Asymptotic Behavior of Near Critical Branching Processes and Modeling of Cell Growth Data

Dominik Reinhold

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Statistics and Operations Research (Statistics).

Chapel Hill
2011

Approved by:

Advisor: Amarjit Budhiraja

Advisor: M. Ross Leadbetter

Reader: Shankar Bhamidi

Reader: Jan Hannig

Reader: Chuanshu Ji

Abstract

DOMINIK REINHOLD: Asymptotic Behavior of Near Critical Branching Processes and Modeling of Cell Growth Data.
(Under the direction of Amarjit Budhiraja and M. Ross Leadbetter.)

This dissertation is composed of two parts, a theoretical part, in which certain asymptotic properties of near critical branching processes are studied, and an applied part, consisting of statistical analysis of cell growth data.

First, near critical single type Bienaymé-Galton-Watson (BGW) processes are considered. It is shown that, under appropriate conditions, Yaglom distributions of suitably scaled BGW processes converge to that of the corresponding diffusion approximation. Convergences of stationary distributions for Q-processes and models with immigration to the corresponding distributions of the associated diffusion approximations are established as well. Moreover, convergence of Yaglom distributions of suitably scaled multitype subcritical BGW processes to that of the associated diffusion model is established.

Next, near critical catalyst-reactant branching processes with controlled immigration are considered. The catalyst population evolves according to a classical continuous time branching process, while the reactant population evolves according to a branching process whose branching rate is proportional to the total mass of the catalyst. Immigration takes place exactly when the catalyst population falls below a certain threshold, in which case the population is instantaneously replenished to the threshold. A diffusion limit theorem for the scaled processes is established, in which the catalyst limit is a reflected diffusion and the reactant limit is a diffusion with coefficients depending on the reactant.

Stochastic averaging under fast catalyst dynamics are considered next. In the setting where both catalyst and reactant evolve according to the above described (reflected)

diffusions, but where the evolution of the catalyst is accelerated by a factor of n, we establish a scaling limit theorem, in which the reactant process is asymptotically described through a one dimensional SDE with coefficients depending on the invariant distribution of the catalyst reflected diffusion. Convergence of the stationary distribution of the scaled catalyst branching process (with immigration) to that of the limit reflected diffusion is established as well.

Finally, results from a collaborative proof-of-principle study, relating cell growth to the stiffness of the surrounding environment, are presented.

Acknowledgments

I am very grateful to my advisor Professor Amarjit Budhiraja for all his encouragement and patience throughout my studies. His enthusiasm for research and mentoring helped making my stay in Chapel Hill an enormously rewarding and enjoyable experience. He was always available for short and long conversations, which were of great importance for my understanding and finishing of this dissertation.

I am thankful to my advisors Professors M. Ross Leadbetter and Budhiraja for giving me the opportunity to participate in the project *Emerging Frontiers in 3-D Breast Cancer Tissue Test Systems*, part of which is discussed in this dissertation. Their experience and advice was of great importance.

Next, I would like to thank the other members of my committee, Professors Shankar Bhamidi, Jan Hannig, and Chuanshu Ji, for their valuable comments on this dissertation. Special thanks go to Professor Bhamidi for all his suggestions and conversations with me.

Faculty, staff, and my fellow students played an important role in making my stay in Chapel Hill a great experience, and I am very thankful to them.

Finally, I am grateful to my family and friends from near and far for their support and encouragement.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

This dissertation is composed of two parts, a theoretical part, in which we study certain asymptotic properties of near critical branching processes, and an applied part, consisting of statistical analysis of cell growth data as part of the *EFRI-CBE*[1] project *Emerging Frontiers in 3-D Breast Cancer Tissue Test Systems*.

Branching processes have been studied extensively (see [1]) and have proved to be useful for modeling population dynamics in a variety of fields (see [15]). We study here the so-called Bienaymé-Galton-Watson (BGW) processes, their multidimensional analogues, and continuous time catalyst-reactant branching processes. Roughly speaking, a BGW process is a Markov chain $\{Z_n\}_{n\in\mathbb{N}_0}$ with state space $\mathbb{N}_0$ with the following behavior. The process starts with Z_0 *particles*; each of the Z_n particles alive at time n lives for one unit of time and then dies, giving rise to l *offspring* particles with probability p_l, $l \in \mathbb{N}_0$, where $\{p_l\}_{l\in\mathbb{N}_0}$ is a probability distribution, the so-called *offspring distribution*. In the multitype setting of k-BGW processes ($k \in \mathbb{N}$), Z_n is a k-dimensional vector representing the number of k different types of particles, each of which can give birth to particles of k different types according to an offspring distribution that is specific to the parent type. In particular, the particles modeled by k-BGW processes evolve independently of each other. Catalyst-reactant branching processes, on the other hand,

[1] Emerging Frontiers in Research and Innovation - Cellular and Biomolecular Engineering

model populations with a certain kind of interaction (see [14] and references therein). More precisely, they describe dynamics where the catalyst population directly affects the activity level of an associated reactant population. In our model, roughly speaking, we consider a population of catalyst particles, which evolves according to a classical continuous time branching process with constant branching rate (the life time of each particle is exponentially distributed, instead of being constant as in the setting of BGW processes) with a specific form of immigration, and a reactant population whose branching rate is proportional to the total mass of the catalyst population.

We first describe our main results for BGW processes in the single and multitype setting. This work has appeared as "Some asymptotic results for near critical branching processes" in *Communications on Stochastic Analysis* in 2010 ([5]). Next, we describe our results for catalyst-reactant branching processes, and finally give a description of a proof-of-principle study from the EFRI-CBE project ([2]). This project involves researchers from multiple departments and universities, and the work presented here has been conducted jointly with multiple collaborators.

Consider first the single type setting (i.e. $k = 1$) of BGW processes. Depending on the mean m of the offspring distribution, the BGW process is referred to as subcritical, critical, or supercritical, according to whether $m < 1$, $m = 1$, or $m > 1$, respectively.

We are concerned with the scaled processes $\hat{Z}_t^{(n)} = \frac{1}{n} Z_{\lfloor nt \rfloor}^{(n)}$, where $\{Z^{(n)}\}_{n\in\mathbb{N}}$ is a sequence of BGW processes, with offspring means m_n tending to 1, as $n \to \infty$. Our primary interest is in the steady state behavior of $\hat{Z}_t^{(n)}$ as $t \to \infty$. However, in order to obtain a meaningful asymptotic limit (as $t \to \infty$), one needs to suitably reformulate this question since, as is well known, for $m_n > 1$, $\hat{Z}_t^{(n)}$ tends to infinity on the set of non-extinction as $t \to \infty$, and for $m_n \le 1$, $\hat{Z}_t^{(n)}$ eventually becomes extinct (see [1]). There are two common approaches to address the problem of certain extinction of $\hat{Z}^{(n)}$ in the (sub-) critical case. The first is to condition the process $\hat{Z}^{(n)}$ on non-extinction, where, loosely speaking, the conditioning can either be on non-extinction at the present time or in the distant future. The state process $\hat{Z}^{(n)}$ under these two conditionings leads to different limiting distributions as $t \to \infty$. The first is the so-called quasi-stationary distribution of $\hat{Z}^{(n)}$, while the second is the stationary distribution of the Q-process associated with $\hat{Z}^{(n)}$ (see Section I.14 of [1]).

The second approach to deal with eventual extinction (in the (sub-) critical setting) is to introduce an immigration component, namely, in each generation a (random) number of particles that are indistinguishable from the original set of particles is added to the population. The immigration component in particular ensures that the resulting scaled state process $\hat{Z}^{(n)}$ has a non-degenerate stationary probability distribution.

In the supercritical case, in order to obtain a nontrivial limiting behavior, one typically conditions on the intersection of the event of non-extinction at the present time and the event of eventual extinction. It turns out that a supercritical BGW process with this conditioning has the same distribution as a certain subcritical process. This observation enables us to transfer the results for subcritical processes to supercritical processes.

It is well known (see [10], [20]) that, under suitable conditions, $\hat{Z}^{(n)}$ converges weakly to a diffusion ξ. Such a result implies convergence of finite time statistics of $\hat{Z}^{(n)}$ to those of ξ, but does not provide any information on the relationship between the time asymptotic behaviors of $\hat{Z}^{(n)}$ and ξ. The main goal of this work is to make such relationships mathematically precise. In particular, we show that the time asymptotic distribution of $\hat{Z}_t^{(n)}$ under suitable conditioning converges to that of ξ_t under a similar conditioning, as $n \to \infty$. An analogous result for subcritical models with immigration (where no conditioning is required) is also established. The results say that the long time behavior of a BGW process suitably conditioned (or with an immigration component) is well approximated by the long time behavior of the diffusion limit ξ under a similar conditioning (or with an immigration term).

In addition to the results in the single type setting, we have results in a multitype setting as well. Here, the mean offspring matrix $\mathbf{M}$ plays an analogous role to the mean m in the single type case. The $(i, j)^{th}$ component of $\mathbf{M}$ is the expected number of type j offspring from a single particle of type i in one generation. The Perron-Frobenius theorem shows that, under suitable conditions, there exists an eigenvalue $\rho \in \mathbb{R}_+$ of $\mathbf{M}$, such that the absolute value of any other eigenvalue is strictly smaller than ρ. Similar to the single type case, the process is referred to as subcritical, critical, or supercritical, according to whether $\rho < 1$, $\rho = 1$, or $\rho > 1$, respectively. We establish a convergence result for quasi-stationary distributions analogous to that in the single type setting.

We next study catalyst-reactant branching processes. As in the BGW setting, we consider sequences of catalyst and reactant processes, which are denoted by $\{X^{(n)}\}_{n\in\mathbb{N}}$ and $\{Y^{(n)}\}_{n\in\mathbb{N}}$, respectively. Both $X^{(n)}$ and $Y^{(n)}$ are subcritical, with offspring means tending to 1, as $n \to \infty$, and both processes start with n particles. In typical settings (see e.g. [14]), the catalyst evolution is modeled through a classical continuous time branching process, and consequently population dynamics are described until the time the catalyst becomes extinct. In contrast, our work considers a setting where the catalyst population is maintained above a positive threshold through a specific form of controlled immigration. More precisely, when the catalyst population $X^{(n)}$ drops below n, it is instantaneously restored to level n.

There are many settings where controlled immigration models of the above form arise naturally. For example, immunotherapy in which the natural immune response is stimulated has been successful in treating cancer. Instillation of bacteria (bacillus Calmette-Guérin), for instance, has been shown to reduce recurrence of bladder carcinoma (see [19] and references therein). Many aspects of such treatments remain poorly understood, and it is of interest to develop minimally invasive plans of treatment that intervene only when the level of certain substances drop below a some threshold. Another class of examples arise from predator-prey models in ecology, where one may be concerned with the restoration of populations that are close to extinction by reintroducing species when they fall below a certain threshold. In our work, the motivation for the study of such controlled immigration models comes from problems in chemical kinetics where one wants to keep the level of a catalyst above a certain threshold in order to maintain a desirable level of reaction activity.

We will establish a diffusion limit theorem for the scaled processes $(\hat{X}_t^{(n)}, \hat{Y}_t^{(n)}) := (\frac{1}{n}X_{nt}^{(n)}, \frac{1}{n}Y_{nt}^{(n)})$. The limit process (X, Y) will be such that X is a reflected diffusion with reflection at 1, and Y is a diffusion with coefficients depending on X. The driving Brownian motions in the two diffusions will be independent.

The catalyst and reactant populations considered in the results described above evolve on a comparable time scale in the sense that the branching rates of both $X^{(n)}$ and $Y^{(n)}$ converge to positive (possibly different) constants, as $n \to \infty$. In situations in which the catalyst evolves "much faster" than the reactant, it is of interest to find simplified diffusion models that capture the parts of the dynamics one is interested in economically. Such model reductions (see [21] and references therein for the related setting of chemical reaction networks) not only help in better understanding the dynamics of the system, but also help to reduce computational costs in simulations. In our work we consider a simplified setting of catalyst and reactant populations that evolve according to the above described (reflected) diffusions X and Y, but where the evolution of the catalyst is accelerated by a factor of n (i.e. drift and diffusion coefficients depend on n). We establish a scaling limit theorem, as $n \to \infty$, in which the reactant process is asymptotically described through a one dimensional SDE with coefficients depending on the invariant distribution of X. Averaging results of a similar form in the more realistic setting where the catalyst and reactant populations are described through branching processes will be a topic for future research. In this dissertation, we take a key step towards such a research program, which is to establish the convergence of the stationary distribution of the scaled catalyst process $\hat{X}^{(n)}$ to that of the limit reflecting diffusion X.

The second part of this dissertation is concerned with the statistical analysis of cell growth, metabolic activity, viability, and morphology data from a proof-of-principle study as part of the aforementioned EFRI-CBE project. The overall goals of the project are twofold. We aim to enhance the knowledge of the relationships between normal and breast cancer cellular behavior and impact on cell growth behavior of factors such as tissue stiffness and oxygen level. The longterm goal of the project is to develop bioengineering tools to build tissues (or their *in vitro* representatives, *hydrogels*) and to assess experimentally and analytically the above relationships. The study presented here ([2]) is concerned with the relationships between cell characteristics (cell growth, metabolic activity, and aggregation) and environment (tissue stiffness), and our contribution is primarily in the modeling of these relationships. The results suggest that not only stiffness, but also other characteristics of the experimental setup influence cell behavior. In a first *Metabolic Activity and Viability* experiment, cells were suspended in a hydrogel, and the cells' metabolic activity and viability were measured. In a second *Morphology* experiment, cells were seeded in a monolayer on top of a hydrogel, and the morphology of cell aggregates was observed. The first experiment is in a 3-dimensional setting, whereas the second experiment is in a "2.5D" setting (the hydrogel is 3D, while the cell layer is approximately 2D). Moreover, the constitution of the hydrogels in the two experiments differed. One difference between the two experiments was in cell aggregation, which was more pronounced in weaker gels in the *Metabolic Activity and Viability* experiment, whereas it appeared to be more pronounced in stiffer gels in the *Morphology* experiment. While substrate stiffness alone in a 2D substrate would often dictate greater cell proliferation, this study indicates the importance of considering additional factors which influence cell behavior in a 3D system, such as proteolysis susceptibility and pore size.

The enormous potential of a 3D test system is in the ability to construct cells in a biomaterial substrate in a spatially meaningful manner that allows cellular behavior that is more indicative of behavior in native tissue than a 2D system, thus allowing rapid discoveries of therapies and preventatives for an array of diseases. This proof-of-principle work demonstrates the possibility of using image processing and statistical modeling to describe pseudo-3D cancer systems in a non-destructive and spatio-temporal manner.

The chapters are organized as follows. At the end of this section we list notation that will be used throughout the dissertation.

In Chapter 2 we study asymptotic behavior of single and multitype BGW processes.

Chapter 3 is concerned with the weak convergence of suitably scaled catalyst-reactant branching processes to (reflected) diffusions.

In Chapter 4 we establish an averaging principle under fast catalyst dynamics and also study convergence of invariant distributions of the catalyst processes.

In Chapter 5, we summarize results from a proof-of-principle study from the EFRI-CBE project.

Chapter 6 is an appendix in which we provide proofs of some auxiliary lemmas that are used in preceding chapters.

1.1 Notation

$\mathbb{N} := \{1, 2, \ldots\}$

$\mathbb{N}_0 := \{0, 1, 2, \ldots\}$

$\mathbb{N}^k := \{\mathbf{i} \equiv (i_1, \ldots, i_k)' | i_\alpha \in \mathbb{N}, 1 \leq \alpha \leq k\}$

$\mathbb{N}_0^k := \{\mathbf{i} \equiv (i_1, \ldots, i_k)' | i_\alpha \in \mathbb{N}_0, 1 \leq \alpha \leq k\}$

$\mathbb{N}_{>0}^k := \mathbb{N}_0^k \backslash \{\mathbf{0}\}$

$\mathbb{R}_+^k := \{\mathbf{i} \equiv (i_1, \ldots, i_k)' | i_\alpha \in [0, \infty), 1 \leq \alpha \leq k\}$

$\mathbb{R}_{\geq 1} := [1, \infty)$

$\mathbb{Z} := \{\ldots, -2, -1, 0, 1, 2, \ldots\}$

$\mathbf{0} := (0, 0, \ldots, 0)'$

$\mathbf{1} := (1, 1, \ldots, 1)'$

$$\delta_{ij} := \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

$$1_A(x) := \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{otherwise} \end{cases}$$

$\mathbf{e}_\alpha := (\delta_{1\alpha}, \ldots, \delta_{k\alpha})'$

$\mathbf{s}^{\mathbf{i}} := \prod_{\alpha=1}^k s_\alpha^{i_\alpha}$, for $\mathbf{i} = (i_1, \ldots, i_k)' \in \mathbb{N}_0^k$ and $\mathbf{s} = (s_1, \ldots, s_k)' \in \mathbb{R}_+^k$.

$C^l(\mathbb{S})$: The set of l-times continuously differentiable, real-valued functions on $\mathbb{S}$.

$C_c^l(\mathbb{S})$: The set of l-times continuously differentiable, real-valued functions on $\mathbb{S}$ that have compact support.

$C(\mathbb{R}_+ : \mathbb{S})$: the space of continuous functions from $\mathbb{R}_+$ to $\mathbb{S}$.

$D(\mathbb{R}_+ : \mathbb{S})$: the space of càdlàg (or RCLL – right continuous with left limits) functions from $\mathbb{R}_+$ to $\mathbb{S}$.

$\mathcal{P}(\Omega)$: The set of probability measures on a fixed measurable space $(\Omega, \mathcal{F})$.

$\mathbb{S}_n := \{\frac{l}{n} | l \in \mathbb{N}_0\}$

$\mathbb{S}_X^{(n)} := \{\frac{l}{n} | l \in \mathbb{N}_0\} \cap [1, \infty)$

$\mathbb{S}_Y^{(n)} := \{\frac{l}{n} | l \in \mathbb{N}_0\}$

$\mathbb{S}_Z^{(n)} := \{\frac{l}{n} | l \in \mathbb{Z}\}$

$\mathbb{W}^{(n)} := \mathbb{S}_X^{(n)} \times \mathbb{S}_Y^{(n)} \times \mathbb{S}_Z^{(n)}$

$\mathbb{W} := \mathbb{R}_{\geq 1} \times \mathbb{R}_+ \times \mathbb{R}$

$|| \cdot ||$: the Euclidean norm on $\mathbb{R}^k$.

$|X|_{*,t} := \sup_{0 \leq s \leq t} |X_s|$

Chapter 2

Asymptotic Behavior of Near Critical Branching Processes

The following results appeared as "Some asymptotic results for near critical branching processes" in *Communications on Stochastic analysis* in 2010 ([5]).

2.1 Introduction and Main Results

Consider a population consisting of k types of particles whose evolution is described in terms of a discrete time multitype (k-type) Bienaymé-Galton-Watson (k-BGW) process – such a process is a Markov chain $\{Z_p\}_{p\in\mathbb{N}_0}$ on $\mathbb{N}_0^k$, with the vector Z_p representing the number of particles of each type in generation p. We are interested in the long time behavior of the scaled process $\frac{1}{p}Z_{\lfloor pt\rfloor}$, $t \geq 0$, when the k-BGW process is close to criticality. More precisely, we consider a sequence of BGW processes $\{Z_p^{(n)}, p \in \mathbb{N}_0\}_{n\in\mathbb{N}}$ such that, as n becomes large, the processes approach criticality. It is well known (see [10], [20]) that, under suitable conditions, the process $X_t^{(n)} = \frac{1}{n}Z_{\lfloor nt\rfloor}^{(n)}$, $t \geq 0$, converges weakly to a diffusion ξ. (We note here that in Chapter 1 the scaled process $\frac{1}{n}Z_{\lfloor nt\rfloor}^{(n)}$ was denoted by $\hat{Z}_t^{(n)}$. However, throughout this chapter we will use $X_t^{(n)}$ to denote this scaled process.) Such a result implies convergence of finite time statistics of $X^{(n)}$ to those of ξ, but does not provide any information on relationships between the time asymptotic behaviors of $X^{(n)}$ and ξ. The main goal of this work is to make such relationships mathematically precise. In particular, we show that, under appropriate assumptions, the time asymptotic distribution of $X_t^{(n)}$ with suitable conditioning converges to that of ξ_t with a similar conditioning, as $n \to \infty$ (see Theorems 2.1.4 and 2.1.7). An analogous result for models with immigration (where no conditioning is required) is also established (Theorem 2.1.10). The results say that the long time behavior of a BGW process is well approximated by that of the corresponding diffusion limit ξ. Most of the results in this work are for single type BGW processes, namely for the case $k = 1$. Similar results can be obtained in multitype settings and we consider one such result in Theorem 2.1.18.

When $k = 1$, the transition probabilities of a BGW process $\{Z_p\}$ can be written as

$$p(i,j) = P(Z_{p+1} = j | Z_p = i) = \begin{cases} p_j^{*i} & \text{if } i \geq 1, \quad j \geq 0, \\ \delta_{0j} & \text{if } i = 0, \quad j \geq 0, \end{cases} \tag{2.1.1}$$

where $\{p_l\}_{l \in \mathbb{N}_0}$ is a given probability function – the offspring distribution of each particle – and $\{p_l^{*i}\}_{l \in \mathbb{N}_0}$ is the i-fold convolution of $\{p_l\}_{l \in \mathbb{N}_0}$. The process starts with Z_0 particles; each of the Z_n particles alive at time n lives for one unit of time and then dies, giving rise to l offspring particles with probability p_l, $l \in \mathbb{N}_0$. The particles behave independently of each other and of the past.

Depending on the mean m of the offspring distribution, BGW processes can be divided into three cases: subcritical, critical, and supercritical, according to whether $m < 1$, $m = 1$, or $m > 1$, respectively.

Consider a sequence of processes $Z^{(n)}$ described as follows. If $Z_0^{(n)} = 1$, then $Z_1^{(n)}$ has the probability generating function (pgf)

$$F^{(n)}(s) = \sum_{l=0}^{\infty} p_l^{(n)} s^l, \quad s \in [0,1], \tag{2.1.2}$$

with mean m_n and variance σ_n^2, where $\{p_l^{(n)}\}_{l \in \mathbb{N}_0}$ is the offspring distribution of $Z^{(n)}$. We denote the p^{th} iterate of $F^{(n)}$ by $F_p^{(n)}$, i.e. for $s \in [0,1]$ and $p \geq 0$

$$F_0^{(n)}(s) = s, \quad F_{p+1}^{(n)}(s) = F^{(n)}(F_p^{(n)}(s)).$$

Let q_n be the extinction probability of $Z^{(n)}$ starting from a single particle, i.e. $q_n = P(Z_p^{(n)} = 0 \text{ for some } p \in \mathbb{N} | Z_0^{(n)} = 1)$. Let

$$X_t^{(n)} := \frac{1}{n} Z_{\lfloor nt \rfloor}^{(n)}, \quad t \in \mathbb{R}_+; \tag{2.1.3}$$

then $\{X_t^{(n)}\}_{t \in \mathbb{R}_+}$ is an $\mathbb{S}_n := \{\frac{l}{n} | l \in \mathbb{N}_0\}$ valued (time inhomogeneous) Markov process with sample paths in $D(\mathbb{R}_+ : \mathbb{S}_n)$, the space of càdlàg functions from $\mathbb{R}_+ := [0,\infty)$ to $\mathbb{S}_n$. Throughout, $\mathbb{S}_n$ is endowed with the discrete topology and, given a metric space S, $D(\mathbb{R}_+ : S)$ is endowed with the usual Skorohod topology. Space of probability measures on a metric space S will be denoted by $\mathcal{P}(S)$.

Condition 2.1.1. *(i) For each n, $p_0^{(n)} > 0$, $p_0^{(n)} + p_1^{(n)} < q_n$, $m_n = 1 + \frac{c_n}{n}$, $c_n \in (-n, \infty) \setminus \{0\}$, and $\sigma_n^2 < \infty$. (ii) As $n \to \infty$, $c_n \to c \in \mathbb{R} \setminus \{0\}$ and $\sigma_n^2 \to \sigma^2 \in (0,\infty)$. (iii) The family of functions $\{F^{(n)''}\}_{n \in \mathbb{N}}$ is equicontinuous at 1. (iv) As $n \to \infty$, $\sum_{l: l > \epsilon\sqrt{n}} (l - m_n)^2 p_l^{(n)} \to 0$, and $X_0^{(n)}$ converges in distribution to some $\mu \in \mathcal{P}(\mathbb{R}_+)$.*

Condition 2.1.1 (i) ensures that, as $n \to \infty$, $m_n \to 1$, and thus the processes approach criticality without being critical. The case where $c < 0$ will be referred to as the subcritical case while $c > 0$ corresponds to the supercritical case. Condition 2.1.1 (iii) will be used in the study of the supercritical case in Theorem 2.1.4. Condition 2.1.1 (iv) is needed for the diffusion approximation result in Theorem 2.1.2.

We now recall a well known weak convergence result for $X^{(n)}$ (see [12], [20, Theorem 4.2.2]), which describes the asymptotic behavior of $X^{(n)}$, as $n \to \infty$, over any fixed finite time horizon. Here we only give the result in a one dimensional setting. The multidimensional result will be presented later in this section.

Theorem 2.1.2. *[12, 20] Assume Condition 2.1.1. Then $X^{(n)}$ converges weakly in $D(\mathbb{R}_+ : \mathbb{R}_+)$ to the unique (in law) diffusion process ξ with generator*

$$(Lf)(x) = xcf'(x) + \frac{1}{2}x\sigma^2 f''(x), \quad f \in C^2(\mathbb{R}_+), \quad x \in \mathbb{R}_+, \tag{2.1.4}$$

and initial distribution (i.e. probability law of ξ_0) equal to μ.

We are concerned with the study of relationships between the steady state behavior of $X^{(n)}$ and that of ξ. However, one needs to suitably interpret the term "steady state" since, as is well known, as $t \to \infty$, for $m_n > 1$, $X_t^{(n)}$ tends to infinity on the set of non-extinction, and for $m_n \le 1$, $X_t^{(n)}$ eventually becomes extinct (see [1]). There are two well studied approaches for formulating time asymptotic questions in the subcritical case. The first is to condition the processes $X^{(n)}$ on non-extinction, where, loosely speaking, the conditioning can either be on non-extinction at the present time or in the distant future. The state process $X^{(n)}$ under these two conditionings has different limiting distributions as $t \to \infty$. The first is called the Yaglom distribution of $X^{(n)}$, while the second is the stationary distribution of the Q-process associated with $X^{(n)}$ (see Section I.14 of [1]). The second approach for obtaining a nontrivial time asymptotic behavior is to introduce an immigration component. Namely, in each generation a (random) number of particles that are indistinguishable from the original set of particles is added to the population. The immigration component in particular ensures that the resulting scaled state process, denoted by $V^{(n)}$, has a non-degenerate stationary distribution. For the supercritical case, a common approach is to reduce the problem to that of a subcritical setting by conditioning on the event of eventual extinction. The so conditioned state process $X^{(n)}$ has the same law as the state process corresponding to a certain subcritical BGW process. In this work we will show that the time asymptotic distribution of $X_t^{(n)}$ (in both subcritical and supercritical settings), under suitable conditioning, converges to that of ξ_t under a similar conditioning, as $n \to \infty$. For models with immigration we will prove convergence of stationary distributions.

We begin by describing results for models without immigration. For a Markov process $\{Y_t\}_{t\in\mathbb{R}_+}$ with initial value $Y_0 = y$, we write $P(Y_t \in \cdot)$ as $P_y(Y_t \in \cdot)$. Similarly, when the distribution of Y_0 is μ, we write $P(Y_t \in \cdot)$ as $P_\mu(Y_t \in \cdot)$. Similar notations will be used for conditional expectations. Let $\mathbb{S}$ be a subset of $\mathbb{R}_+^k$, for some $k \in \mathbb{N}$. When $\mathbb{S}$ is endowed with a topology, we will denote by $\mathcal{B}(\mathbb{S})$ the σ-field generated by the open sets of $\mathbb{S}$. Let $Y \equiv \{Y_t\}_{t\in\mathbb{R}_+}$ be an $\mathbb{S}$-valued Markov process such that $\mathbf{0} \in \mathbb{S}$ is an absorbing state.

Definition 2.1.1. (i) A quasi-stationary distribution (qsd) for Y is a probability distribution μ on $(\mathbb{S}, \mathcal{B}(\mathbb{S}))$ such that $P_\mu(Y_t \in B | t < T_Y < \infty) = \mu(B)$ for all $B \in \mathcal{B}(\mathbb{S})$ and $t \geq 0$, where $T_Y := \inf\{t | Y_t = \mathbf{0}\}$.
(ii) If for all $\mathbf{y} \in \mathbb{S} \setminus \{\mathbf{0}\}$, as $t \to \infty$, $P_{\mathbf{y}}(Y_t \in \cdot | t < T_Y < \infty)$ converges weakly to some probability measure μ on $(\mathbb{S}, \mathcal{B}(\mathbb{S}))$, then μ is called the Yaglom distribution of Y.

The following result follows from [23] and Proposition 2.3.2.1 of [24].

Theorem 2.1.3. *[23, 24] The Yaglom distribution of ξ exists and is Exponential with density*

$$f(x) = \frac{2|c|}{\sigma^2} \exp\left(-\frac{2|c|}{\sigma^2} x\right), \quad x \geq 0. \tag{2.1.5}$$

Our first result, Theorem 2.1.4 below, says that the Yaglom distribution of $X^{(n)}$ approaches that of ξ, as $n \to \infty$. Note that the existence of the Yaglom distribution of $X^{(n)}$ in the subcritical case is a direct consequence of Theorem V.4.2 of [1].

Theorem 2.1.4. *Assume Condition 2.1.1. For each n, $X^{(n)}$ has a Yaglom distribution $\nu^{(n)}$. This distribution is also a qsd, and it converges weakly to the Yaglom distribution ν of ξ.*

We now consider the second form of conditioning where one conditions the process on not being extinct in the "distant future". We will see that in this case a somewhat different asymptotic behavior emerges. For this result we restrict ourselves to the subcritical case (i.e. $c_n < 0$). We begin with the definition of a Q-process (see [1], [23]).

Let $\hat{\Omega} = D(\mathbb{R}_+ : \mathbb{R}_+)$ and $\hat{\mathcal{F}}$ be the corresponding Borel σ-field (with the usual Skorohod topology). Denote by $\{\mathcal{F}_t\}_{t\in\mathbb{R}_+}$ the canonical filtration on $(\hat{\Omega}, \hat{\mathcal{F}})$, i.e. $\mathcal{F}_t = \sigma(\pi_s : s \leq t)$, where $\pi_s(x) = x_s$ for $x \in \hat{\Omega}$. We denote by $\hat{P}^{(n)}_\mu$ the measure induced by $X^{(n)}$ on $(\hat{\Omega}, \hat{\mathcal{F}})$ when $Z^{(n)}_0$ has distribution μ (supported on $\mathbb{N}$). Let $T := \inf\{t | \pi_t = 0\}$.

By Lemma 6.0.3 in the appendix, there is a probability measure $P^{(n)\uparrow}_\mu$ on $(\hat{\Omega}, \hat{\mathcal{F}})$ such that, as $s \to \infty$, $\hat{P}^{(n)}_\mu(\Theta | T > s) \to P^{(n)\uparrow}_\mu(\Theta)$, for all $\Theta \in \mathcal{F}_t$, $t \in \mathbb{R}_+$. Furthermore if $\{Z^{(n)\uparrow}_k\}_{k\in\mathbb{N}_0}$ is a Markov chain with state space $\mathbb{N}$, l-step transition function

$$p^{(n)\uparrow}_l(i,j) = P(Z^{(n)}_l = j | Z^{(n)}_0 = i)\frac{j}{i} m_n^{-l}, \quad i, j \in \mathbb{N},$$

and initial distribution μ, then $P^{(n)\uparrow}_\mu$ is the law of $\{X^{(n)\uparrow}_t\}_{t\in\mathbb{R}_+}$, where $X^{(n)\uparrow}_t := \frac{1}{n} Z^{(n)\uparrow}_{\lfloor nt \rfloor}$, $t \in \mathbb{R}_+$. The process $Z^{(n)\uparrow}$ [respectively $X^{(n)\uparrow}$] is called the Q-process associated with $Z^{(n)}$ [respectively $X^{(n)}$]. Q-processes associated with branching processes can be interpreted as branching processes conditioned on being never extinct.

Next, we introduce the Q-process associated with the diffusion ξ from Theorem 2.1.2. Denote by $P_{\xi,x}$ the measure induced by ξ on $(\hat{\Omega}, \hat{\mathcal{F}})$, where $\xi(0) = x > 0$. The following theorem is contained in [23].

Theorem 2.1.5. *[23] There is a probability measure $P^{\uparrow}_{\xi,x}$ on $(\hat{\Omega}, \hat{\mathcal{F}})$, such that for all $t \in \mathbb{R}_+$ and $\Theta \in \mathcal{F}_t$, $P_{\xi,x}(\Theta | T > s)$ converges to $P^{\uparrow}_{\xi,x}(\Theta)$, as $s \to \infty$. Let $\xi^\uparrow$ be the unique weak solution of the SDE*

$$d\xi^\uparrow_t = c\xi^\uparrow_t dt + \sqrt{\sigma^2 \xi^\uparrow_t} dB_t + \sigma^2 dt, \quad \xi^\uparrow_0 = x,$$

where B is a standard Brownian motion. Then $P^{\uparrow}_{\xi,x}$ equals the measure induced by $\xi^\uparrow$ on

$(\hat{\Omega}, \hat{\mathcal{F}})$.

The process $\xi^{\uparrow}$ is referred to as the Q-process associated with ξ. The following result (see [23], Section 5.2) says that the process $\xi^{\uparrow}$ has a unique stationary distribution, $\nu^{\uparrow}$, which is given as the convolution of two copies of the exponential distribution ν with density as in (2.1.5).

Theorem 2.1.6. *[23] Assume $c < 0$. As $t \to \infty$, for every initial condition x, $\xi_t^{\uparrow}$ converges in distribution to a random variable $\xi_\infty^{\uparrow}$, whose distribution, denoted by $\nu^{\uparrow}$, is the convolution of two copies of the Yaglom distribution ν. In particular, $\nu^{\uparrow}$ has density*

$$f(x) = \left(\frac{2c}{\sigma^2}\right)^2 x \exp\left(\frac{2c}{\sigma^2} x\right), \quad x \geq 0. \tag{2.1.6}$$

Our next result shows that the time asymptotic behavior of the Q-process associated with $X^{(n)}$ can be well approximated by that of the Q-process associated with the diffusion approximation of $X^{(n)}$. Note that the existence of the stationary distribution of the Q-process $X^{(n)\uparrow}$ is immediate from Theorem I.14.2 in [1].

Theorem 2.1.7. *Assume Condition 2.1.1 and that $c_n < 0$ for all $n \in \mathbb{N}$. For each n, $X_t^{(n)\uparrow}$ converges in distribution, as $t \to \infty$, to a random variable $X_\infty^{(n)\uparrow}$. The distribution $\nu^{(n)\uparrow}$ of $X_\infty^{(n)\uparrow}$ is the unique stationary distribution of the $\mathbb{S}_n$ valued Markov process $X^{(n)\uparrow}$. As $n \to \infty$, $\nu^{(n)\uparrow}$ converges weakly to $\nu^{\uparrow}$.*

We now describe the results for BGW processes with immigration. Let F and G be pgf's of $\mathbb{N}_0$ valued random variables. A Bienaymé-Galton-Watson branching process with immigration corresponding to (F,G) (referred to as a DBI(F, G) process), is a Markov chain $\{Y_n\}$ with state-space $\mathbb{N}_0$ and transition probability function described in terms of the corresponding pgf: Given $Y_0 = i \in \mathbb{N}$, the pgf $H(i, \cdot)$ of Y_1 is $H(i, s) = \sum_{j=0}^{\infty} P(Y_1 = j | Y_0 = i) s^j = F(s)^i G(s)$, $s \in [0, 1]$.

Let $G^{(n)}$ be a sequence of pgf's, and consider a sequence of DBI$(F^{(n)}, G^{(n)})$ processes $Y^{(n)}$.

Condition 2.1.8. *(i) There are $\iota_0, \kappa_0 \in (0, \infty)$ such that, for all $n \in \mathbb{N}$, $G^{(n)\prime}(1) = \iota_n \geq \iota_0$ and $G^{(n)\prime\prime}(1) = \kappa_n \leq \kappa_0$. (ii) As $n \to \infty$, $\iota_n \to \iota$. (iii) There is a $\tau_0 \in [0, \infty)$ such that, for all $n \in \mathbb{N}$, $F^{(n)\prime\prime\prime}(1) = \tau_n < \tau_0$.*

Let $V_t^{(n)} := \frac{1}{n} Y_{\lfloor nt \rfloor}^{(n)}$, $t \in \mathbb{R}_+$. The proof of the following theorem is easy to establish using [23] and [25, Theorem 2.1].

Theorem 2.1.9. *Assume Conditions 2.1.1 and 2.1.8 and that $c < 0$. Suppose that $V_0^{(n)}$ converges in distribution to some $\mu \in \mathcal{P}(\mathbb{R}_+)$. Then $V^{(n)}$ converges weakly in $D(\mathbb{R}_+ : \mathbb{R}_+)$ to the process ζ which is the unique weak solution of*

$$d\zeta_t = c\zeta_t dt + \sqrt{\sigma^2 \zeta_t} dB_t + \iota dt, \quad t \geq 0,$$

where ζ_0 has distribution μ. The Markov process ζ has a unique stationary distribution η, which is a gamma distribution with parameters $2\iota/\sigma^2$ and $\sigma^2/(2|c|)$, i.e., η has density g given as

$$g(x) = x^{\frac{2\iota}{\sigma^2} - 1} \frac{\exp\left(-x \frac{2|c|}{\sigma^2}\right)}{\left(\frac{\sigma^2}{2|c|}\right)^{\frac{2\iota}{\sigma^2}} \Gamma\left(\frac{2\iota}{\sigma^2}\right)}, \quad x > 0.$$

We are interested in the long time behavior of the scaled processes $V^{(n)}$ as they approach criticality. Our main result is the following. Note that the existence of the stationary distribution of $V^{(n)}$ is immediate from [34], p. 414.

Theorem 2.1.10. *Assume Conditions 2.1.1 and 2.1.8 and that $c_n < 0$ for all $n \in \mathbb{N}$. For each $n \in \mathbb{N}$, $V^{(n)}$ has a unique stationary distribution $\eta^{(n)}$, and as $n \to \infty$, $\eta^{(n)}$ converges weakly to η.*

As noted earlier in the introduction, results similar to Theorems 2.1.4, 2.1.6, and 2.1.10 can be established for multitype settings as well. To illustrate the key ideas involved, we only discuss one case in detail, namely the convergence of the Yaglom distribution in the setting of a subcritical multitype process. We begin with some notation and definitions. Let $\{Z_j^{(n)}, j \in \mathbb{N}_0\}_{n\in\mathbb{N}}$ be a sequence of k-BGW processes with transition mechanism described below. Let $C := [0,1]^k$, $\mathbf{e}_\alpha := (\delta_{1\alpha}, \ldots, \delta_{k\alpha})'$ be the α^{th} canonical basis vector, and $\mathbf{s}^{\mathbf{i}} := \prod_{\alpha=1}^k s_\alpha^{i_\alpha}$, for $\mathbf{i} = (i_1, \ldots, i_k)' \in \mathbb{N}_0^k$ and $\mathbf{s} = (s_1, \ldots, s_k)' \in \mathbb{R}_+^k$. Similar to the single type case, the evolution of $Z_j^{(n)} = (Z_{j,1}^{(n)}, \cdots Z_{j,k}^{(n)})'$ is described as follows. For any $\alpha = 1, \ldots, k$, each of the $Z_{j,\alpha}^{(n)}$ type α particles alive at time j (if any) lives for one unit of time and then dies, giving rise to a number of offspring particles, represented by $\mathbf{l} = (l_1, \ldots, l_k)$, l_β being the number of type β offspring, with probability $p^{(n)}(\mathbf{e}_\alpha, \mathbf{l})$. The particles behave independently of each other and of the past. The probability law of $Z^{(n)}$ is given in terms of the pgf $F^{(n)}(\mathbf{s}) := (F_{(1)}^{(n)}(\mathbf{s}), \ldots, F_{(k)}^{(n)}(\mathbf{s}))$, $\mathbf{s} \in C$, where $F_{(\alpha)}^{(n)}(\mathbf{s}) := \sum_{\mathbf{j}\in\mathbb{N}_0^k} p^{(n)}(\mathbf{e}_\alpha, \mathbf{j})\mathbf{s}^{\mathbf{j}}$, $1 \le \alpha \le k$, $\mathbf{s} \in C$. Let $m_{\alpha\beta}^{(n)} = E_{\mathbf{e}_\alpha} Z_{1,\beta}^{(n)}$ be the expected number of type β offspring from a single particle of type α in one generation. Then the $k \times k$ matrix $\mathbf{M}^{(n)} = (m_{\alpha\beta}^{(n)})_{\alpha,\beta=1,\ldots,k}$ is called the *mean matrix* of $Z^{(n)}$. Note that $m_{\alpha\beta}^{(n)} = \frac{\partial F_{(\alpha)}^{(n)}}{\partial s_\beta}(\mathbf{1})$, where the partial derivative is understood to be the left hand derivative. The processes $Z^{(n)}$ will be assumed to have a *uniformly strictly positive* mean matrix $\mathbf{M}^{(n)}$, by which we mean that there exist $U \in \mathbb{N}$ and $a \in (0, \infty)$ such that for every $n \ge 1$ $((\mathbf{M}^{(n)})^U)_{\alpha,\beta} \ge a$ for all $1 \le \alpha, \beta \le k$. From the Perron-Frobenius Theorem it then follows that $\mathbf{M}^{(n)}$ has a real, positive maximal eigenvalue ρ_n with associated positive left and right eigenvectors $\mathbf{v}^{(n)}$ and $\mathbf{u}^{(n)}$, respectively, which, without loss of generality, are normalized so that $\mathbf{u}^{(n)'}\mathbf{v}^{(n)} = 1$ and $\mathbf{u}^{(n)'}\mathbf{1} = 1$ (see [1]). The maximal eigenvalue ρ_n plays a similar role in the classification of the k-BGW process as the mean played in classifying the (single type) BGW process. The k-BGW process is called subcritical, critical, or supercritical, according to whether $\rho_n < 1$, $\rho_n = 1$, or $\rho_n > 1$, respectively. We will consider the subcritical case, namely for all $n \geq 1$ $\rho_n \in (0,1)$, and study the behavior of quasi-stationary and Yaglom distributions of the scaled process $X^{(n)}(t) = \frac{Z^{(n)}_{\lfloor nt \rfloor}}{n}$, $t \geq 0$, as $\rho_n \to 1$.

Condition 2.1.11. *For each $n \geq 1$, $E_{\mathbf{1}}(||Z_1^{(n)}|| \log ||Z_1^{(n)}||) < \infty$.*

The existence of the Yaglom distribution of $X^{(n)}$ is assured by the following result, which is a consequence of Theorem V.4.1 of [1] (see also Theorem V.4.2 therein). We give a proof in Section 2.3.

Theorem 2.1.12. *Assume Condition 2.1.11. For each $n \in \mathbb{N}$, $X^{(n)}$ has a Yaglom distribution $\nu^{(n)}$. This distribution is also a qsd.*

Condition 2.1.13. *There exist $b, d \in (0,\infty)$ such that for all $n \in \mathbb{N}$ (i) $\sum_{\alpha\beta\gamma} \partial^2 F^{(n)}_{(\alpha)}(\mathbf{1})/\partial s_\beta \partial s_\gamma \geq b$, and (ii) $\sum_{\alpha,\beta,\gamma,\delta} \partial^3 F^{(n)}_{(\alpha)}(\mathbf{1})/\partial s_\beta \partial s_\gamma \partial s_\delta \leq d$, where $\alpha, \beta, \gamma, \delta$ in the above sums vary over $\{1, \ldots, k\}$.*

Part (i) of the assumption can be interpreted as a non-degeneracy condition, and part (ii) says that the third moments of the offspring distributions are uniformly bounded in n.

The assumption on convergence of means translates into the following requirement in the multitype setting.

Condition 2.1.14. *For some strictly positive matrix $\mathbf{M}$ and each $n \in \mathbb{N}$, $\mathbf{M}^{(n)} = \mathbf{M} + \frac{\mathbf{C}^{(n)}}{n}$, and $\lim_{n\to\infty} \mathbf{C}^{(n)} = \mathbf{C}$. The maximal eigenvalues ρ_n of $\mathbf{M}^{(n)}$ are of the form $\rho_n = 1 + \frac{c_n}{n}$, with $c_n \in (-n, 0)$ and $\lim_{n\to\infty} c_n = c \in (-\infty, 0)$. Moreover, $\mathbf{M}$ has maximal eigenvalue 1 with corresponding eigenvectors $\mathbf{v} = \lim \mathbf{v}^{(n)}$ and $\mathbf{u} = \lim \mathbf{u}^{(n)}$. Finally, $\mathbf{v}'\mathbf{C}\mathbf{u} = c$.*

Example 2.1.1. *Let $\mathbf{C}^{(n)} = c_n\mathbf{I}$, where $\mathbf{I}$ is the identity matrix and $c_n \in (-n, 0)$ such that $c_n \to c \in (0,\infty)$. Let $\mathbf{M}$ be a strictly positive matrix with maximal eigenvalue equal to 1. Then $\mathbf{M}^{(n)} = \mathbf{M} - \frac{\mathbf{C}^{(n)}}{n}$ satisfies Condition 2.1.14.*

Let

$$\sigma_{i,j}^{(n)}(l) = \sum_{\boldsymbol{r} \in \mathbb{N}_0^k} (r_i - m_{li}^{(n)})(r_j - m_{lj}^{(n)}) p^{(n)}(\mathbf{e}_l, \boldsymbol{r}).$$

The following condition is analogous to the assumption on convergence of variances in the single type case.

Condition 2.1.15. *As $n \to \infty$, $\sigma_{i,j}^{(n)}(l) \to \sigma_{i,j}(l)$ for all $1 \le i, j, l \le k$ and $Q := \frac{1}{2}\sum_{l=1}^{k} v_l \mathbf{u}' \boldsymbol{\sigma}(l) \mathbf{u} > 0$, where $\boldsymbol{\sigma}(l)$ is the matrix with $(i,j)^{th}$ entry $\sigma_{i,j}(l)$.*

The following diffusion approximation result can be established along the lines of Theorem 4.3.1 of [20] and Theorem 9.2.1 of [10]. We provide a sketch in Section 2.3.

Theorem 2.1.16. *Assume Conditions 2.1.13, 2.1.14, and 2.1.15. Suppose that the distribution of $X^{(n)}(0)$ converges to some $\mu \in \mathcal{P}(\mathbb{R}_+^k)$. Let $\mu_1 \in \mathcal{P}(\mathbb{R}_+)$ be given as*

$$\mu_1(A) = \mu\{\mathbf{x} \in \mathbb{R}_+^k | \mathbf{x}'\mathbf{u} \in A\}, \quad A \in \mathcal{B}(\mathbb{R}_+). \tag{2.1.7}$$

Let $\zeta^{(n)} = X^{(n)\prime} \mathbf{u}^{(n)}$. Then $\zeta^{(n)}$ converges weakly in $D(\mathbb{R}_+ : \mathbb{R}_+)$ to the unique (in law) diffusion ζ with initial distribution μ_1 and generator $\tilde{L}$ given as

$$(\tilde{L}f)(x) = cxf'(x) + Qxf''(x), \quad f \in C_c^\infty(\mathbb{R}_+), \quad x \in \mathbb{R}_+. \tag{2.1.8}$$

Furthermore, for any $t_0 \in (0, \infty)$, the process $X^{(n,0)}$, defined by $X^{(n,0)}(t) = X^{(n)}(t_0 + t)$, $t \ge 0$, converges weakly to $X^{(0)} = \mathbf{v}\zeta^{(0)}$, where $\zeta^{(0)}(t) = \zeta(t_0 + t)$, $t \ge 0$.

The process $X^{(0)}$ is a Markov process with state space $S_{\mathbf{v}} = \{\theta \mathbf{v} | \theta \ge 0\}$ and can be formally regarded as the limit of $X^{(n)}$. Indeed, if the support of μ is contained in $S_{\mathbf{v}}$, then, noting that $\mathbf{u}'\mathbf{v} = 1$, we see that the law of $\mathbf{v}\zeta(0)$ equals μ, and that in fact $X^{(n)}$ converges weakly to $\mathbf{v}\zeta$, where ζ is as in Theorem 2.1.16. We will be concerned with the Yaglom distribution of the $S_{\mathbf{v}}$ valued Markov process $X^{(0)}$ and its relation to the Yaglom distribution of $X^{(n)}$. For that it will be convenient to regard a probability measure on $S_{\mathbf{v}}$ as one on $\mathbb{R}_+^k$. Denote by $\tilde{\nu}$ the Exponential distribution with density $f(x) = |c|Q^{-1}\exp(-|c|Q^{-1}x)$, $x \geq 0$. Theorem 2.1.3 says that the Yaglom distribution of $\zeta^{(0)}$ is given by $\tilde{\nu}$. Since $X^{(0)} = \mathbf{v}\zeta^{(0)}$, the Yaglom distribution of $X^{(0)}$ exists as well and equals the distribution of $\mathbf{v}Y$, where Y has distribution $\tilde{\nu}$. Thus, we have the following:

Theorem 2.1.17. *Assume Conditions 2.1.13, 2.1.14, and 2.1.15. The Yaglom distribution of $\zeta^{(0)}$ exists and equals $\tilde{\nu}$. Furthermore, the Yaglom distribution of $X^{(0)}$, denoted by $\bar{\nu}$, exists and equals the distribution of $\mathbf{v}Y$, where Y has distribution $\tilde{\nu}$.*

The following is our main result that relates the qsd's and Yaglom distributions of $X^{(n)}$ to that of its "diffusion limit" $X^{(0)}$. Probability distributions similar to $\bar{\nu}$ have previously been noted in the study of qsd's of multitype BGW processes. In [1] (p. 191), a single critical BGW process Z (rather than a sequence of near critical BGW processes) is considered and it is shown that Z_n/n conditioned on non-extinction converges to a random variable that is concentrated on the ray $\{x\mathbf{v}_Z | x \geq 0\}$, where $\mathbf{v}_Z$ is the left eigenvector of the mean matrix of Z corresponding to the eigenvalue 1. In [35] (see Theorem 3 therein) the case where Z is near critical and a somewhat differently (component wise) scaled process Z^* is considered. The asymptotic behavior of Z_n^* conditioned on non-extinction, as $n \to \infty$, and the offspring distribution approaches criticality, is related to the limiting distributions considered here. We remark that none of these results concern the setting of diffusion approximation, where time and space are scaled and one starts with a large number of particles.

Theorem 2.1.18. *Assume Conditions 2.1.13, 2.1.14, and 2.1.15. The Yaglom distribution $\nu^{(n)}$ of $X^{(n)}$ converges weakly to the Yaglom distribution $\bar{\nu}$ of $X^{(0)}$.*

2.2 Proofs: Single Type Case.

In this section we give proofs of Theorems 2.1.4, 2.1.7, and 2.1.10. We begin with Theorem 2.1.4.

Proof of Theorem 2.1.4. In the subcritical case, it is immediate from Theorem V.4.2 of [1] that $X^{(n)}$ has a Yaglom distribution $\nu^{(n)}$. A representation of the Laplace transform of $\nu^{(n)}$ in the subcritical or supercritical case is given in Lemma 2.2.1, below. Moreover, by Lemma 6.0.1 and (2.2.7), $\nu^{(n)}$ is a qsd.

We now show that $\nu^{(n)}$ converges weakly to ν. The first step is to establish the representation for the Laplace transform of $\nu^{(n)}$ given in Lemma 2.2.1 below. In the subcritical case, define

$$Q_k^{(n)}(s) := m_n^{-k}(F_k^{(n)}(s) - 1), \quad s \in [0,1]. \tag{2.2.1}$$

Then $Q_k^{(n)}$ converges pointwise over $[0,1]$, as $k \to \infty$, to a continuous function $Q^{(n)}$ that is positive on $[0,1)$ (see [1], p. 40, Corollary I.11.1), i.e.

$$\lim_{k\to\infty} Q_k^{(n)}(s) =: Q^{(n)}(s), \quad s \in [0,1], \tag{2.2.2}$$

where $Q^{(n)}(s) > 0$ for $s \in [0,1)$. The function $Q^{(n)}$ will determine the Laplace transform of $\nu^{(n)}$ in the subcritical case. In the supercritical case, we proceed as follows. Note that, since $p_0^{(n)} > 0$, we have that $q_n > 0$. Also since $m_n > 1$, we have $q_n \in (0,1)$ and that q_n is the smallest root of $F^{(n)}(t) = t$ (see [1], Theorem I.5.1). Define $\tilde{F}^{(n)}(s) := q_n^{-1} F^{(n)}(q_n s)$, $s \in [0,1]$. Since $F^{(n)}(q_n) = q_n$, each $\tilde{F}^{(n)}$ is again a pgf and thus has a representation $\tilde{F}^{(n)}(s) = \sum_{l=0}^{\infty} \tilde{p}_l^{(n)} s^l$, $s \in [0,1]$, with $\sum_{l=0}^{\infty} \tilde{p}_l^{(n)} = 1$. In fact, $\tilde{p}_l^{(n)} = p_l^{(n)} q_n^{l-1}$, $l \in \mathbb{N}_0$. The probability distribution $\{\tilde{p}_l^{(n)}\}$ has mean $\tilde{m}_n = q_n^{-1} F^{(n)\prime}(q_n) q_n = F^{(n)\prime}(q_n) < 1$ and variance $\tilde{\sigma}_n^2 = \tilde{F}^{(n)\prime\prime}(1) - \tilde{m}_n^2 + \tilde{m}_n = q_n F^{(n)\prime\prime}(q_n) - \tilde{m}_n^2 + \tilde{m}_n$. That $F^{(n)\prime}(q_n) < 1$ is a consequence of $F^{(n)\prime}(1) > 1$, $F^{(n)}(q_n) = q_n$, and the strict convexity of $F^{(n)}$ on $[0,1]$. The latter follows from the assumption that $p_0^{(n)} + p_1^{(n)} < q_n$. Let $\tilde{Q}_k^{(n)}(s) := \tilde{m}_n^{-k}(\tilde{F}_k^{(n)}(s) - 1)$, $s \in [0,1]$. Then

$$\lim_{k\to\infty} \tilde{Q}_k^{(n)}(s) =: \tilde{Q}^{(n)}(s), \quad s \in [0,1], \tag{2.2.3}$$

and $\tilde{Q}^{(n)}$ has the same properties as those of $Q^{(n)}$ in the subcritical case noted earlier.

Lemma 2.2.1. *The Laplace transform of $\nu^{(n)}$, $\int_{[0,\infty)} e^{-\alpha x}\nu^{(n)}(dx)$, in the subcritical case, is given as $[Q^{(n)}(0) - Q^{(n)}(e^{-\alpha/n})]/(Q^{(n)}(0))$ and, in the supercritical case as $[\tilde{Q}^{(n)}(0) - \tilde{Q}^{(n)}(e^{-\alpha/n})]/(\tilde{Q}^{(n)}(0))$.*

Proof. Consider first the subcritical case. Since $T_{X^{(n)}} < \infty$ *a.s.*, it suffices to show that, for each $\alpha \geq 0$,

$$\lim_{t\to\infty} E_{\frac{i}{n}}\left(e^{-\alpha X_t^{(n)}} | X_t^{(n)} > 0\right) = \frac{Q^{(n)}(0) - Q^{(n)}(e^{-\alpha/n})}{Q^{(n)}(0)}. \tag{2.2.4}$$

Elementary calculations give

$$E_{\frac{i}{n}}\left(e^{-\alpha X_t^{(n)}} | X_t^{(n)} > 0\right) = 1 - \frac{A_{n,t}(e^{-\alpha/n})}{A_{n,t}(0)}, \tag{2.2.5}$$

where, for $\theta \in [0,1]$, $A_{n,t}(\theta) = m_n^{-\lfloor nt \rfloor}\left(1 - \left[F_{\lfloor nt \rfloor}^{(n)}(\theta)\right]^i\right)$. Next,

$$\begin{aligned}
\lim_{t\to\infty} A_{n,t}(e^{-\alpha/n}) &= \lim_{t\to\infty} m_n^{-\lfloor nt \rfloor}\left(1 - \sum_{k=0}^{i}\binom{i}{k}\left(F_{\lfloor nt \rfloor}^{(n)}(e^{-\alpha/n}) - 1\right)^k\right) \\
&= -i \lim_{t\to\infty} m_n^{-\lfloor nt \rfloor}\left(F_{\lfloor nt \rfloor}^{(n)}(e^{-\alpha/n}) - 1\right) = -iQ^{(n)}(e^{-\alpha/n}),
\end{aligned} \tag{2.2.6}$$

where the second and third equalities follow from (2.2.2). In exactly the same way one sees that $\lim_{t\to\infty} A_{n,t}(0) = -iQ^{(n)}(0)$. Combining the above observations we have (2.2.4), which proves the lemma for the subcritical case.

Consider now the supercritical case. Similar to the subcritical case

$$E_{\frac{i}{n}}(e^{-\alpha X_t^{(n)}}|t < T_{X^{(n)}} < \infty) = \frac{\left[F_{\lfloor nt\rfloor}^{(n)}(q_n e^{-\alpha/n})\right]^i - \left[F_{\lfloor nt\rfloor}^{(n)}(0)\right]^i}{\left[F_{\lfloor nt\rfloor}^{(n)}(q_n)\right]^i - \left[F_{\lfloor nt\rfloor}^{(n)}(0)\right]^i} \equiv 1 - \frac{\tilde{A}_{n,t}(e^{-\frac{\alpha}{n}})}{\tilde{A}_{n,t}(0)}$$

where, for $\theta \in [0,1]$, $\tilde{A}_{n,t}(\theta) = \tilde{m}_n^{-\lfloor nt\rfloor}\left(1 - \left[\tilde{F}_{\lfloor nt\rfloor}^{(n)}(\theta)\right]^i\right)$. This says in particular that

$$E_{\frac{i}{n}}(e^{-\alpha X_t^{(n)}}|t < T_{X^{(n)}} < \infty) = E_{\frac{i}{n}}(e^{-\alpha \tilde{X}_t^{(n)}}|\tilde{X}_t^{(n)} > 0), \tag{2.2.7}$$

where $\tilde{X}_t^{(n)} := \frac{1}{n}\tilde{Z}_{\lfloor nt\rfloor}^{(n)}$, $t \in \mathbb{R}_+$, and $\tilde{Z}^{(n)}$ is a BGW process with pgf $\tilde{F}^{(n)}$. Now making use of (2.2.3) instead of (2.2.2), the proof for the supercritical case is completed exactly as for the subcritical case. □

We continue with the proof of Theorem 2.1.4, which is based on the fact that the Laplace transform of ν is $G(\alpha) = (1 + \frac{\alpha\sigma^2}{2|c|})^{-1}$, $\alpha \geq 0$. First, we show that $\nu^{(n)}$ converges to ν for a special subcritical model where the pgf is of the so-called linear fractional form (see [1], pp. 6-7, [16], pp. 9-10). We then establish a comparison lemma which allows us to prove the general subcritical result by an approximation argument.

Lemma 2.2.2. *Assume Condition 2.1.1 and that $c_n < 0$ for all n. Let, for each n, $F^{(n)}$ be of the linear fractional form:*

$$F^{(n)}(s) = 1 - \frac{b^{(n)}}{1 - p^{(n)}} + \frac{b^{(n)}s}{1 - p^{(n)}s}, \quad s \in [0,1], \tag{2.2.8}$$

where $b^{(n)}, p^{(n)} \in (0,1)$ and $b^{(n)} < 1 - p^{(n)}$. Then $\nu^{(n)}$ converges weakly to ν.

We note that Condition 2.1.1 imposes certain restrictions on $b^{(n)}$ and $p^{(n)}$ which are not made explicit in the statement of the lemma. See Lemma 6.0.2 for a precise relationship between the parameters $b^{(n)}$, $p^{(n)}$, and the mean and variance of $Z_1^{(n)}$.

Proof. With $A_{n,t}$ as in the proof of Lemma 2.2.1, we have

$$E_{\frac{i}{n}}\left(e^{-\alpha X_t^{(n)}} | X_t^{(n)} > 0\right) = 1 - \frac{A_{n,t}(e^{-\alpha/n})}{A_{n,t}(0)}.$$

In order to prove the lemma it suffices to show that

$$\lim_{n\to\infty}\lim_{t\to\infty}\frac{n}{i}A_{n,t}(0) = -\frac{2c}{\sigma^2} \text{ and } \lim_{n\to\infty}\lim_{t\to\infty}\frac{n}{i}A_{n,t}(e^{-\alpha/n}) = \frac{2c\alpha}{2c-\alpha\sigma^2}. \tag{2.2.9}$$

Since $m_n < 1$ for each n, we get (see [1], p. 7) for each $l \geq 1$

$$\begin{aligned} F_l^{(n)}(s) &= 1 - m_n^l\left(\frac{1-s_{n,0}}{m_n^l - s_{n,0}}\right) + \frac{m_n^l\left(\frac{1-s_{n,0}}{m_n^l-s_{n,0}}\right)^2 s}{1-\left(\frac{m_n^l-1}{m_n^l-s_{n,0}}\right)s} \\ &= 1 - m_n^l a_{n,l} + \frac{m_n^l a_{n,l}^2 s}{1-b_{n,l}s}, \end{aligned} \tag{2.2.10}$$

where $a_{n,l} = \frac{1-s_{n,0}}{m_n^l-s_{n,0}}$, $b_{n,l} = \frac{m_n^l-1}{m_n^l-s_{n,0}}$, and $s_{n,0}$ is the unique root of $F^{(n)}(s) = s$ that is strictly greater than 1. Note that both $a_{n,l}$ and $b_{n,l}$ converge as $l \to \infty$. We get, by using (2.2.10) in the definition of $A_{n,t}$, $\lim_{t\to\infty}\frac{n}{i}A_{n,t}(0) = n\frac{s_{n,0}-1}{s_{n,0}}$. From the explicit form of $F^{(n)}$ (see [1], p. 6) we have $\frac{1-p^{(n)}}{1-p^{(n)}s_{n,0}} = \frac{1}{m_n}$, and thus $s_{n,0} = \frac{1-m_n(1-p^{(n)})}{p^{(n)}}$. As a consequence of Condition 2.1.1, we have that

$$s_{n,0} \to 1,\ p^{(n)} \to p, \text{ and } \sigma^2 = \frac{2p}{1-p}. \tag{2.2.11}$$

Combining these observations we obtain

$$\lim_{n\to\infty}\lim_{t\to\infty}\frac{n}{i}A_{n,t}(0) = \lim_{n\to\infty} n\frac{s_{n,0}-1}{s_{n,0}} = \lim_{n\to\infty} n\frac{(1-p^{(n)})(1-m_n)}{p^{(n)}s_{n,0}} = -\frac{2c}{\sigma^2},$$

which proves the first equality in (2.2.9). Similarly one can show that

$$\lim_{t\to\infty} \frac{n}{i} A_{n,t}(e^{-\alpha/n}) = n\frac{s_{n,0}-1}{s_{n,0}} - n\frac{(s_{n,0}-1)^2 e^{-\alpha/n}}{(s_{n,0}-e^{-\alpha/n})s_{n,0}}.$$

Using (2.2.11) and the above display, we now have

$$\lim_{n\to\infty}\lim_{t\to\infty} \frac{n}{i} A_{n,t}(e^{-\alpha/n}) = -\frac{2c}{\sigma^2} + \frac{2c}{\sigma^2}\lim_{n\to\infty}\frac{s_{n,0}-1}{s_{n,0}-e^{-\alpha/n}} = -\frac{2c}{\sigma^2}\left(1 - \frac{1}{1-\frac{\alpha\sigma^2}{2c}}\right),$$

which proves the second identity in (2.2.9). □

We will next treat the general case and begin with the following comparison lemma, which extends a result due to Spitzer (see [1], p. 22). The latter is concerned with pgf's with mean 1. The lemma given below extends Spitzer's result to a setting where the two pgf's have the same mean m which may be strictly less than 1.

Lemma 2.2.3. *Let $f^{(1)}$ and $f^{(2)}$ be pgf's of two $\mathbb{N}_0$ valued random variables having the same mean $m \in (0,1]$ and variances $\sigma_1^2 < \sigma_2^2 \le \infty$. Then there exist integers n_i, $i = 1,2$, such that for all $n \ge 0$*

$$f^{(1)}_{n+n_1}(t) \le f^{(2)}_{n+n_2}(t), \quad \textit{for } t \in [0,1]. \tag{2.2.12}$$

Proof. The proof is adapted from [1]. Using L'Hospital's rule, we get for $f = f^{(1)}, f^{(2)}$ and $\sigma^2 = \sigma_1^2, \sigma_2^2$

$$\lim_{t\to 1}\frac{f(t) - mt - (1-m)}{(1-t)^2} = \lim_{t\to 1}\frac{f'(t)-m}{2(t-1)} = \frac{\sigma^2 + m^2 - m}{2} =: a. \tag{2.2.13}$$

Note that $a \in (0,\infty]$. Define $\epsilon(t) := \frac{f(t)-mt-(1-m)}{(1-t)^2}$. We are interested in $\epsilon'(t)$ for t close to 1, $t \in (0,1]$. Once more by L'Hospital's rule, $\lim_{t\to 1}\epsilon'(t) = \lim_{t\to 1}\frac{f'''(t)}{6} \in [0,\infty]$. Thus $\epsilon(t)$ is non-decreasing in a (left) neighborhood of 1 and it converges to a. We define for

$f^{(i)}$, $i = 1, 2$, a_i and ϵ_i analogous to a, ϵ, by replacing f by $f^{(i)}$ and σ^2 by σ_i^2. Since $\sigma_1^2 < \sigma_2^2$ and the means of $f^{(1)}$ and $f^{(2)}$ are equal, we have that $a_1 < a_2$. Thus, from (2.2.13) and the monotonicity of ϵ_i near 1, there exists a $\delta \in (0, 1]$, such that $f^{(1)}(t) \leq f^{(2)}(t)$ for all $t \in [1 - \delta, 1]$. Using the monotonicity of $f^{(i)}$ we now have, for all $n \geq 0$ and $n_1 \leq n_2$,

$$f^{(1)}_{n+n_1}(t) \leq f^{(2)}_{n+n_2}(t) \quad \text{for } t \in [1 - \delta, 1]. \tag{2.2.14}$$

To show that (2.2.12) holds, it remains to consider $t \in [0, 1 - \delta]$. We can choose n_1 and $n_2 > n_1$, such that $f^{(1)}_{n_1}(0) \in [1 - \delta, 1]$ and $f^{(1)}_{n_1}(1 - \delta) \leq f^{(2)}_{n_2}(0)$, and thus

$$1 - \delta \leq f^{(1)}_{n_1}(0) \leq f^{(1)}_{n_1}(t) \leq f^{(1)}_{n_1}(1 - \delta) \leq f^{(2)}_{n_2}(0) \leq f^{(2)}_{n_2}(t) < 1.$$

Since $1 - \delta \leq f^{(1)}_{n_1}(t) \leq f^{(2)}_{n_2}(t)$, we get, using the monotonicity of $f^{(i)}$, that for $n \geq 0$, $f^{(1)}_{n+n_1}(t) \leq f^{(2)}_{n+n_2}(t)$, for $t \in [0, 1 - \delta]$. Combining this with (2.2.14) we have (2.2.12). $\square$

Continuing the proof of Theorem 2.1.4, we now establish the convergence of the Yaglom distribution of $X^{(n)}$ to that of ξ in the general setting.

Consider first the subcritical case. From Lemma 6.0.2 in the appendix, it follows that for all $\epsilon > 0$ and $n \in \mathbb{N}$ we can find pgf's of the linear fractional form, $f^{(n,1)}$ and $f^{(n,2)}$, such that their means are m_n and variances are $\sigma^2_{n,1} = \sigma^2_n - \epsilon$ and $\sigma^2_{n,2} = \sigma^2_n + \epsilon$, respectively.

By Lemma 2.2.3, for all $n, i \in \mathbb{N}$, there exist an l_n and a $t_0 := t_0(n)$, such that for all $t \geq t_0$ and all $r \in [0, 1]$

$$[f^{(n,1)}_{\lfloor nt \rfloor - l_n}(r)]^i \leq [F^{(n)}_{\lfloor nt \rfloor}(r)]^i \leq [f^{(n,2)}_{\lfloor nt \rfloor + l_n}(r)]^i,$$

where $f_l^{(n,j)}$ denotes the l^{th} iterate of $f^{(n,j)}$. Thus, with $A_{n,t}$ as before, for all $t \geq t_0$,

$$m_n^{-\lfloor nt \rfloor}\left(1 - \left[f^{(n,2)}_{\lfloor nt \rfloor + l_n}(0)\right]^i\right) \leq A_{n,t}(0) \leq m_n^{-\lfloor nt \rfloor}\left(1 - \left[f^{(n,1)}_{\lfloor nt \rfloor - l_n}(0)\right]^i\right). \tag{2.2.15}$$

Denote by $s_0^{(n,j)}$ the root of $f^{(n,j)}(r) = r$ that is greater 1. Then, for all $n \geq 1$,

$$\frac{n(s_0^{(n,1)} - 1)}{s_0^{(n,1)}} \geq \lim_{t\to\infty} \frac{n}{i} A_{n,t}(0) \geq \frac{n(s_0^{(n,2)} - 1)}{s_0^{(n,2)}}.$$

Similar to the calculation below (2.2.11), we now have, on letting $n \to \infty$ in the above display,

$$-\frac{2c}{\sigma^2 - \epsilon} \geq \limsup_{n\to\infty} \lim_{t\to\infty} \frac{n}{i} A_{n,t}(0) \geq \liminf_{n\to\infty} \lim_{t\to\infty} \frac{n}{i} A_{n,t}(0) \geq -\frac{2c}{\sigma^2 + \epsilon}.$$

Letting $\epsilon \to 0$, we have $\lim_{n\to\infty} \lim_{t\to\infty} \frac{n}{i} A_{n,t}(0) = -\frac{2c}{\sigma^2}$. Similarly, it is seen that

$$\lim_{n\to\infty} \lim_{t\to\infty} \frac{n}{i} A_{n,t}(e^{-\alpha/n}) = -\frac{2c}{\sigma^2}\left(1 - \frac{1}{1 - \frac{\alpha\sigma^2}{2c}}\right). \tag{2.2.16}$$

Combining the above observations, we have

$$\lim_{n\to\infty} \lim_{t\to\infty} E_{\frac{i}{n}}\left(e^{-\alpha X_t^{(n)}} | X_t^{(n)} > 0\right) = \frac{2c}{2c - \alpha\sigma^2} = \left(1 + \frac{\alpha\sigma^2}{2|c|}\right)^{-1},$$

and this proves Theorem 2.1.4 for the subcritical case.

We now consider the supercritical case. From (2.2.7) it follows that the Yaglom distribution $\nu^{(n)}$ of $X^{(n)}$ is the same as the Yaglom distribution $\tilde{\nu}^{(n)}$ of $\tilde{X}^{(n)}$. Thus it suffices, in view of the result for the subcritical case, to show that $\lim_{n\to\infty} n(\tilde{m}_n - 1) = -c$ and $\lim_{n\to\infty} \tilde{\sigma}_n^2 = \sigma^2$.

We begin by showing that $q_n \to 1$ as $n \to \infty$. We argue via contradiction. Suppose $\liminf_{n\to\infty} q_n = q < 1$. Let $\epsilon \in (0, \sigma^2/2)$. By the equicontinuity assumption in Condition

2.1.1, there exist a $\delta \in (0, 1-q)$ and an n_δ such that for $n \geq n_\delta$

$$|F^{(n)''}(1-\delta) - \sigma^2| \leq |F^{(n)''}(1-\delta) - F^{(n)''}(1)| + |F^{(n)''}(1) - \sigma^2| \leq 2\epsilon < \sigma^2.$$

Since $F^{(n)''}$ is nondecreasing, we have

$$F^{(n)}(1-\delta) \geq F^{(n)}(1) - \delta F^{(n)'}(1) + \frac{\delta^2}{2} F^{(n)''}(1-\delta) \geq 1 - \delta - \delta \frac{c_n}{n} + \frac{\delta^2}{2}(\sigma^2 - 2\epsilon).$$

Choose n large enough so that $q_n < 1-\delta$ and $\frac{\delta^2}{2}(\sigma^2 - 2\epsilon) > \delta \frac{c_n}{n}$. Then $F^{(n)}(1-\delta) > 1-\delta$. Since $q_n < 1-\delta$, we arrive at a contradiction because $F^{(n)}(x) < x$ for all $x \in (q_n, 1)$. The convergence of q_n to 1 and equicontinuity of $F^{(n)''}$ now immediately yield the convergence of $\tilde{\sigma}_n^2$ to σ^2.

We next establish the convergence of $n(\tilde{m}_n - 1)$. Observe that $\tilde{m}_n - m_n = F^{(n)'}(q_n) - F^{(n)'}(1) = -\int_{q_n}^1 F^{(n)''}(u)du$ and thus

$$n(\tilde{m}_n - 1) = n(m_n - 1) - n(1-q_n)\left(\int_{q_n}^1 F^{(n)''}(u) \frac{1}{1-q_n} du\right). \tag{2.2.17}$$

Moreover,

$$\begin{aligned} 1 - q_n = 1 - F^{(n)}(q_n) &= \int_{q_n}^1 (F^{(n)'}(u) - F^{(n)'}(1))du + (1-q_n)m_n \\ &= -\int_{q_n}^1 \int_u^1 F^{(n)''}(v)dvdu + (1-q_n)m_n. \end{aligned}$$

Rearranging terms gives

$$(1-q_n)(m_n - 1) = \frac{(1-q_n)^2}{2}\left(\int_{q_n}^1 F^{(n)''}(v)\frac{v-q_n}{(1-q_n)^2/2}dv\right).$$

Thus

$$n(1-q_n) = 2n(m_n - 1)\left(\int_{q_n}^{1} F^{(n)''}(v)\frac{v-q_n}{(1-q_n)^2/2}dv\right)^{-1}. \tag{2.2.18}$$

Combining equations (2.2.17) and (2.2.18), we get

$$n(\tilde{m}_n - 1) = n(m_n - 1)\left(1 - 2\frac{\int_{q_n}^{1} F^{(n)''}(u)g_{n,1}(u)du}{\int_{q_n}^{1} F^{(n)''}(v)g_{n,2}(v)dv}\right),$$

where $g_{n,1}(u) = \frac{1}{1-q_n}$ and $g_{n,2}(v) = \frac{v-q_n}{(1-q_n)^2/2}$. To complete the proof, we will now show that the ratio of integrals in the last display converges to 1, as $n \to \infty$. In fact, we will show that each integral converges to σ^2. Observing that $\int_{q_n}^{1} g_{n,i}(u)du = 1$, $i = 1, 2$, and using the monotinicity of $F^{(n)''}$, we get, for $i = 1, 2$,

$$\begin{aligned}
\limsup_{n\to\infty} \int_{q_n}^{1} F^{(n)''}(u)g_{n,i}(u)du &\leq \limsup_{n\to\infty} \int_{q_n}^{1} F^{(n)''}(1)g_{n,i}(u)du \\
&= \limsup_{n\to\infty} (\sigma_n^2 + m_n^2 - m_n) = \sigma^2. \qquad (2.2.19)
\end{aligned}$$

Similarly,

$$\begin{aligned}
\liminf_{n\to\infty} \int_{q_n}^{1} F^{(n)''}(u)g_{n,i}(u)du &\geq \liminf_{n\to\infty} \int_{q_n}^{1} F^{(n)''}(q_n)g_{n,i}(u)du \\
&= \liminf_{n\to\infty} F^{(n)''}(q_n) = \sigma^2. \qquad (2.2.20)
\end{aligned}$$

This proves $n(\tilde{m}_n - 1) \to -c$ and as argued earlier this proves Theorem 2.1.4 for the supercritical case.

Proof of Theorem 2.1.7. Note that the existence of the stationary distribution of the Q-process $X^{(n)\uparrow}$ is immediate from Theorem I.14.2 in [1]. We will show that for all $i \in \mathbb{N}$

$$\lim_{t\to\infty} E_{\frac{i}{n}}\left(e^{-\alpha X_t^{(n)\uparrow}}\right) = Q^{(n)'}(e^{-\alpha/n})e^{-\alpha/n}, \quad \alpha \geq 0. \tag{2.2.21}$$

Since $Q^{(n)'}$ is continuous at 1 (see [1], p. 40), this will show that $h_n(\alpha)$ defined by the right side of (2.2.21) is a Laplace transform of some random variable $X_\infty^{(n)\uparrow}$ with probability law $\nu^{(n)\uparrow}$. Similar to the calculation in [1], pp. 59-60, we have for $\alpha > 0$

$$E_{\frac{i}{n}} \exp\left(-\alpha X_t^{(n)\uparrow}\right) = \frac{\partial}{\partial \alpha}\left(\frac{n}{i}\left[m_n^{-\lfloor nt \rfloor}\left(1 - \left[F_{\lfloor nt \rfloor}^{(n)}(e^{-\alpha/n})\right]^i\right)\right]\right).$$

Taking the limit, as $t \to \infty$, we get

$$\begin{aligned}\lim_{t\to\infty} E_{\frac{i}{n}} \exp\left(-\alpha X_t^{(n)\uparrow}\right) &= \lim_{t\to\infty}\left(\left[F_{\lfloor nt \rfloor}^{(n)}(e^{-\alpha/n})\right]^{i-1} Q_{\lfloor nt \rfloor}^{(n)'}(e^{-\alpha/n})e^{-\alpha/n}\right) \\ &= Q^{(n)'}(e^{-\alpha/n})e^{-\alpha/n}. \qquad (2.2.22)\end{aligned}$$

This proves (2.2.21) and thus $X_t^{(n)\uparrow}$ converges in distribution, as $t \to \infty$, to $X_\infty^{(n)\uparrow}$. It is easily checked that $\nu^{(n)\uparrow}$ is a stationary distribution.

We now show that, as $n \to \infty$, $\nu^{(n)\uparrow}$ converges weakly to $\nu^\uparrow$. For this it suffices to show that

$$\lim_{n\to\infty}\lim_{t\to\infty} E_{\frac{i}{n}} \exp\left(-\alpha X_t^{(n)\uparrow}\right) = \left(\frac{1}{1-\frac{\alpha\sigma^2}{2c}}\right)^2, \quad \alpha \in (0,\infty). \qquad (2.2.23)$$

From (2.2.22) we have

$$\lim_{n\to\infty}\lim_{t\to\infty} E_{\frac{i}{n}} \exp\left(-\alpha X_t^{(n)\uparrow}\right) = \lim_{n\to\infty}\frac{\partial}{\partial\alpha}\left(-nQ^{(n)}(e^{-\alpha/n})\right).$$

We next show that for $\alpha \in (0,\infty)$

$$\lim_{n\to\infty}\frac{\partial}{\partial\alpha}\left(-nQ^{(n)}(e^{-\alpha/n})\right) = \frac{\partial}{\partial\alpha}\lim_{n\to\infty}\left(-nQ^{(n)}(e^{-\alpha/n})\right). \qquad (2.2.24)$$

Define $g_n(\alpha) := -nQ^{(n)}(e^{-\alpha/n})$. Note that $Q^{(n)}(s) = \sum_{j=0}^\infty v_j^{(n)} s^j$, $0 \le s < 1$, for some $\{v_j^{(n)}\}_{j\in\mathbb{N}_0}$ with $v_0^{(n)} < 0$ and $v_j^{(n)} > 0$, for $j \ge 1$, and $\lim_{s\nearrow 1} Q^{(n)\prime}(s) = 1$ (see [1], pp.

40-41); in particular, $Q^{(n)}$ is convex. Next note that $|g_n'(\alpha)| \le \sup_{s\in(0,1)} \left\{|Q^{(n)\prime}(s)s|\right\} = 1$, which implies that $\{g_n\}_{n\in\mathbb{N}}$ is equicontinuous on $[0,\infty)$. From (2.2.6) and (2.2.16) we have that g_n converges pointwise to g, where $g(\alpha) = \frac{2c\alpha}{2c-\alpha\sigma^2}$, $\alpha \ge 0$. Thus, by equicontinuity and uniform boundedness on compacts of $\{g_n\}$, we have that for every interval $[a,b]$, $0 < a < b < \infty$, there exists a subsequence $\{g_{n_k}\}$ which convergences to g uniformly on $[a,b]$. Thus, by [9], (9.12.1), p. 229, g is analytic on $(0,\infty)$ and $\lim_{k\to\infty} \frac{\partial}{\partial\alpha} g_{n_k}(\alpha) = \frac{\partial}{\partial\alpha} g(\alpha)$, for $\alpha \in (0,\infty)$. This proves equation (2.2.24). Equation (2.2.23) is now immediate on combining the above two displays.

Proof of Theorem 2.1.10. Let $H_l^{(n)}(i,\cdot)$ be the l^{th} iterate of the pgf $H^{(n)}(i,\cdot)$ of $Y_1^{(n)}$ given $Y_0^{(n)} = i$. Then, for all n and $s \in [0,1]$, $H_l^{(n)}(i,s) = [F_l^{(n)}(s)]^i \prod_{r=0}^{l-1} G^{(n)}(F_r^{(n)}(s))$ and $H_l^{(n)}(i,\cdot)$ converges, as $l \to \infty$, to the pgf $\tilde{\Pi}^{(n)}$ given as $\tilde{\Pi}^{(n)}(s) = \prod_{r=0}^{\infty} G^{(n)}(F_r^{(n)}(s))$ (see [34]). This shows that, for each $n \in \mathbb{N}$, $V^{(n)}$ has a unique stationary distribution $\eta^{(n)}$, which is characterized through its pgf $\Pi^{(n)}(s) = \prod_{r=0}^{\infty} G^{(n)}(F_r^{(n)}(s^{1/n}))$. We now show that, as $n \to \infty$, $\eta^{(n)}$ converges weakly to η. Let $\alpha(n,l) = \frac{(m_n^l-1)F^{(n)\prime\prime}(1)}{2(m_n-1)m_n}$. Then $V^{(n)}(t) = W_{\lfloor nt\rfloor}^{(n)} \frac{\alpha(n,\lfloor nt\rfloor)}{n}$, where $W_l^{(n)} = \frac{Y_l^{(n)}}{\alpha(n,l)}$. Theorem 3 of [11] gives the weak convergence, as $t \to \infty$ and $n \to \infty$, of $W_{\lfloor nt\rfloor}^{(n)}$ to W, where W has a $\Gamma\left(\frac{2\iota}{\sigma^2}, 1\right)$ distribution. The result now follows on observing that

$$\lim_{n\to\infty}\lim_{t\to\infty} \frac{\alpha(m_n,\lfloor nt\rfloor)}{n} = \lim_{n\to\infty} \frac{-F^{(n)\prime\prime}(1)}{2n(m_n-1)m_n} = -\frac{\sigma^2}{2c}.$$

2.3 Proofs: Multitype Case.

In this section we prove Theorems 2.1.12, 2.1.16 and 2.1.18.

Proof of Theorem 2.1.12. Denote by $F_p^{(n)} = (F_{p,(1)}^{(n)}, \ldots, F_{p,(k)}^{(n)})$ the p^{th} iterate of $F^{(n)}$, i.e. for $\mathbf{s} \in C$ and $p \in \mathbb{N}_0$, $F_{p+1}^{(n)}(\mathbf{s}) = F^{(n)}(F_p^{(n)}(\mathbf{s}))$, where $F_0^{(n)}(\mathbf{s}) = \mathbf{s}$. Let $\gamma^{(n)}(\mathbf{s}) := \lim_{p\to\infty} \frac{\mathbf{v}^{(n)\prime}[\mathbf{1}-F_p^{(n)}(\mathbf{s})]}{\rho_n^p}$, $\mathbf{s} \in C$. The latter limit exists and defines a positive function on $C \setminus \{\mathbf{1}\}$ that is continuous at $\mathbf{1}$ (see [1, Theorems V.4.1]). We will next show that, for each $\mathbf{s} \in C$,

$$\lim_{t\to\infty} E_{\frac{\mathbf{i}}{n}}(e^{-\mathbf{s}' X_t^{(n)}} | X_t^{(n)} \neq \mathbf{0}) = \frac{\gamma^{(n)}(\mathbf{0}) - \gamma^{(n)}(\mathbf{r}_n)}{\gamma^{(n)}(\mathbf{0})}, \tag{2.3.1}$$

where $\mathbf{r}_n = (e^{-s_1/n}, \cdots, e^{-s_k/n})'$ and $\mathbf{s} = (s_1, \cdots s_k)'$. Denoting by $\nu^{(n)}$ the probability law corresponding to the Laplace transform on the right hand side of the above display, we will then have that $\nu^{(n)}$ is the Yaglom distribution of $X^{(n)}$. The fact that $\nu^{(n)}$ is also a qsd is a consequence of Lemma 6.0.1 in the appendix. We now prove (2.3.1).

Elementary calculations give

$$E_{\frac{\mathbf{i}}{n}}(e^{-\mathbf{s}' X_t^{(n)}} | X_t^{(n)} \neq \mathbf{0}) = E(e^{-\frac{\mathbf{s}'}{n} Z_{\lfloor nt \rfloor}^{(n)}} | Z_0^{(n)} = \mathbf{i}, Z_{\lfloor nt \rfloor}^{(n)} \neq \mathbf{0}) = 1 - \frac{A_{n,t}(\mathbf{r}_n)}{A_{n,t}(0)},$$

where, for $\theta \in C$, $A_{n,t}(\theta) = \rho_n^{-\lfloor nt \rfloor} \left(1 - (F_{\lfloor nt \rfloor}^{(n)}(\theta))^{\mathbf{i}}\right)$. Next note that

$$\begin{aligned} (F_{\lfloor nt \rfloor}^{(n)}(\mathbf{r}_n))^{\mathbf{i}} &= \prod_{\substack{\alpha=1 \\ i_\alpha \neq 0}}^{k} \sum_{r=1}^{i_\alpha} \binom{i_\alpha}{r} 1^{i_\alpha - r} \left(F_{\lfloor nt \rfloor, (\alpha)}^{(n)}(\mathbf{r}_n) - 1\right)^r \\ &= 1 - \mathbf{i}' \left(\mathbf{1} - F_{\lfloor nt \rfloor}^{(n)}(\mathbf{r}_n)\right) + \tilde{R}_{n,t}, \end{aligned} \tag{2.3.2}$$

where the term $\tilde{R}_{n,t}$ is a linear combination of terms of the form $\left(\mathbf{1} - F_{\lfloor nt \rfloor}^{(n)}(\mathbf{r}_n)\right)^{\mathbf{d}}$, where $\mathbf{d} = (d_1, \ldots, d_k)$ and $\sum_{j=1}^{k} d_j > 1$. Since $E_{\mathbf{1}}(||Z_1^{(n)}|| \log ||Z_1^{(n)}||) < \infty$, we have

$$\lim_{t\to\infty} \rho_n^{-\lfloor nt \rfloor} (\mathbf{1} - F_{\lfloor nt \rfloor}^{(n)}(\mathbf{r}_n)) = \gamma^{(n)}(\mathbf{r}_n) \mathbf{u}^{(n)} \tag{2.3.3}$$

(see [1, Theorems V.4.1]), and thus

$$\lim_{t\to\infty} \rho_n^{-\lfloor nt \rfloor} \tilde{R}_{n,t} = 0. \tag{2.3.4}$$

This implies $\lim_{t\to\infty} A_{n,t}(\mathbf{r}_n) = \gamma^{(n)}(\mathbf{r}_n) \mathbf{i}' \mathbf{u}^{(n)}$. In exactly the same way, we see that

$\lim_{t\to\infty} A_{n,t}(0) = \gamma^{(n)}(\mathbf{0})\mathbf{i}'\mathbf{u}^{(n)}$. Combining the above observations, we now have (2.3.1) and the result follows.

Proof of Theorem 2.1.16. The proof is similar to that of Theorem 4.3.1 of [20] and thus only a sketch is provided. Let

$$Y^{(n)}(t) := \mathbf{y}^{(n)} + n\int_0^t (\mathbf{M}' - \mathbf{I})X^{(n)}(\tau-)dA_n(\tau),$$

where $A_n(\tau) = \frac{\lfloor n\tau \rfloor}{n}$, $\tau \geq 0$. Define a Markov chain $\{(\check{X}^{(n)}(k), \check{Y}^{(n)}(k))\}_{k\in\mathbb{N}_0}$ as

$$(\check{X}^{(n)}(k), \check{Y}^{(n)}(k)) = (X^{(n)}(k/n), Y^{(n)}(k/n)), \quad k \in \mathbb{N}_0.$$

This chain has transition probabilities given by

$$\check{P}^{(n)}(\mathbf{x}, \mathbf{y}, \tilde{\mathbf{x}}, \tilde{\mathbf{y}}) = \check{Q}^{(n)}(\mathbf{x}, \tilde{\mathbf{x}})1_{\tilde{\mathbf{y}}=\mathbf{y}+(\mathbf{M}'-\mathbf{I})\mathbf{x}},$$

where $\check{Q}^{(n)}(\cdot,\cdot)$ is the transition probability of the process $\check{X}^{(n)}$. Let

$$(\check{L}^{(n)}f)(\mathbf{x}, \mathbf{y}) = \sum_{\tilde{\mathbf{x}},\tilde{\mathbf{y}}} \check{P}^{(n)}(\mathbf{x}, \mathbf{y}, \tilde{\mathbf{x}}, \tilde{\mathbf{y}})[f(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) - f(\mathbf{x}, \mathbf{y})]$$

and $L^{(n)} = n\check{L}^{(n)}$. Then we have that for each smooth test function f

$$f(X^{(n)}(t), Y^{(n)}(t)) - \int_0^t (L^{(n)}f)(X^{(n)}(\tau-), Y^{(n)}(\tau-))dA^{(n)}(\tau)$$

is a martingale (with respect to the filtration generated by $(X^{(n)}, Y^{(n)})$). Let $f(\mathbf{x}, \mathbf{y}) = \phi(\mathbf{x} - \mathbf{y})$ with $\phi \in C_c^\infty(\mathbb{R}^k)$. Taking $\mathbf{y}^{(n)} = \mathbf{0}$ and using a Taylor series expansion about

$(X^{(n)}(0), Y^{(n)}(0))$, we have that

$$\phi(X^{(n)}(t) - Y^{(n)}(t)) - \phi(X^{(n)}(0)) + E_n(t)$$
$$- \int_0^t \sum_{i=1}^k (\mathbf{C}^{(n)\prime} X^{(n)}(\tau-))_i \frac{\partial \phi}{\partial s_i}(X^{(n)}(\tau-) - Y^{(n)}(\tau-)) dA^{(n)}(\tau)$$
$$- \frac{1}{2} \int_0^t \sum_{i,j,l=1}^k (X^{(n)}(\tau-))_l \sigma_{i,j}^{(n)}(l) \frac{\partial^2 \phi}{\partial s_i \partial s_j}(X^{(n)}(\tau-) - Y^{(n)}(\tau-)) dA^{(n)}(\tau)$$

is a martingale, where the remainder $E_n(t)$ is such that $\sup_{0\le t\le T} |E_n(t)| \to 0$, in probability for all $T \in \mathbb{R}_+$. From Condition 2.1.14 and the Perron-Frobenius Theorem it follows (see Remark 4.3.2 in [20]) that with $P = \mathbf{u}\mathbf{v}'$

$$(\mathbf{I} - P')X^{(n,0)} \text{ converges to } \mathbf{0} \text{ in probability, uniformly on compacts, for all } t_0 > 0. \tag{2.3.5}$$

Also, using the fact that $P'(M' - I) = 0$, we have $P'Y^{(n)}(t) = 0$ for all $t \ge 0$. Using these observations, it can be shown that, for all $t \in \mathbb{R}_+$,

$$\lim_{n\to\infty} \int_0^t E|(\hat{L}^{(n)}\phi)(X^{(n)}(\tau-), \xi^{(n)}(\tau-)) - (L\phi)(\xi^{(n)}(\tau-))| dA_\tau^{(n)} = 0, \tag{2.3.6}$$

where $\xi^{(n)} = X^{(n)} - Y^{(n)}$, and for $(\mathbf{x}, \mathbf{z}) \in \mathbb{R}_+^k \times \mathbb{R}^k$,

$$(\hat{L}^{(n)}\phi)(\mathbf{x}, \mathbf{z}) = \sum_{i=1}^k ((\mathbf{C}^{(n)})'\mathbf{x})_i \frac{\partial \phi}{\partial s_i}(\mathbf{z}) + \frac{1}{2} \sum_{l=1}^k \sum_{i,j=1}^k \mathbf{x}_l \sigma_{i,j}(l) \frac{\partial^2 \phi}{\partial s_i \partial s_j}(\mathbf{z}).$$

and

$$(L\phi)(\mathbf{z}) = \sum_{i=1}^k (\mathbf{C}'\mathbf{P}'\mathbf{z})_i \frac{\partial \phi}{\partial s_i}(\mathbf{z}) + \frac{1}{2} \sum_{l=1}^k \sum_{i,j=1}^k (\mathbf{P}'\mathbf{z})_l \sigma_{i,j}(l) \frac{\partial^2 \phi}{\partial s_i \partial s_j}(\mathbf{z}).$$

Following [20], one can show that $\xi^{(n)}$ is a tight sequence in $D(\mathbb{R}_+ : \mathbb{R}^k)$, and, using

(2.3.6), it follows that if ξ is any weak limit of $\xi^{(n)}$, then

$$\phi(\xi(t)) - \phi(\xi(0)) - \int_0^t (L\phi)(\xi(s))ds$$

is an $\mathcal{F}_t^{\xi} := \sigma(\xi(s) : s \le t)$ martingale. Thus $\xi^{(n)}$ converges weakly to the diffusion ξ with generator L and initial condition μ. Next note that $\zeta^{(n)} = X^{(n)\prime}\mathbf{u}^{(n)} = \xi^{(n)\prime}\mathbf{u}^{(n)}$. The weak convergence of $\xi^{(n)}$ to ξ shows that $\zeta^{(n)}$ converges in distribution to $\xi'\mathbf{u} \equiv \zeta$. Let $g \in C_c^{\infty}(\mathbb{R}_+)$ and define $\phi \in C_c^{\infty}(\mathbb{R}_+^k)$ as $\phi(\mathbf{z}) = g(\mathbf{z}'\mathbf{u})$, $\mathbf{z} \in \mathbb{R}_+^k$. Then

$$\begin{aligned}(Lg)(\mathbf{z}'\mathbf{u}) &= \sum_{i=1}^{k} (\mathbf{z}'\mathbf{PC})_i' g'(\mathbf{z}'\mathbf{u}) u_i + \frac{1}{2}\sum_{l=1}^{k}\sum_{i,j=1}^{k} (\mathbf{z}'\mathbf{P})_l' g''(\mathbf{z}'\mathbf{u}) u_i u_j \sigma_{i,j}^{(n)}(l) \\ &= (\mathbf{z}'\mathbf{u}\mathbf{v}'\mathbf{C}\mathbf{u}) g'(\mathbf{z}'\mathbf{u}) + \frac{1}{2}\sum_{l=1}^{k} (\mathbf{z}'\mathbf{u}\mathbf{v}')_l' \mathbf{u}'\boldsymbol{\sigma}(l)\mathbf{u} g''(\mathbf{z}'\mathbf{u}).\end{aligned}$$

Since $\mathbf{v}'\mathbf{C}\mathbf{u} = c$, we see that ζ is a Markov process with generator

$$(\tilde{L}g)(x) = cxg'(x) + Qxg''(x), \quad x \in \mathbb{R}_+.$$

This proves the first part of the theorem.

Next noting that $P'X^{(n)} = P'\xi^{(n)}$ and recalling (2.3.5) we see that $X^{(n,0)}$ converges weakly to $P'\xi^{(0)}$, where $\xi^{(0)}(t) = \xi(t + t_0)$, $t \ge 0$. Finally, since $P = \mathbf{u}\mathbf{v}'$ and $\zeta = \xi' u$ we have that $P'\xi^{(0)} = \mathbf{v}\zeta^{(0)} = X^{(0)}$ and the result follows.

Proof of Theorem 2.1.18. We begin with some preliminary results. For each $n \in \mathbb{N}, \mathbf{s} \in C$, and $\alpha = 1, \dots, k$, define $q_\alpha^{(n)}[\mathbf{s}] = \frac{1}{2}\sum_{\beta\gamma} \frac{\partial^2 F_{(\alpha)}^{(n)}(\mathbf{1})}{\partial s_\beta \partial s_\gamma} s_\beta s_\gamma$, $Q_n[\mathbf{s}] = \sum_\alpha v_\alpha^{(n)} q_\alpha^{(n)}[\mathbf{s}]$, $Q_n = Q_n[\mathbf{u}^{(n)}]$. Let

$$\pi_{n,p} = \begin{cases} \sum_{r=1}^{p} \rho_n^{r-2} & \text{for } p = 1, 2, \dots \\ 0 & \text{for } p = 0 \end{cases} \tag{2.3.7}$$

and

$$h_{n,p}(\mathbf{s}) = \rho_n^p \mathbf{v}^{(n)\prime}\mathbf{s}/(1+\pi_{n,p}Q_n\mathbf{v}^{(n)\prime}\mathbf{s}), \quad \mathbf{s}\in C. \tag{2.3.8}$$

In what follows, $\{o(p,n)|p,n\in\mathbb{N}\}$ will denote a collection of functions from C to $\mathbb{R}^k$ satisfying the property that for every $\epsilon>0$, there exist $N,P\in\mathbb{N}$, such that for $n\geq N$ and $p\geq P$ we have $\sup_{\mathbf{s}\in C}||o(p,n)(\mathbf{s})||<\epsilon$.

Proposition 2.3.1. *Assume Conditions 2.1.13, 2.1.14, and 2.1.15. For each* $n,p\in\mathbb{N}$ *and* $\mathbf{s}\in C$

$$\mathbf{1}-F_p^{(n)}(\mathbf{s}) = h_{n,p}(\mathbf{1}-\mathbf{s})\{\mathbf{u}^{(n)}+o(p,n)(\mathbf{s})\}. \tag{2.3.9}$$

The proof of the proposition is immediate from Theorem 1 of [35] (see equation (2.3) therein) and is therefore omitted. The following corollary facilitates the proof of the main result.

Corollary 2.3.1. *Assume Conditions 2.1.13, 2.1.14, and 2.1.15. For any convergent sequence* $\{\mathbf{r}_n\}\subset C$,

$$\lim_{n\to\infty}\lim_{t\to\infty}\rho_n^{-\lfloor nt\rfloor}\left(\mathbf{1}-F_{\lfloor nt\rfloor}^{(n)}(\mathbf{r}_n)\right) = \lim_{n\to\infty}\lim_{t\to\infty}\rho_n^{-\lfloor nt\rfloor}h_{n,\lfloor nt\rfloor}(\mathbf{1}-\mathbf{r}_n)\mathbf{u}^{(n)}. \tag{2.3.10}$$

Proof. Let $a(n,t)=\rho_n^{-\lfloor nt\rfloor}h_{n,\lfloor nt\rfloor}(\mathbf{1}-\mathbf{r}_n)\mathbf{u}^{(n)}$ and with $o(p,n)$ as in Proposition 2.3.1, let $b(n,t)=\rho_n^{-\lfloor nt\rfloor}h_{n,\lfloor nt\rfloor}(\mathbf{1}-\mathbf{r}_n)o(\lfloor nt\rfloor,n)(\mathbf{r}_n)$. Note that for each n we have from (2.3.3) and (2.3.9)

$$\lim_{t\to\infty}\rho_n^{-\lfloor nt\rfloor}\left(\mathbf{1}-F_{\lfloor nt\rfloor}^{(n)}(\mathbf{r}_n)\right) = \lim_{t\to\infty}(a(n,t)+b(n,t)) = \gamma^{(n)}(\mathbf{r}_n)\mathbf{u}^{(n)}.$$

Moreover, $\lim_{t\to\infty}a(n,t)$ and $\lim_{t\to\infty}o(\lfloor nt\rfloor,n)(\mathbf{r}_n)$ exist. Denoting the latter limit by $o(\infty,n)(\mathbf{r}_n)$, we have

$$\lim_{t\to\infty}b(n,t) = \frac{\mathbf{v}^{(n)\prime}(\mathbf{1}-\mathbf{r}_n)o(\infty,n)(\mathbf{r}_n)}{1+\pi_{n,\infty}Q_n\mathbf{v}^{(n)\prime}(\mathbf{1}-\mathbf{r}_n)} =: d(n),$$

where $\pi_{n,\infty} := \frac{\rho_n^{-1}}{1-\rho_n}$. Since $\lim_{n\to\infty} o(\infty, n)(\mathbf{r}_n) = 0$, we get that $\lim_{n\to\infty} d(n) = 0$, and thus $\limsup_{n\to\infty} || \lim_{t\to\infty} b(n,t)|| = \limsup_{n\to\infty} ||d(n)|| = 0$. The result follows. $\square$

We now prove Theorem 2.1.18. In view of Theorem 2.1.12, it suffices to show that, for $\mathbf{s} \in C$,

$$\lim_{n\to\infty} \lim_{t\to\infty} E_{\frac{\mathbf{i}}{n}}(e^{-\mathbf{s}' X_t^{(n)}} | X_t^{(n)} \neq \mathbf{0}) = \frac{1}{1 - \frac{Q}{c}\mathbf{v}'\mathbf{s}}. \tag{2.3.11}$$

With $A_{n,t}$ and $\mathbf{r}_n$ as in the proof of Theorem 2.1.12, we have

$$E_{\frac{\mathbf{i}}{n}}(e^{-\mathbf{s}' X_t^{(n)}} | X_t^{(n)} \neq \mathbf{0}) = 1 - \frac{A_{n,t}(\mathbf{r}_n)}{A_{n,t}(0)}.$$

Using Proposition 2.3.1, Corollary 2.3.1, and equations (2.3.2) and (2.3.4), we get

$$\lim_{n\to\infty} \lim_{t\to\infty} nA_{n,t}(\mathbf{r}_n) = \lim_{n\to\infty} n\mathbf{i}'\mathbf{u}^{(n)} \left(\frac{\mathbf{v}^{(n)\prime}(\mathbf{1} - \mathbf{r}_n)}{1 + \pi_{n,\infty} Q_n \mathbf{v}^{(n)\prime}(\mathbf{1} - \mathbf{r}_n)} \right) = \frac{\mathbf{i}'\mathbf{u}}{(\mathbf{v}'\mathbf{s})^{-1} - \frac{Q}{c}},$$

where the last equality follows on noting that $n\mathbf{v}^{(n)\prime}(\mathbf{1} - \mathbf{r}_n) \to \mathbf{v}'\mathbf{s}$, $\frac{\pi_{n,\infty}}{n} \to -\frac{1}{c}$, and using Lemma 6.0.4. Setting $\mathbf{r}_n = 0$ in the above display, we have

$$\lim_{n\to\infty} \lim_{t\to\infty} nA_{n,t}(0) = -\frac{c\mathbf{i}'\mathbf{u}}{Q}.$$

Combining the above observations we have (2.3.11) and the result follows.

Acknowledgment. The proof of Theorem 2.1.4 for the supercritical case is based on an idea communicated to us by Professor K.B. Athreya. We gratefully acknowledge his help in the proof.

Chapter 3

Catalyst-Reactant Branching Processes

3.1 Introduction and Main Results

The particles in the (multitype) Bienaymé-Galton-Watson processes considered in the last chapter evolved independently of each other. In this chapter, we consider catalytic branching processes that model the dynamics of a catalyst population which affects the activity level of an associated reactant population. Roughly speaking, we consider a population of catalyst particles, which evolve according to a continuous time branching process, and a reactant population whose branching rate is proportional to the total mass of the catalyst population. Such interacting branching process models have a long history (see [14] and references therein). Typical settings describe the catalyst evolution through a classical continuous time branching process, which in particular implies that population dynamics are modeled until the time the catalyst becomes extinct. In this chapter, we consider a setting where the catalyst population is maintained above a positive threshold through a specific form of controlled immigration. Branching process models with immigration have also been well studied in literature ([1]). However, typical mechanisms that have been considered correspond to adding a random i.i.d. immigration component to each generation. Here, instead, we consider a model where immigration is allowed only when the population drops below a certain threshold. Roughly speaking, we consider a sequence $\{X^{(n)}\}_{n\in\mathbb{N}}$ of continuous time branching processes, where $X^{(n)}$ starts with n particles. When the population drops below n, it is instantaneously restored to the level n.

There are many settings where controlled immigration models of the above form arise naturally. For example, immunotherapy in which the natural immune response is stimulated has been successful in treating cancer. Instillation of bacteria (bacillus Calmette-Guérin), for instance, has been shown to reduce recurrence of bladder carcinoma (see [19] and references therein). Many aspects of such treatments remain poorly understood, and it is of interest to develop minimally invasive plans of treatment that intervene only when the level of certain substances drop below a some threshold. Another class of examples arise from predator-prey models in ecology, where one may be concerned with the restoration of populations that are close to extinction by reintroducing species when they fall below a certain threshold. In our work, the motivation for the study of such controlled immigration models comes from problems in chemical kinetics where one wants to keep the level of a catalyst above a certain threshold in order to maintain a desirable level of reaction activity.

We now describe the precise mathematical setting. Consider a sequence of pairs of continuous time, discrete space Markov branching processes $(X^{(n)}, Y^{(n)})$, where $X^{(n)}$ and $Y^{(n)}$ represent the number of catalyst and reactant particles, respectively. The mass of each particle in the n^{th} process will be scaled down by a factor of $1/n$, and $(X^{(n)}/n, Y^{(n)}/n)$ is referred to as the joint *total mass* process. Throughout, we will use the subscript 1 and 2 for parameters describing, the catalyst and reactant population, respectively. Each of the $X_t^{(n)}$ particles alive at time t has an exponentially distributed life time with parameter $\lambda_1^{(n)}$ (mean life time $1/\lambda_1^{(n)}$). When it dies, it gives rise to a number of offspring, according to the offspring distribution $\mu_1^{(n)}(\cdot)$. In particular, the catalyst population evolves independently from the reactant population. Additionally, if the catalyst population drops below n, it is instantaneously replenished back to the level n (*controlled immigration*). The branching rate of the reactant process $Y^{(n)}$ is proportional to the total mass of the catalyst process, and we denote the offspring distribution of $Y^{(n)}$ by $\mu_2^{(n)}(\cdot)$. To facilitate some weak convergence arguments, we will consider an auxiliary sequence of processes $Z^{(n)}$ that "shadow" $X^{(n)}$ in a suitable manner. The process $Z^{(n)}$ will be a $\mathbb{Z}$ valued pure jump process whose jump instances and sizes are the same as that of $X^{(n)}$ away from the boundary $\{n\}$, whereas when $X^{(n)}$ is at the boundary, $Z^{(n)}$ has a negative jump of size 1 whenever there is immigration of a catalyst particle into the system. This description is made precise through the definition of the generator given in (3.1.1).

We are interested in establishing a suitable scaling limit theorem for $(X^{(n)}, Y^{(n)})$ as $n \to \infty$. For that purpose, we not only scale mass, but also time, and consider the processes

$$\hat{W}_t^{(n)} := \left(\hat{X}_t^{(n)}, \hat{Y}_t^{(n)}, \hat{Z}_t^{(n)}\right) := \left(\frac{X_{nt}^{(n)}}{n}, \frac{Y_{nt}^{(n)}}{n}, \frac{Z_{nt}^{(n)}}{n}\right), \quad t \in \mathbb{R}_+.$$

For sake of simplicity of the presentation, throughout this chapter, the processes are assumed to start with total mass 1, i.e. $(\hat{X}_0^{(n)}, \hat{Y}_0^{(n)}, \hat{Z}_0^{(n)}) = (1, 1, 1)$. However, analogous results can be established for more general initial configurations.

Let $\mathbb{S}_X^{(n)} := \{\frac{l}{n} | l \in \mathbb{N}_0\} \cap [1, \infty)$, $\mathbb{S}_Y^{(n)} := \{\frac{l}{n} | l \in \mathbb{N}_0\}$, $\mathbb{S}_Z^{(n)} := \{\frac{l}{n} | l \in \mathbb{Z}\}$, and $\mathbb{W}^{(n)} := \mathbb{S}_X^{(n)} \times \mathbb{S}_Y^{(n)} \times \mathbb{S}_Z^{(n)}$. Then $\{\hat{W}_t^{(n)}\}_{t \in \mathbb{R}_+}$ is a $\mathbb{W}^{(n)}$ valued Markov process with sample paths in $D(\mathbb{R}_+ : \mathbb{W}^{(n)})$. When considering an initial condition $\boldsymbol{w} := (x, y, z)$ for $\hat{W}^{(n)}$ or describing its generator, $\boldsymbol{w}$ will always be in $\mathbb{W}^{(n)}$, although this will frequently be suppressed in the notation, and we will write $\boldsymbol{w} \in \mathbb{W} := \mathbb{R}_{\geq 1} \times \mathbb{R}_+ \times \mathbb{R}$ instead.

The process $\hat{W}^{(n)}$ has infinitesimal generator $\hat{\mathcal{A}}^{(n)}$ given as

$$\hat{\mathcal{A}}^{(n)}\phi(\boldsymbol{w}) = \lambda_1^{(n)} n^2 x \sum_{k=0}^{\infty} \Big[\phi\Big(\Big(x + \frac{k-1}{n} - 1\Big)^+ + 1, y, z + \frac{k-1}{n}\Big) - \phi(\boldsymbol{w})\Big]\mu_1^{(n)}(k)$$
$$+ \lambda_2^{(n)} n^2 xy \sum_{k=0}^{\infty} \Big[\phi\Big(x, y + \frac{k-1}{n}, z\Big) - \phi(\boldsymbol{w})\Big]\mu_2^{(n)}(k), \tag{3.1.1}$$

where $\boldsymbol{w} = (x, y, z) \in \mathbb{W}^{(n)}$ and $\phi \in C_c^{\infty}(\mathbb{W})$. For $i = 1, 2$, let $m_i^{(n)} := \sum_{k=0}^{\infty} k\mu_i^{(n)}(k)$ and $\alpha_i^{(n)} = \sum_{k=0}^{\infty}(k-1)^2\mu_i^{(n)}(k)$. Note that $\alpha_i^{(n)}$ is not that variance of $\mu_i^{(n)}$, which we denote by $\kappa_i^{(n)}$, however, under Condition 3.1.1, introduced below, $\lim_{n\to\infty} \alpha_i^{(n)}/\kappa_i^{(n)} = 1$.

In this chapter, we will establish a diffusion limit theorem for $(\hat{X}^{(n)}, \hat{Y}^{(n)})$ as $n \to \infty$. The limit process (X, Y) will be such that X is a reflected diffusion with reflection at 1, and Y is a diffusion with coefficients depending on X. The driving Brownian motions in the two diffusions will be independent.

We make the following basic assumptions on the parameters of the branching rates and offspring distributions and initial configurations of the catalyst and reactant populations.

Condition 3.1.1. *(i) For each $n \in \mathbb{N}$ and $i = 1, 2$, $m_i^{(n)} = 1 + \frac{c_i^{(n)}}{n}$, $c_i^{(n)} \in (-n, 0)$, and $\alpha_i^{(n)} \in (0, \infty)$. (ii) For $i = 1, 2$, $\lim_{n\to\infty} c_i^{(n)} =: c_i \in (-\infty, 0)$, $\lim_{n\to\infty} \alpha_i^{(n)} =: \alpha_i \in (0, \infty)$, and $\lim_{n\to\infty} \lambda_i^{(n)} =: \lambda_i \in (0, \infty)$. (iii) For every $\epsilon \in (0, \infty)$, and $i = 1, 2$, $\lim_{n\to\infty} \sum_{l:l>\epsilon\sqrt{n}}(l - m_i^{(n)})^2\mu_i^{(n)}(l) = 0$, and for all $n \in \mathbb{N}$, $(\hat{X}_0^{(n)}, \hat{Y}_0^{(n)}, \hat{Z}_0^{(n)}) = (1, 1, 1)$.*

Condition 3.1.1 will hold throughout this chapter, without further mentioning in the statements of the results.

In order to present our main weak convergence result for $(\hat{X}^{(n)}, \hat{Y}^{(n)})$, we begin by recalling the definition and some properties of the one dimensional Skorohod map. Let $\Gamma : D(\mathbb{R}_+ : \mathbb{R}) \to D(\mathbb{R}_+ : \mathbb{R}_{\geq 1})$ be defined as

$$\Gamma(\psi)(t) = (\psi(t) + 1) - \inf_{0\leq s\leq t}\{\psi(s) \wedge 1\}, \quad \text{for } \psi \in D(\mathbb{R}_+ : \mathbb{R}).$$

We refer to Γ as the Skorohod map, which can be characterized as follows: If $\psi, \phi, \eta^* \in D(\mathbb{R}_+ : \mathbb{R})$ are such that (i) $\psi(0) \geq 1$, (ii) $\phi = \psi + \eta^*$, (iii) $\phi \geq 1$, (iv) η^* is non-decreasing, $\int_0^\infty 1_{\{\phi(s) \neq 1\}} d\eta^*(s) = 0$, and $\eta^*(0) = 0$, then $\phi = \Gamma(\psi)$ and $\eta^* = \phi - \psi$. The process η^* can be regarded as the reflection term that is applied to the original trajectory ψ to produce a trajectory ϕ that is constrained to $[1, \infty)$. We will make use of the following Lipschitz continuity of the Skorohod map: For $\psi, \tilde{\psi} \in D(\mathbb{R}_+ : \mathbb{R})$, with $\psi(0), \tilde{\psi}(0) \in [1, \infty)$,

$$\sup_{s \leq t} |\Gamma(\psi)(s) - \Gamma(\tilde{\psi})(s)| \leq 2 \sup_{s \leq t} |\psi(s) - \tilde{\psi}(s)|. \tag{3.1.2}$$

Let

$$\hat{\eta}_t^{(n)} := \lambda_1^{(n)} n \mu_1^{(n)}(0) \int_0^t 1_{\{\hat{X}_s^{(n)} = 1\}} ds. \tag{3.1.3}$$

This process will play the role of the reflection term in the dynamics of the catalyst process due to the controlled immigration. The diffusion limit of $(\hat{X}^{(n)}, \hat{Y}^{(n)})$ will be the process (X, Y) which is characterized in the following proposition through a system of SDEs.

Proposition 3.1.1. *Let $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{P}, \{\bar{\mathcal{F}}_t\})$ be a filtered probability space on which are given independent standard $\{\bar{\mathcal{F}}_t\}$ Brownian motions B^X and B^Y. Then the following system of SDEs has a unique strong solution:*

$$X_t = \Gamma\left(X_0 + \int_0^\cdot c_1 \lambda_1 X_s ds + \int_0^\cdot \sqrt{\alpha_1 \lambda_1 X_s} dB_s^X\right)(t), \tag{3.1.4}$$

$$Y_t = Y_0 + \int_0^t c_2 \lambda_2 X_s Y_s ds + \int_0^t \sqrt{\alpha_2 \lambda_2 X_s Y_s} dB_s^Y, \tag{3.1.5}$$

$$\eta_t = X_t - X_0 - \int_0^t c_1 \lambda_1 X_s ds - \int_0^t \sqrt{\alpha_1 \lambda_1 X_s} dB_s^X, \tag{3.1.6}$$

where $X_0 = Y_0 = 1$ and Γ is the Skorohod map described above.

The following is the main result of this chapter.

Theorem 3.1.2. *The process* $(\hat{X}^{(n)}, \hat{Y}^{(n)})$ *converges weakly in* $D(\mathbb{R}_+ : \mathbb{R}_{\geq 1} \times \mathbb{R}_+)$ *to the process* (X, Y) *given in Proposition 3.1.1.*

In Section 3.2 we will prove Proposition 3.1.1, while Theorem 3.1.2 will be proved in Section 3.3.

3.2 Proof of Proposition 3.1.1

We first observe that the unique solvability of (3.1.4) is an immediate consequence of the Lipschitz property of the Skorohod map, Lipschitz coefficients (note that $f(x) = \sqrt{x}$ is a Lipschitz function on $[1, \infty)$), and a standard Picard iteration scheme.

We next show the unique solvability of (3.1.5). For $n \in \mathbb{N}$, let $\sigma^{(n)} := \inf\{t > 0 | X_t \geq n\}$, $\check{X}_t^{(n)} := X_{t \wedge \sigma^{(n)}}$, and $f^{(n)}(y) := y \vee \frac{1}{n}$. Consider the equation

$$\check{Y}_t^{(n)} = Y_0 + c_2 \lambda_2 \int_0^t \check{X}_s^{(n)} f^{(n)}(\check{Y}_s^{(n)}) ds + \sqrt{\alpha_2 \lambda_2} \int_0^t \sqrt{\check{X}_s^{(n)} f^{(n)}(\check{Y}_s^{(n)})} dB_s^Y. \tag{3.2.1}$$

From the Lipschitz property of $f^{(n)}$ and $\sqrt{f^{(n)}}$ it follows that, for each n, the above equation has a unique pathwise solution. Let $\tau^{(n)} := \inf\{t > 0 | \check{Y}_t^{(n)} = \frac{1}{n}\}$ and $\theta^{(n)} := \tau^{(n)} \wedge \sigma^{(n)}$. Note that $\check{Y}^{(n)}$ solves (3.1.5) on $[0, \theta^{(n)}]$. Also, by unique solvability of (3.2.1), we have for all $n \in \mathbb{N}$, $\check{Y}^{(n+1)}(\cdot \wedge \theta^{(n)}) = \check{Y}^{(n)}(\cdot \wedge \theta^{(n)})$. Finally, letting $\theta^{(\infty)} := \lim_{n \to \infty} \theta^{(n)}$, the unique solution of (3.1.5) is given by the following:

$$Y_t(\omega) = \begin{cases} \check{Y}_t^{(n)}(\omega), & \text{if } 0 \leq t \leq \theta^{(n)}(\omega) \text{ for some } n \in \mathbb{N} \\ 0, & \text{if } t \geq \theta^{\infty}(\omega). \end{cases}$$

3.3 Proof of Theorem 3.1.2

We will first show tightness of the family $\{(\hat{X}^{(n)}, \hat{Y}^{(n)}, \hat{\eta}^{(n)})\}_{n\in\mathbb{N}}$.

Lemma 3.3.1. *The family $\{(\hat{X}^{(n)}, \hat{Y}^{(n)}, \hat{\eta}^{(n)})\}_{n\in\mathbb{N}}$ is tight in $D(\mathbb{R}_+ : \mathbb{R}_{\geq 1} \times \mathbb{R}_+ \times \mathbb{R}_+)$.*

Before we begin the proof, we recall here the definitions of quadratic covariation and predictable (or conditional) quadratic covariation for semimartingales (see e.g. [32]). Let X, Y be semimartingales. The quadratic covariation (or bracket process) of X, Y is the process $\{[X, Y]_t\}_{t\in\mathbb{R}_+}$ defined by

$$[X, Y]_t := X_t Y_t - \int_0^t X_{s-} dY_s - \int_0^t Y_{s-} dX_s, \quad t \geq 0,$$

where $X_{0-} := 0, Y_{0-} := 0$. The predictable quadratic covariation of X and Y is the unique predictable process $\{\langle X, Y\rangle_t\}_{t\in\mathbb{R}_+}$ such that $\{[X, Y]_t - \langle X, Y\rangle_t\}_{t\in\mathbb{R}_+}$ is a local martingale. If $X = Y$, then $[X] \equiv [X, X]$ and $\langle X\rangle \equiv \langle X, X\rangle$ are, respectively, the quadratic and predictable quadratic variation processes of X.

Proof of Lemma 3.3.1. The proof is split into the following steps: Finding representations for the processes $\hat{X}^{(n)}, \hat{Y}^{(n)}, \hat{Z}^{(n)}, \hat{X}^{(n)} - \hat{Z}^{(n)}$, and $\hat{\eta}^{(n)}$ in terms of their generators and corresponding martingales (**step 1**). Establishing bounds for the expected values of the squared suprema of the catalyst and reactant processes over finite time horizons and analogous bounds for martingales associated with the catalyst, reactant, and shadow processes (**step 2**). Using these bounds to prove the tightness of $\hat{X}^{(n)}$, $\hat{Y}^{(n)}$, and $\hat{\eta}^{(n)}$ (**step 3**).

Step 1: Recall that $\hat{W}^{(n)} = (\hat{X}^{(n)}, \hat{Y}^{(n)}, \hat{Z}^{(n)})$ and $\boldsymbol{w} = (x, y, z)$. For $\boldsymbol{x} = (x_1, x_2, x_3) \in \mathbb{W}$, let $\phi_i(\boldsymbol{x}) = x_i$, $i = 1, 2, 3$ and $h := \phi_1 - \phi_3$. Then

$$\hat{Z}_t^{(n)} = \phi_3(\hat{W}_t^{(n)}) = \hat{Z}_0^{(n)} + \int_0^t \hat{\mathcal{A}}^{(n)} \phi_3(\hat{W}_s^{(n)}) ds + M_t^{(n)}(\phi_3), \tag{3.3.1}$$

where $M^{(n)}(\phi_3)$ is a local martingale. Using (3.1.1), we get

$$\hat{\mathcal{A}}^{(n)}\phi_3(\hat{W}_t^{(n)}) = \lambda_1^{(n)} n \hat{X}_t^{(n)} \sum_{k=0}^{\infty}(k-1)\mu_1^{(n)}(k) = c_1^{(n)}\lambda_1^{(n)}\hat{X}_t^{(n)} \tag{3.3.2}$$

and

$$\hat{\mathcal{A}}^{(n)}\phi_3^2(\hat{W}_t^{(n)}) = \lambda_1^{(n)} n^2 \hat{X}_t^{(n)} \sum_{k=0}^{\infty}\left[\left(\hat{Z}_t^{(n)} + \frac{k-1}{n}\right)^2 - \left(\hat{Z}_t^{(n)}\right)^2\right]\mu_1^{(n)}(k)$$
$$= \lambda_1^{(n)} n^2 \hat{X}_t^{(n)} \sum_{k=0}^{\infty}\left[\frac{k-1}{n}\left(2\hat{Z}_t^{(n)} + \frac{k-1}{n}\right)\right]\mu_1^{(n)}(k) = 2\lambda_1^{(n)}\hat{X}_t^{(n)}\hat{Z}_t^{(n)}c_1^{(n)} + \lambda_1^{(n)}\hat{X}_t^{(n)}\alpha_1^{(n)}.$$

Define maps $b_3^{(n)}$ and $a_{3,3}^{(n)}$ from $\mathbb{W}$ to $\mathbb{R}$ as

$$b_3^{(n)}(\boldsymbol{w}) := \hat{\mathcal{A}}^{(n)}\phi_3(\boldsymbol{w}) = c_1^{(n)}\lambda_1^{(n)}x, \quad a_{3,3}^{(n)}(\boldsymbol{w}) := \hat{\mathcal{A}}^{(n)}\phi_3^2(\boldsymbol{w}) - 2zb_3^{(n)}(\boldsymbol{w}) = \lambda_1^{(n)}x\alpha_1^{(n)},$$

$\boldsymbol{w} = (x, y, z) \in \mathbb{W}$. A routine calculation then shows (see [20], Lemma 3.1.3) that

$$\langle M^{(n)}(\phi_3)\rangle_t = \int_0^t a_{3,3}^{(n)}(\hat{W}_s^{(n)})ds = \lambda_1^{(n)}\alpha_1^{(n)}\int_0^t \hat{X}_s^{(n)}ds. \tag{3.3.3}$$

Next, since $\hat{X}_0^{(n)} = \hat{Z}_0^{(n)}$, we have

$$\hat{X}_t^{(n)} - \hat{Z}_t^{(n)} = h(\hat{W}_t^{(n)}) = \int_0^t \hat{\mathcal{A}}^{(n)}h(\hat{W}_s^{(n)})ds + M_t^{(n)}(h),$$

where $M^{(n)}(h)$ is a local martingale, and, once more using (3.1.1),

$$\hat{\mathcal{A}}^{(n)}h(\boldsymbol{w}) = \lambda_1^{(n)} n \mu_1^{(n)}(0) 1_{\{x=1\}}.$$

Thus with $\hat{\eta}^{(n)}$ as in (3.1.3) we get

$$\hat{X}_t^{(n)} - \hat{Z}_t^{(n)} = h(\hat{W}_t^{(n)}) = \hat{\eta}_t^{(n)} + M_t^{(n)}(h). \tag{3.3.4}$$

Let $\hat{U}^{(n)} := \hat{X}^{(n)} - \hat{Z}^{(n)}$. Then

$$\left(\hat{U}_t^{(n)}\right)^2 = 2\int_0^t \hat{U}_{s-}^{(n)} d\hat{U}_s^{(n)} + [\hat{U}^{(n)}]_t = 2\int_0^t \hat{U}_s^{(n)} d\hat{\eta}_s^{(n)} + 2\int_0^t \hat{U}_{s-}^{(n)} dM_s^{(n)}(h) + [\hat{U}^{(n)}]_t$$
$$\equiv 2\int_0^t \hat{U}_s^{(n)} d\hat{\eta}_s^{(n)} + L_t^{(n)} + [\hat{U}^{(n)}]_t,$$

where $L^{(n)}$ is again a local martingale. On the other hand, using the generator $\hat{\mathcal{A}}^{(n)}$, we find

$$\left(\hat{U}_t^{(n)}\right)^2 = h^2(\hat{W}_t^{(n)}) = \int_0^t \hat{\mathcal{A}}^{(n)} h^2(\hat{W}_s^{(n)}) ds + M_t^{(n)}(h^2),$$

where $M^{(n)}(h^2)$ is a local martingale. The last two representations of $\left(\hat{U}_t^{(n)}\right)^2$ imply that

$$\langle \hat{U}^{(n)} \rangle_t = \int_0^t \hat{\mathcal{A}}^{(n)} h^2(\hat{W}_s^{(n)}) ds - 2\int_0^t \hat{U}_s^{(n)} d\hat{\eta}_s^{(n)}. \tag{3.3.5}$$

Now, with $\boldsymbol{w} = (x, y, z)$,

$$\hat{\mathcal{A}}^{(n)} h^2(\boldsymbol{w}) = \lambda_1^{(n)} n^2 \mu_1^{(n)}(0) \left(\left(1 - z + \frac{1}{n}\right)^2 - (1-z)^2 \right) 1_{\{x=1\}}$$
$$= \lambda_1^{(n)} n \mu_1^{(n)}(0) \left(\frac{1}{n} + 2(1-z) \right) 1_{\{x=1\}}.$$

Thus, with $\hat{W}^{(n)} = (\hat{X}^{(n)}, \hat{Y}^{(n)}, \hat{Z}^{(n)})$,

$$\int_0^t \hat{\mathcal{A}}^{(n)} h^2(\hat{W}_s^{(n)}) ds = \int_0^t \left(\frac{1}{n} + 2(1 - \hat{Z}_s^{(n)}) \right) d\hat{\eta}_s^{(n)}.$$

Moreover, recalling that $\hat{U}^{(n)} = \hat{X}^{(n)} - \hat{Z}^{(n)}$ and that $\hat{\eta}^{(n)}$ increases only when $\hat{X}^{(n)} = 1$, we have

$$2\int_0^t \hat{U}_s^{(n)} d\hat{\eta}_s^{(n)} = 2\int_0^t (1 - \hat{Z}_s^{(n)}) d\hat{\eta}_s^{(n)}.$$

Combining the last two equations with (3.3.5), we get

$$\langle \hat{U}^{(n)} \rangle_t = \int_0^t \left(\frac{1}{n} + 2(1 - \hat{Z}_s^{(n)}) \right) d\hat{\eta}_s^{(n)} - 2 \int_0^t (1 - \hat{Z}_s^{(n)}) d\hat{\eta}_s^{(n)}$$
$$= \frac{1}{n} \hat{\eta}_t^{(n)} = \lambda_1^{(n)} \mu_1^{(n)}(0) \int_0^t \hat{X}_s^{(n)} 1_{\{\hat{X}_s^{(n)} = 1\}} ds.$$

Also, since $\hat{\eta}^{(n)}$ is a continuous process with bounded variation, $[\hat{\eta}^{(n)}] = [\hat{\eta}^{(n)}, M^{(n)}(h)] = 0$, and consequently

$$\langle M^{(n)}(h) \rangle_t = \langle \hat{U}^{(n)} \rangle_t = \lambda_1^{(n)} \mu_1^{(n)}(0) \int_0^t \hat{X}_s^{(n)} 1_{\{\hat{X}_s^{(n)} = 1\}} ds. \tag{3.3.6}$$

Next, by (3.3.1), (3.3.2), and (3.3.4), we have

$$\hat{X}_t^{(n)} = \hat{X}_0^{(n)} + c_1^{(n)} \lambda_1^{(n)} \int_0^t \hat{X}_s^{(n)} ds + M_t^{(n)}(\phi_3) + M_t^{(n)}(h) + \hat{\eta}_t^{(n)}. \tag{3.3.7}$$

Since $\hat{\eta}^{(n)}$ is non-decreasing and $\int_0^\infty 1_{\{\hat{X}_s^{(n)} \neq 1\}} d\hat{\eta}_s^{(n)} = 0$, we have from the characterization given above (3.1.2) that

$$\hat{X}_t^{(n)} = \Gamma \left(\hat{X}_0^{(n)} + c_1^{(n)} \lambda_1^{(n)} \int_0^\cdot \hat{X}_s^{(n)} ds + M_\cdot^{(n)}(\phi_3) + M_\cdot^{(n)}(h) \right) (t). \tag{3.3.8}$$

We next consider the reactant population. With similar calculations as for $\hat{\mathcal{A}}^{(n)} \phi_3(\hat{W}_t^{(n)})$, we get

$$\hat{Y}_t^{(n)} = \hat{Y}_0^{(n)} + \int_0^t b_2^{(n)}(\hat{W}_s^{(n)}) ds + M_t^{(n)}(\phi_2), \tag{3.3.9}$$

where

$$b_2^{(n)}(\hat{W}_t^{(n)}) := \hat{\mathcal{A}}^{(n)} \phi_2(\hat{W}_t^{(n)}) = c_2^{(n)} \lambda_2^{(n)} \hat{X}_t^{(n)} \hat{Y}_t^{(n)} \tag{3.3.10}$$

and

$$\langle M^{(n)}(\phi_2)\rangle_t = \lambda_2^{(n)}\alpha_2^{(n)}\int_0^t \hat{X}_s^{(n)}\hat{Y}_s^{(n)}ds. \tag{3.3.11}$$

Let $H^{(n)} := M^{(n)}(\phi_3) + M^{(n)}(h)$. Using (3.3.3), (3.3.6), and the Kunita-Watanabe inequality (see e.g. [32]), we have a bound for its predictable quadratic variation:

$$\langle H^{(n)}\rangle_t \leq \left(\sqrt{\langle M^{(n)}(\phi_3)\rangle_t} + \sqrt{\langle M^{(n)}(h)\rangle_t}\right)^2 \leq \lambda_1^{(n)}(\alpha_1^{(n)} + \mu_1^{(n)}(0))\int_0^t \hat{X}_s^{(n)}ds. \tag{3.3.12}$$

Let

$$\hat{N}_t^{(n)} := \hat{X}_0^{(n)} + c_1^{(n)}\lambda_1^{(n)}\int_0^t \hat{X}_s^{(n)}ds + H_t^{(n)}. \tag{3.3.13}$$

Then, from (3.3.8), $\hat{X}_t^{(n)} = \Gamma(\hat{N}_\cdot^{(n)})(t)$. The Lipschitz continuity of the Skorohod map implies

$$\sup_{t\leq T}|\hat{X}_t^{(n)} - 1| \leq 2\sup_{t\leq T}|\hat{N}_t^{(n)}|. \tag{3.3.14}$$

Step 2 We will establish the following bounds, the proofs of which are adapted from Lemma 3.2.2 of [20]. For each $T \geq 0$ there is a $K_T \in (0,\infty)$ such that

$$\sup_{n\in\mathbb{N}} E\left(\sup_{t\leq T}|\hat{X}_t^{(n)}|^2\right) \leq K_T, \tag{3.3.15}$$

$$\sup_{n\in\mathbb{N}} E\left(\sup_{t\leq T}|H_t^{(n)}|^2\right) \leq K_T, \tag{3.3.16}$$

$$\sup_{n\in\mathbb{N}} E\left(\sup_{t\leq T}|\hat{N}_t^{(n)}|^2\right) \leq K_T, \tag{3.3.17}$$

$$\sup_{n\in\mathbb{N}} E\left(\sup_{t\le T}|\hat{\eta}_t^{(n)}|^2\right) \le K_T, \tag{3.3.18}$$

and

$$\sup_{n\in\mathbb{N}} E\left(\sup_{t\le T}\left|\hat{Y}^{(n)}_{\sigma_k^{(n)}\wedge t}\right|^2\right) \le \exp(K_T k^2). \tag{3.3.19}$$

We first consider (3.3.15). Using Doob's inequality, we get

$$E\sup_{t\le T}|H_t^{(n)}|^2 \le 4E|H_T^{(n)}|^2 = 4E[H^{(n)}]_T = 4E\langle H^{(n)}\rangle_T.$$

Combining this with (3.3.12), we have

$$E\sup_{t\le T}|H_t^{(n)}|^2 \le 4\lambda_1^{(n)}(\alpha_1^{(n)}+\mu_1^{(n)}(0))E\left(\int_0^T \hat{X}_s^{(n)}ds\right). \tag{3.3.20}$$

Using (3.3.14), we get

$$\sup_{t\le T}|X_t^{(n)}|^2 \equiv |\hat{X}^{(n)}|^2_{*,T} = \left(|\hat{X}^{(n)}-1|_{*,T}+1\right)^2 \le 2|\hat{X}^{(n)}-1|^2_{*,T}+2 \le 8|\hat{N}^{(n)}|^2_{*,T}+2.$$

Combining this with (3.3.13) and (3.3.20), we obtain

$$\begin{aligned} E\left(|\hat{X}^{(n)}|^2_{*,T}\right) &\le 2+8E\left(|\hat{N}^{(n)}|^2_{*,T}\right) \\ &\le 2+24\left[1+\left(T\left(c_1^{(n)}\lambda_1^{(n)}\right)^2+4\lambda_1^{(n)}(\alpha_1^{(n)}+\mu_1^{(n)}(0))\right)\int_0^T E\big|\hat{X}^{(n)}|^2_{*,s}\big)ds\right]. \end{aligned}$$

Using Gronwall's inequality, we get

$$E\left(|\hat{X}^{(n)}|^2_{*,T}\right) \le 26\exp\left(K_{1,T}^{(n)}\right),$$

where $K_{1,T}^{(n)} := 24T\left(T\left(c_1^{(n)}\lambda_1^{(n)}\right)^2 + 4\lambda_1^{(n)}(\alpha_1^{(n)} + \mu_1^{(n)}(0))\right)$. Since $c_1^{(n)}, \lambda_1^{(n)}$, and $\alpha^{(n)}$ converge as $n \to \infty$, we have (3.3.15).

The estimate in (3.3.16) is now an immediate consequence of (3.3.15) and (3.3.20). Using the estimates in (3.3.15) and (3.3.16), the estimate in (3.3.17) follows immediately from (3.3.13), and that in (3.3.18) follows from (3.3.7).

We next establish (3.3.19). Using Doob's inequality once more, we have

$$E\left(\sup_{t\le T}|M^{(n)}_{\sigma_k^{(n)}\wedge t}(\phi_2)|^2\right) \le 4E\left(\langle M^{(n)}(\phi_2)\rangle_{\sigma_k^{(n)}\wedge T}\right).$$

Combining this with (3.3.11), we obtain

$$E\left(\sup_{t\le T}|M^{(n)}_{\sigma_k^{(n)}\wedge t}(\phi_2)|^2\right) \le 4\lambda_2^{(n)}\alpha_2^{(n)}E\left(\int_0^{\sigma_k^{(n)}\wedge T}\hat{X}_s^{(n)}\hat{Y}_s^{(n)}ds\right).$$

Using (3.3.9) and (3.3.10), we now get

$$E\left(|\hat{Y}^{(n)}|^2_{*,T\wedge\sigma_k^{(n)}}\right) \le 3\left(1+\left[T\left(c_2^{(n)}\lambda_2^{(n)}k\right)^2 + 4\lambda_2^{(n)}\alpha_2^{(n)}k\right]\int_0^T E\left(|\hat{Y}^{(n)}|^2_{*,s\wedge\sigma_k^{(n)}}\right)ds\right).$$

Using Gronwall's inequality, we have

$$E\left(\sup_{t\le T}|\hat{Y}^{(n)}_{\sigma_k^{(n)}\wedge t}|^2\right) \le 3\exp(K_{2,T}^{(n)}k^2),$$

where $K_{2,T}^{(n)} = T\left(T\left(c_2^{(n)}\lambda_2^{(n)}\right)^2 + 4\lambda_2^{(n)}\alpha_2^{(n)}\right)$, and thus (3.3.19) follows.

Step 3 In order to show tightness of $\{\hat{X}^{(n)}\}$ and $\{\hat{\eta}^{(n)}\}$, it suffices, due to the Lipschitz continuity of the Skorohod map, to show that $\{\hat{N}^{(n)}\}$ is tight. For this, in view of (3.3.17), it suffices to show that the following condition (Aldous' criterion) holds: For each $N > 0, \epsilon > 0$, and $\gamma > 0$ there are $\delta > 0$ and n_0 such that for all stopping times

$\{\tau_n\}_{n\in\mathbb{N}}$ with $\tau_n \le N$, we have

$$\sup_{n\ge n_0}\sup_{\theta\le\delta} P(|\hat{N}^{(n)}_{\tau_n+\theta} - \hat{N}^{(n)}_{\tau_n}| \ge \gamma) \le \epsilon. \tag{3.3.21}$$

Let $N, \epsilon, \gamma \in (0,\infty)$ be given. Note that

$$P(|\hat{N}^{(n)}_{\tau_n+\theta} - \hat{N}^{(n)}_{\tau_n}| \ge \gamma) \le P\left(|c_1^{(n)}\lambda_1^{(n)}\int_{\tau_n}^{\tau_n+\theta}\hat{X}^{(n)}_s ds| \ge \frac{\gamma}{2}\right) + P\left(|H^{(n)}_{\tau_n+\theta} - H^{(n)}_{\tau_n}| \ge \frac{\gamma}{2}\right),$$

By (3.3.15), we have $\sup_{n\in\mathbb{N}}\sup_{\theta\le\delta} P\left(|c_1^{(n)}\lambda_1^{(n)}\int_{\tau_n}^{\tau_n+\theta}\hat{X}^{(n)}_s ds| \ge \frac{\gamma}{2}\right) < \frac{\epsilon}{2}$ for δ sufficiently small. It remains to prove that, for some $\delta > 0$,

$$\sup_{n\in\mathbb{N}}\sup_{\theta\le\delta} P\left(|H^{(n)}_{\tau_n+\theta} - H^{(n)}_{\tau_n}| \ge \frac{\gamma}{2}\right) < \frac{\epsilon}{2}. \tag{3.3.22}$$

Note that

$$\begin{aligned} P\Big(|H^{(n)}_{\tau_n+\theta} - H^{(n)}_{\tau_n}| \ge \frac{\gamma}{2}\Big) &\le \frac{E(|H^{(n)}_{\tau_n+\theta} - H^{(n)}_{\tau_n}|^2)}{(\gamma/2)^2} = \frac{E((H^{(n)}_{\tau_n+\theta})^2) - E((H^{(n)}_{\tau_n})^2)}{(\gamma/2)^2} \\ &= \frac{E\langle H^{(n)}\rangle_{\tau_n+\theta} - E\langle H^{(n)}\rangle_{\tau_n}}{(\gamma/2)^2}. \end{aligned}$$

Recalling that $H^{(n)} = M^{(n)}(\phi_3) + M^{(n)}(h)$, and using properties of the predictable quadratic variation process, we get

$$\begin{aligned} &E\langle H^{(n)}\rangle_{\tau_n+\theta} - E\langle H^{(n)}\rangle_{\tau_n} \\ &\quad\le 3E\left(\langle M^{(n)}(\phi_3)\rangle_{\tau_n+\theta} - \langle M^{(n)}(\phi_3)\rangle_{\tau_n} + \langle M^{(n)}(h)\rangle_{\tau_n+\theta} - \langle M^{(n)}(h)\rangle_{\tau_n}\right) \\ &\quad= 3E\left(\lambda_1^{(n)}\alpha_1^{(n)}\int_{\tau_n}^{\tau_n+\theta}\hat{X}^{(n)}_s ds + \lambda_1^{(n)}\mu_1^{(n)}(0)\int_{\tau_n}^{\tau_n+\theta}\hat{X}^{(n)}_s 1_{\{\hat{X}^{(n)}_s=1\}} ds\right). \end{aligned}$$

Using (3.3.15) once more, we can choose $\delta > 0$ such that (3.3.22) holds. This proves tightness of $\{\hat{N}^{(n)}\}$ and consequently that of $\{\hat{X}^{(n)}\}$ and $\{\hat{\eta}^{(n)}\}$.

We next show tightness of $\{\hat{Y}^{(n)}\}$. The calculations are similar to those for $\{\hat{X}^{(n)}\}$ and thus only a sketch is provided. Fix $\epsilon > 0$. Using (3.3.19), we get, for $K \in (0, \infty)$,

$$P\Big(\sup_{t\leq T}|\hat{Y}^{(n)}_t| > K\Big) \leq P\Big(\sup_{t\leq T}|\hat{Y}^{(n)}_{\sigma^{(n)}_k\wedge t}| > K \text{ and } \sigma^{(n)}_k > T\Big) + P\left(\sigma^{(n)}_k \leq T\right)$$
$$\leq \frac{E\Big(\sup_{t\leq T}|\hat{Y}^{(n)}_{\sigma^{(n)}_k\wedge t}|^2\Big)}{K^2} + P\Big(\sup_{t\leq T}|\hat{X}^{(n)}_t| \geq k\Big) \leq \frac{\exp(K_T k^2)}{K^2} + \frac{E\sup_{t\leq T}|\hat{X}^{(n)}_t|^2}{k^2}.$$

Using (3.3.15), we can choose k such that

$$\sup_{n\in\mathbb{N}} \frac{E\sup_{t\leq T}|\hat{X}^{(n)}_t|^2}{k^2} < \frac{\epsilon}{2}. \tag{3.3.23}$$

Now choose K such that

$$\frac{\exp(K_T k^2)}{K^2} < \frac{\epsilon}{2}.$$

The last two displays imply $\sup_{n\in\mathbb{N}} P(\sup_{t\leq T}|\hat{Y}^{(n)}_t| > K) < \epsilon$, and since $\epsilon > 0$ is arbitrary, the tightness of the random variables $\{\hat{Y}^{(n)}_t\}_{n\in\mathbb{N}}$, for each $t \geq 0$, follows. To establish the tightness of the processes $\{\hat{Y}^{n)}\}_{n\in\mathbb{N}}$, we need, as for the family of catalyst processes, a bound for the fluctuations of the processes. It suffices to show that for each $N > 0, \epsilon > 0$, and $\gamma > 0$ there are $\delta > 0$ and n_0 such that for all stopping times $\{\tau_n\}_{n\in\mathbb{N}}$ with $\tau_n \leq N$, we have

$$\sup_{n\geq n_0}\sup_{\theta\leq\delta} P\left(|\hat{Y}^{(n)}_{\tau_n+\theta} - \hat{Y}^{(n)}_{\tau_n}| \geq \gamma\right) \leq \epsilon. \tag{3.3.24}$$

Fix $N, \epsilon, \gamma \in (0, \infty)$. Then, for any $\theta \in (0, 1)$,

$$P\left(|\hat{Y}^{(n)}_{\tau_n+\theta} - \hat{Y}^{(n)}_{\tau_n}| \geq \gamma\right) \leq P\Big(|\hat{Y}^{(n)}_{(\tau_n+\theta)\wedge\sigma^{(n)}_k} - \hat{Y}^{(n)}_{\tau_n\wedge\sigma^{(n)}_k}| \geq \gamma\Big) + P\left(\sigma^{(n)}_k \leq N+1\right).$$

Let $T = N+1$ and k as in (3.3.23), then $P(\sigma_k^{(n)} < N+1) < \epsilon/2$ for all $n \in \mathbb{N}$. For the first term on the right hand side of the last display, we get, using (3.3.9) and that $\sup_{t \le T \wedge \sigma_k^{(n)}} \hat{X}_t^{(n)} \le k$,

$$
\begin{aligned}
&P\left(|\hat{Y}^{(n)}_{(\tau_n+\theta)\wedge\sigma_k^{(n)}} - \hat{Y}^{(n)}_{\tau_n\wedge\sigma_k^{(n)}}| \ge \gamma\right) \qquad (3.3.25)\\
&\le P\Big(|c_2^{(n)}\lambda_2^{(n)} \int_{\tau_n\wedge\sigma_k^{(n)}}^{(\tau_n+\theta)\wedge\sigma_k^{(n)}} \hat{Y}_s^{(n)} ds| \ge \frac{\gamma}{2k}\Big) + P\Big(|M^{(n)}_{(\tau_n+\theta)\wedge\sigma_k^{(n)}}(\phi_2) - M^{(n)}_{\tau_n\wedge\sigma_k^{(n)}}(\phi_2)| \ge \frac{\gamma}{2}\Big).
\end{aligned}
$$

The first term on the right hand side can be bounded as follows:

$$
\begin{aligned}
P\left(\left|c_2^{(n)}\lambda_2^{(n)} \int_{\tau_n\wedge\sigma_k^{(n)}}^{(\tau_n+\theta)\wedge\sigma_k^{(n)}} \hat{Y}_s^{(n)} ds\right| \ge \frac{\gamma}{2k}\right) &\le \left(\frac{2kc_2^{(n)}\lambda_2^{(n)}}{\gamma}\right)^2 E\left(\left(\int_{\tau_n\wedge\sigma_k^{(n)}}^{(\tau_n+\theta)\wedge\sigma_k^{(n)}} \hat{Y}_s^{(n)} ds\right)^2\right)\\
&\le \theta\left(\frac{2kc_2^{(n)}\lambda_2^{(n)}}{\gamma}\right)^2 \exp\left(K_{N+1}k^2\right),
\end{aligned}
$$

where K_{N+1} is the constant from (3.3.19). Thus, for δ sufficiently small, we get

$$
\sup_{n\in\mathbb{N}} \sup_{\theta\le\delta} P\left(\left|c_2^{(n)}\lambda_2^{(n)} \int_{\tau_n\wedge\sigma_k^{(n)}}^{(\tau_n+\theta)\wedge\sigma_k^{(n)}} \hat{Y}_s^{(n)} ds\right| \ge \frac{\gamma}{2k}\right) < \epsilon/4.
$$

The second term on the right hand side of (3.3.25) can be bounded as follows:

$$
\begin{aligned}
&P\left(\left|M^{(n)}_{(\tau_n+\theta)\wedge\sigma_k^{(n)}}(\phi_2) - M^{(n)}_{\tau_n\wedge\sigma_k^{(n)}}(\phi_2)\right| \ge \frac{\gamma}{2}\right)\\
&\quad\le \frac{E\left(\left\langle M^{(n)}(\phi_2)\right\rangle_{(\tau_n+\theta)\wedge\sigma_k^{(n)}} - \left\langle M^{(n)}(\phi_2)\right\rangle_{\tau_n\wedge\sigma_k^{(n)}}\right)}{(\gamma/2)^2}\\
&\quad\le \frac{4}{\gamma^2}\lambda_2^{(n)}\alpha_2^{(n)} E\int_{\tau_n\wedge\sigma_k^{(n)}}^{(\tau_n+\theta)\wedge\sigma_k^{(n)}} \hat{X}_s^{(n)}\hat{Y}_s^{(n)} ds \le \frac{4}{\gamma^2}\lambda_2^{(n)}\alpha_2^{(n)} k\theta E\left(\sup_{0\le s\le N+1} Y^{(n)}_{s\wedge\sigma_k^{(n)}}\right).
\end{aligned}
$$

Using (3.3.19) once more, we have that, for δ sufficiently small, the second term in (3.3.25) is bounded by $\frac{\epsilon}{4}$. Combining the above estimates, we now see that (3.3.24) holds, and thus that $\{\hat{Y}^{(n)}\}_{n\in\mathbb{N}}$ is tight. $\square$

The following martingale characterization result will be useful in the proof of Theorem 3.1.2. The proof of the characterization is standard and is omitted (see Theorem 4.5.2 in [39]).

For $\phi \in C_c^\infty(\mathbb{R}_{\geq 1} \times \mathbb{R}_+)$, let

$$\begin{aligned}\mathcal{L}\phi(x,y) := c_1\lambda_1 x\frac{\partial}{\partial x}\phi(x,y) + \frac{1}{2}\alpha_1\lambda_1 x\frac{\partial^2}{\partial x^2}\phi(x,y)\\ + c_2\lambda_2 xy\frac{\partial}{\partial y}\phi(x,y) + \frac{1}{2}\alpha_2\lambda_2 xy\frac{\partial^2}{\partial y^2}\phi(x,y).\end{aligned}$$

Let $\tilde{\Omega} := D(\mathbb{R}_+ : \mathbb{R}_{\geq 1} \times \mathbb{R}_+^2)$ and $\tilde{\mathcal{F}}$ be the corresponding Borel σ-field (with respect to the Skorohod topology). Denote by $\{\mathcal{F}_t\}_{t\in\mathbb{R}_+}$ the canonical filtration on $(\tilde{\Omega}, \tilde{\mathcal{F}})$, i.e. $\mathcal{F}_t = \sigma(\pi_s | s \leq t)$, where $\pi_s(\tilde{\omega}) = \tilde{\omega}_s = \tilde{\omega}(s)$ for $\tilde{\omega} \in \tilde{\Omega}$. Finally, let $\pi^{(i)}$, $i = 1, 2, 3$, be the coordinate processes, i.e. $(\pi^{(1)}(\tilde{\omega}), \pi^{(2)}(\tilde{\omega}), \pi^{(3)}(\tilde{\omega})) = \pi(\tilde{\omega})$.

Theorem 3.3.1. *Let $\tilde{P}$ be a probability measure on $(\tilde{\Omega}, \tilde{\mathcal{F}})$ under which the following hold a.s.:*

(i) $\pi^{(3)}$ *is a non-decreasing, continuous process, and* $\pi_0^{(3)} = 0$,

(ii) $(\pi^{(1)}, \pi^{(2)})$ *is an* $(\mathbb{R}_{\geq 1} \times \mathbb{R}_+)$ *valued continuous process,*

(iii) $\int_0^\infty 1_{(1,\infty)}(\pi_s^{(1)})d\pi_s^{(3)} = 0$,

(iv) for all $\phi \in C_c^\infty(\mathbb{R}_{\geq 1} \times \mathbb{R}_+)$

$$\phi(\pi_t^{(1)}, \pi_t^{(2)}) - \int_0^t \mathcal{L}\phi(\pi_s^{(1)}, \pi_s^{(2)})ds - \int_0^t \frac{\partial\phi}{\partial x}(1, \pi_s^{(2)})d\pi_s^{(3)}$$

is an $\{\mathcal{F}_t\}$ *martingale, and*

(v) $(\pi_0^{(1)}, \pi_0^{(2)}) = (1, 1)$.

Then $\tilde{P} \circ (\pi^{(1)}, \pi^{(2)})^{-1} = \bar{P} \circ (X, Y)^{-1}$, *where* X, Y, *and* $\bar{P}$ *are as in Proposition 3.1.1.*

Proof of Theorem 3.1.2. Recall that for $\phi \in C_c^\infty(\mathbb{R}_{\geq 1} \times \mathbb{R}_+)$, we have

$$\phi(\hat{X}_t^{(n)}, \hat{Y}_t^{(n)}) = \phi(\hat{X}_0^{(n)}, \hat{Y}_0^{(n)}) + \int_0^t \hat{\mathcal{A}}^{(n)} \phi(\hat{X}_s^{(n)}, \hat{Y}_s^{(n)}) ds + M_t^{(n)}(\phi), \tag{3.3.26}$$

where $M_t^{(n)}(\phi)$ is a local martingale, and $\hat{\mathcal{A}}^{(n)}$, defined in (3.1.1), can be rewritten as

$$\hat{\mathcal{A}}^{(n)} \phi(x, y) = \mathcal{L}^{(n)} \phi(x, y) + \mathcal{D}^{(n)} \phi(y) n \lambda_1^{(n)} \mu_1^{(n)}(0) 1_{\{x=1\}},$$

where

$$\begin{aligned} \mathcal{L}^{(n)} \phi(x, y) := \lambda_1^{(n)} n^2 x \sum_{k=0}^\infty \left[\phi\left(x + \frac{k-1}{n}, y \right) - \phi(x, y) \right] \mu_1^{(n)}(k) \\ + \lambda_2^{(n)} n^2 xy \sum_{k=0}^\infty \left[\phi\left(x, y + \frac{k-1}{n} \right) - \phi(x, y) \right] \mu_2^{(n)}(k) \end{aligned}$$

and

$$\mathcal{D}^{(n)} \phi(y) := n \left(\phi(1, y) - \phi\left(1 - \frac{1}{n}, y \right) \right).$$

Thus, (3.3.26) can be rewritten as

$$\begin{aligned} \phi(\hat{X}_t^{(n)}, \hat{Y}_t^{(n)}) = \phi(\hat{X}_0^{(n)}, \hat{Y}_0^{(n)}) + \int_0^t \mathcal{L}^{(n)} \phi(\hat{X}_s^{(n)}, \hat{Y}_s^{(n)}) ds \\ + \int_0^t \mathcal{D}^{(n)} \phi(\hat{Y}_s^{(n)}) d\hat{\eta}_s^{(n)} + M_t^{(n)}(\phi). \end{aligned}$$

Recall the path space $(\tilde{\Omega}, \tilde{\mathcal{F}})$ introduced above Theorem 3.3.1. We denote by $\tilde{P}^{(n)}$ the measure induced by $(\hat{X}^{(n)}, \hat{Y}^{(n)}, \hat{\eta}^{(n)})$ on $(\tilde{\Omega}, \tilde{\mathcal{F}})$ and by $\tilde{E}^{(n)}$ the corresponding expectation.

From Lemma 3.3.1, $\tilde{P}^{(n)}$ is tight. Let $\tilde{P}$ be a limit point of $\{\tilde{P}^{(n)}\}$ along some subsequence $\{n_k\}$. In order to complete the proof, it suffices to show that under $\tilde{P}$ properties (i)-(v) in Theorem 3.3.1 hold almost surely. Property (i) is immediate from the fact that $\hat{\eta}^{(n)}$ is non-decreasing and continuous with initial value 0 for each n. Also, property (v) is immediate from the fact that $(\hat{X}_0^{(n)}, \hat{Y}_0^{(n)}) = (1,1)$, a.s., for each n. Next, consider property (ii). In order to establish continuity of $\pi^{(1)}$, it suffices (see [18], Proposition VI.3.26, p. 315) to show that

$$\lim_{n\to\infty} P\left(\sup_{0\le t\le T} |\Delta \hat{X}_t^{(n)}| \ge \epsilon\right) = 0, \tag{3.3.27}$$

where $\Delta \hat{X}_t^{(n)} := \hat{X}_t^{(n)} - \hat{X}_{t-}^{(n)}$ and $\Delta \hat{X}_0^{(n)} := 0$. Let $N_T^{(n)}$ be the number of deaths of particles of the (unscaled) process $X^{(n)}$ in the time interval $[0,T]$. Fix $\epsilon, \delta > 0$. Then

$$P\big(\sup_{0\le t\le T} |\Delta \hat{X}_t^{(n)}| \ge \epsilon\big) \le P\big(\sup_{0\le t\le T} |\Delta \hat{X}_t^{(n)}| \ge \epsilon; \sup_{0\le t\le T} X_t^{(n)} \le nL\big) + P\big(\sup_{0\le t\le T} X_t^{(n)} > nL\big).$$

By (3.3.15), we can choose $L \in (0,\infty)$ such that $P\big(\sup_{0\le t\le T} X_t^{(n)} > nL\big) < \frac{\delta}{3}$, for $n \in \mathbb{N}$. Next, consider

$$\begin{aligned}
&P\big(\sup_{0\le t\le T} |\Delta \hat{X}_t^{(n)}| \ge \epsilon; \sup_{0\le t\le T} X_t^{(n)} \le nL\big) \\
&\le P\big(\sup_{0\le t\le T} |\Delta \hat{X}_t^{(n)}| \ge \epsilon; N_T^{(n)} < nCL\big) + P\big(\sup_{0\le t\le T} X_t^{(n)} \le nL; N_T^{(n)} \ge nCL\big).
\end{aligned}$$

Note that on the set $\{\sup_{0\le t\le T} X_t^{(n)} \le nL\}$ the branching rates of $X^{(n)}$ are bounded during the time interval $[0,T]$, uniformly in n, and thus we can choose a $C \in (0,\infty)$ such that

$$P\big(\sup_{0\le t\le T} X_t^{(n)} \le nL; N_T^{(n)} \ge nCL\big) < \frac{\delta}{3}.$$

Finally, let, for $n \in \mathbb{N}$, $\{\xi_i^{(n)}\}_{i\in\mathbb{N}}$ be i.i.d. random variables distributed as $\mu_1^{(n)}$. Then

$$
\begin{aligned}
P\left(\sup_{0\le t\le T}|\Delta\hat{X}_t^{(n)}| \ge \epsilon; N_T^{(n)} < nCL\right) &\le P\left(\max_{1\le i<nCL}\frac{|\xi_i^{(n)}-1|}{n} \ge \epsilon\right) \\
\le \sum_{i=1}^{nCL-1} P\left(|\xi_i^{(n)}-1| \ge n\epsilon\right) &\le \sum_{i=1}^{nCL-1}\frac{E\left(|\xi_i^{(n)}-1|^2\right)}{(n\epsilon)^2} < \frac{\delta}{3},
\end{aligned}
$$

for $n \ge n_0(\delta)$, since the variance of the offspring distribution converges. Combining the above estimates, (3.3.27) follows. The continuity of $\pi^{(2)}$ is established similarly.

To see (iii), consider, for $\delta > 0$, continuous bounded test functions $f_\delta : [1,\infty) \longrightarrow \mathbb{R}_+$ such that

$$
f_\delta(x) = \begin{cases} 1, & \text{if } x \ge 1+2\delta \\ 0, & \text{if } x \le 1+\delta. \end{cases}
$$

Note that, for each $n \in \mathbb{N}$, $\int_0^\infty f_\delta(\hat{X}_s^{(n)})d\hat{\eta}_s^{(n)} = 0$ and thus, for each $\delta > 0$,

$$
0 = \lim_{k\to\infty}\tilde{E}^{(n_k)}\left(\int_0^\infty f_\delta(\pi_s^{(1)})d\pi_s^{(3)} \wedge 1\right) = \tilde{E}\left(\int_0^\infty f_\delta(\pi_s^{(1)})d\pi_s^{(3)} \wedge 1\right).
$$

Consequently, for each $\delta > 0$, $\int_0^\infty 1_{[1+2\delta,\infty)}(\pi_s^{(1)})d\pi_s^{(3)} = 0$, almost surely w.r.t. $\tilde{P}$. The property in (iii) now follows on sending $\delta \to 0$. Finally, we consider part (iv). It suffices to show that for every $0 \le s \le t < \infty$

$$
\tilde{E}\Bigg(\psi(\cdot)\Bigg(\phi(\pi_t^{(1)},\pi_t^{(2)}) - \phi(\pi_s^{(1)},\pi_s^{(2)}) - \int_s^t \mathcal{L}\phi(\pi_u^{(1)},\pi_u^{(2)})du - \int_s^t \frac{\partial\phi}{\partial x}(1,\pi_u^{(2)})d\pi_u^{(3)}\Bigg)\Bigg) = 0,
$$

where $\psi : \tilde{\Omega} \to \mathbb{R}$ is an arbitrary bounded, continuous, $\mathcal{F}_s$ measurable map. Now fix such

s, t, and ψ. Then by weak convergence of $\tilde{P}^{(n_k)}$ to $\tilde{P}$,

$$\begin{aligned}\lim_{k\to\infty} \tilde{E}^{(n_k)}\big(\psi(\cdot)\big(\phi(\pi_t^{(1)}, \pi_t^{(2)}) - \phi(\pi_s^{(1)}, \pi_s^{(2)}) - \int_s^t \mathcal{L}\phi(\pi_u^{(1)}, \pi_u^{(2)})du - \int_s^t \frac{\partial \phi}{\partial x}(1, \pi_u^{(2)})d\pi_u^{(3)}\big)\big) \\ = \tilde{E}\big(\psi(\cdot)\big(\phi(\pi_t^{(1)}, \pi_t^{(2)}) - \phi(\pi_s^{(1)}, \pi_s^{(2)}) - \int_s^t \mathcal{L}\phi(\pi_u^{(1)}, \pi_u^{(2)})du - \int_s^t \frac{\partial \phi}{\partial x}(1, \pi_u^{(2)})d\pi_u^{(3)}\big)\big).\end{aligned}$$

To complete the proof, it suffices to show that the limit on the left side above is 0. In view of the martingale property in (3.3.26), to show this, it suffices to prove that for $\phi \in C_c^\infty(\mathbb{R}_{\geq 1} \times \mathbb{R}_+)$

$$\lim_{n\to\infty} E\left|\int_0^t \left(\mathcal{L}^{(n)}\phi(\hat{X}_s^{(n)}, \hat{Y}_s^{(n)}) - \mathcal{L}\phi(\hat{X}_s^{(n)}, \hat{Y}_s^{(n)})\right) ds\right| = 0 \tag{3.3.28}$$

and

$$\lim_{n\to\infty} E\left|\int_0^t \left(D^{(n)}\phi(\hat{Y}_s^{(n)}) - \frac{\partial \phi}{\partial x}(1, \hat{Y}_s^{(n)})\right) d\hat{\eta}_s^{(n)}\right| = 0.$$

The latter is immediate upon using the smoothness of ϕ and (3.3.18). For (3.3.28), we rewrite $\mathcal{L}^{(n)}\phi$ using a Taylor expansion as follows:

$$\begin{aligned}&\mathcal{L}^{(n)}\phi(x, y) \\ &= \lambda_1^{(n)} n^2 x \sum_{k=0}^{\infty} \left[\frac{k-1}{n}\frac{\partial}{\partial x}\phi(x, y) + \frac{1}{2}\left(\frac{k-1}{n}\right)^2 \frac{\partial^2}{\partial x^2}\phi(x, y)\right] \mu_1^{(n)}(k) \\ &\quad + \lambda_2^{(n)} n^2 xy \sum_{k=0}^{\infty} \left[\frac{k-1}{n}\frac{\partial}{\partial y}\phi(x, y) + \frac{1}{2}\left(\frac{k-1}{n}\right)^2 \frac{\partial^2}{\partial y^2}\phi(x, y)\right] \mu_2^{(n)}(k) + R^{(n)}(x, y) \\ &= c_1^{(n)}\lambda_1^{(n)} x \frac{\partial}{\partial x}\phi(x, y) + \frac{1}{2}\alpha_1^{(n)}\lambda_1^{(n)} x \frac{\partial^2}{\partial x^2}\phi(x, y) \\ &\quad + c_2^{(n)}\lambda_2^{(n)} xy \frac{\partial}{\partial y}\phi(x, y) + \frac{1}{2}\alpha_2^{(n)}\lambda_2^{(n)} xy \frac{\partial^2}{\partial y^2}\phi(x, y) + R^{(n)}(x, y),\end{aligned}$$

where the term $R^{(n)}(x, y)$ is a remainder term, which, using part (iii) of Condition 3.1.1, is seen to converge to 0, as $n \to \infty$. Furthermore, using the compact support property of ϕ, it follows that $\lim_{n\to\infty} E\int_0^t |R^{(n)}(\hat{X}_s^{(n)}, \hat{Y}_s^{(n)})|ds = 0$. Next note that

$$\begin{aligned}
&\mathcal{L}^{(n)}\phi(x,y) - R^{(n)}(x,y) - \mathcal{L}\phi(x,y)\\
&= (\lambda_1^{(n)}c_1^{(n)} - \lambda_1 c_1)x\frac{\partial}{\partial x}\phi(x,y) + \frac{1}{2}(\lambda_1^{(n)}\alpha_1^{(n)} - \lambda_1\alpha_1)x\frac{\partial^2}{\partial x^2}\phi(x,y)\\
&\quad + (\lambda_2^{(n)}c_2^{(n)} - \lambda_2 c_2)xy\frac{\partial}{\partial y}\phi(x,y) + \frac{1}{2}(\lambda_2^{(n)}\alpha_2^{(n)} - \lambda_2\alpha_2)xy\frac{\partial^2}{\partial y^2}\phi(x,y),
\end{aligned}$$

which, in view of Condition 3.1.1, converges to 0, as $n \to \infty$. Once more using the compact support property of ϕ, it follows that

$$\lim_{n\to\infty} E\int_0^t \left|\mathcal{L}^{(n)}\phi(\hat{X}_s^{(n)}, \hat{Y}_s^{(n)}) - R^{(n)}(\hat{X}_s^{(n)}, \hat{Y}_s^{(n)}) - \mathcal{L}\phi(\hat{X}_s^{(n)}, \hat{Y}_s^{(n)})\right| ds = 0.$$

Combining the above estimates, we have (3.3.28), and the result follows. □

Chapter 4

Stochastic Averaging Under Fast Catalyst Dynamics

The catalyst and reactant populations considered in the last chapter evolve on a comparable time scale in the sense that the branching rates of both $X^{(n)}$ and $Y^{(n)}$ converge to positive (possibly different) constants, as $n \to \infty$. In situations in which the catalyst evolves "much faster" than the reactant, in a sense made precise below, it is of interest to find simplified diffusion models that capture the parts of the dynamics one is interested in economically. Such model reductions (see [21] and references therein for the setting of chemical reaction networks) not only help in better understanding the dynamics of the system but also help to reduce computational costs in simulations. When the catalyst population evolves much faster than the reactant population, we expect to obtain a reduced diffusion model, in which the influence of the catalyst on the reactant is only through the catalyst's stationary distribution. In our work, we consider a simplified setting of catalyst and reactant populations that evolve according to the (reflected) diffusions X and Y from Proposition 3.1.1, but where the evolution of the catalyst is accelerated by a factor of n (i.e. drift and diffusion coefficients depend on n). In Section 4.1, we establish a scaling limit theorem, as $n \to \infty$, in which the reactant process is asymptotically described through the solution of the one dimensional SDE given in (4.0.2). This result shows that the reactant evolution, which is given through a coupled two dimensional system, can be well approximated by a one dimensional SDE with coefficients depending on the stationary distribution of the catalyst. Averaging results of a similar form in the more realistic setting where the catalyst and reactant populations are described through branching processes will be a topic for future research. In Section 4.2 we take one key step towards such a research program, which is to establish the convergence of the stationary distribution of the scaled catalyst process $\hat{X}^{(n)}$ to that of the limit reflecting diffusion X.

We begin by describing the unique stationary distribution of X, where X is the reflected diffusion from Proposition 3.1.1, approximating the catalyst dynamics (Theorem 3.1.2).

Proposition 4.0.1. *The process X introduced in Proposition 3.1.1 has a unique stationary distribution, ν_1, which has density*

$$p(x) := \begin{cases} \frac{\theta}{x}\exp(2\frac{c_1}{\alpha_1}x), & \text{if } x \geq 1 \\ 0, & \text{if } x < 1, \end{cases} \tag{4.0.1}$$

where $\theta := \left(\int_1^\infty (\frac{1}{x}\exp(2\frac{c_1}{\alpha_1}x))dx\right)^{-1}$ is some normalizing constant.

Proposition 4.0.1 will be proved in Section 4.3. Let $m_X = \int_1^\infty x\nu_1(dx)$ and $\check{Y}$ be the solution of

$$\check{Y}_t = \check{Y}_0 + \int_0^t c_2\lambda_2 m_X \check{Y}_s ds + \int_0^t \sqrt{\alpha_2\lambda_2 m_X \check{Y}_s} dB_s, \tag{4.0.2}$$

where $\check{Y}_0 = 1$. The above equation will approximate the dynamics of the reactant population under appropriate conditions, as described in Section 4.1.

4.1 Fast Catalyst Diffusion

Consider a system of catalyst and reactant populations that can be represented by a system of SDEs similar to that in Proposition 3.1.1, but with a faster evolving catalyst population. More precisely, consider catalyst and reactant populations $\check{X}^{(n)}$ and $\check{Y}^{(n)}$ of the form

$$\begin{aligned}
\check{X}_t^{(n)} &= \Gamma\left(\check{X}_0^{(n)} + \int_0^{\cdot} nc_1\lambda_1\check{X}_s^{(n)}ds + \int_0^{\cdot}\sqrt{n\alpha_1\lambda_1\check{X}_s^{(n)}}dB_s^X\right)(t)\\
\check{Y}_t^{(n)} &= \check{Y}_0^{(n)} + \int_0^t c_2\lambda_2\check{X}_s^{(n)}\check{Y}_s^{(n)}ds + \int_0^t\sqrt{\alpha_2\lambda_2\check{X}_s^{(n)}\check{Y}_s^{(n)}}dB_s^Y,
\end{aligned}$$

where $\check{X}_0^{(n)} = \check{Y}_0^{(n)} = 1$, $c_i \in (-\infty, 0)$, $\alpha_i, \lambda_i \in (0,\infty)$, B^X and B^Y are independent standard Brownian motions, and Γ is the Skorohod map described above Proposition 3.1.1.

The following result shows that the reactant population process $\check{Y}^{(n)}$ can be well approximated by the one dimensional diffusion $\check{Y}$ in (4.0.2).

Theorem 4.1.1. *The process $\check{Y}^{(n)}$ converges weakly in $C(\mathbb{R}_+ : \mathbb{R}_+)$ to the process $\check{Y}$.*

Theorem 4.1.1 will be proved in Section 4.4.

4.2 Convergence of Invariant Distributions

Towards a more accurate catalyst-reactant dynamical model, we next consider a setting in which the catalyst and reactant populations evolve according to branching processes of the form described in Section 3.1. In particular the catalyst evolution is given by the branching process $X^{(n)}$ introduced in Section 3.1. As a first step towards developing a general averaging theory, we present a result in this section which shows that under suitable conditions the stationary distribution of the scaled branching process $\hat{X}^{(n)}$ converges weakly to the unique stationary distribution ν_1 of X characterized in Proposition

4.0.1. In future work, we will combine this limit theorem with averaging techniques of the form used in the proof of Theorem 4.1.1 to establish weak convergence of suitably scaled reactant branching processes to the solution of the one dimensional SDE given in (4.0.2).

Our main assumption, in addition to Condition 3.1.1, will be that

$$\sup_{n\in\mathbb{N}}\sum_{k=0}^{\infty} e^{\bar{\delta}k}\mu_1^{(n)}(k) < \infty \tag{4.2.1}$$

for some $\bar{\delta} > 0$.

Theorem 4.2.1. *Suppose Condition 3.1.1 and* (4.2.1) *hold. Then, for each $n \in \mathbb{N}$, the process $\hat{X}^{(n)}$ has a unique stationary distribution $\nu_1^{(n)}$, and the family $\{\nu_1^{(n)}\}_{n\in\mathbb{N}}$ is tight.*

Recall from Proposition 4.0.1 that the reflected diffusion X from Proposition 3.1.1 has a unique stationary distribution ν_1.

Theorem 4.2.2. *As $n \to \infty$, $\nu_1^{(n)}$ converges weakly to ν_1.*

Proof of Theorem 4.2.2. Theorem 4.2.1 implies that every subsequence of $\nu_1^{(n)}$ has a convergent subsequence. Call such a limit ν_1^*. The weak convergence of $\hat{X}^{(n)}$ to X (Theorem 3.1.2) and the stationarity of $\nu_1^{(n)}$ imply that ν_1^* is a stationary distribution of X. Since the stationary distribution of X is unique, we have $\nu_1^* = \nu_1$, which completes the proof. $\square$

4.3 Proof of Proposition 4.0.1

Uniqueness of the invariant measure of X is an immediate consequence of the non-degeneracy of the diffusion coefficient (note that $\alpha_2\lambda_2 x \geq \alpha_2\lambda_2 > 0$). For existence, we will apply the Echeverria-Weiss-Kurtz criterion ([3, 22, 41]). This criterion, in the current context, says that in order to establish that a probability measure $\bar{\nu}_1$ is an invariant measure for X, it suffices to verify that for some $C \geq 0$ and all $\phi \in C_c^\infty(\mathbb{R}_{\geq 1})$

$$\int_{[1,\infty)} \mathcal{L}\phi(x)\bar{\nu}_1(dx) + C\alpha_1\lambda_1\phi'(1) = 0, \tag{4.3.1}$$

where $\mathcal{L}\phi(x) = c_1\lambda_1 x\phi'(x) + \frac{1}{2}\alpha_1\lambda_1 x\phi''(x)$. We will check that (4.3.1) holds with $\bar{\nu}_1 = \nu_1$ and $C = \frac{p(1)}{2}$. Note that for $\phi \in C_c^\infty(\mathbb{R}_{\geq 1})$ and p as in (4.0.1)

$$\begin{aligned}
&\int_1^\infty \left(c_1\lambda_1 x\phi'(x) + \frac{1}{2}\alpha_1\lambda_1 x\phi''(x)\right) p(x)dx \\
&\quad = c_1\lambda_1\theta e^{2\frac{c_1}{\alpha_1}x}\phi(x)\Big|_1^\infty - \int_1^\infty 2c_1\lambda_1\theta\frac{c_1}{\alpha_1}e^{2\frac{c_1}{\alpha_1}x}\phi(x)dx \\
&\qquad + \frac{1}{2}\alpha_1\lambda_1\theta e^{2\frac{c_1}{\alpha_1}x}\phi'(x)\Big|_1^\infty - \int_1^\infty \alpha_1\lambda_1\theta\frac{c_1}{\alpha_1}e^{2\frac{c_1}{\alpha_1}x}\phi'(x)dx \\
&\quad = -\frac{1}{2}\alpha_1\lambda_1\theta e^{2\frac{c_1}{\alpha_1}}\phi'(1) = -\frac{p(1)}{2}\alpha_1\lambda_1\phi'(1),
\end{aligned}$$

where the second to last equality can be seen on noting that

$$\int_1^\infty \alpha_1\lambda_1\theta\frac{c_1}{\alpha_1}e^{2\frac{c_1}{\alpha_1}x}\phi'(x)dx = c_1\lambda_1\theta e^{2\frac{c_1}{\alpha_1}x}\phi(x)\Big|_1^\infty - 2\int_0^\infty \lambda_1\theta\frac{c_1^2}{\alpha_1}e^{2\frac{c_1}{\alpha_1}x}\phi(x)dx.$$

Thus, (4.3.1) holds with $C = \frac{p(1)}{2}$ and $\bar{\nu}_1 = \nu_1$. The result follows.

4.4 Proof of Theorem 4.1.1

In order to prove the result, we will verify that the assumptions of Theorem II.1 (more precisely, those in the remark following Theorem II.1) in [38], pp. 78-79, hold. For this, it suffices to show that for any $k \in \mathbb{N}$, $\Phi \in C_c^2(\mathbb{R}_+^k)$, $\phi \in C_c^2(\mathbb{R}_+)$, $0 \leq t_1 < t_2 < \cdots < t_{k+1} < T < \infty$, and for some sequence h_n with $\lim_{n\to\infty} h_n = 0$,

$$\sup_{t\in[t_{k+1},T]} E\left|\Phi\left(\check{Y}_{t_1}^{(n)}, \ldots, \check{Y}_{t_k}^{(n)}\right)\left(\phi\left(\check{Y}_{t+h_n}^{(n)}\right) - \phi\left(\check{Y}_t^{(n)}\right) - h_n\check{L}\phi\left(\check{Y}_t^{(n)}\right)\right)\right| = o(h_n),$$

where $\check{\mathcal{L}}$ is given as

$$\check{\mathcal{L}}\phi(y) := c_2\lambda_2 m_X y\phi'(y) + \frac{1}{2}\alpha_2\lambda_2 m_X y\phi''(y), \quad \phi \in C_c^\infty(\mathbb{R}_+).$$

Letting $X_t^\circ := \check{X}_{t/n}^{(n)}$, $t \geq 0$, we see, using scaling properties of the Skorohod map and elementary martingale characterization properties, that X° has the same probability law as the process X that was introduced in Proposition 3.1.1. The following uniform moment bound will be used in the proof of Theorem 4.1.1.

Lemma 4.4.1. *Let X be as in Proposition 3.1.1. Then for some $\delta_0 \in (0, \infty)$*

$$\sup_{0 \leq t < \infty} E\left(e^{\delta_0 X_t}\right) =: d(\delta_0) < \infty.$$

Proof of Lemma 4.4.1. Throughout Chapter 3 and until now in this chapter, we took the initial value of X to be 1. Here, it will be convenient to allow the initial random variable X_0 to have an arbitrary distribution on $\mathbb{R}_{\geq 1}$. When X_0 has distribution μ on $\mathbb{R}_{\geq 1}$, we will denote the corresponding probability and expectation operator by P_μ and E_μ, respectively. If $\mu = \delta_x$ for some $x \in \mathbb{R}_{\geq 1}$, we will instead write P_x and E_x, respectively.

We begin by establishing exponential moment estimates for the increase of X over time intervals of length $l\rho$ when the process is away from the boundary 1, where $\rho > 0$ and $l \geq 1$. Fix $\rho \in (0, \infty)$. Let $a := c_1\lambda_1$, $b := \alpha_1\lambda_1$, and $\delta \in (0, -\frac{a}{b} \wedge 1)$. Define $\sigma_r := \inf\{t \in [0, \infty) | \int_0^t X_s ds > r\}$ and $\rho_{l,r} := l\rho \wedge \sigma_r$. Then

$$\begin{aligned}
&E_x\left(\exp\left(\delta a \int_0^{\rho_{l,r}} X_s ds + \delta\sqrt{b}\int_0^{\rho_{l,r}} \sqrt{X_s} dB_s^X\right)\right) \\
&\quad = E_x\left(\exp\left((\delta a + \delta^2 b)\int_0^{\rho_{l,r}} X_s ds + \delta\sqrt{b}\int_0^{\rho_{l,r}} \sqrt{X_s} dB_s^X - \delta^2 b\int_0^{\rho_{l,r}} X_s ds\right)\right) \\
&\quad \leq \left(E_x e^{2\rho_{l,r}(\delta a + \delta^2 b)}\right)^{\frac{1}{2}} \left(E_x \exp\left(2\delta\sqrt{b}\int_0^{\rho_{l,r}} \sqrt{X_s} dB_s^X - \frac{(2\delta\sqrt{b})^2}{2}\int_0^{\rho_{l,r}} X_s ds\right)\right)^{\frac{1}{2}},
\end{aligned}$$

where the inequality follows on noting that $\rho_{l,r} \leq \int_0^{\rho_{l,r}} X_s ds$ and $\delta \in (0, -\frac{a}{b})$. Using the martingale property of the stochastic exponential (see e.g. [32], Theorem III.45, p. 141), we have that the second term on the right hand side of the last display equals 1. Thus, sending $r \to \infty$, we have, with $-\theta := \delta a + \delta^2 b < 0$,

$$E_x \left(\exp \left(\delta a \int_0^{l\rho} X_s ds + \delta \sqrt{b} \int_0^{l\rho} \sqrt{X_s} dB_s^X \right) \right) \leq e^{l\rho(\delta a + \delta^2 b)} = e^{-\theta l \rho}. \tag{4.4.1}$$

Next, for $x \in (1, \infty)$, we have by application of Itô's formula that for $t \leq \rho$ and $\tilde{\delta} \leq \delta$

$$E_x \left(e^{\tilde{\delta} X_t} \right) \leq e^{\tilde{\delta} x} + \tilde{\delta} e^{\tilde{\delta}} E_x \eta_\rho \equiv e^{\tilde{\delta} x} + x C_1(\rho), \tag{4.4.2}$$

where $C_1(\rho) \in (0, \infty)$ and the last equivalence can be checked by an application of Gronwall's inequality and the Lipschitz property of the Skorohod map.

Using the above estimates, we will now establish certain uniform estimates on the tail behavior of $X_{k\rho}$, which will lead to exponential moment estimates at these time points. Let

$$\tau_j := \inf\{t \geq (j-1)\rho | X_t \leq L\} \wedge j\rho, \quad j \geq 1,$$

and

$$e_j := X_{j\rho} - X_{\tau_j}, \quad j \geq 1,$$

$e_0 = 0$. Fix $k \in \mathbb{N}$ and $L > 1$, and let

$$M := \max \left\{ j = 1, \ldots, k \middle| \inf_{(j-1)\rho \leq s \leq j\rho} X_s \leq L \right\}$$

if there is an $s \in [0, k\rho]$ such that $X_s \le L$ and 0 otherwise. Let

$$v_j := \int_{(j-1)\rho}^{j\rho} aX_s ds + \int_{(j-1)\rho}^{j\rho} \sqrt{bX_s} dB_s^X, \quad j \ge 1.$$

If $M = 0$ then $X_{k\rho} = X_0 + \sum_{j=1}^k v_j$, and if $M > 0$ then $X_{k\rho} = X_{M\rho} + \sum_{j=M+1}^k v_j$. In both cases, we have, using (4.4.1) and with $\zeta_i := e_i + \sum_{j=i+1}^k v_j$,

$$\begin{aligned}
P_x(X_{k\rho} > K) &\le P_x\left(X_{M\rho} + \sum_{j=M+1}^k v_j > K - x\right) \le P_x\left(\max_{0\le i\le k} \zeta_i > K - L - x\right) \\
&\le \sum_{i=0}^k P_x(\zeta_i > K - L - x) \le \sum_{i=0}^k E_x\left(e^{\frac{\delta}{2}\zeta_i}\right) e^{-\frac{\delta}{2}(K-L-x)} \\
&\le \sum_{i=0}^k \left(E_x e^{\delta e_i}\right)^{\frac{1}{2}} \left(E_x e^{\delta \sum_{j=i+1}^k v_j}\right)^{\frac{1}{2}} e^{-\frac{\delta}{2}(K-L-x)} \le \sum_{i=0}^k \left(E_x e^{\delta e_i}\right)^{\frac{1}{2}} e^{-\frac{1}{2}(k-i)\theta\rho} e^{-\frac{\delta}{2}(K-L-x)}.
\end{aligned}$$

Next note that

$$\begin{aligned}
E_x e^{\delta e_i} &\le E_x\left(e^{\delta[X_{i\rho} - X_{\tau_i}]} 1_{\tau_i < i\rho}\right) + 1 = E_x\left(e^{-\delta X_{\tau_i}} 1_{\tau_i < i\rho} E_{X_{\tau_i}}\left(e^{\delta X_{i\rho}}\right)\right) + 1 \\
&\le E_x\left(e^{\delta X_{\tau_i}} + X_{\tau_i} C_1(\rho)\right) + 1 \equiv C_2(L, \rho) + 1.
\end{aligned}$$

Hence,

$$P_x(X_{k\rho} > K) \le (C_2(L, \rho) + 1)^{\frac{1}{2}} e^{-\frac{\delta}{2}(K-L-x)} \sum_{l=0}^k e^{-\frac{1}{2}l\rho\theta} \le (C_2(L, \rho) + 1)^{\frac{1}{2}} \frac{e^{-\frac{\delta}{2}(K-L-x)}}{1 - e^{-\frac{1}{2}\rho\theta}}.$$

The last estimate yields

$$\sup_{k\in\mathbb{N}_0} E_x(e^{\frac{\delta}{4}X_{k\rho}}) \le Ce^{\frac{\delta}{2}x}, \tag{4.4.3}$$

for some $C \in (0, \infty)$. Let $\delta_0 := \frac{\delta}{4}$. Finally, for $t \in ((k-1)\rho, k\rho]$, $k \geq 1$,

$$
\begin{aligned}
E_x\left(e^{\delta_0 X_t}\right) = E_x\left(E_{X_{(k-1)\rho}}\left(e^{\delta_0 X_t}\right)\right) &\leq E_x\left(e^{\delta_0 X_{(k-1)\rho}} + X_{(k-1)\rho} C_1(\rho)\right) \\
&\leq C e^{\frac{\delta}{2} x}\left(1 + \frac{1}{\delta_0} C_1(\rho)\right).
\end{aligned}
$$

The result follows.

$\square$

Remark 4.4.1. Note that Lemma 4.4.1 and the scaling property noted above that lemma say that $\sup_{n\in\mathbb{N}} \sup_{0\leq t<\infty} E\left(e^{\delta_0 \check{X}_t^{(n)}}\right) < \infty$.

Let, for $\phi \in C_c^\infty(\mathbb{R}_+)$

$$
\mathcal{L}_x \phi(y) := c_2 \lambda_2 x y \phi'(y) + \frac{1}{2} \alpha_2 \lambda_2 x y \phi''(y), \quad (x, y) \in \mathbb{R}_{\geq 1} \times \mathbb{R}_+.
$$

Then

$$
\begin{aligned}
&E\left[\Phi\left(\check{Y}_{t_1}^{(n)}, \ldots, \check{Y}_{t_k}^{(n)}\right)\left(\phi(\check{Y}_{t+h_n}^{(n)}) - \phi(\check{Y}_t^{(n)})\right)\right] \\
&= E\left[\Phi\left(\check{Y}_{t_1}^{(n)}, \ldots, \check{Y}_{t_k}^{(n)}\right) \int_t^{t+h_n} \mathcal{L}_{\check{X}_s^{(n)}} \phi(\check{Y}_t^{(n)}) ds\right] \\
&\quad + E\Bigg[\Phi\left(\check{Y}_{t_1}^{(n)}, \ldots, \check{Y}_{t_k}^{(n)}\right) \int_t^{t+h_n} \Big[c_2 \lambda_2 \check{X}_s^{(n)}\left(\check{Y}_s^{(n)} \phi'(\check{Y}_s^{(n)}) - \check{Y}_t^{(n)} \phi'(\check{Y}_t^{(n)})\right) \\
&\qquad + \frac{1}{2} \alpha_2 \lambda_2 \check{X}_s^{(n)}\left(\check{Y}_s^{(n)} \phi''(\check{Y}_s^{(n)}) - \check{Y}_t^{(n)} \phi''(\check{Y}_t^{(n)})\right)\Big] ds\Bigg].
\end{aligned} \tag{4.4.4}
$$

For the second term, we have

$$
\begin{aligned}
\sup_{t\in[t_{k+1},T]} E\Bigg|&\Phi\left(\check{Y}_{t_1}^{(n)}, \ldots, \check{Y}_{t_k}^{(n)}\right) \int_t^{t+h_n} \Bigg(c_2 \lambda_2 \check{X}_s^{(n)}\left(\check{Y}_s^{(n)} \phi'(\check{Y}_s^{(n)}) - \check{Y}_t^{(n)} \phi'(\check{Y}_t^{(n)})\right) \\
&+ \frac{1}{2} \alpha_2 \lambda_2 \check{X}_s^{(n)}\left(\check{Y}_s^{(n)} \phi''(\check{Y}_s^{(n)}) - \check{Y}_t^{(n)} \phi''(\check{Y}_t^{(n)})\right)\Bigg) ds\Bigg| = o(h_n)
\end{aligned} \tag{4.4.5}
$$

since the functions Φ, ϕ, and the derivatives of ϕ are continuous with bounded support and, by Remark 4.4.1, $\sup_{n\in\mathbb{N}}\sup_{s\geq 0} E(\check{X}_s^{(n)}) < \infty$. Recalling the definition of X° above Lemma 4.4.1, the first expected value on the right hand side in (4.4.4) equals

$$h_n E\left[\Phi\left(\check{Y}_{t_1}^{(n)}, \ldots, \check{Y}_{t_k}^{(n)}\right) \frac{1}{h_n n} \int_{tn}^{tn+h_n n} \mathcal{L}_{X_s^\circ} \phi(\check{Y}_t^{(n)}) ds\right].$$

Thus

$$\begin{aligned}
&E\left[\Phi\left(\check{Y}_{t_1}^{(n)}, \ldots, \check{Y}_{t_k}^{(n)}\right)\left(\phi(\check{Y}_{t+h_n}^{(n)}) - \phi(\check{Y}_t^{(n)}) - h_n \check{\mathcal{L}}\phi(\check{Y}_t^{(n)})\right)\right] \\
&\quad = E\left[\Phi\left(\check{Y}_{t_1}^{(n)}, \ldots, \check{Y}_{t_k}^{(n)}\right) h_n \left(\frac{1}{h_n n}\int_{tn}^{tn+h_n n} \mathcal{L}_{X_s^\circ}\phi(\check{Y}_t^{(n)}) ds - \check{\mathcal{L}}\phi(\check{Y}_t^{(n)})\right)\right] + o(h_n) \\
&\quad = E\Bigg[\Phi\left(\check{Y}_{t_1}^{(n)}, \ldots, \check{Y}_{t_k}^{(n)}\right) h_n \left(c_2\lambda_2 \check{Y}_t^{(n)} \phi'(\check{Y}_t^{(n)}) + \frac{1}{2}\alpha_2\lambda_2 \check{Y}_t^{(n)} \phi''(\check{Y}_t^{(n)})\right) \\
&\qquad\qquad \left(\frac{1}{h_n n}\int_{tn}^{tn+h_n n} X_s^\circ ds - m_X\right)\Bigg] + o(h_n).
\end{aligned}$$

To complete the proof, it thus remains to show that for some sequence $\{h_n\}$ with $\lim_{n\to\infty} h_n = 0$

$$\lim_{n\to\infty} E\left|\frac{1}{h_n n}\int_{tn}^{tn+h_n n} X_s^\circ ds - m_X\right| = E\left|\frac{1}{h_n n}\int_{tn}^{tn+h_n n} X_s ds - m_X\right| = 0, \qquad (4.4.6)$$

uniformly in $t \in [t_{k+1}, T]$. Note that the first equality is due to the fact that X and X° have the same distribution. Next, the law of large numbers (see e.g. [28], Theorem 17.0.1) yields, for any $t \in [t_{k+1}, T]$,

$$E\left|\frac{1}{tn}\int_0^{tn} X_s ds - m_X\right| \to 0, \quad \text{as } n \to \infty. \qquad (4.4.7)$$

The above result, along with Lemma 4.4.2, below, implies that there is a sequence $\{h_n\}$ such that $\lim_{n\to\infty} h_n = 0$ and (4.4.6) holds, uniformly in $t \in [t_{k+1}, T]$. This completes the proof.

The proof of the following lemma is adapted from Lemma II.9, p. 137, in [38].

Lemma 4.4.2. *Fix $t_{k+1}, T \in (0,\infty)$. If for all $t \in [t_{k+1}, T]$*

$$E\left|\frac{1}{tn}\int_0^{tn} X_s ds - m_X\right| \to 0, \quad \text{as } n \to \infty,$$

then there is a sequence $\{h_n\}$ such that $h_n \to 0$ as $n \to \infty$, and

$$\sup_{t\in[t_{k+1},T]} E\left|\frac{1}{h_n n}\int_{tn}^{tn+h_n n} X_s ds - m_X\right| \to 0, \quad \text{as } n \to \infty.$$

Proof. Let $\alpha(\tau) := \sup_{u>\tau} E\left|\frac{1}{u}\int_0^u X_s ds - m_X\right|$. For $t \in [t_{k+1}, T]$ we have

$$\begin{aligned}
&E\left|\frac{1}{h_n n}\int_{tn}^{tn+h_n n} X_s ds - m_X\right| \\
&\quad = E\left|\frac{tn+h_n n}{h_n n}\frac{1}{tn+h_n n}\int_0^{tn+h_n n} X_s ds - \frac{tn}{h_n n}\frac{1}{tn}\int_0^{tn} X_s ds - m_X\right| \\
&\quad \le \frac{tn+h_n n}{h_n n}\alpha(tn+h_n n) + \frac{tn}{h_n n}\alpha(tn) \le \frac{3T}{h_n}\alpha(t_{k+1}n)
\end{aligned}$$

for all n such that $h_n \le T$. Note that the right hand side of the last display is independent of $t \in [t_{k+1}, T]$. Choosing $h_n = \sqrt{\alpha(t_{k+1}n)}$, the lemma follows. □

4.5 Proof of Theorem 4.2.1

Throughout this section we assume that Condition 3.1.1 and (4.2.1) hold. This will not be explicitly noted in the statements of the results.

Existence of a stationary distribution $\nu_1^{(n)}$ of the $\mathbb{S}_X^{(n)} = \{\frac{l}{n} | l \in \{n, n+1, \dots\}\}$ valued Markov process $\hat{X}^{(n)}$ follows from the tightness of $\hat{X}^{(n)}$, which is a consequence of (4.5.3), below. The uniqueness of the stationary distribution follows from the irreducibility of $\hat{X}^{(n)}$.

In order to establish the tightness of the sequence $\{\nu_1^{(n)}\}_{n\in\mathbb{N}}$, we will use the following uniform in n moment stability estimate for $\hat{X}^{(n)}$.

In order to study tightness and convergence of invariant measures, it will be convenient, as in the proof of Lemma 4.4.1, to allow the initial random variable $\hat{X}_0^{(n)}$ to have an arbitrary distribution on $\mathbb{S}_X^{(n)}$. When $X_0^{(n)}$ has distribution μ on $\mathbb{R}_{\geq 1}$, we will denote the corresponding probability and expectation operator by P_μ and E_μ, respectively. If $\mu = \delta_x$ for some $x \in \mathbb{R}_{\geq 1}$, we will instead write P_x and E_x, respectively. When considering an initial condition x for $\hat{X}^{(n)}$, then x will always be in $\mathbb{S}_X^{(n)}$, although this will frequently be suppressed in the notation.

Theorem 4.5.1. *There is a $t_0 \in \mathbb{R}_+$ such that for all $t \geq t_0$ and $p > 0$*

$$\lim_{x\to\infty} \sup_{n\in\mathbb{N}} \frac{1}{x^p} E_x\left(\left(\hat{X}_{tx}^{(n)}\right)^p\right) = 0. \tag{4.5.1}$$

The proof is based on the following lemmas.

Lemma 4.5.1. *There exist $\delta, \rho \in (0,\infty)$ such that for every $M > 0$*

$$\sup_{n\in\mathbb{N}, x\leq M} E_x\left(\sup_{0\leq t\leq \rho} e^{\delta \hat{X}_t^{(n)}}\right) =: d(\delta,\rho,M) < \infty. \tag{4.5.2}$$

Lemma 4.5.2. *There exist $\delta, \tilde{d}(\delta) \in (0,\infty)$ such that for every $x \in \mathbb{S}_X^{(n)}$, $n \in \mathbb{N}$, and $t \geq 0$*

$$E_x(e^{\delta \hat{X}_t^{(n)}}) \leq \tilde{d}(\delta) e^{\delta x}. \tag{4.5.3}$$

Proof of Lemma 4.5.1. Condition 3.1.1 and the assumption in (4.2.1) imply, using a Taylor series expansion, that there are $\delta_0, d_1, d_2 \in (0,\infty)$ such that for all $\delta \in [0,\delta_0]$ and

$n \in \mathbb{N}$

$$-\delta d_2 \leq \sum_{k=0}^{\infty} n^2 \left[e^{\frac{(k-1)\delta}{n}} - 1 \right] \mu_1^{(n)}(k) \leq -\delta d_1. \tag{4.5.4}$$

Recall that

$$\hat{\eta}_t^{(n)} = n\lambda_1^{(n)} \mu_1^{(n)}(0) \int_0^t 1_{\{\hat{X}_s^{(n)}=1\}} ds.$$

Calculations similar to those leading to (3.3.18) give

$$\sup_{x \leq M} \sup_{n \in \mathbb{N}} E_x \left((\hat{\eta}_\rho^{(n)})^2 \right) = d_3(\rho, M) < \infty \tag{4.5.5}$$

for any $\rho, M > 0$.

For δ_0 as above and $\delta \leq \delta_0$, let

$$\alpha_\delta^{(n)} := ne^\delta \sum_{k=1}^{\infty} \left(e^{\frac{(k-1)\delta}{n}} - 1 \right) \frac{\mu_1^{(n)}(k)}{\mu_1^{(n)}(0)}$$

and

$$\beta_t^{(n),\delta} := n^2 \lambda_1^{(n)} \int_0^t \hat{X}_s^{(n)} \sum_{k=0}^{\infty} \left(\left[e^{\frac{(k-1)\delta}{n}} - 1 \right] \mu_1^{(n)}(k) \right) 1_{\{\hat{X}_s^{(n)}>1\}} ds.$$

Note that, for any $t \geq 0$,

$$-\delta d_2 \lambda_1^{(n)} \int_0^t \hat{X}_s^{(n)} 1_{\{\hat{X}_s^{(n)}>1\}} ds \leq \beta_t^{(n),\delta} \leq -\delta d_1 \lambda_1^{(n)} \int_0^t \hat{X}_s^{(n)} 1_{\{\hat{X}_s^{(n)}>1\}} ds \tag{4.5.6}$$

and

$$0 \leq \alpha_\delta^{(n)} \leq e^\delta \delta. \tag{4.5.7}$$

We will now show that

$$M_t^{(n),\delta} := \exp\left(\delta \hat{X}_t^{(n)} - \beta_t^{(n),\delta}\right) - \alpha_\delta^{(n)} \hat{\eta}_t^{(n)}$$

is a martingale. Let

$$q(x) := \frac{\hat{\mathcal{L}}^{(n)} f(x)}{f(x)} 1_{\{x>1\}}, \quad \text{where} \quad f(x) = e^{\delta x},$$

and $\hat{\mathcal{L}}^{(n)}$ is the generator of $\hat{X}^{(n)}$, that is

$$\hat{\mathcal{L}}^{(n)} f(x) = \lambda_1^{(n)} n^2 x \sum_{k=0}^{\infty} \left[f\left(\left(x + \frac{k-1}{n} - 1 \right)^+ + 1 \right) - f(x) \right] \mu_1^{(n)}(k).$$

Moreover, let

$$V_t^{(n)} := \left(\hat{X}_t^{(n)}, \exp\left(- \int_0^t q(\hat{X}_s^{(n)}) ds \right) \right).$$

Then for functions of the form $f(x)g(y)$ with $g(y) = y$, the generator $\mathcal{L}^{(n)}$ of $V^{(n)}$ is of the form

$$\mathcal{L}^{(n)}(f(x)g(y)) = \hat{\mathcal{L}}^{(n)} f(x) - q(x) f(x) = \hat{\mathcal{L}}^{(n)} f(x) 1_{\{x=1\}}.$$

Thus

$$\begin{aligned}
M_t^{(n),\delta} &= (fg)(V_t^{(n)}) - \int_0^t \hat{\mathcal{L}}^{(n)} (fg)(V_s^{(n)}) ds \\
&= e^{\delta \hat{X}_t^{(n)}} \exp\left(- \int_0^t \frac{\hat{\mathcal{L}}^{(n)} f(\hat{X}_s^{(n)})}{f(\hat{X}_s^{(n)})} 1_{\{\hat{X}_s^{(n)} > 1\}} ds \right) - \int_0^t \hat{\mathcal{L}}^{(n)} f(1) 1_{\{\hat{X}_s^{(n)} = 1\}} ds \\
&= e^{\delta \hat{X}_t^{(n)} - \beta_t^{(n),\delta}} - \alpha_\delta^{(n)} \hat{\eta}_t^{(n)}
\end{aligned}$$

is a local martingale ([37], pp. 65-66).

We next show that for every $\delta \leq \delta_0$

$$d_4(\delta, \rho, M) := \sup_{x \leq M} \sup_{n \in \mathbb{N}} E_x \left(e^{\delta \hat{X}_\rho^{(n)}} \right) < \infty. \tag{4.5.8}$$

From (4.5.6), $\beta_\rho^{(n),\delta} \leq 0$ and consequently

$$\begin{aligned} e^{\delta \hat{X}_\rho^{(n)}} &= \left(e^{\delta \hat{X}_\rho^{(n)} - \beta_\rho^{(n),\delta}} - \alpha_\delta^{(n)} \hat{\eta}_\rho^{(n)} + \alpha_\delta^{(n)} \hat{\eta}_\rho^{(n)} \right) e^{\beta_\rho^{(n),\delta}} \\ &= \left(M_\rho^{(n),\delta} + \alpha_\delta^{(n)} \hat{\eta}_\rho^{(n)} \right) e^{\beta_\rho^{(n),\delta}} \leq M_\rho^{(n),\delta} + \alpha_\delta^{(n)} \hat{\eta}_\rho^{(n)}. \end{aligned}$$

Thus, using (4.5.5), (4.5.7), and that $M^{(n),\delta}$ is a local martingale, we get

$$E_x \left(e^{\delta \hat{X}_\rho^{(n)}} \right) \leq e^{\delta x} + e^\delta \delta [d_3(\rho, M)]^{\frac{1}{2}}.$$

This proves (4.5.8).

Next, using (4.5.5) once more, for $x \leq M$ and $\delta \leq \frac{\delta_0}{4}$,

$$\begin{aligned} &E_x \left(\sup_{0 \leq t \leq \rho} e^{\delta \hat{X}_t^{(n)}} \right) \leq E_x \left(\sup_{0 \leq t \leq \rho} \left(M_t^{(n),\delta} + \alpha_\delta^{(n)} \hat{\eta}_t^{(n)} \right) \right) \\ &\quad \leq E_x \left(\sup_{0 \leq t \leq \rho} M_t^{(n),\delta} \right) + e^\delta \delta \sqrt{d_3(\rho, M)} \leq 2 \left(E_x \left(\left(M_\rho^{(n),\delta} \right)^2 \right) \right)^{\frac{1}{2}} + e^\delta \delta \sqrt{d_3(\rho, M)}. \end{aligned}$$

Moreover,

$$E_x \left((M_\rho^{(n),\delta})^2 \right) \leq 2 E_x \left(e^{2\delta \hat{X}_\rho^{(n)} - 2\beta_\rho^{(n),\delta}} \right) + 2 \left(e^\delta \delta \right)^2 d_3(\rho, M),$$

and from (4.5.6) and (4.5.8), we have for $x \leq M$

$$\begin{aligned} E_x \left(e^{2\delta \hat{X}_\rho^{(n)} - 2\beta_\rho^{(n),\delta}} \right) &\leq \left(E_x \left(e^{4\delta \hat{X}_\rho^{(n)}} \right) \right)^{\frac{1}{2}} \left(E_x \left(e^{-4\beta_\rho^{(n),\delta}} \right) \right)^{\frac{1}{2}} \\ &\leq \left(d_4(4\delta, \rho, M) \right)^{\frac{1}{2}} E_x \left(\exp \left(4\delta d_2 \lambda_1^{(n)} \rho \sup_{0 \leq t \leq \rho} \hat{X}_t^{(n)} \right) \right). \end{aligned}$$

Choose $\rho < \left(8d_2 \sup_{n\in\mathbb{N}} \lambda_1^{(n)}\right)^{-1}$, then there is a $d(\delta,\rho,M)$ such that for $x \leq M$

$$E_x\left(\sup_{0\leq t\leq\rho} e^{\delta \hat{X}_t^{(n)}}\right) \leq d_5(\delta,\rho,M) E_x\left(\exp\left(4\delta d_2 \lambda_1^{(n)} \rho \sup_{0\leq t\leq\rho} \hat{X}_t^{(n)}\right)\right)$$
$$\leq d_5(\delta,\rho,M) E_x\left(\exp\left(\frac{\delta}{2} \sup_{0\leq t\leq\rho} \hat{X}_t^{(n)}\right)\right) \leq d_5(\delta,\rho,M)\left[E_x\left(\sup_{0\leq t\leq\rho} e^{\delta \hat{X}_t^{(n)}}\right)\right]^{\frac{1}{2}}.$$

Dividing both sides by $\left[E_x\left(\sup_{0\leq t\leq\rho} e^{\delta \hat{X}_t^{(n)}}\right)\right]^{\frac{1}{2}}$ yields

$$\left[E_x\left(\sup_{0\leq t\leq\rho} e^{\delta \hat{X}_t^{(n)}}\right)\right]^{\frac{1}{2}} \leq d_5(\delta,\rho,M)$$

for any $x \leq M$ and $n \in \mathbb{N}$. The result follows. □

Proof of Lemma 4.5.2. For $\delta \in (0,1), n \in \mathbb{N}$, define

$$b_\delta^{(n),1}(x) := \lambda_1^{(n)} n^2 x \sum_{k=0}^{\infty}\left(e^{\delta\frac{k-1}{n}} - 1\right)\mu_1^{(n)}(k),$$
$$b_\delta^{(n),2}(x) := \lambda_1^{(n)} n^2 x \sum_{k=1}^{\infty}\left(e^{\delta\frac{k-1}{n}} - 1\right)\mu_1^{(n)}(k), \quad \text{and}$$
$$b_\delta^{(n)}(x) := b_\delta^{(n),1}(x) 1_{\{x>1\}} + b_\delta^{(n),2}(x) 1_{\{x=1\}}.$$

From (4.5.4), we have, for some $\kappa \in (0,\infty)$,

$$\sup_{n\in\mathbb{N}} b_\delta^{(n),1}(x) \leq -\delta d_1 x \inf_{n\in\mathbb{N}} \lambda^{(n)} \leq -\delta\kappa x \leq -\delta\kappa$$

for all $\delta \leq \delta_0$ (with δ_0 as in (4.5.4)) and $n \in \mathbb{N}$. We note that

$$U_t^{(n)} := e^{\delta \hat{X}_t^{(n)} - \int_0^t b_\delta^{(n)}(\hat{X}_s^{(n)}) ds}, \quad t \geq 0, \tag{4.5.9}$$

is a martingale ([37]). Fix δ and ρ as in the statement of Lemma 4.5.1. Without loss of generality, we can assume that $\delta \leq \delta_0$. Note that on the set

$$\{\omega : \hat{X}_s^{(n)}(\omega) > 1 \text{ for all } s \in [(j-1)\rho, j\rho)\},$$

we have

$$\delta[\hat{X}_{j\rho}^{(n)} - \hat{X}_{(j-1)\rho}^{(n)}] \leq \delta[\hat{X}_{j\rho}^{(n)} - \hat{X}_{(j-1)\rho}^{(n)}] - \int_{(j-1)\rho}^{j\rho} b_\delta^{(n)}(\hat{X}_s^{(n)})ds - \delta\kappa\rho \equiv v_j^{(n)} - \delta\kappa\rho.$$

Fix $t > 0$ and let $N \in \mathbb{N}$ be such that $(N-1)\rho \leq t < N\rho$. Then on the set

$$\{\omega : \hat{X}_t^{(n)}(\omega) > 1 \text{ for all } s \in [(N-1)\rho, t)\},$$

$\delta[\hat{X}_t^{(n)} - \hat{X}_{(N-1)\rho}^{(n)}] \leq v_N^{(n)}(t)$, where

$$v_j^{(n)}(t) := \delta[\hat{X}_t^{(n)} - \hat{X}_{(j-1)\rho}^{(n)}] - \int_{(j-1)\rho}^{t} b_\delta^{(n)}(\hat{X}_s^{(n)})ds.$$

Now, for a fixed ω, let $m \equiv m(\omega)$ be such that $[(m-1)\rho, m\rho)$ is the last interval in which $\hat{X}^{(n)}$ visits 1 before time $N\rho$. We distinguish between the cases $m < N$, $m = N$, and $m = 0$, where the latter corresponds to the case where 1 is not visited before time $N\rho$.

Case 1: $m < N$.

In this case

$$\delta\hat{X}_t^{(n)} \leq \delta\hat{X}_{m\rho}^{(n)} + \sum_{j=m+1}^{N-1}(v_j^{(n)} - \delta\kappa\rho) + v_N^{(n)}(t).$$

For $j \in \mathbb{N}$, let

$$\gamma_j^{(n)} := \inf\{t \geq (j-1)\rho | \hat{X}_t^{(n)} = 1\} \wedge j\rho$$

and

$$\theta_j^{(n)} := \sup_{0 \leq t \leq \rho}[\hat{X}_{(t+\gamma_j^{(n)})\wedge j\rho}^{(n)} - \hat{X}_{\gamma_j^{(n)}}^{(n)}]. \tag{4.5.10}$$

Then

$$\delta \hat{X}^{(n)}_{m\rho} \leq \delta\theta^{(n)}_m + \delta.$$

Combining the above estimates, we have

$$\delta \hat{X}^{(n)}_t \leq \delta\theta^{(n)}_m + \delta + \sum_{j=m+1}^{N-1} (v^{(n)}_j - \delta\kappa\rho) + v^{(n)}_N(t). \tag{4.5.11}$$

Case 2: $m = N$.

In this case $\hat{X}^{(n)}_s = 1$ for some $s \in [(N-1)\rho, N\rho)$. Suppose first that there is no such $s \in [(N-1)\rho, t)$. In that case, we see, exactly as in Case 1, that (4.5.11) holds. Now consider the case where there is an $s \in [(N-1)\rho, t]$ such that $\hat{X}^{(n)}_s = 1$. It then follows that $\delta\hat{X}^{(n)}_t \leq \delta\theta^{(n)}_m + \delta$.

Case 3: $m = 0$.

In this case, 1 is not visited before time $N\rho$ and thus

$$\delta \hat{X}^{(n)}_t = \delta \hat{X}^{(n)}_0 + \sum_{j=m+1}^{N-1} (v^{(n)}_j - \delta\kappa\rho) + v^{(n)}_N(t).$$

Combining the three cases, we have

$$\begin{aligned}
\delta \hat{X}^{(n)}_t &\leq v^{(n)}_N(t) + \sum_{j=m+1}^{N-1} (v^{(n)}_j - \delta\kappa\rho) + \delta\theta^{(n)}_m + \delta + \delta\hat{X}^{(n)}_0 \\
&\leq v^{(n)}_N(t) + \max_{1\leq l\leq N} \left\{ \sum_{j=l}^{N-1} (v^{(n)}_j - \delta\kappa\rho) + \delta\theta^{(n)}_l \right\} + \delta + \delta\hat{X}^{(n)}_0.
\end{aligned}$$

Thus, for any $M_0 > 0$,

$$\begin{aligned}
P_x(\delta\hat{X}^{(n)}_t \geq M_0) &\leq \sum_{l=1}^{N} P_x\left(v^{(n)}_N(t) + \sum_{j=l}^{N} (v^{(n)}_j - \delta\kappa\rho) + \delta\theta^{(n)}_l + \delta + \delta\hat{X}^{(n)}_0 \geq M_0 \right) \\
&\leq e^{-M_0} \sum_{l=1}^{N} \left[e^{\delta(1+x)} E_x \left(\exp\left[\delta\theta^{(n)}_l + \sum_{j=l}^{N-1} v^{(n)}_j + v^{(n)}_N(t) \right] \right) e^{-\delta\kappa\rho(N-l+1)} \right].
\end{aligned}$$

Recalling $U^{(n)}$ from (4.5.9) and using its martingale property, we get

$$P_x(\delta \hat{X}_t^{(n)} \geq M_0) \leq e^{-M_0} e^{\delta(1+x)} \sum_{l=1}^{N} e^{-\delta\kappa\rho(N-l+1)} E_x\left(e^{\delta\theta_l^{(n)}}\right)$$
$$\leq d(\delta,\rho,1) e^{-M_0} \frac{e^{(1+x)\delta}}{1-e^{-\delta\kappa\rho}}, \tag{4.5.12}$$

where the last inequality follows from Lemma 4.5.1 and the observation that

$$\sup_{n\in\mathbb{N}} E_x(e^{\delta\theta_l^{(n)}}) \leq \sup_{n\in\mathbb{N}} E_1\left(\sup_{0\leq t\leq\rho} e^{\delta\hat{X}_t^{(n)}}\right) \leq d(\delta,\rho,1) < \infty, \tag{4.5.13}$$

where $\theta_l^{(n)}$ is as in (4.5.10). Since the constant $d(\delta,\rho,1)$ in (4.5.13) is independent of $n \in \mathbb{N}$, $t \geq 0$, and N, we get from (4.5.12) for all $t \geq 0$ and $n \in \mathbb{N}$

$$E_x(e^{\delta\hat{X}_t^{(n)}}) = \int_0^\infty P_x(\delta\hat{X}_t^{(n)} > y)dy \leq d(\delta,\rho,1)\frac{e^{(1+x)\delta}}{1-e^{-\delta\kappa\rho}} \int_0^\infty e^{-y}dy = \tilde{d}(\delta)e^{\delta x},$$

where $\tilde{d}(\delta) = d(\delta,\rho,1)\frac{e^\delta}{1-e^{-\delta\kappa\rho}}$. The result follows. $\square$

Proof of Theorem 4.5.1. Fix an $L > 1$, and let $\tau^{(n)} := \inf\{t : \hat{X}_t^{(n)} \leq L\}$. Observe that if $t \in [(N-1)\rho, N\rho)$ for some $N \in \mathbb{N}$, then, following arguments as in the proof of Lemma 4.5.2,

$$P_x(\tau^{(n)} > t) \leq P_x\left(\sum_{j=1}^{N-1}(v_j^{(n)} - \delta\kappa\rho) > \delta(L-x)\right) \leq e^{\delta(x-L-\delta\kappa\rho(N-1))}.$$

Thus we have that

$$\sup_{n\in\mathbb{N}} P_x(\tau^{(n)} > t) \leq \gamma_1 e^{\delta x} e^{-\gamma_2 t},$$

where $\gamma_i \in (0,\infty)$, $i = 1,2$. The above estimate, along with Lemma 4.5.2, implies, for

$n \in \mathbb{N}$,

$$\begin{aligned}E_x e^{\frac{\delta}{2}\hat{X}_t^{(n)}} &= E_x\left(1_{\{\tau^{(n)}\le t\}}e^{\frac{\delta}{2}\hat{X}_t^{(n)}}\right) + E_x\left(1_{\{\tau^{(n)}>t\}}e^{\frac{\delta}{2}\hat{X}_t^{(n)}}\right)\\ &\le \tilde{d}(\delta)e^{\delta L} + \left(\gamma_1 e^{\delta x}e^{-\gamma_2 t}\right)^{\frac{1}{2}}\left(E_x\left(e^{\delta \hat{X}_t^{(n)}}\right)\right)^{\frac{1}{2}} \le \tilde{d}(\delta)e^{\delta L} + \left(\gamma_1 e^{\delta x}e^{-\gamma_2 t}\right)^{\frac{1}{2}}(\tilde{d}(\delta)e^{\delta x})^{\frac{1}{2}}\\ &\le d_1(1+e^{\delta x}e^{-\frac{\gamma_2}{2}t}),\end{aligned}$$

where $\tilde{d}(\delta)$ is as in (4.5.3) and $d_1 \in (0,\infty)$ is some constant, independent of n. Fix $p > 0$. Then, for some $d_2 \in (0,\infty)$, we have

$$\sup_{n\in\mathbb{N}} \frac{E_x(\hat{X}_{tx}^{(n)})^p}{x^p} \le \sup_{n\in\mathbb{N}} \frac{d_2 E_x e^{\frac{\delta}{2}\hat{X}_{tx}^{(n)}}}{x^p} \le \sup_{n\in\mathbb{N}} \frac{d_1 d_2(1+e^{\delta x}e^{-\frac{\gamma_2}{2}tx})}{x^p}.$$

Choose t_0 large enough such that $\frac{\gamma_2}{2}t_0 > \delta$. Then for $t \ge t_0$

$$\lim_{x\to\infty}\sup_{n\in\mathbb{N}} \frac{E_x\left((\hat{X}_{tx}^{(n)})^p\right)}{x^p} = 0.$$

The result follows $\square$

As a consequence of Theorem 4.5.1, we have the following result. The proof is adapted from that of Theorem 3.4 in [4] (see also references therein). For $\delta \in (0,\infty)$, define the return time to a compact set $C \subset \mathbb{R}_{\ge 1}$ by $\tau_C^{(n)}(\delta) := \inf\{t \ge \delta | \hat{X}_t^{(n)} \in C\}$.

Theorem 4.5.2. *There are $\tilde{c}, \delta \in (0,\infty)$ and a compact set $C \in \mathbb{R}_{\ge 1}$ such that*

$$\sup_n E_x\left(\int_0^{\tau_C^{(n)}(\delta)} (\hat{X}_t^{(n)})^2 dt\right) \le \tilde{c}x^3, \quad x \ge 1.$$

Proof. Applying Theorem 4.5.1 with $p = 3$, we have that there is an $L \in (1,\infty)$ such that with $C := \{x \in \mathbb{R}_+ | x \leq L\}$, for all $x \in C^c$,

$$\sup_n E_x \left(\left(\hat{X}^{(n)}_{t_0 x} \right)^3 \right) \leq \frac{1}{2} x^3, \tag{4.5.14}$$

where t_0 is as in Theorem 4.5.1. Let $\delta := t_0 L$ and $\tau^{(n)} := \tau^{(n)}_C(\delta) := \inf\{t \geq \delta | \hat{X}^{(n)}_t \leq L\}$. Define stopping times as follows:

$$\sigma^{(n)}_0 := 0, \quad \sigma^{(n)}_m := \sigma^{(n)}_{m-1} + t_0 \left(\hat{X}^{(n)}_{\sigma^{(n)}_{m-1}} \vee L \right), \quad m \in \mathbb{N}.$$

Let $m^{(n)}_0 = \min\{m \geq 1 | \hat{X}^{(n)}_{\sigma^{(n)}_m} \leq L\}$, and

$$\hat{V}^{(n)}(x) := E_x \left(\int_0^{\tau^{(n)}} (\hat{X}^{(n)}_t)^2 dt \right).$$

Then

$$\hat{V}^{(n)}(x) \leq E_x \left(\int_0^{\left(\sigma^{(n)}_{m^{(n)}_0}\right)} (\hat{X}^{(n)}_t)^2 dt \right) = \sum_{k=0}^{\infty} E_x \left(\int_{\sigma^{(n)}_k}^{\sigma^{(n)}_{k+1}} (\hat{X}^{(n)}_t)^2 dt 1_{\{k < m^{(n)}_0\}} \right).$$

Let $\mathcal{F}^{(n)}_t := \sigma\{\hat{X}^{(n)}_s | 0 \leq s \leq t\}$, then we claim that there is a $c_0 \in (0, \infty)$ such that for all $n, k \in \mathbb{N}$, $x \geq 1$

$$E_x \left(\int_{\sigma^{(n)}_k}^{\sigma^{(n)}_{k+1}} (\hat{X}^{(n)}_t)^2 dt | \mathcal{F}^{(n)}_{\sigma^{(n)}_k} \right) 1_{\{k < m^{(n)}_0\}} \leq c_0 \left(\hat{X}^{(n)}_{\sigma^{(n)}_k} \right)^3 1_{\{k < m^{(n)}_0\}}. \tag{4.5.15}$$

To prove the claim it suffices to show, by the strong Markov property, that there is a $c_0 \in (0, \infty)$ such that for all $n \in \mathbb{N}, x \geq 1$

$$E_x \left(\int_0^{\sigma^{(n)}_1} (\hat{X}^{(n)}_t)^2 dt \right) \leq c_0 x^3.$$

Note that $\sigma^{(n)}_1 = t_0(x \vee L) \leq \tilde{c}_0 x$, for some $\tilde{c}_0 \in (0, \infty)$. Using this bound along with arguments similar to those leading to (3.3.15) (see also Lemma 3.2.2 of [20]), we get, for some $\hat{c}_0 \in (0,\infty)$,

$$\sup_{n\in\mathbb{N}} E_x\left(\sup_{t\le\sigma_1^{(n)}} (\hat{X}_t^{(n)})^2\right) \le \hat{c}_0 x^2.$$

The claim follows.

From the estimate (4.5.15), we now have

$$\sup_n \hat{V}^{(n)}(x) \le c_0 \sup_n E_x\Bigg(\sum_{k=0}^{m_0^{(n)}-1} \left(\hat{X}^{(n)}_{\sigma_k^{(n)}}\right)^3 \Bigg). \tag{4.5.16}$$

Note that $\left\{\hat{X}^{(n)}_{\sigma_k^{(n)}}\right\}_{k\in\mathbb{N}_0}$ is a Markov chain with transition probability kernel

$$\check{P}^{(n)}(x,A) := P^{(n)}_{t_0(x\vee L)}(x,A), \quad x\in\mathbb{R}_{\ge 1}, \quad A\in\mathcal{B}(\mathbb{R}_{\ge 1}),$$

where $P_t^{(n)}$ is the transition probability kernel for $\hat{X}^{(n)}$. Using (4.5.14), we get, for all $x\in[1,\infty)$,

$$\begin{aligned}
&\sup_n \int_1^\infty y^3 \check{P}^{(n)}(x,dy) = \sup_n \int_1^\infty y^3 P^{(n)}_{t_0(x\vee L)}(x,dy) \qquad (4.5.17)\\
&= \sup_n \int_1^\infty y^3 P^{(n)}_{t_0 x}(x,dy) 1_{\{x>L\}} + \sup_n \int_1^\infty y^3 P^{(n)}_{t_0 L}(x,dy) 1_{\{x\le L\}} \le x^3 - \frac{1}{2}x^3 + \tilde{b}1_{[0,L]}(x),
\end{aligned}$$

where $\tilde{b} := \tilde{b}(L)\in(0,\infty)$. The above inequality along with Theorem 14.2.2 of [28] yields

$$\begin{aligned}
\sup_n E_x\Bigg(\sum_{k=0}^{m_0^{(n)}-1} \left(\hat{X}^{(n)}_{\sigma_k^{(n)}}\right)^3 \Bigg) &\le 2\Bigg(x^3 + \sup_n E_x\Bigg(\sum_{k=0}^{m_0^{(n)}-1} \tilde{b}1_{[0,L]}\left(\hat{X}^{(n)}_{\sigma_k^{(n)}}\right)\Bigg)\Bigg)\\
&= 2\left(x^3 + \tilde{b}1_{[0,L]}(x)\right) \le \tilde{c}x^3,
\end{aligned}$$

where the equality in the last display follows from the fact that $\hat{X}^{(n)}_{\sigma_k^{(n)}} > L$ for $1\le k < m_0^{(n)}$. Using the last estimate together with (4.5.16), the result follows. $\square$

The following theorem is proved exactly as Proposition 5.4 of [8] (see also Theorem 3.5 of [4]). The proof is omitted.

Theorem 4.5.3. *Let $f : \mathbb{R}_{\geq 1} \to \mathbb{R}_+$ be a measurable function. Define for $\hat{\delta} \in (0, \infty)$ and a compact set $C \subset \mathbb{R}_{\geq 1}$*

$$V^{(n)}(x) := E_x \left(\int_0^{\tau_C^{(n)}(\hat{\delta})} f(\hat{X}_t^{(n)}) dt \right), \quad x \in \mathbb{R}_{\geq 1}.$$

If $\sup_{n \in \mathbb{N}} V^{(n)}$ is everywhere finite and uniformly bounded on C, then there exists a $\hat{\kappa} \in (0, \infty)$ such that for all $n \in \mathbb{N}, t > 0, x \in \mathbb{R}_{\geq 1}$

$$\frac{1}{t} E_x \left(V^{(n)}(\hat{X}_t^{(n)}) \right) + \frac{1}{t} \int_0^t E_x \left(f(\hat{X}_s^{(n)}) \right) ds \leq \frac{1}{t} V^{(n)}(x) + \hat{\kappa}.$$

We now complete the proof of Theorem 4.2.1, which is adapted from that of Theorem 3.2 in [4]. Recall that it only remains to establish the tightness of $\{\nu_1^{(n)}\}_{n \in \mathbb{N}}$. We will apply Theorem 4.5.3 with $f(x) := x^2$, and $\hat{\delta}, C$ as in Theorem 4.5.2. We will show that for all $n \in \mathbb{N}$ and $\hat{\kappa}$ as in Theorem 4.5.3

$$\langle \nu_1^{(n)}, f \rangle := \int_{\mathbb{R}_{\geq 1}} f(x) \nu_1^{(n)}(dx) \leq \hat{\kappa},$$

from which tightness is immediate. Since $\nu_1^{(n)}$ is an invariant measure for $\hat{X}^{(n)}$, we have for non-negative, real valued, measurable functions Φ on $\mathbb{R}_{\geq 1}$

$$\int_{\mathbb{R}_{\geq 1}} E_x \left(\Phi(\hat{X}_t^{(n)}) \right) \nu_1^{(n)}(dx) = \langle \nu_1^{(n)}, \Phi \rangle. \tag{4.5.18}$$

Fix $k \in \mathbb{N}$ and let $V_k^{(n)}(x) := V^{(n)}(x) \wedge k$. Let

$$\Psi_k^{(n)}(x) := \frac{1}{t} V_k^{(n)}(x) - \frac{1}{t} E_x \left(V_k^{(n)}(\hat{X}_t^{(n)}) \right).$$

By (4.5.18), we have that $\int_{\mathbb{R}_{\geq 1}} \Psi_k^{(n)}(x)\nu_1^{(n)}(dx) = 0$. Let

$$\Psi^{(n)}(x) := \frac{1}{t}V^{(n)}(x) - \frac{1}{t}E_x\left(V^{(n)}(\hat{X}_t^{(n)})\right).$$

By the monotone convergence theorem $\Psi_k^{(n)}(x) \to \Psi^{(n)}(x)$ as $k \to \infty$. We next show that $\Psi_k^{(n)}(x)$ is bounded from below for all $x \in \mathbb{R}_{\geq 1}$: If $V^{(n)}(x) \leq k$

$$\Psi_k^{(n)}(x) = \frac{1}{t}V_k^{(n)}(x) - \frac{1}{t}E_x\left(V_k^{(n)}(\hat{X}_t^{(n)})\right) \geq \frac{1}{t}V^{(n)}(x) - \frac{1}{t}E_x\left(V^{(n)}(\hat{X}_t^{(n)})\right) \geq -\hat{\kappa}$$

where the last inequality follows from Theorem 4.5.3. If $V^{(n)}(x) \geq k$

$$\Psi_k^{(n)}(x) = \frac{1}{t}k - \frac{1}{t}E_x\left(V_k^{(n)}(\hat{X}_t^{(n)})\right) \geq 0.$$

Thus $\Psi_k^{(n)}(x) \geq -\hat{\kappa}$ for all $x \geq 1$. By Fatou's lemma, we have

$$\int_{\mathbb{R}_{\geq 1}} \Psi^{(n)}(x)\nu_1^{(n)}(dx) \leq \liminf_{k\to\infty} \int_{\mathbb{R}_{\geq 1}} \Psi_k^{(n)}(x)\nu_1^{(n)}(dx) = 0.$$

By Theorem 4.5.3 we have $\Psi^{(n)}(x) \geq \frac{1}{t}\int_0^t E_x\left(f(\hat{X}_s^{(n)})\right)ds - \hat{\kappa}$. Combining this with the last display, we have

$$0 \geq \int_{\mathbb{R}_{\geq 1}} \Psi^{(n)}(x)\nu_1^{(n)}(dx) \geq \frac{1}{t}\int_0^t \int_{\mathbb{R}_{\geq 1}} E_x\left(f(\hat{X}_s^{(n)})\right)\nu_1^{(n)}(dx)ds - \hat{\kappa}.$$

Using the invariance property of $\nu_1^{(n)}$ once more, we have $\langle \nu_1^{(n)}, f\rangle \leq \hat{\kappa}$, which completes the proof of tightness and thus of Theorem 4.2.1.

Chapter 5

Modeling Cancer Cell Behavior as a Function of Substrate Stiffness

5.1 Introduction

The final topic of this dissertation is concerned with the statistical analysis of cell growth, metabolic activity, viability, and morphology data as part of the *EFRI-CBE*[1] project *Emerging Frontiers in 3-D Breast Cancer Tissue Test Systems*. This is an NSF project that involves researchers from multiple universities and departments: the Bioengineering, Chemistry, Electrical and Computer Engineering, Animal and Veterinary Science, and Biological Sciences Departments at Clemson University, the Biology Department at the University of North Carolina at Charlotte, and the Department of Statistics and Operations Research at the University of North Carolina at Chapel Hill. The Clemson University Bioengineering Department is taking the lead and coordinating role.

The overall aims of the project are twofold: (i) to enhance the knowledge of the relationships between normal and breast cancer cellular behavior, and (ii) to study the impact on cell growth behavior of factors such as tissue stiffness and oxygen level. The goal is to develop bioengineering tools to build tissues (or their *in vitro* representatives, *hydrogels*) and to assess experimentally and analytically the aforementioned relationships.

---

[1] Emerging Frontiers in Research and Innovation - Cellular and Biomolecular Engineering

The research is multidisciplinary in nature with a large component concerned with engineering and experimentation efforts. The above mentioned goals can be divided into the following specific tasks:

1. To build a tissue fabricator which can produce 3-D cellular breast tissues of different (controlled) stiffness and oxygen levels. Moreover, the fabricator should be able to produce tissues that have a stiffness gradient.

2. To develop nanoparticle oxygen sensors to measure site specific oxygen levels.

3. To develop a closed-loop control mechanism for oxygen regulation within an *in vitro* tissue sample.

4. To develop cellular assays to measure cell state.

5. Modeling the relationships between cells (cell growth, metabolic activity, and migration) and environment (tissue stiffness and oxygen levels). Analyzing the data obtained from experiments and providing suggestions for improvements of the designs of future experiments.

Our role in this project is in the last item on this list.

The proof-of-principle study presented here is based on a working paper ([2]) involving multiple collaborators. This study will lead to further experiments addressing the above mentioned questions. Large parts of the following are as in [2]. However, any imprecision resulting from adapting the material is the authors responsibility.

It is well known that 3D culture systems are more complex and therefore evoke very different cellular behavior or interactions than those found in 2D systems; hence, carefully crafted 3D systems have the potential to unlock more mysteries of disease processes. The objectives of this proof-of-principle study were to build a simple 3D breast cancer tissue test system, using cells and a biomaterial substrate, and to correlate substrate stiffness to cancer cell behavior. The latter was part of the aim to show the potential for the development of mathematical models to better define substrate-cell interactions in a breast cancer tissue test system (and thus show the potential for mathematical models to improve the design of the test system). Determining the microenvironmental differences between cancer and non-cancerous tissues may yield vital information for diagnosing and treating cancer. It has been observed that tumor cells and nearby tissues have greater stiffness than healthy tissue; however, the cause and effect of these stiffness differences are not fully understood ([29, 31], see also [2] for a more detailed description of the biological background and further references).

Two-dimensional cultures have been shown to be poor indicators of cell behavior *in vivo*, where three-dimensional models (also termed test systems) often more accurately simulate natural conditions [6, 17, 42, 43]. Three-dimensional models afford replicable testing of tumorigenesis while avoiding many of the inaccuracies of two-dimensional models. Tumor invasiveness has been strongly tied to the migration of cancer cells into healthy tissue [33]. In two-dimensional cell cultures, migration rates of cells have been widely studied and have been found to be dependent on both culture substrate mechanical properties and cell-adhesion ligand densities. Migration rates are greatest at intermediate cell-adhesion ligand densities as compared to high densities (which oppose cell detachment, inhibiting movement) and low densities (which discourage the cell attachment necessary for movement), and at higher levels of substrate stiffness. In three-dimensional matrices, however, the relationship between ligand concentration and matrix stiffness, and cell migration, is not as clearly understood [43]. Additionally, it has been suggested that substrate pore size [36] and sensitivity to matrix proteolysis [43] also influence migration and proliferation rates of cells in 3D matrices.

In this study, two experimental systems were developed by investigators at Clemson University to address the physiological inadequacies of two-dimensional systems. A

3D hydrogel system was used to examine the effects of substrate stiffness on tumorigenesis. However, due to the numerous parameters which may affect cell migration and proliferation in true 3D systems (e.g. stiffness, ligand density, pore size, and proteolysis susceptibility [36, 43]), a simplified hydrogel system was also used. In the latter system, termed "2.5-dimensional" (2.5D), cells are cultured on top of a soft hydrogel. This may reduce restrictions to cell migration and proliferation based on substrate pore size and proteolysis while simultaneously providing a physiologically soft and compliant matrix absent in true two-dimensional culture. The 2.5D hydrogel test system was developed to control substrate stiffness. Microscope image processing techniques were used to obtain morphological information about the population of cultured tumor cells. Both *in vitro* systems allow control of ligand density and substrate stiffness, which are both important in regulating tumorigenesis and cancer invasiveness [33]. Using the morphological data obtained from image processing, regression models were developed to correlate characteristics such as tumor size, perimeter, and number of tumors to substrate stiffness. Previous work has employed measurements of tumor size for assessment of tumor progression [7, 30]. The imaging-based measurements of tumor morphological parameters were developed to serve here as a non-destructive, spatially-sensitive alternative to direct tests of cell proliferation or migration.

5.2 Metabolic Activity and Viability

5.2.1 Preparation and Measurements

Cancer cells (MCF-7) were suspended in "strong" and "weak" hydrogels (containing 2% and 1.1% agarose, respectively). The gels also contained collagen as compared to gelatin in the *Morphology* experiment, below. The latter might be a reason for contrasting results as discussed below, and should be kept in mind. The sample size was $n = 6$ for each of the weak and strong gel groups.

The cells were tested daily for metabolic activity by measuring glucose consumption, lactic acid production, and Alamar Blue uptake. Higher values in these measurements indicate higher metabolic activity. After 14 days in culture, two samples from each of the two experimental groups (strong gel and weak gel) were used for qualitative Live/Dead cell viability assays and imaged via a fluorescent microscope (see Figure 5.3).

For a more detailed description of the experimental setup, measurements, materials, and instruments used, see [2].

5.2.2 Lactic Acid Production and Glucose Consumption

Lactic acid production and glucose consumption were measured over 24 hour intervals in g/L. The measurements were normalized by subtracting the average measurement of the acellular controls, as described below.

Since weak gels were mechanically unstable, they were continually destroyed by routine handling (media change, etc.), and no weak gel metabolic samples remained by day 11. Note that for the control group for the weak hydrogel only one sample per day (up to day 6) was available. As a result, the averaged measurements from the strong gel (acellular) control group were used to normalize the measurements of the weak and strong gel groups. A justification for this normalization was a comparison (paired t-test) between the six measurements from the weak gels and the six averaged measurements from the strong gels (based on days 1 through 6), which showed no significant difference between the averaged measurements from the strong gel and the measurements for the weak gel ($p = 0.51$ and $p = 0.89$ for the difference in lactic acid production and glucose consumption, respectively).

Lactic acid production by cells in strong gels was nearly constant and approximately the same as that in acellular controls from day 4 on, whereas production in the weak cellular gels increased over time and was, based on t-tests, significantly greater ($p < 0.05$) than that measured in control samples by day 3 (Figure 5.1). Cells in weak gels consumed significantly more glucose than cells in strong gels on days 7 and 8 ($p < 0.05$). Even though the difference is statistically significant on these days, the overall difference in glucose consumption between weak and strong gels appears to be small. Note that all t-tests were performed on a day-by-day basis.

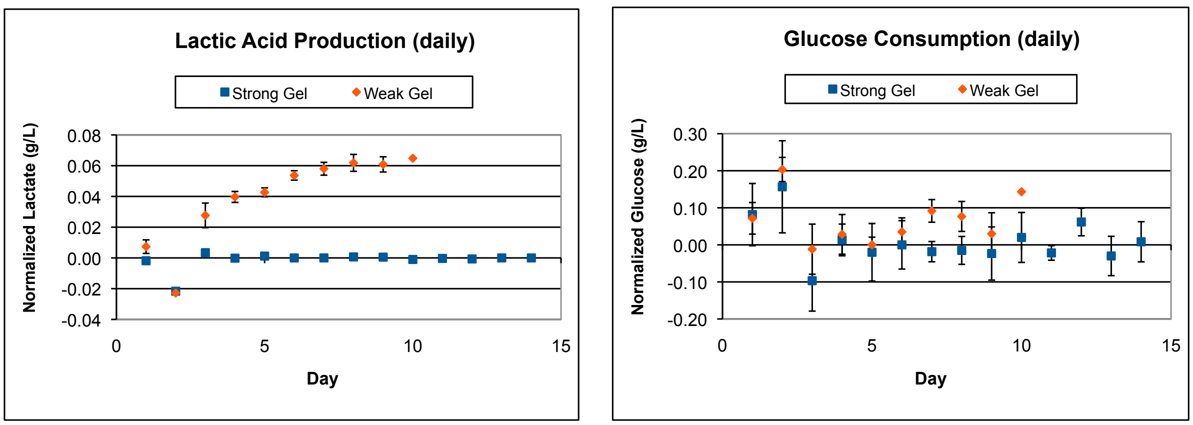

Figure 5.1: Sample means and standard deviations of the normalized lactic acid production and glucose consumption measured over 24 hour intervals (from day $i - 1$ to day i).

5.2.3 Alamar Blue Assay

All Alamar Blue absorbance measurements were normalized by subtracting the average absorbance of the acellular controls (Figure 5.2). Note that we again used, as for the glucose and lactic acid data, the average values of the strong control group to normalize the weak and strong experimental groups. This time, the difference between the strong and weak control groups appeared to be different. However, on average (for the first 6 days), the measurements for the strong control group are only 0.05 g/L larger than those for the weak control group. This is very small compared to the difference between the experimental groups (see Figure 5.2).

The absorbance values resulting from weak gels are statistically significantly larger (at significance level $\alpha = 0.05$) than those of both strong gels and control gels at all time points. Absorbance values from strong gels are, based on day-by-day t-tests, not significantly different ($\alpha = 0.05$) from those of control gels at all time points excluding days 7, 9, and 11. Note that while the sample size was initially 6, the sample size in the weak gel group declined over time due to gel fracture and was as follows: 5 on days 1-3, 4 on day 4, 3 on days 5-8, 2 on day 9, 1 on day 10, and 0 by day 11. The sample size was 6 on all days in the strong gel group.

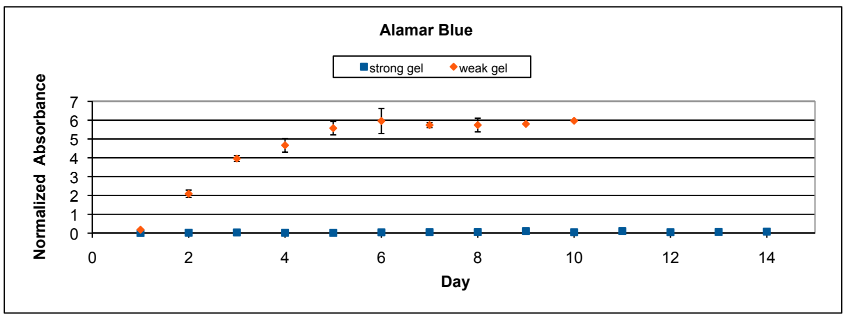

Figure 5.2: Alamar Blue assay results, showing the increase in metabolic activity over time of cells in weak gels and the absence of metabolic activity of cells in strong gels. Sample means and standard deviations of the normalized absorbance levels of the weak and strong gel groups are given.

5.2.4 Live/Dead Cell Viability Assay

Live/Dead cell staining on day 14 showed the presence of live MCF-7 aggregates (on the order of 100-200 μm diameter) as well as live, single cells (non-aggregated) in weak gels (Figure 5.3). No aggregates formed in strong gels, which only had single cells.

5.3 Morphology Imaging

In this experiment, MCF-7 cells were seeded in a "2.5D" hydrogel system, that is, cells were seeded in monolayers on hydrogels. The agarose content in the different groups was,

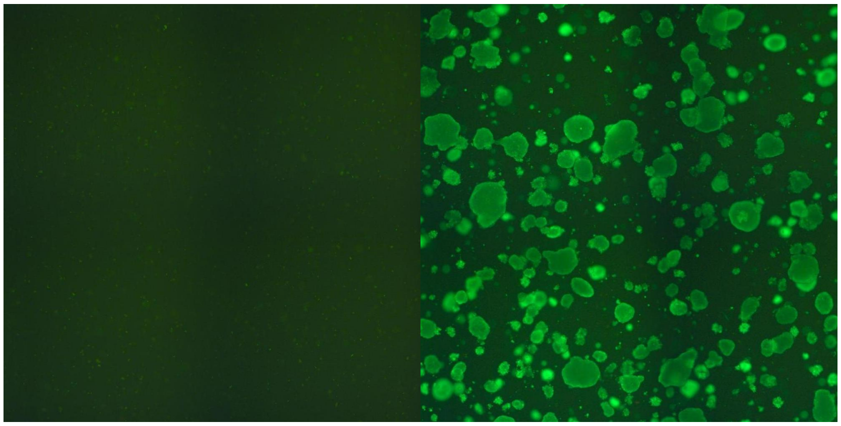

Figure 5.3: Live/Dead assay. On the left: strong gels showed no MCF-7 aggregate formation over 14 days. On the right: weak gels developed many large cell aggregates.

respectively, 0.75%, 1%, 1.25%, 1.5%, 2%, 2.25%, 2.5%, and 3%. The gels also contained gelatin as compared to collagen in the *Metabolic Activity and Viability* experiment, above. As mentioned before, this might be a reason for contrasting results as discussed below, and should be kept in mind. A sample size of three for each group was used for this experiment.

Medium was added following imaging on days 1, 2, and 3. Additionally, medium was completely replaced following imaging on days 5, 8, and 12. The time points of medium replenishment and change are noted since these procedures may have caused aggregate displacements and disturbances.

ImageJ software was used to convert images to black and white format and for data collection from the binary processed images. In this analysis, each isolated cellular region is referred to as an aggregate. An aggregate could be composed of a single cell, several cells clustered together, or even multiple touching cell aggregates. The large connected black area in Figure 5.4 is thus counted as one rather than several aggregates. Processed images qualitatively showed a high fidelity, that is few acellular regions were incorrectly marked as cells (black), and all aggregates were appropriately marked as cells. However, late time point images are an exception to this rule (approximately days 10-14), when aggregates were noted to cluster. In this event, due to the image processing procedures, intra-cluster spaces (areas enclosed by aggregates) were erroneously counted as part of the surrounding aggregates (Figure 5.4).

For each image, the following data were collected: size / coverage area of each aggregate, perimeter of each aggregate, and total number of aggregates.

For a more detailed description of the setup and instruments used, see [2].

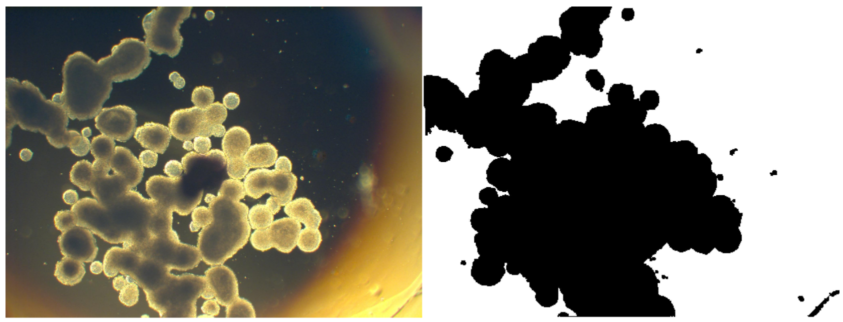

Figure 5.4: Example of errors introduced by the image processing protocol at late time points when multi-aggregate MCF-7 clusters were present. The image displayed is from the 3% agarose group on day 10. Note that intra-cluster spaces are erroneously represented as part of the surrounding aggregate.

5.3.1 Statistical Modeling

Regression analyses were performed to assess the feasibility of modeling aggregate measurements as a function of substrate stiffness (i.e. agarose content) and time. The measurements examined were coverage area over threshold (cumulative area of aggregates that are larger than a threshold), number of aggregates, average coverage area of aggregates, and aggregate perimeter. In the following analyses, the experimental groups with 0.75%, 1%, and 1.5% agarose were excluded. The gels of the two former groups were too weak and fell apart; moreover cells sank through the gels. In most groups, cells sank through the gels and were not consistently in focus for image processing. However, only in the group with 1.5% agarose this effect seemed to be strong enough to warrant exclusion of the group.

Throughout the following discussion, $\mathtt{S}$ denotes the stiffness of the gel as measured by % agarose used, and $\mathtt{D}$ denotes the day of the measurement. The interaction effect between stiffness and day is denoted by $\mathtt{S} \cdot \mathtt{D}$. Other interaction terms are denoted analogously. Length and area are both measured in pixels, with 1 pixel corresponding to 2.635 μm and 6.925 μm^2, respectively.

Coverage Area over Threshold Coverage area was measured over a size threshold of 15,000 pixels, both to reduce noise and eliminate metabolically inactive single (non-aggregated) cells. Generally, coverage area over the size threshold (referred to as Coverage over Threshold, COT) increased over time more on gels with higher agarose content. This result was especially prominent over days 6-8 (see Figure 5.5), during which time culture media were not exchanged. The latter time period is of special importance since it was noted that handling, including medium exchange, could displace weakly-attached aggregates, so that they were not consistently in the microscope's focus, thus increasing day-to-day variability of measurements.

Although culture medium exchange introduces variables such as rapidly changing biomolecule concentrations and cell displacement via handling and fluid shear, a model describing COT as a function of agarose content is still promising over days 1-10, with multiple $R^2 = 0.79$. The fitted model is as follows (see also Table 5.1):

$$\mathtt{COT} = -111695.1 + 67944.8\,\mathtt{S} - 2362\,\mathtt{D}^2 + 4659.5\,\mathtt{S} \cdot \mathtt{D}^2.$$

The errors, which are omitted from this equation, are assumed to be independent and identically normally distributed. Note that the assumption of equal variance appears to be violated. On the other hand, the errors appear to be approximately normally distributed. Transformations of the data were considered with the aim of obtaining a variance that is closer to being constant. However, after transforming the data, the errors appeared to be less normally distributed. In the trade off between normality and equal

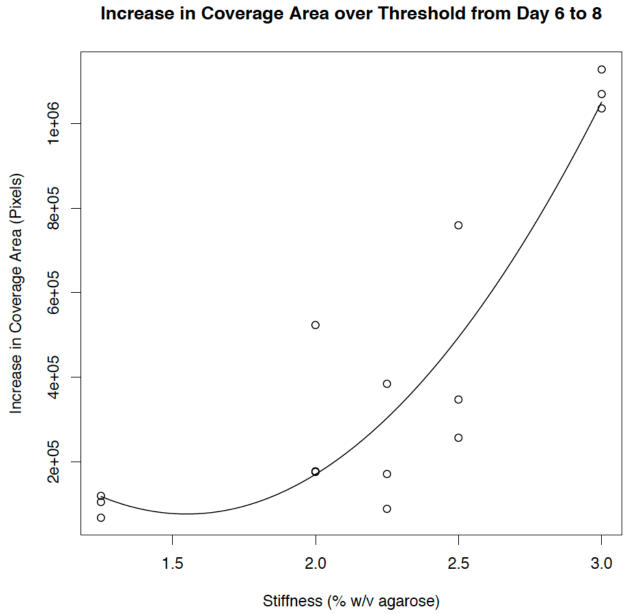

Figure 5.5: Quadratic model demonstrating the increase in `COT` as a function of agarose content from day 6 - 8, during which time culture medium was not exchanged. Multiple $R^2 = 0.83$.

variance of the errors, the unequal variance is of lesser concern in the present context; if the variance is not constant, the estimators are still unbiased, and we use this model only for data description rather than for inference. The coefficient of the interaction $\texttt{S} \cdot \texttt{D}^2$ is the only one that is statistically significant ($p < 0.05$). To understand the model equation, and in particular the interaction term, we give the following example. For day 4 and a gel with 3% agarose, the model for `COT` predicts that

$$\texttt{COT} = -111695.1 + 67944.8 \cdot 3 - 2362 \cdot 4^2 + 4659.5 \cdot 3 \cdot 4^2 = 278003.3.$$

The `COT` increases over time faster on stiffer gels, which indicates that on stiffer gels cells aggregate more and/or that there is greater cell growth than on weaker gels.

Number of Aggregates The natural logarithm of the number of aggregates, ln(`N`), was modeled as a function of agarose content (`S`) and time (`D`), with the result that ln(`N`) decreased over time, and faster so on stiffer gels. The logarithm was taken to linearize

Table 5.1: Estimates of the coefficients of the model for COT are shown, along with the corresponding standard errors, t-statistics, and p-values.

	Estimate	Std Error	t-statistic	p-value
Intercept	-111695.1	111790	-0.999	0.319
Stiffness	67944.8	49141.4	1.383	0.169
Day^2	2362	2221.1	1.063	0.289
Stiffness $\cdot$ Day^2	4659.5	976.3	4.772	4.38e-06

data.

A weighted least squares regression model was fitted, with weights $1/\mathtt{D}$, which means that in this model, the errors are assumed to be independent and normally distributed, with mean 0 and variance proportional to D. The model equation is as follows (see also Figure 5.6 and Table 5.2):

$$\ln(\mathtt{N}) = 7.217 - 0.256\,\mathtt{S} - 0.518\,\mathtt{D} + 0.022\,\mathtt{D}^2 - 0.035\,\mathtt{S}\cdot\mathtt{D}.$$

This simple model fits the data quite well ($R^2 = 0.93$), and the statistically significant ($p = 0.031$) interaction term suggests that the logarithm of the number of aggregates decreases over time faster on stiffer gels. The residuals are larger in absolute value for smaller fitted values, indicating a possible violation of the model assumption. Note that, some potential outliers were removed before fitting the model. Although this model suggests an influence of stiffness on the change of the number of aggregates, the effect seems to be small.

Table 5.2: Estimates of the coefficients of the model for $\ln(\mathtt{N})$ are shown, along with the corresponding standard errors, t-statistics, and p-values.

	Estimate	Std Error	t-statistic	p-value
Intercept	7.21743	0.16479	43.798	$< 2 \times 10^{-16}$
Stiffness	-0.25644	0.06891	-3.721	0.000285
Day	-0.51821	0.05134	-10.094	$< 2 \times 10^{-16}$
Day^2	0.02180	0.00375	5.813	3.89×10^{-08}
Stiffness $\cdot$ Day	-0.03540	0.01627	-2.176	0.031192

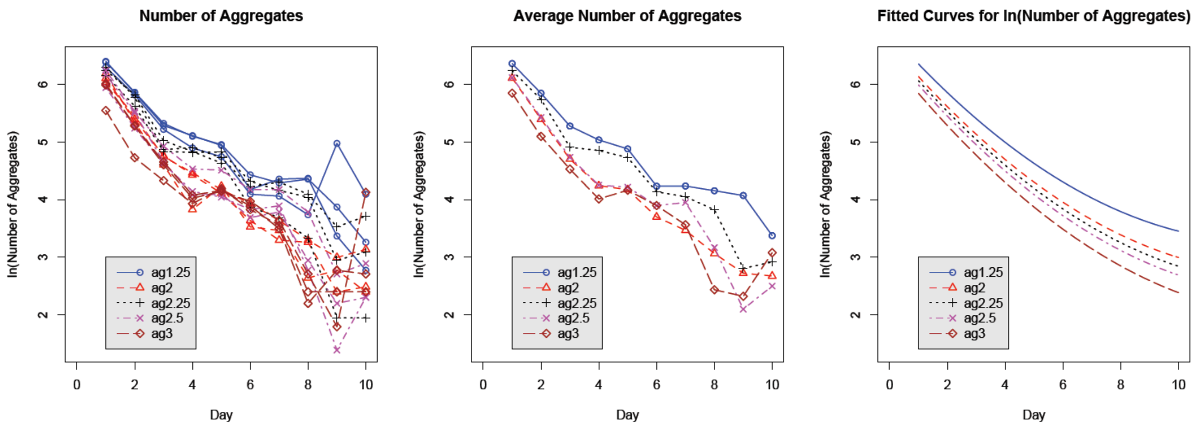

Figure 5.6: The stiffness groups are referred to by their agarose content (ag). Left panel: the natural logarithm (ln) of the number of aggregates as a function of time for each sample for days 1 - 10. Middle panel: averages were taken in each stiffness group (3 samples per group). Right panel: the fitted curves from the model for each of the groups.

Average Coverage Area The average coverage area of aggregates per sample (`AC`) increased over time, and faster so on stiffer gels. The natural logarithm of `AC` was taken to linearize the data. The following least squares regression model was fitted over days 1-8 (see also Figure 5.7 and Table 5.3):

$$\ln(\mathtt{AC}) = 7.061 + 0.269\,\mathtt{D} + 0.023\,\mathtt{S}^2 \cdot \mathtt{D}.$$

The errors are assumed to be independent and identically normally distributed. Note that again the assumption of equal variance appears to be violated. On the other hand, the errors appear to be approximately normally distributed. Overall, the model describes the data quite well, with $R^2 = 0.81$. The significant interaction term ($\mathtt{S}^2 \cdot \mathtt{D}$) indicates, as mentioned above, that the average coverage area per aggregate increases over time in a manner significantly different for different stiffness cases.

We also considered the increase of $\ln(\mathtt{AC})$ from day 6-8 (Figure 5.8). As discussed earlier, when considering the `COT` data, this is an important time period since the aggregates are not disturbed by media changes. The plot shows a faster increase of the average coverage area, giving evidence that aggregate formation might be more prominent in stiffer gels when not disturbed by media changes.

Table 5.3: Estimates of the coefficients of the model for ln(`AC`) are shown, along with the corresponding standard errors, t-statistics, and p-values.

	Estimate	Std Error	t-statistic	p-value
Intercept	7.061	0.09093	77.654	$< 2 \times 10^{-16}$
Day	0.269	0.02490	10.787	$< 2 \times 10^{-16}$
Stiffness^2 · Day	0.023	0.00332	6.958	2.13×10^{-10}

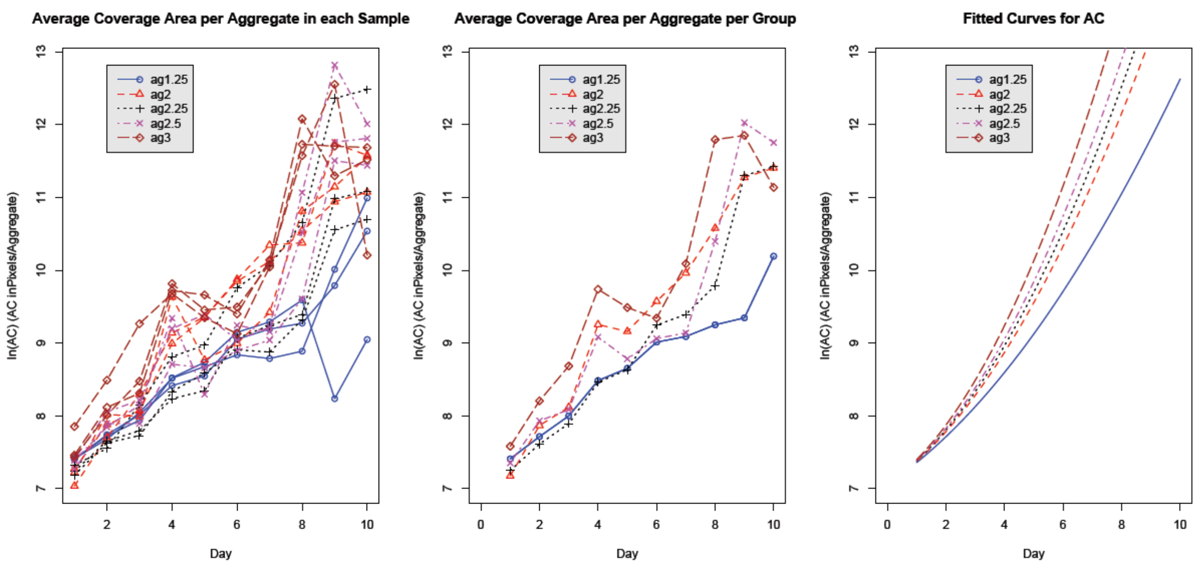

Figure 5.7: The stiffness groups are referred to by their agarose content (ag). Left panel: the natural logarithm of the average coverage area of aggregates (aggregate size) as a function of time for each sample for days 1 to 8. Middle panel: averages of the logarithms were taken in each stiffness group. Right panel: the fitted curves from the model for each of the groups.

Aggregate Perimeter For each sample, the mean aggregate perimeter (`MP`) was calculated. The `MP` was shown to increase over time, and quicker so on stiffer gels. In Figure 5.9 the natural logarithms of the mean perimeters are plotted. The following least squares

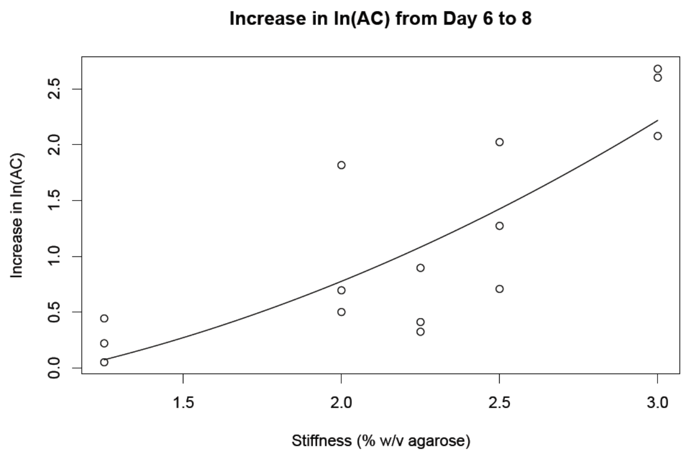

Figure 5.8: Quadratic model demonstrating the increase in $\ln(\mathtt{AC})$ as a function of agarose content from day 6 to 8, during which time culture medium was not exchanged. Multiple $R^2 = 0.67$.

regression model was fitted over days 1-8 (see also Table 5.4):

$$\ln(\mathtt{MP}) = 4.963 + 0.121\,\mathtt{D} + 0.032\,\mathtt{S}\cdot\mathtt{D}.$$

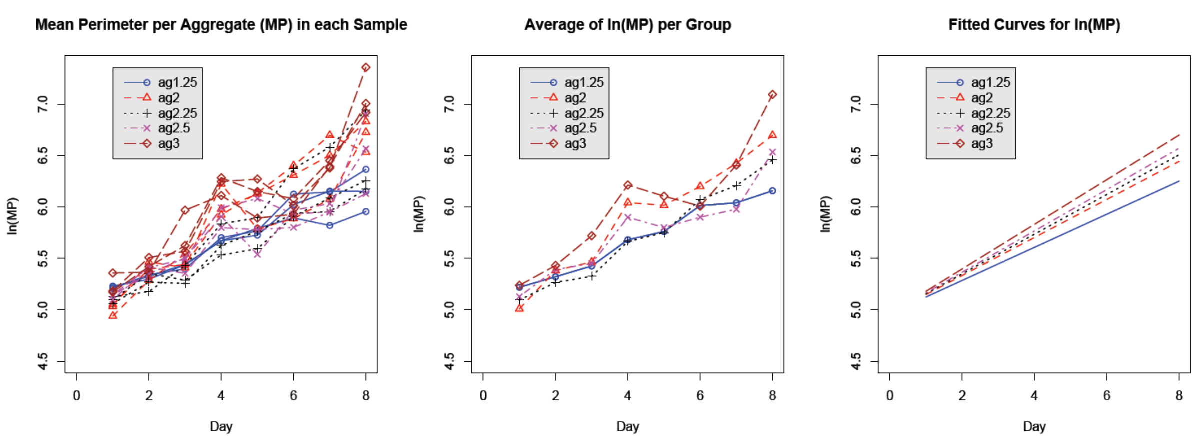

Figure 5.9: The stiffness groups are referred to by their agarose content (ag). Left panel: the natural logarithm of the mean aggregate perimeter, $\ln(\mathtt{MP})$, as a function of time for each sample for days 1-8. Middle panel: averages of the logarithms were taken in each stiffness group. Right panel: the fitted curves from the model for each of the groups.

Table 5.4: Estimates of the coefficients of the model for ln(MP) are shown, along with the corresponding standard errors, t-statistics, and p-values.

	Estimate	Std Error	t-statistic	p-value
Intercept	4.963	0.04442	111.720	$< 2 \times 10^{-16}$
Day	0.121	0.01754	6.914	2.65×10^{-10}
Stiffness · Day	0.032	0.00690	4.632	9.45×10^{-06}

5.4 Discussion

A primary objective of this work was to develop statistical models correlating the behavior and development of cancer cells in response to culture substrate stiffness, with the long-term goal of developing tunable *in vitro* breast cancer test systems. Previous work has shown that in two-dimensional cultures, surface stiffness influences cell morphology ("round" cells on compliant surfaces and "spread" cells on stiff surfaces) [13], migration direction (towards a higher surface stiffness [26]), migration speed [40], and cancer development [27]. However, this influence of hydrogel stiffness on cell behavior does not fully translate to 3D substrates. In our 3D metabolic activity studies, MCF-7 cells cultured within weak hydrogels formed aggregates whereas those in strong hydrogels remained round and in single cell distribution over 14 days of culture. Furthermore, cell metabolic activity tests (glucose and lactic acid measurements and Alamar Blue assay) indicated that cells in strong gels were nearly or completely metabolically inactive (not statistically significantly different from acellular controls), even though Live/Dead assay results demonstrated that cells were alive. Since the agarose component of the hydrogel composite is not degradable by matrix metalloproteinases, it is possible that pore size and proteolysis susceptibility was sufficiently small to limit cell motility, effectively locking cells in place in the stronger gel. Note that, work by Rolli and coworkers [36] suggests that pore sizes of less than roughly 10 μm will inhibit cell migration.

While the hydrogel system employed in the morphology experiment was not a true 3D model, as cells were seeded in a monolayer on top of a hydrogel, the hydrogel substrate selected permits independent control of surface rigidity and cell-adhesion ligand density. Furthermore, while the aqueous hydrogel substrate mimics the natural ECM more closely than does a rigid, flat, anhydrous surface, the system also allows free cell migration and proliferation, independent of substrate pore size and proteolysis susceptibility, similar to that of a two-dimensional system. To assess the feasibility of modeling cancer cell position and morphological parameters in a 3D system, a simplified 2.5D system was used instead. This approach improved the ability to characterize the system while still maintaining biomimetic properties such as the gel-like, fibrillar nature of natural ECM.

The regression models in the morphology study suggest that the coverage area (over threshold) increases over time, and faster so on stiffer gels. This was particularly true for the model of the increase in coverage area over threshold from day 6-8. Moreover, the number of aggregates decreased over time, and there was some evidence that the decrease was faster on stiffer gels. These models (for coverage area and number of aggregates) suggest that cell aggregate formation is more pronounced on stiffer gels. Further evidence for this was obtained from a regression model that showed that the average aggregate size increased over time faster on stiffer gels. Note that this is in contrast to the metabolic activity and viability study, in which aggregate formation appeared to be stronger in weak gels. Evidence for this can be seen in Figure 5.3, where aggregates formed only in weak gels. Additionally, higher metabolic activity (higher lactate production and higher absorbance values for the Alamar Blue assay) were observed for cells in the weak gels. This higher metabolic activity is usually seen in aggregated cells rather than single cells. The contrasting evidence of the effect of stiffness on aggregate formation might be due to the different experimental setups, indicating that not only stiffness is an important factor for the degree of aggregate formation, but also the composition of the gel (agarose/collagen in the metabolic activity experiment versus agarose/gelatin in the morphology experiment) and the configuration of cells within the gel (three dimensional distribution of cells in the metabolic activity experiment versus a monolayer in the morphology experiment).

The model for the average aggregate perimeter suggests that the average perimeters increase over time, and faster so on stiffer gels. A larger aggregate perimeter could mean that the aggregate is larger and/or that it is more irregularly shaped. Since the coverage area is increasing faster on stiffer gels (in the *Cell Morphology Analysis* experiment), a future question is whether it is possible to separate the effects of larger aggregates and more irregularly shaped aggregates on the perimeter.

The enormous potential of a 3D test system is in the ability to construct cells in a biomaterial substrate in a spatially meaningful manner that allows cellular behavior that is more indicative of behavior in native tissue than a 2D system, thus allowing rapid discoveries of therapies and preventatives for an array of diseases. This proof-of-principle work has demonstrated the possibility of using image processing and statistical modeling to describe pseudo-3D cancer systems in a non-destructive and spatio-temporal manner. While substrate stiffness alone in a 2D substrate would often dictate greater cell proliferation, cancer cells were found to form colonies only in “weak” hydrogels in the 3D metabolic activity and viability experiment – this indicates the importance of considering additional factors which influence cell behavior in a 3D system, such as proteolysis susceptibility and pore size.

Chapter 6

Appendix

Here we collect some results used in Chapter 2.

Lemma 6.0.1. *Let* $\mathbb{S}_n^k = \{\mathbf{x} \in \mathbb{R}_+^k | n\mathbf{x} \in \mathbb{N}_0^k\}$ *and* $\{X_t\}_{t\in\mathbb{R}_+}$ *be an* $\mathbb{S}_n^k$ *valued Markov process with* $\mathbf{0}$ *as an absorbing state, such that* $X_t = X_{\lfloor nt\rfloor/n}$, $t \geq 0$. *Suppose for some* $\nu \in \mathcal{P}(\mathbb{S}_n^k)$, $P_{\mathbf{y}}(X_t \in \cdot | X_t \neq \mathbf{0})$ *converges weakly, as* $t \to \infty$, *to* ν *for all* $\mathbf{y} \in \mathbb{S}_n^k$. *Then* ν *is a qsd for* $\{X_t\}_{t\in\mathbb{R}_+}$.

Proof. We need to show that for each $A \subseteq \mathbb{S}_n^k$ and $t \geq 0$

$$P_\nu(X_t \in A | X_t \neq \mathbf{0}) = \nu(A). \tag{6.0.1}$$

The left hand side of (6.0.1) equals

$$\frac{P_\nu(X_t \in A, X_t \neq \mathbf{0})}{P_\nu(X_t \neq \mathbf{0})}. \tag{6.0.2}$$

Letting $A^\circ := A \setminus \{\mathbf{0}\}$, and denoting the measure $P_{\mathbf{y}}(X_t \in \cdot | X_t \neq \mathbf{0})$ by $\nu_t^{\mathbf{y}}$,

$$
\begin{aligned}
P_\nu(X_t \in A, X_t \neq \mathbf{0}) &= \int P(X_t \in A^\circ | X_0 = \mathbf{x}) \nu(d\mathbf{x}) \\
&= \lim_{\substack{s\to\infty \\ s\in\mathbb{S}_n}} \int P(X_t \in A^\circ | X_0 = \mathbf{x}) \nu_s^{\mathbf{y}}(d\mathbf{x}) = \lim_{\substack{s\to\infty \\ s\in\mathbb{S}_n}} \int P(X_{t+s} \in A^\circ | X_s = \mathbf{x}) \nu_s^{\mathbf{y}}(d\mathbf{x}) \\
&= \lim_{\substack{s\to\infty \\ s\in\mathbb{S}_n}} \frac{1}{P_{\mathbf{y}}(X_s \neq \mathbf{0})} P_{\mathbf{y}}(X_{t+s} \in A, X_{t+s} \neq \mathbf{0}),
\end{aligned}
$$

where $\mathbb{S}_n = \{\frac{j}{n} | j \in \mathbb{N}_0\}$, the second equality follows from the assumption in the lemma while the third and fourth use the Markov property of X and the observation that $P(X_{t+s} \in A^\circ | X_s = \mathbf{0}) = 0$.

Setting $A = \mathbb{S}_n$, we have

$$
P_\nu(X_t \neq \mathbf{0}) = \lim_{\substack{s\to\infty \\ s\in\mathbb{S}_n}} \frac{1}{P_{\mathbf{y}}(X_s \neq \mathbf{0})} P_{\mathbf{y}}(X_{t+s} \neq \mathbf{0}).
$$

Combining the above, we have

$$
\begin{aligned}
P_\nu(X_t \in A | X_t \neq \mathbf{0}) &= \lim_{\substack{s\to\infty \\ s\in\mathbb{S}_n}} \frac{P_{\mathbf{y}}(X_{t+s} \in A, X_{t+s} \neq \mathbf{0})}{P_{\mathbf{y}}(X_{t+s} \neq \mathbf{0})} \\
&= \lim_{\substack{s\to\infty \\ s\in\mathbb{S}_n}} P_{\mathbf{y}}(X_{t+s} \in A | X_{t+s} \neq \mathbf{0}) = \nu(A),
\end{aligned}
$$

from which the result follows. □

Lemma 6.0.2. *Let $m \in (0,1)$ and $\sigma^2 \in (0,\infty)$. Then there exists a pgf f of linear fractional form,*

$$
f(s) = 1 - \frac{b}{1-p} + \frac{bs}{1-ps}, \quad s \in [0,1], \tag{6.0.3}
$$

with $b, p \in (0,1)$, $b < 1-p$, such that the corresponding probability distribution has mean

m and variance σ^2. Specifically,

$$p = \frac{\sigma^2/m + m - 1}{2 + \sigma^2/m + m - 1} \quad \textit{and} \quad b = m\left(1 - \frac{\sigma^2/m + m - 1}{2 + \sigma^2/m + m - 1}\right)^2. \tag{6.0.4}$$

Proof. Fix $b, p \in (0,1)$, $b < 1 - p$. Define f by (6.0.3). Then $f'(s) = \frac{b}{(1-ps)^2}$ and $f''(s) = \frac{2bp}{(1-ps)^3}$. The mean $\bar{m}$ and the variance $\bar{\sigma}^2$ of the probability distribution corresponding to f is given as $\bar{m} = f'(1) = \frac{b}{(1-p)^2}$ and

$$\begin{aligned} \bar{\sigma}^2 &= f''(1) - [f'(1)]^2 + [f'(1)] = \frac{2bp(1-p) - b^2 + b(1-p)^2}{(1-p)^4} \\ &= \bar{m}\left(2\frac{p}{(1-p)} - \bar{m} + 1\right). \end{aligned}$$

Solving the last two equations for p and b, we get (6.0.4) with $m = \bar{m}$ and $\sigma^2 = \bar{\sigma}^2$. $\square$

Recall the notation introduced below Theorem 2.1.4.

Lemma 6.0.3. *Assume that $Z_0^{(n)}$ has distribution μ (supported on $\mathbb{N}$). Then there exist a probability measure $P_\mu^{(n)\uparrow}$ on $(\hat{\Omega}, \hat{\mathcal{F}})$ such that as $s \to \infty$*

$$\hat{P}_\mu^{(n)}(\Theta | T > s) \to P_\mu^{(n)\uparrow}(\Theta), \quad \textit{for all } \Theta \in \mathcal{F}_t, \quad t \in \mathbb{R}_+.$$

Furthermore if $\{Z_k^{(n)\uparrow}\}_{k\in\mathbb{N}_0}$ is a Markov chain with state space $\mathbb{N}$, l-step transition function

$$p_l^{(n)\uparrow}(i,j) = P(Z_l^{(n)} = j | Z_0^{(n)} = i)\frac{j}{i} m_n^{-l}, \tag{6.0.5}$$

and initial distribution μ, then $P_\mu^{(n)\uparrow}$ is the law of $\{X_t^{(n)\uparrow}\}_{t\in\mathbb{R}_+}$, where

$$X_t^{(n)\uparrow} := \frac{1}{n} Z_{\lfloor nt \rfloor}^{(n)\uparrow}, \quad t \in \mathbb{R}_+.$$

Proof. The proof is along the lines of [1], p. 58. Fix $\alpha \in \mathbb{N}$ and $0 \le t_1 < \cdots < t_\alpha < t_\alpha + s$. Let $k_l = \lfloor nt_l \rfloor$, $l = 1, \ldots, \alpha$, and $\tilde{k} = \lfloor n(t_\alpha + s) \rfloor$. First assume $Z_0 = i$. Then

$$\begin{aligned}
&P_{i/n}\left(X_{t_1}^{(n)} = \frac{i_1}{n}, \ldots, X_{t_\alpha}^{(n)} = \frac{i_\alpha}{n} | t_\alpha + s < T_{X^{(n)}} < \infty\right) \\
&\quad = P_i(Z_{k_1}^{(n)} = i_1, \ldots, Z_{k_\alpha}^{(n)} = i_\alpha | \tilde{k} < T_{Z^{(n)}} < \infty) \\
&\quad = P_i(Z_{k_1}^{(n)} = i_1, \ldots, Z_{k_\alpha}^{(n)} = i_\alpha) \frac{\sum_{j=1}^\infty P_{\tilde{k}-k_\alpha}(i_\alpha, j)}{\sum_{j=1}^\infty P_{\tilde{k}}(i, j)} \\
&\quad = P_i(Z_{k_1}^{(n)} = i_1, \ldots, Z_{k_\alpha}^{(n)} = i_\alpha) \frac{P_{\tilde{k}-k_\alpha}(1,1)}{P_{\tilde{k}}(1,1)} \frac{\sum_{j=1}^\infty \frac{P_{\tilde{k}-k_\alpha}(i_\alpha, j)}{P_{\tilde{k}-k_\alpha}(1,1)}}{\sum_{j=1}^\infty \frac{P_{\tilde{k}}(i,j)}{P_{\tilde{k}}(1,1)}}. \qquad (6.0.6)
\end{aligned}$$

Using Theorem I.7.4 of [1], we get that the right hand side of (6.0.6) converges, as $\tilde{k} \to \infty$, to

$$P_{i/n}\left(X_{t_1}^{(n)} = \frac{i_1}{n}, \ldots, X_{t_\alpha}^{(n)} = \frac{i_\alpha}{n}\right) m^{-k_\alpha} \frac{i_\alpha}{i} =: P_{i/n}^{(n)\uparrow}\left(\pi_{t_1} = \frac{i_1}{n}, \ldots, \pi_{t_\alpha} = \frac{i_\alpha}{n}\right),$$

where $\pi_t(x) = x_t$ for $x \in \hat{\Omega}$ and $t \in \mathbb{R}_+$. The right hand side of the last display determines a probability measure $P_i^{(n)\uparrow}$ on $\bigcup_{t>0} \sigma\{\pi_t\}$, which extends uniquely to a measure $P_i^{(n)\uparrow}$ on $\hat{\mathcal{F}}$. The measure $P_\mu^{(n)\uparrow}$ on $(\hat{\Omega}, \hat{\mathcal{F}})$ for a general initial distribution μ of $Z_0^{(n)}$ is defined as $\sum_{i=1}^\infty \mu(i) P_i^{(n)\uparrow}$. Let $Z^{(n)\uparrow}$ be a Markov chain on a probability space $(\tilde{\Omega}^{(n)}, \tilde{\mathcal{F}}^{(n)}, \tilde{P}^{(n)})$ as in the statement of the lemma. Then

$$\tilde{P}^{(n)}\left(X_{t_1}^{(n)\uparrow} = \frac{i_1}{n}, \ldots, X_{t_\alpha}^{(n)\uparrow} = \frac{i_\alpha}{n}\right) = \sum_{i=1}^\infty P_{i/n}^{(n)\uparrow}\left(\pi_{t_1} = \frac{i_1}{n}, \ldots, \pi_{t_\alpha} = \frac{i_\alpha}{n}\right) \mu(i),$$

which implies that $P_\mu^{(n)\uparrow}$ is the law of $X^{(n)\uparrow}$. □

Lemma 6.0.4. *Assume Conditions 2.1.13, 2.1.14, 2.1.15. Let Q_n be as introduced above (2.3.7). Then $Q_n \to Q$.*

Proof. For $l = 1, \ldots, k$ and $n \in \mathbb{N}$, let $\{\gamma_{l,j}^{(n)}\}_{1 \le j \le k}$ denote a random variable representing the offspring count (in a single generation) of a particle of type l for the n-th BGW process. Then, since $\mathbf{u}^{(n)} = (u_1^{(n)}, \ldots, u_k^{(n)})'$ and $\mathbf{v}^{(n)} = (v_1^{(n)}, \ldots, v_k^{(n)})'$ are the right and left eigenvectors of $\mathbf{M}^{(n)}$, and $m_{l,j}^{(n)} = E(\gamma_{l,j}^{(n)})$, we get

$$
\begin{aligned}
\sum_{l=1}^{k} v_l^{(n)} \mathbf{u}^{(n)\prime} \boldsymbol{\sigma}^{(n)}(l) \mathbf{u}^{(n)} - 2Q_n &= \sum_{l=1}^{k} v_l^{(n)} \mathbf{u}^{(n)\prime} \boldsymbol{\sigma}^{(n)}(l) \mathbf{u}^{(n)} - \sum_{l=1}^{k} v_l^{(n)} q_{n,i}[\mathbf{u}^{(n)}] \\
&= \sum_{l=1}^{k} v_l^{(n)} \left(\sum_{i,j=1}^{k} u_i^{(n)} E(\gamma_{l,i}^{(n)} \gamma_{l,j}^{(n)}) u_j^{(n)} - \sum_{i,j=1}^{k} u_i^{(n)} E(\gamma_{l,i}^{(n)}) E(\gamma_{l,j}^{(n)}) u_j^{(n)} \right) \\
&\quad - \sum_{l=1}^{k} v_l^{(n)} \left(\sum_{i,j=1}^{k} u_i^{(n)} E(\gamma_{l,i}^{(n)} \gamma_{l,j}^{(n)}) u_j^{(n)} - \sum_{i=1}^{k} \left(u_i^{(n)} \right)^2 E(\gamma_{l,i}^{(n)}) \right) \\
&= \sum_{i=1}^{k} \left(u_i^{(n)} \right)^2 \left(1 + \frac{c_n}{n} \right) v_i^{(n)} - \sum_{l=1}^{k} v_l^{(n)} \left(1 + \frac{c_n}{n} \right)^2 \left(u_l^{(n)} \right)^2 \\
&= -\frac{c_n}{n} \left(1 + \frac{c_n}{n} \right) \sum_{i=1}^{k} \left(u_i^{(n)} \right)^2 v_i^{(n)}.
\end{aligned}
$$

The result follows on sending $n \to \infty$ in the above display. $\square$

Bibliography

[1] K. B. Athreya and P. Ney. *Branching Processes*. Springer-Verlag, 1972.

[2] E. Bland, D. Reinhold, A. Budhiraja, M. R. Leadbetter, and K. J. L. Burg. Modeling cancer cell behavior as a function of substrate stiffness in 3D test systems. Under preparation.

[3] A. Budhiraja. An ergodic control problem for constrained diffusion processes: Existence of optimal markov control. *SIAM J. Control. Optim.*, Vol. 42, No. 2:532–558, 2003.

[4] A. Budhiraja and C. Lee. Stationary distribution convergence for generalized jackson networks in heavy traffic. *Mathematics of Operations Research*, Vol. 34, No. 1:45–56, 2009.

[5] A. Budhiraja and D. Reinhold. Some asymptotic results for near critical branching processes. *Communications on Stochastic Analysis*, Vol. 4, No. 1:91–113, 2010.

[6] T. Burg, C. A. Cass, R. Groff, M. Pepper, and K. J. Burg. Building off-the-shelf tissue-engineered composites. *Philos Transact A Math Phys Eng Sci*, 368(1917):1839–1862, 2010.

[7] M. A. Cifone. In vitro growth characteristics associated with benign and metastatic variants of tumor cells. *Cancer Metastasis Rev*, 1(4):335–347, 1982.

[8] J. G. Dai and S. P. Meyn. Stability and convergence of moments for multiclass queueing networks via fluid limit models. *IEEE Transactions on Automatic Control*, 40:1889–1904, 1995.

[9] J. Dieudonne. *Foundations of Modern Analysis*. READ BOOKS, 2006.

[10] S. N. Ethier and T. G. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley, 1986.

[11] K. S. Fahady, M. P. Quine, and D. Vere-Jones. Heavy traffic approximation for the Galton-Watson process. *Advances in Applied Probability*, 3:282–300, 1971.

[12] W. Feller. Diffusion processes in genetics. *Proc. Second Berkeley Symp. on Math. Statist. and Prob.*, pages 227–246, 1951.

[13] P. Friedl and K. Wolf. Plasticity of cell migration: a multiscale tuning model. *J Cell Biol*, 188(1):11–19, 2010.

[14] A. Greven, L. Popovic, and A. Winter. Genealogy of catalytic branching models. *The Annals of Applied Probability*, 19:1232–1272, 2009.

[15] P. Haccou, P. Jagers, and V. A. Vatutin. *Branching processes: variation, growth, and extinction of populations.* Cambridge University Press, 2005.

[16] T. E. Harris. *The Theory of Branching Processes.* Courier Dover Publications, 2002.

[17] C. P. Huang, J. Lu, H. Seon, A. P. Lee, L. A. Flanagan, H. Y. Kim, A. J. Putnam, and N. L. Jeon. Engineering microscale cellular niches for three-dimensional multicellular co-cultures. *Lab Chip*, 9(12):1740–1748, 2009.

[18] J. Jacod and A. N. Shiryaev. *Limit Theorems for Stochastic Processes.* Springer-Verlag, 1987.

[19] R. Järvinen, A. Kaasinen, A. Sankila, and E. Rintala. Long-term efficacy of maintenance bacillus Calmette-Guérin versus maintenance mitomycin C instillation therapy in frequently recurrent TaT1 tumours without carcinoma in situ: A subgroup analysis of the prospective, randomised finnBladder I study with a 20-year follow-up. *European Urology*, 56 (2):260–265, 2009.

[20] A. Joffe and M. Metivier. Weak convergence of sequences of semimartingales with applications to multitype branching processes. *Advances in Applied Probability*, Vol. 18:pp. 20–65, 1986.

[21] H.-W. Kang and T. G. Kurtz. Separation of time-scales and model reduction for stochastic reaction networks. *Submitted*.

[22] T. G. Kurtz. A control formulation for constrained markov processes. In *Mathematics of Random Media, Lectures in Appl. Math., 27, AMS, Providence, RI*, pages 139–150, 1991.

[23] A. Lambert. Quasi-stationary distributions and the continuous branching process conditioned to be never extinct. *Electronic Journal of Probability*, 12:420–446, 2007.

[24] A. Lambert. Population dynamics and random genealogies. *Stochastic Models*, 24:45–163, 2008.

[25] Z. Li. A limit theorem for discrete Galton-Watson branching processes with immigration. *Journal of Applied Probability*, 43:289–295, 2006.

[26] C. M. Lo, H. B. Wang, M. Dembo, and Y. L. Wang. Cell movement is guided by the rigidity of the substrate. *Biophys J*, 79(1):144–152, 2000.

[27] M. Makale. Cellular mechanobiology and cancer metastasis. *Birth Defects Res C Embryo Today*, 81(4):329–343, 2007.

[28] S. P. Meyn and R. L. Tweedie. *Markov Chains and Stohastic Stability.* Springer-Verlag, 1993.

[29] M. R. Ng and J. S. Brugge. A stiff blow from the stroma: collagen crosslinking drives tumor progression. *Cancer Cell*, 16(6):455–457, 2009.

[30] C. K. Osborne, D. D. V. Hoff, and K. Mullins. Endocrine therapy testing of human breast cancers in the soft agar clonogenic assay. *Breast Cancer Res Treat*, 6(3):229–235, 1985.

[31] M. J. Paszek, N. Zahir, K. Johnson, J. N. Lakins, G. I. Rozenberg, A. Gefen, C. A. Reinhart-King, S. S. Margulies, M. Dembo, D. Boettiger, D. A. Hammer, and V. M. Weaver. Tensional homeostasis and the malignant phenotype. *Cancer Cell*, 8(3):241–254, 2005.

[32] P. E. Protter. *Stochastic Integration and Differential Equations*. Springer-Verlag, 2005.

[33] P. P. Provenzano, D. R. Inman, K. W. Eliceiri, J. G. Knittel, L. Yan, C. T. Rueden, J. G. White, and P. J. Keely. Collagen density promotes mammary tumor initiation and progression. *BMC Medicine*, 6:11, 2008.

[34] M. P. Quine. The multi-type Galton-Watson process with immigration. *Journal of Applied Probability*, 7:411–422, 1970.

[35] M. P. Quine. The multitype Galton-Watson process with ρ near 1. *Advances in Applied Probability*, 4:pp.429–452, 1972.

[36] C. G. Rolli, T. Seufferlein, R. Kemkemer, and J. P. Spatz. Impact of tumor cell cytoskeleton organization on invasiveness and migration: a microchannel-based approach. *PLoS One*, 5(1):e8726, 2010.

[37] A. Shwartz and A. Weiss. *Large Deviations For Performance Analysis*. Chapman & Hall, 1995.

[38] A. V. Skorokhod. *Asymptotic Methods in the Theory of Stochastic Differential Equations*. American Mathematical Society, 1989.

[39] D. Stroock and S. R. S. Varadhan. *Multidimensional Diffusion Processes*. Springer-Verlag, 1979.

[40] T. Tzvetkova-Chevolleau, A. Stéphanou, D. Fuard, J. Ohayon, P. Schiavone, and P. Tracqui. The motility of normal and cancer cells in response to the combined influence of the substrate rigidity and anisotropic microstructure. *Biomaterials*, 29(10):1541–1551, 2008.

[41] A. Weiss. *Invariant Measures of Diffusions in Bounded Domains*. PhD thesis, New York University, New York, 1981.

[42] F. Xu and K. J. Burg. Three-dimensional polymeric systems for cancer cell studies. *Cytotechnology*, 54(3):135–143, 2007.

[43] M. H. Zaman, L. M. Trapani, A. L. Sieminski, D. Mackellar, H. Gong, R. D. Kamm, A. Wells, D. A. Lauffenburger, and P. Matsudaira. Migration of tumor cells in 3D matrices is governed by matrix stiffness along with cell-matrix adhesion and proteolysis. *Proc Natl Acad Sci U S A*, 103(29):10889–10894, 2006.