 4, to get that it scales as $n^{t\lambda_k}\log^{\theta(k-1)}(n)$. Hence, asymptotically it becomes negligible and vanishes in the limit. Heuristically, the selectiveness of the trait implies that only the one generated around time $t(k)\log(n)$ are asymptotically contributing at time $(t(k)+t)\log(n)$. At time $t(k)\log(n)$, the number of mutant cells of trait k generated from mutational events from trait $k-1$ is of order $\mathcal{O}\left(\log^{\theta(k-1)}(n)\right)$. Then their lineages are growing exponentially fast at rate λ_k during a time $t\log(n)$, giving a size of order $\mathcal{O}\left(\log^{\theta(k-1)}(n)n^{t\lambda_k}\right)$ at time $(t(k)+t)\log(n)$. Biologically meaning, after time $t(k)\log(n)$ the generated selective mutant cells and their lineages have too much time delay compared to the first generated mutant cells such that their contribution appears negligible.

Before going into the mathematical definition of the model used to prove Theorem 5.2.3, we are making some interesting remarks on such result.

Remark 5.2.4.

- ***First-order asymptotics:*** *In Chapter 4, Theorem 4.2.9, one obtains that the stochastic exponent associated to the mutant subpopulation of trait k asymptotically converges, conditioning on $\{W > 0\}$, to $t\lambda_k$, which corresponds to the power of n captured in Theorem 5.2.3. Notably, such result means that in the case of the selective mutant trait, growth stems not from the mutational process but rather from the inherent expansion of selective mutant cell lineages. This stands in contrast to deleterious mutations, where growth is actually driven by the mutational process, and neutral mutations, where growth represents a*

balance between mutational process and inherent lineage expansions. Moreover the power of $\log(n)$ *is also captured by this theorem, which is a generalization of Theorem 4.2.7 allowing for a first selective mutation. An interesting pattern for such power occurs. For deleterious and neutral mutations, we previously obtained that a neutral one increases such power by* 1 *where a deleterious one has no effect. For a selective mutation, it is a bit more intricate: the power obtained for the previous trait is accelerated by the ratio between the growth rate of the selective trait and the actual one of the previous trait (that is the one of the wild-type subpopulation). In particular if there are only deleterious mutations before the first selective one, then the power of* $\log(n)$ *stays equal to 0 for the first selective trait.*

- ***Time dependence:*** *An interesting characteristic of the limit for the selective mutant trait, that contrasts with neutral and deleterious mutations, is its independence from the time parameter* t*. In Theorem 4.2.7 we have shown that alongside a mono-directional graph composed of neutral or deleterious mutations only, the asymptotic limit of a trait depends on the time parameter* t *as a polynomial function of degree the number of neutral mutations up to this considered trait. With Theorem 5.2.3 we show that this time dependence is lost when the first selective mutant trait is encountered in the mutational pathway. In a certain sense, one can say that it resets the time dependence. Observing such differences depending on the type of the mutation seems more than intuitively natural. Indeed, with a neutral mutation, the balance between the mutational process and the inherent growth makes clear that the lineages have equal contributions on the size order, generating a dependence on the time parameter via an extra integral operator from the time dependence of the previous trait. For a deleterious mutation, driven by the mutational process, the time dependence comes from taking into account only clones generated at the considered time, meaning that the dependence is identically forwarded from the one of the previous trait. For a selective mutation, the inherent growth prevails the mutational process, implying that only the lineages generated around the typical time to observe such mutant trait are asymptotically contributing, and so all the time dependency from the previous trait is completely lost.*

- ***Stochasticity of the limit:*** *In the case of neutral (and deleterious) cancer evolution, we obtained in Theorem 4.2.7 that the stochasticity captured by the first order asymptotics of the mutant subpopulation sizes is fully given by the random variable* W*, which quantifies the large time stochasticity associated to the primary subpopulation. In Remark 4.2.8 point 3., we argued that such result could not be obtained for a selective mutation, explaining the failure of the martingale approach. In Theorem 5.2.3 we show that the latter conjecture is true. In the limiting random variables,* $Z^{\infty}_{k,1}$ *and* $Z^{\infty}_{k,2}$*, the stochasticity of* W *is present, but another layer of stochasticity is also given due to the growth of the lineages of the selective mutant cells. The independence, when conditioning on* W*, of these two limiting random variables comes from the fact that they correspond to the asymptotic limits of two subpopulations of mutant cell lineages that are independent due to the mutational process which is Poissonian.*

- ***Dealing with the next mutant trait:*** *Asymptotically we obtain that the limiting random variable* $Z^{\infty}_{k,1}$ *has an infinite expectation, coming from the Luria-Delbrück part of the compound random variable, a consequence of the selectiveness of the considered mutation. Such specificity entails mathematical difficulties, explaining why the martingale approach developed in Chapter 4 fails. Indeed, taking an expectation balances the probability of an event with the number of cells seen if the event occurs. Consequently, unlikely events, such as rapid mutations to the selective trait, can lead to extremely large population sizes far exceeding typical*

asymptotic sizes implying the expectation order is given by these unlikely event. Moreover, it hampers from dealing with the next mutant trait. In particular the cells of trait $k+1$, and their lineages, produced by $Z_{k,1}^{(n)}$ are, for now, untractable. More specifically, dealing with the latter subpopulation of trait $k+1$ using the same approach as for capturing the asymptotic limit of trait k, the martingale approach developed to show that $Z_{k+1,2}^{(n)}$ and the adapted approximate model are sufficiently close to transfer the limiting behaviour from the approximate model to $Z_{k+1,2}^{(n)}$, fails. Indeed, a L^1 approach is needed for this step. Instead of approximating a mutant trait by its asymptotic limit to tackle the next mutant trait, a potential solution to deal with such difficulty is to directly deal with all the mutant subpopulations together without using the approximations given by the limits at each step. This procedure would normally allow for overstepping the infinite expectation of $Z_{k,1}^{\infty}$.

- ***Two steps limit:*** *Instead of being interested in the double limit given by the large population and power law mutation rates regime, another interesting approach is to consider a two steps asymptotics where first the large time limit is obtained with fixed mutation rates and then the limit when these rates tend to 0 is obtained. In [24], Nicholson, Cheek and Antal are studying this two steps limit. An interesting fact is that the results are different, meaning that there is no equivalence between performing the double limit directly or in two steps. But some similarities are actually obtained for the structure of the asymptotic limits. Notably, both asymptotic results can be decomposed into the product of a time-independent random variable (which are different) and a simple time-dependent deterministic function controlled by the growth rate of the selective mutant trait (with the same deterministic function but taken at different times). More precisely, in their case, the stochasticity of the asymptotic limit that is captured (see [24], Theorem 1) is a Mittag-Leffler distribution with tail parameter the ratio between the growth rate of the wild-type subpopulation and the one of the first selective mutant trait, and a scale parameter satisfying a specific equation depending on the birth, growth, and mutation rates of the preceeding mutant traits. Such asymptotic random variable is not captured by Theorem 5.2.3, as mentioned above. Moreover, the deterministic function is in both case the exponential function whose growth rate is the one of the selective mutant trait, but taken at different times. In their case, when looking at the limit when $t \to \infty$, the time that is taken for the deterministic function is actually t, where in our case, at time $(t(k)+t)\log(n)$ it is $t\log(n)$ that is taken.*

5.3 Mathematical definition of the model

In this section we mathematically construct the model in the case of the finite labelled mono-directional graph of (5.1.1). The sequence $\left(\left(Z_i^{(n)}\right)_{i\in\{0,\cdots,k\}}\right)_{n\in\mathbb{N}}$ is mathematically constructed using independent Poisson Point Measures (PPMs) and birth and death branching processes. Such construction differs from the one given in Chapter 4, Subsection 4.3, but remains obviously equal in law. More specifically, it depends on whether $\theta(k-1)=0$ or $\theta(k-1)\geq 1$, biologically meaning whether there are neutral mutation on the trait space before the first selective one or not. We start by the more complex, but more interesting case that is $\theta(k-1)\geq 1$.

5.3.1 Case with at least one neutral mutation

Assume in this subsection that $\theta(k-1) \geq 1$. In particular this case implies that $k \geq 2$, because there is at least one neutral mutation before encountering the first selective one. Consequently we have that $t_-^{(n)}(n) < t(k)\log(n)$. Hence, introduce the difference between these two times as

$$\ell_-^{(n)}(k) := t(k)\log(n) - t_-^{(n)}(k) = \frac{\theta(k-1)}{\lambda_0}\log(\log(n)) > 0. \tag{5.3.1}$$

The $k-2$ first subpopulations $\left(Z_i^{(n)}\right)_{i\in\{0,\cdots,k-2\}}$ are constructed exactly as in Chapter 4, Section 4.3. More precisely, let

$$\begin{aligned}&Q_0^b(ds,d\theta), Q_0^d(ds,d\theta),\\ &(Q_i(ds,d\theta))_{i\in\{1,\cdots,k-2\}}, (N_i(ds,d\theta))_{i\in\{0,\cdots,k-2\}}, \text{ and } (Q_i^m(ds,d\theta))_{i\in\{0,\cdots,k-2\}},\end{aligned}$$

be independent PPMs with intensity $dsd\theta$. The subpopulation of wild-type cells is

$$Z_0^{(n)}(t) := 1 + \int_0^t\int_{\mathbb{R}^+} \mathbb{1}_{\left\{\theta\leq\alpha_0^{(n)}Z_0^{(n)}(s^-)\right\}}Q_0^b(ds,d\theta) - \int_0^t\int_{\mathbb{R}^+}\mathbb{1}_{\left\{\theta\leq\beta_0 Z_0^{(n)}(s^-)\right\}}Q_0^d(ds,d\theta) - H_0^{(n)}(t), \tag{5.3.2}$$

and for all $i \in \{1,\cdots,k-2\}$

$$\begin{aligned}Z_i^{(n)}(t) := &\int_0^t\int_{\mathbb{R}^+}\left(\mathbb{1}_{\left\{\theta\leq\alpha_i^{(n)}Z_i^{(n)}(s^-)\right\}} - \mathbb{1}_{\left\{\alpha_i^{(n)}Z_i^{(n)}(s^-)\leq\theta\leq\left(\alpha_i^{(n)}+\beta_i\right)Z_i^{(n)}(s^-)\right\}}\right)Q_i(ds,d\theta)\\ &+ K_{i-1}^{(n)}(t) + 2H_{i-1}^{(n)}(t) - H_i^{(n)}(t),\end{aligned} \tag{5.3.3}$$

where for all $i \in \{0,\cdots,k-2\}$

$$\begin{aligned}K_i^{(n)}(t) &:= \int_0^t\int_{\mathbb{R}^+}\mathbb{1}_{\left\{\theta\leq 2\alpha_i\mu_i^{(n)}\left(1-\mu_i^{(n)}\right)Z_i^{(n)}(s^-)\right\}}N_i(ds,d\theta),\\ H_i^{(n)}(t) &:= \int_0^t\int_{\mathbb{R}^+}\mathbb{1}_{\left\{\theta\leq\alpha_i\left(\mu_i^{(n)}\right)^2 Z_i^{(n)}(s^-)\right\}}Q_i^m(ds,d\theta).\end{aligned} \tag{5.3.4}$$

The subpopulation $Z_{k-1}^{(n)}$ is also constructed using PPMs, but the process denoted by $H_{k-1}^{(n)}$ counting the number of cell divisions of trait $k-1$ where both daughter cells are mutating to trait k is constructed in a more complex way. More precisely, we are separating this process into a sum of three different processes, where each of them are contributing for a specific time interval: the first one $H_{k-1,1}^{(n)}$ up to time $t_-^{(n)}(k)$, the second one $H_{k-1,2}^{(n)}$ between $t_-^{(n)}(k)$ and $t(k)\log(n)$, and the third one $H_{k-1,3}^{(n)}$ after time $t(k)\log(n)$. Then, let $Q_{k-1}(ds,d\theta)$, $\left(Q_{k-1,i}^m(ds,d\theta)\right)_{i\in\{1,2,3\}}$ be independent PPMs with intensity $dsd\theta$, that are also independent from $\left(Z_i^{(n)}\right)_{i\in\{0,\cdots,k-2\}}$. The subpopulation

$Z_{k-1}^{(n)}$ is

$$
\begin{aligned}
Z_{k-1}^{(n)}(t) &:= K_{k-2}^{(n)}(t) + 2H_{k-2}^{(n)}(t) - H_{k-1}^{(n)}(t) \\
&\quad + \int_0^t \int_{\mathbb{R}^+} \left(\mathbb{1}_{\left\{\theta \le \alpha_{k-1}^{(n)} Z_{k-1}^{(n)}(s^-)\right\}} - \mathbb{1}_{\left\{\alpha_{k-1}^{(n)} Z_{k-1}^{(n)}(s^-) \le \theta \le \left(\alpha_{k-1}^{(n)} + \beta_{k-1}\right) Z_{k-1}^{(n)}(s^-)\right\}} \right) Q_{k-1}(ds, d\theta) \\
H_{k-1}^{(n)}(t) &:= H_{k-1,1}^{(n)}(t) + H_{k-1,2}^{(n)}\left(\left(t - t_-^{(n)}(k)\right) \vee 0\right) + H_{k-1,3}^{(n)}\left((t - t(k)\log(n)) \vee 0\right), \\
H_{k-1,1}^{(n)}(t) &:= \int_0^{t \wedge t_-^{(n)}(k)} \int_{\mathbb{R}^+} \mathbb{1}_{\left\{\theta \le \alpha_{k-1}\left(\mu_{k-1}^{(n)}\right)^2 Z_{k-1}^{(n)}(s^-)\right\}} Q_{k-1,1}^m(ds, d\theta), \\
H_{k-1,2}^{(n)}(t) &:= \int_0^{t \wedge \ell_-^{(n)}(k)} \int_{\mathbb{R}^+} \mathbb{1}_{\left\{\theta \le \alpha_{k-1}\left(\mu_{k-1}^{(n)}\right)^2 Z_{k-1}^{(n)}\left(t_-^{(n)}(k)+s\right)\right\}} Q_{k-1,2}^m(ds, d\theta), \\
H_{k-1,3}^{(n)}(t) &:= \int_0^t \int_{\mathbb{R}^+} \mathbb{1}_{\left\{\theta \le \alpha_{k-1}\left(\mu_{k-1}^{(n)}\right)^2 Z_{k-1}^{(n)}(t(k)\log(n)+s)\right\}} Q_{k-1,3}^m(ds, d\theta).
\end{aligned} \tag{5.3.5}
$$

The construction of the subpopulation $Z_k^{(n)}$ is the most involved of all. In order to use different proof techniques, we are using PPMs, birth and death branching processes and a Poisson process. In a same way as for the construction of $H_{k-1}^{(n)}$, the process $K_{k-1}^{(n)}$, counting the number of cell divisions of trait $k-1$ where exactly one of the daughter cells mutates to trait k, is constructed as a sum of three processes, each of them contributing specifically for one of the aforementioned time interval. Then, let $N_{k-1,1}$ be a Poisson process with intensity 1, $(N_{k-1,i}(ds, d\theta))_{i\in\{2,3\}}$, $Q_{k,2}^b(ds, d\theta)$, $Q_{k,2}^d(ds, d\theta)$, and $Q_{k,3}(ds, d\theta)$ be PPMs with intensity $dsd\theta$ and $(U_i, Y_i, X_{1,i}, X_{2,i}, V_i)_{i\in\mathbb{N}}$ be birth-death branching processes with rates α_k and β_k all starting with 1 individual at time $t = 0$ almost surely. All such processes are independent from each others and independent from $\left(Z_i^{(n)}\right)_{i\in\{0,\cdots,k-1\}}$. The subpopulation $Z_k^{(n)}$ and the process $K_{k-1}^{(n)}$ are constructed as

$$
\begin{aligned}
Z_k^{(n)}(t) &:= Z_{k,1}^{(n)}(t) + Z_{k,2}^{(n)}\left(\left(t - t_-^{(n)}(k)\right) \vee 0\right) + Z_{k,3}^{(n)}\left((t - t(k)\log(n)) \vee 0\right), \\
K_{k-1}^{(n)}(t) &:= K_{k-1,1}^{(n)}(t) + K_{k-1,2}^{(n)}\left(\left(t - t_-^{(n)}(k)\right) \vee 0\right) + K_{k-1,3}^{(n)}\left((t - t(k)\log(n)) \vee 0\right).
\end{aligned} \tag{5.3.6}
$$

The subpopulation $Z_{k,1}^{(n)}$ satisfies

$$
\begin{aligned}
Z_{k,1}^{(n)}(t) &:= \sum_{i=1}^{B_{k-1,1}^{(n)}(t)} U_i\left(\left(t - t_-^{(n)}(k)\right) \vee 0\right), \\
B_{k-1}^{(n)}(t) &:= \sum_{i=1}^{K_{k-1,1}^{(n)}(t)} Y_i\left(\left(t - T_i^{(n)}\right) \wedge \left(t_-^{(n)}(k) - T_i^{(n)}\right)\right), \\
&\quad + \sum_{i=1}^{H_{k-1,1}^{(n)}(t)} X_{i,1}\left(\left(t - S_i^{(n)}\right) \wedge \left(t_-^{(n)}(k) - S_i^{(n)}\right)\right) + X_{i,2}\left(\left(t - S_i^{(n)}\right) \wedge \left(t_-^{(n)}(k) - S_i^{(n)}\right)\right),
\end{aligned} \tag{5.3.7}
$$

$$K^{(n)}_{k-1,1}(t) := N_{k-1,1}\left(\int_0^{t\wedge t^{(n)}_-(k)} 2\alpha_{k-1}\mu^{(n)}_{k-1}\left(1-\mu^{(n)}_{k-1}\right)Z^{(n)}_{k-1}(s)ds\right),$$
$$T^{(n)}_i := \inf\left\{t\geq 0 : K^{(n)}_{k-1,1}\geq i\right\},$$
$$S^{(n)}_i := \inf\left\{t\geq 0 : H^{(n)}_{k-1,1}\geq i\right\}.$$

The subpopulation $Z^{(n)}_{k,2}$ is defined as

$$Z^{(n)}_{k,2}(t) := \sum_{i=1}^{B^{(n)}_{k-1,2}(t)} V_i\left(\left(t-\ell^{(n)}_-(k)\right)\vee 0\right), \tag{5.3.8}$$
$$B^{(n)}_{k-1,2}(t) := K^{(n)}_{k-1,2}(t) + 2H^{(n)}_{k-1,2}(t) + \int_0^{t\wedge\ell^{(n)}_-(k)}\int_{\mathbb{R}^+}\mathbb{1}_{\left\{\theta\leq\alpha_k Z^{(n)}_{k,2}(s)\right\}}Q^b_{k,2}(ds,d\theta)$$
$$- \int_0^{t\wedge\ell^{(n)}_-(k)}\int_{\mathbb{R}^+}\mathbb{1}_{\left\{\theta\leq\beta_k Z^{(n)}_{k,2}(s)\right\}}Q^d_{k,2}(ds,d\theta),$$
$$K^{(n)}_{k-1,2}(t) := \int_0^{t\wedge\ell^{(n)}_-(k)}\int_{\mathbb{R}^+}\mathbb{1}_{\left\{\theta\leq 2\alpha_{k-1}\mu^{(n)}_{k-1}\left(1-\mu^{(n)}_{k-1}\right)Z^{(n)}_{k-1}\left(t^{(n)}_-(k)+s\right)\right\}}N_{k-1,2}(ds,d\theta).$$

And finally the subpopulation $Z^{(n)}_{k,3}$ is constructed as

$$Z^{(n)}_{k,3}(t) := K^{(n)}_{k-1,3}(t) + 2H^{(n)}_{k-1,3}(t) \tag{5.3.9}$$
$$+\int_0^t\int_{\mathbb{R}^+}\mathbb{1}_{\left\{\theta\leq\alpha_k Z^{(n)}_{k,3}(s)\right\}} - \mathbb{1}_{\left\{\alpha_k Z^{(n)}_{k,3}(s)\leq\theta\leq(\alpha_k+\beta_k)Z^{(n)}_{k,3}(s)\right\}}Q_{k,3}(ds,d\theta),$$
$$K^{(n)}_{k-1,3}(t) := \int_0^t\int_{\mathbb{R}^+}\mathbb{1}_{\left\{\theta\leq 2\alpha_{k-1}\mu^{(n)}_{k-1}\left(1-\mu^{(n)}_{k-1}\right)Z^{(n)}_{k-1}(t(k)\log(n)+s)\right\}}N_{k-1,3}(ds,d\theta).$$

5.3.2 Case without any neutral mutation

Assume in this subsection that $\theta(k-1)=0$. For the simplicity, we give a construction when $k\geq 2$, although the case $k=1$ is included in this specific case $\theta(k-1)=0$, and can be mathematically constructed in a similar way. The specificity of this case is that $t^{(n)}_-(k) = t(k)\log(n)$. Let

$$Q^b_0(ds,d\theta), Q^d_0(ds,d\theta),$$
$$(Q_i(ds,d\theta))_{i\in\{1,\cdots,k-2\}}, (N_i(ds,d\theta))_{i\in\{0,\cdots,k-2\}}, \text{ and } (Q^m_i(ds,d\theta))_{i\in\{0,\cdots,k-2\}},$$

be independent PPMs with intensity $dsd\theta$, and construct the wild-type subpopulations $Z^{(n)}_0$, the mutant subpopulations $Z^{(n)}_i$ for $i\in\{1,\cdots,k-2\}$ and the processes $H^{(n)}_i$ and $K^{(n)}_i$ for $i\in\{0,\cdots,k-2\}$

as in the case $\theta(k-1) \geq 1$. The subpopulation $Z_{k-1}^{(n)}$ is constructed as

$$
\begin{aligned}
Z_{k-1}^{(n)}(t) &:= \int_0^t \int_{\mathbb{R}^+} \left(\mathbb{1}_{\left\{\theta \leq \alpha_{k-1}^{(n)} Z_{k-1}^{(n)}(s^-)\right\}} - \mathbb{1}_{\left\{\alpha_{k-1}^{(n)} Z_{k-1}^{(n)}(s^-) \leq \theta \leq \left(\alpha_{k-1}^{(n)} + \beta_{k-1}\right) Z_{k-1}^{(n)}(s^-)\right\}} \right) Q_{k-1}(ds, d\theta) \\
&\quad + K_{k-2}^{(n)}(t) + 2H_{k-2}^{(n)}(t) - H_{k-1}^{(n)}(t), \\
H_{k-1}^{(n)}(t) &:= H_{k-1,1}^{(n)}(t) + H_{k-1,2}^{(n)}\left((t - t(k)\log(n)) \vee 0\right), \\
H_{k-1,1}^{(n)}(t) &:= \int_0^{t \wedge t(k)\log(n)} \int_{\mathbb{R}^+} \mathbb{1}_{\left\{\theta \leq \alpha_{k-1}\left(\mu_{k-1}^{(n)}\right)^2 Z_{k-1}^{(n)}(s^-)\right\}} Q_{k-1,1}^m(ds, d\theta), \\
H_{k-1,2}^{(n)}(t) &:= \int_0^t \int_{\mathbb{R}^+} \mathbb{1}_{\left\{\theta \leq \alpha_{k-1}\left(\mu_{k-1}^{(n)}\right)^2 Z_{k-1}^{(n)}(t(k)\log(n)+s)\right\}} Q_{k-1,2}^m(ds, d\theta).
\end{aligned}
$$

Let $N_{k-1,1}$ be a Poisson process with intensity 1, $N_{k-1,2}(ds, d\theta)$, and $Q_{k,2}(ds, d\theta)$ be PPMs with intensity $dsd\theta$ and $(U_i, Y_i, X_{1,i}, X_{2,i}, V_i)_{i\in\mathbb{N}}$ be birth-death branching processes with rates α_k and β_k all starting with 1 individual at time $t = 0$ almost surely. All such processes are independent from each others and independent from $\left(Z_i^{(n)}\right)_{i\in\{0,\cdots,k-1\}}$. We write the process $Z_k^{(n)}$ as the sum of two sub-processes $Z_{k,1}^{(n)}$ and $Z_{k,2}^{(n)}$. For all $t > 0$ define

$$
\begin{aligned}
Z_k^{(n)}(t) &:= Z_{k,1}^{(n)}(t) + Z_{k,2}^{(n)}\left((t - t(k)\log(n)) \vee 0\right), \\
Z_{k,1}^{(n)}(t) &:= \sum_{i=1}^{B_{k-1}^{(n)}(t)} U_i \left((t - t(k)\log(n)) \vee 0\right), \\
B_{k-1}^{(n)}(t) &:= \sum_{i=1}^{K_{k-1,1}^{(n)}(t)} Y_i \left(\left(t - T_i^{(n)}\right) \wedge \left(t(k)\log(n) - T_i^{(n)}\right)\right) \\
&\quad + \sum_{i=1}^{H_{k-1,1}^{(n)}(t)} X_{i,1}\left(\left(t - S_i^{(n)}\right) \wedge \left(t(k)\log(n) - S_i^{(n)}\right)\right) + X_{i,2}\left(\left(t - S_i^{(n)}\right) \wedge \left(t(k)\log(n) - S_i^{(n)}\right)\right),
\end{aligned}
$$

$$
\begin{aligned}
K_{k-1,1}^{(n)}(t) &:= N_{k-1,1}\left(\int_0^{t \wedge t_-^{(n)}(k)} 2\alpha_{k-1}\mu_{k-1}^{(n)}\left(1 - \mu_{k-1}^{(n)}\right) Z_{k-1}^{(n)}(s)ds\right), \\
T_i^{(n)} &:= \inf\left\{t \geq 0 : K_{k-1,1}^{(n)} \geq i\right\}, \\
S_i^{(n)} &:= \inf\left\{t \geq 0 : H_{k-1,1}^{(n)} \geq i\right\}, \\
Z_{k,2}^{(n)}(t) &:= K_{k-1,2}^{(n)}(t) + 2H_{k-1,2}^{(n)}(t) \\
&\quad + \int_0^t \int_{\mathbb{R}^+} \mathbb{1}_{\left\{\theta \leq \alpha_k Z_{k,2}^{(n)}(s)\right\}} - \mathbb{1}_{\left\{\alpha_k Z_{k,2}^{(n)}(s) \leq \theta \leq (\alpha_k + \beta_k) Z_{k,2}^{(n)}(s)\right\}} Q_{k,2}(ds, d\theta),
\end{aligned}
$$

$$K_{k-1,2}^{(n)}(t) := \int_0^t \int_{\mathbb{R}^+} \mathbb{1}_{\left\{\theta \leq 2\alpha_{k-1}\mu_{k-1}^{(n)}\left(1-\mu_{k-1}^{(n)}\right)Z_{k-1}^{(n)}(t(k)\log(n)+s)\right\}} N_{k-1,2}(ds, d\theta),$$

where U_i are birth and death branching processes with rates α_k and β_k, as well as Y_i and $X_{i,1}$ and $X_{i,2}$. $T_i^{(n)}$, respectively $S_i^{(n)}$, are the jumping times of the process $K_{k-1,1}^{(n)}$, respectively $H_{k-1,1}^{(n)}$.

5.3.3 Approximate primary subpopulation

Let $(Z_0(t))_{t\in\mathbb{R}^+}$ be the birth-death branching process with rates α_0 and β_0 respectively, constructed in the following way

$$Z_0(t) := 1 + \int_0^t \int_{\mathbb{R}^+} \mathbb{1}_{\{\theta \leq \alpha_0 Z_0(s^-)\}} Q_0^b(ds, d\theta) - \int_0^t \int_{\mathbb{R}^+} \mathbb{1}_{\{\theta \leq \beta_0 Z_0(s^-)\}} Q_0^d(ds, d\theta).$$

Notice that with such a construction it immediately follows, independently of the value of $\theta(k-1)$, the monotone coupling

$$\forall t \geq 0, Z_0^{(n)}(t) \leq Z_0(t) \ a.s.$$

Denote by

$$W := \lim_{t\to\infty} e^{-\lambda_0 t} Z_0(t), \tag{5.3.10}$$

the almost sure limit of the positive martingale $\left(e^{-\lambda_0 t} Z_0(t)\right)_{t\in\mathbb{R}^+}$, whose law is

$$W \overset{law}{=} Ber\left(\frac{\lambda_0}{\alpha_0}\right) \otimes Exp\left(\frac{\lambda_0}{\alpha_0}\right),$$

see [22], Section 1.1, or [94], Theorem 1.

5.4 Proofs of the first-order asymptotics of the first-selective mutant subpopulation size (Theorem 5.2.3)

This subsection is devoted to the proof of Theorem 5.2.3. To this aim we start by giving an interesting result on the speed of convergence for the martingale associated to a supercritical birth and death branching process to its asymptotic large time limit.

Lemma 5.4.1. *Let $(Z(t))_{t\geq 0}$ be a birth and death branching process with rates α and β satisfying $Z(0) = 1$ almost surely. Denote by $\lambda := \alpha - \beta$ its growth rate and assume $\lambda > 0$. Let W the almost sure large time limit of the martingale $\left(e^{-\lambda t} Z(t)\right)_{t\geq 0}$. Let $f_1(n) \underset{n\to\infty}{\longrightarrow} \infty$ and $h(n) \underset{n\to\infty}{\longrightarrow} \infty$. Denote by $\varepsilon_n := e^{-\lambda \frac{f_1(n)}{2}} \sqrt{h(n)}$. Then we have*

$$\mathbb{P}\left(\sup_{t\in[f_1(n),\infty]} \left|e^{-\lambda t} Z(t) - W\right| \geq \varepsilon_n\right) \leq C h^{-1}(n),$$

where $C = 20\frac{\alpha+\beta}{\lambda}$.

Proof. Let $f_2(m)$ be a function tending to infinity. We have

$$\mathbb{P}\left(\sup_{t\in[f_1(n),f_2(m)]}\left|e^{-\lambda t}Z(t)-W\right|\geq\varepsilon_n\right)\leq\mathbb{P}\left(\sup_{t\in[f_1(n),f_2(m)]}\left|e^{-\lambda t}Z(t)-e^{-\lambda f_1(n)}Z(f_1(n))\right|\geq\frac{\varepsilon_n}{2}\right)$$
$$+\mathbb{P}\left(\left|e^{-\lambda f_1(n)}Z(f_1(n))-W\right|\geq\frac{\varepsilon_n}{2}\right).$$

Then using Doob's inequality to the martingale $\left(e^{-\lambda t}Z(t)-e^{-\lambda f_1(n)}Z(f_1(n))\right)_{t\geq f_1(n)}$ and Chebyshev's inequality for the random variable $e^{-\lambda f_1(n)}Z(f_1(n))-W$, we get

$$\begin{aligned}
&\mathbb{P}\left(\sup_{t\in[f_1(n),f_2(m)]}\left|e^{-\lambda t}Z(t)-W\right|\geq\varepsilon_n\right)\\
&\leq 4\varepsilon_n^{-2}\left(4\mathbb{E}\left[\left(e^{-\lambda f_2(m)}Z(f_2(m))-e^{-\lambda f_1(n)}Z(f_1(n))\right)^2\right]+\mathbb{E}\left[\left(e^{-\lambda f_1(n)}Z(f_1(n))-W\right)^2\right]\right)\\
&\leq 20\varepsilon_n^{-2}\int_{f_1(n)}^{\infty}(\alpha+\beta)\,e^{-2\lambda s}\mathbb{E}\left[Z(s)\right]ds\\
&=20\frac{\alpha+\beta}{\lambda}\varepsilon_n^{-2}e^{-\lambda f_1(n)}.
\end{aligned}$$

Taking the limit when m tends to infinity, due to continuity of measure, we get

$$\mathbb{P}\left(\sup_{t\in[f_1(n),\infty]}\left|e^{-\lambda t}Z(t)-W\right|\geq\varepsilon_n\right)\leq\frac{\alpha+\beta}{\lambda}h^{-1}(n),$$

which ends the proof. □

Analogously to [8], we show that at time $t_-^{(n)}(k)$ only a finite number of mutants cells of trait k are presents in the population of cells, whose law is asymptotically distributed as a Luria-Delbrück random variable.

Lemma 5.4.2. *We have*

$$B_{k-1,1}^{(n)}\left(t_-^{(n)}(k)\right)\underset{n\to\infty}{\longrightarrow}B_{k-1,1}^{\infty},$$

in probability, where $B_{k-1,1}^{\infty}$ follows a Luria-Delbrück distribution with parameters

$$\left(\lambda_0,\alpha_k,\beta_k,\frac{2\alpha_{k-1}\mu_{k-1}}{\lambda_0}w_{k-1}W\right).$$

In particular due to $\lambda_k>\lambda_0$, we have $\mathbb{E}\left[B_{k-1,1}^{\infty}\right]=\infty$.

Proof. The proof is sufficiently close to the mathematical analysis made in [8, 10] such that we are only giving the different steps without details:

- One start to show that it is unlikely to observe up to time $t_-^{(n)}(k)$ cell divisions of trait $k-1$ whose both daughter cells mutate to trait k, that is

$$\mathbb{P}\left(H_{k-1,1}^{(n)}\left(t_-^{(n)}(k)\right)=0\right)\underset{n\to\infty}{\longrightarrow}1.$$

  Then it immediately follows that

$$\mathbb{P}\left(\sum_{i=1}^{H_{k-1,1}^{(n)}\left(t_-^{(n)}(k)\right)}X_{i,1}\left(t_-^{(n)}(k)-S_i^{(n)}\right)+X_{i,2}\left(t_-^{(n)}(k)-S_i^{(n)}\right)=0\right)\underset{n\to\infty}{\longrightarrow}1.$$

- Then adapting techniques from [8, 10] one shows that the number of cell divisions of trait $k-1$ whose only one daughter cell mutates to trait k up to time $t_-^{(n)}(k)$ is asymptotically Poisson distributed with parameter $\frac{2\alpha_{k-1}\mu_{k-1}}{\lambda_0}w_{k-1}W$, that is

$$K_{k-1,1}^{(n)}\left(t_-^{(n)}(k)\right)\underset{n\to\infty}{\longrightarrow}N_{k-1,1}\left(\frac{2\alpha_{k-1}\mu_{k-1}}{\lambda_0}w_{k-1}W\right),$$

  in probability.

- The last step consists in showing that conditioned on $K_{k-1,1}^{(n)}\left(t_-^{(n)}(k)\right)=j$, the law of the vector $\left(t_-^{(n)}(k)-T_i^{(n)}\right)_{i\in\{1,\cdots,j\}}$ is asymptotically distributed as the statistic order of a vector of j i.i.d. exponential random variables with parameter λ_0, which is a classical result.

□

We start by showing that the first order asymptotics at time $(t(k)+t)\log(n)$ of the lineages of cells of trait k issued from mutational events up to time $t_-^{(n)}(k)$ is of order $n^{t\lambda_k}\log^{\theta(k-1)\frac{\lambda_k}{\lambda_0}}(n)$. Such order comes from the exponential growth of the lineages of the mutant cells at rate λ_k between times $[t_-^{(n)}(k),(t+t(k))\log(n)]$. Asymptotically, it is a compound Luria-Delbrück distribution (associated with the distribution of the large time limit of the martingale associated to a birth and death branching process with rates α_k and β_k) that is captured. Basically such result comes from Lemma 5.4.2 giving asymptotically how many mutant cells there are at time $t_-^{(n)}(k)$ and then we control the exponential growth at rate λ_k of each lineage for a time $(t+t(k))\log(n)-t_-^{(n)}(k)=t\log(n)+\ell_-^{(n)}(k)$.

Proposition 5.4.3. *For all* $0<T_1<T_2$

$$\left(\frac{Z_{k,1}^{(n)}\left((t(k)+t)\log(n)\right)}{n^{t\lambda_k}\log^{\theta(k-1)\frac{\lambda_k}{\lambda_0}}(n)}\right)_{t\in[T_1,T_2]}\underset{n\to\infty}{\longrightarrow}\left(Z_{k,1}^{\infty}:=\sum_{i=1}^{B_{k-1,1}^{\infty}}U_i^{\infty}\right)_{t\in[T_1,T_2]},$$

in probability in $L^{\infty}([T_1,T_2])$*, where the sequence* $(U_i^{\infty})_{i\in\mathbb{N}}$ *is the almost sure limits of the positive martingales* $\left(e^{-\lambda_k t}U_i(t)\right)_{i\in\mathbb{N}}$*. Namely* $U_i^{\infty}\overset{law}{=}Ber\left(\frac{\lambda_k}{\alpha_k}\right)\otimes Exp\left(\frac{\lambda_k}{\alpha_k}\right)$.

Proof. By definition of the process $Z_{k,1}^{(n)}$, given in Equation (5.3.6), we have for all $t > 0$

$$\frac{Z_{k,1}^{(n)}\left((t+t(k))\log(n)\right)}{n^{t\lambda_k}\log^{\theta(k-1)\frac{\lambda_k}{\lambda_0}}(n)} = \sum_{i=1}^{B_{k-1,1}^{(n)}\left(t_-^{(n)}(k)\right)} U_i\left(\ell_-^{(n)}(k)+t\log(n)\right)e^{-\lambda_k\left(\ell_-^{(n)}(k)+t\log(n)\right)},$$

because $(t+t(k))\log(n)-t_-^{(n)}(k) = t\log(n)+\ell_-^{(n)}(k)$ and $n^{t\lambda_k}\log^{\theta(k-1)\frac{\lambda_k}{\lambda_0}}(n) = e^{-\lambda_k\left(\ell_-^{(n)}(k)+t\log(n)\right)}$, by definition of $\ell_-^{(n)}(k)$ given in Equation (5.3.1). Then we have

$$\mathbb{P}\left(\sup_{t\in[T_1,T_2]}\left|\frac{Z_{k,1}^{(n)}\left((t+t(k))\log(n)\right)}{n^{t\lambda_k}\log^{\theta(k-1)\frac{\lambda_k}{\lambda_0}}(n)} - \sum_{i=1}^{B_{k-1,1}^{\infty}} U_i^{\infty}\right| \geq \varepsilon\right) \leq \mathbb{P}\left(B_{k-1,1}^{(n)}\left(t_-^{(n)}(k)\right) \neq B_{k-1,1}^{\infty}\right)$$
$$+\mathbb{P}\left(B_{k-1,1}^{\infty} \geq g_n\right) + \sum_{i=1}^{g_n}\mathbb{P}\left(B_{k-1,1}^{\infty}=i\right) i\mathbb{P}\left(\sup_{s\in[T_1\log(n),\infty]}\left|U_1(s)e^{-\lambda_k s}-U_1^{\infty}\right| \geq \frac{\varepsilon}{i}\right).$$

According to Lemma 5.4.2, we have that $\mathbb{P}\left(B_{k-1,1}^{(n)}\left(t_-^{(n)}(k)\right) \neq B_{k-1,1}^{\infty}\right) \underset{n\to\infty}{\longrightarrow} 0$ because they are integer random variables.

Applying Lemma 5.4.1 to the birth and death branching process U_1 with rates α_k and β_k one obtain that

$$\mathbb{P}\left(\sup_{s\in[T_1\log(n),\infty]}\left|U_1(s)e^{-\lambda_k s}-U_1^{\infty}\right| \geq \frac{\varepsilon}{i}\right) \leq \frac{C}{\varepsilon} i n^{-\frac{\lambda_k T_1}{2}}.$$

Then choosing $g_n \underset{n\to\infty}{\longrightarrow} \infty$ such that $g_n^3 = o\left(n^{\frac{\lambda_k T_1}{2}}\right)$ we get

$$\sum_{i=1}^{g_n}\mathbb{P}\left(B_{k-1,1}^{\infty}=i\right) i\mathbb{P}\left(\sup_{s\in[T_1\log(n),\infty]}\left|U(s)e^{-\lambda_k s}-U^{\infty}\right| \geq \frac{\varepsilon}{i}\right) \leq \frac{C}{\varepsilon} g_n^3 n^{-\frac{\lambda_k T_1}{2}} \underset{n\to\infty}{\longrightarrow} 0,$$

which concludes the proof. □

In the next proposition we capture the first-order asymptotic of the subpopulation $Z_{k,2}^{(n)}$ at time $t\log(n)+\ell_-^{(n)}(k)$. We first show that its first-order asymptotic at time $\ell_-^{(n)}(k)$ is of order $\log^{\theta(k-1)\frac{\lambda_k}{\lambda_0}}(n)$ which is obtained using the approximate model of Definition 5.2.2 combined with a martingale approach. Then, using a law of large numbers we control the exponential growth at rate λ_k of the lineages of the mutant cells present at time $\ell_-^{(n)}(k)$ during a time $t\log(n)$ giving the renormalizing factor $n^{t\lambda_k}$.

Proposition 5.4.4. *For all* $0 < T_1 < T_2$

$$\left(\frac{Z_{k,2}^{(n)}\left(t\log(n)+\ell_-^{(n)}(k)\right)}{n^{t\lambda_k}\log^{\theta(k-1)\frac{\lambda_k}{\lambda_0}}(n)}\right)_{t\in[T_1,T_2]} \underset{n\to\infty}{\longrightarrow} \left(Z_{k,2}^{\infty}\right)_{t\in[T_1,T_2]},$$

in probability in $L^{\infty}([T_1,T_2])$*, where* $Z_{k,2}^{\infty}$*, properly defined in* (5.4.1)*, is the asymptotic limit of the approximate model of Definition 5.2.2 with parameters* $(\lambda_0, 2\alpha_{k-1}\mu_{k-1}w_{k-1}W, \alpha_k, \beta_k)$.

Proof. Introduce the following approximate model

$$\widetilde{Z}_{k,2}(t) := \widetilde{K}_{k-1,2}(t) + \int_0^t \int_{\mathbb{R}^+} \mathbb{1}_{\left\{\theta \leq \alpha_k \widetilde{Z}_{k,2}(s)\right\}} Q^b_{k,2}(ds, d\theta) - \int_0^t \int_{\mathbb{R}^+} \mathbb{1}_{\left\{\theta \leq \widetilde{Z}_{k,2}(s)\right\}} Q^d_{k,2}(ds, d\theta),$$
$$\widetilde{K}_{k-1,2}(t) := \int_0^t \int_{\mathbb{R}^+} \mathbb{1}_{\left\{\theta \leq 2\alpha_{k-1}\mu_{k-1}w_{k-1}We^{\lambda_0 s}\right\}} N_{k-1,2}(ds, d\theta),$$

where $w_{k-1} = \widetilde{w}_{(0,\cdots,k-1)}(t(k))$. According to Definition 5.2.2 we have

$$\frac{\widetilde{Z}_{k-1,2}\left(\ell^{(n)}_-(k)\right)}{\log^{\theta(k-1)\frac{\lambda_k}{\lambda_0}}(n)} \underset{n\to\infty}{\longrightarrow} Z^\infty_{k,2}, \tag{5.4.1}$$

almost surely, where the law of $Z^\infty_{k,2}$ is given by the distribution given in this latter definition with parameters $(\lambda_0, 2\alpha_{k-1}\mu_{k-1}w_{k-1}W, \alpha_k, \beta_k)$. The proof is done in two steps. The first one is to show that up to time $\ell^{(n)}_-(k)$ the processes $Z^{(n)}_{k,2}$ and $\widetilde{Z}_{k,2}$ are close using a martingale method, and the second step is applying a law of large numbers method to get the asymptotic limit.

Step 1: We are going to show that

$$\frac{Z^{(n)}_{k,2}\left(\ell^{(n)}_-(k)\right)}{\log^{\theta(k-1)\frac{\lambda_k}{\lambda_0}}(n)} \underset{n\to\infty}{\longrightarrow} Z^\infty_{k,2}, \tag{5.4.2}$$

in probability. In the next Lemma, we are introducing the martingales that we are going to use, as well as we compute their quadratic variations.

Lemma 5.4.5. *Define*

$$M^{(n)}_{k,2}(t) := e^{-\lambda_k t}Z^{(n)}_{k,2}(t) - \int_0^t 2\alpha_{k-1}\mu^{(n)}_{k-1}e^{-\lambda_k s}Z^{(n)}_{k-1}\left(t^{(n)}_-(k)+s\right)ds,$$
$$\widetilde{M}_{k,2}(t) := e^{-\lambda_k t}\widetilde{Z}_{k,2}(t) - \int_0^t 2\alpha_{k-1}\mu_{k-1}e^{-(\lambda_k-\lambda_0)s}w_{k-1}W ds.$$

Then $\left(M^{(n)}_{k,2}(t)\right)_{t\in\mathbb{R}^+}$ *and* $\left(\widetilde{M}_{k,2}(t)\right)_{t\in\mathbb{R}^+}$ *are martingales, with quadratic variations*

$$\left\langle M^{(n)}_{k,2}\right\rangle_t = \int_0^t 2\alpha_{k-1}\mu^{(n)}_{k-1}e^{-2\lambda_k s}Z^{(n)}_{k-1}\left(t^{(n)}_-(k)+s\right)ds + (\alpha_k+\beta_k)\int_0^t e^{-2\lambda_k s}Z^{(n)}_{k,2}(s)ds,$$
$$\left\langle \widetilde{M}_{k,2}\right\rangle_t = \int_0^t 2\alpha_{k-1}\mu_{k-1}e^{-(2\lambda_k-\lambda_0)s}w_{k-1}W ds + (\alpha_k+\beta_k)\int_0^t e^{-2\lambda_k s}\widetilde{Z}_{1,2}(s)ds.$$

We also have that

$$\left\langle M^{(n)}_{k,2}, \widetilde{M}_{k,2}\right\rangle_t = \int_0^t 2\alpha_{k-1}e^{-2\lambda_k s}\left\{\mu^{(n)}_{k-1}\left(1-\mu^{(n)}_{k-1}\right)Z^{(n)}_{k-1}\left(t^{(n)}_-(k)+s\right) \wedge \mu_{k-1}e^{\lambda_0 s}w_{k-1}W\right\}ds$$
$$+ (\alpha_k+\beta_k)\int_0^t e^{-2\lambda_k s}\left(Z^{(n)}_{k,2}(s) \wedge \widetilde{Z}_{k,2}(s)\right)ds.$$

Proof. The proof that $M_{k,2}^{(n)}$ and $\widetilde{M}_{k,2}$ are martingales as well as the computation of their quadratic variations can be obtained easily by adapting the proof of Chapter 4, Lemma 4.3.5. We are going to compute $\left\langle M_{k,2}^{(n)}, \widetilde{M}_{k,2} \right\rangle_t$. Let $F(t,x,y,\widetilde{x},\widetilde{y}) = \left(e^{-\lambda_k t}x - y\right)\left(e^{-\lambda_k t}\widetilde{x} - \widetilde{y}\right)$. We have that

$$\begin{aligned}
\frac{\partial F}{\partial t}(t,x,y,\widetilde{x},\widetilde{y}) &= -\lambda_k e^{-\lambda_k t}\left[x\left(e^{-\lambda_k t}\widetilde{x} - \widetilde{y}\right) + \widetilde{x}\left(e^{-\lambda_k t}x - y\right)\right], \\
\frac{\partial F}{\partial y}(t,x,y,\widetilde{x},\widetilde{y}) &= -\left(e^{-\lambda_k t}\widetilde{x} - \widetilde{y}\right), \\
\frac{\partial F}{\partial \widetilde{y}}(t,x,y,\widetilde{x},\widetilde{y}) &= -\left(e^{-\lambda_k t}x - y\right).
\end{aligned}$$

Applying Ito's formula we get

$$\begin{aligned}
M_{k,2}^{(n)}(t)\widetilde{M}_{k,2}(t) &= F\Bigg(t, Z_{k,2}^{(n)}(t), \int_0^t 2\alpha_{k-1}\mu_{k-1}^{(n)} e^{-\lambda_k s} Z_{k-1}^{(n)}\left(t_-^{(n)}(k)+s\right) ds, \\
&\qquad\qquad \widetilde{Z}_{k,2}(t), \int_0^t 2\alpha_{k-1}\mu_{k-1} e^{-(\lambda_k-\lambda_0)s} w_{k-1} W ds\Bigg) \\
&= F(0,0,0,0) - \lambda_k \int_0^t e^{-\lambda_k s}\left(Z_{k,2}^{(n)}(s)\widetilde{M}_{k,2}(s) + \widetilde{Z}_{k,2}(s)M_{k,2}^{(n)}(s)\right) ds \\
&- \int_0^t \widetilde{M}_{k,2}(s) 2\alpha_{k-1}\mu_{k-1}^{(n)} Z_{k-1}^{(n)}\left(t_-^{(n)}(k)+s\right) e^{-\lambda_k s} ds - \int_0^t M_{k,2}^{(n)}(s) 2\alpha_{k-1}\mu_{k-1} w_{k-1} W e^{-(\lambda_k-\lambda_0)s} ds \\
&+ \int_0^t \int_{\mathbb{R}^+} \Bigg[\left(M_{k,2}^{(n)}(s) + e^{-\lambda_k s}\mathbb{1}_{\left\{\theta \le 2\alpha_{k-1}\mu_{k-1}^{(n)}\left(1-\mu_{k-1}^{(n)}\right) Z_{k-1}^{(n)}\left(t_{k-1}^{(n)}+s\right)\right\}}\right) \\
&\qquad \cdot \left(\widetilde{M}_{k,2}(s) + e^{-\lambda_k s}\mathbb{1}_{\left\{\theta \le 2\alpha_{k-1}\mu_{k-1} w_{k-1} W e^{\lambda_0 s}\right\}}\right) - M_{k,2}^{(n)}(s)\widetilde{M}_{k,2}(s)\Bigg] N_{k-1,2}(ds, d\theta) \\
&+ \int_0^{t\wedge \ell_-^{(n)}(k)} \int_{\mathbb{R}^+} \Bigg[\left(M_{k,2}^{(n)}(s) + 2e^{-\lambda_k s}\mathbb{1}_{\left\{\theta \le \alpha_{k-1}\left(\mu_{k-1}^{(n)}\right)^2 Z_{k-1}^{(n)}\left(t_{k-1}^{(n)}+s\right)\right\}}\right)\widetilde{M}_{k,2}(s) \\
&\qquad - M_{k,2}^{(n)}(s)\widetilde{M}_{k,2}(s)\Bigg] Q_{k-1,2}^{m}(ds, d\theta) \\
&+ \int_0^t \int_{\mathbb{R}^+} \Bigg[\left(M_{k,2}^{(n)}(s) + e^{-\lambda_k s}\mathbb{1}_{\left\{\theta \le \alpha_k Z_{k,2}^{(n)}(s^-)\right\}}\right)\left(\widetilde{M}_{k,2}(s) + e^{-\lambda_k s}\mathbb{1}_{\left\{\theta \le \alpha_k \widetilde{Z}_{k,2}(s^-)\right\}}\right) \\
&\qquad - M_{k,2}^{(n)}(s)\widetilde{M}_{k,2}(s)\Bigg] Q_{k,2}^{b}(ds, d\theta) \\
&+ \int_0^t \int_{\mathbb{R}^+} \Bigg[\left(M_{k,2}^{(n)}(s) - e^{-\lambda_k s}\mathbb{1}_{\left\{\theta \le \beta_k Z_{k,2}^{(n)}(s^-)\right\}}\right)\left(\widetilde{M}_{k,2}(s) - e^{-\lambda_k s}\mathbb{1}_{\left\{\theta \le \beta_k \widetilde{Z}_{k,2}(s^-)\right\}}\right) \\
&\qquad - M_{k,2}^{(n)}(s)\widetilde{M}_{k,2}(s)\Bigg] Q_{k,2}^{d}(ds, d\theta)
\end{aligned}$$

$$
\begin{aligned}
&= -\lambda_k \int_0^t e^{-\lambda_k s} \left(Z_{k,2}^{(n)}(s) \widetilde{M}_{k,2}(s) + \widetilde{Z}_{k,2}(s) M_{k,2}^{(n)}(s) \right) ds \\
&- \int_0^t 2\alpha_{k-1} e^{-\lambda_k s} \left(\widetilde{M}_{k,2}(s) \mu_{k-1}^{(n)} Z_{k-1}^{(n)} \left(t_-^{(n)}(k) + s \right) + M_{k,2}^{(n)}(s) \mu_{k-1} w_{k-1} W e^{\lambda_0 s} \right) ds \\
&+ \int_0^t \left(M_{k,2}^{(n)}(s) 2\alpha_{k-1} \mu_{k-1} w_{k-1} W e^{\lambda_0 s} e^{-\lambda_k s} + \widetilde{M}_{k,2}(s) 2\alpha_{k-1} \mu_{k-1}^{(n)} Z_{k-1}^{(n)} \left(t_-^{(n)}(k) + s \right) e^{-\lambda_k s} \right) ds \\
&+ \int_0^t e^{-2\lambda_k s} 2\alpha_{k-1} \left(\mu_{k-1}^{(n)} \left(1 - \mu_{k-1}^{(n)} \right) Z_{k-1}^{(n)} \left(t_-^{(n)}(k) + s \right) \wedge \mu_{k-1} w_{k-1} W e^{\lambda_0 s} \right) ds \\
&+ \lambda_k \int_0^t e^{-\lambda_k s} \left(Z_{k,2}^{(n)}(s) \widetilde{M}_{k,2}(s) + \widetilde{Z}_{k,2}(s) M_{k,2}^{(n)}(s) \right) ds \\
&+ \int_0^t e^{-2\lambda_k s} \left(\alpha_k + \beta_k \right) \left(Z_{k,2}^{(n)}(s) \wedge \widetilde{Z}_{k,2}(s) \right) ds \\
&+ N^{(n)}(s),
\end{aligned}
$$

where $N^{(n)}$ is a martingale. Finally we obtain that

$$
\begin{aligned}
\left\langle M_{k,2}^{(n)}, \widetilde{M}_{k,2} \right\rangle_t = \int_0^t 2\alpha_{k-1} e^{-2\lambda_k s} \left(\mu_{k-1}^{(n)} \left(1 - \mu_{k-1}^{(n)} \right) Z_{k-1}^{(n)} \left(t_-^{(n)}(k) + s \right) \wedge \mu_{k-1} e^{\lambda_0 s} w_{k-1} W \right) ds \\
+ \left(\alpha_k + \beta_k \right) \int_0^t e^{-2\lambda_k s} \left(Z_{k,2}^{(n)}(s) \wedge \widetilde{Z}_{k,2}(s) \right) ds.
\end{aligned}
$$

□

In the next Lemma, we are computing the asymptotic limit of the finite variation process associated to the subpopulation $Z_{k,2}^{(n)}$.

Lemma 5.4.6. *Let $0 < T_1 < T_2 \leq \frac{\theta(k-1)}{\lambda_0}$ and $\varepsilon > 0$. Then we have*

$$
\mathbb{P} \left(\sup_{t \in [T_1, T_2]} \int_0^{t \log(\log(n))} 2\alpha_{k-1} e^{-\lambda_k s} \left| \mu_{k-1}^{(n)} Z_{k-1}^{(n)} \left(t_-^{(n)}(k) + s \right) - \mu_{k-1} e^{\lambda_0 s} w_{k-1} W \right| ds \geq \varepsilon \right) \underset{n \to \infty}{\longrightarrow} 0.
$$

Proof. Introduce

$$
A_{\widetilde{\varepsilon}}^{(n)} := \left\{ \sup_{s \in [0, T_2 \log(\log(n))]} \left| \frac{Z_{k-1}^{(n)} \left(t_-^{(n)}(k) + s \right)}{n^{\lambda_0 \left(\frac{\ell(k-1)}{\lambda_0} - \ell_-^{(n)}(k) \right)} \log^{\theta(k-1)}(n) e^{\lambda_0 s}} - w_{k-1} W \right| \leq \widetilde{\varepsilon} \right\}.
$$

We have that $\mathbb{P}\left(A_{\widetilde{\varepsilon}}^{(n)}\right) \underset{n\to\infty}{\longrightarrow} 1$ according to Proposition 5.1.2. Under the event $A_{\widetilde{\varepsilon}}^{(n)}$ we have

$$\begin{aligned}
&\int_0^{t\log(\log(n))} 2\alpha_{k-1}e^{-\lambda_k s}\left|\mu_{k-1}^{(n)}Z_{k-1}^{(n)}\left(t_-^{(n)}(k)+s\right)-\mu_{k-1}e^{\lambda_0 s}w_{k-1}W\right|ds\\
&=\int_0^{t\log(\log(n))} 2\alpha_{k-1}e^{-(\lambda_k-\lambda_0)s}\\
&\qquad\qquad\cdot\left|\left(\mu_{k-1}^{(n)}n^{\ell(k-1)}\right)\frac{Z_{k-1}^{(n)}\left(t_-^{(n)}(k)+s\right)}{n^{\lambda_0\left(\frac{\ell(k-1)}{\lambda_0}-\ell_-^{(n)}(k)\right)}\log^{\theta(k-1)}(n)e^{\lambda_0 s}}-\mu_{k-1}w_{k-1}W\right|ds\\
&\leq\int_0^{t\log(\log(n))} 2\alpha_{k-1}e^{-(\lambda_k-\lambda_0)s}\left(\left|\left(\mu_{k-1}^{(n)}n^{\ell(k-1)}\right)-\mu_{k-1}\right|w_{k-1}W+\left(\mu_{k-1}^{(n)}n^{\ell(k-1)}\right)\widetilde{\varepsilon}\right)ds\\
&\leq\frac{2\alpha_{k-1}}{\lambda_k-\lambda_0}\left(\left|\left(\mu_{k-1}^{(n)}n^{\ell(k-1)}\right)-\mu_{k-1}\right|w_{k-1}W+\left(\mu_{k-1}^{(n)}n^{\ell(k-1)}\right)\widetilde{\varepsilon}\right)\\
&\leq\varepsilon
\end{aligned}$$

for n large enough and taking a well chosen $\widetilde{\varepsilon}$. So we get that under $A_{\widetilde{\varepsilon}}^{(n)}$ we have

$$\sup_{t\in[T_1,T_2]}\int_0^{t\log(\log(n))} 2\alpha_{k-1}e^{-\lambda_k s}\left|\mu_{k-1}^{(n)}Z_{k-1}^{(n)}\left(t_-^{(n)}(k)+s\right)-\mu_{k-1}e^{\lambda_0 s}w_{k-1}W\right|ds\leq\varepsilon,$$

and because $\mathbb{P}\left(A_{\widetilde{\varepsilon}}^{(n)}\right) \underset{n\to\infty}{\longrightarrow} 1$ the result is obtained. □

Now we are going to show that well renormalized the processes $Z_{k,2}^{(n)}$ and $\widetilde{Z}_{k,2}$ are closed in probability up to time $\ell_-^{(n)}(k)$.

Lemma 5.4.7. *Let $0<T_1<T_2\leq\frac{\theta(k-1)}{\lambda_0}$ and $\varepsilon>0$. Then we have*

$$\mathbb{P}\left(\sup_{t\in[T_1,T_2]}\log^{-t\lambda_k}(n)\left|Z_{k,2}^{(n)}\left(t\log(\log(n))\right)-\widetilde{Z}_{k,2}\left(t\log(\log(n))\right)\right|\geq\varepsilon\right)\underset{n\to\infty}{\longrightarrow}0.$$

Proof. We have

$$\begin{aligned}
&\mathbb{P}\left(\sup_{t\in[T_1,T_2]}\log^{-t\lambda_k}(n)\left|Z_{k,2}^{(n)}\left(t\log(\log(n))\right)-\widetilde{Z}_{k,2}\left(t\log(\log(n))\right)\right|\geq 2\varepsilon\right)\\
&\qquad\leq\mathbb{P}\left(\sup_{t\in[T_1,T_2]}\left|M_{k,2}^{(n)}\left(t\log(\log(n))\right)-\widetilde{M}_{k,2}\left(t\log(\log(n))\right)\right|\geq\varepsilon\right) \qquad (5.4.3)\\
&\qquad+\mathbb{P}\left(\sup_{t\in[T_1,T_2]}\int_0^{t\log(\log(n))}2\alpha_{k-1}e^{-\lambda_k s}\left|\mu_{k-1}^{(n)}Z_{k-1}^{(n)}\left(t_-^{(n)}(k)+s\right)\right.\right. \qquad (5.4.4)\\
&\qquad\qquad\left.\left.-\mu_{k-1}e^{\lambda_0 s}w_{k-1}W\right|ds\geq\varepsilon\right).
\end{aligned}$$

The convergence to 0 of the term (5.4.4) is given by Lemma 5.4.6. For the convergence to 0 of the term (5.4.3) we start by using Doob's Inequality in L^2 (see [95] Proposition 3.15) applied to the martingale $\left(M_{k,2}^{(n)}(t) - \widetilde{M}_{k,2}(t)\right)_{t\in\mathbb{R}^+}$ and then we use the expression of the quadratic variations of the martingales $M_{k,2}^{(n)}$ and $\widetilde{M}_{k,2}$, proved in Lemma 5.4.5, to get

$$
\begin{aligned}
&\mathbb{P}\left(\sup_{t\in[T_1,T_2]}\left|M_{k,2}^{(n)}\left(t\log(\log(n))\right) - \widetilde{M}_{k,2}\left(t\log(\log(n))\right)\right| \geq \varepsilon\right)\\
&\leq \frac{4}{\varepsilon^2}\mathbb{E}\Big[M_{k,2}^{(n)}\left(T_2\log(\log(n))\right)^2 + \widetilde{M}_{k,2}\left(T_2\log(\log(n))\right)^2\\
&\qquad\qquad - 2M_{k,2}^{(n)}\left(T_2\log(\log(n))\right)\widetilde{M}_{k,2}\left(T_2\log(\log(n))\right)\Big]\\
&= \frac{4}{\varepsilon^2}\Bigg(\int_0^{T_2\log(\log(n))} 2\alpha_{k-1}\mu_{k-1}^{(n)}e^{-2\lambda_k s}\mathbb{E}\left[Z_{k-1}^{(n)}\left(t_-^{(n)}(k)+s\right)\right]ds\\
&+ (\alpha_k+\beta_k)\int_0^{T_2\log(\log(n))} e^{-2\lambda_k s}\mathbb{E}\left[Z_{k,2}^{(n)}(s)\right]ds\\
&+ \int_0^{T_2\log(\log(n))} 2\alpha_{k-1}\mu_{k-1}e^{-(2\lambda_k-\lambda_0)s}w_{k-1}\mathbb{E}\left[W\right]ds\\
&+ (\alpha_k+\beta_k)\int_0^{T_2\log(\log(n))} e^{-2\lambda_k s}\mathbb{E}\left[\widetilde{Z}_{k,2}(s)\right]ds\\
&- 2\int_0^{T_2\log(\log(n))} 2\alpha_{k-1}e^{-2\lambda_k s}\mathbb{E}\left[\mu_{k-1}^{(n)}\left(1-\mu_{k-1}^{(n)}\right)Z_{k-1}^{(n)}\left(t_-^{(n)}(k)+s\right)\wedge\mu_{k-1}e^{\lambda_0 s}w_{k-1}W\right]ds\\
&- 2(\alpha_k+\beta_k)\int_0^{T_2\log(\log(n))} e^{-2\lambda_k s}\mathbb{E}\left[Z_{k,2}^{(n)}(s)\wedge\widetilde{Z}_{k,2}(s)\right]ds\Bigg)\\
&= \frac{4}{\varepsilon^2}\Bigg(\int_0^{T_2\log(\log(n))} 2\alpha_{k-1}e^{-2\lambda_k s}\mathbb{E}\left[\left|\mu_{k-1}^{(n)}\left(1-\mu_{k-1}^{(n)}\right)Z_{k-1}^{(n)}\left(t_-^{(n)}(k)+s\right) - \mu_{k-1}e^{\lambda_0 s}w_{k-1}W\right|\right]ds \qquad (5.4.5)\\
&\qquad + (\alpha_k+\beta_k)\int_0^{T_2\log(\log(n))} e^{-2\lambda_k s}\mathbb{E}\left[\left|Z_{k,2}^{(n)}(s)-\widetilde{Z}_{k,2}(s)\right|\right]ds \qquad (5.4.6)\\
&\qquad + \int_0^{T_2\log(\log(n))} 2\alpha_{k-1}\left(\mu_{k-1}^{(n)}\right)^2 e^{-2\lambda_k s}\mathbb{E}\left[Z_{k-1}^{(n)}\left(t_-^{(n)}(k)+s\right)\right]ds\Bigg), \qquad (5.4.7)
\end{aligned}
$$

where for the last equality we used that for all $(a,b)\in\mathbb{R}_+^2$ we have $a+b-2(a\wedge b) = |a-b|$, applied with the couples $(a,b) = \left(\mu_{k-1}^{(n)}\left(1-\mu_{k-1}^{(n)}\right)Z_{k-1}^{(n)}\left(t_-^{(n)}(k)+s\right), \mu_{k-1}e^{\lambda_0 s}w_{k-1}W\right)$ and $(a,b) = \left(Z_{k,2}^{(n)}(s), \widetilde{Z}_{k,2}(s)\right)$. To show that the terms (5.4.5) and (5.4.6) converge to 0, we are going to apply the dominated convergence theorem. Using that $\mathbb{E}[W] = 1$ and that for all $s\in\mathbb{R}^+$ we

have according to Lemma 4.3.10

$$\mathbb{E}\left[Z_{k-1}\left(t_-^{(n)}(k)+s\right)\right] \leq C_{k-1}\left(\mu_{\otimes,k-1}^{(n)} n^{t(k-1)}\right) \cdot \left(t_-^{(n)}(k)+s\right)^{\theta(k-1)} e^{\lambda_0 s} n^{\ell(k-1)} \log^{-\theta(k-1)}(n), \tag{5.4.8}$$

we get that for all $s \leq T_2 \log(\log(n))$ and for n large enough

$$\begin{aligned} e^{-2\lambda_k s}\mathbb{E}\left[\left|\mu_{k-1}^{(n)}\left(1-\mu_{k-1}^{(n)}\right) Z_{k-1}^{(n)}\left(t_-^{(n)}(k)+s\right) - \mu_{k-1}e^{\lambda_0 s} w_{k-1} W\right|\right] \\ \leq e^{-(2\lambda_k-\lambda_0)s} 2\left(t(k)\right)^{\theta(k-1)}\left(C_{k-1}\left(\mu_{\otimes,k}^{(n)} n^{t(k)}\right) + \mu_{k-1}w_{k-1}\right) \\ \leq C(k) e^{-(2\lambda_k-\lambda_0)s}, \end{aligned}$$

because $\mu_{\otimes,k}^{(n)} n^{t(k)} \underset{n\to\infty}{\longrightarrow} \prod_{i=0}^{k-1}\mu_i$ and where $C(k)$ is a constant that depends on k. We obtained an upper bound that is integrable on $\mathbb{R}^+$. Moreover we have for n large enough

$$\begin{aligned} \mathbb{E}\left[e^{-\lambda_k s} Z_{k,2}^{(n)}(s)\right] &= \int_0^s 2\alpha_{k-1}\mu_{k-1}^{(n)} e^{-\lambda_k u}\mathbb{E}\left[Z_{k-1}^{(n)}\left(t_-^{(n)}(k)+u\right)\right] du \\ &\leq C(k)\left(\mu_{\otimes,k}^{(n)} n^{t(k)}\right)\int_0^s e^{-(\lambda_k-\lambda_0)u} du \\ &\leq C(k), \end{aligned}$$

where $C(k)$ is a constant depending on k that may change from line to line, as well as

$$\mathbb{E}\left[e^{-\lambda_k s}\widetilde{Z}_{k,2}(s)\right] = \int_0^s 2\alpha_{k-1}\mu_{k-1} e^{-(\lambda_k-\lambda_0)u} w_{k-1} du = \frac{2\alpha_{k-1}\mu_{k-1}w_{k-1}}{\lambda_k-\lambda_0}\left(1-e^{-(\lambda_k-\lambda_0)s}\right).$$

It implies that it exists $C(k)$ such that for n large enough we have

$$e^{-2\lambda_k s}\mathbb{E}\left[\left|Z_{k,2}^{(n)}(s) - \widetilde{Z}_{k,2}(s)\right|\right] \leq e^{-2\lambda_k s}\left(\mathbb{E}\left[Z_{k,2}^{(n)}(s)\right] + \mathbb{E}\left[\widetilde{Z}_{k,2}(s)\right]\right) \leq C(k)e^{-\lambda_k s},$$

which is integrable on $\mathbb{R}^+$. To apply the dominated convergence theorem, it suffices now to show that for all $s \in \mathbb{R}^+$

$$\mathbb{E}\left[\left|\mu_{k-1}^{(n)}\left(1-\mu_{k-1}^{(n)}\right) Z_{k-1}^{(n)}\left(t_-^{(n)}(k)+s\right) - \mu_{k-1}w_{k-1}We^{\lambda_0 s}\right|\right] \underset{n\to\infty}{\longrightarrow} 0, \tag{5.4.9}$$

$$\mathbb{E}\left[\left|Z_{k,2}^{(n)}(s) - \widetilde{Z}_{k,2}(s)\right|\right] \underset{n\to\infty}{\longrightarrow} 0. \tag{5.4.10}$$

For the convergence to 0 of the term (5.4.9) we use that

$$\begin{aligned} &\mathbb{E}\left[\left|\mu_{k-1}^{(n)}\left(1-\mu_{k-1}^{(n)}\right) Z_{k-1}^{(n)}\left(t_-^{(n)}(k)+s\right) - \mu_{k-1}w_{k-1}We^{\lambda_0 s}\right|\right] \\ &\quad\leq \left|\left(\mu_{k-1}^{(n)} n^{\ell(k-1)}\right)\left(1-\mu_{k-1}^{(n)}\right) - \mu_{k-1}\right|\mathbb{E}\left[n^{-\ell(k-1)} Z_{k-1}^{(n)}\left(t_-^{(n)}(k)+s\right)\right] \\ &\quad+ \mu_{k-1}e^{\lambda_0 s}\mathbb{E}\left[\left|\frac{Z_{k-1}^{(n)}\left(t_-^{(n)}(k)+s\right)}{n^{\lambda_0\left(\frac{\ell(k-1)}{\lambda_0}-\ell_-^{(n)}(k)\right)}\log^{\theta(k-1)}(n)e^{\lambda_0 s}} - w_{k-1}W\right|\right] \\ &\quad\underset{n\to\infty}{\longrightarrow} 0, \end{aligned}$$

because $\mu_{k-1}^{(n)} n^{\ell(k-1)} \underset{n\to\infty}{\longrightarrow} \mu_{k-1}$ and because one can get with some extra efforts the convergence in L^1 for the subpopulation of trait k-1.
For the convergence of the term (5.4.10) we use that under the event $\left\{\forall u \in [0,s], K_{k-1,2}^{(n)}(u) = \widetilde{K}_{k-1,2}(u) \text{ and } H_{k-1,2}^{(n)}(u) = 0\right\}$ we have that $Z_{k,2}^{(n)}(u) = \widetilde{Z}_{k,2}(u), \forall u \in [0,s]$ almost surely. Then we apply the Chauchy-Schwarz inequality to get

$$\begin{aligned}
\mathbb{E}\left[\left|Z_{k,2}^{(n)}(s) - \widetilde{Z}_{k,2}(s)\right|\right] &= \mathbb{E}\left[\left|Z_{k,2}^{(n)}(s) - \widetilde{Z}_{k,2}(s)\right| \mathbb{1}_{\left\{\forall u\in[0,s], K_{k-1,2}^{(n)}(u)=\widetilde{K}_{k-1,2}(u) \text{ and } H_{k-1,2}^{(n)}(u)=0\right\}}\right] \\
&\quad + \mathbb{E}\left[\left|Z_{k,2}^{(n)}(s) - \widetilde{Z}_{k,2}(s)\right| \mathbb{1}_{\left\{\exists u\in[0,s], K_{k-1,2}^{(n)}(u)\neq\widetilde{K}_{k-1,2}(u) \text{ or } H_{k-1,2}^{(n)}(u)\geq 1\right\}}\right] \\
&\leq \sqrt{\mathbb{E}\left[\left|Z_{k,2}^{(n)}(s) - \widetilde{Z}_{k,2}(s)\right|\right]} \\
&\quad \cdot \sqrt{\mathbb{P}\left(\exists u \in [0,s], K_{k-1,2}^{(n)}(u) \neq \widetilde{K}_{k-1,2}(u)\right) + \mathbb{P}\left(\exists u \in [0,s], H_{k-1,2}^{(n)}(u) \geq 1\right)} \\
&\leq Ce^{s\frac{\lambda_0}{2}} \left\{1 - \exp\left(\int_0^s 2\alpha_{k-1}\mathbb{E}\left[\left|\mu_{k-1}^{(n)}\left(1-\mu_{k-1}^{(n)}\right) Z_{k-1}^{(n)}\left(t_-^{(n)}(k)+u\right) - \mu_{k-1}e^{\lambda_0 u} W w_{k-1}\right|\right] du\right)\right. \\
&\quad \left. + 1 - \exp\left(\int_0^s \alpha_{k-1}\left(\mu_{k-1}^{(n)}\right)^2 \mathbb{E}\left[Z_{k-1}^{(n)}\left(t_-^{(n)}(k)+u\right)\right] du\right)\right\}^{\frac{1}{2}} \\
&\underset{n\to\infty}{\longrightarrow} 0,
\end{aligned}$$

where we use the dominated convergence theorem for the last convergence. We conclude the proof by showing that (5.4.7) converges to 0. Using the computation of (5.4.8) we obtain that

$$\begin{aligned}
(5.4.7) &\leq C_{k-1}\left(\mu_{\otimes,k-1}^{(n)} n^{t(k-1)}\right)\left(t(k)+T_2\right)^{\theta(k-1)}\left(\mu_{k-1}^{(n)} n^{\ell(k-1)}\right)\mu_{k-1}^{(n)}\int_0^{T_2\log(\log(n))} e^{-(2\lambda_k-\lambda_0)s} ds \\
&\underset{n\to\infty}{\longrightarrow} 0,
\end{aligned}$$

because $\mu_{\otimes,k-1}^{(n)} n^{t(k-1)} \underset{n\to\infty}{\longrightarrow} \prod_{i=0}^{k-2}\mu_i < \infty$ and $\mu_{k-1}^{(n)} n^{\ell(k-1)} \underset{n\to\infty}{\longrightarrow} \mu_{k-1} < \infty$.

Finally we obtained the convergence to 0 of the terms (5.4.5), (5.4.6) and (5.4.7) which ends the proof. □

Combining Lemma 5.4.7 and Equation (5.4.1) we obtain Equation (5.4.2).

Step 2: By construction of $Z_{k,2}^{(n)}$ we have that

$$Z_{k,2}^{(n)}\left(t\log(n) + \ell_-^{(n)}(k)\right) = \sum_{i=1}^{Z_{k,2}^{(n)}\left(\ell_-^{(n)}(k)\right)} V_i\left(t\log(n)\right),$$

where the $(V_i)_{i\in\mathbb{N}}$ are independent birth and death branching processes with rates α_k and β_k

satisfying $V_i(0) = 1$ almost surely. So it comes that

$$\frac{Z_{k,2}^{(n)}\left(t\log(n) + \ell_-^{(n)}(k)\right)}{n^{t\lambda_k}\log^{\theta(k-1)\frac{\lambda_k}{\lambda_0}}(n)} = \frac{1}{\log^{\theta(k-1)\frac{\lambda_k}{\lambda_0}}(n)} \sum_{i=1}^{Z_{k,2}^{(n)}\left(\ell_-^{(n)}(k)\right)} e^{-\lambda_k t\log(n)} Y_i\left(t\log(n)\right).$$

In order to show the result, we are going to use the following lemma, which is a variant of the law of large numbers.

Lemma 5.4.8 (Law of Large Numbers)**.** *Let $g_n \underset{n\to\infty}{\longrightarrow} \infty$ be a deterministic function and $\left(G^{(n)}\right)_{n\in\mathbb{N}}$ be a sequence of natural numbers valued random variables such that $\left(\frac{G^{(n)}}{g_n}\right)_{n\in\mathbb{N}}$ converges in probability to $G^\infty \in L^1$ having a density with respect to the Lebesgue measure. Let $\left(\left(X_i^{(n)}\right)_{i\in\mathbb{N}}\right)_{n\in\mathbb{N}}$ be a sequence of independent and identically distributed stochastic processes, independent from $\left(G^{(n)}\right)_{n\in\mathbb{N}}$. Let $f_n \underset{n\to\infty}{\longrightarrow} \infty$, $h_n \underset{n\to\infty}{\longrightarrow} \infty$ and $0 < T_1 < T_2$, assume that for all $\varepsilon > 0$ and for all $i \in \mathbb{N}$*

$$\mathbb{P}\left(\sup_{t\in[T_1 f_n, T_2 f_n]} \left|X_i^{(n)}(t) - X_i^\infty\right| \geq \varepsilon\right) \leq h_n^{-1},$$

where $X_i^\infty \in L^1$. Then if $g_n = o(h_n)$ we have

$$\frac{1}{g_n}\sum_{i=1}^{G^{(n)}} X_i^{(n)}(tf_n) \underset{n\to\infty}{\longrightarrow} \mathbb{E}[X^\infty]G^\infty,$$

in probability in $L^\infty([T_1, T_2])$, where X^∞ follows the distribution of any X_i^∞.

Proof. We have

$$\begin{aligned}
&\mathbb{P}\left(\sup_{t\in[T_1,T_2]}\left|\frac{1}{g_n}\sum_{i=1}^{G^{(n)}} X_i^{(n)}(tf_n) - \mathbb{E}[X^\infty]G^\infty\right| \geq 4\varepsilon\right) \\
&\leq \mathbb{P}\left(\sup_{t\in[T_1,T_2]}\left|\frac{1}{g_n}\sum_{i=1}^{G^{(n)}}\left(X_i^{(n)}(tf_n) - X_i^\infty\right)\right| \geq 2\varepsilon\right) \\
&\qquad + \mathbb{P}\left(\left|\frac{1}{g_n}\sum_{i=1}^{G^{(n)}} X_i^\infty - \mathbb{E}[X^\infty]G^\infty\right| \geq 2\varepsilon\right) \\
&\leq \mathbb{P}\left(\frac{1}{g_n}\sum_{i=1}^{\lfloor G^\infty g_n\rfloor}\sup_{t\in[T_1,T_2]}\left|X_i^{(n)}(tf_n) - X_i^\infty\right| \geq \varepsilon\right) \qquad (5.4.11) \\
&\qquad + \mathbb{P}\left(\frac{1}{g_n}\sum_{i=1}^{\left|G^{(n)} - \lfloor G^\infty g_n\rfloor\right|}\sup_{t\in[T_1,T_2]}\left|X_i^{(n)}(tf_n) - X_i^\infty\right| \geq \varepsilon\right)
\end{aligned}$$

$$+\mathbb{P}\left(\left|\frac{1}{g_n}\sum_{i=1}^{\lfloor G^\infty g_n\rfloor} X_i^\infty - \mathbb{E}[X^\infty]G^\infty\right| \geq \varepsilon\right) + \mathbb{P}\left(\frac{1}{g_n}\sum_{i=1}^{\left|G^{(n)}-\lfloor G^\infty g_n\rfloor\right|} |X_i^\infty| \geq \varepsilon\right). \quad (5.4.12)$$

The second term of (5.4.12) converges to 0 according to the law of large numbers because the sequence of random variables $\left(\frac{\left|G^{(n)}-G^\infty g_n\right|}{g_n}\right)_{n\in\mathbb{N}}$ converges in probability to 0. The first term of (5.4.12) satisfies

$$\mathbb{P}\left(\left|\frac{1}{g_n}\sum_{i=1}^{\lfloor G^\infty g_n\rfloor} X_i^\infty - \mathbb{E}[X^\infty]G^\infty\right| \geq \varepsilon\right)$$
$$= \int_0^\infty \mathbb{P}\left(G^\infty = x\right)\mathbb{P}\left(\left|\frac{1}{xg_n}\sum_{i=1}^{\lfloor xg_n\rfloor} X_i^\infty - \mathbb{E}\left[X^\infty\right]\right| \geq \frac{\varepsilon}{x}\right)dx,$$

which converges to 0 according to the dominated convergence theorem, where for $x > 0$ we use the law of large numbers to prove that the integrand converges to 0 when n goes to infity. The first term of (5.4.11) satisfies

$$\mathbb{P}\left(\frac{1}{g_n}\sum_{i=1}^{\lfloor G^\infty g_n\rfloor}\sup_{t\in[T_1,T_2]}\left|X_i^{(n)}(tf_n) - X_i^\infty\right| \geq \varepsilon\right)$$
$$= \int_0^\infty \mathbb{P}\left(G^\infty = x\right)\lfloor xg_n\rfloor\mathbb{P}\left(\sup_{t\in[T_1,T_2]}\left|X_1^{(n)}(tf_n) - X_1^\infty\right| \geq \frac{\varepsilon g_n}{\lfloor xg_n\rfloor}\right)dx,$$

which converges to 0 according to the dominated convergence theorem, where for all $x > 0$ we use the hypothesis $g_n = o(h_n)$ to show that the integrand converges to 0, as well as we use that G^∞ is in L^1. The second term of (5.4.11) converges to 0 using similar techniques with the assumptions. □

The proof is ended by an application of Lemma 5.4.8 with $g_n = \log^{\theta(k-1)\frac{\lambda_k}{\lambda_0}}(n)$, $G^{(n)} = Z_{k,2}^{(n)}\left(\ell_-^{(n)}\right)$, $G^\infty = Z_{k,2}^\infty$, $\left(X_i^{(n)}\right)_{i\in\mathbb{N}} = \left(e^{-\lambda_k t}Y_i(t)\right)_{i\in\mathbb{N}}$, $f_n = \log(n)$, and where we use Lemma 5.4.1 to get the scaling $h_n = n^{T_1\lambda_k}$.

□

In the next proposition, we capture the first-order asymptotics of the subpopulation $Z_{k,3}^{(n)}$ at time $t\log(n)$ using a martingale approach. Initially, the mutation process producing cells of trait k is of order $\mathcal{O}\left(\log^{\theta(k-1)}(n)\right)$, and then the mutant cells have a time $t\log(n)$ to growth exponentially fast at rate λ_k. These two heuristics combined gives that the typical order of $Z_{k,3}^{(n)}(t\log(n))$ is $\mathcal{O}\left(n^{t\lambda_k}\log^{\theta(k-1)}(n)\right)$.

Proposition 5.4.9. *For all* $0 < T_1 < T_2$

$$\left(\frac{Z_{k,3}^{(n)}\left(t\log(n)\right)}{n^{t\lambda_k}\log^{\theta(k-1)}(n)}\right)_{t\in[T_1,T_2]} \underset{n\to\infty}{\longrightarrow} \left(Z_{k,3}^\infty\right)_{t\in[T_1,T_2]},$$

in probability in $L^\infty([T_1,T_2])$*, where* $Z^\infty_{k,3} := \frac{2\alpha_{k-1}\mu_{k-1}}{\lambda_k-\lambda_0}Wa_0$*, where* a_0 *is the constant coefficient of the polynomial function* $u \mapsto \widetilde{w}_{(0,\cdots,k-1)}(t(k)+u)$*.*

Proof. Define the following martingale

$$M^{(n)}_{k,3}(t) := e^{-\lambda_k t}Z^{(n)}_{k,3}(t) - \int_0^t 2\alpha_{k-1}\mu^{(n)}_{k-1}e^{-\lambda_k s}Z^{(n)}_{k-1}\left(t(k)\log(n)+s\right)ds,$$

whose quadratic variation is

$$\left\langle M^{(n)}_{k,3}\right\rangle_t = \int_0^t 2\alpha_{k-1}\mu^{(n)}_{k-1}e^{-2\lambda_k s}Z^{(n)}_{k-1}\left(t(k)\log(n)+s\right)ds + (\alpha_k+\beta_k)\int_0^t e^{-2\lambda_k s}Z^{(n)}_{k,3}(s)ds.$$

These two facts follows from adapting the proof of Lemma 4.3.5 to this specific situation. Then we have

$$\mathbb{P}\left(\sup_{t\in[T_1,T_2]}\left|\frac{Z^{(n)}_{k,3}\left(t\log(n)\right)}{n^{t\lambda_k}\log^{\theta(k-1)}(n)} - Z^\infty_{k,3}\right| \geq 2\varepsilon\right) \leq \mathbb{P}\left(\sup_{t\in[T_1,T_2]}\frac{\left|M^{(n)}_{k,3}(t)\right|}{\log^{\theta(k-1)}(n)} \geq \varepsilon\right) \tag{5.4.13}$$

$$+\mathbb{P}\left(\sup_{t\in[T_1,T_2]}\left|\frac{\int_0^{t\log(n)}2\alpha_{k-1}\mu^{(n)}_{k-1}e^{-\lambda_k s}Z^{(n)}_{k-1}\left(t(k)\log(n)+s\right)ds}{\log^{\theta(k-1)}(n)} - Z^\infty_{k,3}\right| \geq \varepsilon\right). \tag{5.4.14}$$

We start by dealing with the term (5.4.13), where in the computations C is a constant that may change from line to line. By using Doob's Inequality, Lemma 4.3.9 and that $\mu^{(n)}_{\otimes,k}n^{\lambda_0 t(k)} \underset{n\to\infty}{\to} \mu_{\otimes,k} = \prod_{i=0}^{k-1}\mu_i < \infty$ as well as $\mathbb{E}\left[e^{-\lambda_k s}Z^{(n)}_{k,3}(s)\right] \leq C\log^{\theta(k-1)}(n)$ for all $s \leq T_2\log(n)$ that

$$\begin{aligned}
\mathbb{P}\left(\sup_{t\in[T_1,T_2]}\frac{\left|M^{(n)}_{k,3}(t)\right|}{\log^{\theta(k-1)}(n)} \geq \varepsilon\right) &\leq \frac{C}{\log^{2\log(k-1)}(n)}\mathbb{E}\left[\left\langle M^{(n)}_{k,3}\right\rangle_{T_2\log(n)}\right] \\
&\leq \frac{C}{\log^{2\theta(k-1)}(n)}\left(\int_0^{T_2\log(n)}2\alpha_{k-1}\mu^{(n)}_{k-1}e^{-2\lambda_k s}\mathbb{E}\left[Z^{(n)}_{k-1}\left(t(k)\log(n)+s\right)\right]ds\right. \\
&\qquad\left. + (\alpha_k+\beta_k)\int_0^{T_2\log(n)}e^{-2\lambda_k s}\mathbb{E}\left[Z^{(n)}_{k,3}(s)\right]ds\right) \\
&\leq \frac{C}{\log^{2\theta(k-1)}(n)}\left(\int_0^{T_2\log(n)}\left(\mu^{(n)}_{\otimes,k}n^{\lambda_0 t(k)}\right)\left(t(k)\log(n)+s\right)^{\theta(k-1)}e^{-(2\lambda_k-\lambda_0)s}ds\right. \\
&\qquad\left. + \int_0^{T_2\log(n)}\log^{\theta(k-1)}(n)e^{-\lambda_k s}ds\right) \\
&\leq \mathcal{O}\left(\frac{1}{\log^{\theta(k-1)}(n)}\right) \\
&\underset{n\to\infty}{\longrightarrow} 0.
\end{aligned}$$

Now we are going to deal with the term (5.4.14). Using the change of variable $s = u\log(n)$ we obtain that

$$\begin{aligned}
&\int_0^{t\log(n)} 2\alpha_{k-1}\mu_{k-1}^{(n)} e^{-\lambda_k s}\frac{Z_{k-1}^{(n)}\left((t(k)\log(n)+s\right)}{\log^{\theta(k-1)}(n)}ds \\
&\qquad = \int_0^t 2\alpha_{k-1}\mu_{k-1}^{(n)} n^{-\lambda_k u}\frac{Z_{k-1}^{(n)}\left((t(k)+u)\log(n)\right)}{\log^{\theta(k-1)}(n)}\log(n)du \\
&\qquad = \int_0^t 2\alpha_{k-1}\left(\mu_{k-1}^{(n)} n^{\ell(k-1)}\right)\frac{Z_{k-1}^{(n)}\left((t(k)+u)\log(n)\right)}{n^{\lambda_0(u+\ell(k-1)/\lambda_0)}\log^{\theta(k-1)}(n)}n^{-(\lambda_k-\lambda_0)u}\log(n)du.
\end{aligned}$$

Now introduce the event

$$A_{\widetilde{\varepsilon}}^{(n)} := \left\{\sup_{u\in[0,T_2]}\left|\frac{Z_{k-1}^{(n)}\left((t(k)+u)\log(n)\right)}{n^{\lambda_0(u+\ell(k-1)/\lambda_0)}\log^{\theta(k-1)}(n)} - \widetilde{w}_{(0,\cdots,k-1)}(t(k)+u)W\right| \leq \widetilde{\varepsilon}\right\}.$$

We have according to Proposition 5.1.2 that $\mathbb{P}\left(A_{\widetilde{\varepsilon}}^{(n)}\right) \underset{n\to\infty}{\longrightarrow} 1$. Moreover $\widetilde{w}_{(0,\cdots,k-1)}(t(k)+u)$ is a polynomial of degree $\theta(k-1)$, so for the sake of simplicity we will denote it by $P(u) := \sum_{i=0}^{\theta(k-1)} a_i u^i$. We have that

$$\begin{aligned}
&\mathbb{P}\Bigg(\sup_{t\in[T_1,T_2]}\Bigg|\int_0^t 2\alpha_{k-1}\left(\mu_{k-1}^{(n)}n^{\ell(k-1)}\right) \\
&\qquad\qquad \cdot\frac{Z_{k-1}^{(n)}\left((t(k)+u)\log(n)\right)}{n^{\lambda_0(u+\ell(k-1)/\lambda_0)}\log^{\theta(k-1)}(n)}n^{-(\lambda_k-\lambda_0)u}\log(n)du - Z_{k,3}^{\infty}\Bigg| \geq 2\varepsilon\Bigg) \\
&\leq \mathbb{P}\Bigg(\sup_{t\in[T_1,T_2]}\Bigg|\int_0^t 2\alpha_{k-1}\left(\mu_{k-1}^{(n)}n^{\ell(k-1)}\right) \\
&\qquad\qquad \cdot\left(\frac{Z_{k-1}^{(n)}\left((t(k)+u)\log(n)\right)}{n^{\lambda_0(u+\ell(k-1)/\lambda_0)}\log^{\theta(k-1)}(n)} - P(u)W\right)n^{-(\lambda_k-\lambda_0)u}\log(n)du\Bigg| \geq \varepsilon\Bigg) \\
&+ \mathbb{P}\left(\sup_{t\in[T_1,T_2]}\left|\int_0^t 2\alpha_{k-1}\left(\mu_{k-1}^{(n)}n^{\ell(k-1)}\right)P(u)Wn^{-(\lambda_k-\lambda_0)u}\log(n)du - Z_{k,3}^{\infty}\right| \geq \varepsilon\right). \qquad (5.4.15)
\end{aligned}$$

Under the event $A_{\widetilde{\varepsilon}}^{(n)}$ we have

$$\begin{aligned}
&\int_0^t 2\alpha_{k-1}\left(\mu_{k-1}^{(n)}n^{\ell(k-1)}\right)\left(\frac{Z_{k-1}^{(n)}\left((t(k)+u)\log(n)\right)}{n^{\lambda_0(u+\ell(k-1)/\lambda_0)}\log^{\theta(k-1)}(n)}-P(u)W\right)n^{-(\lambda_k-\lambda_0)u}\log(n)du\\
&\leq \int_0^t 2\alpha_{k-1}\left(\mu_{k-1}^{(n)}n^{\ell(k-1)}\right)\widetilde{\varepsilon}n^{-(\lambda_k-\lambda_0)u}\log(n)du\\
&= 2\alpha_{k-1}\left(\mu_{k-1}^{(n)}n^{\ell(k-1)}\right)\widetilde{\varepsilon}\int_0^{t\log(n)}e^{-(\lambda_k-\lambda_0)s}ds\\
&\leq \frac{2\alpha_{k-1}\left(\mu_{k-1}^{(n)}n^{\ell(k-1)}\right)}{\lambda_k-\lambda_0}\widetilde{\varepsilon}\\
&\leq C\widetilde{\varepsilon},
\end{aligned}$$

where C is a constant strictly positive, that comes from the fact that $\mu_{k-1}^{(n)}n^{\ell(k-1)} \underset{n\to\infty}{\longrightarrow} \mu_{k-1} < \infty$. Due to $\mathbb{P}\left(A_{\widetilde{\varepsilon}}^{(n)}\right) \underset{n\to\infty}{\longrightarrow} 1$ for all $\widetilde{\varepsilon} > 0$, one obtains that

$$\begin{aligned}
&\mathbb{P}\Bigg(\sup_{t\in[T_1,T_2]}\Bigg|\int_0^t 2\alpha_{k-1}\left(\mu_{k-1}^{(n)}n^{\ell(k-1)}\right)\\
&\qquad\cdot\left(\frac{Z_{k-1}^{(n)}\left((t(k)+u)\log(n)\right)}{n^{\lambda_0(u+\ell(k-1)/\lambda_0)}\log^{\theta(k-1)}(n)}-P(u)W\right)n^{-(\lambda_k-\lambda_0)u}\log(n)du\Bigg|\geq\varepsilon\Bigg)\\
&\underset{n\to\infty}{\longrightarrow} 0.
\end{aligned}$$

Now we have that

$$\begin{aligned}
&\int_0^t 2\alpha_{k-1}\left(\mu_{k-1}^{(n)}n^{\ell(k-1)}\right)P(u)Wn^{-(\lambda_k-\lambda_0)u}\log(n)du\\
&\qquad=\sum_{i=0}^{\theta(k-1)}\int_0^t 2\alpha_{k-1}\left(\mu_{k-1}^{(n)}n^{\ell(k-1)}\right)a_iWu^in^{-(\lambda_k-\lambda_0)u}\log(n)du
\end{aligned}$$

For the case $i = 0$ we get

$$\begin{aligned}
\int_0^t 2\alpha_{k-1}\left(\mu_{k-1}^{(n)}n^{\ell(k-1)}\right)a_0Wn^{-(\lambda_k-\lambda_0)u}\log(n)du &= \frac{2\alpha_{k-1}\left(\mu_{k-1}^{(n)}n^{\ell(k-1)}\right)a_0W}{\lambda_k-\lambda_0}\left(1-n^{-t(\lambda_k-\lambda_0)}\right)\\
&\underset{n\to\infty}{\longrightarrow}\frac{2\alpha_{k-1}\mu_{k-1}a_0}{\lambda_k-\lambda_0}W.
\end{aligned}$$

For the case $i \in \{1,\cdots,\theta(k-1)\}$ applying i integration by part subsequently, one gets that

$$\sum_{i=1}^{\theta(k-1)}\int_0^t 2\alpha_{k-1}\left(\mu_{k-1}^{(n)}n^{\ell(k-1)}\right)a_iWu^in^{-(\lambda_k-\lambda_0)u}\log(n)du \underset{n\to\infty}{\longrightarrow} 0.$$

Combining these two previous computations, one can show with some extra efforts that the term (5.4.15) converges to 0, which concludes the proof. □

Appendix

Proof of Equation (5.2.2). We define the following function, corresponding to the Laplace transform of the well-renormalized population $Z(t)$

$$L(\xi,t) := \mathbb{E}\left[e^{-\xi e^{-\widetilde{\lambda}t}Z(t)}\right].$$

We have that

$$L(\xi,t) = \mathbb{E}\left[\exp\left(\int_0^t \mu e^{\lambda s}\left(r\left(\xi e^{-\widetilde{\lambda}t}\right)-1\right)ds\right)\right],$$

where

$$r(\xi) := \int_0^t \frac{\mu e^{\lambda s}}{\int_0^t \mu e^{\lambda u}du}\mathbb{E}\left[e^{-\xi Y(t-s)}\right]ds,$$

with Y is a birth and death branching process with birth and death rates α and β respectively, starting with one cell at time 0.
Combining the two previous equations we have

$$L(\xi,t) = \mathbb{E}\left[\exp\left(\mu\int_0^t e^{\lambda s}\left(\mathbb{E}\left[e^{-\xi e^{-\widetilde{\lambda}t}Y(t-s)}\right]-1\right)ds\right)\right].$$

Notice that

$$\mathbb{E}\left[e^{-\xi e^{-\widetilde{\lambda}t}Y(t-s)}\right] = \mathbb{E}\left[e^{-\xi e^{-\widetilde{\lambda}s}\left\{e^{-\widetilde{\lambda}(t-s)}Y(t-s)\right\}}\right],$$

which gives using the Laplace transform of the martingale associated to a birth and death process that

$$\mathbb{E}\left[e^{-\xi e^{-\widetilde{\lambda}t}Y(t-s)}\right]-1 \underset{t\to\infty}{\longrightarrow} -\frac{\xi e^{-\widetilde{\lambda}s}}{1+\xi e^{-\widetilde{\lambda}s}\frac{\alpha}{\widetilde{\lambda}}}.$$

Applying the dominated convergence theorem we obtain that

$$\begin{aligned}
L(\xi,t) &\underset{t\to\infty}{\longrightarrow} \mathbb{E}\left[\exp\left(-\mu\xi\int_0^\infty e^{-(\widetilde{\lambda}-\lambda)s}\frac{1}{1+\xi e^{-\widetilde{\lambda}s}\frac{\alpha}{\widetilde{\lambda}}}ds\right)\right]\\
&= \mathbb{E}\left[\exp\left(-\frac{\mu\xi}{\widetilde{\lambda}}\int_0^\infty e^{-\frac{\widetilde{\lambda}-\lambda}{\widetilde{\lambda}}s}\frac{1}{1+\xi\frac{\alpha}{\widetilde{\lambda}}e^{-s}}ds\right)\right]\\
&= \mathbb{E}\left[\exp\left(-\frac{\mu\xi}{\widetilde{\lambda}}\Phi\left(-\xi\frac{\alpha}{\widetilde{\lambda}},1,\frac{\widetilde{\lambda}-\lambda}{\widetilde{\lambda}}\right)\right)\right],
\end{aligned}$$

where Φ is the Lerch transcendent defined as $\forall\ \Re(s)>0,\ \Re(a)>0$

$$\Phi(z,s,a) = \frac{1}{\Gamma(s)}\int_0^\infty \frac{t^{s-1}e^{-at}}{1-ze^{-t}}dt.$$

□

Bibliography

[1] Brouard, V. and Pokalyuk C. (2022). Invasion of cooperative parasites in moderately structured host populations. *Stochastic Process. Appl.* **153** 221–263 2, 55, 101, 102, 103, 105, 106, 114, 119, 136, 137, 142, 153, 155, 158, 168, 169

[2] Brouard, V., Pokalyuk, C., Seiler, M. and Tran, H. (2024). Spatial invasion of cooperative parasite. *Theoretical Population Biology* **159** 35–58 2, 101

[3] Williams, M. J., Werner, B., Barnes, C. P., Graham, T. A. and Sottoriva, A. (2016). Identification of neutral tumor evolution across cancer types. *Nature Genetics* **48** 238–244 3, 26, 176

[4] Jones, S., Chen, W., Parmigiani, G., Diehl, F., Beerenwinkel, N., Antal, T., Traulsen, A., Nowak, M., Siegel, C., Velculescu, V., and others (2008). Comparative lesion sequencing provides insights into tumor evolution *Proceedings of the National Academy of Sciences* **105** 4283–4288 3, 174

[5] Bozic, I. and Nowak, M. (2014). Timing and heterogeneity of mutations associated with drug resistance in metastatic cancers. *Proceedings of the National Academy of Sciences* **45** 15964–15968 3, 174

[6] Durrett, R. and Mayberry, J. (2011). Traveling waves of selective sweeps. *Ann. Appl. Probab.* **21** 699–744 3, 27, 38, 174, 176, 177, 191

[7] Smadi, C. (2017). The effect of recurrent mutations on genetic diversity in a large population of varying size. *Acta Applicandae Mathematicae* **149** 11–51 3, 174, 177

[8] Cheek,D. and Antal, T. (2018). Mutation frequencies in a birth-death branching process. *Ann. Appl. Probab.* **28** 3922–3947 3, 27, 30, 43, 45, 46, 51, 174, 177, 180, 238, 240, 241, 250, 251

[9] Bovier, A., Coquille, L. and Smadi, C. (2019). Crossing a fitness valley as a metastable transition in a stochastic population model. *Ann. Appl. Probab.* **29** 3541–3589 3, 174, 176, 177

[10] Cheek,D. and Antal, T. (2020). Genetic composition of an exponentially growing cell population. *Stochastic Process. Appl.* **130** 6580–6624 3, 27, 30, 43, 46, 51, 174, 177, 180, 238, 241, 250, 251

[11] Champagnat, N., Méléard, S. and Tran, V. C. (2021). Stochastic analysis of emergence of evolutionary cyclic behavior in population dynamics with transfer. *Ann. Appl. Probab.* **31** 1820–1867 3, 38, 174, 176, 177

[12] Coquille, L., Kraut, A. and Smadi, C. (2021). Stochastic individual-based models with power law mutation rate on a general finite trait space. *Electron. J. Probab.* **26** 1–37 3, 38, 50, 174, 176, 177

[13] Esser, M. and Kraut, A. (2021). A general multi-scale description of metastable adaptive motion across fitness valleys. *arXiv preprint arXiv:2112.12675* 3, 174, 176, 177

[14] Blath, J., Paul, T. and Tóbiás, A. (2023). A Stochastic Adaptive Dynamics Model for Bacterial Populations with Mutation, Dormancy and Transfer. *ALEA, Lat. Am. J. Probab. Math. Stat.* **20** 313–357 3, 174, 176, 177

[15] Paul, T. (2023). The canonical equation of adaptive dynamics in individual-based models with power law mutation rates. *arXiv preprint arXiv:2309.02148* 3, 174, 176, 177

[16] Gamblin, J., Gandon, S., Blanquart, F. and Lambert, A. (2023). Bottlenecks can constrain and channel evolutionary paths. *Genetics* **224** iyad001 3, 174, 177, 178

[17] Esser, M. and Kraut, A. (2023). Effective growth rates in a periodically changing environment: From mutation to invasion. *arXiv preprint arXiv:2310.20509* 3, 50, 174, 176, 177

[18] Bonnet, C. and Leman, H. (2023). Site frequency spectrum of a rescued population under rare resistant mutations. *arXiv preprint arXiv:2303.04069.* 3, 27

[19] Li, A., Kibby, D. and Foo, J. (2023). A comparison of mutation and amplification-driven resistance mechanisms and their impacts on tumor recurrence. *Journal of Mathematical Biology*, **87**(4), 59. 3, 27

[20] Leder, K., Sun, R., Wang, Z., and Zhang, X. (2024). Parameter Estimation from Single Patient, Single Time-Point Sequencing Data of Recurrent Tumors. *arXiv preprint arXiv:2403.13081.* 3, 27

[21] Brouard, V. (2023). Genetic composition of supercritical branching populations under power law mutation rates. *arXiv preprint arXiv:2309.12055* 3, 173

[22] Durrett, R. and Moseley, S. (2010). Evolution of resistance and progression to disease during clonal expansion of cancer. *Theoretical Population Biology* **77** 42–48 3, 27, 43, 45, 194, 225, 238, 240, 249

[23] Nicholson, M. and Tibor, A. (2019). Competing evolutionary paths in growing populations with applications to multidrug resistance. *PLoS Computational Biology* **15** e1006866 3, 27, 43, 177, 238

[24] Nicholson, M., Cheek, D. and Tibor, A. (2023). Sequential mutations in exponentially growing populations. *PLoS Computational Biology* **19** e1011289 3, 27, 43, 45, 48, 49, 178, 238, 240, 244

[25] Tolker-Nielsen, T. and Molin, S. (2000). Spatial Organization of Microbial Biofilm Communities. *Microb Ecol* **40** 75–84 4, 56

[26] Rath, R., Amlinger, L., Rath, A. and Lundgren, M. (2015). The CRISPR-Cas immune system: Biology, mechanisms and applications. *Biochimie* **117** 119–128 4, 55

[27] Camara-Wilpert, S., Mayo-Muñoz, D., Russel, J., Fagerlund, R.D., Madsen, J.S., Fineran, P.C., Sørensen, S. J. and Pinilla-Redondo, R. (2023). Bacteriophages suppress CRISPR–Cas immunity using RNA-based anti-CRISPRs. *Nature* **623** 601–607 5

[28] Landsberger, M., Gandon, S., Meaden, S., Rollie, C., Chevallereau, A., Chabas, H., Buckling, A., Westra, E.R. and van Houte, S. (2018). Anti-CRISPR phages cooperate to overcome CRISPR-Cas immunity. *Cell* **174** 908–916 5, 56

[29] Borges, A.L., Zhang, J.Y., Rollins, M.F., Osuna, B.A., Wiedenheft, B. and Bondy-Denomy, J. (2018). Bacteriophage cooperation suppresses CRISPR-Cas3 and Cas9 immunity. *Cell* **174** 917–925 5, 56

[30] Volovat, S.R., Scripcariu, D.V., Vasilache, I.A., Stolniceanu, C.R., Volovat, C., Augustin, I.G., Volovat, C.C., Ostafe, M., Andreea-Voichita, S., Bejusca-Vieriu, T. and Others (2024). Oncolytic Virotherapy: A New Paradigm in Cancer Immunotherapy. *International Journal of Molecular Sciences* **25** 1180 5

[31] Rodriguez-Brenes, I.A., Hofacre, A., Hung, F. and Wodarz, D. (2017). Complex dynamics of virus spread from low infection multiplicities: Implications for the spread of oncolytic viruses. *PLoS Computational Biology* **13**(1) 5, 26, 56

[32] Lion, S. and van Baalen, M. (2008). Self-structuring in spatial evolutionary ecology. *Ecol Lett* **11** 277–295 10, 59

[33] Sturm, A. and Swart, J.M. (2015). A particle system with cooperative branching and coalescence. *Ann. Appl. Probab.* **25**(3) pp. 1616–1649 26, 102

[34] Vogelstein, B. and Kinzler, K. W. (2004). Cancer genes and the pathways they control. *Nature Medicine* **10** 789–799 26

[35] Vogelstein, B., Papadopoulos, N., Velculescu, V. E., Zhou, S., Diaz Jr, L. A., and Kinzler, K. W. (2013). Cancer genome landscapes. *Science* **339** 1546–1558 26

[36] Fearon, E. R. (2011). Molecular genetics of colorectal cancer. *Annual Review of Pathology: Mechanisms of Disease* **6** 479–507 26

[37] Morin, P. J., Sparks, A. B., Korinek, V., Barker, N., Clevers, H., Vogelstein, B., and Kinzler, K. W. (1997). Activation of β-catenin-Tcf signaling in colon cancer by mutations in β-catenin or APC. *Science*, **275**(5307), 1787–1790 26

[38] Tomasetti, C., Marchionni, L., Nowak, M. A., Parmigiani, G., and Vogelstein, B. (2015). Only three driver gene mutations are required for the development of lung and colorectal cancers. *Proceedings of the National Academy of Sciences*, **112**(1), 118–123 26

[39] Deininger, M. W. N., Goldman, J. M., and Melo, J. V. (2000). The molecular biology of chronic myeloid leukemia. *Blood, The Journal of the American Society of Hematology*, **96**(10), 3343–3356 26

[40] Sottoriva, A., Kang, H., Ma, Z., Graham, T., Salomon, M., Zhao, J., Marjoram, P., Siegmund, K., Press, M., Shibata, D., and others (2015). A Big Bang model of human colorectal tumor growth. *Nature Genetics* **47** 209–216 26, 176

[41] Ling, S., Hu, Z., Yang, Z., Yang, F., Li, Y., Li, P., Chen, K., Dong, L., Cao, L., Tao, Y., and others (2015). Extremely high genetic diversity in a single tumor points to prevalence of non-Darwinian cell evolution. *Proceedings of the National Academy of Sciences* **112** E6496–E6505 26, 176

[42] Venkatesan, S. and Swanton, C. (2016). Tumor evolutionary principles: how intratumor heterogeneity influences cancer treatment and outcome. *American Society of Clinical Oncology Educational Book* **36** e141–e149 26, 176

[43] Davis, A., Gao, R. and Navin, N. (2017). Tumor evolution: Linear, branching, neutral or punctuated? *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer* **36** e141–e149 26, 176

[44] Zeng, K., Fu, Y., Shi, S. and Wu, C. (2006). Statistical tests for detecting positive selection by utilizing high-frequency variants. *Genetics* **174** 1431–1439 27, 176

[45] Achaz, G. (2009). Frequency spectrum neutrality tests: one for all and all for one. *Genetics* **183** 249–258 27, 176

[46] Gunnarsson, E. B., Leder, K. and Foo, J. (2021). Exact site frequency spectra of neutrally evolving tumors: A transition between power laws reveals a signature of cell viability. *Theoretical Population Biology* **142** 67–90 27, 176

[47] Haeno, H., Iwasa, Y., and Michor, F. (2007). The evolution of two mutations during clonal expansion. *Genetics* **177**(4), 2209–2221 27

[48] Durrett, R., Schmidt, D., and Schweinsberg, J. (2009). A waiting time problem arising from the study of multi-stage carcinogenesis. *Ann. Appl. Probab.* **19**(2) 676–718 27

[49] Durrett, R., Foo, J., Leder, K., Mayberry, J., and Michor, F. (2010). Evolutionary dynamics of tumor progression with random fitness values. *Theoretical Population Biology* **78**(1), 54–66 27

[50] Durrett, R., Foo, J., Leder, K., Mayberry, J., and Michor, F. (2011). Intratumor heterogeneity in evolutionary models of tumor progression. *Genetics* **188**(2), 461–477 27

[51] Foo, J., Leder, K., and Schweinsberg, J. (2020). Mutation timing in a spatial model of evolution. *Stochastic Process. Appl.* **130**(10), 6388–6413 27

[52] Dinh, K. N., Jaksik, R., Kimmel, M., Lambert, A., and Tavaré, S. (2020). Statistical inference for the evolutionary history of cancer genomes. *Statistical Science* **35**(1), 129–144 27

[53] Chao, B., and Schweinsberg, J. (2023). A spatial mutation model with increasing mutation rates. *J Appl. Probab.* **60**(4), 1157–1180 27

[54] Johnson, B., Shuai, Y., Schweinsberg, J., and Curtius, K. (2023). Estimating single cell clonal dynamics in human blood using coalescent theory. *bioRxiv*, 2023–02 27

[55] Zhang, R., and Bozic, I. (2024). Accumulation of Oncogenic Mutations During Progression from Healthy Tissue to Cancer. *bioRxiv*, 2024–02 27

[56] Foo, J. and Leder, K. (2013). Dynamics of cancer recurrence. *Ann. Appl. Probab.* **23** 1437–1468 27, 37, 188, 198, 216

[57] Komarova, N. L. (2005). Drug resistance in cancer: principles of emergence and prevention. *Proceedings of the National Academy of Sciences* **102**(27), 9714–9719 27

[58] Bozic, I., Reiter, J. G., Allen, B., Antal, T., Chatterjee, K., Shah, P., Moon, Y. S., Yaqubie, A., Kelly, N., Le, D. T. and others (2013). Evolutionary dynamics of cancer in response to targeted combination therapy. *eLife* **2**, e00747 27

[59] Avanzini, S., and Antal, T. (2019). Cancer recurrence times from a branching process model. *PLoS Computational Biology* **15**(11), e1007423 27

[60] Durrett, R. (2013). Population genetics of neutral mutations in exponentially growing cancer cell populations. *The Annals of Applied Probability: An Official Journal of the Institute of Mathematical Statistics* **23**(1), 230 27

[61] Tung, H.-R., and Durrett, R. (2021). Signatures of neutral evolution in exponentially growing tumors: A theoretical perspective. *PLoS Computational Biology* **17**(2), e1008701 27

[62] Gunnarsson, E. B., Leder, K. and Zhang, X. (2023). Limit theorems for the site frequency spectrum of neutral mutations in an exponentially growing population. *arXiv preprint arXiv:2307.03346* 27, 178

[63] Schweinsberg, J., and Shuai, Y. (2023). Asymptotics for the site frequency spectrum associated with the genealogy of a birth and death process. *arXiv preprint arXiv:2304.13851* 27

[64] Shuai, Y. (2024). Transition in the ancestral reproduction rate and its implications for the site frequency spectrum. *arXiv preprint arXiv:2402.16306* 27

[65] Thiel, K. (2004). Old dogma, new tricks–21st Century phage therapy. *Nat Biotechnol.* **22** 31–36 55

[66] Allen, B., Lippner, G., Chen, YT., Fotouhi, B., Momeni, N., Yau, SY. and Nowak, M.A. (2017). Evolutionary dynamics on any population structure. *Nature* **544** 227–230 56

[67] Czuppon, P. and Pfaffelhuber, P. (2017). A spatial model for selection and cooperation. *J Appl. Probab.* **54** 522–539 56

[68] Barbour, A. and Reinert, G. (2013). Approximating the epidemic curve. *Electron J Probab* **18** 1–30 56, 57, 63

[69] Britton, T. and Pardoux, E. (2019). *Stochastic Epidemic Models with Inference.* Springer 56, 57, 60

[70] Bernstein, E., Hamblen, C., Junge, M. and Ree, L. (2022). Chase-escape on the configuration model. *Electron. Commun. Probab.* **27** 1–14 56

[71] Levin, D.A., Peres, Y. and Elizabeth, L. (2017). *Markov Chains and Mixing Times.* AMS, 2nd. rev. ed. 73

[72] Rubenstein, D. and Kealey, J. (2010). Cooperation, Conflict, and the Evolution of Complex Animal Societies. *Nature Education Knowledge* 3(10):78 102

[73] Landsberger, M., Gandon, S., Meaden, S., Rollie, C., Chevallereau, A., Chabas, H., Buckling, A., Westra, E. R., and van Houte, S.(2018). Anti-CRISPR phages cooperate to overcome CRISPR-Cas immunity. *Cell*, **174**(4):908 –916 102

[74] Borges, A. L., Zhang, J. Y., Rollins, M. F., Osuna, B. A., Wiedenheft, B., and Bondy-Denomy, J. (2018). Bacteriophage cooperation suppresses Crispr-Cas3 and Cas9 immunity. *Cell*, **174**(4):917–925 102

[75] González Casanova, A., Pardo, J. and Pérez, J. (2021). Branching processes with interactions: Subcritical cooperative regime. *Advances in Applied Probability* **53**(1) 251–278 102, 104

[76] Neuhauser, C. (1994). A long range sexual reproduction process. *Stochastic Process. Appl.* **53**(2) 193–220 102

[77] Noble, C. (1992). Equilibrium behavior of the sexual reproduction process with rapid diffusion. *The Annals of Probability* **20**(2) 724–745 102

[78] Mach, T., Sturm, A. and Swart, J. M. (2020). Recursive tree processes and the mean-field limit of stochastic flows. *Electron. J. Probab.* **25** 1–63 102

[79] Sevast'yanov, B. A. (1949). On certain types of Markov processes. *Uspekhi Mat. Nauk* **4**(4) 104

[80] Kalinkin, A. V. (2002). Markov branching processes with interaction. *Russian Math. Surveys.* **57** (2) 104

[81] Kalinkin, A. V. (2003). Extinction probability of a branching process with two kinds of interaction of particles. *Theory Probab. Appl.* **46** (2) 104

[82] Berzunza Ojeda, G. and Pardo, J. (2020). Branching processes with pairwise interactions. *Preprint* Arxiv: 2009.11820v2 104

[83] Chen, A., Li, J., Chen, Y. and Zhou, D. (2012). Extinction probability of interacting branching collision processes. *Adv. Appl. Prob.* **44** 104

[84] Chavel, I. (1984). Eigenvalues in Riemannian geometry. *Pure and Applied Mathematics* **115**, Orlando 110

[85] Samet, H. (1984). The quadtree and related hierarchical data structures. *ACM Comput. Surv.* **16**(2) 111

[86] Van Der Hofstad, R. (2016). *Random graphs and complex networks.* Cambridge university press. 161

[87] Gao, F. (1996). Moderate deviations for martingales and mixing random processes. *Stochastic Process. Appl.* **61**(2) 263–275 169, 170

[88] Luria, S. E. and Delbrük, M. (1943). Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* **28** 491 177

[89] Lea, D. E. and Coulson, C. A. (1949). The distribution of the numbers of mutants in bacterial populations. *Journal of Genetics* **49** 264–285 177

[90] Kendall, D. G. (1960). Birth-and-death processes, and the theory of carcinogenesis. *Biometrika* **47** 113–21 177

[91] Hamon, A. and Ycart, B. (2012). Statistics for the Luria-Delbrück distribution. *Electronic Journal of Statistics* **6** 1251–1272 177

[92] Keller, P. and Tibor, A. (2015). Mutant number distribution in an exponentially growing population. *Journal of Statistical Mechanics: Theory and Experiment* **2015** P01011 177

[93] Kessler, D. A. and Levine, H. (2015). Scaling solution in the large population limit of the general asymmetric stochastic Luria–Delbrück evolution process. *Journal of statistical physics* **158** 783–805 177

[94] Durrett, R. (2015). *Branching process models of cancer.* Springer. 194, 225, 249

[95] Le Gall, J. (2016). *Brownian motion, martingales, and stochastic calculus.* Springer. 195, 257

[96] Dellacherie, C. and Mayer, P. (1975). *Probabilités et Potentiel.* Collection Enseignement des Sciences. Hermann, Paris. 208

Abstract:

This thesis focuses on the study of two stochastic models related to medical problems. The first one lies on understanding infection spread of cooperating bacteriophages on a structured multi-drug resistant bacterial host population. Motivated by this example, we introduce an epidemiological model where infections are generated by cooperation of parasites in a host population structured on a configuration model. We analysed the invasion probability for which we obtain a phase transition depending on the connectivity degree of the vertices and the offspring number of parasites during an infection of a host. At the critical scaling, the invasion probability is identified as the survival probability of a Galton-Watson process.

With the aim to get a biological more relevant model, we analysed a similar model where a spatial structure is added for the host population using a random geometric graph. We have shown that such spatial structure facilitates cooperation of parasites. A similar phase transition occurs where at the same critical scaling the invasion probability is upper and lower bounded by the survival probabilities of two discrete branching processes with cooperation.

The second medical question deals with understanding the evolution of the genetic composition of a tumour under carcinogenesis, using multitype birth and death branching process models on a general finite trait space. In the case of neutral and deleterious cancer evolution, we provide first-order asymptotics results on all mutant subpopulation sizes. In particular such results capture the randomness of all cell trait sizes when a tumour is clinically observed, and mostly it allows to characterise the effective evolutionary pathways, providing information on the past, present, and future of tumour evolution.

Moving beyond this restrictive neutral and deleterious cancer evolution framework, we provide a new method to understand the first selective mutant trait size.

Résumé:

Cette thèse porte sur l'étude de deux modèles stochastiques liés à des problèmes médicaux. Le premier vise à comprendre le processus épidémique généré par des bactériophages coopératifs dans une population de bactéries résistantes aux antibiotiques. Pour cela, nous introduisons un modèle épidémiologique où les infections sont générées par la coopération de parasites dans une population d'hôtes structurée selon un modèle de configuration. Une transition de phase est observée pour la probabilité d'invasion dépendant du degré de connectivité des sommets et du nombre de parasites générés lors d'une infection d'un hôte. Au seuil critique, la probabilité d'invasion est identifiée comme la probabilité de survie d'un processus de Galton-Watson.

Dans le but d'obtenir un modèle biologiquement plus pertinent, nous avons analysé un modèle similaire où une structure spatiale est ajoutée à la population d'hôtes en utilisant un "random geometric graph". Nous avons montré qu'une telle structure spatiale facilite la coopération des parasites. Une transition de phase similaire se produit où au seuil critique, des bornes supérieure et inférieure sont obtenues pour la probabilité d'invasion en tant que probabilités de survie de deux processus de branchement avec coopération.

La deuxième question médicale concerne la compréhension de l'évolution de la composition génétique d'une tumeur en formation, en utilisant des processus de naissance et de mort multitypes branchants sur un espace de traits fini. Considérant une évolution neutre et délétère, nous fournissons des résultats au premier ordre asymptotique pour toutes les tailles des sous-populations mutantes. En particulier, nous capturons la stochasticité associée aux tailles des sous-populations mutantes lorsqu'une tumeur est observée cliniquement, et surtout nous caractérisons les chemins évolutifs effectifs, fournissant des informations sur le passé, le présent et le futur de l'évolution tumorale.

Au-delà de ce cadre restrictif d'évolution neutre et délétère, nous proposons une nouvelle méthode pour comprendre le premier ordre asymptotique du premier trait mutant sélectif.