THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Two-Sex Branching Populations

Ali Falahati

Department of Mathematical Statistics
CHALMERS UNIVERSITY OF TECHNOLOGY
AND GÖTEBORG UNIVERSITY
Göteborg, Sweden 1999

Two-Sex Branching Populations
Ali Falahati
ISBN 91-7197-788-0

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie nr 1493
ISSN 0346-718x

Department of Mathematical Statistics
Chalmers University of Technology and Göteborg University
SE-412 96 Göteborg
Sweden
Telephone +46 (0)31-772 1000

This work is part of the Bank of Sweden
Tercentenary Foundation project Dependence and
Interaction in Stochastic Population Dynamics

To
my mother and father,
Jelveh and *Jafar*,
with love.

Acknowledgements

The completion of this thesis gives me the opportunity to express deep gratitude to my advisor, Professor Peter Jagers, for introducing me to the area of branching processes, for his constant encouragement and support. My warm thanks go to Serik Sagitov, for inspiring discussions and many helpful suggestions which lead to completion of the thesis. I am also very grateful to Ziad Taib who always welcome me and my questions and being a friend to me. I am indebted to these people who spent a lot of their time to read the thesis and gave me valuable comments which always lead to improvement of the work.

I also would like to thanks my dear friend, Nibia Aires, for reading the Introduction and her valuable comments on it.

I would like to extend my thanks to Lena Lööv, for being very kind, understanding and patient with me in the last two years which I was assisting her in the library.

Also, I wish to express my sincere appreciation to all my friends and colleagues at the Department of Mathematics at Chalmers University of Technology and the University of Göteborg for creating an inspiring and friendly research environment.

It is a great pleasure to express gratitude to all my friends, in particular my dear Lyudmila Turowska, Sander van den Hark, Daniel Fridlender, Mia Indrika, Ilona Szarvadi, Boris Stoew and Per-Gösta Andersson, for their encouragements and supports and for being always there for me.

Finally, and most of all, I wish to thank from the bottom of my heart my dearest mother, father and my sisters who always love me and believe in me. An especial thank to my dear sister and my best friend Sorour who supports and backs me up in whatever I do.

Ali Falahati,
Göteborg, March 1999

Contents

Chapter 1

Introduction

Understanding the dynamics of populations has been the subject of many scientific investigations in the past. The theory of branching processes is a tool for this and can be used to study, among other things, the extinction, growth and composition of a population and its long-run behavior. It emphasises the role of an individual in determining the fate of a population. This theory has been extensively studied for asexual populations, i.e. where we have only one sex in the population such as cell populations.

In comparison with asexual populations there has not been much work on two-sex models. Apart from the additional technical difficulties which arise when dealing with two-sex populations and hinder the development in this field, the sexual interaction between individual makes a conceptual difference from the classical case of independently reproducing individual.

In this work, we introduce a model, a Two-Sex Galton-Watson Branching Process, to study the dynamics of two-sex branching populations. This model allows individuals to interact with each other in order to form couples (or mating units), and each of these couples produces children for the new generation. In Section 1.1 we give a short summary of other models used to study two-sex branching populations. In Sections 1.2 and 1.3 we present an overview of our models and highlight our main results concerning these models.

1.1. Previous Models for Two-Sex Branching Populations

The classical Galton-Watson branching process considers a population of females which evolves in the following manner. Start the population at generation zero with a number of females, F_0. Independently of each other and according to a common (and fixed) distribution, say G_0, each of them gives birth to a number of daughters and at the same time dies. If $X_{1,j}$ denotes the number of daughters of the jth female, for $j = 1, 2, \ldots, F_0$, the size of the female population in the first generation will be equal to $F_1 = \sum_{j=1}^{F_0} X_{1,j}$. In the same way, if F_n denotes the size of the female population in generation n, for $n = 0, 1, 2, \ldots$ and $X_{n,j}$ denotes the number of daughters of the jth female, for $j = 1, 2, \ldots, F_{n-1}$, we have

$$F_n = \sum_{j=1}^{F_{n-1}} X_{n,j}, \tag{1.1.1}$$

for $n = 1, 2, \ldots$.

The theory of branching processes studies the growth and the extinction of a population. Most natural populations consist of two sexes, female and male. Females have the potential of giving birth to children (of both sexes), whereas males are unable to reproduce but however their existence is necessary for the survival of a population. However, they do not play any role in the classical Galton-Watson process. In most studies the population has been treated as asexual.

Daley [7] presented a model for bisexual populations, the so-called *Bisexual* Galton-Watson Branching Process (**B-GWBP** for short). This model has been the center of most recent studies on bisexual populations. In it, he considers a population which consists of both females and males. These combine into mating units (or couples) who give birth to the next generation. To be more precise, in Daley's model a population and its evolution are determined by two characteristics:

(i) A fixed distribution, say G on $\mathcal{Z}_+^2$, which is the reproduction law of each mating unit (or couple).

(ii) A fixed non-negative, integer-valued function ζ on $\mathcal{Z}_+^2$ which determines the number of mating units of a generation, given the total numbers of females and males in that generation.

Let F_n and M_n denote the numbers of females and males in the nth generation, for $n = 0, 1, 2, \ldots$. Start a population by F_0 females and M_0 males. The total number of mating units formed by these will be equal to $C_0 = \zeta(F_0, M_0)$. Then, each of these mating units independently of each other, produces daughters and sons, according to the fixed distribution G. Hence, if $X_{1,j}$ and $Y_{1,j}$ denote the numbers of daughters and sons of the jth mating unit (of the zero generation), for $j = 1, 2, \ldots, C_0$, then the total number of females and males in the first generation is $(F_1, M_1) = \sum_{j=1}^{C_0} (X_{1,j}, Y_{1,j})$.

This model assumes non-overlapping generations, i.e. an individual born in a generation, will die at the beginning of the next generation. We can define $X_{n,j}$, $Y_{n,j}$ and C_n similarly, for the nth generation, as

$$\begin{aligned}
&\{(X_{n,j}, Y_{n,j})\}_{j=1}^{C_{n-1}} \text{ are i.i.d. and } (X_{n,1}, Y_{n,1}) \sim G, \\
&(F_n, M_n) = \textstyle\sum_{j=1}^{C_{n-1}} (X_{n,j}, Y_{n,j}), \qquad (1.1.2) \\
&\text{and } C_n = \zeta(F_n, M_n),
\end{aligned}$$

for $n = 1, 2, \ldots$.

With x denoting the number of females and y the number of males, examples of mating functions could be:

(i) $\zeta(x, y) = x \min(1, y)$, promiscuous mating,

(ii) $\zeta(x, y) = \min(x, y)$, mating with fidelity,

(iii) $\zeta(x, y) = \min(x, dy)$, polygamous mating with perfect fidelity (d is a positive integer),

(iv) and $\zeta(x, y) = x$, mating function for the standard Galton-Watson branching process.

The extinction probability, i.e. $P(\lim_{n\to\infty} C_n = 0)$, for some simple B-GWBP models has been studied by Daley [7]. Hull [15] drew attention to the class of mating functions which are superadditive, i.e.

$$\zeta(x_1 + x_2, y_1 + y_2) \geq \zeta(x_1, y_1) + \zeta(x_2, y_2) \qquad (1.1.3)$$

for $(x_i, y_i) \in \mathcal{Z}_+^2$, $i = 1, 2$. Since then, B-GWBP with superadditive mating functions have been the subject of many studies. The extinction probability is further investigated in [8] and [1]. The asymptotic behaviour of B-GWBP in the supercritical case has been further studied by [12]-[14].

However, in the bisexual model introduced by Daley [7], the number of mating units (couples) in a generation is determined as soon as we know the total numbers of the females and the males in the population. Asmussen and Hering [3], made a more general approach towards the two-sex model. As in the B-GWBP, individuals develop in two stages:

- the first stage corresponds to reproduction,
- the second stage is the mating stage.

Interaction between individuals (females and males) born in the first stage, results in a number of couples (or mating units) in the second stage. The development of the population is determined by two main characteristics, a fixed distribution, G on $\mathcal{Z}_+^2$, which determines the reproduction law of each mating unit (or couple), and a class of probability distributions say $\mathcal{C} = \{H_{f,m} : (f,m) \in \mathcal{Z}_+^2\}$ on $\mathcal{Z}_+$ ($H_{f,m}$ will be the probability distribution with mass one at zero if $\min(f,m) = 0$). Furthermore, the number of couples in the second stage is a random variable with the probability distribution $H_{f,m} \in \mathcal{C}$, where f and m, as before are the total numbers of females and males in the first stage, respectively. Thus, in this model, a population starts with C_0 couples. Each of these couples, independently of each other, and according to the law G, produces $X_{1,j}$ females and $Y_{1,j}$ males, for $j = 1, 2, \ldots, C_0$. Let $(F_1, M_1) = \sum_{j=1}^{C_0}(X_{1,j}, Y_{1,j})$ be the total number of children. Given F_1 and M_1, the number of couples is a result of the interaction between these F_1 females and M_1 males, i.e. C_1, is a random variable with the distribution $H_{F_1,M_1} \in \mathcal{C}$. Hence for the nth generation we have

$$\begin{aligned} &(F_n, M_n) = \textstyle\sum_{j=1}^{C_{n-1}}(X_{n,j}, Y_{n,j}), \\ \text{and} \quad &C_n \mid (F_n, M_n) \sim H_{F_n,M_n}, \end{aligned} \tag{1.1.4}$$

for $n = 1, 2, \ldots$. Asmussen and Hering [3] also studied the asymptotic behaviour of this model in the supercritical - as well as subcritical case. They showed that in the supercritical case, the process $\{C_n\}_{n=0}^{\infty}$ (after appropriate scaling), converges to a non-degenerate random variable. However, they did not give sufficient conditions for the process to grow to infinity with positive probability.

1.2. Two-Sex Galton-Watson Branching Populations and Summary of the Results

In Chapters 2-3, we present and analyse our main model for studying a bisexual population, which we call a Two-Sex Galton-Watson Branching Process (**2S-GWBP**). In this model, as in Daley [7] and Asmussen and Hering [3], we assume that our two-sex population develops in two stages: first reproduction and then mating. The new feature of our model is that we allow the possibility of using the information which exists during a generation in each of these two stages.

Consider a population starting with C_0 couples (or mating units).

Stage 1: Reproduction. In the first stage each couple produces a number of daughters and sons, $(X_{1,j}, Y_{1,j})$, independently of the others, $j =$

$1, 2, \ldots, C_0$. It is reasonable to assume that the number of couples should have some effect on their reproduction law. This effect might be caused by the availability of resources, food, etc. Hence, for some class of probability distribution, $\mathcal{C}_1 = \{H_z : z \in \mathcal{Z}_+\}$ on $\mathcal{Z}_+^2$, we have

$$\text{for given } C_0, \ (X_{1,1}, Y_{1,1}), \ldots, (X_{1,C_0}, Y_{1,C_0}), \text{ are i.i.d and}$$
$$(X_{1,1}, Y_{1,1}) \mid C_0 \sim H_{C_0}, \text{ where } H_{C_0} \in \mathcal{C}_1. \tag{1.2.1}$$

The dependence of the reproduction law of each couple on the total number of couples present is a new feature.

Stage 2: Mating. Entering the mating stage we have the following information from the previous stage.

(i) We know the number of couples involved in the reproduction stage, i.e. C_0.

(ii) We also know the number of daughters and sons of each couple in the first stage, i.e. $\{(j, (X_{1,j}, Y_{1,j}))\}_{j=1}^{C_0}$.

Call the children of each couple a *family* (their parents have died as soon as they completed their reproduction). Different families can be more or less successful in the mating competition. It might be so that daughters of a family with fewer children are more successful (or on the contrary, they are less successful). Hence, we can assume that the daughters of a family, after knowing how many families are around, i.e. C_0, and the size of each of the families, i.e. $\{(j, (X_{1,j}, Y_{1,j}))\}_{j=1}^{C_0}$, independently of other families (but not necessarily independently of each other) try to marry a number of times. Denote the total number of marriages of the daughters of the jth family by $U_{1,j}$, for $j = 1, 2, \ldots, C_0$. The total number of couples will be the sum of these marriages, i.e. $C_1 = \sum_{j=1}^{C_0} U_{1,j}$.

Comparing this to the models by Daley [7] and Asmussen and Hering [3], one can see that in their models only the total numbers of females and males, play a determining role in the mating stage, while our model uses more information at that stage.

The information obtained at the first stage can be used in the second stage in many different ways. In Sections 1.2.1 and 1.2.2 we give overviews of our results concerning Models 2.1 and 2.2. In the following we refer to $U_{1,j}$, $X_{1,j}$, $Y_{1,j}$, F_1 and M_1 as U_j, X_j, Y_j, F and M, respectively, for $j = 1, 2, \ldots, C_0$, and we let $S_j = (X_j, Y_j)$.

1.2.1. Model 2.1

In our first simple model, namely Model 2.1, introduced in Section 2.1, from the total information available at stage one, i.e. C_0 and $\{(j, S_j)\}_{j=1}^{C_0}$, we have used only C_0, (F, M) and X_j to determine the distribution of U_j. To be more clear, in Sections 2.2-2.3, we have studied the case where each female marries a random number of times independently of the number of marriages of other females, with a distribution which determines only by C_0, F and M or equivalently by C_0, $\overline{S} = (\overline{X}, \overline{Y}) = (\frac{F}{C_0}, \frac{M}{C_0})$. If we consider a simple case where a female marries at most once in her life, then $U_1, \ldots, U_{C_0}$ are independent and

$$U_j \mid \{C_0, \overline{S}, X_j\} \sim \mathrm{Bin}(X_j, \varphi_{C_0}(\overline{S})), \tag{1.2.2}$$

$j = 1, 2, \ldots, C_0$, for some sequence of non-negative function $\{\varphi_z\}_{z=0}^{\infty}$ on $\mathcal{R}_+^2$ which are bounded by one.

The more general case for Model 2.1 is obtained if multiple marriages are allowed. We study the supercritical as well as critical case of this model and we find sufficient conditions for non-extinction of the process $\{C_n\}$. Perhaps it is the results for the critical case which are most important in Chapter 2 (as well as in Chapter 3), where we will see how the relation between the different parameters involved in the characterisation of the process determine the asymptotic behavior of the process $\{C_n\}$.

In the reproduction stage, with $\mu(z) = (\mu_1(z), \mu_2(z)) \in \mathcal{R}_+^2$ and $\Sigma(z) = [\sigma_{i,j}(z)]_{2\times 2}$ let $E[S_1 \mid C_0 = z] = \mu(z)$ and $\mathrm{Cov}[S_1 \mid C_0 = z] = \Sigma(z)$ denote the conditional expectation and the conditional covariance of the number of the offspring of a couple given that there are z of them, respectively. We assume that $\mu(z) \to \mu^* = (\mu_1^*, \mu_2^*)$ and $\Sigma(z) \to \Sigma^* = [\sigma_{i,j}^*]_{2\times 2}$ as $z \to \infty$.

In the mating stage, let us denote the conditional expectation of the number of marriages of a female, given that there are $C_0 = z$ couples around and these couples in the average gave birth to $\overline{S} = x \in \mathcal{R}_+^2$ children, is $\varphi_z(x)$. We assume that as $z \to \infty$ there exists a function φ^* on $\mathcal{R}_+^2$ such that $\varphi_z \to \varphi^*$.

Now, if we let $m := \mu_1^* \varphi^*(\mu^*)$ then one can show that (see Lemma 2.8(i) and Remark 2.9)

$$E[C_1 \mid C_0 = z] = mz + o(z). \tag{1.2.3}$$

If $m > 1$ (=1) we say that the process $\{C_n\}$ is a supercritical (critical) process (see Definition 2.1).

In Section 2.2 we study the supercritical case, i.e. when $m > 1$. Under suitable conditions we establish the dichotomy between extinction or explosion, show that the latter can occur and prove exponential growth in that case, almost surely and in $\mathcal{L}_1$ (as well as in $\mathcal{L}_2$ under some additional conditions). These results are given in Theorem 2.3.

In Section 2.3 we study the critical case, i.e. when $m = 1$. In this case the conditional expectation and the conditional variance of the number of couples in the first generation, C_1, given that we start the population with $C_0 = z$ couples, can be written in the following form.

$$\begin{aligned} E[C_1 \mid C_0 = z] &= z + b^* z^{\beta_0} + o(z^{\beta_0}), \\ \mathrm{Var}[C_1 \mid C_0 = z] &= c^* z + o(z), \end{aligned} \tag{1.2.4}$$

where the parameters b^*, β_0 and c^* are such that $b^* \in \mathcal{R}$, $0 \leq \beta_0 < 1$ and $c^* > 0$. The different values of these parameters determine the limiting behavior of the process $\{C_n\}$. Under suitable conditions we have:

- If $b^* \leq 0$, then the process $\{C_n\}$ dies out almost surely, i.e. $P(C_n \to 0) = 1$.
- If either $\beta_0 = 0$ and $2b^* > c^*$ or $0 < \beta_0 < 1$ and $b^* > 0$, then we have $P(C_n \to 0 \text{ or } \infty) = 1$ and $P(C_n \to \infty) > 0$.
- If $\beta_0 = 0$ and $2b^* > c^*$, then as $n \to \infty$ on the non-extinction set, i.e. $\{C_n \to \infty\}$, the distribution of C_n/n converges to a Gamma distribution with the scale and shape parameters which determine only by b^* and c^*.
- If $b^* > 0$ and $0 < \beta_0 \leq 2/3$, then C_n converges to a deterministic sequence in probability. And we can translate and scale C_n appropriately such that the conditional distribution of the obtained random variable converges weakly to a standard normal distribution.
- If $b^* > 0$ and $2/3 < \beta_0 < 1$, then C_n converges to a deterministic sequence almost surely.

The reader should see Theorem 2.15 for full description of these results. For a description of the way that the limiting parameters of Model 2.1, i.e. μ^*, Σ^*, φ^* and ν^*, from both the reproduction and mating stages determine the values of β_0, b^* and c^* see Remark 2.10. Examples 2.2 and 2.3 are devoted to the role of these parameters in a simple case where for $x \in \mathcal{R}_+^2$, the function $\varphi_z(x)$ has the following form

$$\varphi_z(x) = \left(\frac{x_2}{x_1 + x_2}\right)^{d_0 + d_1 z^{-h_1}} \exp\{-d_2(x_1 + x_2) z^{-h_2}\} \mathbf{1}_{\{\min(z, x_1, x_2) > 0\}}, \tag{1.2.5}$$

where the parameters $d_i \in \mathcal{R}$ and $h_j \in \mathcal{R}_+$, for $i = 0, 1, 2$ and $j = 1, 2$. Notice that d_0 and μ^* should be such that $m = \mu_1^* \varphi^*(\mu^*) = 1$.

1.2.2. Model 2.2

In Chapter 3, we study the process $\{C_n\}$ for Model 2.2, both in the supercritical and the critical case, where the conditional distribution of U_j could depend on C_0, $\overline{S}$ and (X_j, Y_j), for $j = 1, 2, \ldots, C_0$. However, from the point of view of the daughters of the jth family it seems reasonable to separate the information concerning their family from the rest of the population. For this purpose, we let $\overline{X}^{(j)}$ and $\overline{Y}^{(j)}$ denote the average numbers of daughters and sons born from all couples except the jth couple (see (3.0.1) for the definition). Since knowing C_0, $\overline{S}$ and S_j, is equivalent to knowing C_0, $\overline{S}^{(j)} = (\overline{X}^{(j)}, \overline{Y}^{(j)})$ and S_j, for $j = 1, 2, \ldots$, in the following we shall use the latter. For $z \in \mathcal{Z}_+$ and $x, y \in \mathcal{R}_+^2$, let $\psi_z(x, y)$ denote the conditional expectation of the number of marriages of the daughters of the jth family given $C_0 = z$, $S_j = x$ and $\overline{S}^{(j)} = y$. We assume that there exists a function ψ^* on $\mathcal{R}_+^2 \times \mathcal{R}_+^2$ such that $\psi_z \to \psi^*$ as $z \to \infty$. With $\mu(z)$, μ^*, $\Sigma(z)$ and Σ^* as before, define the parameter $\tilde{m}$ as $\tilde{m} = \lim_{z\to\infty} E[\psi^*(S_1, \mu^*) \mid C_0 = z]$. Then we can show (see Lemma 3.3(i) and Remark 3.3) that

$$E[C_1 \mid C_0 = z] = \tilde{m}z + o(z). \tag{1.2.6}$$

If $\tilde{m} > 1$ ($\tilde{m} = 1$) we say that the process $\{C_n\}$ is supercritical (critical) (see Definition 3.1).

In Section 3.1 we studied the supercritical case, i.e. when $\tilde{m} > 1$. For this model, similarly to Model 2.1, under suitable conditions we establish the dichotomy between extinction or explosion, show that the latter can occur and prove exponential growth in that case, almost surely and in $\mathcal{L}_1$ (as well as in $\mathcal{L}_2$ under some additional conditions). These results are given in Theorem 3.5.

In Section 3.2 we study the critical case, i.e. when $\tilde{m} = 1$. As for Model 2.1, we can write the conditional expectation and the conditional variance of the number of couples of the first generation as

$$\begin{aligned} E[C_1 \mid C_0 = z] &= z + \tilde{b}^* z^{\tilde{\beta}_0} + o(z^{\tilde{\beta}_0}), \\ \mathrm{Var}[C_1 \mid C_0 = z] &= \tilde{c}^* z + o(z), \end{aligned} \tag{1.2.7}$$

where the parameters $\tilde{b}^*$, $\tilde{\beta}_0$ and $\tilde{c}^*$ are such that $\tilde{b}^* \in \mathcal{R}$, $0 \le \tilde{\beta}_0 < 1$ and $\tilde{c}^* > 0$. Under some assumptions Theorem 3.8 presents our results for Model 2.2 in this case. The assertions of Theorem 3.8 are the same as those of Theorem 2.15 (after replacing β_0, b^* and c^* by $\tilde{\beta}_0$, $\tilde{b}^*$ and $\tilde{c}^*$, respectively).

In Example 3.2 we derived the values of $\tilde{b}^*$, $\tilde{\beta}_0$ and $\tilde{c}^*$ for an example of Model 2.2, where for $x, y \in \mathcal{R}_+^2$, the function $\psi_z(x, y)$ has the following form

$$\psi_z(x,y) = x_1\{\alpha(\frac{x_2}{x_1+x_2})^{d_0+d_1z^{-h_1}} + (1-\alpha)(\frac{y_2}{y_1+y_2})^{d_2+d_3z^{-h_2}}\}\mathbf{1}_{\{z>0\}}, \tag{1.2.8}$$

for some parameters $0 \le \alpha \le 1$, $h_1, h_2, d_0, d_1 \in \mathcal{R}_+$ and $d_1, d_3 \in \mathcal{R}$. Notice that d_0, d_2 and α in (1.2.8) together with μ^* should be such that $\tilde{m} = 1$.

In Example 3.3 we study the relation between the parameters of Model 2.1 with those of Model 2.2.

1.3. Bisexual Populations and Summary of the Results for Model 4.1

Chapter 4 is devoted to a particular simple bisexual population, namely Model 4.1. In contrary to the 2S-GWBP studied in Chapters 2 and 3, this bisexual population evolves in one stage: in each generation all females independently produce children (of both sexes) according to a reproduction law. It is this reproduction law which reflects the effect of the males in a generation. To be more precise, let $\mathcal{C}_2 = \{G_z : z \in \mathcal{Z}_+\}$ be a class of distributions on $\mathcal{Z}_+^2$ (where $G_0 = \mathbf{1}_{\mathcal{Z}_+^2}$ and $G_z(x) \ge G_{z+1}(x)$ for all $z \in \mathcal{Z}_+$ and $x \in \mathcal{Z}_+^2$). Start the population with F_0 females and M_0 males. Then the jth of these females, independently of others and according to the distribution $G_{M_0} \in \mathcal{C}_2$, produces $X_{1,j}$ daughters and $Y_{1,j}$ sons, for $j = 1, 2, \ldots, F_0$. If we let $S_{1,j} = (X_{1,j}, Y_{1,j})$, the total number of females and males in the first generation is equal to $(F_1, M_1) = \sum_{j=1}^{F_0} S_{1,j}$.

We consider the case where $G_z \in \mathcal{C}_2$, converges monotonously to a limiting distribution, say G^*, as $z \to \infty$. One can think of G^* as the case where females reproduce in a population with an abundance of males and hence they reproduce with their full capacity. We have used the Coupling Method to study the asymptotic behavior of this model for the supercritical case, and coupled the process $\{F_n\}$ to a B-GWBP with a promiscuous mating function, and we have shown that the coupling is successful. Hence, after a finite coupling time, the two processes would have the same behavior.

Loosely speaking, we can construct two sequences of random vectors, namely $\{(\hat{F}_n, \hat{M}_n)\}_{n=0}^{\infty}$ and $\{(\hat{F}_n^*, \hat{M}_n^*)\}_{n=0}^{\infty}$, on the same probability space such that (for detailed of their construction see (4.1.7)-(4.1.9)):

(i) $\{(\hat{F}_n, \hat{M}_n)\}_{n=0}^{\infty}$ has the same marginal distribution as $\{(F_n, M_n)\}_{n=0}^{\infty}$.

(ii) Almost surely, $\hat{F}_n \le \hat{F}_n^*$ and $\hat{M}_n \le \hat{M}_n^*$, for all $n = 0, 1, 2, \ldots$.

(iii) Define the random time τ as $\tau := \{k : \hat{F}_n = \hat{F}_n^* \text{ for all } n \ge k\}$. If we assume that $P(\tau < \infty) = 1$, then we can show that $\{(\hat{F}_n^*, \hat{M}_n^*)\}_{n\ge\tau}$ is a **B-GWBP** with the promiscuous mating function.

Theorems 4.7 and 4.8 give sufficient conditions for the random time τ to be finite almost surely, i.e. $P(\tau < \infty) = 1$. To highlight our results we need to introduce some parameters. For $z \in \mathcal{Z}_+$, let the vectors $\mu(z)$ and μ^* be the expectations of random vectors with the distributions G_z and G^*, respectively.

Theorem 4.1 gives sufficient conditions which guarantee that the process $\{\hat{F}_n\}_{n=0}^{\infty}$ grows to infinity with positive probability, i.e. $P(\inf_n \hat{F}_n = 0) < 1$.

Notice that by our assumption $\mu_i(z) \nearrow \mu_i^*$ as $z \to \infty$, for $i = 1, 2$. Under some assumptions such as $\mu_1(1) > 1$ and $\sum_{z=1}^{\infty}(\mu_1^* - \mu_1(z)) < \infty$, Theorem 4.7 says that $\tau < \infty$ almost surely.

Next, define the function ρ on $\mathcal{Z}_+$ as $\rho(z) := (\mu_1^* - \mu_1(z))(1 - \mathbf{1}_{\{z=0\}})$ for $z \in \mathcal{Z}_+$. Theorem 4.7 requires that $\mu_1(1) > 1$, i.e. the expected number of daughters born from a female is greater than one as long as there is at least one male around. Theorem 4.8 aims to weaken this assumption. Under some assumptions such as $\mu_1(1) > 0$, $\mu_1^* > 1$ and that the function ρ is a sub-multiplicative, Theorem 4.8 says that $\tau < \infty$ almost surely.

Remark 4.4 at the end of Section 4.2 answers to the question why one can not treat a two-sex branching population as a two-types branching population. Examples 4.1 and 4.2 present cases where they satisfy the assumptions of Theorem 4.7 and 4.8, respectively, however neither of these examples satisfy the assumptions of both of these two theorems. Example 4.3 aims to connect Model 4.1 of this Chapter to **2S-GWBP** of Chapters 2 and 3.

In Chapter 5, we present a way of constructing a maximal coupling (or γ-coupling) between two stochastically ordered, $\mathcal{Z}_+$-valued random variables. Hence this construction can be used to study the model of Chapter 4, whenever the numbers of daughters and sons of a couple are independent of each other.

In the above models, individuals are born at discrete times and live only one unit of time, as in the case with the processes of the Galton-Watson type. However, there are other approaches to study the behaviour of a two-sex population. Kendall [21] introduced a stochastic growth process (a birth-death type of process) and studied some simple cases where the birth rates (the death rates, respectively) for females and males are the same. This type of model has been further studied by others. For more advanced models see Goodman [10]-[11], Mode [32]-[33] and Asmussen [2].

Many interesting examples of mating systems in nature are given by Krebs and Davies [29].

Chapter 2

Two-Sex Galton-Watson Branching Processes: Conditional Independent Mating for Females

Consider a two-sex population, i.e. a population which consists of two sexes, female and male. The fate of such a population is determined by many factors, among them are the behaviour of the members of the population, their relationship towards each other which could manifest itself through the competition for food, mating, etc. To present a model for such a population which is both realistic and mathematically tractable, one must, in the first place, determine the kind of characteristic to be investigated with the help of the model. In this Chapter, we present the mathematical structure of a simple model for two-sex populations. Our objective is to study the asymptotic properties of such a population. As one will see, we discard the effects caused by the variation in the environment as well as the spatial positions of the individuals (these aspects are suitable for future studies).

In our model, we have non-overlapping generations, i.e. each female and male lives only one unit of time (or in other words, the individual born in each generation dies at the end of that generation). A population starts with a number of families (or couples or mating units). Each of these families produces a number of children, independently of each other and according to a probability law which could depend only on the total number of families present at that moment. Since the size of a population varies, one would expect that this variation would have some effect on the reproduction law of families, and hence we allow such dependence. In comparison with asexual branching process, we have some kind of size dependent reproduction law for each family.

After giving birth to new children, the parents die and it is time for the newly born children to continue the reproduction process. However, to do so, these children (of both sexes, female and male) should somehow interact with each other to form their own family (or couple or mating unit). After defining the necessary notation one could clearly see what kind of interaction we have in mind. Having formed these families and given their number, again each of these families, independently of each other, produces children of the next generation according to a probability law which depends on the total number of families. Hence, our two-sex population evolves in two stages, in the first stage reproduction and in the second stage mating. In the rest of this chapter, we explain the nature of these two stages in details.

In Section 2.1, we present the notation as well as our general model, followed by descriptions of the two cases, i.e. Models 2.1 and 2.2. While Model 2.1 is a special case of Model 2.2, studying its properties (both in the supercritical - as well as critical case) presents us with the methodology which can be used later on in Chapter 3 to study Model 2.2. In Sections 2.2 and 2.3, we have studied the supercritical and critical case of Model 2.1, respectively, and in Section 2.4 we present some simple examples to illustrate the obtained results.

2.1. The General Model

Assume that each individual dies at the end of the generation they were born into and let us start a population with a number of couples, say C_0. Without loss of generality, through the rest of this work, we assume that C_0 is fixed (i.e. non-random). In the first stage, we assume that each of these C_0 couples (or families) reproduce independently, according to some probability law which depends only on C_0. Let $X_{1,j}$ and $Y_{1,j}$ be the numbers of daughters and sons of the jth couple of the zeroth generation, respectively, and also let $S_{1,j} := (X_{1,j}, Y_{1,j})$, for $j = 1, 2, \ldots, C_0$. After reproduction, the old individual will die and the newly born children enter the next stage, i.e. mating. We assume that given the total numbers of couples C_0, and the sizes of their family, i. e. $\{(j, S_{1,j})\}_{j=1}^{C_0}$, the total number of marriages by the daughters of each family is independent of each other, with a distribution law which depends only on C_0 and $\{(j, S_{1,j})\}_{j=1}^{C_0}$. We shall refer to the latter information as the *population configuration* (here as the population configuration in the first generation). Let $U_{1,j}$ denote the total number of marriages formed by daughters of the jth couple of the zeroth generation.

Hence the total number of couples in the first generation will be equal to sum of these marriages for all families, i.e. $C_1 = \sum_{j=1}^{C_0} U_{1,j}$.

Similarly, let C_n denote the number of couples in the nth generation and $S_{n,j} = (X_{n,j}, Y_{n,j})$ be the total number of the daughters and sons of the jth couple in the $(n-1)$th generation, for $j = 1, 2, \ldots, C_{n-1}$. Then in each generation, we can present the information available from the past at the beginning of each of these two stages by the σ-algebras $\mathcal{F}_n$ and $\mathcal{G}_n$ which are defined by

$$\begin{aligned} \mathcal{F}_n &= \sigma(C_0, C_1, \ldots, C_n), \\ \mathcal{G}_n &= \sigma(C_0, C_1, \ldots, C_n, \{(j, S_{n+1,j})\}_{j=1}^{C_n}), \end{aligned} \tag{2.1.1}$$

for $n = 0, 1, 2, \ldots$.

Hence, we assume that given $\mathcal{F}_n$, $\{S_{n+1,j}\}_{j=1}^{C_n}$ are i.i.d. random vectors with common probability distribution, say H_{C_n}, which depends on the past only through C_n. For example, when we have a cooperative (or competitive) population, one could expect that the fertility rate increases (or decreases) in the $(n+1)$th generation as the number of couples increases (or decreases).

Let $U_{n+1,j}$ denote the total number of couples formed by the daughters of the jth couple in the nth generation during their life. We assume that given $\mathcal{G}_n$, $U_{n+1,1}, \ldots, U_{n+1,C_n}$ are independent, with a distribution which is determined by C_n and the population *configuration* in the $(n+1)$th generation, i.e. $\{(j, S_{n+1,j})\}_{j=1}^{C_n}$. The total number of couples in the $(n+1)$th generation will be equal to

$$C_{n+1} = \sum_{j=1}^{C_n} U_{n+1,j}. \tag{2.1.2}$$

In the sequel, we refer to this model as 2S-GWBP (Two-Sex Galton-Watson Branching Process) whenever it is appropriate.

As one can see from the construction of 2S-GWBP model, the first stage in each generation, reminds us of the kind of dependence one would have in the models related to population size dependent branching processes (for an asexual population). At first sight, the second stage may appear to be an additional complication. However, not only does it help us to model a variety of scenarios for the mating stage by using the information available on the family level (the population configuration in that generation), but it also helps to adapt some of the methods used for population size dependent branching processes (for an asexual population) in the study of 2S-GWBP.

We shall focus on the behaviour of the process $\{C_n\}_{n=0}^{\infty}$. As one can see from the presentation of our general model given above, the number of couples (or families) in the $(n+1)$th generation, i.e. C_{n+1}, is equal to

$C_{n+1} = \sum_{j=1}^{C_n} U_{n+1,j}$. Given $\mathcal{F}_n$, i.e. the information available at the end of the nth generation, the r.v.s $\{U_{n+1,j}\}_{j=1}^{C_n}$ are no longer independent, but they are interchangeable r.v.s. This dependence is due to the second step, i.e. the mating step, and this distinguishes our model from the models studied in the population size-dependent branching processes. This point will become clearer when we present some examples.

As we will see, the above construction of the 2S-GWBP has some advantages and disadvantages. We hope that the population configuration helps us to study the behaviour of some interesting mating scenarios, and hence a wider class of models will be available.

In the following, we describe two models which will be the subject of our study in the present chapter as well as the next chapter.

Model 2.1. *Assume that, given $\mathcal{G}_n$, in the $(n+1)$th generation each female forms a number of marriages, independently of other females. Denote the number of marriages for the kth daughter in the jth family by $\eta_{n+1,j,k}$. Then we have $U_{n+1,j} = \sum_{k=1}^{X_{n+1,j}} \eta_{n+1,j,k}$ and C_{n+1} will be*

$$\begin{aligned} C_{n+1} &= \sum_{j=1}^{C_n} U_{n+1,j} \\ &= \sum_{j=1}^{C_n} \sum_{k=1}^{X_{n+1,j}} \eta_{n+1,j,k}. \end{aligned} \tag{2.1.3}$$

Further, assume that given $\mathcal{G}_n$, the conditional distribution of $\eta_{n+1,j,k}$ depends on the number of couples in the nth generation, i.e. *C_n as well as the average number of children per family,* i.e.

$$\overline{S}_{n+1} = \sum_{j=1}^{C_n} S_{n+1,j}/C_n = (\overline{X}_{n+1}, \overline{Y}_{n+1}). \tag{2.1.4}$$

Remark 2.1. First notice that, knowing C_n and $\sum_{j=1}^{C_n} S_{n+1,j}$ in the $(n+1)$th generation is equivalent to knowing C_n and $\overline{S}_{n+1} = \sum_{j=1}^{C_n} S_{n+1,j}/C_n$ (i.e. the total number of females and males in $(n+1)$th generation). Secondly, by a Law of Large Numbers the average number of children born in the $(n+1)$th generation $\overline{S}_{n+1}$, converges almost surely to a fixed point in $\mathbf{R}_+^2$ as C_n goes to infinity. For these reasons, we prefer to express the conditional distribution of $\eta_{n+1,j,k}$ given $\mathcal{G}_n$ in terms of C_n and $\overline{S}_{n+1}$ rather than C_n and $\sum_{j=1}^{C_n} S_{n+1,j}$. □

The Model 2.1 reflects the situation where the number of marriages of the females in a generation is not affected by the size of the family they were born into or by the number of marriages of other females in that generation. But rather it depends on the number of families and the average numbers of children per family in that generation. One can think of the situation where with respect to population resources, each female will be equally supported by the population in search of partner(s).

Remark 2.2. Let $\hat{\mathbf{P}}_{n+1}$ denote the empirical distribution of the number of children born into each family in the nth generation, i.e.

$$\hat{\mathbf{P}}_{n+1}(A) := \frac{1}{C_n} \sum_{j=1}^{C_n} \mathbf{1}_A(S_{n+1,j}), \tag{2.1.5}$$

for any $A \subset \mathcal{Z}_+^2$. Then one can extend the scope of Model 2.1 and assume that given $\mathcal{G}_n$, the distribution of $\eta_{n+1,j,k}$ (as in Model 2.1) depends only on C_n and $\hat{\mathbf{P}}_{n+1}$ or some parameter of this distribution, such as its mean (i.e. $\overline{S}_{n+1}$), variance and etc. □

Let us define some parameters for Model 2.1. Given $C_0 = z$, $S_{1,1} = (X_{1,1}, Y_{1,1})$ has the mean vector $\mu(z)$ and the covariance matrix $\Sigma(z)$, i.e.

$$\begin{aligned} \mu(z) &:= (\mu_1(z), \mu_2(z)) = E[S_{1,1} \mid C_0 = z], \\ \Sigma(z) &:= [\sigma_{ij}(z)]_{2\times 2} = \mathrm{Cov}[S_{1,1} \mid C_0 = z]. \end{aligned} \tag{2.1.6}$$

And, given $C_0 = z$, we denote the conditional expectation and variance of $\eta_{1,1,1}$ (the number of marriages of the first daughter in the first family in the first generation) given $\mathcal{G}_0$, will be denoted by

$$\begin{aligned} \varphi_z(\overline{S}_1) &:= E[\eta_{1,1,1} \mid \mathcal{G}_0, C_0 = z], \\ \nu_z(\overline{S}_1) &:= \mathrm{Var}[\eta_{1,1,1} \mid \mathcal{G}_0, C_0 = z]. \end{aligned} \tag{2.1.7}$$

Without loss of generality, we assume that φ_z and ν_z are defined on all $\mathcal{R}_+^2$.

In the following, we present some assumptions corresponding to the sequence of vectors $\{\mu(z)\}_{z=0}^{\infty}$ and the sequences of real-valued functions $\{\varphi_z\}_{z=0}^{\infty}$ and $\{\nu_z\}_{z=0}^{\infty}$.

A. 1. *We assume that there is $\mu^* \in \mathcal{R}_+^2$ and real-valued functions φ^* and ν^* on $\mathcal{R}_+^2$ such that as $z \to \infty$ we have*

$$\begin{aligned} &(i) \quad \mu(z) \longrightarrow \mu^*, \\ &(ii) \quad \varphi_z \longrightarrow \varphi^* \\ &(iii) \quad \nu_z \longrightarrow \nu^*. \end{aligned} \tag{2.1.8}$$

Further, we assume that $\{\varphi_z\}_{z=1}^{\infty}$ and $\{\nu_z\}_{z=1}^{\infty}$ are sequences of uniformly bounded functions on $\mathcal{R}_+^2$.

Now we are ready for the following definition.

Definition 2.1. *We say that Model 2.1 is a subcritical, critical or supercritical if*

$$m := \mu_1^* \varphi^*(\mu^*) <, \quad =, \quad \text{or} > 1, \tag{2.1.9}$$

respectively.

Remark 2.3. One can think of the parameter m as the expected number of the marriages of the daughters of a family when the population size is large. □

In Sections 2.2 and 2.3, we study the supercritical and the critical case of Model 2.1.

The description of our next model is presented below.

Model 2.2. *Assume that, given $\mathcal{G}_n$, in the $(n+1)$th generation, the distribution of the number of marriages of the daughters of the jth couple,* i.e. *$U_{n+1,j}$, depends on the population configuration only through C_n, $S_{n+1,j}$ and $\overline{S}_{n+1}$, for $j = 1, 2, \dots, C_n$. Remember that by the general construction of 2S-GWBP, given $\mathcal{G}_n$, $\{U_{n+1,j}\}_{j=1}^{C_n}$ are independent* r.v.s.

Model 2.1 is a special case of Model 2.2. Further, Model 2.2 considers the case where the family size, i.e. the number of female and male children of each family effects the distribution law of the number of marriages of the daughters of that family. As an example of such a relation, a determining factor for the number of marriages in the jth family, could be the difference between the sex ratio in that family and the overall sex ratio in that generation, i.e. $D_{n+1,j} := \frac{X_{n+1,j}}{X_{n+1,j}+Y_{n+1,j}} - \frac{\overline{X}_{n+1}}{\overline{X}_{n+1}+\overline{Y}_{n+1}}$. And one allows that the distributions of the number of marriages in the jth and j'th families in the $(n+1)$th generation satisfy the following relation. Given $\mathcal{G}_n$, if $D_{n+1,j} > D_{n+1,j'}$, then

$$U_{n+1,j} \overset{\mathrm{D}}{\geq} U_{n+1,j'} \, . \tag{2.1.10}$$

This means that one could expect that the daughters of a family with a higher proportion of daughters, i.e. $\frac{X_{n+1,j}}{X_{n+1,j}+Y_{n+1,j}}$, stochastically would have more chances to get married than those with a smaller proportion of daughters.

Another example could be provided by considering models where we could have only inbreeding (or inbreeding is prohibited). This would imply that: $U_{n+1,j} = 0$ if we have $Y_{n+1,j} = 0$ (for the no inbreeding case this would imply that: $U_{n+1,j} = 0$ if $\sum_{i \neq j}^{C_n} Y_{n+1,i} = 0$).

Remark 2.4. Similar to Model 2.1 (as in Remark 2.2), in Model 2.2 one can allow the conditional distribution of $U_{n+1,j}$ given $\mathcal{G}_n$, depends on C_n, $S_{n+1,j}$ and $\hat{\mathbf{P}}_{n+1}$, with $\hat{\mathbf{P}}_{n+1}$ as in (3.0.2). However, we shall not study a more complicated model beyond Model 2.2 in this work. $\square$

We postpone the definitions of the subcriticality, criticality and supercriticality of Model 2.2 to Section 3.1 where we present some notation and related assumptions.

2.2. Model 2.1: Supercritical Case

In this section we study Model 2.1 in the supercritical case. For this purpose, we need some additional assumptions on Model 2.1, besides Assumption **A.1**. For convenience, in the rest of this chapter we shall refer to $\eta_{1,j,k}$, $S_{1,j} = (X_{1,j}, Y_{1,j})$, $\overline{S}_1 = (\overline{X}_1, \overline{Y}_1)$ and $U_{1,j}$ as $\eta_{j,k}$, $S_j = (X_j, Y_j)$, $\overline{S} = (\overline{X}, \overline{Y})$ and U_j, respectively, for $j = 1, 2, \ldots, C_0$ and $k = 1, 2, \ldots, X_{1,j}$.

A. 2. *With $\mu(\cdot)$ and μ^* as in Assumption* **A.1***, assume that for some $\delta_1 > 0$*

$$\mu_1^* \varphi^*(\mu^*) > 1, \quad \text{and} \quad \| \mu(z) - \mu^* \| = O(z^{-\delta_1}), \tag{2.2.1}$$

and that for some $\delta_2 > 0$

$$| \varphi_z(\mu(z)) - \varphi^*(\mu^*) | = O(z^{-\delta_2}). \tag{2.2.2}$$

A. 3. *Assume that for some $k_0 \geq 3$, $E[X_1^{k_0} + Y_1^{k_0} \mid C_0 = z]$ is uniformly bounded in $z \in \mathcal{Z}_+$. Also, let $\{t_z\}_{z=0}^{\infty}$ be a sequence of functions on $\mathcal{R}_+^2$ such that $t_{C_n}(\overline{S}_1) = E[\eta_{1,1}^{k_0} \mid \mathcal{G}_0]$. Assume that $\{t_z\}_{z=0}^{\infty}$ is a sequence of uniformly bounded functions on $\mathcal{R}_+^2$.*

A. 4. *For the sequences of functions $\{\varphi_z\}_{z=0}^{\infty}$ and $\{\nu_z\}_{z=0}^{\infty}$, given in Assumption* **A.1***, assume that there exist $\epsilon > 0$ and $z_\epsilon \in \mathcal{Z}_+$, such that if we let $B = B_\epsilon(\mu^*)$, then for some constant $K > 0$ the sequences $\{\varphi_z\}_{z=0}^{\infty}$ and $\{\nu_z\}_{z=0}^{\infty}$ satisfy*

$$\begin{gathered} | \varphi_z(x) - \varphi_z(\mu(z)) | \leq K \| x - \mu(z) \|, \\ \text{and} \\ | \nu_z(x) - \nu_z(\mu(z)) | \leq K \| x - \mu(z) \|, \end{gathered} \tag{2.2.3}$$

for all $x \in B$ and $z \geq z_\epsilon$.

We have the following definition.

Definition 2.2. *Let m be as in the definition 2.1. Define the sequences of real numbers $\{m_z\}_{z=0}^{\infty}$ and random variables $\{W_n\}_{n=0}^{\infty}$ such that*

$$m_z := E[U_1 \mid C_0 = z], \quad \text{and} \quad W_n := C_n/m^n, \tag{2.2.4}$$

for $z, n \in \mathcal{Z}_+$.

The following theorem presents our results on Model 2.1 in the supercritical case.

Theorem 2.3. (Exponential Growth) *Consider Model 2.1 in the supercritical case, with assumptions* **A.1**-**A.4** *satisfied. Then we have*

(i) $P(C_n \to 0 \text{ or } \infty) = 1$ *and* $P(C_n \to \infty) > 0$,

(ii) $E[C_n] \sim Km^n$, *as $n \to \infty$ for some constant $K > 0$,*

(iii) there exists a non-degenerate r.v. *W such that $W_n \to W$ almost surely and in $\mathcal{L}_1$.*

(iv) and if further the δ_1 and δ_2 in Assumption **A.2** *satisfy $(\delta_1 \wedge \delta_2) \geq 1/2$, then W_n converges in $\mathcal{L}_2$.*

Remark 2.5. Asmussen and Hering [3] studied a supercritical two-sex branching process. Their model is the two-stage chain (see pages 409-414 of [3])

$$C_0 \to \cdots \to (F_n, M_n) \to C_n \to (F_{n+1}, M_{n+1}) \to C_{n+1} \to \cdots \tag{2.2.5}$$

where F_n, M_n and C_n are the number of females, males and couples in the nth generation, respectively. Start the chain with C_0 couples. Each of these couples, independently of each other and according to a common distribution, produces children of both sexes. Denote the expected number of daughters and sons for a couple by a and b. The distribution of the number of couples in that generation is determined by the number of females and males in that generation. Let us define $r(f,m) := E[C_1 \mid F_1 = f, M_1 = m]/(f+m)$, for $f, m \in \mathcal{Z}_+$. Then they assumed that:

- For all $K, c \in \mathcal{Z}_+$ we have

$$P(\sup_{n \in \mathcal{Z}_+} C_n > K \mid C_0 = c) > 0. \tag{2.2.6}$$

- The limit

$$r_0 := \lim_{f+m \to \infty, \frac{m}{f+m} \to \frac{b}{a+b}} r(f,m), \tag{2.2.7}$$

  exists.

- For some $d \in (0,2)$ we have

$$\limsup_{f+m\to\infty, \frac{m}{f+m}\to\frac{b}{a+b}} \operatorname{Var}[C_1 \mid F_1 = f, M_1 = m]/(f+m)^d < \infty. \quad (2.2.8)$$

- $E[(F_1 + M_1)^2 \mid C_0 = 1] < \infty$ and $r_0(a+b) > 1$.

If we let $E = \{C_n = 0 \text{ eventually}\}$, then, among other things, they have shown that $P(E) + P(C_n \to \infty) = 1$ and almost surely on E^c for all $\epsilon > 0$, $C_n/(r_0(a+b)+\epsilon)^n \to 0$ and $C_n/(r_0(a+b)-\epsilon)^n \to \infty$. These results are given in Theorem 2.1 (see page 409 of [3]). In addition to the above if we assume that:

- $r(f,m)$ converges to r_0 with the rate

$$O(\mid \log(\frac{m}{f+m} - \frac{b}{a+b}) \mid^{-1-x} + (\log(f+m))^{-1-x}), \qquad (2.2.9)$$

  for some $x > 0$ as $f+m \to \infty$ and $m/(f+m) \to b/(a+b)$.

Then Theorem 2.2 (see page 409 of [3]) gives that $W := \lim_{n\to\infty} C_n/((r_0(a+b))^n$ exists almost surely and $0 < W < \infty$ on E^c.

In comparing their model with Model 2.1 we like to draw your attention to the following points.

(1) If, for a moment, we add assumption that in Model 2.1 the conditional distribution of S_1 given $\mathcal{F}_0$ is fixed and does not depend on C_0 and further the conditional distribution of $\eta_{1,1}$ given $\mathcal{G}_0$ depends only on the total number of daughters and sons (i. e. depends only on F_1 and M_1 but not on C_0), then Model 2.1 will be a special case of Asmussen and Hering's model, otherwise when the number of couples in each generation is the sum of the number of marriages of each female, our model is more general than that of Asmussen and Hering.

(2) However, in the Asmussen and Hering work they have given sufficient condition for the almost sure extinction of the process $\{C_n\}$ (see Proposition 2.4, page 413 [3]), but they did not give any sufficient condition for the survival of the process $\{C_n\}$, i. e. a condition which would imply that $P(E) < 1$ or equivalently $P(C_n \to \infty) > 0$. However we can see from Theorem 2.3(i) that we gave conditions under which $P(C_n \to \infty) > 0$ for Model 2.1.

(3) In our work we have studied the critical case which has not been considered in the work by Asmussen and Hering.

□

Before the proof of Theorem 2.3, we need some lemma.

Lemma 2.4. *For given $i, k \geq 1$ and all $a_j \geq 0$ for $j = 1, \ldots, k$, we have*

$$(a_1 + \ldots + a_k)^i \leq k^{i-1}(a_1^i + \ldots + a_k^i). \tag{2.2.10}$$

Proof of Lemma 2.4: Define the function h on $\mathcal{R}_+$ as $h(x) := x^i$. Then it is easy to see that h is a convex function.For given $a_j \geq 0$, for $j = 1, \ldots, k$, define the probability measure $\mathbf{F}$ as

$$\mathbf{F}(A) := \frac{1}{k}\sum_{j=1}^{k} \mathbf{1}_A(a_j), \tag{2.2.11}$$

for all $A \in \mathcal{B}(\mathbf{R}_+)$. Let R be a r.v. with the distribution $\mathbf{F}$. Now, since h is convex by the Jensen inequality we have

$$\begin{aligned} & h(E[R]) \leq E[h(R)] \\ & \Longleftrightarrow h(\tfrac{a_1 + \ldots + a_k}{k}) \leq \tfrac{1}{k}(h(a_1) + \ldots + h(a_k)) \\ & \Longleftrightarrow (a_1 + \ldots + a_k)^i \leq k^{i-1}(a_1^i + \ldots + a_k^i). \end{aligned} \tag{2.2.12}$$

■

Lemma 2.5. *Let $\{Z_j\}_{j=1}^{\infty}$ be a sequence of independent random variables with means zero and uniformly bounded qth absolute moments, for $q = 1$ or $q \geq 2$. Then*

$$E[| \sum_{j=1}^{n} Z_j |^q] = \begin{cases} O(n), & \text{if } q = 1, \\ O(n^{q/2}), & \text{if } q \geq 2, \end{cases} \tag{2.2.13}$$

for $n = 1, 2, \ldots$.

Proof of Lemma 2.5: The assertion of the lemma is obvious for $q = 1, 2$. For $q > 2$, following [26], one can write

$$\begin{aligned} & E[| \textstyle\sum_{j=1}^{n} Z_j |^q] \leq K_q E[\{\textstyle\sum_{j=1}^{n} Z_j^2\}^{q/2}] \\ & \leq K_q E[\{n^{\frac{q-2}{q}}(\textstyle\sum_{j=1}^{n} | Z_j |^q)^{2/q}\}^{q/2}] \\ & = K_q n^{\frac{q-2}{2}} \textstyle\sum_{j=1}^{n} E[| Z_j |^q] = O(n^{q/2}). \end{aligned} \tag{2.2.14}$$

For the first inequality we used Marcinkiewicz-Zygmund's inequality and for the second Hölder's inequality. ■

The following lemma gives some properties of $\overline{S}$.

Lemma 2.6. *Assume that for some integer k, $E[X_1^{k\vee 2} + Y_1^{k\vee 2} \mid C_0 = z]$ is uniformly bounded in z. Then as $z \to \infty$,*

$$E[\| \overline{S} - \mu(z) \|^k \mid C_0 = z] = O(z^{-k/2}), \tag{2.2.15}$$

for $k \geq 1$. Further, for any $\epsilon > 0$ and $z \in \mathcal{Z}_+$, let D be a closed disk in $\mathcal{R}_+^2$ centred at $\mu(z)$ and radius $\epsilon > 0$. Then we have

$$P(\overline{S} \in D^c \mid C_0 = z) = O(z^{-k/2}), \tag{2.2.16}$$

for $k \geq 1$.

Proof of Lemma 2.6: For $k = 2$ it is obvious. For $k = 1$ by an application of the Hölder inequality we have

$$\begin{aligned} & E[\| \overline{S} - \mu(z) \| \mid C_0 = z] \\ & \leq E[\| \overline{S} - \mu(z) \|^2 \mid C_0 = z]^{1/2} = O(z^{-1/2}). \end{aligned} \tag{2.2.17}$$

For $k > 2$, by the triangle inequality and Lemma 2.4 we have

$$\begin{aligned} \| \overline{S} - \mu(z) \|^k & \leq \{ | \overline{X} - \mu_1(z) | + | \overline{Y} - \mu_2(z) | \}^k \\ & \leq 2^{k-1} \{ | \overline{X} - \mu_1(z) |^k + | \overline{Y} - \mu_2(z) |^k \}. \end{aligned} \tag{2.2.18}$$

Now using Lemma 2.5 we have

$$\begin{aligned} & E[| \overline{X} - \mu_1(z) |^k \mid C_0 = z] = E[| \tfrac{1}{z} \textstyle\sum_{j=1}^{z} (X_j - \mu_1(z)) |^k \mid C_0 = z] \\ & = \tfrac{1}{z^k} E[| \textstyle\sum_{j=1}^{z} (X_j - \mu_1(z)) |^k \mid C_0 = z] = \tfrac{1}{z^k} O(z^{k/2}) = O(z^{-k/2}). \end{aligned} \tag{2.2.19}$$

Similarly we have $E[| \overline{Y} - \mu_2(z) |^k \mid C_0 = z] = O(z^{-k/2})$. Hence from these and (2.2.18) we have

$$E[\| \overline{S} - \mu(z) \|^k \mid C_0 = z] = O(z^{-k/2}). \tag{2.2.20}$$

To show the second assertion of the lemma we use Chebyshev's inequality and the first part of this lemma. We get

$$\begin{aligned} & P(\overline{S} \in D^c \mid C_0 = z) = P(\| \overline{S} - \mu(z) \| > \epsilon \mid C_0 = z) \\ & \leq \tfrac{1}{\epsilon^k} E[\| \overline{S} - \mu(z) \|^k \mid C_0 = z] = O(z^{-k/2}). \end{aligned} \tag{2.2.21}$$

■

From the sequences of $\{\mu(z)\}_{z=0}^{\infty}$ and $\{\varphi_z\}_{z=0}^{\infty}$ given in the assumption **A.1**, define the sequence of functions $\{f_z\}_{z=0}^{\infty}$ on $\mathcal{R}^2$ as in the following:

$$f_z(x) := \begin{cases} x_1\varphi_z(x) - \mu_1(z)\varphi_z(\mu(z)), & \text{if } x \in \mathcal{R}_+^2, \\ 0, & \text{otherwise.} \end{cases} \tag{2.2.22}$$

Then we have the following lemma.

Lemma 2.7. *Assume that assumptions* **A.1**-**A.4** *hold. Then, with the function f_z as in (2.2.22), we have*

$$E[\mid f_z(\overline{S}) \mid^k \mid C_0 = z] = O(z^{-k/2}), \tag{2.2.23}$$

for $1 \le k \le k_0$ and $z > 0$.

Proof of Lemma 2.7: In the following, as in Assumption **A.4**, let B be a closed disk in $\mathcal{R}_+^2$ centred at μ^* with radius $\epsilon > 0$ and $z_\epsilon > 0$ such that for all $z \ge z_\epsilon$, $\mu(z)$ is an interior point of B. Further, take $\epsilon_0 > 0$ small enough such that the closed disk B contains $\cup_{z \ge z_\epsilon} B_z$, where $B_z := \{x \in \mathcal{R}_+^2 :\parallel x - \mu(z) \parallel \le \epsilon_0\}$.

Without loss of generality we can assume that the sequence of functions $\{\varphi_z\}_{z=1}^{\infty}$ is bounded by one.

With f_z as in (2.2.22) and $x = (x_1, x_2) \in \mathcal{R}_+^2$, we can write

$$\begin{aligned} f_z(x) &= x_1\varphi_z(x) - \mu_1(z)\varphi_z(\mu(z)) \\ &= (x_1 - \mu_1(z))(\varphi_z(x) - \varphi_z(\mu(z))) \\ &+\mu_1(z)(\varphi_z(x) - \varphi_z(\mu(z))) + \varphi_z(\mu(z))(x_1 - \mu_1(z)). \end{aligned} \tag{2.2.24}$$

By Lemma 2.4 and (2.2.24) we can write $E[\mid f_z(\overline{S}) \mid^k \mid C_0 = z]$ as

$$\begin{aligned} &E[\mid f_z(\overline{S}) \mid^k \mid C_0 = z] \\ &\le 3^{k-1}\{E[\mid \overline{X} - \mu_1(z) \mid^k \mid \varphi_z(\overline{S}) - \varphi_z(\mu(z)) \mid^k \mid C_0 = z] \\ &+\mu_1^k(z)E[\mid \varphi_z(\overline{S}) - \varphi_z(\mu(z)) \mid^k \mid C_0 = z] \\ &+\varphi_z^k(\mu(z))E[\mid \overline{X} - \mu_1(z) \mid^k \mid C_0 = z]\}. \end{aligned} \tag{2.2.25}$$

Now by the assumption φ_z is bounded by one, hence $\mid \varphi_z(x) - \varphi_z(y) \mid \le 1$ for all $x, y \in \mathcal{R}_+^2$. Thus by this and Lemma 2.6 we have

$$\begin{aligned} &E[\mid \overline{X} - \mu_1(z) \mid^k \mid \varphi_z(\overline{S}) - \varphi_z(\mu(z)) \mid^k \mid C_0 = z] \\ &\le E[\parallel \overline{S} - \mu(z) \parallel^k \mid C_0 = z] = O(z^{-k/2}). \end{aligned} \tag{2.2.26}$$

By Assumption **A.4** and Lemma 2.6 we have

$$
\begin{aligned}
&E[\mid \varphi_z(\overline{S}) - \varphi_z(\mu(z)) \mid^k \mid C_0 = z] \\
&= E[\mid \varphi_z(\overline{S}) - \varphi_z(\mu(z)) \mid^k ; \overline{S} \in B_z \mid C_0 = z] \\
&+ E[\mid \varphi_z(\overline{S}) - \varphi_z(\mu(z)) \mid^k ; \overline{S} \in B_z^c \mid C_0 = z] \qquad (2.2.27) \\
&\leq KE[\parallel \overline{S} - \mu(z) \parallel^k ; \overline{S} \in B_z \mid C_0 = z] + P(\overline{S} \in B_z^c \mid C_0 = z) \\
&= O(z^{-k/2}) + O(z^{-k_0/2}).
\end{aligned}
$$

And finally by Lemma 2.6 we have

$$E[\mid \overline{X} - \mu_1(z) \mid^k \mid C_0 = z] \leq E[\parallel \overline{S} - \mu(z) \parallel^k \mid C_0 = z] = O(z^{-k/2}). \quad (2.2.28)$$

By replacing (2.2.26)-(2.2.28) in (2.2.25), we get

$$E[\mid f_z(\overline{S}) \mid^k \mid C_0 = z] = O(z^{-k/2}). \qquad (2.2.29)$$

■

Lemma 2.8. *Assume that the assumptions* **A.1**-**A.4** *hold. With m as in (2.1.8), then we have*

$$
\begin{aligned}
&(i) \quad m = \lim_{z\to\infty} \tfrac{1}{z} E[C_1 \mid C_0 = z], \\
&(ii) \quad E[\mid C_1 - E[C_1 \mid C_0 = z] \mid^k \mid C_0 = z] = O(z^{k/2}),
\end{aligned}
\qquad (2.2.30)
$$

for $k \leq k_0$ where k_0 is as in Assumption **A.3**.

Proof of Lemma 2.8: We have $C_1 = \sum_{j=1}^{C_0} U_j$. Then given $C_0 = z$, from the description of Model 2.1 we have $U_j = \sum_{l=1}^{X_j} \eta_{j,l}$. Hence we can write

$$
\begin{aligned}
&E[C_1 \mid \mathcal{G}_0] = \sum_{j=1}^{z} E[U_j \mid \mathcal{G}_0] = \sum_{j=1}^{z} E[\sum_{l=1}^{X_j} \eta_{j,l} \mid \mathcal{G}_0] \\
&= \sum_{j=1}^{z} X_j \varphi_z(\overline{S}) = z\overline{X}\varphi_z(\overline{S}).
\end{aligned}
\qquad (2.2.31)
$$

To prove (2.2.30)(i), first we can write $\frac{1}{z}E[C_1 \mid C_0 = z]$ as

$$
\begin{aligned}
\tfrac{1}{z}E[C_1 \mid C_0 = z] &= \tfrac{1}{z}E[E[C_1 \mid \mathcal{G}_0] \mid C_0 = z] \\
&= \tfrac{1}{z}E[z\overline{X}\varphi_z(\overline{S}) \mid C_0 = z] \qquad (2.2.32) \\
&= E[\overline{X}\varphi_z(\overline{S}) \mid C_0 = z].
\end{aligned}
$$

Since $m = \mu_1^* \varphi^*(\mu^*)$ we can write $\frac{1}{z} E[C_1 \mid C_0 = z] - m$ as

$$\tfrac{1}{z} E[C_1 \mid C_0 = z] - m = E[\overline{X} \varphi_z(\overline{S}) \mid C_0 = z] - \mu_1^* \varphi^*(\mu^*). \tag{2.2.33}$$

Now, notice that: by Lemma 2.7 we have $E[\overline{X}\varphi_z(\overline{S}) \mid C_0 = z] = \mu_1(z)\varphi_z(\mu(z)) + O(z^{-1/2})$; by Assumption **A.1** we have $\varphi_z \to \varphi^*$ as $z \to \infty$; by Assumption **A.2** we have $\| \mu(z) - \mu^* \| = O(z^{-\delta_1})$ for some $\delta_1 > 0$; and by Assumption **A.4** for z large and for some constant $K > 0$ we have $| \varphi_z(\mu(z)) - \varphi_z(\mu^*) | \leq K \| \mu(z) - \mu^* \|$. Hence by adding and subtracting $\mu_1(z)\varphi_z(\mu(z))$, $\mu_1(z)\varphi_z(\mu^*)$ and $\mu_1(z)\varphi^*(\mu^*)$ to (2.2.33) we obtain

$$\begin{aligned}
&| \tfrac{1}{z} E[C_1 \mid C_0 = z] - m | \leq | E[\overline{X}\varphi_z(\overline{S}) \mid C_0 = z] - \mu_1(z)\varphi_z(\mu(z)) | \\
&+ | \mu_1(z)\varphi_z(\mu(z)) - \mu_1(z)\varphi_z(\mu^*) | + | \mu_1(z)\varphi^*(\mu^*) - \mu_1^*\varphi^*(\mu^*) | \\
&\leq O(z^{-1/2}) + K\mu_1(z) \| \mu(z) - \mu^* \| \\
&+ \mu_1(z) | \varphi_z(\mu^*) - \varphi^*(\mu^*) | + \varphi^*(\mu^*) | \mu_1(z) - \mu_1^* | .
\end{aligned} \tag{2.2.34}$$

The right hand side of (2.2.34) goes to zero as $z \to \infty$, hence we have shown the assertion (i) of the lemma.

For (2.2.30)(ii), from (2.2.31) $E[C_1 \mid \mathcal{G}_0] = z\overline{X}\varphi_z(\overline{S})$. With the function f_z defined in (2.2.22), we can decompose $C_1 - E[C_1 \mid C_0 = z]$ into the three terms as

$$\begin{aligned}
&C_1 - E[C_1 \mid C_0 = z] = \textstyle\sum_{j=1}^{z} U_j - E[U_j \mid C_0 = z] \\
&= \textstyle\sum_{j=1}^{z} \{U_j - E[U_j \mid \mathcal{G}_0]\} \\
&+ \textstyle\sum_{j=1}^{z} \{E[U_j \mid \mathcal{G}_0] - \mu_1(z)\varphi_z(\mu(z))\} \\
&+ \textstyle\sum_{j=1}^{z} \{\mu_1(z)\varphi_z(\mu(z)) - E[U_j \mid C_0 = z]\} \\
&= \textstyle\sum_{j=1}^{z} (U_j - E[U_j \mid \mathcal{G}_0]) + z f_z(\overline{S}) - zE[f_z(\overline{S}_1)] \\
&= \xi_1^{(z)} + \xi_2^{(z)} + \xi_3^{(z)}.
\end{aligned} \tag{2.2.35}$$

Now for some constant K' we have

$$E[| C_1 - E[C_1 \mid C_0 = z] |^k | C_0 = z] \leq K' \sum_{i=1}^{3} E[| \xi_i^{(z)} |^k | C_0 = z]. \tag{2.2.36}$$

For $i = 2, 3$, using Lemma 2.7 we have

$$E[| \xi_2^{(z)} |^k | C_0 = z] = z^k E[| f_z(\overline{S}) |^k | C_0 = z] = O(z^{k/2}), \tag{2.2.37}$$

and

$$\begin{aligned} & E[| \xi_3^{(z)} |^k | C_0 = z] = z^k \mid E[f_z(\overline{S}) \mid C_0 = z] \mid^k \\ & \leq z^k E[| f_z(\overline{S}) || C_0 = z]^k = O(z^{k/2}). \end{aligned} \tag{2.2.38}$$

Hence, to prove the lemma, we only need to show that $E[| \xi_1^{(z)} |^k | C_0 = z] = O(z^{k/2})$.

To show this, first notice that since $U_j = \sum_{l=1}^{X_j} \eta_{j,l}$, Lemma 2.5 and Assumption **A.3** (concerning the uniform boundedness of $E[\eta_{1,1}^k \mid \mathcal{G}_0]$) will give us

$$E[| U_j - E[U_j \mid \mathcal{G}_0] |^k | \mathcal{G}_0] = \begin{cases} O(X_j), & \text{if } k = 1, \\ \\ O(X_j^{k/2}), & \text{if } 2 \leq k \leq k_0, \end{cases} \tag{2.2.39}$$

where k_0 is as in Assumption **A.3**. Thus by the assumptions of the lemma and (2.2.39), $E[| U_j - E[U_j \mid \mathcal{G}_0] |^k | C_0 = z]$ is uniformly bounded.

For $k = 1$ by using Marcinkiewicz-Zygmund's inequality and concavity of the square root function we have

$$\begin{aligned} & E[| \xi_1^{(z)} || C_0 = z] = E[| \textstyle\sum_{j=1}^z U_j - E[U_j \mid \mathcal{G}_0] || C_0 = z] \\ & \leq E[\{\textstyle\sum_{j=1}^z (U_j - E[U_j \mid \mathcal{G}_0])^2\}^{1/2} \mid C_0 = z] \\ & \leq \{E[\textstyle\sum_{j=1}^z (U_j - E[U_j \mid \mathcal{G}_0])^2 \mid C_0 = z]\}^{1/2}. \end{aligned} \tag{2.2.40}$$

Now by (2.2.39) and Assumption **A.3**, for some constant $K'' > 0$ we can write (2.2.40) as

$$\begin{aligned} & E[| \xi_1^{(z)} || C_0 = z] \leq \{E[\textstyle\sum_{j=1}^z E[(U_j - E[U_j \mid \mathcal{G}_0])^2 \mid \mathcal{G}_0] \mid C_0 = z]\}^{1/2} \\ & \leq \{K'' E[\textstyle\sum_{j=1}^z X_j]\}^{1/2} = O(z^{1/2}). \end{aligned} \tag{2.2.41}$$

For $2 \leq k \leq k_0$ by using Marcinkiewicz-Zygmund's inequality and the Hölder inequality (as in the proof of Lemma 2.5) together with (2.2.39) we have

$$\begin{aligned} & E[| \xi_1^{(z)} || C_0 = z] = E[| \textstyle\sum_{j=1}^z U_j - E[U_j \mid \mathcal{G}_0] |^k | C_0 = z] \\ & \leq K_1 z^{\frac{k-2}{2}} \textstyle\sum_{j=1}^z E[| U_j - E[U_j \mid \mathcal{G}_0] |^k | C_0 = z] \\ & \leq K_2 z^{\frac{k-2}{2}} \textstyle\sum_{j=1}^z E[X_j^{k/2} \mid C_0 = z] = O(z^{k/2}). \end{aligned} \tag{2.2.42}$$

Hence we have proved our lemma. ■

Now, consider the following decomposition for the process $\{C_n\}_{n=0}^{\infty}$,

$$C_{n+1} = C_n + g(C_n) + \xi_{n+1}, \tag{2.2.43}$$

for $n \geq 0$, where g is a real-valued function and $g(C_n)$ corresponds to the drift of the process at the $(n+1)$th generation from the previous generation, further ξ_{n+1} is the residual term. Define the function g and the random residuals $\{\xi_n\}_{n=1}^{\infty}$ as

$$\begin{aligned} g(z) &:= E[C_1 \mid C_0 = z] - z, \quad \text{and} \\ \xi_{n+1} &:= C_{n+1} - E[C_{n+1} \mid \mathcal{F}_n], \end{aligned} \tag{2.2.44}$$

for $z \in \mathcal{Z}_+$ and $n \geq 0$.

The following lemma highlights some of the properties of the terms in the decomposition (2.2.43).

Lemma 2.9. *Assume that the assumptions* **A.1**-**A.4** *for Model 2.1 hold. Consider the following decomposition for the process* $\{C_n\}_{n=0}^{\infty}$,

$$C_{n+1} = C_n + g(C_n) + \xi_{n+1}, \tag{2.2.45}$$

with the function g and the residual ξ_{n+1} as in (2.2.44), for $n \geq 0$. Then we have

$$\begin{aligned} &(i) && E[\xi_{n+1} \mid \mathcal{F}_n] = 0, \\ &(ii) && E[\xi_{n+1}^2 \mid \mathcal{F}_n] = \sigma^2(C_n), \ \text{ for some function } \sigma \geq 0, \\ &(iii) && \text{for some constant } K > 0, \ \text{as } z \to \infty, \\ & && \text{we have } Kz \leq \sigma^2(z) = O(z), \ \text{ and } g(z) = O(z) \\ &(iv) && E[\mid \xi_1 \mid^{k_0} \mid C_0 = z] = O(\sigma^{k_0}(z)), \ \text{ for all } z \geq 1, \end{aligned} \tag{2.2.46}$$

with k_0 as in the assumption **A.3**.

Proof of Lemma 2.9: One can easily see that ξ_{n+1} satisfy (2.2.46)(i).

Given $\mathcal{F}_n$, the distribution of $\{U_{n+1,j}\}_{j=1}^{C_n}$ depends only on C_n, and since $C_{n+1} = \sum_{j=1}^{C_n} U_{n+1,j}$, (2.2.46)(ii) follows immediately.

By Lemma 2.8 and (2.2.44) we have $\sigma^2(z) = O(z)$. On the other hand, we can write $\sigma^2(z)$ as

$$
\begin{aligned}
\sigma^2(z) &= \mathrm{Var}[C_1 \mid C_0 = z] \\
&= E[\mathrm{Var}[C_1 \mid \mathcal{G}_0] \mid C_0 = z] + \mathrm{Var}[E[C_1 \mid \mathcal{G}_0] \mid C_0 = z] \\
&\geq E[\mathrm{Var}[C_1 \mid \mathcal{G}_0] \mid C_0 = z] \\
&= E[\mathrm{Var}[\textstyle\sum_{j=1}^{z} U_j \mid \mathcal{G}_0] \mid C_0 = z].
\end{aligned} \tag{2.2.47}
$$

By the conditional independence of $\{U_j\}_{j=1}^{z}$ given $\mathcal{G}_0$, we can write (2.2.47) as

$$
\sigma^2(z) \geq E[\sum_{j=1}^{z} \mathrm{Var}[U_j \mid \mathcal{G}_0] \mid C_0 = z] \tag{2.2.48}
$$

Since $U_j = \sum_{k=1}^{X_j} \eta_{j,k}$, with v_z as in (2.1.7), we have

$$
\begin{aligned}
\sigma^2(z) \quad &\geq E[\textstyle\sum_{j=1}^{z} X_j \nu_z(\overline{S}) \mid C_0 = z] \\
&= zE[\overline{X}\nu_z(\overline{S}) \mid C_0 = z].
\end{aligned} \tag{2.2.49}
$$

Now, by a method similar to that used in Lemma 2.7, we can show that $E[\overline{X}\nu_z(\overline{S}) \mid C_0 = z] = \mu_1(z)\nu_z(\mu(z)) + O(z^{-1/2})$. According to Assumption **A.1** for z large enough, there is a constant K such that

$$
\sigma^2(z) \geq K_1 z(\mu_1(z)\nu_z(\mu(z)) + O(z^{-1/2})) \geq Kz. \tag{2.2.50}
$$

This completes the proof of (2.2.46)(iii).

With the functions φ_z, φ^* as in Assumption **A.1**, respectively, and with the function f_z defined in (2.2.22). The function g defined in (2.2.44) is equal to

$$
\begin{aligned}
g(z) \quad &= E[C_1 \mid C_0 = z] - z = \textstyle\sum_{j=1}^{z}(E[U_j \mid C_0 = z] - 1) \\
&= z\{E[f_z(\overline{S}) \mid C_0 = z] + (\mu_1(z)\varphi_z(\mu(z)) - \mu_1^*\varphi(\mu^*)) \\
&+(\mu_1^*\varphi_z(\mu^*) - 1)\},
\end{aligned} \tag{2.2.51}
$$

where $\mu(z)$ and μ^* are as in Assumption **A.1**. Then by Lemma 2.7 we have

$$
\mid E[f_z(\overline{S}) \mid C_0 = z] \mid = O(z^{-1/2}). \tag{2.2.52}
$$

On the other hand, for z large enough, by using Assumption **A.4** together with Assumptions **A.1** and **A.2**, we can show that

$$
\begin{aligned}
&\mid \mu_1(z)\varphi_z(\mu(z)) - \mu_1^*\varphi_z(\mu^*) \mid = O(z^{-1/2}), \quad \text{and} \\
&\mu_1^*\varphi_z(\mu^*) - 1 = \mu_1^*\varphi(\mu^*) - 1 + o(1) > 0.
\end{aligned} \tag{2.2.53}
$$

Thus, by using (2.2.52) and (2.2.53) in (2.2.51), we get $g(z) = O(z)$ as $z \to \infty$ and this completes the proof of (2.2.46)(iii).

Finally, (2.2.46)(iv) follows from Lemma 2.8 and (2.2.46)(iii). ■

The following important lemmas are due to Klebaner. They play a crucial role in the proof of our main results.

Lemma 2.10. (Klebaner, 1984) *Let $h(x)$ be a positive, non-increasing function. Then for any constants $\alpha > 0$ and $\beta > 1$, we have*

$$\sum_{l=1}^{\infty} h(\alpha\beta^l) < \infty \quad \text{if and only if} \quad \sum_{l=1}^{\infty} \frac{h(l)}{l} < \infty. \tag{2.2.54}$$

Remark 2.6. This is Lemma 1 of Klebaner [24]. □

Lemma 2.11. (Klebaner, 1989) *Let h be a non-negative and non-increasing function on the interval $[l_0, \infty)$ with $l_0 > 0$ such that*

$$\sum \frac{h(n)}{n} < \infty. \tag{2.2.55}$$

Then there exists a $h^\diamond$ on $[l_0, \infty)$ such that $h^\diamond(x) \geq h(x)$, and $h^\diamond(x)$ is non-increasing, $xh^\diamond(x)$ is non-decreasing and concave for $x \geq l_0$ and $\sum h^\diamond(n)/n < \infty$.

Remark 2.7. This is Lemma 2 of Klebaner [28] where he proved this for $l_0 = 1$. For general $l_0 > 0$, his proof can be carried on if we take $h^\diamond(x)$, for $x \geq l_0$, to be

$$h^\diamond(x) = \frac{h(l_0)}{x} + \frac{1}{x} \cdot \int_{l_0}^{x} h(t)\, dt. \tag{2.2.56}$$

□

Lemma 2.12. (Klebaner, 1984) *Let $h(x)$ be a positive non-increasing function on $\mathcal{R}_+$, such that $xh(x)$ is non-decreasing and*

$$\sum_{n=1}^{\infty} \frac{h(n)}{n} < \infty. \tag{2.2.57}$$

Let $\{a_n\}$ be a sequence of positive numbers satisfying for some $K > 1$ and all $n \geq 0$

$$| a_{n+1} - a_n | \leq a_n h(a_n K^n). \tag{2.2.58}$$

Then

(i) $\lim_{n\to\infty} a_n = a$ *exist.*

(ii) There exists a constant z_0 depending only on the function h and K such that if $a_0 > z_0$ then $a > 0$.

Remark 2.8. This is Lemma 2 of [24]. □

Now we are ready for the proof of Theorem 2.3.

Proof of Theorem 2.3: Assertion (i) follows from Theorem 2(i) of Kersting [22], if we show that its conditions are satisfied by the process $\{C_n\}$. (2.2.43)-(2.2.44) together with Lemma 2.9 give the assumptions A1- A5 of Theorem 2(i) of [22]. It remains to show that for $\gamma > 2/k_0$, we have

$$\liminf_{z\to\infty} \frac{2zg(z)}{\sigma^2(z)}(\log(z))^{-\gamma} > 0, \tag{2.2.59}$$

where k_0 is as in **A.3**, the functions g and σ as in (2.2.44) and Lemma 2.9(ii), respectively.

Now, by Lemma 2.9(iii) we have $g(z) = O(z)$ as $z \to \infty$. But, by Lemma 2.9(iii) we have $Kz \le \sigma^2(z) = O(z)$. Thus $\frac{2zg(z)}{\sigma^2(z)} = O(z)$ and hence (2.2.59) holds. Therefore we proved the assertion (i) of the theorem.

With m and m_z as in (2.1.9) and (2.2.4), respectively, define the following function

$$\rho(x) := \mid m_1 - m \mid \mathbf{1}_{\{0\le x<1\}} + \mid m_{[x]} - m \mid \mathbf{1}_{\{x\ge 1\}}. \tag{2.2.60}$$

First we show that $\sum_n \rho(n)/n < \infty$. For $z > 0$, we have

$$\begin{aligned}\rho(z) &= \mid E[U_1 \mid C_0 = z] - m \mid \\ &= \mid E[E[U_1 \mid \mathcal{G}_0] \mid C_0 = z] - m \mid \\ &= \mid E[X_1\varphi_z(\overline{S}) \mid C_0 = z] - m \mid \\ &= \mid E[\overline{X}\varphi_z(\overline{S}) \mid C_0 = z] - m \mid .\end{aligned} \tag{2.2.61}$$

Now with f_z as in (2.2.22), by Lemma 2.7 and Assumption **A.2**, we have

$$\begin{aligned}\rho(z) &= \mid E[f_z(\overline{S}) \mid C_0 = z] + \mu_1(z)\varphi_z(\mu(z)) - \mu_1^*\varphi^*(\mu^*) \mid \\ &\le \mid E[f_z(\overline{S}) \mid C_0 = z] \mid + K_1 \mid \mu_1(z) - \mu_1^* \mid \\ &\quad + K_2 \mid \varphi_z(\mu(z)) - \varphi^*(\mu^*) \mid \\ &= O(z^{-1/2}) + O(z^{-\delta_1}) + O(z^{-\delta_2}) \\ &= O(z^{-(\delta_1\wedge\delta_2\wedge\frac{1}{2})}).\end{aligned} \tag{2.2.62}$$

Hence $\sum_{z=1}^{\infty} \rho(z)/z < \infty$.

To show the assertion (ii) of the theorem, we apply mostly the method used in Theorem 1(i) of Klebaner [25]. We can write the conditional expectation of $W_{n+1} = C_{n+1}/m^{n+1}$ as

$$\begin{aligned} & E[W_{n+1} \mid \mathcal{F}_n] = E[C_{n+1} \mid \mathcal{F}_n]/m^{n+1} \\ & = \{mC_n + C_n(m_{C_n} - m)\}/m^{n+1} = W_n + m^{-1}W_n(m_{C_n} - m). \end{aligned} \tag{2.2.63}$$

Now, from (2.2.63) we can write the absolute value of the difference of the expectation of $W_{n+1} - W_n$ as

$$\begin{aligned} & \mid E[W_{n+1}] - E[W_n] \mid = m^{-1} \mid E[W_n(m_{C_n} - m)] \mid \\ & \leq m^{-1}E[W_n\rho(C_n)] = m^{-1}E[W_n\rho(m^nW_n)]. \end{aligned} \tag{2.2.64}$$

Since by (2.2.62) ρ is a non-increasing function on $[1, \infty)$ and further we have $\sum_n \rho(n)/n < \infty$. Hence by Lemma 2.11 there exists a non-increasing function $\rho^\diamond$ on $[1, \infty)$, such that $\rho(x) \leq \rho^\diamond(x)$ and $x\rho^\diamond(x)$ in non-decreasing and concave for $x \geq 1$. By the same lemma $\sum_n \rho(n)/n < \infty$ implies that $\sum_n \rho^\diamond(n)/n < \infty$. Using these properties of $\rho^\diamond$, we can write (2.2.64) as

$$\begin{aligned} \mid E[W_{n+1}] - E[W_n] \mid \quad & \leq m^{-1}E[W_n\rho^\diamond(m^nW_n)] \\ & = m^{-n-1}E[m^nW_n\rho^\diamond(m^nW_n)] \\ & \leq m^{-1}E[W_n]\rho^\diamond(m^nE[W_n]). \end{aligned} \tag{2.2.65}$$

Let $\mathbf{m} = \inf_{n\geq 1} m_n > 0$. Then we have

$$E[W_{n+1}] \geq \frac{\mathbf{m}}{m}E[W_n] \geq \frac{\mathbf{m}^{n+1}}{m^{n+1}}C_0 \geq 0 \quad \text{a.s.,} \tag{2.2.66}$$

since C_0 is non-random. Now we use Lemma 2.12 with $a_{C_0,n} := E[W_n \mid C_0]$ and the function $h(x) = m^{-1}\rho^\diamond(x)$ for $x \geq 1$. Thus by Lemma 2.12, with $\alpha_{C_0} := \lim_{n\to\infty} a_{C_0,n}$ there exists a constant $c_0 > 0$ such that for all $C_0 > c_0$ we have $0 < \alpha_{C_0} < \infty$ and for all $C_0 \leq c_0$ we have $0 \leq \alpha_{C_0} < \infty$. We show that $\alpha_i > 0$ for all $i \geq 1$. For all $k \geq 1$, by the Markov property of $\{C_n\}_{n=0}^{\infty}$

and Fatou's lemma we get

$$\begin{aligned}
\alpha_i &= \lim_{n\to\infty} E[W_n \mid C_0 = i] \\
&= \lim_{n\to\infty} E[E[W_n \mid C_k, C_0 = i] \mid C_0 = i] \\
&= \lim_{n\to\infty} E[m^{-k} E[W_{n-k} \mid C_k] \mid C_0 = i] \qquad (2.2.67)\\
&\geq m^{-k} E[\lim_{n\to\infty} E[W_{n-k} \mid C_k] \mid C_0 = i] \\
&= m^{-k} E[\alpha_{C_k} \mid C_0 = i].
\end{aligned}$$

Now by the assertion (i) of the theorem we can take k large enough such that $P(C_k > c_0 \mid C_0 = i) > 0$ and on the set $\{C_k > c_0\}$ we have $0 < \alpha_{C_k} < \infty$. Hence (2.2.67) will be greater than

$$\begin{aligned}
\alpha_i \quad &\geq m^{-k} E[\alpha_{C_k}; C_k > c_0 \mid C_0 = i] \\
&= m^{-k} E[\alpha_{C_k} \mid C_k > c_0, C_0 = i] P(C_k > c_0 \mid C_0 = i) > 0.
\end{aligned} \qquad (2.2.68)$$

Thus $0 < \alpha_i < \infty$ for all $i \geq 1$.

To show the assertion (iii) of the theorem we show that the proof of Theorem 1(ii) of Klebaner [25] can be applied.

Taking expectation of both sides of (2.2.63), we obtain

$$\begin{aligned}
E[W_{n+1}] \quad &= E[W_n] + m^{-1} E[W_n(m_{C_n} - m)] \\
&= C_0 + m^{-1} E[\textstyle\sum_{v=0}^{n} W_v(m_{C_v} - m)].
\end{aligned} \qquad (2.2.69)$$

Now let $T_n := m^{-1} \sum_{v=0}^{n-1} W_v(m_{C_v} - m)$, and define the sequence $\{V_n\}$ by $V_n = W_n - T_n$. One can easily see that $\{V_n, \mathcal{F}_n\}_{n=0}^{\infty}$ is a martingale. We show that T_n and V_n converge a.s. to non-degenerate limits and this implies that W_n converges a.s.

Since, we have

$$\lim_{n\to\infty} \mid T_n \mid \leq \sum_{v=0}^{\infty} W_v \mid m_{C_v} - m \mid, \qquad (2.2.70)$$

to show that $\{T_n\}$ converges a.s., it is enough to show the right hand side of (2.2.70) converges a.s. If the function ρ is as in (2.2.60) and $\rho^\diamond$ its corresponding function as in Lemma 2.11 ($\rho(x) \leq \rho^\diamond(x)$ and $x\rho^\diamond(x)$ is non-decreasing and concave for all $x \geq 1$). Then by using this we can write

$$\begin{aligned}
&E[\textstyle\sum_{v=0}^{\infty} W_v \mid m_{C_v} - m \mid] = \sum_{v=0}^{\infty} E[W_v \rho(m^v W_v)] \\
&\leq \textstyle\sum_{v=0}^{\infty} E[W_v \rho^\diamond(m^v W_v)] \leq \sum_{v=0}^{\infty} E[W_v] \rho^\diamond(m^v E[W_v]).
\end{aligned} \qquad (2.2.71)$$

Since by the assertion (i) of the theorem, $E[W_n]$ converges to a non-zero limit, hence $E[W_n]$ is bounded from above by some constant $K^{(C_0)} > 0$. Thus by using the non-decreaseness of $x\rho^\diamond(x)$, we have

$$E[\sum_{v=0}^{\infty} W_v \mid m_{C_v} - m \mid] \leq \sum_{v=0}^{\infty} K_{C_0}\rho^\diamond(m^v K^{(C_0)}). \tag{2.2.72}$$

By Lemma 2.10, $\sum_{v=1}^{\infty} \rho^\diamond(m^v K^{(C_0)}) < \infty$ if and only if $\sum_{v=1}^{\infty} \rho^\diamond(v)/v < \infty$ and further by Lemma 2.11 $\sum_{v=1}^{\infty} \rho^\diamond(v)/v < \infty$ if $\sum_{v=1}^{\infty} \rho(v)/v < \infty$. But we have shown in (2.2.62) that $\sum_{v=1}^{\infty} \rho(v)/v < \infty$. Thus this and (2.2.72) imply that

$$E[\sum_{v=0}^{\infty} W_v \mid m_{C_v} - m \mid] < \infty, \tag{2.2.73}$$

and hence $\{T_n\}_{n=0}^{\infty}$ converges a.s. Since $\{V_n, \mathcal{F}_n\}$ is a martingale and

$$\sup_n \mid E[V_n] \mid < \infty, \tag{2.2.74}$$

the martingale convergence theorem implies that $V_n \to V$ a.s. for some r.v. V. Hence $W_n \to W$ a.s. and $E[W] < \infty$ by Fatou's lemma.

To show $\mathcal{L}_1$ convergence of $\{W_n\}$, we show that the proof of Theorem 2 of [25] holds. First notice that for $z \geq 1$, from Lemma 2.8 and (2.2.62), we have

$$\begin{aligned} &E[\mid C_1 - zm \mid\mid C_0 = z] \\ &\leq E[\mid C_1 - zm_z \mid\mid C_0 = z] + z \mid m_z - m \mid \\ &= E[\mid C_1 - E[C_1 \mid C_0 = z] \mid\mid C_0 = z] + z\rho(z) = O(z^{1/2}). \end{aligned} \tag{2.2.75}$$

If we define the function θ on $\mathcal{R}_+$ as

$$\theta(x) := \begin{cases} E[\mid C_1 - m \mid\mid C_0 = 1] + \rho(x), & \text{if } \ 0 \leq x < 1, \\ E[\mid C_1 - zm \mid\mid C_0 = z]/z + \rho(x), & \text{if } \ z = [x] \geq 1. \end{cases} \tag{2.2.76}$$

Then (2.2.75) and (2.2.62) imply that $\{\theta(z)\}_{z=1}^{\infty}$ is a non-increasing sequence and further we have $\sum_{z=1}^{\infty} \theta(z)/z < \infty$. Now we can write $E[\mid W_{n+1} - W_n \mid]$ as

$$\begin{aligned} &E[\mid W_{n+1} - W_n \mid] \\ &\leq E[\mid W_{n+1} - E[W_{n+1} \mid \mathcal{F}_n] \mid] + E[\mid E[W_{n+1} \mid \mathcal{F}_n] - W_n \mid] \\ &= \tfrac{1}{m^{n+1}} E[\mid C_{n+1} - E[C_{n+1} \mid \mathcal{F}_n] \mid] + \tfrac{1}{m^{n+1}} C_n \mid m_{C_n} - m \mid] \\ &= \tfrac{1}{m^{n+1}} E[C_n\theta(C_n)] = \tfrac{1}{m} E[W_n\theta(m^n W_n)]. \end{aligned} \tag{2.2.77}$$

By Lemma 2.11, there is a non-increasing function $\theta^\diamond$ such that $\theta(x) \le \theta^\diamond(x)$ and $x\theta^\diamond(x)$ is concave for all $x \ge 1$. Applying this to $E[W_n\theta(m^n W_n)]$, we get

$$\begin{aligned} E[W_n\theta(m^n W_n)] &\le E[W_n\theta^\diamond(m^n W_n)] \\ &\le E[W_n]\theta^\diamond(m^n E[W_n]). \end{aligned} \tag{2.2.78}$$

As before, $E[W_n]$ is bounded by some constant $K^{(C_0)} > 0$. Now by Lemma 2.10 and Lemma 2.11 we have:

$$\begin{aligned} \textstyle\sum_n \theta^\diamond(m^n K^{(C_0)}) < \infty \quad &\Longleftrightarrow \textstyle\sum_n \theta^\diamond(n)/n < \infty \\ &\Longleftarrow \textstyle\sum_n \theta(n)/n < \infty. \end{aligned} \tag{2.2.79}$$

Since we have $\sum_n \theta(n)/n < \infty$ this implies that $\sum_n \theta^\diamond(m^n K^{(C_0)}) < \infty$. This and (2.2.77)-(2.2.78) imply that

$$\begin{aligned} \sup_{j>n} E[\mid W_j - W_n \mid] \quad &\le \textstyle\sum_{j\ge n} E[\mid W_{j+1} - W_j \mid] \\ &\le K^{(C_0)} \textstyle\sum_{j>n} \theta^\diamond(m^j K^{(C_0)}) \to 0 \end{aligned} \tag{2.2.80}$$

as $n \to \infty$. Hence by Theorem 4.2.3(ii) of [6] (page 98), $\{W_n\}$ converges in $\mathcal{L}_1$ and almost surely to r.v. W and hence $E[W] = \lim_{n\to\infty} E[W_n] > 0$.

To show the last assertion of the theorem, we decompose $W_{n+1} = C_{n+1}/m^{n+1}$ as

$$\begin{aligned} W_{n+1} &= \tfrac{1}{m^{n+1}}\{mC_n + (E[C_{n+1} \mid \mathcal{F}_n] - mC_n) + \xi_{n+1}\} \\ &= W_n + m^{-1}W_n(m_{C_n} - m) + \xi_{n+1}/m^{n+1}, \end{aligned} \tag{2.2.81}$$

where $\xi_{n+1} = C_{n+1} - E[C_{n+1} \mid \mathcal{F}_n]$ as in (2.2.44). With the function $\rho(\cdot)$ as in (2.2.60) we can write $E[(W_{n+1} - W_n)^2 \mid \mathcal{F}_n]$ as

$$\begin{aligned} &E[(W_{n+1} - W_n)^2 \mid \mathcal{F}_n] = E[(m^{-1}W_n(m_{C_n} - m) + \xi_{n+1}/m^{n+1})^2 \mid \mathcal{F}_n] \\ &= W_n^2\rho^2(C_n)/m^2 + 2W_n(m_{C_n} - m)E[\xi_{n+1} \mid \mathcal{F}_n]/m^{n+2} \\ &+E[\xi_{n+1}^2 \mid \mathcal{F}_n]/m^{2n+2}. \end{aligned} \tag{2.2.82}$$

Now, by Lemma 2.12 we have $E[\xi_{n+1} \mid \mathcal{F}_n] = 0$ and $E[\xi_{n+1}^2 \mid \mathcal{F}_n] = \sigma^2(C_n) = O(C_n)$. Also from the additional assumption we made in the assertion (iv) (i.e, $\delta_1 \wedge \delta_2 \ge 1/2$) and (2.2.62) we have $\rho(z) = O(z^{-1/2})$. Hence we can write (2.2.82) as

$$\begin{aligned} E[(W_{n+1} - W_n)^2 \mid \mathcal{F}_n] \quad &\le K_1 C_n^{-1} W_n^2 + K_2 C_n/m^{2n} \\ &\le K_3 W_n/m^n, \end{aligned} \tag{2.2.83}$$

for some constants K_1, k_2 and K_3. Now by the assertion (i) of the theorem $E[W_n]$ converges. Hence, for some constant $K^{'} > 0$, $E[(W_{n+1} - W_n)^2]$ is bounded by

$$E[(W_{n+1} - W_n)^2] \leq K^{'}/m^n. \tag{2.2.84}$$

Since $m > 1$, we have $\sum_{n=0}^{\infty} E[(W_{n+1} - W_n)^2] < \infty$ and this implies that $\sup_{j>n} E[(W_j - W_n)^2]$ goes to zero as $n \to \infty$. Again by Theorem 4.2.3(ii) of [6], $\{W_n\}$ converges in $\mathcal{L}_2$ and almost surely to a r.v. W. ■

2.3. Model 2.1: Critical Case

In this section we study Model 2.1 in the critical case (see the definition 2.1). We need to present an assumption.

A. 5. *With $\Sigma(\cdot)$ as in (2.1.6) assume that there is a square matrix $\Sigma^* = [\sigma^*_{ij}]_{2\times 2}$, such that*

$$(i) \quad \Sigma(z) \to \Sigma^*, \quad \text{as} \quad z \to \infty. \tag{2.3.1}$$

Further, with $\mu(\cdot)$, μ^*, φ_z, φ^* and ν_z as in (2.1.6)-(2.1.8), assume that for some constants $\beta_i, K > 0$, a^*_i and $\delta_1 > 0$ as $z \to \infty$, we have

$$\begin{aligned} &(ii) \quad m = \mu^*_1 \varphi^*(\mu^*) = 1, \quad \text{and} \quad \| \mu(z) - \mu^* \| = O(z^{-\delta_1}), \\ &(iii) \quad \mu_1(z)\varphi_z(\mu(z)) = \mu_1(z)\varphi^*(\mu(z)) + a^*_1 z^{-\beta_1} + o(z^{-\beta_1}), \\ &(iv) \quad \mu_1(z)\varphi^*(\mu(z)) = 1 + a^*_2 z^{-\beta_2} + o(z^{-\beta_2}), \end{aligned} \tag{2.3.2}$$

where we take $\beta_i = \infty$ (or any value greater than one) whenever $a^*_i = 0$, for $i = 1, 2$. Also assume that there exist $K > 0$, $\epsilon > 0$ and $z_\epsilon > 0$, such that if *we let* $B = B_\epsilon(\mu^*)$, *then for all* $z \geq z_\epsilon$ *we have*

$$(v) \quad \sup_{y\in B} \left| \frac{\partial^3 \varphi_z(x)}{\partial x_{i_1}\partial x_{i_2}\partial x_{i_3}} \right|_{x=y} \Big| < K,$$

$$\sup_{y\in B} \left| \frac{\partial^2 \nu_z(x)}{\partial x_{i_1}\partial x_{i_2}} \right|_{x=y} \Big| < K,$$

$$\sup_{y\in B} \left| \frac{\partial^3 \nu_z(x)}{\partial x_{i_1}\partial x_2\partial x_{i_3}} \right|_{x=y} \Big| < K,$$

and as $z \to \infty$ we have (2.3.3)

$$(vi) \quad \frac{\partial \varphi_z(x)}{\partial x_{i_1}} \Big|_{x=\mu(z)} \longrightarrow \frac{\partial \varphi^*(x)}{\partial x_{i_1}} \Big|_{x=\mu^*},$$

and

$$\frac{\partial^2 \varphi_z(x)}{\partial x_{i_1}\partial x_{i_2}} \Big|_{x=\mu(z)} \longrightarrow \frac{\partial^2 \varphi^*(x)}{\partial x_{i_1}\partial x_{i_2}} \Big|_{x=\mu^*},$$

where $i_l \in \{1,2\}$ *for* $l = 1,2,3$.

We have the following lemma.

Lemma 2.13. *Assume* **A.1**, **A.3**-**A.5** *hold. With the functions* f_z *and* ν_z *as in (2.2.22) and (2.1.7), respectively, there exist constants* a_3^* *and* a_4^* *such that*

$$(i) \quad E[f_z(\overline{S}) \mid C_0 = z] = a_3^* z^{-1} + o(z^{-1}),$$

$$(ii) \quad E[f_z^2(\overline{S}) \mid C_0 = z] = a_4^* z^{-1} + o(z^{-1}), \tag{2.3.4}$$

$$(iii) \quad E[\overline{X}\nu_z(\overline{S}) \mid C_0 = z] = \mu_1(z)\nu_z(\mu(z)) + o(z^{-1}),$$

for $z \in \mathcal{Z}_+$.

Proof of Lemma 2.13: For a real-valued function h on $\mathcal{R}^2$, let

$$Dh(y) := [\frac{\partial h(x)}{\partial x_i} \mid_{x=y}]_{2\times 1}, \quad and \quad D^2h(y) := [\frac{\partial^2 h(x)}{\partial x_i \partial x_j} \mid_{x=y}]_{2\times 2}, \tag{2.3.5}$$

for $y \in \mathcal{R}^2$. For the simplicity, whenever we multiply a vector from the left, we assume it is a row vector and whenever it is multiplied from the right, we assume it is a column vector. Define f to be the real function such that

$$f(x) := x_1\varphi^*(x) - \mu_1^*\varphi^*(\mu^*), \tag{2.3.6}$$

for $x \in \mathcal{R}_+^2$. Then by Assumption **A.5**, for enough large z, we can write

$$D^2 f_z(\mu(z)) = D^2 f(\mu^*) + o_{2\times 2}(1) \tag{2.3.7}$$

where $o_{2\times 2}(1)$ is a 2×2 matrix which its elements are of order $o(1)$.

Now, we use the Taylor expansion of the function f_z in order to show the assertion (i). Let the function f_z be as in (2.2.22). As in Assumption **A.4**, let B be a closed disk in $\mathcal{R}_+^2$ centred at μ^* with radius $\epsilon > 0$ and $z_\epsilon > 0$ such that for all $z \geq z_\epsilon$, $\mu(z)$ is an interior point of B. Further, take $\epsilon_0 > 0$ small enough such that the closed disk B contains $\cup_{z\geq z_\epsilon} B_z$, where $B_z = B_{\epsilon_0}(\mu(z))$.

Then for all $x \in \mathcal{R}_+^2$ and $z > 0$, we can write $f_z(x)$ on B_z as

$$\begin{aligned} f_z(x) \quad &= (x-\mu(z))Df_z(\mu(z)) \\ &+\tfrac{1}{2}(x-\mu(z))D^2 f_z(\mu(z))(x-\mu(z)) \\ +R_z(x), \end{aligned} \tag{2.3.8}$$

where the remainder term, R_z, is equal to

$$R_z(x) = \frac{1}{6}\sum_{i_1=1}^{2}\sum_{i_2=1}^{2}\sum_{i_3=1}^{2} \frac{\partial^3 f_z(y)}{\partial y_{i_1}\partial y_{i_2}\partial y_{i_3}} \mid_{y=l_z(x)} \prod_{j=1}^{3}(x_{i_j} - \mu_{i_j}(z)), \tag{2.3.9}$$

where for given z and x, $l_z(x)$ is a point on the line segment joining x and $\mu(z)$.

By (2.3.7) and (2.3.8), we can rewrite $f_z(x)$ as

$$f_z(x) = (x-\mu(z))Df_z(\mu(z)) + \frac{1}{2}(x-\mu(z))D^2 f(\mu^*)(x-\mu(z)) + R_z^*(x), \tag{2.3.10}$$

where $R_z^*(x)$ is equal to

$$\begin{aligned} R_z^*(x) &= (x-\mu(z))o_{2\times 2}(1)(x-\mu(z)) + R_z(x)\mathbf{1}_{B_z}(x) \\ &+f_z(x)\mathbf{1}_{B_z^c}(x) - (x-\mu(z))Df_z(\mu(z))\mathbf{1}_{B_z^c}(x) \\ &-\tfrac{1}{2}(x-\mu(z))D^2 f_z(\mu(z))(x-\mu(z))\mathbf{1}_{B_z^c}(x). \end{aligned} \tag{2.3.11}$$

To show (i) we first show that

$$\mid E[R_z^*(\overline{S}) \mid C_0 = z] \mid = o(z^{-1}). \tag{2.3.12}$$

To do this, first notice that by the assumption **A.5** the first two terms of the right hand side of (2.3.11) are bounded by

$$\begin{aligned} &\mid (x-\mu(z))o_{2\times 2}(1)(x-\mu(z)) \mid \leq o(1) \parallel x-\mu(z) \parallel^2, \\ &\mid R_z(x)\mathbf{1}_{B_z}(x) \mid \leq K \parallel x-\mu(z) \parallel^3, \end{aligned} \tag{2.3.13}$$

for some constant $K > 0$. We can decompose $f_z(x)$ as

$$\begin{aligned} f_z(x) &= x_1\varphi_z(x) - \mu_1(z)\varphi_z(\mu(z)) \\ &= (x_1 - \mu_1(z))\varphi_z(x) + \mu_1(z)(\varphi_z(x) - \varphi_z(\mu(z))). \end{aligned} \tag{2.3.14}$$

Hence by the uniform boundedness of $\{\varphi_z\}$ (see Assumption **A.1**) and Assumption **A.5** for some constant $K > 0$ we have

$$\begin{aligned} &| f_z(x) | \leq K(\| x - \mu(z) \| + 1), \\ &| (x - \mu(z))Df_z(\mu(z)) | \leq K \| x - \mu(z) \| \quad \text{and} \\ &| (x - \mu(z))D^2 f_z(\mu(z))(x - \mu(z)) | \leq K \| x - \mu(z) \|^2 . \end{aligned} \tag{2.3.15}$$

By using (2.3.11) and (2.3.15) we can write

$$\begin{aligned} &| E[R_z^*(\overline{S}) \mid C_0 = z] | \leq E[| R_z^*(\overline{S}) || C_0 = z] \\ &\leq o(1)E[\| \overline{S} - \mu(z) \|^2 | C_0 = z] + KE[\| \overline{S} - \mu(z) \|^3 | C_0 = z] \\ &+KE[(\| \overline{S} - \mu(z) \| + 1); \overline{S} \in B_z^c \mid C_0 = z] \\ &+KE[\| \overline{S} - \mu(z) \|; \overline{S} \in B_z^c \mid C_0 = z] \\ &+KE[\| \overline{S} - \mu(z) \|^2; \overline{S} \in B_z^c \mid C_0 = z] \end{aligned} \tag{2.3.16}$$

Now by Lemma 2.6 and the Hölder inequality, we have

$$\begin{aligned} &(i) \quad E[\| \overline{S} - \mu(z) \|^2 | C_0 = z] = O(z^{-1}), \\ &(ii) \quad E[\| \overline{S} - \mu(z) \|^3 | C_0 = z] = O(z^{-3/2}), \\ &(iii) \quad P(\overline{S} \in B_z^c \mid C_0 = z) = O(z^{-k_0/2}), \\ &(iv) \quad E[\| \overline{S} - \mu(z) \|; \overline{S} \in B_z^c \mid C_0 = z] \\ &\qquad \leq (E[\| \overline{S} - \mu(z) \|^2 | C_0 = z])^{1/2}(P(\overline{S} \in B_z^c \mid C_0 = z))^{1/2} \\ &\qquad = O(z^{-1/2})O(z^{-k_0/4}) = O(z^{-(2+k_0)/4}), \\ &(v) \quad E[\| \overline{S} - \mu(z) \|^2; \overline{S} \in B_z^c \mid C_0 = z] \\ &\qquad \leq (E[\| \overline{S} - \mu(z) \|^3 | C_0 = z])^{2/3}(P(\overline{S} \in B_z^c \mid C_0 = z))^{1/3} \\ &\qquad = O(z^{-3/2})^{2/3}O(z^{-k_0/2})^{1/3} = O(z^{-(1+k_0/6)}), \end{aligned} \tag{2.3.17}$$

where k_0 is as in Assumption **A.3**. Since $k_0 \geq 3$, $k_0/2$, $(2+k_0)/4$ and $1+k_0/6$ are greater than one. Replacing (2.3.17) in (2.3.16) we get (2.3.12).

Thus we can write $E[f_z(\overline{S}) \mid C_0 = z]$ as

$$\begin{aligned}
&E[f_z(\overline{S}) \mid C_0 = z] = E[(\overline{S} - \mu(z))Df(\mu^*) \mid C_0 = z] \\
&+\tfrac{1}{2}E[(\overline{S} - \mu(z))D^2 f(\mu^*)(\overline{S} - \mu(z)) \mid C_0 = z] \\
&+E[R_z^*(\overline{S}) \mid C_0 = z] \\
&= \tfrac{1}{2}E[(\overline{S} - \mu(z))D^2 f(\mu^*)(\overline{S} - \mu(z)) \mid C_0 = z] + o(z^{-1}) \\
&= \tfrac{1}{2z}E[(S_1 - \mu(z))D^2 f(\mu^*)(S_1 - \mu(z)) \mid C_0 = z] + o(z^{-1}) \\
&= \tfrac{1}{2z}\{\textstyle\sum_{i=1}^2 \sum_{j=1}^2 \sigma_{i,j}(z)(D^2 f(\mu^*))_{ij}\} + o(z^{-1}) \\
&= \tfrac{1}{2z}\textstyle\sum_{i=1}^2 \sum_{j=1}^2 (\sigma_{i,j}^* + o(1))(D^2 f(\mu^*))_{ij} + o(z^{-1}) \\
&= \tfrac{1}{2z}\textstyle\sum_{i=1}^2 \sum_{j=1}^2 \sigma_{i,j}^* D^2 f(\mu^*))_{ij} + o(z^{-1}).
\end{aligned} \tag{2.3.18}$$

If we let zero mean random vector Z have covariance matrix Σ^*, then the constant a_3^* in (2.3.4) is equal to

$$\begin{aligned}
a_3^* &= \tfrac{1}{2}\textstyle\sum_{i=1}^2 \sum_{j=1}^2 \sigma_{i,j}^*(D^2 f(\mu^*))_{ij} \\
&= \tfrac{1}{2}E[ZD^2 f(\mu^*)Z].
\end{aligned} \tag{2.3.19}$$

Similarly, we can write $E[f_z^2(\overline{S}) \mid C_0 = z]$ as

$$E[f_z^2(\overline{S}) \mid C_0 = z] = z^{-1}a_4^* + o(z^{-1}). \tag{2.3.20}$$

where with Z a random vector as above, a_4^* is equal to

$$\begin{aligned}
a_4^* &= \tfrac{1}{2}\textstyle\sum_{i=1}^2 \sum_{j=1}^2 \sigma_{i,j}^*(D^2 f^2(\mu^*))_{ij} \\
&= \tfrac{1}{2}E[ZD^2 f^2(\mu^*)Z].
\end{aligned} \tag{2.3.21}$$

It is easy to see that

$$(D^2 f^2(\mu^*))_{ij} = 2(Df(\mu^*))_i(Df(\mu^*))_j. \tag{2.3.22}$$

Hence we have

$$\begin{aligned}
a_4^* &= \textstyle\sum_{i=1}^2 \sum_{j=1}^2 \sigma_{i,j}^*(Df(\mu^*))_i(Df(\mu^*))_j \\
&= E[(ZDf(\mu^*))^2].
\end{aligned} \tag{2.3.23}$$

Finally, in the similar way, the Taylor expansion can be used to prove the assertion (iii). ■

Remark 2.9. As one can easily see the assumptions on Model 2.1 in the critical case are stronger than those of the supercritical case. Hence the assertions of Lemma 2.8 also hold for the critical case. □

Now, consider the decomposition of C_{n+1} as it is given in (2.2.43), i.e. $C_{n+1} = C_n + g(C_n) + \xi_{n+1}$. The following lemma, similar to Lemma 2.9, presents some of the properties of the elements of this decomposition for Model 2.1 in the critical case.

Lemma 2.14. *Assume* **A.1**, **A.3**-**A.5** *hold. Consider the following decomposition for the process* $\{C_n\}_{n=0}^{\infty}$,

$$C_{n+1} = C_n + g(C_n) + \xi_{n+1}, \tag{2.3.24}$$

with the function g *and the residuals* ξ_{n+1} *as in (2.2.44), for* $n \geq 0$. *Then*

$$\begin{aligned}
&(i) \quad && E[\xi_{n+1} \mid \mathcal{F}_n] = 0, \\
&(ii) \quad && E[\xi_{n+1}^2 \mid \mathcal{F}_n] = \sigma^2(C_n), \;\; \text{for some function } \sigma \geq 0, \\
&(iii) \quad && E[\mid \xi_1 \mid^{k_0} \mid C_0 = z] = O(\sigma^{k_0}(z)), \;\; \text{for all } z \geq 1, \\
&(iv) \quad && \text{for some constants } b^* \text{ and } 0 \leq \beta_0 < 1, \\
& && \quad \text{as } z \to \infty, \text{ we have } g(z) = b^* z^{\beta_0} + o(z^{\beta_0}), \\
&(v) \quad && \text{and for some constant } c^* > 0, \text{ we have} \\
& && \sigma^2(z) = c^* z + o(z),
\end{aligned} \tag{2.3.25}$$

with k_0 *as in Assumption* **A.3**.

Proof of Lemma 2.14: Since the assumptions of Lemma 2.14 are stronger than the assumptions of Lemma 2.9, the proofs of (2.3.25)(i)-(iii) are the same as (2.2.46)(i)-(ii) and (iv) in Lemma 2.9. Hence we only need to prove (2.3.25)(iv)-(v).

To show (2.3.25)(iv), we start with the function g (as in (2.2.44)). We have

$$\begin{aligned}
g(z) &= E[C_1 \mid C_0 = z] - z \\
&= E[E[\textstyle\sum_{j=1}^{z} U_j \mid \mathcal{G}_0] \mid C_0 = z] - z \\
&= z E[\overline{X} \varphi_z(\overline{S}) \mid C_0 = z] - z.
\end{aligned} \tag{2.3.26}$$

First notice that since Model 2.1 is in the critical case we have $\mu_1^*\varphi^*(\mu^*) = 1$. Now, with the function f_z as in (2.2.22) and using (2.3.2)(ii)-(iii) and (2.3.4)(i), we can write (2.3.26) as

$$\begin{aligned} g(z) &= z\{E[f_z(\overline{S}) \mid C_0 = z] + (\mu_1(z)\varphi_z(\mu(z)) - \mu_1(z)\varphi^*(\mu(z))) \\ &+(\mu_1(z)\varphi^*(\mu(z)) - 1)\} \\ &= b^* z^{\beta_0} + o(z^{\beta_0}), \end{aligned} \tag{2.3.27}$$

where β_0 and b^* are equal to

$$\begin{aligned} &\beta_0 = \max\{0, 1-\beta_1, 1-\beta_2\}, \\ \text{and} \quad &b^* = \textstyle\sum_{i=1}^2 a_i^* \mathbf{1}_{\{\beta_0 = 1-\beta_i\}} + a_3^* \mathbf{1}_{\{\beta_0 = 0\}}, \end{aligned} \tag{2.3.28}$$

where a_i^* and β_j are as in (2.3.2)(ii)-(iii) and (2.3.4)(i), for $i = 1, 2, 3$ and $j = 1, 2$. This proves (2.3.25)(iv).

To show (2.3.25)(v), we can write $\sigma^2(z)$ as

$$\begin{aligned} \sigma^2(z) &= \mathrm{Var}[\xi_1 \mid C_0 = z] \\ &= \mathrm{Var}[E[\xi_1 \mid \mathcal{G}_0] \mid C_0 = z] + E[\mathrm{Var}[\xi_1 \mid \mathcal{G}_0] \mid C_0 = z]. \end{aligned} \tag{2.3.29}$$

With the function f_z as in (2.2.22) and ξ_1 as in (2.2.44), and together with the fact that $E[\xi_1 \mid \mathcal{G}_0] = z(f_z(\overline{S}) - E[f_z(\overline{S}) \mid C_0 = z])$, we can write $\mathrm{Var}[E[\xi_1 \mid \mathcal{G}_0] \mid C_0 = z]$ as

$$\begin{aligned} &\mathrm{Var}[E[\xi_1 \mid \mathcal{G}_0] \mid C_0 = z] = \mathrm{Var}[z f_z(\overline{S}) \mid C_0 = z] \\ &= z^2\{E[f_z^2(\overline{S}) \mid C_0 = z] - E[f_z(\overline{S}) \mid C_0 = z]^2\}. \end{aligned} \tag{2.3.30}$$

Now, using Lemma 2.13(i)-(ii), (2.3.30) will be equal to

$$\begin{aligned} &\mathrm{Var}[E[\xi_1 \mid \mathcal{G}_0] \mid C_0 = z] = z^2\{a_4^* z^{-1} + o(z^{-1}) \\ &-(a_3^* z^{-1} + o(z^{-1}))^2\} = a_4^* z + o(z). \end{aligned} \tag{2.3.31}$$

It is easy to show that $\mathrm{Var}[\xi_1 \mid \mathcal{G}_0] = z\overline{X}\nu_z(\overline{S})$. Using this and Lemma 2.13(iii) and noticing that we have $\mu_1(z)\nu_z(\mu(z)) = \mu_1^*\nu^*(\mu^*) + o(1)$, we can write $E[\mathrm{Var}[\xi_1 \mid \mathcal{G}_0] \mid C_0 = z]$ as

$$\begin{aligned} &E[\mathrm{Var}[\xi_1 \mid \mathcal{G}_0] \mid C_0 = z] = zE[\overline{X}\nu_z(\overline{S}) \mid C_0 = z] \\ &= z\{\mu_1(z)\nu_z(\mu(z)) + O(z^{-1})\} = z\{\mu_1^*\nu^*(\mu^*) + o(1)\} \\ &= \mu_1^*\nu^*(\mu^*) z + o(z). \end{aligned} \tag{2.3.32}$$

Thus by replacing (2.3.31) and (2.3.32) in (2.3.29) we get

$$\sigma^2(z) = c^* z + o(z), \tag{2.3.33}$$

with c^* equal to

$$c^* = a_4^* + \mu_1^* \nu^*(\mu^*). \tag{2.3.34}$$

■

Remark 2.10. Let us now have a closer look at the parameters b^*, β_0 and c^*.

For $z \in \mathcal{Z}_+$ and $x = (x_1, x_2) \in \mathcal{R}_+^2$ we have $E[C_1 \mid C_0 = z, \overline{S} = x] = z x_1 \varphi_z(x)$. On one hand as z becomes larger, a Law of Large Numbers says that $\overline{S}$ converges to $\mu(z)$ (or in fact to μ^*) almost surely, and by our assumption $\mu(z)$ converges to μ^*. On the other hand, the rate of convergence of $z\mu_1(z)\varphi_z(\mu(z))$, $z\mu_1(z)\varphi^*(\mu(z))$ and $E[z\overline{X}_1\varphi_z(\overline{S}) \mid C_0 = z]$ to each other and to $z\mu_1^*\varphi^*(\mu^*) = mz = z$, determine the values of the parameters b^* and β_0. By Assumption **A.5** and Lemma 2.13, we have

$$\begin{aligned}
&(i) \quad & z\mu_1(z)\varphi_z(\mu(z)) &= z\mu_1(z)\varphi^*(\mu(z)) + a_1^* z^{1-\beta_1} + o(z^{1-\beta_1}),\\
&(ii) \quad & z\mu_1(z)\varphi^*(\mu(z)) &= z\mu_1^*\varphi^*(\mu^*) + a_2^* z^{1-\beta_2} + o(z^{1-\beta_2})\\
& & &= z + a_2^* z^{1-\beta_2} + o(z^{1-\beta_2}),\\
&(iii) \quad & E[C_1 \mid C_0 = z] &= E[z\overline{X}\varphi_z(\overline{S}) \mid C_0 = z]\\
& & &= z\mu_1(z)\varphi_z(\mu(z)) + a_3^* + o(1),
\end{aligned} \tag{2.3.35}$$

where

$$a_3^* = \frac{1}{2}\sum_{i,j=1}^{2} \sigma_{i,j}^* \left(\frac{\partial^2 x_1 \varphi^*(x)}{\partial x_i \partial x_j} \,|_{x=\mu^*}\right). \tag{2.3.36}$$

One can interpret (2.3.35)(i)-(iii) as follows.

(i) Assume for a moment that the average numbers of children born to z couples, i.e. $\overline{S}$, is equal to the expected one, i.e. $\mu(z)$. Then $z^{-\beta_1}$ in (2.3.35)(i) is the rate of convergence for the conditional expected number of marriages for a female when there are z families present, i.e. $\varphi_z(\mu(z))$, to the case when there are infinity many families present, i.e. $\varphi^*(\mu(z))$.

(ii) Assume for a moment that there is an infinite many families present. Then $z^{-\beta_2}$ in (2.3.35)(ii) is the rate of convergence for the conditional expected number of marriages for a female when the average number of children born, i.e. $\overline{S}$, is equal to $\mu(z)$, to the case when this average is equal to μ^*.

(iii) a_3^* is the deviation of the expected number of marriages of the daughters of z families from $z\mu_1(z)\varphi_z(\mu(z))$. From (2.3.36) one can see that a_3^* determines by a combination of variance and covariance of the limiting reproduction law, i.e. $\Sigma^* = [\sigma_{i,j}^*]$, and the behavior of the function φ^* in the neighbourhood of μ^*, the limiting expected number of the children born from a couple when infinity number of couples are around.

As in the proof of Lemma 2.14(iv), by replacing (2.3.35)(i) and (ii) to the right hand side of (2.3.35)(iii) we get

$$\begin{aligned} \beta_0 &= \max\{0, 1-\beta_1, 1-\beta_2\}, \\ b^* &= \textstyle\sum_{i=1}^2 a_i^* \mathbf{1}_{\{\beta_0=1-\beta_i\}} + a_3^* \mathbf{1}_{\{\beta_0=0\}}. \end{aligned} \tag{2.3.37}$$

Now we explore the value of the parameter c^*. We can break up $\mathrm{Var}[C_1 \mid C_0 = z]$ into the two parts as follows

$$\begin{aligned} \mathrm{Var}[C_1 \mid C_0 = z] &= \mathrm{Var}[z\overline{X}\varphi_z(\overline{S}) \mid C_0 = z] \\ &\quad + E[z\overline{X}\nu_z(\overline{S}) \mid C_0 = z]. \end{aligned} \tag{2.3.38}$$

We can show that each of the terms in the right hand side of (2.3.38) are equal to (see Lemma 2.13(v))

$$\begin{aligned} &(i) \quad \mathrm{Var}[z\overline{X}\varphi_z(\overline{S}) \mid C_0 = z] = a_4^* z + o(z), \\ &(ii) \quad E[z\overline{X}\nu_z(\overline{S}) \mid C_0 = z] = \mu_1^* \nu^*(\mu^*) z + o(z), \end{aligned} \tag{2.3.39}$$

where a_4^* in (2.3.39)(i) is equal to

$$a_4^* = \sum_{i,j=1}^{2} \sigma_{i,j}^* \left(\frac{\partial x_1 \varphi^*(x)}{\partial x_i} \Big|_{x=\mu^*} \frac{\partial x_1 \varphi^*(x)}{\partial x_j} \Big|_{x=\mu^*} \right). \tag{2.3.40}$$

From (2.3.35)-(2.3.40) one can see how the values of the parameters μ^* and $\Sigma^* = [\sigma_{i,j}^*]$ (corresponding to the reproduction stage) together with behaviors of the limiting functions φ^* and ν^* at the point μ^* and its neighbourhood (corresponding to the mating stage), determine the values of the parameters β_0, b^* and c^*. □

Consider the following deterministic equation.

$$a_{n+1} = a_n + g(a_n), \quad \text{with } a_0 = 1. \tag{2.3.41}$$

Let $\{a_n\}_{n=0}^{\infty}$ be the solution to this equation.

The following theorem presents the main results of this section.

Theorem 2.15. *Consider Model 2.1 in the critical case, which satisfies the assumptions* **A.1** *and* **A.3**-**A.5**. *By Lemma 2.14(iv)-(v), $g(x) = b^* x^{\beta_0} + o(x^{\beta_0})$ and $\sigma^2(z) = c^* z + o(z)$, where $b^* \in \mathcal{R}$, $0 \le \beta_0 < 1$ and $c^* > 0$.*

(i) If $b^* \le 0$, then the process $\{C_n\}_n$ dies out almost surely, i.e. $P(C_n \to 0) = 1$.

(ii) If either $\beta_0 = 0$ and $2b^* > c^*$ or $0 < \beta_0 < 1$ and $b^* > 0$, then we have

$$P(C_n \to 0 \text{ or } \infty) = 1 \text{ and } P(C_n \to \infty) > 0. \tag{2.3.42}$$

With $\{a_n\}_{n=0}^{\infty}$ as in (2.3.41), and $\mathcal{E}_\infty = \{C_n \to \infty\}$, we have the following.

(iii) If $\beta_0 = 0$ and $2b^* > c^*$, then

$$P(C_n/n \le t \mid \mathcal{E}_\infty) \to \frac{2}{c^*\Gamma(\frac{2b^*}{c^*})} \int_0^t (\tfrac{2x}{c^*})^{\frac{2b^*}{c^*}-1} e^{-\frac{2x}{c^*}} dx, \tag{2.3.43}$$

as $n \to \infty$.

(iv) If $b^* > 0$ and $0 < \beta_0 \le 2/3$, then $C_n/a_n \to 1$ in probability and the conditional distribution of $K_{\beta_0,n}^{-1}(C_n - a_n)/g(a_n)$ given $\mathcal{E}_\infty$ converges weakly to a standard normal distribution $\mathcal{N}(0,1)$, i.e.

$$\mathcal{D}(K_{\beta_0,n}^{-1} \tfrac{C_n - a_n}{g(a_n)} \mid \mathcal{E}_\infty) \overset{\text{weakly}}{\longrightarrow} \mathcal{N}(0,1), \tag{2.3.44}$$

where $K_{\beta_0,n} > 0$ is equal to

$$K_{\beta_0,n} \approx \begin{cases} \sqrt{\frac{c^*}{2-3\beta_0}} (b^*)^{\frac{1}{2(1-\beta_0)}} (n(1-\beta_0))^{\frac{2-3\beta_0}{2(1-\beta_0)}} & \text{if } 0 < \beta_0 < 2/3, \\ \sqrt{\frac{3c^*}{(b^*)^3}} \sqrt{\log n}, & \text{if } \beta_0 = 2/3. \end{cases} \tag{2.3.45}$$

(v) If $b^* > 0$ and $2/3 < \beta_0 < 1$, then

$$C_n/a_n \to 1 \text{ and } \frac{C_n - a_n}{g(a_n)} \text{ converges}, \tag{2.3.46}$$

a.s. on $\mathcal{E}_\infty$ as $n \to \infty$.

Before presenting the proof of Theorem 2.15 we have the following remark.

Remark 2.11. **(1)** The process $\{C_n\}$ is a Markov chain, i.e. given C_n we can determine the distribution of C_{n+1} for $n = 0, 1, 2, \ldots$.

(2) The function g, given in the decomposition (2.3.24), can be thought of as the drift of the Markov chain $\{C_n\}$ at each step. Lemma 2.14 gives the form of this drift to be $g(z) = b^* z^{\beta_0} + o(z^{\beta_0})$ where b^* is a real number and $0 \leq \beta_0 < 1$. The assertions (i)-(v) of Theorem 2.15 characterise the kinds of limiting behaviours the process $\{C_n\}$ might have.

(3) The assertion (i) says that a process which has a negative drift does not survive and this is irrespective of the size of the variance of the residuals, i.e. c^*.

(4) The first conclusion of the assertion (ii) is this, that either the process dies out or it grows to infinity. Secondly, it gives conditions under which the process can grow (to infinity) with positive probability.

If $\beta_0 = 0$ and consequently $g(z) \simeq b^*$, then the process grows to infinity whenever the relative size of the variance of the residuals, i.e. $\frac{1}{z}E[\xi_1^2 \mid C_0 = z] \simeq c^*$, does not exceed twice the size of the drift of the process, i.e. $2b^*$. Otherwise, with probability one, the process eventually hits the zero and dies out.

If $0 < \beta_0 < 1$, then with positive probability the process survives.

(5) If $\beta_0 = 0$, then the assertion (iii) says that C_n/n would have approximately a Gamma distribution with the scale parameter $c^*/2$ and the shape parameter $2b^*/c^*$.

(6) If $0 < \beta_0 \leq 2/3$, then the assertion (iv) says that in probability the process C_n converges to its deterministic counterpart, i.e. a_n (given in (2.3.41)). And its deviation from $\{a_n\}$, after it has been appropriately normalised, converges weakly to a standard normal distribution. Hence, for a given generation we can construct an interval which contains the process with certain probability. This interval is centered at the deterministic solution, i.e. $\{a_n\}$.

(7) If $2/3 < \beta_0 < 1$, then the assertion (v) says that almost surely the process converges to the deterministic solution, with a rate which is not slower than $g(a_n)$.

□

Proof of Theorem 2.15: Part (i) of the Theorem follows from Theorem 1(i) of Kersting [22]. To see this, notice that Lemma 2.14 gives the assumption A.1-A.5 of [23]. To use Kersting's theorem, it is enough to show that

$$\limsup_{x\to\infty} \frac{2xg(x)}{\sigma^2(x)} \le 0, \quad \text{and} \quad g(x) = O(x). \tag{2.3.47}$$

But from Lemma 2.14 we have

$$\lim_{x\to\infty} \frac{2xg(x)}{\sigma^2(x)} = \lim_{x\to\infty} \frac{2b^* x^{\beta_0}}{c^*} \le 0, \tag{2.3.48}$$

since $b^* \le 0$ and $g(x) = O(x^{\beta_0})$ with $0 \le \beta_0 < 1$.

Part (ii) of the Theorem follows from Theorem 2(ii) of Kersting [22]. To see this, notice that Lemma 2.14 gives the assumption A.1-A.5 of [23]. It is enough to show that

$$\liminf_{x\to\infty} \frac{2xg(x)}{\sigma^2(x)} > 1, \quad \text{and} \quad g(x) = O(\frac{x}{\log^2(x)}). \tag{2.3.49}$$

But from Lemma 2.14 we have

$$\lim_{x\to\infty} \frac{2xg(x)}{\sigma^2(x)} = \begin{cases} \frac{2b^*}{c^*}, & \text{if } \beta_0 = 0, \\ \\ \infty, & \text{if } 0 < \beta_0 < 1. \end{cases} \tag{2.3.50}$$

Since by the assumption of the theorem $2b^* > c^*$ when $\beta_0 = 0$, we have $\liminf_{x\to\infty} \frac{2xg(x)}{\sigma^2(x)} > 1$. The second assertion in (2.3.49) is obvious. Hence, by Theorem 2(ii) of [22], for some constant $c \ge 0$ we have

$$P(C_n \to \infty \text{ or } \limsup_{n\to\infty} C_n \le c) = 1 \text{ and } P(C_n \to \infty) > 0. \tag{2.3.51}$$

But since zero is the only absorbing state we have $\{\limsup_{n\to\infty} C_n \le c\} \equiv \{\lim_{n\to\infty} C_n = 0\}$.

Part (iii) of the theorem follows from Theorem 1 of Kersting [23]. Define the function G on $[1, \infty)$ as follows

$$G(x) = \int_1^x \frac{dy}{g(y)}. \tag{2.3.52}$$

With the function g as in Lemma 2.14, for large x we have $g(x) \approx b^* x^{\beta_0}$. Hence for large x, $G(x)$ is approximately equal to

$$G(x) \approx \frac{x^{1-\beta_0}}{b^*(1-\beta_0)}. \tag{2.3.53}$$

To show that the assumptions of Theorem 1 of [23] hold, we first notice that $(G(x))^\lambda$ is ultimately convex for $\lambda \ge \frac{1}{1-\beta_0}$ and ultimately concave for $0 \le \lambda < \frac{1}{1-\beta_0}$. To see this, let us define the function H as follows

$$H(x) := \frac{x^{\lambda(1-\beta_0)}}{(b^*(1-\beta_0))^\lambda}. \tag{2.3.54}$$

Then, for $x > 0$,

$$\frac{dH(x)}{dx} \begin{cases} \text{is increasing if } \lambda \geq \frac{1}{1-\beta_0}, \\ \\ \text{is decreasing if } 0 \leq \lambda < \frac{1}{1-\beta_0}. \end{cases} \tag{2.3.55}$$

Now, by Theorem 1.4. of [34], the function $H(x)$ is convex for $\lambda \geq \frac{1}{1-\beta_0}$ and concave for $0 \leq \lambda < \frac{1}{1-\beta_0}$.

For the case when $\beta_0 = 0$, this together with Lemma 2.14 gives the assumptions A.1-A.4 of Theorem 1 of [23]. Also we have

$$\frac{\sigma^2(x)}{g^2(x)G(x)} \approx \frac{c^* x}{(b^*)^2 \frac{1}{b^*}(x-1)} \to \frac{c^*}{b^*} < 2, \tag{2.3.56}$$

as $x \to \infty$. By Theorem 1 of [23] we have

$$P(G(C_n)/n \leq t \mid \mathcal{E}_\infty) \to \frac{2b^*}{c^*\Gamma(\frac{2b^*}{c^*})} \int_0^t (\frac{2b^* x}{c^*})^{\frac{2b^*}{c^*}-1} e^{-\frac{2b^* x}{c^*}} dx, \tag{2.3.57}$$

as $n \to \infty$. But since for large x, $G(x) \approx x/b^*$, we have

$$P(C_n/n \leq t \mid \mathcal{E}_\infty) \to \frac{2}{c^*\Gamma(\frac{2b^*}{c^*})} \int_0^t (\frac{2x}{c^*})^{\frac{2b^*}{c^*}-1} e^{-\frac{2x}{c^*}} dx \tag{2.3.58}$$

as $n \to \infty$. This completes the proof of part (ii) of the theorem.

Parts (iv) and (v) of the theorem follow from Theorem 3, parts (b) and (a) of Keller et al. [20], respectively. To show that the conditions of their theorem hold, we need to introduce some notations. Let us define the functions s^2 and I on $\mathcal{R}_+$ as follows. For $x \geq 1$ we have

$$\begin{aligned} s^2(x) &:= \frac{\sigma^2(x)}{g^2(x)} \approx \frac{c^*}{(b^*)^2} x^{1-2\beta_0}, \\ I(x) &:= \int_1^x \frac{s^2(y)}{g(y)} dy \\ &\approx \begin{cases} \frac{c^*}{(b^*)^3(2-3\beta_0)}(x^{2-3\beta_0} - 1), & \text{if } 0 < \beta_0 < 1 \\ & \text{and } \beta_0 \neq \frac{2}{3} \\ \frac{c^*}{(b^*)^3} \ln(x), & \text{if } \beta_0 = \frac{2}{3}. \end{cases} \end{aligned} \tag{2.3.59}$$

(Notice that the functions s^2 and I are the same as the functions σ^2 and ψ as in [20], respectively.) In the above we assume that σ and g are smooth functions. If we let G^{-1} denotes the inverse function of the function G defined in (2.3.52), for any function f on $\mathcal{R}_+$, $\hat{f}$ denotes the function given by $\hat{f}(x) = f \circ G^{-1}(x)$. It is easy to see that for x large enough, and g as before we have

$$\hat{I}(x) = I \circ G^{-1}(x)$$

$$\approx \begin{cases} \frac{c^*}{2-3\beta_0}(b^*)^{\frac{1}{1-\beta_0}}\{(1-\beta_0)x\}^{\frac{2-3\beta_0}{1-\beta_0}}, & \text{if } 0<\beta_0<1 \\ & \text{and } \beta_0 \neq \frac{2}{3} \\ \frac{3c^*}{(b^*)^3}\log x, & \text{if } \beta_0 = \frac{2}{3}. \end{cases} \tag{2.3.60}$$

Now we show that the assumptions of Theorem 3 of [20] hold when $0 < \beta_0 < 1$.

- With $g(x) = b^* x^{\beta_0} + o(x^{\beta_0})$ as given in (2.3.25)(iv) which is positive and twice differentiable on $\mathcal{R}_+$ (we let $g(x) = 0$ for $x < 0$). g is ultimately concave and $g^{'}$ is ultimately convex. Thus the assumption A.1 of Theorem 3 of [20] holds.
- Obviously, the function s^2 given in (2.3.59) is continuously differentiable. Further we have

$$\int_1^\infty \hat{s}^2(x)x^{-2}dx = \int_1^\infty s^2 \circ G^{-1}(x)x^{-2}dx$$

$$\approx \int_1^\infty \frac{c^*}{(b^*)^2x^2}\{b^*(1-\beta_0)x+1\}^{\frac{1-2\beta_0}{1-\beta_0}}dx \tag{2.3.61}$$

$$\leq K_1\int_1^\infty x^{-1-\frac{\beta_0}{1-\beta_0}}dx < \infty,$$

  for some constant $K_1 > 0$. Thus the Assumption A.2 of Theorem 3 of [20] holds.
- For the case where $0 < \beta_0 \leq \frac{2}{3}$, from (2.3.59) we have $I(\infty) = \infty$ and further $\mid g^{''}(x)g(x)/s^2(x) \mid = O(x^{3-4\beta_0})$ which is ultimately decreasing. For the case where $\frac{2}{3} < \beta_0 < 1$, form (2.3.59) we have $I(\infty) < \infty$ and $\mid g^{''}(x)g(x) \mid = O(x^{-2(1-\beta_0)})$ which is ultimately decreasing. Thus the Assumption A.4 of Theorem 3 of [20] holds.
- From (2.3.59), we have $g^{'}(x)I^{1/2}(x) = O(x^{-\beta_0/2}) \to 0$ as $x \to \infty$. With this we see that all the assumptions of Theorem 3 of [20] hold.

Now, for $0 < \beta_0 \leq 2/3$, from (2.3.60) we can see that, $\hat{I}(\infty) = \infty$. Thus, part (iv) of the theorem follows from Theorem 3(b) of [20], where $\hat{I}(n)^{-1/2} \approx K^{-1}_{\beta_0,n}$, with the $K_{\beta_0,n}$ as in (2.3.45).

When $2/3 < \beta_0 < 1$, (2.3.60) implies that $\hat{I}(\infty) < \infty$. Thus, part (v) of the theorem follows from Theorem 3(a) of [20]. And this completes the proof of the theorem. ■

2.4. Examples

In this section we present some examples for Model 2.1 presented and investigated in this chapter. With the help of these examples, we try to highlight some points about the role played by the different parameters involved.

Example 2.1. Consider Model 2.1. In the notation of Section 2.1, assume that $\min(\mu_1^*, \mu_2^*) > 1$. If F_n and M_n denote the numbers of females and males in the nth generation, then a female in the nth generation either marries or not according to the following probability:

$$P(\eta_{n,1,1} = 1 \mid \mathcal{G}_{n-1}) = \begin{cases} 1, & \text{if } 0 < F_n \leq M_n, \\ \frac{M_n}{F_n}, & \text{if } F_n > M_n > 0, \\ 0, & \text{otherwise.} \end{cases} \tag{2.4.1}$$

One can think of this model according to the following. When there are more males than females in a generation, all the females get married. When there are more females than males, the males are not faithful to their partners (possibility of cheating or having a unsuccessful marriage). It is interesting to see that the expected number of married females in the nth generation is then equal to $E[C_n \mid \mathcal{G}_{n-1}] = \min(F_n, M_n)$. Now, by the assumption $m = \min(\mu_1^*, \mu_2^*) > 1$, we are in the supercritical case. Theorem 2.3 provides almost sure convergence and in $\mathcal{L}_2$ convergence of $W_n = \frac{C_n}{m^n}$ to a non-degenerate random variable W.

□

The next two examples are for Model 2.1 in the critical case.

Example 2.2. For fixed $a > 0$, let the reproduction law for a couple in the zeroth generation, given the total number of couples in that generation, be:

$$P(S_{1,j} = u \mid C_0 = z) = \begin{cases} \frac{1}{4}(1 + \frac{1}{8}z^{-a}), & \text{if } \ 0 < z < 4^{1/a} \ \text{ and,} \\ & u \in \{(1,2),(2,1),(2,2)\}, \\ \\ \frac{1}{4}(1 - \frac{3}{8}z^{-a}), & \text{if } \ 0 < z < 4^{1/a} \\ & \text{and } \ u = (0,0), \\ \\ \frac{1}{4}(1 + \frac{1}{2}z^{-a}), & \text{if } \ z \geq 4^{1/a} \ \text{ and,} \\ & u \in \{(1,2),(2,1),(2,2)\}, \\ \\ \frac{1}{4}(1 - \frac{3}{2}z^{-a}), & \text{if } \ z \geq 4^{1/a} \\ & \text{and } \ u = (0,0), \\ \\ 0, & \text{otherwise,} \end{cases} \tag{2.4.2}$$

Now, in the nth generation, given $C_{n-1} = z$ and $\overline{S}_n = (x, y)$, we assume that the number of marriages of a female is a Bernolli random variable with probability of success equal to

$$\varphi_z(x,y) = (\frac{y}{x+y})^{d_0 + d_1 z^{-h_1}} \exp\{-d_2(x+y)z^{-h_2}\}\mathbf{1}_{\{\min(z,x,y)>0\}}, \tag{2.4.3}$$

where the parameters $d_i \in \mathcal{R}$, for $i = 0,1,2$ and $h_1, h_2 > 0$ should be such that the range of the function φ_z is between zero and one. Since we are interested in the critical case, we take $d_0 = \frac{\log 5}{\log 2} - 2$, $d_1 > d_0$ and $d_2 > 0$. For $z \geq 4^{1/a}$, simple calculations give the values of the vectors $\mu(z) = (\mu_1(z), \mu_2(z))$, $\mu^* = (\mu_1^*, \mu_2^*)$ and matrices $\Sigma(z) = [\sigma_{ij}(z)]_{2\times 2}$ and $\Sigma^* = [\sigma^*_{ij}]_{2\times 2}$ (see (2.1.6) for definitions),

$$\begin{aligned} &\mu_1(z) = \mu_2(z) = \tfrac{5}{4}(1 + \tfrac{1}{2}z^{-a}), \quad \mu_1^* = \mu_2^* = \tfrac{5}{4}, \\ &\sigma_{11}(z) = \sigma_{22}(z) = \tfrac{11}{16} - \tfrac{7}{16}z^{-a} + o(z^{-a}), \\ &\sigma_{12}(z) = \sigma_{21}(z) = \tfrac{7}{16} - \tfrac{9}{16}z^{-a} + o(z^{-a}), \\ &\sigma^*_{11} = \sigma^*_{22} = \tfrac{11}{16}, \quad \sigma^*_{12} = \sigma^*_{21} = \tfrac{7}{16}. \end{aligned} \tag{2.4.4}$$

Also the functions φ^*, ν_z and ν^* given (2.1.7) and (2.1.8) will be equal to

$$\begin{aligned} &\varphi^*(x,y) = (\tfrac{y}{x+y})^{d_0}\mathbf{1}_{\{\min(x,y)>0\}}, \\ &\nu_z(x,y) = \varphi_z(x,y)(1 - \varphi_z(x,y))\mathbf{1}_{\{\min(z,x,y)>0\}}, \\ &\nu^*(x,y) = \varphi^*(x,y)(1 - \varphi^*(x,y))\mathbf{1}_{\{\min(x,y)>0\}}, \end{aligned} \tag{2.4.5}$$

where $x, y \in \mathcal{R}_+$ and $z \in \mathcal{Z}_+$.

It is easy to see that Assumption **A.1** holds. For Assumption **A.5**, notice that with $d_0 = \frac{\log 5}{\log 2} - 2$ and μ^* as in (2.4.4), we have $\mu_1^* \varphi^*(\mu^*) = \frac{5}{4}(\frac{1}{2})^{d_0} = 1$. Also one can easily see that $\| \mu(z) - \mu^* \| = \frac{\sqrt{2}}{2} z^{-a}$, where by our assumption $a > 0$. By simple calculation one can show that the parameters a_i^* and β_i, for $i = 1, 2$, in (2.3.2) are equal to

$$\begin{aligned} &\beta_1 = \min\{h_1, h_2\}, \qquad \beta_2 = a, \\ &a_1^* = -d_1 \log(2) \mathbf{1}_{\{h_1 = \beta_1\}} - \tfrac{5}{2} d_2 \mathbf{1}_{\{h_2 = \beta_1\}} \text{ and } a_2^* = \tfrac{1}{2}. \end{aligned} \tag{2.4.6}$$

Further the parameters a_3^* and a_4^* in Lemma 2.13 (given in (2.3.19) and (2.3.21)) are equal to

$$\begin{aligned} a_3^* &= \tfrac{1}{2} \textstyle\sum_{i,j=1}^2 \sigma_{ij}^* (D^2 f(\mu^*))_{ij} \cong 0.036 \\ a_4^* &= \textstyle\sum_{ij=1}^2 \sigma_{ij}^* (Df(\mu^*))_i (Df(\mu^*))_j \cong 0.577, \end{aligned} \tag{2.4.7}$$

where $f(x) = x_1 \varphi^*(x)$ for $x \in \mathcal{R}_+^2$.

Now by using (2.4.6)-(2.4.7), we can calculate the parameters β_0, b^* and c^* of Lemma 2.14. We have

$$\begin{aligned} \beta_0 &= \max\{0, 1 - \beta_1, 1 - \beta_2\} \\ &= \max\{0, 1 - h_1, 1 - h_2, 1 - a\}, \\ b^* &= \textstyle\sum_{j=1}^2 a_j^* \mathbf{1}_{\{\beta_0 = 1 - \beta_j\}} + a_3^* \mathbf{1}_{\{\beta_0 = 0\}} \\ &= -d_1 \log(2) \cdot \mathbf{1}_{\{\beta_0 = 1 - h_1\}} - \tfrac{5}{2} d_2 \cdot \mathbf{1}_{\{\beta_0 = 1 - h_2\}} \\ &+ \tfrac{1}{2} \cdot \mathbf{1}_{\{\beta_0 = 1 - a\}} + 0.036 \cdot \mathbf{1}_{\{\beta_0 = 0\}}, \\ c^* &= a_4^* + \mu_1^* \nu^*(\mu^*) \\ &= 0.577 + 0.2 = 0.777. \end{aligned} \tag{2.4.8}$$

Hence, by using Theorem 2.15 we can describe the limiting behavior of the process $\{C_n\}$ for different values of the parameters b^* and c^*.

$\square$

Example 2.3. In this example we consider again Model 2.1 in the critical case. Our aim is to show how the variation in elements of the limiting covariance matrix Σ^* (see (2.3.1)(i) for definition) affect the limiting behaviour of the process $\{C_n\}$.

For simplicity, assume that the number of marriages of each female is a Bernolli random variable, like in Example 2.2. Also let $\mu(z)$, μ^*, φ_z, φ^*, ν_z and ν^* (defined in (2.1.6), (2.1.7) and (2.1.8), respectively) be equal to

$$\begin{aligned}
&\mu_1(z) = \mu_2(z) = \tfrac{5}{4}(1 + \tfrac{1}{2}z^{-1}),\\
&\mu_1 = \mu_2 = \tfrac{5}{4},\\
&\varphi_z(x,y) = (\tfrac{y}{x+y})^{d_0}\mathbf{1}_{\{\min(x,y,z)>0\}},\\
&\varphi^*(x,y) = (\tfrac{y}{x+y})^{d_0}\mathbf{1}_{\{\min(x,y)>0\}},\\
&\nu_z(x,y) = \varphi_z(x,y)(1-\varphi_z(x,y)),\\
&\nu^*(x,y) = \varphi^*(x,y)(1-\varphi^*(x,y)),
\end{aligned} \tag{2.4.9}$$

where in the above we take $d_0 = \frac{\log(5)}{\log(2)} - 2$. We can easily see that with the function $f(x) = x_1\varphi^*(x)$, we have

$$Df(\mu^*) \cong \begin{bmatrix} 0.671 \\ 0.129 \end{bmatrix}, \quad D^2f(\mu^*) \cong \begin{bmatrix} -0.036 & 0.138 \\ 0.138 & -0.036 \end{bmatrix}. \tag{2.4.10}$$

Now with matrix $\Sigma^* = [\sigma^*_{ij}]_{2\times 2}$ (see (5)(i) for the definition), in a similar way as in Example 2.2, we can show that the parameters β_0, b^* and c^*, which determine the limiting behaviour of the process $\{C_n\}$ (see Theorem 2.15), are equal to

$$\begin{aligned}
\beta_0 &= 0,\\
b^* &= \tfrac{1}{2} + a^*_3 = \tfrac{1}{2} + \tfrac{1}{2}\textstyle\sum_{i,j=1}^2 \sigma^*_{ij}(D^2f(\mu^*))_{ij},\\
c^* &= a^*_4 + \mu^*_1\nu^*(\mu^*)\\
&= \textstyle\sum_{i,j=1}^2 \sigma^*_{ij}(Df(\mu^*))_i(Df(\mu^*))_j + 0.2.
\end{aligned} \tag{2.4.11}$$

Since $\beta_0 = 0$, for the process $\{C_n\}$ grows to infinity with positive probability if and only if $2b^* > c^*$. Replacing values of b^* and c^* from (2.4.11) gives us

$$2b^* > c^* \iff 0.49\sigma^*_{11} - 0.10\sigma^*_{12} + 0.05\sigma^*_{22} < 0.8. \tag{2.4.12}$$

From (2.4.12) one can see that how variation in the elements of the limiting covariance matrix Σ^* could effect the asymptotic behaviour of the process $\{C_n\}$.

□

Chapter 3

Two-Sex Galton-Watson Branching Processes: Conditional Independent Mating for Families

In this chapter, we shall study Model 2.2 in the supercritical - as well as critical case. A careful look at the steps we needed to take to prove Theorems 2.3 and 2.15, shows that under appropriate assumptions one can derive similar results for Model 2.2. In Sections 3.1 and 3.2 we see what kind of assumptions are needed to obtain these results. In doing so, we show that similar to Lemmas 2.8, 2.9, 2.13 and 2.14 hold for Model 2.2. These lemmas path the way for the proofs of Theorems 3.5 and 3.8 which are given in Sections 3.1 and 3.2, respectively.

We start by defining some notations for this model. Let $\overline{S}_n^{(i)}$ be defined as follows,

$$\overline{S}_n^{(i)} := \begin{cases} \frac{1}{C_{n-1}-1} \sum_{j \neq i}^{C_{n-1}} S_{n,j}, & \text{if } 1 \leq i \leq C_{n-1} \\ & \text{and } C_{n-1} > 1, \\ & \\ (0,0), & \text{otherwise}, \end{cases} \tag{3.0.1}$$

for $n, i = 1, 2, \ldots$. Notice that, given $\mathcal{F}_{n-1}$, $S_{n,i}$ and $\overline{S}_n^{(i)}$ are independent and further that knowing C_{n-1}, $S_{n,i}$ and $\overline{S}_n$ is equivalent to knowing C_{n-1}, $S_{n,i}$ and $\overline{S}_n^{(i)}$.

Denote the conditional expectation and the centered conditional absolute moments of the number of marriages of daughters of the jth family in the $(n+1)$th generation by

$$\begin{aligned} \psi_{C_n}(S_{n+1,j}, \overline{S}_{n+1}^{(j)}) &:= E[U_{n+1,j} \mid \mathcal{G}_n], \\ \tilde{\nu}_{k,C_n}(S_{n+1,j}, \overline{S}_{n+1}^{(j)}) &:= E[\mid U_{n+1,j} - E[U_{n+1,j} \mid \mathcal{G}_n] \mid^k \mid \mathcal{G}_n], \end{aligned} \tag{3.0.2}$$

for $j = 1, 2, \ldots, C_n$ and $k = 1, 2, \ldots$.

A. 6. *Let $\{\mu(z)\}_{z=0}^{\infty}$ be as in (2.1.6). Assume that there exist a point $\mu^* \in \mathcal{R}_+^2$ and real-valued functions ψ^* and $\tilde{\nu}_2^*$ on $\mathcal{R}_+^2 \times \mathcal{R}_+^2$ such that*

$$\begin{aligned} &(i) \quad \mu(z) \to \mu^*, \\ &(ii) \quad \psi_z \longrightarrow \psi^*, \\ &(iii) \quad \tilde{\nu}_{2,z} \to \tilde{\nu}_2^*, \end{aligned} \tag{3.0.3}$$

as $z \to \infty$.

Then we can define the following.

Definition 3.1. *We say that Model 2.2 is subcritical, critical or supercritical if*

$$\tilde{m} := \lim_{z\to\infty} E[\psi^*(S_{1,1}, \mu^*) \mid C_0 = z] <, \ =, \ \text{or} > 1, \tag{3.0.4}$$

respectively (we assume the limit in (3.0.4) exists).

Remark 3.1. With Model 2.1 as a special case of Model 2.2, let us derive the functions ψ_z, ψ^*, $\tilde{\nu}_{2,z}$ and $\tilde{\nu}_2$ for Model 2.1 in terms of the functions φ_z, φ^*, ν_z and ν^* given in (2.1.7)-(2.1.9).

Since in Model 2.1 $U_{1,1} = \sum_{l=1}^{C_0} \eta_{1,1,l}$, and given $\mathcal{G}_0$, $\eta_{1,1,1}, \ldots, \eta_{1,1,C_0}$ are i.i.d. random variables with the mean $\varphi_{C_0}(\overline{S}_1) = E[\eta_{1,1,1} \mid \mathcal{G}_0]$ and the variance $\nu_{C_0}(\overline{S}_1) = \mathrm{Var}[\eta_{1,1,1} \mid \mathcal{G}_0]$. Hence for $z > 0$ and $x, y \in \mathcal{R}_+^2$ we can write ψ_z and $\tilde{\nu}_{2,z}$ as

$$\begin{aligned} \psi_z(x,y) &= x_1 \varphi_z(\tfrac{1}{z}x + \tfrac{z-1}{z}y), \\ \tilde{\nu}_{2,z}(x,y) &= x_1 \nu_z(\tfrac{1}{z}x + \tfrac{z-1}{z}y). \end{aligned} \tag{3.0.5}$$

If we assume that the sequence of functions $\{\varphi_z\}$ and $\{\nu_z\}$ are smooth enough, then ψ^* and $\tilde{\nu}_2$ can be written as

$$\begin{aligned} \psi^*(x,y) &= \lim_{z\to\infty} \psi_z(x,y) \\ &= \lim_{z\to\infty} x_1 \varphi_z(\tfrac{1}{z}x + \tfrac{z-1}{z}y) = x_1 \varphi^*(y), \end{aligned} \tag{3.0.6}$$

and similarly, we have $\tilde{\nu}_2(x,y) = x_1\nu^*(y)$, for $x, y \in \mathcal{R}_+^2$.

Further, the parameter $\tilde{m}$ for Model 2.1 is equal to the parameter m given in (2.1.9), i.e.

$$\begin{aligned}\tilde{m} &= \lim_{z\to\infty} E[\psi^*(S_{1,1},\mu^*) \mid C_0 = z] \\ &= \lim_{z\to\infty} E[X_{1,1}\varphi^*(\mu^*) \mid C_0 = z] \qquad (3.0.7) \\ &= \lim_{z\to\infty} \mu_1(z)\varphi^*(\mu^*) = \mu_1^*\varphi^*(\mu^*) = m.\end{aligned}$$

□

3.1. Model 2.2: Supercritical Case

We start by studying Model 2.2 in the supercritical case, i.e. $\tilde{m} > 1$. Some assumptions are then needed. The first is similar to Assumption **A.2** for Model 2.1.

Similar to Chapter 2, for convenience we refer to $S_{1,j} = (X_{1,j}, Y_{1,j})$, $\overline{S}_1^{(j)} = (\overline{X}_1^{(j)}, \overline{Y}_1^{(j)})$ and $U_{1,j}$ as $S_j = (X_j, Y_j)$, $\overline{S}^{(j)} = (\overline{X}^{(j)}, \overline{Y}^{(j)})$ and U_j, respectively, for $j = 1, 2, \ldots, C_0$.

A. 7. *With $\mu(\cdot)$ and μ^* as in Assumption* **A.6** *and $\tilde{m}$ as in (3.0.4), assume that for some positive $\tilde{\delta}_0$, $\tilde{\delta}_1$ and $\tilde{\delta}_2$,*

$$| E[\psi^*(S_1,\mu^*) \mid C_0 = z] - \tilde{m} | = O(z^{-\tilde{\delta}_0}), \tag{3.1.1}$$

$$\tilde{m} > 1, \text{ and } \| \mu(z) - \mu^* \| = O(z^{-\tilde{\delta}_1}), \tag{3.1.2}$$

$$| E[\psi_z(S_1,\mu^*) - \psi^*(S_1,\mu^*) \mid C_0 = z] | = O(z^{-\tilde{\delta}_2}). \tag{3.1.3}$$

Remark 3.2. One way to interpret (3.1.3) is the following. By a Law of Large Numbers argument one expects that $\overline{S}^{(1)}$ goes to μ^* as $C_0 \to \infty$ almost surely. Assume for a moment that $\overline{S}^{(1)} \equiv \mu^*$. Then (3.1.3) tells us about the rate of convergence of the expected number of marriages of the first family when in the second stage we have $C_0 = z$ couples to the case where we have infinity number of couples. However in both cases we still assume that the reproduction law of S_1 depends on $C_0 = z$. □

The next assumption is similar to Assumption **A.3**.

A. 8. *With $k_0 \geq 3$, assume that $E[X_1^{2k_0} + Y_1^{2k_0} \mid C_0 = z]$ is uniformly bounded on $z \in \mathcal{Z}_+$ and the same for $E[\tilde{\nu}_{k,z}(S_1, \overline{S}^{(1)}) \mid C_0 = z]$ for all $k = 1, 2, \ldots, 2k_0$. Further assume that for some constant $K > 0$ we have*

$$\begin{aligned} \psi_z(x,y) &\leq K \parallel x \parallel (1+ \parallel y \parallel)^{1/2}, \\ \tilde{\nu}_{2,z}(x,y) &\leq K \parallel x \parallel (1+ \parallel y \parallel)^{1/2}, \end{aligned} \tag{3.1.4}$$

for all $x, y \in \mathcal{R}_+^2$ and $z \in \mathcal{Z}_+$.

The third assumption concerns the behaviours of the functions ψ_z and $\tilde{\nu}_{2,z}$ in a neighbourhood of μ^*. It is similar to Assumption **A.4** for Model 2.1.

A. 9. *Consider the sequences of functions $\{\psi_z\}_z$ and $\{\tilde{\nu}_{2,z}\}_z$ given in (3.0.2) with $k_0 \geq 3$ (as in Assumption* **A.8***). Assume that there exist constants $K > 0$, $\epsilon > 0$ and $z_\epsilon > 0$ such that if $B = B_\epsilon(\mu^*)$, then we have*

$$| \psi_z(x,y) - \psi_z(x, \mu(z)) | \leq K \parallel x \parallel \parallel y - \mu(z) \parallel,$$

and (3.1.5)

$$| \tilde{\nu}_{2,z}(x,y) - \tilde{\nu}_{2,z}(x, \mu(z)) | \leq K \parallel x \parallel \parallel y - \mu(z) \parallel,$$

for all $x \in \mathcal{R}_+^2$, $y \in B$ and $z \geq z_\epsilon$.

Let $E = \mathcal{R}_+^2$ and for given $z \in \mathcal{Z}_+$, define the function $\tilde{f}_z$ on E^z by

$$\tilde{f}_z(x) = \begin{cases} \frac{1}{z} \sum_{i=1}^{z} (\psi_z(x_i, \frac{1}{z-1} \sum_{j=1, j \neq i}^{z} x_j) & \\ -E[\psi_z(S_1, \mu(z)) \mid C_0 = z]), & \text{if } z \geq 2, \text{ and} \\ & x = (x_1, \ldots, x_z) \in E^z, \\ 0, & \text{otherwise.} \end{cases} \tag{3.1.6}$$

This function is similar to the function f_z defined in (2.2.22) and in the following, Lemmas 3.2, 3.3 and 3.4 are the counterparts of Lemmas 2.7, 2.8 and 2.9, respectively.

Lemma 3.2. *Let Assumptions* **A.6**-**A.9** *hold. Then, with the function $\tilde{f}_z$ as in (3.1.6), we have*

$$\begin{aligned} (i) \quad & E[\tilde{f}_z^{2k}(S_1, \ldots, S_z) \mid C_0 = z] = O(z^{-k}), \\ (ii) \quad & E[\tilde{f}_z^{k}(S_1, \ldots, S_z) \mid C_0 = z] = O(z^{-k/2}), \end{aligned} \tag{3.1.7}$$

for $1 \leq k \leq k_0$ and $z \geq 2$, where k_0 is as in Assumption **A.8***.*

Proof of Lemma 3.2: We can write the left hand side of (3.1.7)(i) as

$$E[\tilde{f}_z^{2k}((S_1,\ldots,S_z)) \mid C_0 = z]$$

$$= E[(\tfrac{1}{z}\textstyle\sum_{i=1}^{z}(\psi_z(S_i,\overline{S}^{(i)}) - E[\psi_z(S_1,\mu(z)) \mid C_0 = z]))^{2k} \mid C_0 = z]$$

$$\leq 2^{2k-1}E[(\tfrac{1}{z}\textstyle\sum_{i=1}^{z}(\psi_z(S_i,\overline{S}^{(i)}) - \psi_z(S_i,\mu(z))))^{2k} \mid C_0 = z]$$

$$+2^{2k-1}E[(\tfrac{1}{z}\textstyle\sum_{i=1}^{z}(\psi_z(S_i,\mu(z)) - E[\psi_z(S_1,\mu(z)) \mid C_0 = z]))^{2k} \mid C_0 = z]. \tag{3.1.8}$$

Now we consider each of the last two terms in the right hand side of (3.1.8) separately. By using Lemma 2.4 we can write the first term as

$$E[(\tfrac{1}{z}\textstyle\sum_{i=1}^{z}(\psi_z(S_i,\overline{S}^{(i)}) - \psi_z(S_i,\mu(z))))^{2k} \mid C_0 = z]$$

$$\leq \tfrac{1}{z^{2k}}z^{2k-1}\textstyle\sum_{i=1}^{z}E[(\psi_z(S_i,\overline{S}^{(i)}) - \psi_z(S_i,\mu(z)))^{2k} \mid C_0 = z] \tag{3.1.9}$$

$$= E[(\psi_z(S_1,\overline{S}^{(1)}) - \psi_z(S_1,\mu(z)))^{2k} \mid C_0 = z].$$

Now if B is the closed disk as in Assumption **A.9**, we can write (3.1.9) as

$$E[(\tfrac{1}{z}\textstyle\sum_{i=1}^{z}(\psi_z(S_i,\overline{S}^{(i)}) - \psi_z(S_i,\mu(z))))^{2k} \mid C_0 = z]$$

$$\leq E[(\psi_z(S_1,\overline{S}^{(1)}) - \psi_z(S_1,\mu(z)))^{2k};\overline{S}^{(1)} \in B \mid C_0 = z] \tag{3.1.10}$$

$$+E[(\psi_z(S_1,\overline{S}^{(1)}) - \psi_z(S_1,\mu(z)))^{2k};\overline{S}^{(1)} \in B^c \mid C_0 = z].$$

By Assumption **A.9** and the conditional independence of S_1 and $\overline{S}^{(1)}$ we can write $E[(\psi_z(S_1,\overline{S}^{(1)}) - \psi_z(S_1,\mu(z)))^{2k};\overline{S}^{(1)} \in B \mid C_0 = z]$ in (3.1.10) as

$$E[(\psi_z(S_1,\overline{S}^{(1)}) - \psi_z(S_1,\mu(z)))^{2k};\overline{S}^{(1)} \in B \mid C_0 = z]$$

$$\leq KE[\| S_1 \|^{2k}\| \overline{S}^{(1)} - \mu(z) \|^{2k};\overline{S}^{(1)} \in B \mid C_0 = z]$$

$$\leq KE[\| S_1 \|^{2k}\mid C_0 = z]E[\| \overline{S}^{(1)} - \mu(z) \|^{2k}\mid C_0 = z] \tag{3.1.11}$$

$$\leq KE[\| S_1 \|^{2k}\mid C_0 = z]E[\| \overline{S}^{(1)} - \mu(z) \|^{2k}\mid C_0 = z]$$

Now, by Assumption **A.8** $E[\| S_1 \|^{2k}\mid C_0 = z]$ is uniformly bounded for $k \leq k_0$ and by Lemma 2.6 $E[\| \overline{S}^{(1)} - \mu(z) \|^{2k}\mid C_0 = z] = O(z^{-k})$. Hence we can write (3.1.11) as

$$E[(\psi_z(S_1,\overline{S}^{(1)})-\psi_z(S_1,\mu(z)))^{2k};\overline{S}^{(1)}\in B\mid C_0=z]=O(z^{-k}). \quad (3.1.12)$$

Further, by Lemma 2.4, Assumption **A.8**, the conditional independence of S_1 and $\overline{S}^{(1)}$ and the Hölder inequality, we can write the term $E[(\psi_z(S_1,\overline{S}^{(1)})-\psi_z(S_1,\mu(z)))^{2k};\overline{S}^{(1)}\in B^c\mid C_0=z]$ in (3.1.10) as

$$\begin{aligned}
&E[(\psi_z(S_1,\overline{S}^{(1)})-\psi_z(S_1,\mu(z)))^{2k};\overline{S}^{(1)}\in B^c\mid C_0=z]\\
&\leq 2^{2k-1}E[(\psi_z^{2k}(S_1,\overline{S}^{(1)})+\psi_z^{2k}(S_1,\mu(z)));\overline{S}^{(1)}\in B^c\mid C_0=z]\\
&\leq 2^{2k-1}K^{2k}E[\| S_1\|^{2k}(1+\|\overline{S}^{(1)}\|)^k;\overline{S}^{(1)}\in B^c\mid C_0=z]\\
&+2^{2k-1}K^{2k}E[\| S_1\|^{2k}(1+\|\mu(z)\|)^k;\overline{S}^{(1)}\in B^c\mid C_0=z]\\
&\leq 2^{2k-1}K^{2k}E[\| S_1\|^{2k}\mid C_0=z]\{E[(1+\|\overline{S}^{(1)}\|)^{2k}\mid C_0=z]^{1/2}\\
&P(\overline{S}^{(1)}\in B^c\mid C_0=z)^{1/2}+(1+\|\mu(z)\|)^kP(\overline{S}^{(1)}\in B^c\mid C_0=z)\}.
\end{aligned} \quad (3.1.13)$$

Again, by Assumption **A.8** $E[\| S_1\|^{2k}\mid C_0=z]$ and $E[(1+\|\overline{S}^{(1)}\|)^{2k}\mid C_0=z]$ are uniformly bounded for $k\leq k_0$ and by Lemma 2.6 $P(\overline{S}^{(1)}\in B^c\mid C_0=z)=O(z^{-k_0})$. Hence, we can write (3.1.13) as

$$E[(\psi_z(S_1,\overline{S}^{(1)})-\psi_z(S_1,\mu(z)))^{2k};\overline{S}^{(1)}\in B^c\mid C_0=z]=O(z^{-k_0}). \quad (3.1.14)$$

By replacing (3.1.12) and (3.1.14) in (3.1.10) we have

$$E[(\frac{1}{z}\sum_{i=1}^{z}(\psi_z(S_i,\overline{S}^{(i)})-\psi_z(S_i,\mu(z))))^{2k}\mid C_0=z]=O(z^{-k}). \quad (3.1.15)$$

Now consider the second term in the left hand side of (3.1.8). First notice that by Assumption **A.8** for some constant K, we can write the $2k$th absolute moment of $\psi_z(S_1,\mu(z))-E[\psi_z(S_1,\mu(z))\mid C_0=z]$ as

$$\begin{aligned}
&E[\mid\psi_z(S_1,\mu(z))-E[\psi_z(S_1,\mu(z))\mid C_0=z]\mid^{2k}\mid C_0=z]\\
&\leq 2^{2k-1}E[\psi_z^{2k}(S_1,\mu(z))+E[\psi_z(S_1,\mu(z))\mid C_0=z]^{2k}\mid C_0=z]\\
&\leq 2^{2k}E[\psi_z^{2k}(S_1,\mu(z))\mid C_0=z]\\
&\leq 2^{2k}E[\| S_1\|^{2k}(1+\|\mu(z)\|)^k\mid C_0=z]<K.
\end{aligned} \quad (3.1.16)$$

Hence by Lemma 2.5 we have

$$E[(\tfrac{1}{z}\textstyle\sum_{i=1}^{z}(\psi_z(S_i,\mu(z)) - E[\psi_z(S_1,\mu(z)) \mid C_0 = z]))^{2k} \mid C_0 = z]$$

$$= \tfrac{1}{z^{2k}} E[(\textstyle\sum_{i=1}^{z}(\psi_z(S_i,\mu(z)) - E[\psi_z(S_1,\mu(z)) \mid C_0 = z]))^{2k} \mid C_0 = z]$$

$$= O(z^{-k}). \tag{3.1.17}$$

By replacing (3.1.15) and (3.1.17) in (3.1.8) we get

$$\mid E[\tilde{f}_z^{2k}(S_1,\ldots,S_z) \mid C_0 = z] \mid = O(z^{-k}). \tag{3.1.18}$$

To show (3.1.7)(ii) we use (3.1.7)(i) and the Hölder inequality to the left hand side of (3.1.7)(ii) to get

$$\begin{aligned} &\mid E[\tilde{f}_z^{k}(S_1,\ldots,S_z) \mid C_0 = z] \mid \\ &\leq E[\tilde{f}_z^{2k}(S_1,\ldots,S_z) \mid C_0 = z]^{1/2} \\ &= O(z^{-k/2}). \end{aligned} \tag{3.1.19}$$

■

Lemma 3.3. *Assume* **A.6**-**A.9** *hold. Then with* $\tilde{m}$ *as in (3.0.4) we have*

$$\begin{aligned} &(i) \quad \tilde{m} = \lim_{z\to\infty} \tfrac{1}{z} E[C_1 \mid C_0 = z], \\ &(ii) \quad E[\mid C_1 - E[C_1 \mid C_0 = z] \mid^k \mid C_0 = z] = O(z^{k/2}), \end{aligned} \tag{3.1.20}$$

for k_0 *as in Assumption* **A.8**, *and* $k \leq k_0$.

Proof of Lemma 3.3: To show (3.1.20)(i), with the function $\tilde{f}_z$ as in (3.1.6) we have

$$\tfrac{1}{z}E[C_1 \mid C_0 = z] = \tfrac{1}{z}\textstyle\sum_{j=1}^{z} E[U_j \mid C_0 = z]$$

$$= \tfrac{1}{z}\textstyle\sum_{j=1}^{z} E[\psi_z(S_1, \overline{S}^{(1)}) \mid C_0 = z]$$

$$= \tfrac{1}{z}\textstyle\sum_{j=1}^{z} E[\{\psi_z(S_1, \overline{S}^{(1)}) - E[\psi_z(S_1,\mu(z)) \mid C_0 = z]\} \mid C_0 = z]$$

$$+E[\psi_z(S_1,\mu(z)) \mid C_0 = z]$$

$$= E[\tilde{f}_z(S_1,\ldots,S_z) \mid C_0 = z] + E[\psi_z(S_1,\mu(z)) \mid C_0 = z]. \tag{3.1.21}$$

Now, by Lemma 3.2 we have $E[\tilde{f}_z(S_1,\ldots,S_z) \mid C_0 = z] = O(z^{-1/2})$; by Assumption **A.7** for some $\tilde{\delta}_1, \tilde{\delta}_2 > 0$ we have $\| \mu(z) - \mu^* \| = O(z^{\tilde{\delta}_1})$ and $| E[\psi_z(S_1,\mu^*) - \psi^*(S_1,\mu^*) \mid C_0 = z] | = O(z^{-\tilde{\delta}_2})$; and by Assumption **A.9** for z large and some constant $K > 0$ we have $| \psi_z(S_1,\mu(z)) - \psi_z(S_1,\mu^*) | \leq K \| S_1 \| \| \mu(z) - \mu^* \|$. Hence by these and (3.1.21) together with adding and subtracting $E[\psi_z(S_1,\mu(z)) \mid C_0 = z]$ and $E[\psi_z(S_1,\mu^*) \mid C_0 = z]$ to $\frac{1}{z}E[C_1 \mid C_0 = z] - E[\psi^*(S_1,\mu^*) \mid C_0 = z]$ we have

$$\begin{aligned}
& | \tfrac{1}{z}E[C_1 \mid C_0 = z] - E[\psi^*(S_1,\mu^*) \mid C_0 = z] | \\
& \leq | E[\tilde{f}_z(S_1,\ldots,S_z) \mid C_0 = z] | \\
& + | E[\psi_z(S_1,\mu(z)) \mid C_0 = z] - E[\psi_z(S_1,\mu^*) \mid C_0 = z] | \qquad (3.1.22) \\
& + | E[\psi_z(S_1,\mu^*) \mid C_0 = z] - E[\psi^*(S_1,\mu^*) \mid C_0 = z] | \\
& \leq O(z^{-1/2}) + K \| \mu(z) - \mu^* \| E[\| S_1 \| \mid C_0 = z] + O(z^{-\tilde{\delta}_2}).
\end{aligned}$$

The right hand side of (3.1.22) goes to zero as $z \to \infty$, hence $\lim_{z\to\infty} \frac{1}{z}E[C_1 \mid C_0 = z] = \lim_{z\to\infty} E[\psi^*(S_1,\mu^*) \mid C_0 = z] = \tilde{m}$.

The proof of the assertion (ii) of the lemma would be similar to that of the assertion (ii) of Lemma 2.8. We can decompose $C_1 - E[C_1 \mid C_0]$ into the two terms as following.

$$C_1 - E[C_1 \mid C_0] = \sum_{i=1}^{2} \tilde{\xi}_i^{(z)}, \qquad (3.1.23)$$

where we have

$$\begin{aligned}
\tilde{\xi}_1^{(z)} &= \textstyle\sum_{j=1}^{z}(U_j - E[U_j \mid \mathcal{G}_0]), \\
\tilde{\xi}_2^{(z)} &= \textstyle\sum_{j=1}^{z}(E[U_j \mid \mathcal{G}_0] - E[\psi_z(S_1,\overline{S}^{(1)}) \mid C_0 = z]), \qquad (3.1.24) \\
&= \textstyle\sum_{j=1}^{z}(\psi_z(S_j,\overline{S}^{(j)}) - E[\psi_z(S_1,\overline{S}^{(1)}) \mid C_0 = z]).
\end{aligned}$$

We need to show that $E[| \tilde{\xi}_i^{(z)} |^k | C_0 = z] = O(z^{k/2})$ for $i = 1,2$ and $1 \leq k \leq k_0$. Then, since for some constant $K > 0$, we have

$$E[C_1 - E[C_1 \mid C_0] \mid C_0 = z] \leq K \sum_{i=1}^{2} E[| \tilde{\xi}_i^{(z)} |^k | C_0 = z], \qquad (3.1.25)$$

then we have proved the lemma.

For $k = 1$, with $\tilde{\nu}_{2,z}$ as in (3.0.2) and by using Marcinkiewicz-Zygmund's inequality, concavity of the square root function and Assumption **A.8**, we have

$$\begin{aligned}
&E[\mid \tilde{\xi}_1^{(z)} \mid\mid C_0 = z] = E[\mid \textstyle\sum_{j=1}^{z}(U_j - E[U_j \mid \mathcal{G}_0]) \mid\mid C_0 = z] \\
&\leq KE[(\textstyle\sum_{j=1}^{z}(U_j - E[U_j \mid \mathcal{G}_0])^2)^{1/2} \mid C_0 = z] \\
&\leq K(\textstyle\sum_{j=1}^{z} E[(U_j - E[U_j \mid \mathcal{G}_0])^2 \mid C_0])^{1/2} \\
&= Kz^{1/2}E[\tilde{\nu}_{2,z}(S_1, \overline{S}^{(1)}) \mid C_0 = z]^{1/2} = O(z^{1/2}).
\end{aligned} \tag{3.1.26}$$

For $2 \leq k \leq k_0$, again using Marcinkiewicz-Zygmund's inequality and the Hölder inequality (as in (2.2.42)) we can show that

$$\begin{aligned}
&E[\mid \tilde{\xi}_1^{(z)} \mid^k\mid C_0 = z] = E[\mid \textstyle\sum_{j=1}^{z}(U_j - E[U_j \mid \mathcal{G}_0]) \mid^k\mid C_0 = z] \\
&\leq Kz^{\frac{k}{2}-1}\textstyle\sum_{j=1}^{z} E[\mid U_j - E[U_j \mid \mathcal{G}_0] \mid^k\mid C_0 = z] \\
&= Kz^{k/2}E[\tilde{\nu}_{k,z}(S_1, \overline{S}^{(1)}) \mid C_0 = z] = O(z^{k/2}).
\end{aligned} \tag{3.1.27}$$

Thus we have shown $E[\mid \tilde{\xi}_1^{(z)} \mid^k\mid C_0 = z] = O(z^{k/2})$ for $1 \leq k \leq k_0$.

By Lemma 3.2, we can show the similar assertion for $\tilde{\xi}_2^{(z)}$. We have

$$\begin{aligned}
&E[\mid \tilde{\xi}_2^{(z)} \mid^k\mid C_0 = z] = z^kE[\mid \tilde{f}_z(S_1, \ldots, S_z) \mid^k\mid C_0 = z] \\
&\leq z^kE[\tilde{f}_z^{2k}(S_1, \ldots, S_z) \mid C_0 = z]^{1/2} = O(z^{k/2}),
\end{aligned} \tag{3.1.28}$$

for $1 \leq k \leq k_0$. ∎

Consider the decomposition which was given for C_{n+1} in (2.2.43)-(2.2.44). Then we have the following lemma.

Lemma 3.4. *Assume* **A.6**-**A.9**. *Consider the following decomposition for the process* $\{C_n\}_{n=0}^{\infty}$,

$$C_{n+1} = C_n + g(C_n) + \xi_{n+1}, \tag{3.1.29}$$

with the function g and the residuals ξ_{n+1} as in (2.2.44), for $n \geq 0$. Then

$$\begin{aligned}
&(i) \quad E[\xi_{n+1} \mid \mathcal{F}_n] = 0, \\
&(ii) \quad E[\xi_{n+1}^2 \mid \mathcal{F}_n] = \sigma^2(C_n), \ \text{for some function} \ \sigma \geq 0, \\
&(iii) \quad \text{for some constant} \ K > 0, \ \text{as} \ z \to \infty, \qquad (3.1.30) \\
&\qquad \text{we have} \ Kz \leq \sigma^2(z) = O(z), \ \text{and} \ g(z) = O(z) \\
&(iv) \quad E[\mid \xi_1 \mid^{k_0} \mid C_0 = z] = O(\sigma^{k_0}(z)), \ \text{for all} \ z \geq 1,
\end{aligned}$$

with k_0 as in the assumption **A.9**.

Proof of Lemma 3.4: We only show that $g(z) = O(z)$. The proofs of the rest of the assertions are similar to that of Lemma 2.9.

With the functions ψ_z, ψ^* and $\tilde{f}_z$ as in (3.0.2), (3.0.3) and (3.1.6), respectively, we can decompose $g(z) = E[C_1 \mid C_0 = z] - z$ as

$$\begin{aligned}
g(z) \quad &= z(E[U_1 \mid C_0 = z] - 1) \\
&= zE[\tilde{f}_z((S_1, \ldots, S_z)) \mid C_0 = z] \\
&+ zE[\psi_z(S_1, \mu(z)) - \psi_z(S_1, \mu^*) \mid C_0 = z] \qquad (3.1.31) \\
&+ zE[\psi_z(S_1, \mu^*) - \psi^*(S_1, \mu^*) \mid C_0 = z] \\
&+ z(E[\psi^*(S_1, \mu^*) \mid C_0 = z] - 1).
\end{aligned}$$

Take z to be large enough. Then by Lemma 3.2

$$E[\tilde{f}_z((S_1, \ldots, S_z)) \mid C_0 = z] = O(z^{-1/2}). \qquad (3.1.32)$$

By Assumptions **A.7**-**9**, for $\tilde{\delta}_1 > 0$ we have

$$\begin{aligned}
&\mid E[\psi_z(S_1, \mu(z)) - \psi_z(S_1, \mu^*) \mid C_0 = z] \mid \\
&\leq E[\mid \psi_z(S_1, \mu(z)) - \psi_z(S_1, \mu^*) \mid\mid C_0 = z] \qquad (3.1.33) \\
&\leq K \parallel \mu(z) - \mu^* \parallel E[\parallel S_1 \parallel\mid C_0 = z] = O(z^{-\tilde{\delta}_1}),
\end{aligned}$$

for some $K > 0$. By Assumption **A.7**,

$$\mid E[\psi_z(S_1, \mu^*) - \psi^*(S_1, \mu^*) \mid C_0 = z] \mid = O(z^{-\tilde{\delta}_2}), \qquad (3.1.34)$$

for $\tilde{\delta}_2 > 0$. And finally by Assumption **A.7**, $\tilde{m} > 1$ and $E[\psi^*(S_1, \mu^*) \mid C_0 = z] = \tilde{m} + o(1)$. Therefore we have $g(z) = O(z)$. ■

The following theorem is the counterpart of Theorem 2.3 for Model 2.2.

Theorem 3.5. (Exponential Growth) *Consider Model 2.2 in the super-critical case, and let* **A.6**-**A.8** *hold. Then,*

(i) $P(C_n \to 0 \text{ or } \infty) = 1$ *and* $P(C_n \to \infty) > 0$,

(ii) $E[C_n] \sim K\tilde{m}^n$, *as* $n \to \infty$ *for some constant* $K > 0$,

(iii) there exists a non-degenerate r.v. W *such that* $W_n := \frac{C_n}{\tilde{m}^n} \to W$ *almost surely and in* $\mathcal{L}_1$.

(iv) and if further the $\tilde{\delta}_0$, $\tilde{\delta}_1$ *and* $\tilde{\delta}_2$ *in* **A.7** *satisfy* $(\tilde{\delta}_0 \wedge \tilde{\delta}_1 \wedge \tilde{\delta}_2) \geq 1/2$, *then* W_n *converges to* W *almost surely and in* $\mathcal{L}_2$.

Proof of Theorem 3.5: The proof of Theorem 2.3 can be carried out here as well. To see this, first notice that Assumptions **A.6**, **A.7**, **A.8**, **A.9** are counter part to Assumptions **A.1**, **A.2**, **A.3** and **A.4** for Model 2.1, respectively. Furthermore, Lemmas 3.2 and 3.4 correspond to Lemmas 2.7 and 2.9, respectively.

Replace the use of Lemma 2.9 by Lemma 3.4 in the proof of the assertion (i) of Theorem 2.3 to prove the assertion (i) of the theorem.

For the assertion (ii), first define the function $\tilde{\rho}$ (similar to ρ defined in (2.2.60)) as

$$\begin{aligned} \tilde{\rho}(x) \quad &:= \mid E[U_1 \mid C_0 = 1] - \tilde{m} \mid \mathbf{1}_{\{0 \leq x < 1\}} \\ &+ \mid E[U_1 \mid C_0 = [x]] - \tilde{m} \mid \mathbf{1}_{\{x \geq 1\}}, \end{aligned} \tag{3.1.35}$$

for $x \geq 0$, where $\tilde{m}$ is as in (3.0.4). Similar to (2.2.61)-(2.2.62), for $z > 0$ we can write $\tilde{\rho}(z)$ as

$$\begin{aligned} \tilde{\rho}(z) \quad &= \mid E[U_1 \mid C_0 = z] - \tilde{m} \mid \\ &= \mid E[\psi_z(S_1, S^{(1)}) \mid C_0 = z] - \tilde{m} \mid . \end{aligned} \tag{3.1.36}$$

Now with the function ψ^* as in (3.0.3), we can write (3.1.36) as

$$\begin{aligned} \tilde{\rho}(z) \quad &\leq \mid E[\psi_z(S_1, S^{(1)}) - \psi_z(S_1, \mu(z)) \mid C_0 = z] \mid \\ &+ \mid E[\psi_z(S_1, \mu(z)) - \psi_z(S_1, \mu^*) \mid C_0 = z] \mid \\ &+ \mid E[\psi_z(S_1, \mu^*) - \psi^*(S_1, \mu^*) \mid C_0 = z] \mid \\ &+ \mid E[\psi^*(S_1, \mu^*) \mid C_0 = z] - \tilde{m} \mid \end{aligned} \tag{3.1.37}$$

Now, with the function $\tilde{f}_z$ as in (3.1.6) and Lemma 3.2 we have

$$
\begin{aligned}
&\mid E[\psi_z(S_1, S^{(1)}) - \psi_z(S_1, \mu(z)) \mid C_0 = z] \mid \\
&= \mid E[\tilde{f}_z(S_1, \ldots, S_z) \mid C_z = z] \mid \\
&= O(z^{-1/2}).
\end{aligned} \tag{3.1.38}
$$

Further, for $z > 0$ large enough by Assumption **A.9** we have $\mid \psi_z(S_1, \mu(z)) - \psi_z(S_1, \mu^*) \mid \leq K \parallel S_1 \parallel\parallel \mu(z) - \mu^* \parallel$. By Assumption **A.8** $E[\parallel S_1 \parallel \mid C_0 = z]$ is uniformly bounded and by Assumption **A.7** $\parallel \mu(z) - \mu^* \parallel = O(z^{-\tilde{\delta}_1})$. Hence

$$
\mid E[\psi_z(S_1, \mu(z)) - \psi_z(S_1, \mu^*) \mid C_0 = z] \mid = O(z^{-\tilde{\delta}_1}). \tag{3.1.39}
$$

By Assumption **A.7** we have

$$
\begin{aligned}
&\mid E[\psi_z(S_1, \mu^*) - \psi^*(S_1, \mu^*) \mid C_0 = z] \mid = O(z^{-\tilde{\delta}_2}), \\
&\mid E[\psi^*(S_1, \mu^*) \mid C_0 = z] - \tilde{m} \mid = O(z^{-\tilde{\delta}_0}).
\end{aligned} \tag{3.1.40}
$$

Replacing (3.1.38)-(3.1.40) in (3.1.37) we get

$$
\tilde{\rho}(z) = O(z^{-(\tilde{\delta}_0 \wedge \tilde{\delta}_1 \wedge \tilde{\delta}_2 \wedge \frac{1}{2})}). \tag{3.1.41}
$$

And hence $\sum_z \tilde{\rho}(z)/z < \infty$. Take $W_n = C_n/\tilde{m}^n$ and repeat the proof of Theorem 2.3(ii) to show the assertion (ii) of the theorem.

Proofs of the assertions (iii)-(iv) of the theorem are similar to those of Theorem 2.3(iii)-(iv).

■

3.2. Model 2.2: Critical Case

After presenting some assumptions, we prove lemmas similar to Lemmas 2.13 and 2.14. Then under appropriate assumption we can see that the same results given in Theorem 2.15 can be obtained for Model 2.2 in the critical case. In the following, whenever it is convenient for a point $x \in \mathcal{R}_+^2$, we shall denote its ith coordinate as x_i or $(x)_i$, for $i = 1, 2$.

A. 10. *With $k_0 \geq 3$, assume that $E[X_1^{2k_0} + Y_1^{2k_0} \mid C_0 = z]$ is uniformly bounded on $z \in \mathcal{Z}_+$ and the same for $E[\tilde{\nu}_{k,z}(S_1, \overline{S}^{(1)}) \mid C_0 = z]$ for all*

$k = 1, 2, \ldots, 2k_0$. *Further assume that there exists constants* $K > 0$, $\epsilon > 0$ *and* $z_\epsilon > 0$ *such that if* $B = B_\epsilon(\mu^*)$, *then for* $z > z_\epsilon$ *we have*

$$
\begin{aligned}
(i) \quad & \psi_z(x,y) \le K \parallel x \parallel (1+ \parallel y \parallel)^{1/2}, \\
& \mid \psi_z(x,y) - \psi_z(x,\mu(z)) \mid \le K \parallel x \parallel \parallel y - \mu(z) \parallel, \\
& \mid \tfrac{\partial^j \psi_z(x,u)}{\partial u_{i_1} \ldots \partial u_{i_j}} \mid_{u=y} \mid \le K \parallel x \parallel, \\
(ii) \quad & \tilde{\nu}_{2,z}(x,y) \le K \parallel x \parallel (1+ \parallel y \parallel)^{1/2}, \\
& \mid \tilde{\nu}_{2,z}(x,y) - \tilde{\nu}_{2,z}(x,\mu(z)) \mid \le K \parallel x \parallel \parallel y - \mu(z) \parallel, \\
& \mid \tfrac{\partial^j \tilde{\nu}_{2,z}(x,u)}{\partial u_{i_1} \ldots \partial u_{i_j}} \mid_{u=y} \mid \le K \parallel x \parallel,
\end{aligned}
\tag{3.2.1}
$$

for all $x \in \mathcal{R}_+^2$, $y \in B$ *and* $z \ge z_\epsilon$, $j = 1, 2$, *and 3, and* $i_l \in \{1, 2\}$ *(for* $l = 1, 2, 3$*). And assume there exist* $\Delta_{i_1}^{(1)}, \Delta_{i_1}^{(2)}$ *and* $\Delta_{i_1,i_2}^{(1)}$ *such that as* $z \to \infty$ *we have*

$$
\begin{aligned}
(iii) \quad & E[\tfrac{\partial \psi_z(S_1,y)}{\partial y_{i_1}} \mid_{y=\mu(z)} \mid C_0 = z] \to \Delta_{i_1}^{(1)}, \quad \text{and} \\
& E[\tfrac{\partial^2 \psi_z(S_1,y)}{\partial y_{i_1} \partial y_{i_2}} \mid_{y=\mu(z)} \mid C_0 = z] \to \Delta_{i_1,i_2}^{(1)}, \\
(iv) \quad & \mathrm{Cov}[(S_1)_{i_1}, \psi_z(S_1,\mu(z)) \mid C_0 = z] \to \Delta_{i_1}^{(2)},
\end{aligned}
\tag{3.2.2}
$$

and the limits are finite, where $i_l \in \{1, 2\}$, *for* $l = 1, 2$.

A. 11. *With* $\Sigma(\cdot)$ *be as in (2.1.6) assume there is a* 2×2 *matrix* Σ^* *such that*

$$
(i) \quad \Sigma(z) \to \Sigma^*, \quad \text{as } z \to \infty. \tag{3.2.3}
$$

Further with $\mu(\cdot)$, μ^*, $\tilde{m}$, ψ_z, ψ^* *and* $\tilde{\nu}_{2,z}$ *as in the previous section, assume that for some constants* $\tilde{\beta}_j^*$, $\tilde{a}_i^*$ *(for* $i = 1, 2, 3, 5, 6$, *and* $j = 1, 2, 3$*) and*

$\tilde{\delta}_1 > 0$ *as* $z \to \infty$, *we have*

$$
\begin{aligned}
&(ii) && \tilde{m} = 1, \ \text{ and } \ \| \mu(z) - \mu^* \| \leq O(z^{-\tilde{\delta}_1}), \\
&(iii) && E[\psi_z(S_1, \mu(z)) - \psi^*(S_1, \mu(z)) \mid C_0 = z] \\
& && = \tilde{a}_1^* z^{-\tilde{\beta}_1} + o(z^{-\tilde{\beta}_1}), \\
&(iv) && E[\psi^*(S_1, \mu(z)) - \psi^*(S_1, \mu^*) \mid C_0 = z] \\
& && = \tilde{a}_2^* z^{-\tilde{\beta}_2} + o(z^{-\tilde{\beta}_2}), \\
&(v) && E[\psi^*(S_1, \mu^*) \mid C_0 = z] = 1 + \tilde{a}_3^* z^{-\tilde{\beta}_3} + o(z^{-\tilde{\beta}_3}), \\
&(vi) && \{\tilde{\nu}_{2,z}\}_{z=1}^{\infty} \ \text{ is uniformly bounded,} \\
&(vii) && E[\tilde{\nu}_{2,z}(S_1, \mu(z)) - \tilde{\nu}_2^*(S_1, \mu^*) \mid C_0 = z] = o(1), \\
&(viii) && E[\tilde{\nu}_2^*(S_1, \mu^*) \mid C_0 = z] = \tilde{a}_5^* + o(1), \\
&(ix) && E[\psi_z^2(S_1, \mu(z)) \mid C_0 = z] = 1 + \tilde{a}_6^* + o(1),
\end{aligned}
\tag{3.2.4}
$$

where we take $\tilde{\beta}_i = \infty$ *(or any value greater than one) whenever* $\tilde{a}_i^* = 0$, *for* $i = 1, 2, 3$.

Lemmas 3.6 and 3.7 below are the counterpart to Lemmas 2.13 and 2.14.

Lemma 3.6. *Assume* **A.6**, **A.10** *and* **A.11** *hold. With the functions* ψ_z

and $\tilde{\nu}_{2,z}$ *as in (3.0.2), we have*

$$
\begin{aligned}
(i) \quad & E[\psi_z(S_1, \overline{S}^{(1)}) \mid C_0 = z] \\
& = E[\psi_z(S_1, \mu(z)) \mid C_0 = z] + \tilde{a}_4^* z^{-1} + o(z^{-1}), \\
(ii) \quad & E[\psi_z^2(S_1, \overline{S}^{(1)}) \mid C_0 = z] \\
& = E[\psi_z^2(S_1, \mu(z)) \mid C_0 = z] + O(z^{-1}), \\
(iii) \quad & \mathrm{Cov}[\psi_z(S_1, \overline{S}^{(1)}), \psi_z(S_2, \overline{S}^{(2)}) \mid C_0 = z] \\
& = \tilde{a}_7^* z^{-1} + o(z^{-1}), \\
(iv) \quad & E[\tilde{\nu}_{2,z}(S_1, \overline{S}^{(1)}) \mid C_0 = z] \\
& = E[\tilde{\nu}_{2,z}(S_1, \mu(z)) \mid C_0 = z] + O(z^{-1}),
\end{aligned}
\tag{3.2.5}
$$

for $z > 0$.

Proof of Lemma 3.6: We use Taylor's expansion to show the assertions of the lemma.

To show part (i) of the lemma, we first write down a Taylor expansion for the function ψ_z, for $z \in \mathcal{Z}_+$ large enough. Let B be a sphere centred at μ^* with radius $\epsilon > 0$ (as in Assumption **A.10**). Choose z_ϵ large enough such that $\mu(z)$ belong to the interior of the disk B for $z \geq z_\epsilon$. Further, take $\epsilon_0 > 0$ small enough such that the closed disk B contains $\cup_{z \geq z_\epsilon} B_z$, where $B_z = B_{\epsilon_0}(\mu(z))$. Then for $x, y \in \mathcal{R}_+^2$ (with x fixed), we can write $\psi_z(x, y)$ as

$$
\begin{aligned}
& \psi_z(x,y) = \{\psi_z(x, \mu(z)) + \textstyle\sum_{i=1}^2 (\frac{\partial \psi_z(x,u)}{\partial u_i} \mid_{u=\mu(z)} (y - \mu(z))_i \\
& + \tfrac{1}{2} \textstyle\sum_{i_1, i_2 = 1}^2 (\frac{\partial^2 \psi_z(x,u)}{\partial u_{i_1} \partial u_{i_2}} \mid_{u=\mu(z)} (y - \mu(z))_{i_1} (y - \mu(z))_{i_2} \\
& + R_z(y;x)\} \mathbf{1}_{B_z}(y) + \psi_z(x,y) \mathbf{1}_{B_z^c}(y),
\end{aligned}
\tag{3.2.6}
$$

where the remainder term $R_z(y;x)$ is equal to

$$
\begin{aligned}
R_z(y;x) \quad & = \tfrac{1}{6} \textstyle\sum_{i_1, i_2, i_3 = 1}^2 \{ \frac{\partial^3 \psi_z(x,u)}{\partial u_{i_1} \partial u_{i_2} \partial u_{i_3}} \mid_{u = l_z(y;x)} \\
& \cdot \textstyle\prod_{j=1}^3 (y - \mu(z))_{i_j} \}.
\end{aligned}
\tag{3.2.7}
$$

In (3.2.7), for given z and x, $l_z(y;x)$ is a point on the line segment joining $\mu(z)$ and y. We can rewrite (3.2.6) in the following form.

$$\begin{aligned}\psi_z(x,y) &= \psi_z(x,\mu(z)) + \sum_{i=1}^{2} \frac{\partial\psi_z(x,u)}{\partial u_i}\mid_{u=\mu(z)} (y-\mu(z))_i \\ &+\frac{1}{2}\sum_{i_1,i_2=1}^{2} \frac{\partial^2\psi_z(x,u)}{\partial u_{i_1}\partial u_{i_2}}\mid_{u=\mu(z)} \prod_{j=1}^{2}(y-\mu(z))_{i_j} \\ &+\hat{R}_z(y;x),\end{aligned} \tag{3.2.8}$$

where $\hat{R}_z(y;x)$ is equal to

$$\begin{aligned}\hat{R}_z(y;x) &= R_z(y;x)\mathbf{1}_{B_z}(y) + \{-\psi_z(x,\mu(z)) \\ &+\sum_{i=1}^{2} \frac{\partial\psi_z(x,u)}{\partial u_i}\mid_{u=\mu(z)} (y-\mu(z))_i \\ &-\frac{1}{2}\sum_{i_1,i_2=1}^{2} \frac{\partial^2\psi_z(x,u)}{\partial u_{i_1}\partial u_{i_2}}\mid_{u=\mu(z)} \prod_{j=1}^{2}(y-\mu(z))_{i_j} \\ &+\psi_z(x,y)\}\mathbf{1}_{B_z^c}(y).\end{aligned} \tag{3.2.9}$$

By **A.10** and (3.2.9) we can see that for some constant $K>0$ we have

$$\begin{aligned}\mid \hat{R}_z(y;x)\mid &\leq\mid R_z(y;x)\mid \mathbf{1}_{B_z}(y) + \{\mid \psi_z(x,\mu(z))\mid \\ &+\mid \sum_{i=1}^{2} \frac{\partial\psi_z(x,u)}{\partial u_i}\mid_{u=\mu(z)} (y-\mu(z))_i \mid \\ &+\mid \sum_{i_1,i_2=1}^{2} \frac{\partial^2\psi_z(x,u)}{\partial u_{i_1}\partial u_{i_2}}\mid_{u=\mu(z)} \prod_{j=1}^{2}(y-\mu(z))_{i_j} \mid \\ &+\psi_z(x,y)\}\mathbf{1}_{B_z^c}(y) \\ &\leq K \parallel x\parallel \{\parallel y-\mu(z)\parallel^3 + \{1+\parallel y-\mu(z)\parallel \\ &+\parallel y-\mu(z)\parallel^2 + (1+\parallel y\parallel)^{1/2}\}\mathbf{1}_{B_z^c}(y)\}.\end{aligned} \tag{3.2.10}$$

Then, for the assertion (i) of the lemma, by taking the conditional expectation of (3.2.8) and by the independence of S_1 and $\overline{S}^{(1)}$ we have

$$E[\psi_z(S_1,\overline{S}^{(1)}) \mid C_0 = z] = E[\psi_z(S_1,\mu(z)) \mid C_0 = z]$$
$$+\sum_{i=1}^2 E[\frac{\partial\tilde{\psi}_z(S_1,u)}{\partial u_i} \mid_{u=\mu(z)} (\overline{S}^{(1)} - \mu(z))_i \mid C_0 = z]$$
$$+\frac{1}{2}\sum_{i_1,i_2=1}^2 E[\frac{\partial^2\tilde{\psi}_z(S_1,u)}{\partial u_{i_1}\partial u_{i_2}} \mid_{u=\mu(z)} \prod_{j=1}^2(\overline{S}^{(1)} - \mu(z))_{i_j} \mid C_0 = z]$$
$$+E[\hat{R}_z(\overline{S}^{(1)};S_1) \mid C_0] \tag{3.2.11}$$
$$= E[\psi_z(S_1,\mu(z)) \mid C_0 = z]$$
$$+\frac{1}{2(z-1)}\sum_{i_1,i_2=1}^2 E[\frac{\partial^2\tilde{\psi}_z(S_1,u)}{\partial u_{i_1}\partial u_{i_2}} \mid_{u=\mu(z)}\mid C_0 = z]\sigma_{i_1,i_2}(z)$$
$$+E[\hat{R}_z(\overline{S}^{(1)};S_1) \mid C_0].$$

Now by (3.2.10), the independence of S_1 of $\overline{S}^{(1)}$ and the Hölder inequality, we have

$$\mid E[\hat{R}_z(\overline{S}^{(1)};S_1) \mid C_0 = z] \mid$$
$$\leq KE[\| S_1 \|\mid C_0 = z]E[\| \overline{S}^{(1)} - \mu(z) \|^3\mid C_0 = z]$$
$$+KE[\| S_1 \|\mid C_0 = z]\{P(\overline{S}^{(1)} \in B^c \mid C_0 = z) \tag{3.2.12}$$
$$+E[\| \overline{S}^{(1)} - \mu(z) \|^2\mid C_0 = z]^{1/2}P(\overline{S}^{(1)} \in B^c \mid C_0 = z)^{1/2}$$
$$+E[(1+ \| \overline{S}^{(1)} \|) \mid C_0 = z]^{1/2}P(\overline{S}^{(1)} \in B^c \mid C_0 = z)^{1/2}\}.$$

By Assumption **A.10** $E[\| S_1 \|\mid C_0 = z]$ and $E[(1+ \| \overline{S}^{(1)} \|) \mid C_0 = z]$ are uniformly bounded. Lemma 2.6 gives $E[\| \overline{S}^{(1)} - \mu(z) \|^j\mid C_0 = z] = O(z^{-j/2})$ for $j = 1,2,3$ and $P(\overline{S}^{(1)} \in B^c \mid C_0 = z) = O(z^{-k_0})$. By replacing these in (3.2.12) we get $E[\hat{R}_z(\overline{S}^{(1)};S_1) \mid C_0] = o(z^{-1})$. On the other hand by **A.11** we have

$$E[\frac{\partial^2\tilde{\psi}_z(S_1,u)}{\partial u_{i_1}\partial u_{i_2}} \mid_{u=\mu(z)}\mid C_0 = z] \to \Delta^{(1)}_{i_1,i_2},$$
$$\text{and } \sigma_{i_1,i_2}(z) \to \sigma^*_{i_1,i_2}, \tag{3.2.13}$$

as $z \to \infty$. These and (3.2.11) give the assertion (i) of the lemma with $\tilde{a}_4^*$ equal to

$$\tilde{a}_4^* = \frac{1}{2} \sum_{i_1,i_2=1}^{2} \Delta^{(1)}_{i_1,i_2} \sigma^*_{i_1,i_2}. \tag{3.2.14}$$

In the similar way the assertions (ii) and (iv) can be proven.

What remains is assertion (iii). Again we use the Taylor expansion of ψ_z given in (3.2.11), to decompose $\psi_z(S_i, \overline{S}^{(i)}) - E[\psi_z(S_i, \mu(z)) \mid C_0 = z] = \sum_{j=1}^{4} \xi_{i,j}$ where $\xi_{i,j}$ are equal to

$$\begin{aligned}
\xi_{i,1} &:= \psi_z(S_i, \mu(z)) - E[\psi_z(S_i, \mu(z)) \mid C_0 = z], \\
\xi_{i,2} &:= \textstyle\sum_{j=1}^{2} \frac{\partial \psi_z(S_i,u)}{\partial u_j} \mid_{u=\mu(z)} (\overline{S}^{(i)} - \mu(z))_j, \\
\xi_{i,3} &:= \textstyle\frac{1}{2} \sum_{j_1,j_2} \frac{\partial^2 \psi_z(S_i,u)}{\partial u_{j_1} \partial u_{j_2}} \mid_{u=\mu(z)} \prod_{l=1}^{2} (\overline{S}^{(i)} - \mu(z))_{j_l}, \\
\xi_{i,4} &:= \hat{R}_z(\overline{S}^{(i)}; S_i),
\end{aligned} \tag{3.2.15}$$

for $j = 1, 2$. Also for $i \neq j$, lets define $\overline{S}_n^{(i,j)}$ as following

$$\overline{S}_n^{(i,j)} = \begin{cases} \frac{1}{z-2} \sum_{l \in \{1,\ldots,C_{n-1}\} \setminus \{i,j\}} S_{n,l}, & \text{if } C_{n-1} > 2,\ i \neq j \\ & \text{and } 1 \leq i,j \leq C_{n-1}, \\ (0,0), & \text{otherwise,} \end{cases} \tag{3.2.16}$$

for $i, j = 1, 2, \ldots$ and $n = 1, 2, \ldots$. In the following, we refer to $\overline{S}_1^{(i,j)} = (\overline{X}_1^{(i,j)}, \overline{Y}_1^{(i,j)})$ as $\overline{S}^{(i,j)} = (\overline{X}^{(i,j)}, \overline{Y}^{(i,j)})$, for $i, j = 1, 2, \ldots, C_0$.

First notice by the assertion (i) of the lemma we have $E[\psi_z(S_1, \overline{S}^{(1)}) \mid C_0 = z] = E[\psi_z(S_1, \mu(z)) \mid C_0 = z] = O(z^{-1})$. Hence we can write the left hand side of (3.2.5)(iii) as

$$\begin{aligned}
&\operatorname{Cov}[\psi_z(S_1, \overline{S}^{(1)}), \psi_z(S_2, \overline{S}^{(2)}) \mid C_0 = z] \\
&= E[\textstyle\prod_{i=1}^{2} (\psi_z(S_i, \overline{S}^{(i)}) - E[\psi_z(S_i, \mu(z)) \mid C_0 = z]) \mid C_0 = z] + O(z^{-2}) \\
&= E[\textstyle\prod_{i=1}^{2} (\sum_{j=1}^{4} \xi_{i,j}) \mid C_0 = z] + O(z^{-2}) \\
&= \textstyle\sum_{j_1,j_2=1}^{4} E[\xi_{1,j_1} \xi_{2,j_2} \mid C_0 = z] + O(z^{-2}) \\
&= \textstyle\sum_{j=1}^{4} E[\xi_{1,j} \xi_{2,j} \mid C_0 = z] + 2 \sum_{i<j} E[\xi_{1,i} \xi_{2,j} \mid C_0 = z] + O(z^{-2}).
\end{aligned} \tag{3.2.17}$$

By the conditional independence of S_1, S_2 we have

$$E[\xi_{1,1}\xi_{2,1} \mid C_0 = z] = E[\xi_{1,1} \mid C_0 = z]E[\xi_{2,1} \mid C_0 = z] = 0. \qquad (3.2.18)$$

Using (3.2.16) and the conditional independence of S_1, S_2 and $\overline{S}^{(1,2)}$, together with the assumption **A.11** we can write $E[\xi_{1,1}\xi_{2,2} \mid C_0 = z]$ as

$$\begin{aligned}
&E[\xi_{1,1}\xi_{2,2} \mid C_0 = z] \\
&= \textstyle\sum_{i=1}^{2}\{\mathrm{Cov}[\psi_z(S_1,\mu(z)),(\overline{S}^{(2)})_i \mid C_0 = z] \\
&\cdot E[\tfrac{\partial\psi_z(S_2,u)}{\partial u_i} \mid_{u=\mu(z)} \mid C_0 = z]\} \\
&= \textstyle\sum_{i=1}^{2}\{\mathrm{Cov}[\psi_z(S_1,\mu(z)),\{\tfrac{1}{z-1}(S_1)_i + \tfrac{z-2}{z-1}(\overline{S}^{(1,2)})_i\} \mid C_0 = z] \\
&\cdot E[\tfrac{\partial\psi_z(S_2,u)}{\partial u_i} \mid_{u=\mu(z)} \mid C_0 = z]\} \\
&= \tfrac{1}{z-1}\textstyle\sum_{i=1}^{2}\{\mathrm{Cov}[\psi_z(S_1,\mu(z)),(S_1)_i \mid C_0 = z] \\
&\cdot E[\tfrac{\partial\psi_z(S_2,u)}{\partial u_i} \mid_{u=\mu(z)} \mid C_0 = z]\} \\
&= z^{-1}\textstyle\sum_{i=1}^{2}\Delta_i^{(1)}\Delta_i^{(2)} + o(z^{-1}).
\end{aligned} \qquad (3.2.19)$$

Also the term $E[\xi_{1,2}\xi_{2,2} \mid C_0 = z]$ is equal to

$$\begin{aligned}
&E[\xi_{1,2}\xi_{2,2} \mid C_0 = z] = E[\{\{\textstyle\sum_{j_1=1}^{2}\tfrac{\partial\psi_z(S_1,u)}{\partial u_{j_1}} \mid_{u=\mu(z)} (\overline{S}^{(1)} - \mu(z))_{j_1}\} \\
&\cdot\{\textstyle\sum_{j_2=1}^{2}\tfrac{\partial\psi_z(S_2,u)}{\partial u_{j_2}} \mid_{u=\mu(z)} (\overline{S}^{(2)} - \mu(z))_{j_2}\}\} \mid C_0 = z] \\
&= \textstyle\sum_{j_1,j_2=1}^{2} E[\{\tfrac{\partial\psi_z(S_1,u)}{\partial u_{j_1}} \mid_{u=\mu(z)} \tfrac{\partial\psi_z(S_2,u)}{\partial u_{j_2}} \mid_{u=\mu(z)} \\
&\cdot\{\tfrac{1}{z-1}(S_2 - \mu(z))_{j_1} + \tfrac{z-2}{z-1}(\overline{S}^{(1,2)} - \mu(z))_{j_1}\} \\
&\cdot\{\tfrac{1}{z-1}(S_1 - \mu(z))_{j_2} + \tfrac{z-2}{z-1}(\overline{S}^{(1,2)} - \mu(z))_{j_2}\}\} \mid C_0 = z]
\end{aligned} \qquad (3.2.20)$$

And (3.2.20) gives

$$
\begin{aligned}
& E[\xi_{1,2}\xi_{2,2} \mid C_0 = z] \\
& = \tfrac{1}{(z-1)^2} \textstyle\sum_{j_1,j_2=1}^{2} \{E[\tfrac{\partial \psi_z(S_1,u)}{\partial u_{j_1}} \mid_{u=\mu(z)} (S_1 - \mu(z))_{j_2} \mid C_0 = z] \\
& \cdot E[\tfrac{\partial \psi_z(S_2,u)}{\partial u_{j_2}} \mid_{u=\mu(z)} (S_2 - \mu(z))_{j_1} \mid C_0 = z]\} \\
& +2\tfrac{z-2}{(z-1)^2} \textstyle\sum_{j_1,j_2=1}^{2} \{E[\tfrac{\partial \psi_z(S_1,u)}{\partial u_{j_1}} \mid_{u=\mu(z)} \mid C_0 = z] \\
& \cdot E[\tfrac{\partial \psi_z(S_2,u)}{\partial u_{j_2}} \mid_{u=\mu(z)} (S_2 - \mu(z))_{j_1} \mid C_0 = z] \\
& \cdot E[(\overline{S}^{(1,2)} - \mu(z))_{j_2} \mid C_0 = z]\} \\
& +(\tfrac{z-2}{z-1})^2 \textstyle\sum_{j_1,j_2=1}^{2} \{E[\tfrac{\partial \psi_z(S_1,u)}{\partial u_{j_1}} \mid_{u=\mu(z)} \mid C_0 = z] \\
& \cdot E[\tfrac{\partial \psi_z(S_2,u)}{\partial u_{j_2}} \mid_{u=\mu(z)} \mid C_0 = z] E[\textstyle\prod_{l=1}^{2} (\overline{S}^{(1,2)} - \mu(z))_{j_l} \mid C_0 = z]\}.
\end{aligned}
\tag{3.2.21}
$$

Now, we consider each of the three terms in the right hand side of (3.2.21) separately. By Assumption **A.10**(i), the absolute value of the first term in (3.2.21) can be written as

$$
\begin{aligned}
& \mid \tfrac{1}{(z-1)^2} \textstyle\sum_{j_1,j_2=1}^{2} \{E[\tfrac{\partial \psi_z(S_1,u)}{\partial u_{j_1}} \mid_{u=\mu(z)} (S_1 - \mu(z))_{j_2} \mid C_0 = z] \\
& \cdot E[\tfrac{\partial \psi_z(S_2,u)}{\partial u_{j_2}} \mid_{u=\mu(z)} (S_2 - \mu(z))_{j_1} \mid C_0 = z]\} \mid \\
& \leq \tfrac{1}{(z-1)^2} \textstyle\sum_{j_1,j_2=1}^{2} \{E[\mid \tfrac{\partial \psi_z(S_1,u)}{\partial u_{j_1}} \mid_{u=\mu(z)} \mid\mid (S_1 - \mu(z))_{j_2} \mid\mid C_0 = z] \\
& \cdot E[\mid \tfrac{\partial \psi_z(S_2,u)}{\partial u_{j_2}} \mid_{u=\mu(z)} \mid\mid (S_2 - \mu(z))_{j_1} \mid\mid C_0 = z]\} \\
& \leq \tfrac{4K^2}{(z-1)^2} E[\| S_1 \|\| S_1 - \mu(z) \|| C_0 = z]^2 = O(z^{-2}).
\end{aligned}
\tag{3.2.22}
$$

Since $E[(\overline{S}^{(1,2)} - \mu(z))_{j_2} \mid C_0 = z] = 0$ the second term in (3.2.21) is equal to zero. By Assumption **A.10**(iii) for large z, $E[\frac{\partial \psi_z(S_1,u)}{\partial u_j} \mid_{u=\mu(z)} \mid C_0] = \Delta_j^{(1)} + o(1)$ and also we have $E[\prod_{l=1}^{2} (\overline{S}^{(1,2)} - \mu(z))_{j_l} \mid C_0 = z] = \frac{1}{z-2}\sigma_{j_1,j_2}(z) = \frac{1}{z-2}\sigma^*_{j_1,j_2} + o(1)$. Hence the last term in (3.2.21) can be written as

$$(\tfrac{z-2}{z-1})^2 \sum_{j_1,j_2=1}^{2} \{E[\tfrac{\partial \psi_z(S_1,u)}{\partial u_{j_1}} \mid_{u=\mu(z)} \mid C_0 = z]$$

$$\cdot E[\tfrac{\partial \psi_z(S_2,u)}{\partial u_{j_2}} \mid_{u=\mu(z)} \mid C_0 = z] \cdot E[\textstyle\prod_{l=1}^{2} (\overline{S}^{(1,2)} - \mu(z))_{j_l} \mid C_0 = z]\}$$

$$= z^{-1} \sum_{j_1,j_2=1}^{2} \Delta_{j_1}^{(1)} \Delta_{j_2}^{(1)} \sigma^*_{j_1,j_2} + o(z^{-1}). \tag{3.2.23}$$

By replacing the values of these three terms back into (3.2.21) we obtain

$$E[\xi_{1,2}\xi_{2,2} \mid C_0 = z] = z^{-1} \sum_{j_1,j_2=1}^{2} \Delta_{j_1}^{(1)} \Delta_{j_2}^{(1)} \sigma^*_{j_1,j_2} + o(z^{-1}). \tag{3.2.24}$$

Similarly we can write $E[\xi_{1,1}\xi_{2,3} \mid C_0 = z]$ as

$$E[\xi_{1,1}\xi_{2,3} \mid C_0 = z]$$

$$= \tfrac{1}{2} \sum_{i_1,i_2=1}^{2} \{\mathrm{Cov}[\psi_z(S_1, \mu(z)), \textstyle\prod_{j=1}^{2} (\overline{S}^{(2)} - \mu(z))_{i_j} \mid C_0 = z]$$

$$\cdot E[\tfrac{\partial^2 \psi_z(S_2,u)}{\partial u_{i_1} \partial u_{i_2}} \mid_{u=\mu(z)} \mid C_0 = z]\}$$

$$= \tfrac{1}{2} \sum_{i_1,i_2=1}^{2} \{\mathrm{Cov}[\psi_z(S_1, \mu(z)), \{\textstyle\prod_{j=1}^{2} (\tfrac{1}{z-1}(S_1 - \mu(z))_{i_j}$$

$$+ \tfrac{z-2}{z-1}(\overline{S}^{(1,2)} - \mu(z))_{i_j})\} \mid C_0 = z] \cdot E[\tfrac{\partial^2 \psi_z(S_2,u)}{\partial u_{i_1} \partial u_{i_2}} \mid_{u=\mu(z)} \mid C_0 = z]\}$$

$$= \tfrac{1}{2(z-1)^2} \sum_{i_1,i_2=1}^{2} \{\mathrm{Cov}[\psi_z(S_1, \mu(z)), \textstyle\prod_{j=1}^{2} (S_1 - \mu(z))_{i_j} \mid C_0 = z]$$

$$\cdot E[\tfrac{\partial^2 \psi_z(S_2,u)}{\partial u_{i_1} \partial u_{i_2}} \mid_{u=\mu(z)} \mid C_0 = z]\}. \tag{3.2.25}$$

Since by the assumption **A.10** we have $\mid \mathrm{Cov}[\psi_z(S_1, \mu(z)), \prod_{j=1}^{2} (S_1 - \mu(z))_{i_j} \mid C_0 = z] \mid$ and $\mid E[\frac{\partial^2 \psi_z(S_2,u)}{\partial u_{i_1} \partial u_{i_2}} \mid_{u=\mu(z)} \mid C_0 = z] \mid$ are uniformly bounded, we have $E[\xi_{1,1}\xi_{2,3} \mid C_0 = z] = o(z^{-1})$.

To show $E[\xi_{1,1}\xi_{2,4} \mid C_0 = z] = o(z^{-1})$ we use the assumption **A.10** and we can write

$$\mid E[\xi_{1,1}\xi_{2,4} \mid C_0 = z] \mid$$

$$\leq E[\psi_z(S_1, \mu(z)) \mid \hat{R}_z(\overline{S}^{(2)}; S_2) \mid\mid C_0 = z] \tag{3.2.26}$$

$$\leq K E[\mid\mid S_{1,1} \mid\mid\mid \hat{R}_z(\overline{S}^{(2)}; S_2) \mid\mid C_0 = z].$$

By using (3.2.10), for some constant $K > 0$ we have

$$
\begin{aligned}
&\mid \hat{R}_z(\overline{S}^{(2)}; S_2) \mid \\
&\leq K \parallel S_2 \parallel\parallel \overline{S}^{(2)} - \mu(z) \parallel^3 \\
&+K \parallel S_2 \parallel \{1+ \parallel \overline{S}^{(2)} - \mu(z) \parallel \\
&+ \parallel \overline{S}^{(2)} - \mu(z) \parallel^2 +(1+ \parallel \overline{S}^{(2)} \parallel)^{1/2}\}\mathbf{1}_{B_z^c}(\overline{S}^{(2)}).
\end{aligned}
\tag{3.2.27}
$$

But on the other hand we have

$$
\begin{aligned}
&\parallel \overline{S}^{(2)} - \mu(z) \parallel^k \\
&=\parallel \tfrac{1}{z-1}(S_1 - \mu(z)) + \tfrac{z-2}{z-1}(\overline{S}^{(1,2)} - \mu(z)) \parallel^k \\
&\leq 2^{k-1}\{(\tfrac{1}{z-1})^k \parallel S_1 - \mu(z) \parallel^k +(\tfrac{z-2}{z-1})^k \parallel \overline{S}^{(1,2)} - \mu(z) \parallel^k),
\end{aligned}
\tag{3.2.28}
$$

for $k = 1, 2, 3$. Replacing (3.2.27) and (3.2.28) in (3.2.26) for some constant $K > 0$ we get

$$
\begin{aligned}
&\mid E[\xi_{1,1}\xi_{2,4} \mid C_0 = z] \mid \\
&\leq KE[\parallel S_2 \mid C_0 = z]\{E[\parallel S_1 \parallel ((\tfrac{1}{z-1})^3 \parallel S_1 - \mu(z) \parallel^3 \\
&+(\tfrac{z-2}{z-1})^3 \parallel \overline{S}^{(1,2)} - \mu(z) \parallel^3) \mid C_0 = z] \\
&+E[\parallel S_1 \parallel; \overline{S}^{(2)} \in B_z^c \mid C_0 = z] \\
&+E[\parallel S_1 \parallel (\tfrac{1}{z-1} \parallel S_1 - \mu(z) \parallel \\
&+\tfrac{z-2}{z-1} \parallel \overline{S}^{(1,2)} - \mu(z) \parallel); \overline{S}^{(2)} \in B_z^c \mid C_0 = z] \\
&+E[\parallel S_1 \parallel ((\tfrac{1}{z-1})^2 \parallel S_1 - \mu(z) \parallel^2 \\
&+(\tfrac{z-2}{z-1})^2 \parallel \overline{S}^{(1,2)} - \mu(z) \parallel^2); \overline{S}^{(2)} \in B_z^c \mid C_0 = z] \\
&+E[\parallel S_1 \parallel (1+ \parallel \overline{S}^{(1,2)} \parallel)^{1/2}; \overline{S}^{(2)} \in B_z^c \mid C_0 = z]\}.
\end{aligned}
\tag{3.2.29}
$$

Now by the conditional independence of S_1, S_2 and $\overline{S}^{(1,2)}$ and also by the

Hölder inequality we can write (3.2.29) as

$$
\begin{aligned}
&| E[\xi_{1,1}\xi_{2,4} \mid C_0 = z] | \\
&\leq KE[\| S_2 \| | C_0 = z]\{(\tfrac{1}{z-1})^3 E[\| S_1 \| \| S_1 - \mu(z) \|^3 | C_0 = z] \\
&+(\tfrac{z-2}{z-1})^3 E[\| S_1 \| | C_0 = z] E[\| \overline{S}^{(1,2)} - \mu(z) \|^3 | C_0 = z]\} \\
&+KE[\| S_2 \| | C_0 = z] P(\overline{S}^{(2)} \in B_z^c \mid C_0 = z)^{1/2} \\
&\cdot\{E[\| S_1 \|^2 | C_0 = z]^{1/2} \\
&+\tfrac{1}{z-1} E[\| S_1 \|^2 \| S_1 - \mu(z) \|^2 | C_0 = z] \\
&+\tfrac{z-2}{z-1} E[\| S_1 \|^2 | C_0 = z]^{1/2} E[\| \overline{S}^{(1,2)} - \mu(z) \|^2 | C_0 = z]^{1/2} \\
&+(\tfrac{1}{z-1})^2 E[\| S_1 \|^2 \| S_1 - \mu(z) \|^4 | C_0 = z]^{1/2} \\
&+(\tfrac{z-2}{z-1})^2 E[\| S_1 \|^2 | C_0 = z]^{1/2} E[\| \overline{S}^{(1,2)} - \mu(z) \|^4 | C_0 = z]^{1/2} \\
&+E[\| S_1 \|^2 | C_0 = z]^{1/2} E[1 + \| \overline{S}^{(1,2)} \| | C_0 = z]^{1/2}\}
\end{aligned}
\tag{3.2.30}
$$

Now by the assumption **A.10** $E[\| S_1 \|^{2k_0} | C_0 = z]$ is uniformly bounded and further by Lemma 2.6 we have

$$
\begin{aligned}
&E[\| \overline{S}^{(1,2)} - \mu(z) \|^k | C_0 = z] = O(z^{-k/2}), \\
&P(\overline{S}^{(1,2)} \in B^c \mid C_0 = z) = O(z^{-k_0}),
\end{aligned}
\tag{3.2.31}
$$

for $1 \leq k \leq 2k_0$. Replacing these in (3.2.30) we get $| E[\xi_{1,1}\xi_{2,4} \mid C_0 = z] |= o(z^{-1})$. In a similar way we can show that the rest of the terms in (3.2.17) are of order $o(z^{-1})$. Hence,

$$
\mathrm{Cov}[\psi_z(S_1, \overline{S}^{(1)}), \psi_z(S_2, \overline{S}^{(2)}) \mid C_0 = z] = z^{-1}\tilde{a}_7^* + o(z^{-1}), \tag{3.2.32}
$$

where $\tilde{a}_7^*$ is equal to

$$
\tilde{a}_7^* = 2\sum_{i=1}^{2} \Delta_i^{(1)}\Delta_i^{(2)} + \sum_{i,j=1}^{2} \Delta_i^{(1)}\Delta_j^{(1)}\sigma_{i,j}^*. \tag{3.2.33}
$$

Proof of the assertion (iii) can be carried out in a similar way as that of Lemma 3.2.

■

Remark 3.3. As one can easily see the assumptions on Model 2.2 in the critical case are stronger than those of the supercritical case. Hence the assertions of Lemma 3.3 also hold for the critical case. □

Lemma 3.7. *Assume* **A.6**, **A.10** *and* **A.11** *hold. Consider the following decomposition for the process* $\{C_n\}_{n=0}^{\infty}$,

$$C_{n+1} = C_n + g(C_n) + \xi_{n+1}, \tag{3.2.34}$$

where the function g *and the residuals* ξ_{n+1} *are as in (2.2.44), for* $n \geq 0$. *Then*

$$\begin{aligned} &(i) && E[\xi_{n+1} \mid \mathcal{F}_n] = 0, \\ &(ii) && E[\xi_{n+1}^2 \mid \mathcal{F}_n] = \tilde{\sigma}^2(C_n), \ \text{ for some function } \tilde{\sigma} \geq 0, \\ &(iii) && E[\mid \xi_1 \mid^{k_0} \mid C_0 = z] = O(\tilde{\sigma}^{k_0}(z)), \ \text{ for all } z \geq 1, \\ &(iv) && \text{for some constant } \tilde{\mathrm{b}}^* \text{ and } 0 \leq \tilde{\beta}_0 < 1 \\ & && \text{as } z \to \infty, \text{ we have } g(z) = \tilde{b}^* z^{\tilde{\beta}_0} + o(z^{\tilde{\beta}_0}), \\ &(v) && \text{and for some constant } \tilde{c}^* > 0, \text{ we have} \\ & && \tilde{\sigma}^2(z) = \tilde{c}^* z + o(z), \end{aligned} \tag{3.2.35}$$

with k_0 *as in Assumption* **A.10**.

Proof of Lemma 3.7: The proof of parts (i)-(iii) is similar to that of Lemma 2.14. To show part (iv) of the lemma, notice that by Assumption **A.11** and Lemma 3.6, we can decompose $g(z)$ into the following parts.

$$\begin{aligned} g(z) &= E[C_1 \mid C_0 = z] - z = zE[\psi_z(S_1, \overline{S}^{(1)}) \mid C_0 = z] - z \\ &= zE[\psi_z(S_1, \overline{S}^{(1)}) - \psi_z(S_1, \mu(z)) \mid C_0 = z] \\ &+ zE[\psi_z(S_1, \mu(z)) - \psi^*(S_1, \mu(z)) \mid C_0 = z] \\ &+ zE[\psi^*(S_1, \mu(z)) - \psi^*(S_1, \mu^*) \mid C_0 = z] \\ &+ z(E[\psi^*(S_1, \mu^*) \mid C_0 = z] - 1) = \tilde{b}^* z^{\tilde{\beta}_0} + o(z^{\tilde{\beta}_0}), \end{aligned} \tag{3.2.36}$$

where

$$\begin{aligned} & \tilde{\beta}_0 = \max\{0, 1 - \tilde{\beta}_i : i = 1, 2, 3\} \\ \text{and} \quad & \tilde{b}^* = \textstyle\sum_{i=1}^{3} \tilde{a}_i^* \mathbf{1}_{\{\tilde{\beta}_0 = 1 - \tilde{\beta}_i\}} + \tilde{a}_4^* \mathbf{1}_{\{\tilde{\beta}_0 = 0\}}. \end{aligned} \tag{3.2.37}$$

To show part (v), we can write $\tilde{\sigma}^2(z)$ as

$$\begin{aligned}
\tilde{\sigma}^2(z) &= \mathrm{Var}[\xi_1 \mid C_0 = z] \\
&= \mathrm{Var}[E[\xi_1 \mid \mathcal{G}_0] \mid C_0 = z] + E[\mathrm{Var}[\xi_1 \mid \mathcal{G}_0] \mid C_0 = z] \\
&= \mathrm{Var}[\textstyle\sum_{j=1}^{z} E[U_j \mid \mathcal{G}_0] \mid C_0 = z] \\
&+ E[\textstyle\sum_{j=1}^{z} \mathrm{Var}[U_j \mid \mathcal{G}_0] \mid C_0 = z].
\end{aligned} \tag{3.2.38}$$

Since $E[U_j \mid \mathcal{G}_0] = \psi_z(S_j, \overline{S}^{(j)})$ and $\mathrm{Var}[U_j \mid \mathcal{G}_0] = \tilde{\nu}_{2,z}(S_j, \overline{S}^{(1)})$, (3.2.38) can be written as

$$\begin{aligned}
\tilde{\sigma}^2(z) &= z\mathrm{Var}[\psi_z(S_1, \overline{S}^{(1)}) \mid C_0 = z] \\
&+ z(z-1)\mathrm{Cov}[\psi_z(S_1, \overline{S}^{(1)}), \psi_z(S_2, \overline{S}^{(2)}) \mid C_0 = z] \\
&+ zE[\tilde{\nu}_{2,z}(S_1, \overline{S}^{(1)}) \mid C_0 = z].
\end{aligned} \tag{3.2.39}$$

Now, by Lemmas 3.6 and Assumption **A.11** , we have

$$\begin{aligned}
&\mathrm{Var}[\psi_z(S_1, \overline{S}^{(1)}) \mid C_0 = z] \\
&= E[\psi_z^2(S_1, \overline{S}^{(1)}) \mid C_0 = z] - E[\psi_z(S_1, \overline{S}^{(1)}) \mid C_0 = z]^2 \\
&= (1 + \tilde{a}_6^* + o(1)) - (1 + o(1))^2 = \tilde{a}_6^* + o(1),
\end{aligned} \tag{3.2.40}$$

and

$$\mathrm{Cov}[\psi_z(S_1, \overline{S}^{(1)}), \psi_z(S_2, \overline{S}^{(2)}) \mid C_0 = z] = \tilde{a}_7^* z^{-1} + o(z^{-1}). \tag{3.2.41}$$

On the other hand from Assumption **A.11** and Lemma 3.6, we have

$$E[\tilde{\nu}_{2,z}(S_1, \overline{S}^{(1)}) \mid C_0 = z] = \tilde{a}_5^* + O(z^{-1}). \tag{3.2.42}$$

Now by replacing (3.2.40), (3.2.41) and (3.2.42) in (3.2.39) we would get

$$\tilde{\sigma}^2(z) = (\tilde{a}_5^* + \tilde{a}_6^* + \tilde{a}_7^*)z + o(z), \tag{3.2.43}$$

where $\tilde{a}_5^*$, $\tilde{a}_6^*$ and $\tilde{a}_7^*$ are as in (3.2.4)(viii), (3.2.4)(ix) and (3.2.5), respectively. Thus $\tilde{c}^*$ in (3.2.35) is equal to

$$\tilde{c}^* = \tilde{a}_5^* + \tilde{a}_6^* + \tilde{a}_7^*. \tag{3.2.44}$$

■

Remark 3.4. From Assumption **A.11** and Lemma 3.6(i) we have

$$
\begin{aligned}
&(i) \quad E[\psi_z(S_1,\mu(z)) - \psi^*(S_1,\mu(z)) \mid C_0 = z] = \tilde{a}_1^* z^{-\tilde{\beta}_1} + o(z^{-\tilde{\beta}_1}),\\
&(ii) \quad E[\psi^*(S_1,\mu(z)) - \psi^*(S_1,\mu^*) \mid C_0 = z] = \tilde{a}_2^* z^{-\tilde{\beta}_2} + o(z^{-\tilde{\beta}_2}),\\
&(iii) \quad E[\psi^*(S_1,\mu^*) \mid C_0 = z] = \tilde{m} + \tilde{a}_3^* z^{-1} + o(z^{-1}),\\
&\qquad\qquad\qquad\qquad\qquad\; = 1 + \tilde{a}_3^* z^{-1} + o(z^{-1}),\\
&(iv) \quad E[\psi_z(S_1,\overline{S}^{(1)}) - \psi_z(S_1,\mu(z)) \mid C_0 = z] = \tilde{a}_4^* z^{-1} + o(z^{-1}),
\end{aligned}
\tag{3.2.45}
$$

where $\tilde{a}_4^*$ in (3.2.45)(iv) is equal to

$$
\tilde{a}_4^* = \frac{1}{2}\sum_{i,j=1}^{2} \sigma_{i,j}^* \Delta_{i,j}^{(1)}, \tag{3.2.46}
$$

and the parameter $\Delta_{i,j}^{(1)}$ are equal to

$$
\Delta_{i,j}^{(1)} = \lim_{z\to\infty} E[\frac{\partial^2 \psi_z(S_1,y)}{\partial y_i \partial y_j} \mid_{y=\mu(z)} \mid C_0 = z]. \tag{3.2.47}
$$

Let us have a closer look at (3.2.45)(i)-(iv). The distribution of S_1 in (3.2.45)(i)-(iv), corresponding to the reproduction stage, depends as before on $C_0 = z$. (3.2.45)(i)-(iii) give various rates of convergence for the expected number of the marriages of a family, say the first one, when what is happening outside the family (this determines by $\overline{S}^{(1)}$) is taken to be fixed.

(i) Assume that $\overline{S}^{(1)} \equiv \mu(z)$, then (3.2.45)(i) gives the rate of convergence of the expected number of marriages of the daughters of the first family when in the mating stage there are z families around, i.e. $E[\psi_z(S_1,\mu(z)) \mid C_0 = z]$, to the case when in the mating stage there are infinity number of families around, i.e. $E[\psi^*(S_1,\mu(z)) \mid C_0 = z]$.

(ii) Assume that the daughters of the first family want to marry and that in the mating stage there are infinity many families around. Then (3.2.45)(ii) gives the rate of convergence of the expected number of marriages of the daughters of the first family given $\overline{S}^{(1)} \equiv \mu(z)$, i.e. $E[\psi^*(S_1,\mu(z)) \mid C_0 = z]$, to the case when $\overline{S}^{(1)} \equiv \mu^*$, i.e. $E[\psi^*(S_1,\mu^*) \mid C_0 = z]$.

(iii) Assume that the daughters of the first family want to marry and that in the mating stage there are infinity many families around. Take

$\overline{S}^{(1)} \equiv \mu^*$. Then (3.2.45)(iii) gives the rate of convergence of the expected number of marriages of the daughters of the first family, i.e. $E[\psi^*(S_1, \mu^*) \mid C_0 = z]$, to the case where the number couples in the reproduction stage is infinity, i.e. $\tilde{m} = \lim_{z\to\infty} E[\psi^*(S_1, \mu^*) \mid C_0 = z]$.

(iv) (3.2.45)(iv) gives the rate of convergence of the expected number of marriages of the daughters of the first family when z there are families around, i.e. $E[U_1 \mid C_0 = z] = E[\psi_z(S_1, \overline{S}^{(1)}) \mid C_0 = z]$, to the case where in the mating stage we have $\overline{S}^{(1)} \equiv \mu(z)$, i.e. $E[\psi_z(S_1, \mu(z)) \mid C_0 = z]$.

Since by symmetry we have $E[C_1 \mid C_0 = z] = zE[\psi_z(S_1, \overline{S}^{(1)}) \mid C_0 = z]$, as in the proof of Lemma 3.7(iv) we have

$$\begin{aligned} &\tilde{\beta}_0 = \max\{0, 1 - \tilde{\beta}_i : i = 1, 2, 3\}, \\ &\tilde{b}^* = \textstyle\sum_{i=1}^{3} \tilde{a}_i^* \mathbf{1}_{\{\tilde{\beta}_0 = 1 - \tilde{\beta}_i\}} + \tilde{a}_4^* \mathbf{1}_{\{\tilde{\beta}_0 = 0\}}. \end{aligned} \tag{3.2.48}$$

We can break up $\mathrm{Var}[C_1 \mid C_0 = z]$ into three terms (see the proof of Lemma 3.4(v)) as follows

$$\begin{aligned} &\mathrm{Var}[C_1 \mid C_0 = z] = zE[\tilde{\nu}_{2,z}(S_1, \overline{S}^{(1)}) \mid C_0 = z] \\ &+ z\mathrm{Var}[\psi_z(S_1, \overline{S}^{(1)}) \mid C_0 = z] \\ &+ z(z-1)\mathrm{Cov}[\psi_z(S_1, \overline{S}^{(1)}), \psi_z(S_2, \overline{S}^{(2)}) \mid C_0 = z]. \end{aligned} \tag{3.2.49}$$

We have

$$\begin{aligned} &(i) \quad E[\tilde{\nu}_{2,z}(S_1, \overline{S}^{(1)}) \mid C_0 = z] = \tilde{a}_5^* + o(1), \\ &(ii) \quad \mathrm{Var}[\psi_z(S_1, \overline{S}^{(1)}) \mid C_0 = z] = \tilde{a}_6^* + o(1), \\ &(iii) \quad \mathrm{Cov}[\psi_z(S_1, \overline{S}^{(1)}), \psi_z(S_2, \overline{S}^{(2)}) \mid C_0 = z] \\ &\qquad = \tilde{a}_7^* z^{-1} + o(z^{-1}), \end{aligned} \tag{3.2.50}$$

where for the parameters $\tilde{a}_5^*$ and $\tilde{a}_6^*$ see Assumption **A.11**(viii) and (ix), respectively, and from Lemma 3.6(iii) the value of the parameter $\tilde{a}_7^*$ is equal to

$$\tilde{a}_7^* = 2\textstyle\sum_{i=1}^{2} \Delta_i^{(1)} \Delta_i^{(2)} + \sum_{i,j=1}^{2} \sigma_{i,j}^* \Delta_i^{(1)} \Delta_j^{(1)}, \tag{3.2.51}$$

where $\Delta_i^{(1)}$ and $\Delta_i^{(2)}$ are equal to

$$\begin{aligned}\Delta_i^{(1)} &= \lim_{z\to\infty} E[\tfrac{\partial \psi_z(S_1,y)}{\partial y_i}\ |_{y=\mu(z)}|\ C_0 = z],\\ \Delta_i^{(2)} &= \lim_{z\to\infty} \mathrm{Cov}[(S_1)_i, \psi_z(S_1,\mu(z)) \mid C_0 = z],\end{aligned} \tag{3.2.52}$$

where $(S_1)_i$ is the ith component of the vector S_1, for $i = 1, 2$.

From (3.2.45)-(3.2.52) one can see how the values of the parameters μ^* and $\Sigma^* = [\sigma^*_{i,j}]$ (corresponding to the reproduction stage) together with behaviors of the limiting functions ψ^* and $\tilde{\nu}_2^*$ at the point μ^* and its neighbourhood (corresponding to the mating stage), determine the values of the parameters $\tilde{\beta}_0$, $\tilde{b}*$ and $\tilde{c}^*$. □

The following theorem summarises the property of the process $\{C_n\}$ in the critical case for the Model 2.2 and is the counterpart of Theorem 2.15.

Theorem 3.8. *Let Model 2.2 in the critical case satisfy* **A.6**, **A.10** *and* **A.11**. *By Lemma 3.7(iv)-(v),* $g(x) = \tilde{b}^* x^{\tilde{\beta}_0} + o(x^{\tilde{\beta}_0})$ *and* $\sigma^2(z) = \tilde{c}^* z + o(z)$, *where* $\tilde{b}^* \in \mathcal{R}$, $0 \le \tilde{\beta}_0 < 1$ *and* $\tilde{c}^* > 0$.

(i) If $\tilde{b}^* \le 0$, *then the process* $\{C_n\}$ *dies out almost surely,* i.e. $P(C_n \to \infty) = 0$.

(ii) If either $\tilde{\beta}_0 = 0$ *and* $2\tilde{b}^* > \tilde{c}^*$ *or* $0 < \tilde{\beta}_0 < 1$, *then*

$$P(C_n \to 0 \ \text{ or } \ \infty) = 1 \ \text{ and } \ P(C_n \to \infty) > 0. \tag{3.2.53}$$

With $\{a_n\}_{n=0}^{\infty}$ *as in (2.3.41) and* $\mathcal{E}_\infty = \{C_n \to \infty\}$, *we have the following.*

(iii) If $\tilde{\beta}_0 = 0$ *and* $2\tilde{b}^* > \tilde{c}^*$ *then*

$$P(C_n/n \le t \mid \mathcal{E}_\infty) \to \frac{2}{\tilde{c}^*\Gamma(\frac{2\tilde{b}^*}{\tilde{c}^*})} \int_0^t (\tfrac{2x}{\tilde{c}^*})^{\frac{2\tilde{b}^*}{\tilde{c}^*}-1} e^{-\frac{2x}{\tilde{c}^*}} dx, \tag{3.2.54}$$

as $n \to \infty$.

(iv) If $\tilde{b}^*$ *and* $0 < \tilde{\beta}_0 \le 2/3$ *then* $C_n/a_n \to 1$ *in probability and*

$$\tilde{K}^{-1}_{\tilde{\beta}_0,n} \frac{C_n - a_n}{g(a_n)} \stackrel{\text{weakly}}{\longrightarrow} \mathcal{N}(0,1), \tag{3.2.55}$$

where $\tilde{K}_{\tilde{\beta}_0,n} > 0$ *is equal to*

$$\tilde{K}_{\tilde{\beta}_0,n} \approx \begin{cases} \sqrt{\frac{\tilde{c}^*}{2-3\tilde{\beta}_0}} (\tilde{b}^*)^{\frac{1}{2(1-\tilde{\beta}_0)}} (n(1-\tilde{\beta}_0))^{\frac{2-3\tilde{\beta}_0}{2(1-\tilde{\beta}_0)}} & \text{if } \ 0 < \tilde{\beta}_0 < 2/3, \\ \sqrt{\frac{3\tilde{c}^*}{(\tilde{b}^*)^3}} \sqrt{\log n}, & \text{if } \ \tilde{\beta}_0 = 2/3. \end{cases} \tag{3.2.56}$$

(v) If $\tilde{b}^* > 0$ and $2/3 < \tilde{\beta}_0 < 1$ then

$$C_n/a_n \to 1 \quad \text{and} \quad \tfrac{C_n - a_n}{g(a_n)} \quad \text{converges,} \tag{3.2.57}$$

a.s. on $\mathcal{E}_\infty$ as $n \to \infty$.

Proof of Theorem 3.8: First notice that Assumption **A.6** corresponds to Assumption **A.1**; Assumption **A.10** corresponds to Assumptions **A.3** and **A.4**; and Assumption **A.11** corresponds to Assumptions **A.5** for Model 2.1. Secondly, Lemma 3.7 is the counter part to Lemma 2.14.

To prove Theorem 3.8 it is enough to replace the use of Lemma 2.14 in the proof of Theorem 2.15 by Lemma 3.7 and the rest of the proof is the same. ■

3.3. Examples

In this section, similar to Section 2.4, we present some examples for Model 2.2 in the supercritical - as well as the critical case.

The first example is related to Model 2.2 in the supercritical case.

Example 3.1. Consider Model 2.2. With the notations of this chapter, we present some examples for the conditional disribution of $U_{1,1}$ given $\mathcal{G}_0$.

(1) Define the function p_z on $\mathcal{R}_+^2 \times \mathcal{R}_+^2$ as follows:

$$p_z(x,y) := \frac{x_2}{x_1 + x_2} \mathbf{1}_{\{x_1 > 0, z > 0\}}, \tag{3.3.1}$$

for $x, y \in \mathcal{R}_+^2$. Then assume that $U_1 \mid \mathcal{G}_0$ is distributed as

$$U_1 \mid \mathcal{G}_0 \sim \mathrm{Bin}(X_1, p_z(S_1, \overline{S}^{(1)})). \tag{3.3.2}$$

The function p_z in (3.3.1) is an example of the situation where one has only inbreeding in the population, i.e. the daughters of each family mate only with their brothers.

(2) Define the function p_z on $\mathcal{R}_+^2 \times \mathcal{R}_+^2$ as follows:

$$p_z(x,y) := \frac{y_2}{y_1 + y_2} \mathbf{1}_{\{x_1 > 0, z > 0\}}, \tag{3.3.3}$$

for $x, y \in \mathcal{R}_+^2$. Then assume that $U_1 \mid \mathcal{G}_0$ is distributed as

$$U_1 \mid \mathcal{G}_0 \sim \mathrm{Poisson}(X_1 p_z(S_1, \overline{S}^{(1)})). \tag{3.3.4}$$

The function p_z in (3.3.3) is an example of the situation where inbreeding is not allowed in the population, i.e. the daughters of each family could only mate with the sons of the other families.

(3) Define the function p_z on $\mathcal{R}_+^2 \times \mathcal{R}_+^2$ as follows:

$$p_z(x,y) := \frac{x_2}{x_1+x_2} \cdot \frac{\| x-y \|}{h+ \| x-y \|} + \frac{y_2}{y_1+y_2} \cdot \frac{h}{h+ \| x-y \|}, \quad (3.3.5)$$

for $x, y \in \mathcal{R}_+^2$ and some constant $h > 0$. Then assume that $U_1 \mid \mathcal{G}_0$ is distributed as

$$U_1 \mid \mathcal{G}_0 \sim \text{Poisson}(X_1 p_z(S_1, \overline{S}^{(1)})). \quad (3.3.6)$$

Having a close look at the function $p_z(S_1, \overline{S}^{(1)})$ in (3.3.5)-(3.3.6) we can see that the chance of a female getting married depends on the proportion of the males in her family as well as the rest of the family in her generation.

□

The next example is for Model 2.2 in the critical case.

Example 3.2. For fixed $a > 0$, let the reproduction law for a couple in the zeroth generation, given the total number of couples in that generation, be that of Example 2.2 (see (2.4.2)). From (2.4.4) one can see that the values of the vectors $\mu(z) = (\mu_1(z), \mu_2(z))$, $\mu^* = (\mu_1^*, \mu_2^*)$ and the matrices $\Sigma(z) = [\sigma_{i,j}(z)]$ and $\Sigma^* = [\sigma_{i,j}^*]$ are equal to

$$\begin{aligned}
&\mu_1(z) = \mu_2(z) = \tfrac{5}{4}(1 + \tfrac{1}{2} z^{-a}), \quad \mu_1^* = \mu_2^* = \tfrac{5}{4}, \\
&\sigma_{11}(z) = \sigma_{22}(z) = \tfrac{11}{16} - \tfrac{7}{16} z^{-a} + o(z^{-a}), \\
&\sigma_{12}(z) = \sigma_{21}(z) = \tfrac{7}{16} - \tfrac{9}{16} z^{-a} + o(z^{-a}), \\
&\sigma_{11}^* = \sigma_{22}^* = \tfrac{11}{16}, \quad \sigma_{12}^* = \sigma_{21}^* = \tfrac{7}{16},
\end{aligned} \quad (3.3.7)$$

for $z \geq 4^{1/a}$.

Define the functions R_z and R^* on $\mathcal{R}_+^2 \times \mathcal{R}_+^2$ as

$$\begin{aligned}
&R_z(x,y) = \{\alpha(\tfrac{x_2}{x_1+x_2})^{d_0+d_1 z^{-h_1}} + (1-\alpha)(\tfrac{y_2}{y_1+y_2})^{d_2+d_3 z^{-h_2}}\} \mathbf{1}_{\{x_1, z>0\}}, \\
&R^*(x,y) = \{\alpha(\tfrac{x_2}{x_1+x_2})^{d_0} + (1-\alpha)(\tfrac{y_2}{y_1+y_2})^{d_2}\} \mathbf{1}_{\{x_1>0\}},
\end{aligned} \quad (3.3.8)$$

where the parameters $0 \leq \alpha \leq 1$, $h_1, h_2, d_0, d_1 \in \mathcal{R}_+$ and $d_1, d_3 \in \mathcal{R}$, are such that $0 \leq R_z(x,y), R^*(x,y) \leq 1$, for all $z \in \mathcal{Z}_+$ and all $(x,y) \in \mathcal{R}_+^2 \times \mathcal{R}_+^2$.

Let $C_{n-1} = z \in \mathcal{Z}_+$, then assume that the conditional distribution of the number of marriages of the daughters of the jth family in the nth generation, given $\mathcal{G}_{n-1}$, is equal to

$$U_{n,j} \mid \mathcal{G}_{n-1} \sim \mathrm{Bin}(X_{n,j}, R_z(S_{n,j}, \overline{S}_n^{(j)})), \tag{3.3.9}$$

for $j = 1, 2, \ldots, z$ and $n = 1, 2, \ldots$.

The functions ψ_z, ψ^*, $\tilde{\nu}_{2,z}$ and $\tilde{\nu}_2^*$ defined in (3.0.2)-(3.0.3) for this example are equal to

$$\begin{aligned} \psi_z(x,y) &= x_1 R_z(x,y), \\ \psi^*(x,y) &= x_1 R^*(x,y), \\ \tilde{\nu}_{2,z}(x,y) &= x_1 R_z(x,y)(1 - R_z(x,y)), \\ \tilde{\nu}_2^*(x,y) &= x_1 R^*(x,y)(1 - R^*(x,y)), \end{aligned} \tag{3.3.10}$$

for $z > 0$ and $x, y \in \mathcal{R}_+^2$. The parameter $\tilde{m}$ given in (3.0.4) will be equal to

$$\tilde{m} = \frac{1}{4}\alpha\{(2/3)^{d_0} + 2(1/3)^{d_0} + 2(1/2)^{d_0}\} + \frac{5}{4}(1-\alpha)(1/2)^{d_2}. \tag{3.3.11}$$

Since we are considering the critical case, the parameters α, d_0 and d_2 should be such that $\tilde{m} = 1$. Some possible values are give in the following table:

	d_0	d_2	α
(1)	0.2	0.4	0.44
(2)	0.2	1.0	0.85
(3)	0.2	1.8	0.91
(4)	0.4	0.2	0.51
(5)	1.0	0.2	0.17
(6)	1.8	0.2	0.12

(3.3.12)

The values of the parameters $\tilde{a}_i^*$ and $\tilde{\beta}_j$, for $i = 1, 2, 3, 5, 6$, and $j =$

$1, 2, 3$, presented in Assumption **A.11** are

$$\tilde{a}_1^* = \begin{cases} \frac{\alpha}{4} d_1 \{(2/3)^{d_0} \log(2/3) + 2(1/3)^{d_0} \log(1/3) \\ +2(1/2)^{d_0} \log(1/2)\}, & \text{if } \ h_1 < h_2, \\ \frac{5}{4}(1-\alpha) d_3 (1/2)^{d_3} \log(1/2), & \text{if } \ h_1 > h_2, \\ \frac{\alpha}{4} d_1 \{(2/3)^{d_0} \log(2/3) + 2(1/3)^{d_0} \log(1/3) \\ +2(1/2)^{d_0} \log(1/2)\} \\ +\frac{5}{4}(1-\alpha) d_3 (1/2)^{d_3} \log(1/2), & \text{if } \ h_1 = h_2, \end{cases} \tag{3.3.13}$$

$$\begin{aligned} \tilde{a}_2^* &= 0, \\ \tilde{a}_3^* &= 1/2, \\ \tilde{a}_5^* &= \tfrac{1}{4} R^*((1,2),\mu^*)(1 - R^*((1,2),\mu^*)) \\ &\quad + \tfrac{1}{2} R^*((2,1),\mu^*)(1 - R^*((2,1),\mu^*)) \\ &\quad + \tfrac{1}{2} R^*((2,2),\mu^*)(1 - R^*((2,2),\mu^*)), \end{aligned} \tag{3.3.14}$$

and

$$\begin{aligned} \tilde{a}_6^* &= \tfrac{\alpha^2}{4}\{(2/3)^{2d_0} + 4(1/3)^{2d_0} + 4(1/2)^{2d_0}\} + \tfrac{(1-\alpha)^2}{4}(1/2)^{2d_2} \\ &\quad + \tfrac{\alpha(1-\alpha)}{2}(1/2)^{d_2}\{(2/3)^{d_0} + 4(1/3)^{d_0} + 4(1/2)^{d_0}\} - 1. \end{aligned} \tag{3.3.15}$$

Further we have $\tilde{\beta}_1 = \min(h_1, h_2)$, $\tilde{\beta}_2 = \infty$ and $\tilde{\beta}_3 = a$.

The values of parameters $\Delta_i^{(1)}$, $\Delta_{i,j}^{(1)}$ and $\Delta_i^{(2)}$ presented in Assumption **A.10** are

$$\begin{cases} \Delta_1^{(1)} = -(1-\alpha) d_2 (1/2)^{d_2+1}, \\ \Delta_2^{(1)} = (1-\alpha) d_2 (1/2)^{d_2+1}, \end{cases}$$

$$\begin{cases} \Delta_{1,1}^{(1)} = \frac{1}{5}(1-\alpha) d_2 (d_2 - 3)(1/2)^{d_2}, \\ \Delta_{2,2}^{(1)} = \frac{1}{5}(1-\alpha) d_2 (d_2 - 3)(1/2)^{d_2}, \\ \Delta_{1,2}^{(1)} = -\frac{1}{5}(1-\alpha) d_2 (d_2 - 1)(1/2)^{d_2}, \end{cases} \tag{3.3.16}$$

and

$$
\begin{cases}
\Delta_1^{(2)} & = \frac{\alpha}{16}\{-(2/3)^{d_0} + 6(1/3)^{d_0} + 6(1/2)^{d_0}\} \\
& +\frac{11}{16}(1-\alpha)(1/2)^{d_2}, \\
\Delta_2^{(2)} & = \frac{\alpha}{16}\{3(2/3)^{d_0} - 2(1/3)^{d_0} + 6(1/2)^{d_0}\} \\
& +\frac{7}{16}(1-\alpha)(1/2)^{d_2}.
\end{cases}
\tag{3.3.17}
$$

The values of the parameters $\tilde{a}_4^*$ and $\tilde{a}_7^*$ presented in Lemma 3.6 are equal to

$$
\begin{aligned}
\tilde{a}_4^* & = \tfrac{1}{2}\textstyle\sum_{i,j=1}^{2} \Delta_{i,j}^{(1)}\sigma_{i,j}^* \\
& = \tfrac{1}{5}(1-\alpha)d_2(2d_2-13)(1/2)^{d_2+3}, \\
\tilde{a}_7^* & = 2\textstyle\sum_{i=1}^{2} \Delta_i^{(1)}\Delta_i^{(2)} + \textstyle\sum_{i,j=1}^{2} \Delta_i^{(1)}\Delta_j^{(1)}\sigma_{i,j}^* \\
& = \alpha(1-\alpha)d_2(1/2)^{d_2+2}\{(2/3)^{d_0} - 2(1/3)^{d_0}\} \\
& +(1-\alpha)^2(d_2^2-2d_2)(1/2)^{2d_2+3}.
\end{aligned}
\tag{3.3.18}
$$

Finally, the values of the parameters $\tilde{b}^*$, $\tilde{\beta}_0$ and $\tilde{c}^*$ are equal to

$$
\begin{aligned}
& \tilde{\beta}_0 = \max\{0, 1-\tilde{\beta}_i : i = 1,2,3\}, \\
& \tilde{b}^* = \textstyle\sum_{i=1}^{3} \tilde{a}_i^* \mathbf{1}_{\{\tilde{\beta}_0 = 1-\tilde{\beta}_i\}} + \tilde{a}_4^* \mathbf{1}_{\{\tilde{\beta}_0 = 0\}}, \\
& \tilde{c}^* = \tilde{a}_5^* + \tilde{a}_6^* + \tilde{a}_7^*,
\end{aligned}
\tag{3.3.19}
$$

(see Lemma 3.7). The behavors of the process $\{C_n\}$ in the critical case have been studied in Theorem 3.8, where the different values of the parameters $\tilde{b}^*$, $\tilde{\beta}_0$ and $\tilde{c}^*$ determine its asymptotic behavior.

In the following we consider different numerical values for the parameters of our example, i.e. a, α, $d_0, \ldots, d_3$, and h_1 and h_2, where for each case, we calculate the values of the parameters $\tilde{\beta}_0$, $\tilde{b}^*$ and $\tilde{c}^*$.

Case #	(1)	(2)	(3)	(4)	(5)
a	1	1	0.5	0.5	0.5
α	0.4395	0.4395	0.4395	0.4395	0.4395
d_0	0.2	0.2	0.2	0.2	0.2
d_1	1	-1	-1	-1	-1
d_2	0.4	0.4	0.4	0.4	0.4
d_3	1	-1	-1	-1	-1
h_1	1	1	1	1/3	1/4
h_2	1	1	1	1/3	1/4
$\tilde{\beta}_0$	0	0	0.5	2/3	3/4
$\tilde{b}^*$	-0.1621	1.7869	0.5	1.3388	1.3388
$\tilde{c}^*$	0.2402	0.2402	0.2402	0.2402	0.2402

(3.3.20)

The five different cases considered in Table 3.3.20 correspond to the five different possible asymptotic limits which the process $\{C_n\}$ might have.

Case #1: In this case $\tilde{b}^* = -0.1621 < 0$ and hence the assertion (i) of Theorem 3.8 says that the process $\{C_n\}$ becomes extinct with probability one, i.e. $P(C_n = 0 \text{ eventually}) = 1$.

Case #2: In this case, $\tilde{b}^* = 1.7869 > 0$ and $\tilde{\beta}_0 = 0$, hence by the assertion (ii) of Theorem 3.8, for the process $\{C_n\}$ to survive with positive probability we should have $2\tilde{b}^* > \tilde{c}^*$. Since $2\tilde{b}^* = 3.5738 > \tilde{c}^* = 0.2763$, the process $P(C_n \to \infty) > 0$. The assertion (iii) of Theorem 3.8 says that the distribution of C_n/n, given the process $\{C_n\}$ survives, converges toward the distribution of a Gamma random variable with the scale parameter $\tilde{c}^*/2 = 0.13815$ and the shape parameter $2\tilde{b}^*/\tilde{c}^* = 12.9345$.

Case #3: In this case, since $\tilde{b}^* = 0.5 > 0$ and $\tilde{\beta}_0 = 0.5 \in (0, 2/3]$, by the assertion (ii) of Theorem 3.8 the process $\{C_n\}$ survives with a positive probability. With the function g equal to $g(x) = \tilde{b}^* x^{\tilde{\beta}_0} + o(x^{\tilde{\beta}_0})$ (see Lemma 3.7) and further with the deterministic sequence $\{a_n\}_{n=0}^{\infty}$ such that $a_0 = 1$ and $a_{n+1} = a_n + g(a_n)$, the assertion (iv) of Theorem 3.8 says that C_n converges in probability to a_n, i.e. for any $\epsilon > 0$ we have $P(| C_n/a_n - 1 |> \epsilon) \to 0$ as $n \to \infty$.

For given n, the assertion (iv) of Theorem 3.8 describes the random variation of C_n (the number of couples in the nth generation) around a_n. Take $C_0 = 1$, then an asymptotic 95% symmetric confidence interval for the number of the couples in the nth generation when

$n = 100$, is equal to

$$\begin{aligned} C_n &= a_n \pm 1.96\tilde{K}_{\tilde{\beta}_0,n} g(a_n) \\ &= 655 \pm 61, \end{aligned} \tag{3.3.21}$$

where $\tilde{K}_{\tilde{\beta}_0,n} = \sqrt{\frac{\tilde{c}^*}{2-3\tilde{\beta}_0}}(\tilde{b}^*)^{\frac{1}{2(1-\tilde{\beta}_0)}}(n(1-\tilde{\beta}_0))^{\frac{2-3\tilde{\beta}_0}{2(1-\tilde{\beta}_0)}}$.

Case #4: This case is similar to the previous case except that since $\tilde{\beta}_0 = 2/3$, the constant $\tilde{K}_{\tilde{\beta}_0,n}$ used to describe the random variation of C_n around a_n will be equal to $\tilde{K}_{\tilde{\beta}_0,n} = \sqrt{\frac{3\tilde{c}^*}{(\tilde{b}^*)^3}}\sqrt{\log n}$. An asymptotic 95% symmetric confidence interval for C_n when $n = 100$, is equal to

$$\begin{aligned} C_n &= a_n \pm 1.96\tilde{K}_{\tilde{\beta}_0,n} g(a_n) \\ &= 85006 \pm 5966. \end{aligned} \tag{3.3.22}$$

Case #5: In this case, since $\tilde{b}^* = 1.3387 > 0$ and $\tilde{\beta}_0 = 3/4 \in (2/3, 1)$, again by the assertion (ii) of Theorem 3.8 the process $\{C_n\}$ survives with positive probability. With the deterministic sequence $\{a_n\}_{n=0}^{\infty}$ as before, by the assertion (iv) of Theorem 3.8 the process C_n converges almost surely to a_n as $n \to \infty$, i.e. for any $\epsilon > 0$ we have $P(\mid C_n/a_n - 1 \mid > \epsilon \text{ infinity often}) = 0$.

By the assertion (v) of Theorem 3.8 the rate of convergence of the process $\{C_n\}$ to $\{a_n\}$ is not slower than $O(a_n^{\tilde{\beta}_0})$. For $n = 100$, we have $a_n = 1.15264 \times 10^6$.

□

Example 3.3. We know that Model 2.1 is an special case of Model 2.2. In Chapter 3 we used different notations for the parameters of Model 2.2 than those of Chapter 2. In Remark 3.1 we already give some relations between different functions defined for these two models. In this example, we consider a simple example of Model 2.1 and then we derive the relationships between the parameters β_i, a_j^*, b^* and c^*, for $i = 0, 1, 2$ and $j = 1, \ldots, 4$, of Section 2.3 to the parameters $\tilde{\beta}_i$, $\tilde{a}_j^*$, $\tilde{b}^*$ and $\tilde{c}^*$, for $i = 0, 1, 2, 3$ and $j = 1, \ldots, 7$, of Section 3.2.

Let the reproduction law for a couple in the zeroth generation, given the total number of couples in that generation, be that of Example 2.2. Let U_1 given $\mathcal{G}_0$ has the following conditional distribution

$$U_1 \mid \mathcal{G}_0 \sim \text{Bin}(X_1, \varphi_{C_0}(\overline{S})), \tag{3.3.23}$$

where the sequence of the functions $\{\varphi_z\}$ is given by

$$\varphi_z(x) = \exp\{-(d_0 + d_1 z^{-h})(x_1 + x_2)\}\mathbf{1}_{\{\min\{z,x_1,x_2\}>0\}}, \tag{3.3.24}$$

where $d_0 = \frac{2}{5}\log(\frac{5}{4})$ and d_1, h are positive constants.

First we derive the parameters β_i, a_j^*, b^* and c^*, for $i = 0, 1, 2$ and $j = 1, \ldots, 4$, as it given in Section 2.3 of Chapter 2. It is easy to see that the functions φ^*, ν_z and ν^* given in (2.1.7)-(2.1.8), are equal to

$$\begin{aligned} \varphi^*(x) &= \exp\{-d_0(x_1 + x_2)\}\mathbf{1}_{\{x_1,x_2>0\}}, \\ \nu_z(x) &= \varphi_z(x)(1 - \varphi_z(x)), \\ \nu^*(x) &= \varphi^*(x)(1 - \varphi^*(x)), \end{aligned} \tag{3.3.25}$$

for $x \in \mathcal{R}_+^2$. As in Example 2.2, for some fixed constant $a > 0$, we have

$$\begin{aligned} &\mu_1(z) = \mu_2(z) = \tfrac{5}{4}(1 + \tfrac{1}{2}z^{-a}), \\ &\mu_1^* = \mu_2^* = \tfrac{5}{4}, \\ &\sigma_{1,1}^* = \sigma_{2,2}^* = \tfrac{11}{16}, \\ &\sigma_{1,2}^* = \sigma_{2,1}^* = \tfrac{7}{16}. \end{aligned} \tag{3.3.26}$$

We are considering the critical case (see (2.1.9), i.e.

$$\begin{aligned} m &= \mu_1^*\varphi^*(\mu^*) \\ &= \mu_1^* \exp\{-d_0(\mu_1^* + \mu_2^*)\} \\ &= \tfrac{5}{4}\exp\{-\tfrac{5}{2}d_0\} = 1. \end{aligned} \tag{3.3.27}$$

From (2.3.2), (2.3.4), (2.3.19), and (2.3.23), together with some calculations we obtain

$$(i) \quad \mu_1(z)\varphi_z(\mu(z)) = \mu_1(z)\varphi^*(\mu(z)) + a_1^* z^{-\beta_1} + o(z^{-\beta}),$$

where $\beta_1 = h$, and $a_1^* = -\frac{5}{2}d_1$,

$$(ii) \quad \mu_1(z)\varphi^*(\mu(z)) = 1 + a_2^* z^{-\beta_2} + o(z^{-\beta_2}),$$

where $\beta_2 = a$, and $a_2^* = \frac{1}{2} - \frac{5}{4}d_0$,

$$(iii) \quad E[\overline{X}\varphi_z(\overline{S}) \mid C_0 = z] = \mu_1(z)\varphi_z(\mu(z)) + a_3^* z^{-1} + o(z^{-1}),$$

where $a_3^* = \frac{1}{2}\sum_{i,j=1}^2 \sigma_{i,j}^* \{\frac{\partial^2 x_1\varphi^*(x)}{\partial x_i \partial x_j} \mid_{x=\mu^*}\} = \frac{9}{8}d_0^2 - \frac{9}{10}d_0$,

$$(iv) \quad E[(\overline{X}\varphi_z(\overline{S}) - \mu_1(z)\varphi_z(\mu(z)))^2 \mid C_0 = z] = a_4^* z^{-1} + o(z^{-1}),$$

$$\begin{aligned} \text{where} \quad a_4^* &= \textstyle\sum_{i,j=1}^2 \sigma_{i,j}^* \{\frac{\partial x_1\varphi^*(x)}{\partial x_i} \mid_{x=\mu^*} \frac{\partial x_1\varphi^*(x)}{\partial x_j} \mid_{x=\mu^*}\} \\ &= \tfrac{9}{4}d_0^2 - \tfrac{9}{5}d_0 + \tfrac{11}{25} \end{aligned}$$

$$(v) \quad \mu_1^*\nu^*(\mu^*) = \mu_1^*\varphi^*(\mu^*)(1 - \varphi^*(\mu^*)) = \tfrac{1}{5}. \tag{3.3.28}$$

From (2.3.28), (2.3.34) and the values of $a_1^*, \ldots a_4^*$, β_1 β_2 given above, we can calculate the values for the parameters β_0, b^* and c^* given in (2.3.28) and (2.3.34) as follows

$$\begin{aligned} \beta_0 &= \max\{0, 1-\beta_1, 1-\beta_2\} \\ &= \max\{0, 1-\min\{1,h\}, 1-a\} \\ &= \max\{0, 1-h, 1-a\}, \\ b^* &= \textstyle\sum_{i=1}^2 a_i^* \mathbf{1}_{\{\beta_0 = 1-\beta_i\}} + a_3^* \mathbf{1}_{\{\beta_0=0\}} \\ &= -\tfrac{5}{2}d_1 \mathbf{1}_{\{\beta_0=1-h\}} + (\tfrac{1}{2} - \tfrac{5}{4}d_0)\mathbf{1}_{\{\beta_0=1-a\}} \\ &\quad + (\tfrac{9}{8}d_0^2 - \tfrac{9}{10}d_0)\mathbf{1}_{\{\beta_0=0\}} \\ c^* &= a_4^* + \mu_1^*\nu^*(\mu^*) = \tfrac{9}{4}d_0^2 - \tfrac{9}{5}d_0 + \tfrac{16}{25}. \end{aligned} \tag{3.3.29}$$

Now, we treat this example as in Model 2.2 and we derive the values for the parameters $\tilde{m}$, $\tilde{a}_i^*$, $\tilde{\beta}_j$, $\tilde{b}^*$ and $\tilde{c}^*$, for $i = 1, \ldots, 7$, and $j = 0, \ldots, 3$.

The functions ψ_z, ψ^* and $\tilde{\nu}_2^*$ defined on $\mathcal{R}_+^2 \times \mathcal{R}_+^2$ are equal to

$$\begin{aligned} \psi_z(x,y) &= x_1 \varphi_z(\tfrac{1}{z}x + \tfrac{z-1}{z}y), \\ \psi^*(x,y) &= x_1 \varphi^*(y), \\ \tilde{\nu}_2^*(x,y) &= x_1 \varphi^*(y)(1 - \varphi^*(y)), \end{aligned} \tag{3.3.30}$$

where the function φ_z and φ^* are as in (3.3.24) and (3.3.25), respectively.

From (3.0.4) we have

$$\begin{aligned} \tilde{m} &= \lim_{z\to\infty} E[\psi(S_1, \mu^*) \mid C_0 = z] \\ &= \lim_{z\to\infty} E[X_1 \varphi^*(\mu^*) \mid C_0 = z] \\ &= \lim_{z\to\infty} \mu_1(z) \varphi^*(\mu^*) \\ &= \mu_1^* \varphi^*(\mu^*) = m = 1. \end{aligned} \tag{3.3.31}$$

The values of the parameters $\Delta_i^{(1)}$, $\Delta_{i,j}^{(1)}$ and $\Delta_i^{(2)}$ defined in (3.2.2) are equal to

$$\begin{aligned} &\lim_{z\to\infty} E[\tfrac{\partial \psi_z(S_1,y)}{\partial y_i} \mid_{y=\mu(z)} \mid C_0 = z] = \Delta_i^{(1)} = -d_0, \\ &\lim_{z\to\infty} E[\tfrac{\partial^2 \psi_z(S_1,y)}{\partial y_i \partial y_j} \mid_{y=\mu(z)} \mid C_0 = z] = \Delta_{i,j}^{(1)} = d_0^2, \\ &\lim_{z\to\infty} \mathrm{Cov}[(S_1)_i, \psi_z(S_1, \mu(z)) \mid C_0 = z] = \Delta_i^{(2)} = \tfrac{4}{5}\sigma_{1,i}^*, \end{aligned} \tag{3.3.32}$$

where $i, j = 1, 2$.

Then from (3.2.4) and some calculation we obtain

$$
\begin{aligned}
(i)\quad & E[\psi_z(S_1,\mu(z)) - \psi^*(S_1,\mu(z)) \mid C_0 = z] \\
& = \tilde{a}_1^* z^{-\tilde{\beta}_1} + o(z^{-\tilde{\beta}_1}), \text{ where } \tilde{\beta}_1 = \min\{1,h\}, \\
& \text{and } \tilde{a}_1^* = -\tfrac{9}{10} d_0 \mathbf{1}_{\{\tilde{\beta}_1 = 1\}} - \tfrac{5}{2} d_1 \mathbf{1}_{\{\tilde{\beta}_1 = h\}}, \\
(ii)\quad & E[\psi^*(S_1,\mu(z)) - \psi^*(S_1,\mu^*) \mid C_0 = z] \\
& = \tilde{a}_2^* z^{-\tilde{\beta}_2} + o(z^{-\tilde{\beta}_2}), \\
& \text{where } \tilde{\beta}_2 = a, \text{ and } \tilde{a}_2^* = -\tfrac{5}{4} d_0, \\
(iii)\quad & E[\psi^*(S_1,\mu^*) \mid C_0 = z] = 1 + \tilde{a}_3^* z^{-\tilde{\beta}_3} + o(z^{-\tilde{\beta}_3}), \\
& \text{where } \tilde{\beta}_3 = a, \text{ and } \tilde{a}_3^* = \tfrac{1}{2}, \\
(iv)\quad & E[\tilde{\nu}_2^*(S_1,\mu^*) \mid C_0 = z] = \tilde{a}_5^* + o(1), \\
& \text{where } \tilde{a}_5^* = \tfrac{1}{5}, \\
(v)\quad & E[\psi_z^2(S_1,\mu(z)) \mid C_0 = z] = 1 + \tilde{a}_6^* + o(1), \\
& \text{where } \tilde{a}_6^* = \tfrac{11}{25}.
\end{aligned}
\tag{3.3.33}
$$

Further, from (3.2.5), (3.2.14), (3.2.33) and (3.3.32) we have

$$
\begin{aligned}
(i)\quad & E[\psi_z(S_1,\overline{S}^{(1)}) - \psi_z(S_1,\mu(z)) \mid C_0 = z] \\
& = \tilde{a}_4^* z^{-1} + o(z^{-1}), \\
& \text{where} \quad \tilde{a}_4^* \quad = \tfrac{1}{2} \textstyle\sum_{i,j=1}^2 \Delta_{i,j}^{(1)} \sigma_{i,j}^* = \tfrac{9}{8} d_0^2 \\
(ii)\quad & \mathrm{Cov}[\psi_z(S_1,\overline{S}^{(1)}), \psi_z(S_2,\overline{S}^{(2)}) \mid C_0 = z] = \tilde{a}_7^* z^{-1} + o(z^{-1}), \\
& \text{where } \tilde{a}_7^* \quad = 2 \textstyle\sum_{i=1}^2 \Delta_i^{(1)} \Delta_i^{(2)} + \sum_{i,j=1}^2 \Delta_i^{(1)} \Delta_j^{(1)} \sigma_{i,j}^* \\
& \qquad\qquad = \tfrac{9}{4} d_0 - \tfrac{9}{5} d_0.
\end{aligned}
\tag{3.3.34}
$$

Now, from (3.2.37) and (3.2.44), together with the values of the parameters

$\tilde{a}_i^*$, $\tilde{\beta}_j$, for $i = 1, \ldots, 7$ and $j = 1, 2, 3$, given above, we obtain

$$
\begin{aligned}
\tilde{\beta}_0 &= \max\{0, 1 - \tilde{\beta}_i : i = 1, 2, 3\}, \\
&= \max\{0, 1 - \min\{1, h\}, 1 - a, 1 - a\} \\
&= \max\{0, 1 - h, 1 - a\}, \\
\tilde{b}^* &= \textstyle\sum_{i=1}^3 \tilde{a}_i^* \mathbf{1}_{\{\tilde{\beta}_0 = 1 - \tilde{\beta}_i\}} + \tilde{a}_4^* \mathbf{1}_{\{\tilde{\beta}_0 = 0\}} \\
&= -\tfrac{5}{2} d_1 \mathbf{1}_{\{\tilde{\beta}_0 = 1 - h\}} + (\tfrac{1}{2} - \tfrac{5}{4} d_0) \mathbf{1}_{\{\tilde{\beta}_0 = 1 - a\}} \\
&\quad + (\tfrac{9}{8} d_0^2 - \tfrac{9}{10} d_0) \mathbf{1}_{\{\tilde{\beta}_0 = 0\}}.
\end{aligned}
\tag{3.3.35}
$$

Hence, from (3.3.27), (3.3.29), (3.3.31) and (3.3.35) we can see that $m = \tilde{m} = 1$, $\beta_0 = \tilde{\beta}_0$, $b^* = \tilde{b}^*$ and $c^* = \tilde{c}^*$.

Further, from (3.3.28)(i) and (iii) we have

$$
\begin{aligned}
E[\overline{X} \varphi_z(\overline{S}) \mid C_0 = z] &= \mu_1(z) \varphi^*(\mu(z)) \\
&\quad + a_1^* z^{-\beta_1} + a_3^* z^{-1} + o(z^{-(\beta_1 \wedge 1)}).
\end{aligned}
\tag{3.3.36}
$$

Similarly, from (3.3.33)(i) and (3.3.34)(i) we have

$$
\begin{aligned}
E[\psi_z(S_1, \overline{S}^{(1)}) \mid C_0 = z] &= E[\psi^*(S_1, \mu(z)) \mid C_0 = z] \\
&\quad + \tilde{a}_1^* z^{-\tilde{\beta}_1} + \tilde{a}_4^* z^{-1} + o(z^{-(\tilde{\beta}_1 \wedge 1)}).
\end{aligned}
\tag{3.3.37}
$$

On the other hand it is easy to see that $E[\overline{X} \varphi_z(\overline{S}) \mid C_0 = z]$ is equal to $E[\psi_z(S_1, \overline{S}^{(1)}) \mid C_0 = z]$ and $\mu_1(z) \varphi^*(\mu(z))$ is equal to $E[\psi^*(S_1, \mu(z)) \mid C_0 = z]$. Hence, with $\beta_0 = \tilde{\beta}_0 = \max\{0, 1 - a, 1 - h\}$, we have

$$
a_1^* \mathbf{1}_{\{\beta_0 = 1 - \beta_1\}} + a_3^* \mathbf{1}_{\{\beta_0 = 0\}} = \tilde{a}_1^* \mathbf{1}_{\{\tilde{\beta}_0 = 1 - \tilde{\beta}_1\}} + \tilde{a}_4^* \mathbf{1}_{\{\tilde{\beta}_0 = 0\}}.
\tag{3.3.38}
$$

In the same way, from (3.3.28)(ii) and also from (3.3.33)(ii) and (iii) we have

$$
\begin{aligned}
&\mu_1(z) \varphi^*(\mu(z)) = 1 + a_2^* z^{-\beta_2} + o(z^{-\beta_2}), \\
&\text{and} \\
&E[\psi^*(S_1, \mu(z)) \mid C_0 = z] = \tilde{a}_2^* z^{-\tilde{\beta}_2} + \tilde{a}_3^* z^{-\tilde{\beta}_3} + o(z^{-(\tilde{\beta}_2 \wedge \tilde{\beta}_3)}).
\end{aligned}
\tag{3.3.39}
$$

As we mentioned before, $\mu_1(z)\varphi^*(\mu(z))$ is equal to $E[\psi^*(S_1, \mu(z)) \mid C_0 = z]$ which leads to the following

$$a_2^* \mathbf{1}_{\{\beta_0 = 1 - \beta_2\}} = \tilde{a}_2^* \mathbf{1}_{\{\tilde{\beta}_0 = 1 - \tilde{\beta}_2\}} + \tilde{a}_3^* \mathbf{1}_{\{\tilde{\beta}_0 = 1 - \tilde{\beta}_3\}}. \tag{3.3.40}$$

In a similar way, from (3.3.28)(iv), (3.3.33)(v) and (3.3.34)(ii) we have

$$a_4^* = \tilde{a}_6^* + \tilde{a}_7^*, \tag{3.3.41}$$

and from (3.3.28)(v) and (3.3.33)(iv) we have $\mu_1^* \nu^*(\mu^*) = \tilde{a}_5^*$. □

Chapter 4

The Coupling Method in Bisexual Populations

In this chapter we study a simple two-sex population where females are the reproductive sex and they reproduce children of both sexes. As in Chapter 2, we assume that individuals live only one unit of time (non-overlapping generations). In this chapter we assume that given the numbers of females and males in each generation, each female reproduces independently of the other females, according to a reproduction law which could depend only on the total number of males in that generation. We are interested in the case where males in a generation improve the reproductivity of the females in that generation. We studied the asymptotic behaviour of this model for a supercritical case. A successful coupling was found between this process and a Bisexual Galton-Watson Branching Process (B-GWBP for short) with promiscuous mating function. In other words, we have shown, under some assumptions, that there is a random time which is almost surely finite, after which these two processes will coincide. Therefore, they have the same limiting behaviour (on the set of non-extinction), and hence we can benefit from the knowledge available for B-GWBP with promiscuous mating function, to describe the asymptotic behaviour of our process.

In Section 4.1 we give the necessary notation as well as the model. Notice that the notations of this chapter are totally independent of those of Chapter 2. In Section 4.2 we present our main results. Section 4.3 provides us with some examples, in particular, Example 4.3 where we see the connection between this model and Model 2.1.

4.1. The Model

We start this section with the definition of the model which will be the subject of our study during the present chapter.

Model 4.1. (i) *Assume that both males and females live only one unit of time,* i.e. *all males and females live during the generation that they were born into.*

(ii) *Let (F_n, M_n) denote the number of females and males in the nth generation, for $n = 0, 1, 2, \ldots$. Define the σ-algebra $\mathcal{F}_n$ as*

$$\mathcal{F}_n := \sigma(F_0, M_0, \ldots, F_n, M_n). \tag{4.1.1}$$

Then given $\mathcal{F}_n$, the jth female in the nth generation produces $S_{n+1,j} := (X_{n+1,j}, Y_{n+1,j})$ number of daughters and sons, independently of all the other females in that generation. We assume that the conditional distribution of $S_{n+1,j}$, given $\mathcal{F}_n$, depends on the past only through the size of the male population in the present generation, i.e. *M_n. Thus given $\mathcal{F}_n$, $\{S_{n+1,j}\}_{j=1}^{F_n}$ are* i.i.d. *random vectors with the common distribution, say, G_{M_n}.*

(iii) *With G_m given as in (ii), we assume that $G_0 = \mathbf{1}_{\mathcal{Z}_+^2}$ and $\{G_m\}_{m=0}^{\infty}$ is a decreasing sequence of probability distributions on $\mathcal{Z}_+^2$,* i.e. *for all $x \in \mathcal{Z}_+^2$ we have*

$$G_m(x) \geq G_{m+1}(x) \tag{4.1.2}$$

and further there exists a distribution G^* on $\mathcal{Z}_+^2$ such that G_m converges to G^* in the total variation norm, i.e.

$$\sup_{A \subset \mathcal{Z}_+^2} \left| \int_A dG_m(x) - \int_A dG^*(x) \right| \to 0, \quad \text{as } m \to \infty. \tag{4.1.3}$$

(iv) *The population sizes in the $(n+1)$th generation are equal to*

$$(F_{n+1}, M_{n+1}) = \sum_{j=1}^{F_n} S_{n+1,j} = \sum_{j=1}^{F_n} (X_{n+1,j}, Y_{n+1,j}). \tag{4.1.4}$$

Let $\{G_m^*\}_{m=0}^{\infty}$ and $\{G_m^{**}\}_{m=0}^{\infty}$ be two sequences of probability distribution functions on $\mathcal{Z}_+^2$,

$$G_m^* = \begin{cases} G_0, & \text{if } m = 0, \\ G^*, & \text{if } m > 0, \end{cases} \quad \text{and } G_m^{**} = \begin{cases} G_0, & \text{if } m = 0, \\ G_1, & \text{if } m > 0. \end{cases} \tag{4.1.5}$$

With this construction, we can easily see that $G_m^{**}(x) \geq G_m(x) \geq G_m^*(x)$ for all $m \in \mathcal{Z}_+$ and all $x \in \mathcal{Z}_+^2$. Hence (4.1.5) gives an upper and lower bound for $\{G_m\}_{m=0}^{\infty}$, the set of possible reproduction laws of our bisexual model. When the population size increases, i.e. (F_n, M_n) goes to infinity, the reproduction law G_{M_n} converges to G^* by (4.1.3). Hence it is reasonable to expect that under some conditions, the limit behaviour of the population when it grows, would be similar to the behaviour of a population governed by the law G^* from the beginning. To investigate this idea more precisely, we need some notation.

Let $\preceq$ denote a partial ordering on $\mathcal{E} := \mathcal{R}^2$ such that $(x, y) \preceq (x^{'}, y^{'})$ if and only if $x \leq x^{'}$ and $y \leq y^{'}$ and let $\Delta := \{(s, s^{'}) \in \mathcal{E}^2 : s \preceq s^{'}\}$ be a subset of $\mathcal{E}^2$. Then using *Strassen's* theorem (see Lindvall [31]) for any $m^{'} \geq m \geq m^{''} \geq 0$, there exists a coupling $\hat{P}^*_{m,m^{'}}$ of G_m and $G^*_{m^{'}}$ and a coupling $\hat{P}^{**}_{m^{''},m}$ of $G^{**}_{m^{''}}$ and G_m such that $\hat{P}^*_{m,m^{'}}(\Delta) = \hat{P}^{**}_{m^{''},m}(\Delta) = 1$. In other words, when two random vectors, say ξ and η, are stochastically ordered, i.e. $\xi \overset{\mathrm{D}}{\leq} \eta$, *Strassen's* theorem gives the existence of two random vectors , say, $\hat{\xi}$ and $\hat{\eta}$, on the same probability space such that $\hat{\xi} \overset{\mathrm{D}}{=} \xi$ and $\hat{\eta} \overset{\mathrm{D}}{=} \eta$ and further $\hat{\xi} \preceq \hat{\eta}$ holds almost surely.

In the following, we shall define five processes on $\mathcal{Z}^2_+$, namely $\{(\hat{F}_n, \hat{M}_n)\}$, $\{(\hat{F}^*_n, \hat{M}^*_n)\}$, $\{(\hat{F}^{**}_n, \hat{M}^{**}_n)\}$, $\{(\tilde{F}^*_n, \tilde{M}^*_n)\}$, and $\{(\tilde{F}^{**}_n, \tilde{M}^{**}_n)\}$, recursively. First we define the σ-algebras $\hat{\mathcal{F}}^*_n$, $\hat{\mathcal{F}}^{**}_n$ and $\tilde{\mathcal{F}}^*_n$ as

$$
\begin{aligned}
\hat{\mathcal{F}}^*_n &:= \sigma(\{\hat{F}_k, \hat{M}_k, \hat{F}^*_k, \hat{M}^*_k\}_{k=0}^n) \\
\hat{\mathcal{F}}^{**}_n &:= \sigma(\{\hat{F}_k, \hat{M}_k, \hat{F}^{**}_k, \hat{M}^{**}_k, \tilde{F}^{**}_k, \tilde{M}^{**}_k\}_{k=0}^n), \qquad (4.1.6) \\
\tilde{\mathcal{F}}^*_n &:= \sigma(\{\tilde{F}^*_k, \tilde{M}^*_k, \hat{F}_k, \hat{M}_k\}_{k=0}^n).
\end{aligned}
$$

Now let all the above five processes take the same value at the zeroth generation

$$(\hat{F}_0, \hat{M}_0) = (\hat{F}^*_0, \hat{M}^*_0) = (\hat{F}^{**}_0, \hat{M}^{**}_0) = (\tilde{F}^*_0, \tilde{M}^*_0) = (\tilde{F}^{**}_0, \tilde{M}^{**}_0). \qquad (4.1.7)$$

Then our processes are defined for the nth generation $(n > 0)$ in two steps as follows:

Step I: Given $\hat{\mathcal{F}}^*_{n-1}$, $\hat{\mathcal{F}}^{**}_{n-1}$ and $\tilde{\mathcal{F}}^*_{n-1}$, let $\{\hat{S}_{n,j} = (\hat{X}_{n,j}, \hat{Y}_{n,j})\}_{j=1}^{\hat{F}_{n-1}}$, $\{\hat{S}^*_{n,j} = (\hat{X}^*_{n,j}, \hat{Y}^*_{n,j})\}_{j=1}^{\hat{F}_{n-1}}$, $\{\hat{S}^{**}_{n,j} = (\hat{X}^{**}_{n,j}, \hat{Y}^{**}_{n,j})\}_{j=1}^{\hat{F}^{**}_{n-1}}$, $\{\tilde{S}^*_{n,j} = (\tilde{X}^*_{n,j}, \tilde{Y}^*_{n,j})\}_{j=1}^{\tilde{F}^*_{n-1}}$, and $\{\tilde{S}^{**}_{n,j} = (\tilde{X}^{**}_{n,j}, \tilde{Y}^{**}_{n,j})\}_{j=1}^{\tilde{F}^*_{n-1}}$ be sequences of conditionally independent random vectors with the following conditional joint distributions

$$\begin{aligned}
&(\hat{S}_{n,j}, \hat{S}^*_{n,j}) \mid \hat{\mathcal{F}}^*_{n-1} \sim \hat{P}^*_{\hat{M}_{n-1}, \hat{M}^*_{n-1}},\\
&(\hat{S}^{**}_{n,j}, \hat{S}_{n,j}, \tilde{S}^{**}_{n,j}) \mid \hat{\mathcal{F}}^{**}_{n-1} \sim \hat{P}^{**}_{\hat{M}_{n-1}, \hat{M}_{n-1}} \times G_1, \\
&(\tilde{S}^*_{n,j}, \hat{S}_{n,j}) \mid \tilde{\mathcal{F}}^*_{n-1} \sim \hat{P}^{**}_{\tilde{M}^*_{n-1}, \hat{M}_{n-1}}.
\end{aligned} \tag{4.1.8}$$

Step II: In the second step, by using the random vectors defined in the first step, we can write down our processes as

$$\begin{aligned}
&(\mathbf{A}) \quad (\hat{F}_n, \hat{M}_n) = \textstyle\sum_{j=1}^{\hat{F}_{n-1}} \hat{S}_{n,j},\\
&(\mathbf{B}) \quad (\hat{F}^*_n, \hat{M}^*_n) = \textstyle\sum_{j=1}^{\hat{F}_{n-1}} \hat{S}^*_{n,j},\\
&(\mathbf{C}) \quad (\hat{F}^{**}_n, \hat{M}^{**}_n) = \textstyle\sum_{j=1}^{\hat{F}^{**}_{n-1}} \hat{S}^{**}_{n,j},\\
&(\mathbf{D}) \quad (\tilde{F}^*_n, \tilde{M}^*_n) = \textstyle\sum_{j=1}^{\tilde{F}^*_{n-1}} \tilde{S}^*_{n,j},\\
&(\mathbf{E}) \quad (\tilde{F}^{**}_n, \tilde{M}^{**}_n) = \begin{cases} (\hat{F}^{**}_n, \hat{M}^{**}_n), & \text{if } \hat{M}_{n-1} > 0, \\ \sum_{j=1}^{\tilde{F}^{**}_{n-1}} \tilde{S}^{**}_{n,j}, & \text{if } \hat{M}_{n-1} = 0. \end{cases}
\end{aligned} \tag{4.1.9}$$

Remark 4.1. **(i)** First notice that by using (4.1.8) we have the following marginal distributions

$$\begin{aligned}
&\hat{S}_{n,1} \mid \hat{\mathcal{F}}^*_{n-1} \sim G_{\hat{M}_{n-1}},\\
&\hat{S}^*_{n,1} \mid \hat{\mathcal{F}}^*_{n-1} \sim G^*_{\hat{M}^*_{n-1}},\\
&\hat{S}^{**}_{n,1} \mid \hat{\mathcal{F}}^{**}_{n-1} \sim G^{**}_{\hat{M}_{n-1}},\\
&\tilde{S}^*_{n,1} \mid \tilde{\mathcal{F}}^*_{n-1} \sim G^{**}_{\tilde{M}^*_{n-1}},\\
&\tilde{S}^{**}_{n,1} \mid \hat{\mathcal{F}}^{**}_{n-1} \sim G_1.
\end{aligned} \tag{4.1.10}$$

And further we can see from (4.1.8) that $\{\tilde{S}^{**}_{n,j}\}_{n,j=1}^{\infty}$ and $\{(\hat{S}^{**}_{n,j}, \hat{S}_{n,j})\}_{n,j=1}^{\infty}$ are independent of each other. We come back to this in the part (ii) below.

(ii) Secondly, from (4.1.2), (4.1.5), (4.1.8) and (4.1.9), one can easily see that the following inequalities hold

$$
\begin{aligned}
(\hat{F}_n, \hat{M}_n) &\overset{\text{a.s.}}{\preceq} (\hat{F}^*_n, \hat{M}^*_n) \\
(\hat{F}^{**}_n, \hat{M}^{**}_n) &\overset{\text{a.s.}}{\preceq} (\hat{F}_n, \hat{M}_n) \qquad (4.1.11) \\
(\tilde{F}^*_n, \tilde{M}^*_n) &\overset{\text{a.s.}}{\preceq} (\hat{F}_n, \hat{M}_n).
\end{aligned}
$$

As we pointed out in (i) $\{\hat{S}^{**}_{n,j}\}^{\infty}_{n,j=1}$ and $\{\tilde{S}^{**}_{n,j}\}^{\infty}_{n,j=1}$ are independent of each other. From (4.1.9)**(E)** one can see that the $\{\tilde{F}^{**}_n\}^{\infty}_{n=0}$ is a standard Galton-Watson branching process. If we let $T = \max\{n : \hat{M}_n > 0\}$, then $\hat{F}^{**}_n = \tilde{F}^{**}_n$ for $0 \leq n \leq T$ and $\hat{F}^{**}_n \leq \tilde{F}^{**}_n$ for $n > T$. In other words, these two processes coincide at least up to a random time T, and after this time the process $\{\hat{F}^{**}_n\}$ dies out and the process $\{\tilde{F}^{**}_n\}$ may grow to infinity. Hence we have $\hat{F}^{**}_n \leq \tilde{F}^{**}_n$ a.s. for all $n \geq 0$.

(iii) In this work, it is enough for us to define the joined distributions of these random vectors as we did in (4.1.8). However, it is possible to define all these random vectors on a common probability space.

□

It is the coupling between the two processes given in (4.1.9)**(A)**-**(B)** which is our primary interest. To be more precise, we want to show that with probability one, the two processes $\{\hat{F}_n\}$ and $\{\hat{F}^*_n\}$ coincide after a finite random time. If we denote the random time from which these two processes coincide with τ. Then we have

$$\tau := \inf\{k : \hat{F}_n = \hat{F}^*_n \text{ for all } n \geq k\}. \qquad (4.1.12)$$

Assume for a moment that $P(\tau < \infty) = 1$, and let us have a closer look on the process $\{(\hat{F}^*_n, \hat{M}^*_n)\}^{\infty}_{n=\tau}$. Let ζ be a non-negative, integer valued function on $\mathcal{Z}^2_+$ such that $\zeta(x, y) = x \min\{y, 1\}$, for $(x, y) \in \mathcal{Z}^2_+$. If for some $n \geq \tau$, $\hat{M}^*_n = 0$ then we have $\zeta(\hat{F}^*_n, \hat{M}^*_n) = 0$. However, from (4.1.8) and (4.1.9) we have $\hat{M}^*_n = 0$ implies $\hat{F}^*_{n+1} = 0$. Therefore, $\zeta(\hat{F}^*_k, \hat{M}^*_k) = 0 = \hat{F}^*_k$, for all $k > n \geq \tau$. If for all $n \geq \tau$, $\hat{M}^*_n > 0$ then we have $\zeta(\hat{F}^*_n, \hat{M}^*_n) = \hat{F}^*_n$ for all $n \geq \tau$. This shows that $\{\hat{F}^*_n\}_{n \geq \tau}$ has the same limiting behaviour as $\{\zeta(\hat{F}^*_n, \hat{M}^*_n)\}_{n \geq \tau}$. On the other hand, after random time τ we have

$$(\hat{F}^*_{n+1}, \hat{M}^*_{n+1}) = \sum_{j=1}^{\hat{F}^*_n} \hat{S}^*_{n+1,j}, \qquad (4.1.13)$$

when $\hat{M}_n^* > 0$ for $n \geq \tau$. And $(\hat{F}_{n+1}^*, \hat{M}_{n+1}^*) = (0,0)$ if $\hat{M}_{n-1}^* = 0$ for $n > \tau$. Hence, if we show that $P(\tau < \infty) = 1$, then the process $\{\zeta(\hat{F}_n^*, \hat{M}_n^*)\}_{n\geq\tau}$ will be the number of mating units in a B-GWBP with a promiscuous mating function, with reproduction distribution G^* and a random number of initial mating units equal to $\hat{F}_\tau^*$.

Therefore, showing that $\tau < \infty$ with probability one, is equivalent to saying that the two processes $\{\hat{F}_n\}$ and $\{\hat{F}_n^*\}$ have the same limiting behavior. This will be investigated in the next section. However, to be able to obtain the limiting behaviour of the process $\{\hat{F}_n^*\}_{n\geq\tau}$, we need to know the distribution of $\hat{F}_\tau^*$. In Chapter 5, we present an explicit coupling construction for a class of distributions on $\mathcal{Z}_+^2$, namely the case where the number of daughters and sons is independent.

Before closing this section, we need to define some parameters for later use. Define the conditional mean vectors of the random vectors $\hat{S}_{1,1}$, $\hat{S}_{1,1}^*$ and $\hat{S}_{1,1}^{**}$ by $\mu(m) = E[\hat{S}_{1,1} \mid \hat{M}_0 = m]$, $\mu^*(m) = E[\hat{S}_{1,1}^* \mid \hat{M}_0^* = m]$ and $\mu^{**}(m) = E[\hat{S}_{1,1}^{**} \mid \hat{M}_0 = m]$, respectively. Then

$$\mu(0) = \mu^*(0) = \mu^{**}(0) = (0,0), \tag{4.1.14}$$

and for $m > 0$

$$\begin{aligned} \mu(m) &= (\mu_1(m), \mu_2(m)) = E[\hat{S}_{1,1} \mid \hat{M}_0 = m], \\ \mu^*(m) &= (\mu_1^*, \mu_2^*) = E[\hat{S}_{1,1}^* \mid \hat{M}_0^* = m], \\ \mu^{**}(m) &= (\mu_1^{**}, \mu_2^{**}) = E[\hat{S}_{1,1}^{**} \mid \hat{M}_0 = m]. \end{aligned} \tag{4.1.15}$$

From (4.1.15) we can see that for all $m > 0$, $\mu^*(m)$ and $\mu^{**}(m)$ are fixed values. This can be seen from (4.1.8) and (4.1.5), where we have $\hat{S}_{1,1}^* \mid \{\hat{M}_0^* = m\} \sim G^*$ and $\hat{S}_{1,1}^{**} \mid \{\hat{M}_0 = m\} \sim G_1$ for $m > 0$.

Finally, let $(\hat{X}^{**}, \hat{Y}^{**}) \sim G_1$ and f be the probability generating function of $\hat{X}^{**}$, i.e.

$$f(s) := E[s^{\hat{X}^{**}}] \tag{4.1.16}$$

for $0 \leq s \leq 1$. Let q be smallest non-negative real number such that

$$q = f(q), \quad \text{and} \quad \text{also} \quad \gamma := f'(q). \tag{4.1.17}$$

These parameters play an important role for us. q is the extinction probability for a standard Galton-Watson branching process with probability generating function f. And $q < 1$ corresponds to a supercritical standard Galton-Watson process.

4.2. The Results

We start this section by presenting Theorem 4.1, which shows that under some assumptions, our bisexual population, presented in the previous section, can grow to infinity with a positive probability. Theorems 4.7-4.8 study the growth of the population under different assumptions, in other words, these theorems give almost sure finiteness of coupling time τ (for a definition see (4.1.12)) under various conditions.

Theorem 4.1. *Assume that in (4.1.15) we have* $1 < \mu_1(m) < \infty$ *and* $\mu_2(m) > 0$ *for all* $m > 0$ *and let* $\{(\hat{F}_n, \hat{M}_n)\}_{n=0}^{\infty}$ *be as in (4.1.9)***(A)***. Then*

$$P(\inf_n \hat{F}_n = 0) < 1. \tag{4.2.1}$$

Proof of Theorem 4.1: Let $\zeta(x,y) = x\min\{y,1\}$, for $(x,y) \in \mathcal{Z}_+^2$. From (4.1.11) we have $(\tilde{F}_n^*, \tilde{M}_n^*) \overset{\text{a.s.}}{\preceq} (\hat{F}_n, \hat{M}_n)$ and this implies that $\zeta(\tilde{F}_n^*, \tilde{M}_n^*) \overset{\text{a.s.}}{\leq} \hat{F}_n$. Therefore to prove the theorem, we only need to show that $\{\zeta(\tilde{F}_n^*, \tilde{M}_n^*)\}$ is supercritical B-GWBP. Define the mean growth rate of the process $\{\zeta(\tilde{F}_n^*, \tilde{M}_n^*)\}$ to be equal to

$$r_k := \frac{1}{k}E[\zeta(\tilde{F}_{n+1}^*, \tilde{M}_{n+1}^*) \mid \zeta(\tilde{F}_n^*, \tilde{M}_n^*) = k], \tag{4.2.2}$$

for $k > 0$. The function ζ is superadditive, i.e. $\zeta(x_1 + x_2, y_1 + y_2) \geq \zeta(x_1, y_1) + \zeta(x_2, y_2)$ for all $x_i, y_i \in \mathcal{Z}_+$, and for $i = 1, 2$. To show that $\{\zeta(\tilde{F}_n^*, \tilde{M}_n^*)\}$ is a supercritical B-GWBP, by a theorem of [7] it is enough to show that $\sup_k r_k > 1$. But r_k can be written as

$$\begin{aligned} r_k &= \tfrac{1}{k}E[\zeta(\tilde{F}_{n+1}^*, \tilde{M}_{n+1}^*) \mid \zeta(\tilde{F}_n^*, \tilde{M}_n^*) = k] \\ &= \tfrac{1}{k}(E[\tilde{F}_{n+1}^* \mid \zeta(\tilde{F}_n^*, \tilde{M}_n^*) = k] \\ &-E[\tilde{F}_{n+1}^*; \tilde{M}_{n+1}^* = 0 \mid \zeta(\tilde{F}_n^*, \tilde{M}_n^*) = k]) \end{aligned} \tag{4.2.3}$$

Given $\zeta(\tilde{F}_n^*, \tilde{M}_n^*) = \tilde{F}_n^* = k > 0$, by (4.1.8) $\{(\tilde{X}_{n+1,j}^*, \tilde{Y}_{n+1,j}^*)\}_{j=1}^k$ and $(\tilde{X}^*, \tilde{Y}^*)$ are i.i.d. random vectors with the common distribution G_1. And further, since $\tilde{F}_{n+1}^* = \sum_{j=1}^{\tilde{F}_n^*} \tilde{X}_{n+1,j}^*$ and by symmetry we can write (4.2.3) as

$$\begin{aligned} r_k &= \tfrac{1}{k}(k\mu_1^* - kE[\tilde{X}_{n+1,1}^*; \cap_{j=1}^k \{\tilde{Y}_{n+1,j}^* = 0\} \mid \zeta(\tilde{F}_n^*, \tilde{M}_n^*) = k]) \\ &= \mu_1^* - E[\tilde{X}^*; \{\tilde{Y}^* = 0\}](P(\tilde{Y}^* = 0))^{k-1}. \end{aligned} \tag{4.2.4}$$

By (4.1.5) we have $\mu_1^* = \mu_1(1) > 1$. Thus $\sup_{k>0} r_k = \mu_1^* > 1$ hence there exists a q_0, $0 \le q_0 < 1$, such that

$$q_0 = P(\inf_n \zeta(\tilde{F}_n^*, \tilde{M}_n^*) = 0). \tag{4.2.5}$$

On the other hand by (4.1.5)-(4.1.9) we have

$$\begin{aligned} \{\inf_n \hat{F}_n = 0\} &= \{\inf_n \zeta(\hat{F}_n, \hat{M}_n) = 0\} \\ &\subset \{\inf_n \zeta(\tilde{F}_n^*, \tilde{M}_n^*) = 0\}. \end{aligned} \tag{4.2.6}$$

Now by (4.2.5) and (4.2.6) we have

$$\begin{aligned} 1 > q_0 = P(\inf_n \zeta(\tilde{F}_n^*, \tilde{M}_n^*) = 0) &\ge P(\inf_n \zeta(\hat{F}_n, \hat{M}_n) = 0) \\ &= P(\inf_n \hat{F}_n = 0). \end{aligned} \tag{4.2.7}$$

■

One can easily see that our bisexual model fits in the framework of the two-sex model studied by Asmussen and Hering [3], Chapter XI, section 2. In their work, the main concern is about the asymptotic growth of their process. They show that, under suitable assumptions, their process either becomes extinct or grows to infinity with the probability one (see Theorem 2.1. of Chapter XI of [3]). However, they only gave a sufficient condition for extinction of their process (see Proposition 2.4. of Chapter XI of [3]), but not for the survival of their process. Therefore, with our bisexual model as a special case of their process, Theorem 4.1 provides us with a sufficient condition for the survival of our process.

We need to present some lemmas before we are able to present Theorem 4.7. The first lemma is due to Badalbaev and Muhitdinov. For the proof see [5], page 16-21.

Lemma 4.2. (Badalbaev and Muhitdinov, 1990) *Let $\{Z_n\}$ be a Galton-Watson process with the probability generating function h, the extinction probability a, $b := h^{'}(1) > 1$ (the mean number of offspring), $h^{''}(1) < \infty$ and $c := h^{'}(a) < 1/b$. Then as $n \to \infty$,*

$$E[1/Z_n; Z_n \ge 1] \sim b^{-n} \int_0^\infty E[e^{-tw}; w > 0]\, dt, \tag{4.2.8}$$

where $w := \lim_{n\to\infty} Z_n/b^n$ and the integral converges.

Remark 4.2. The probability generation function h of the number of offspring of an individual satisfies three conditions; $b = h^{'}(1) > 1$ which implies the supercriticality of the Galton-Watson process, $h^{''}(1) < \infty$ implies that the distribution of the number of offspring of an individual should have a finite variance and the last condition is $bc = h^{'}(a)h^{'}(1) < 1$ does not have a clear interpretation.

□

Lemma 4.3 below, prepares us for Lemma 4.4 in which we obtain an upper bound for the expectation of an inverse of sum of $\mathcal{Z}_+$-valued, i.i.d. random variables. We shall make use of Lemma 4.4 during the proofs of the theorems below.

Lemma 4.3. *If U has a binomial distribution with parameters n and p, then for $j = 1, 2, \ldots, n$ we have*

$$E[\prod_{k=1}^{j} \frac{1}{k+U}] = \frac{n!}{(n+j)!p^j}(1 - \sum_{k=0}^{j-1} \binom{n+j}{k} p^k (1-p)^{n+j-k}). \quad (4.2.9)$$

Proof of Lemma 4.3: Since $\prod_{k=1}^{j} \frac{1}{i+k} = \frac{i!}{(i+j)!}$, the left hand side of (4.2.9) is equal to

$$E[\prod_{k=1}^{j} \tfrac{1}{k+U}] = \sum_{i=0}^{n} \prod_{k=1}^{j} \tfrac{1}{k+i} \binom{n}{i} \cdot p^i (1-p)^{n-i}$$

$$= \sum_{i=0}^{n} \tfrac{i!}{(i+j)!} \tfrac{n!}{i!(n-i)!} \cdot p^i (1-p)^{n-i}$$

$$= \tfrac{n!}{(n+j)!p^j} \sum_{i=0}^{n} \binom{n+j}{j+i} \cdot p^{i+j} \cdot (1-p)^{n-i} \quad (4.2.10)$$

$$= \tfrac{n!}{(n+j)!p^j} \sum_{i=j}^{n+j} \binom{n+j}{i} \cdot p^i \cdot (1-p)^{n+j-i}$$

$$= \tfrac{n!}{(n+j)!p^j} (1 - \sum_{i=0}^{j-1} \binom{n+j}{i} p^i (1-p)^{n+j-i}).$$

■

Lemma 4.4. *Assume that $\{\xi_j\}_{j=1}^{\infty}$ is a sequence of* i.i.d. *random variables taking values in the set $\mathcal{Z}_+$. Then for all n, $n = 1, 2, \ldots$ and $k = 1, \ldots, n$ there exists a constant K_k such that*

$$E[\frac{1}{(\sum_{j=1}^{n} \xi_j)^k}; \{\sum_{j=1}^{n} \xi_j > 0\}] < K_k \prod_{i=1}^{k} \frac{1}{n+i}. \quad (4.2.11)$$

Proof of Lemma 4.4: Let $Y_j = \mathbf{1}_{\{\xi_j>0\}}$ and $p = P(\xi_j > 0) > 0$. Then

$$E[\tfrac{1}{(\sum_{j=1}^n \xi_j)^k}; \{\sum_{j=1}^n \xi_j > 0\}] \le E[\tfrac{1}{(\sum_{j=1}^n Y_j)^k}; \{\sum_{j=1}^n Y_j \ge 1\}]. \quad (4.2.12)$$

But for $U \ge 1$, we have $\frac{1}{U} \le \frac{k+1}{k+U}$. Hence (4.2.12) can be written as

$$\begin{aligned} &E[\tfrac{1}{(\sum_{j=1}^n \xi_j)^k}; \{\sum_{j=1}^n \xi_j > 0\}] \le E[\textstyle\prod_{i=1}^k \frac{i+1}{i+\sum_{j=1}^n Y_j}] \\ &\le K_k \textstyle\prod_{i=1}^k \frac{n!}{(n+k)!} = K_k \prod_{i=1}^k \frac{1}{(n+i)}, \end{aligned} \quad (4.2.13)$$

where $K_k = \frac{(k+1)!}{p^k}$. In the last inequality we have used Lemma 4.3.

■

With $\hat{F}_n^{**}$ and $\tilde{F}_n^{**}$ defined in (4.1.9)**(C)** and **(E)**, respectively, and since $\hat{F}_n^{**} = \tilde{F}_n^{**}$ whenever $\hat{F}_n^{**} > 0$, we have the following lemma.

Lemma 4.5. *If h is a non-negative real-valued function such that $h(0) = 0$. Then with $\{\hat{F}_n^{**}\}$ and $\{\tilde{F}_n^{**}\}$ given in (4.1.9), respectively, we have*

$$E[h(\hat{F}_n^{**})] \le E[h(\tilde{F}_n^{**})]. \quad (4.2.14)$$

Proof of Lemma 4.5: Define the random time $T := \min\{n : \hat{M}_n = 0\}$, where the process $\{\hat{M}_n\}$ is as in (4.1.9)**(A)**. In terms of random time T one can write the left hand side of (4.2.14) as

$$\begin{aligned} E[h(\hat{F}_n^{**})] &= E[h(\hat{F}_n^{**}); \{n-1 \le T\}] \\ &= E[h(\tilde{F}_n^{**}); \{n-1 \le T\}] \le E[h(\tilde{F}_n^{**})], \end{aligned} \quad (4.2.15)$$

since by our construction in (4.1.9) $\hat{F}_n^{**} = \tilde{F}_n^{**}$ on the set $\{n-1 \le T\}$. And also $\hat{F}_n^{**} = \hat{F}_n = 0$ and $\tilde{F}_n^{**} \ge 0$ on the set $\{n-1 > T\}$. ■

Let μ_1^*, μ_2^{**} and $\mu_1(\cdot)$ as in (4.1.15), and $a_0 := \min\{1, \mu_2^{**}\}$. Then for any fixed $u \in \mathcal{R}$, let g_u be a non-negative and non-increasing real-valued function on $[a_0, \infty)$ such that

$$g_u(m) = \begin{cases} m^u(\mu_1^* - \mu_1(m)), & \text{if } m = 1, 2, \dots, \\ 0, & \text{if } m = 0. \end{cases} \quad (4.2.16)$$

Let $g_u^\diamond$ be corresponding function to g_u, given in Lemma 2.11, defined on the interval $[a_0, \infty)$. Define the function $\tilde{g}_u^\diamond$ on $\mathcal{R}_+$ as

$$\tilde{g}_u^\diamond(x) = g_u^\diamond(x)\mathbf{1}_{[a_0,\infty)}(x), \quad (4.2.17)$$

for $x \ge 0$.

Then we have the following lemma.

Lemma 4.6. *Let the processes* $\{(\hat{F}_n, \hat{M}_n)\}$, $\{(\hat{F}_n^{**}, \hat{M}_n^{**})\}$ *and* $\{(\tilde{F}_n^{**}, \tilde{M}_n^{**})\}$ *be as in (4.1.9), and* (μ_1^{**}, μ_2^{**}), f *and* γ *be as in (4.1.15), (4.1.16) and (4.1.17), respectively. Assume that* $1 < \mu_1^{**} < \infty$, $\mu_2^{**} > 0$, $f^{''}(1) < \infty$ *and* $\mu_1^{**}\gamma < 1$. *For some fixed* $u \geq 0$, *let* g_u *and* $\tilde{g}_u^{\diamond}$ *be as in (4.2.16) and (4.2.17), respectively. Then if* $\sum_{n=1}^{\infty} g_u(n)/n < \infty$ *then we have*

$$\sum_{n=1}^{\infty} E[\tilde{g}_u^{\diamond}(\mu_2^{**}\hat{F}_n)] < \infty. \tag{4.2.18}$$

Proof of Lemma 4.6: First notice that since $\tilde{g}_u^{\diamond}$ is a non-increasing function and it satisfies the condition of Lemma 4.5 and further by (4.1.11) we have $\hat{F}_n \geq \hat{F}_n^{**}$ almost surely. We have

$$\begin{aligned} E[\tilde{g}_u^{\diamond}(\mu_2^{**}\hat{F}_n)] &\leq E[\tilde{g}_u^{\diamond}(\mu_2^{**}\hat{F}_n^{**})] \\ &\leq E[\tilde{g}_u^{\diamond}(\mu_2^{**}\tilde{F}_n^{**})]. \end{aligned} \tag{4.2.19}$$

Hence to show the assertion of the lemma it is enough to show

$$\sum_{n=1}^{\infty} E[\tilde{g}_u^{\diamond}(\mu_2^{**}\tilde{F}_n^{**})] < \infty. \tag{4.2.20}$$

By the assumption $\mu_1^{**}\gamma < 1$. Since $\mu_1^{**} > 1$, by Remark 4.1, $\{\tilde{F}_n^{**}\}$ is a supercritical standard Galton-Watson branching process. Hence by Lemma 4.2 we have $E[\tilde{F}_n^{**}]E[(1/\tilde{F}_n^{**}); \tilde{F}_n^{**} > 0] < K$, for some $K > 0$. By using this and non-increasingness of $\tilde{g}_u^{\diamond}$, we can write

$$\begin{aligned} &\textstyle\sum_n^{\infty} E[\tilde{g}_u^{\diamond}(\mu_2^{**}\tilde{F}_{n-1}^{**})] \leq \sum_n^{\infty} E[\tilde{g}_u^{\diamond}(a_0\tilde{F}_{n-1}^{**})] \\ &= \textstyle\sum_n E[(1/\tilde{F}_n^{**})\tilde{F}_n^{**}\tilde{g}_u^{\diamond}(a_0\tilde{F}_{n-1}^{**}); \tilde{F}_n^{**} \geq 1] \end{aligned} \tag{4.2.21}$$

Now, since by Lemma 2.11 the function $x\tilde{g}_u^{\diamond}(x)$ is non-decreasing and concave for $x \geq a_0$. Thus (4.2.21) can be written as

$$\begin{aligned} &\textstyle\sum_n^{\infty} E[\tilde{g}_u^{\diamond}(\mu_2^{**}\tilde{F}_{n-1}^{**})] \\ &\leq \textstyle\sum_n E[(1/\tilde{F}_n^{**}); \tilde{F}_n^{**} \geq 1]E[\tilde{F}_n^{**}\tilde{g}_u^{\diamond}(a_0\tilde{F}_{n-1}^{**})] \\ &\leq \textstyle\sum_n E[(1/\tilde{F}_n^{**}); \tilde{F}_n^{**} \geq 1]E[\tilde{F}_n^{**}]\tilde{g}_u^{\diamond}(a_0E[\tilde{F}_{n-1}^{**}]) \end{aligned} \tag{4.2.22}$$

Now by Lemma 4.2, $E[(1/\tilde{F}_n^{**}); \tilde{F}_n^{**} \geq 1]E[\tilde{F}_n^{**}]$ is bounded and further since $\{\tilde{F}_n^{**}\}$ is a standard Galton-Watson branching process, we have $E[\tilde{F}_n^{**}] = \hat{F}_0(\mu_1^{**})^n$ (with $\hat{F}_0$ fixed by assumption). Hence (4.2.22) can be written as

$$\begin{aligned} &\textstyle\sum_n^{\infty} E[\tilde{g}_u^{\diamond}(\mu_2^{**}\tilde{F}_{n-1}^{**})] \leq K\sum_n \tilde{g}_u^{\diamond}(a_0\hat{F}_0(\mu_1^{**})^{n-1}) \\ &\leq K\textstyle\sum_{n=1}^{\infty} g_u^{\diamond}(a_0(\mu_1^{**})^{n-1}) \end{aligned} \tag{4.2.23}$$

Now, since by the assumptions of the lemma, $\mu_1^{**} > 1$, then by Lemma 2.10, $\sum_{n=1}^{\infty} g_u^{\diamond}(a_0(\mu_1^{**})^{n-1}) < \infty$ if and only if $\sum_{n=1}^{\infty} g_u^{\diamond}(n)/n < \infty$. By Lemma 2.11, $\sum_{n=1}^{\infty} g_u^{\diamond}(n)/n < \infty$ if $\sum_{n=1}^{\infty} g_u(n)/n < \infty$. On the other hand, by the assumptions of the lemma $\sum_{n=1}^{\infty} g_u(n)/n < \infty$. Hence, this implies that $\sum_{n=1}^{\infty} g_u^{\diamond}(a_0(\mu_1^{**})^{n-1}) < \infty$ and we have proved the lemma. ■

Theorem 4.7. *Make the following assumptions*

(i) *With μ_1^{**}, μ_2^{**}, f and γ as in (4.1.15), (4.1.16) and (4.1.17), respectively, assume that $\mu_1^{**} > 1$, $\mu_2^{**} > 0$, $f^{''}(1) < \infty$ and $\mu_1^{**}\gamma < 1$.*

(ii) *With μ_1^* and $\mu_1(\cdot)$ as in (4.1.15), assume that $\{m(\mu_1^* - \mu_1(m))\}_{m=1}^{\infty}$ is a non-negative and decreasing sequence such that*

$$\sum_{m=1}^{\infty} (\mu_1^* - \mu_1(m)) < \infty. \qquad (4.2.24)$$

Then we have a successful coupling, i.e. $P(\tau < \infty) = 1$, *with τ as in (4.1.12).*

Remark 4.3. For some comments on the assumption (i) see Remark 4.2. Assumption (ii) of the theorem gives information about the speed of convergence of $\mu_1(m)$ to μ_1^* as m (the total number of males) tends to infinity. Notice that the assumption (ii) considers only the convergence of the sum of the differences of the expected numbers of daughters born into each family when we have abounded number of males around, i.e. μ_1^*, to the case where the number of males are finite, i.e. $\mu_1(m)$ for $m = 1, 2, \ldots$. □

Proof of Theorem 4.7: From (4.1.12) we have $\tau := \min\{k \geq 1 : \hat{F}_n^* = \hat{F}_n$ for all $n \geq k\}$. Define the random time τ^* as following:

$$\tau^* := \min\{k : \hat{F}_n^* \mathbf{1}_{\{\hat{M}_{n-1} \geq 1\}} = \hat{F}_n \mathbf{1}_{\{\hat{M}_{n-1} \geq 1\}} \text{ for all } n \geq k\}. \qquad (4.2.25)$$

Then from the construction of the processes $\{(\hat{F}_n, \hat{M}_n)\}$ and $\{(\hat{F}_n^*, \hat{M}_n^*)\}$ given in (4.1.9), one can see that

$$\{\hat{F}_n^* > \hat{F}_n \text{ i.o.}\} \subset \{\hat{M}_n \geq 1 \text{ for all } n \geq 0\}. \qquad (4.2.26)$$

This implies that $\{\tau = \infty\} \equiv \{\tau^* = \infty\}$. By the Borel-Cantelli lemma we have

$$\begin{aligned} P(\tau = \infty) = 0 \quad &\Longleftrightarrow P(\tau^* = \infty) = 0 \\ &\Longleftrightarrow P(\limsup_n \{\hat{F}_n^* > \hat{F}_n, \hat{M}_{n-1} \geq 1\}) = 0 \qquad (4.2.27) \\ &\Longleftarrow \textstyle\sum_n P(\hat{F}_n^* > \hat{F}_n, \hat{M}_{n-1} \geq 1) < \infty. \end{aligned}$$

Notice by (4.1.11) $\hat{F}^*_n \overset{\text{a.s.}}{\geq} \hat{F}_n$. Hence to prove the theorem it is enough to show that the sum in (4.2.27) is finite. Using the Markov inequality

$$\begin{aligned}
&\textstyle\sum_n P(\hat{F}^*_{n+1} > \hat{F}_{n+1} \ \text{ and } \ \hat{M}_n \geq 1) \\
&\leq \textstyle\sum_n E[E[\hat{F}^*_{n+1} - \hat{F}_{n+1} \mid \hat{\mathcal{F}}^*_n]; \hat{M}_n \geq 1] \\
&= \textstyle\sum_n E[\hat{F}_n(\mu_1^* \mathbf{1}_{\{\hat{M}^*_n \geq 1\}} - \mu_1(\hat{M}_n)); \hat{M}_n \geq 1]
\end{aligned} \tag{4.2.28}$$

Since by (4.1.11) $\hat{M}^*_n \overset{\text{a.s.}}{\geq} \hat{M}_n$, $\mu_1^* \mathbf{1}_{\{\hat{M}^*_n \geq 1\}} = \mu_1^*$ on the set $\{\hat{M}_n \geq 1\}$ as in (4.1.15). We can write (4.2.28) as

$$\textstyle\sum_n P(\hat{F}^*_{n+1} > \hat{F}_{n+1}, \hat{M}_n \geq 1) \leq \sum_n E[\hat{F}_n(\mu_1^* - \mu_1(\hat{M}_n)); \hat{M}_n \geq 1]. \tag{4.2.29}$$

Let g_1 be as in (4.2.16), $g_1(m) = m(\mu_1^* - \mu_1(m))$ for $m = 1, 2, \ldots$ and further by Assumption (ii) of the theorem, $\{g_1(m)\}_{m=1}^\infty$ is a non-negative and non-increasing sequence. By symmetry we can write (4.2.29) as

$$\begin{aligned}
&\textstyle\sum_n P(\hat{F}^*_{n+1} > \hat{F}_{n+1}, \hat{M}_n \geq 1) \\
&\leq \textstyle\sum_n E[E[\sum_{j=1}^{\hat{F}_{n-1}} \hat{X}_{n,j} \frac{g_1(\hat{M}_n)}{\hat{M}_n}; \hat{M}_n \geq 1 \mid \hat{\mathcal{F}}^*_{n-1}]] \\
&= \textstyle\sum_n E[\hat{F}_{n-1} E[\hat{X}_{n,1} \frac{g_1(\hat{M}_n)}{\hat{M}_n}; \hat{M}_n \geq 1 \mid \hat{\mathcal{F}}^*_{n-1}]]
\end{aligned} \tag{4.2.30}$$

Let $\hat{Z}_n := \hat{M}_n - \hat{Y}_{n,1}$. Since $x \geq 1$ is equivalent to $\frac{2}{x+1} \geq \frac{1}{x}$ and also $g_1(x)/(x+1)$ is non-increasing, we can write (4.2.30) as

$$\begin{aligned}
&\textstyle\sum_n P(\hat{F}^*_{n+1} > \hat{F}_{n+1}, \hat{M}_n \geq 1) \\
&\leq \textstyle\sum_n E[\hat{F}_{n-1} E[\hat{X}_{n,1} \frac{2g_1(\hat{M}_n)}{\hat{M}_n + 1}; \hat{M}_n \geq 1 \mid \hat{\mathcal{F}}^*_{n-1}]] \\
&\leq 2 \textstyle\sum_n E[\hat{F}_{n-1} E[\hat{X}_{n,1} \frac{g_1(\hat{Z}_n)}{\hat{Z}_n + 1}; \hat{M}_n \geq 1 \mid \hat{\mathcal{F}}^*_{n-1}]] \\
&\leq 2 \textstyle\sum_n E[\hat{F}_{n-1} E[\hat{X}_{n,1} \frac{g_1(\hat{Z}_n)}{\hat{Z}_n + 1} \mid \hat{\mathcal{F}}^*_{n-1}]].
\end{aligned} \tag{4.2.31}$$

Now since given $\hat{\mathcal{F}}^*_{n-1}$, $\hat{X}_{n,1}$ and $\hat{Z}_n$ are independent of each other and we have $\mu_1(m) \le \mu_1^*$. We can write (4.2.31) as

$$\begin{aligned}
&\textstyle\sum_n P(\hat{F}^*_{n+1} > \hat{F}_{n+1}, \hat{M}_n \ge 1)\\
&\le 2\textstyle\sum_n E[\hat{F}_{n-1} E[\hat{X}_{n,1} \mid \hat{\mathcal{F}}^*_{n-1}] E[\frac{g_1(\hat{Z}_n)}{\hat{Z}_n+1} \mid \hat{\mathcal{F}}^*_{n-1}]]\\
&= 2\textstyle\sum_n E[\hat{F}_{n-1} \mu_1(\hat{M}_n) E[\frac{g_1(\hat{Z}_n)}{\hat{Z}_n+1} \mid \hat{\mathcal{F}}^*_{n-1}]]\\
&\le 2\mu_1^* E[\hat{F}_{n-1} E[\frac{g_1(\hat{Z}_n)}{\hat{Z}_n+1} \mid \hat{\mathcal{F}}^*_{n-1}]].
\end{aligned} \tag{4.2.32}$$

Now let $\tilde{g}_1^\diamond$ be the function corresponding to g_1 as in (4.2.17). Then, $x\tilde{g}_1^\diamond(x)$ is a non-decreasing function and further given $\hat{\mathcal{F}}^*_{n-1}$, $\frac{1}{\hat{Z}_n(\hat{Z}_n+1)}\mathbf{1}_{\{\hat{Z}_n \ge 1\}}$ and $\hat{Z}_n\tilde{g}_1^\diamond(\hat{Z}_n)$ have a negative covariance. Thus we can write (4.2.32) as

$$\begin{aligned}
&\textstyle\sum_n P(\hat{F}^*_{n+1} > \hat{F}_{n+1}, \hat{M}_n \ge 1)\\
&\le 2\mu_1^* E[\hat{F}_{n-1} E[\frac{\tilde{g}_1^\diamond(\hat{Z}_n)}{\hat{Z}_n+1} \mid \hat{\mathcal{F}}^*_{n-1}]]\\
&\le 2\mu_1^* E\left[\hat{F}_{n-1} E[\frac{\hat{Z}_n\tilde{g}_1^\diamond(\hat{Z}_n)}{\hat{Z}_n(\hat{Z}_n+1)}; \hat{Z}_n \ge 1 \mid \hat{\mathcal{F}}^*_{n-1}]\right]\\
&\le 2\mu_1^* E\left[\hat{F}_{n-1} E[\frac{1}{\hat{Z}_n(\hat{Z}_n+1)}; \hat{Z}_n \ge 1 \mid \hat{\mathcal{F}}^*_{n-1}]\right.\\
&\left.\cdot\ E[\hat{Z}_n\tilde{g}_1^\diamond(\hat{Z}_n) \mid \hat{\mathcal{F}}^*_{n-1}]\right]\\
&\le 2\mu_1^* E\left[\hat{F}_{n-1}\ \cdot E[\frac{1}{\hat{Z}_n(\hat{Z}_n+1)}; \hat{Z}_n \ge 1 \mid \hat{\mathcal{F}}^*_{n-1}]\right.\\
&\left.\cdot E[\hat{M}_n\tilde{g}_1^\diamond(\hat{M}_n) \mid \hat{\mathcal{F}}^*_{n-1}]\right].
\end{aligned} \tag{4.2.33}$$

Now by Lemma 4.4, for some constant $K > 0$, we have

$$E[\frac{1}{\hat{Z}_n(\hat{Z}_n+1)}; \hat{Z}_n \ge 1 \mid \hat{\mathcal{F}}^*_{n-1}] \le \frac{K}{\hat{F}_{n-1}(\hat{F}_{n-1}+1)}. \tag{4.2.34}$$

By concavity of $x\tilde{g}_1^\diamond(x)$, non-increasingness of $\tilde{g}_1^\diamond(x)$ and also since $\mu_2(m) \ge \mu_2^{**}$ and $\mu_2(m) \le \mu_2^*$, by (4.1.11) for all $m > 0$, we can write

$$\begin{aligned}
&E[\hat{M}_n\tilde{g}_1^\diamond(\hat{M}_n) \mid \hat{\mathcal{F}}^*_{n-1}] \le E[\hat{M}_n \mid \hat{\mathcal{F}}^*_{n-1}]\tilde{g}_1^\diamond(E[\hat{M}_n \mid \hat{\mathcal{F}}^*_{n-1}])\\
&\le \mu_2(\hat{M}_{n-1})\hat{F}_{n-1}\tilde{g}_1^\diamond(\mu_2(\hat{M}_{n-1})\hat{F}_{n-1}) \le \mu_2^*\hat{F}_{n-1}\tilde{g}_1^\diamond(\mu_2^{**}\hat{F}_{n-1}).
\end{aligned} \tag{4.2.35}$$

Now by replacing (4.2.34) and (4.2.35) in (4.2.33) we obtain

$$
\begin{aligned}
&\textstyle\sum_n P(\hat{F}^*_{n+1} > \hat{F}_{n+1}, \hat{M}_n \geq 1) \\
&\leq K_1 \textstyle\sum_n E[\frac{\hat{F}^2_{n-1}}{\hat{F}_{n-1}(\hat{F}_{n-1}+1)} \tilde{g}^\diamond_1(\mu^{**}_2 \hat{F}_{n-1}); \hat{F}_{n-1} \geq 1] \qquad (4.2.36)\\
&\leq K_1 \textstyle\sum_n E[\tilde{g}^\diamond_1(\mu^{**}_2 \hat{F}_{n-1})].
\end{aligned}
$$

By Lemma 4.6, the sum $\sum_n E[\tilde{g}^\diamond_1(\mu^{**}_2 \hat{F}_{n-1})]$ is finite and hence we have a successful coupling. ■

With $\mu_1(\cdot)$, μ^*_1, and μ^{**}_1 as in (4.1.15), define the function ρ on $\mathcal{Z}_+$ as follows.

$$
\rho(x) := \begin{cases} \mu^*_1 - \mu_1(x), & \text{if } x = 1, 2, \ldots, \\ \\ 1, & \text{if } x = 0. \end{cases} \tag{4.2.37}
$$

This function is equal to the function g_u with $u = 0$ on $\mathcal{Z}_+$ except at the point zero. In the next and last theorem of this chapter we show that we have a successful coupling without assuming that $\mu^{**}_1 > 1$. This means that at the start, the process $\{(\hat{F}_n, \hat{M}_n)\}$ could be a subcritical process. However, as we see from the assertion of Theorem 4.8 below, this requires that as m (the size of male sub-population) increases, $\mu_1(m)$ converges very rapidly (exponentially fast) to $\mu^*_1 > 1$.

Theorem 4.8. *Assume the following:*

(i) *With μ^{**}_1, μ^*_1 and μ^{**}_2 given in (4.1.15). Assume that $\mu^{**}_1 \neq 1$, $\mu^*_1 > 1$ and $\mu^{**}_2 > 0$.*

(ii) *Assume that, ρ given in (4.2.37) is a sub-multiplicative function on $\mathcal{Z}_+$,* i.e. *for $x, y \in \mathcal{Z}_+$ we have $\rho(x + y) \leq \rho(x)\rho(y)$, with $\rho(1) < 1$.*

(iii) *With $(\hat{X}^{**}, \hat{Y}^{**}) \sim G_1$, and G_1 as in Model 4.1 and let $\delta := E[\rho(\hat{Y}^{**})]$. Assume that $\delta \leq e^{-1}$.*

Then we have a successful coupling, i.e. *$P(\tau < \infty) = 1$, with τ as in (4.1.12).*

Proof of Theorem 4.8: Let the random time τ^* be defined as

$$
\tau^* := \min\{k \geq 1 : \hat{F}^*_n \mathbf{1}_{\{\hat{M}_{n-1} \geq 1\}} = \hat{F}_n \mathbf{1}_{\{\hat{M}_{n-1} \geq 1\}} \text{ for all } n \geq k\}. \tag{4.2.38}
$$

Like in (4.2.25)-(4.2.28) we can see that $\{\tau = \infty\} \equiv \{\tau^* = \infty\}$ and hence

$$
P(\tau^* = \infty) = 0 \quad \Longleftarrow \textstyle\sum_n E[\hat{F}_n(\mu^*_1 - \mu_1(\hat{M}_n)); \hat{M}_{n-1} \geq 1] < \infty. \tag{4.2.39}
$$

Now by symmetry we can write the last sum in (4.2.39) as

$$\begin{aligned} &\textstyle\sum_n E[\hat{F}_n(\mu_1^* - \mu_1(\hat{M}_n)); \hat{M}_{n-1} \geq 1] \\ &\leq \mu_1^* E[\hat{F}_{n-1} E[\hat{X}_{n,1}\rho(\hat{M}_n) \mid \hat{\mathcal{F}}_{n-1}^*]; \hat{M}_{n-1} \geq 1] \end{aligned} \tag{4.2.40}$$

By the sub-multiplicative property of ρ we can write (4.2.40) as

$$\begin{aligned} &\textstyle\sum_n E[\hat{F}_n(\mu_1^* - \mu_1(\hat{M}_n)); \hat{M}_{n-1} \geq 1] \\ &\leq \mu_1^* \textstyle\sum_n E[\hat{F}_{n-1} E[\hat{X}_{n,1}\rho(\sum_{k=1}^{\hat{F}_{n-1}} \hat{Y}_{n,k}) \mid \hat{\mathcal{F}}_{n-1}^*]; \hat{M}_{n-1} \geq 1] \\ &\leq \mu_1^* \textstyle\sum_n E[\hat{F}_{n-1} E[\hat{X}_{n,1} \prod_{k=1}^{\hat{F}_{n-1}} \rho(\hat{Y}_{n,k}) \mid \hat{\mathcal{F}}_{n-1}^*]; \hat{M}_{n-1} \geq 1]. \end{aligned} \tag{4.2.41}$$

By conditional independence of $\hat{Y}_{n,1}$ from $\{\hat{Y}_{n,k}\}_{k=2}^{\hat{F}_{n-1}}$ and boundedness of the function ρ, we can write (4.2.41) as

$$\begin{aligned} &\textstyle\sum_n E[\hat{F}_n(\mu_1^* - \mu_1(\hat{M}_n)); \hat{M}_{n-1} \geq 1] \\ &\leq \mu_1^* \textstyle\sum_n E[\hat{F}_{n-1} E[\hat{X}_{n,1}\rho(\hat{Y}_{n,1}) \mid \hat{\mathcal{F}}_{n-1}^*] \\ &\textstyle\prod_{k=2}^{\hat{F}_{n-1}} E[\rho(\hat{Y}_{n,k}) \mid \hat{\mathcal{F}}_{n-1}^*]; \hat{M}_{n-1} \geq 1] \\ &\leq K_1 \textstyle\sum_n E[\hat{F}_{n-1}(E[\rho(\hat{Y}_{n,1}) \mid \hat{\mathcal{F}}_{n-1}^*])^{\hat{F}_{n-1}-1}; \hat{M}_{n-1} \geq 1]. \end{aligned} \tag{4.2.42}$$

With δ as in Assumption (iii), let us define the function $\varphi(x) := x\delta^{x-1}$ which is a decreasing function on the interval $(0, \infty)$ and also $\varphi(0) = 0$. Since by (4.1.8) we have $\hat{Y}_{n,1}^{**} \overset{\text{a.s.}}{\leq} \hat{Y}_{n,1}$, hence by using these we can write (4.2.42) as

$$\begin{aligned} &\textstyle\sum_n E[\hat{F}_n(\mu_1^* - \mu_1(\hat{M}_n)); \hat{M}_{n-1} \geq 1] \\ &\leq K_2 E[\hat{F}_{n-1}(E[\rho(\hat{Y}_{n,1}^{**}) \mid \hat{\mathcal{F}}_{n-1}^*])^{\hat{F}_{n-1}-1}; \hat{M}_{n-1} \geq 1] \\ &\leq K_2 E[\hat{F}_{n-1}\delta^{\hat{F}_{n-1}-1}] = K_2 \textstyle\sum_n E[\varphi(\hat{F}_{n-1}); \hat{M}_{n-1} \geq 1]. \end{aligned} \tag{4.2.43}$$

With the function f as in (4.1.16), let $f_0(s) := s$ and $f_{n+1}(s) := f(f_n(s))$, for $n \geq 0$ and $0 \leq s \leq 1$, be the probability generating function of a standard Galton-Watson branching process $\{\hat{F}^{**}\}$ (with reproduction generating function f). Then by Lemma 2.11 we have

$$
\begin{aligned}
&\textstyle\sum_n E[\hat{F}_n(\mu_1^* - \mu_1(\hat{M}_n)); \hat{M}_{n-1} \geq 1] \\
&\leq K_2 \textstyle\sum_n E[\varphi(\hat{F}_{n-1}^{**}); \hat{M}_{n-1} \geq 1] \qquad (4.2.44)\\
&\leq K_2 \textstyle\sum_n E[\varphi(\tilde{F}_{n-1}^{**})] = K_2 \sum_n f_n^{'}(\delta)
\end{aligned}
$$

With γ as in (4.1.17), we have $\gamma = f^{'}(q) < 1$. To see this, first notice that f is strictly increasing function on the interval $[0,1]$ and secondly $0 \leq q < 1$ is the only solution to the equation $f(s) = s$ for $0 \leq s \leq 1$. Therefore, we have $\gamma = f^{'}(q) < 1$. Hence $\frac{1}{\gamma^n} f_n^{'}(\delta)$ converges to a finite limit as $n \to \infty$ (see [4], page 38) and therefore it is bounded. Hence (4.2.44) will be bounded by

$$\textstyle\sum_n E[\hat{F}_n(\mu_1^* - \mu_1(\hat{M}_n)); \hat{M}_{n-1} \geq 1] \leq K_3 \sum_n \gamma^n < \infty. \qquad (4.2.45)$$

Therefore, we have a successful coupling. ■

Remark 4.4. One can ask the following question. Why can not we think of a bisexual population as a two-type population size dependent Galton-Watson branching process as defined and studied in Klebaner [28]? Consider Model 4.1 and define the matrices $\mathbf{M}(\cdot,\cdot)$ and $\mathbf{M}^*$ as following.

$$
\begin{aligned}
\mathbf{M}(x,y) &= \begin{bmatrix} E[\hat{X}_{1,1} \mid \hat{F}_0 = x,\ \hat{M}_0 = y] & E[\hat{Y}_{1,1} \mid \hat{F}_0 = x,\ \hat{M}_0 = y] \\ 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} \mu_1(y)\mathbf{1}_{\{x>0\}} & \mu_2(y)\mathbf{1}_{\{x>0\}} \\ 0 & 0 \end{bmatrix},
\end{aligned}
\qquad (4.2.46)
$$

for $x, y \in \mathcal{Z}_+$ and

$$\mathbf{M}^* = \begin{bmatrix} \mu_1^* & \mu_2^* \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} E[\hat{X}_{1,1}^*] & E[\hat{Y}_{1,1}^*] \\ 0 & 0 \end{bmatrix}. \qquad (4.2.47)$$

If the females and the males in a bisexual population are two types, the first and the second rows of the matrix $\mathbf{M}(\cdot,\cdot)$ give the expected number of children of each sex born from a female and a male, respectively, when there are x females and y males in the population. Since, the males do not reproduce, the second row is equal to zero. There are similar interpretations for the matrix $\mathbf{M}^*$. Now we can write $\mathbf{M}(\cdot,\cdot)$ as in the following form

$$\mathbf{M}(x,y) = \mathbf{M}^* + \mathbf{D}(x,y), \qquad (4.2.48)$$

for $x, y = 0, 1, 2, \ldots$. Theorem 1 of Klebaner [28] gives the asymptotic behaviour of a multitype population size dependent Galton-Watson branching process, under the assumption that there exists a non-negative, non-increasing function d on $\mathcal{Z}_+$, such that

$$\| \mathbf{D}(x, y) \| \leq d(x + y) \text{ and } \sum_n d(n)/n < \infty. \tag{4.2.49}$$

Hence, to use the results of [28] for a bisexual population, there should exist a non-negative, non-increasing function of total population size, i.e. $x + y$, such that this function satisfies (4.2.49). However, from (4.2.46) one can see that $\| \mathbf{D}(x, y) \|$ is only a function of $(\mathbf{1}_{\{x>0\}}, y)$, and further $\mathbf{M}(x, y) \to \mathbf{M}^*$ as $y\mathbf{1}_{\{x>0\}} \to \infty$. One can see from this and (4.2.46) that on one hand for any $\epsilon > 0$ we can take y large enough and $x > 0$, such that $\| \mathbf{D}(x, y) \| < \epsilon$ and on the other hand $\mathbf{D}(0, y) = -\mathbf{M}^*$. Thus, having $x + y$ large, does not imply that $\| \mathbf{D}(x, y) \|$ is small and therefore there does not exist a non-increasing function d of total population size which satisfies (4.2.49). Hence, this assumption (and a similar assumption in Theorem 2 of [28]) can not hold and therefore we can not use the results for the multitype population size dependent Galton-Watson branching processes to study Model 4.1.

For a general bisexual model, if the decomposition (4.2.48) holds then, since in a bisexual population there would not be any reproduction if at least one of the two sexes die out, this implies that the residual matrix $\mathbf{D}(x, y)$ in (4.2.48) would be such that $\mathbf{D}(0, y) = -\mathbf{M}^*$. Again, take $\epsilon > 0$ such that $\epsilon < \| \mathbf{M}^* \|$, and k large enough such that $\| \mathbf{D}(k, k) \| < \epsilon$. But, on the other hand we have $\| \mathbf{D}(0, 2k) \| = \| \mathbf{M}^* \| > \epsilon$. Hence, an arbitrary bisexual population can not be treated as a two type population size dependent Galton-Watson branching process. $\square$

4.3. Examples

In this section we present some examples for the B-GWBP model studied in Section 4.2. In the first and second examples presented below, we consider two simple cases where the conditions of Theorem 4.7 and Theorem 4.8 are fulfilled, respectively. The third example, presents a case where the method developed for the 2S-GWBP, given in Chapter 2 and 3, can not be used in that case.

Example 4.1. Let $\theta^{(1)}(\infty) > 2/3$ and $\theta^{(1)}(m) := \theta^{(1)}(\infty)(1 - \frac{1}{4m^2})\mathbf{1}_{\{m>0\}}$ for $m \in \mathcal{Z}_+$. Let $S_{n,j} = (X_{n,j}, Y_{n,j})$, the number of daughters and sons of the jth female in the $(n-1)$th generation which satisfies the following:

$$X_{n,j} \stackrel{\text{a.s.}}{=} 2Y_{n,j}, \quad \text{and}$$

$$Y_{n,j} \mid \mathcal{F}_{n-1} \sim \begin{cases} \text{Bernolli}(\theta^{(1)}(M_{n-1})), & \text{if } M_{n-1} > 0, \\ \\ \mathbf{1}_{\{0\}}, & \text{if } M_{n-1} = 0, \end{cases} \tag{4.3.1}$$

where $\mathcal{F}_{n-1}$ is the σ-algebra defined in (4.1.1). From (4.3.1) we can see the values for the vectors $\mu(m)$, $\mu^*(m)$ and $\mu^{**}(m)$, given in (4.1.14)-(4.1.15) are equal to

$$\begin{aligned} \mu(m) &= (2\theta^{(1)}(m), \theta^{(1)}(m)), \\ \mu^*(m) &= (\mu_1^*, \mu_2^*)\mathbf{1}_{\{m>0\}} = (2\theta^{(1)}(\infty), \theta^{(1)}(\infty))\mathbf{1}_{\{m>0\}}, \\ \mu^{**}(m) &= (\mu_1^{**}, \mu_2^{**})\mathbf{1}_{\{m>0\}} = (2\theta^{(1)}(1), \theta^{(1)}(1))\mathbf{1}_{\{m>0\}}, \end{aligned} \tag{4.3.2}$$

for $m \in \mathcal{Z}_+$. Hence we have $\mu_1^{**} = 2\theta^{(1)}(1) = \frac{3}{2}\theta^{(1)}(\infty) > 1$ and $\mu_2^{**} = \theta^{(1)}(1) > 0$. With f as in (4.1.16) we have

$$f(s) = (1 - \theta^{(1)}(1)) + s^2\theta^{(1)}(1), \tag{4.3.3}$$

for $0 \le s \le 1$. The smallest non-negative solution to the equation $f(s) = s$ is equal to $q := \frac{1-\theta^{(1)}(1)}{\theta^{(1)}(1)}$. We have $\gamma := f^{'}(q) = 2(1 - \theta^{(1)}(1))$. Hence we have $\mu_1^{**}\gamma = 4\theta^{(1)}(1)(1 - \theta^{(1)}(1)) < 1$. Thus the assumption (i) of Theorem 4.7 holds.

One can easily see that $\mu_1^* - \mu_1(m) = \frac{1}{4m^2}$. Hence the assumption (ii) of Theorem 4.7 holds. □

Example 4.2. Let $\theta^{(2)}(m) := (\theta^{(2)}(\infty) - e^{-m})\mathbf{1}_{\{m>0\}}$ for $m \in \mathcal{Z}_+$ where $e^{-1} < \theta^{(2)}(\infty) < 1/2$. Let $S_{n,j} = (X_{n,j}, Y_{n,j})$, the number of daughters and sons of the jth female in the $(n-1)$th generation which satisfies the following:

$$X_{n,j} \text{ and } Y_{n,j} \text{ are i.i.d., and}$$

$$Y_{n,j} \mid \mathcal{F}_{n-1} \sim \begin{cases} \text{Geo}(\theta^{(2)}(M_{n-1})), & \text{if } M_{n-1} > 0, \\ \\ \mathbf{1}_{\{0\}}, & \text{if } M_{n-1} = 0, \end{cases} \tag{4.3.4}$$

where $\mathcal{F}_{n-1}$ is the σ-algebra defined in (4.1.1). From (4.3.4) we can see the values for the vectors $\mu(m)$, $\mu^*(m)$ and $\mu^{**}(m)$, given in (4.1.14)-(4.1.15) are equal to

$$\mu(m) = (\tfrac{1-\theta^{(2)}(m)}{\theta^{(2)}(m)}, \tfrac{1-\theta^{(2)}(m)}{\theta^{(2)}(m)})\mathbf{1}_{\{m>0\}},$$

$$\mu^*(m) = (\mu_1^*, \mu_2^*)\mathbf{1}_{\{m>0\}} = (\tfrac{1-\theta^{(2)}(\infty)}{\theta^{(2)}(\infty)}, \tfrac{1-\theta^{(2)}(\infty)}{\theta^{(2)}(\infty)})\mathbf{1}_{\{m>0\}}, \tag{4.3.5}$$

$$\mu^{**}(m) = (\mu_1^{**}, \mu_2^{**})\mathbf{1}_{\{m>0\}} = (\tfrac{1-\theta^{(2)}(1)}{\theta^{(2)}(1)}, \tfrac{1-\theta^{(2)}(1)}{\theta^{(2)}(1)})\mathbf{1}_{\{m>0\}},$$

for $m \in \mathcal{Z}_+$. The function ρ defined in (4.2.37) is equal to $\rho(m) = e^{-m}$ for $m \in \mathcal{Z}_+$.

It is easy to see that $\mu_1^{**} \neq 1$, $\mu_1^* > 1$ and $\mu_2^{**} > 0$. Hence the assumption (i) of Theorem 4.8 holds.

Also we have $\delta := E[\rho(Y^{**})] = \frac{\theta^{(2)}(1)}{1-e^{-1}(1-\theta^{(2)}(1))}$ which is strictly less than e^{-1}. Hence the assumption (ii) of Theorem 4.8, and therefore the assertion of that theorem holds.

□

Examples 4.1 and 4.2 present cases where Example 4.1 does not satisfy the assumptions of Theorem 4.8 and Example 4.2 does not satisfy the assumptions of Theorem 4.7. To see this, with the function ρ defined as in (4.2.37), and $Y^{**} \sim$ Bernolli($\theta^{(1)}(1)$), we have $E[\rho(Y^{**})] = 1 - \theta^{(1)} + \frac{\theta^{(1)}(\infty)}{4}\theta^{(1)}(1) > e^{-1}$. Thus Assumption (ii) of Theorem 4.8 does not hold.

In the notation defined in (4.1.15)-(4.1.17), for Example 4.2, we have $f(s) = \frac{\theta^{(2)}(1)}{1-s(1-\theta^{(2)}(1))}$ which gives $q = \frac{\theta^{(2)}(1)}{1-\theta^{(2)}(1)}$. From (4.3.5), we have $\mu_1^{**} = \frac{1-\theta^{(2)}(1)}{\theta^{(2)}(1)}$, which gives $\gamma = \mu_1^{**}f'(q) = 1$. Thus, Assumption (i) of Theorem 4.7 which requires $\gamma < 1$ does not hold.

In the following example we present a case which fits in the framework of Model 2.1 as well as Model 4.1 which satisfies the assumptions of Model 4.1 but not Model 2.1.

Example 4.3. Let $0 \leq K_0 < 1$ and $K_1 > 0$ be fixed constants such that $K_0 + K_1\sum_{x=1}^{\infty} 1/x^4 = 1$ and let $\eta(y) := K_0\mathbf{1}_{\{0\}}(y) + K_1\frac{1}{y^4}\mathbf{1}_{(0,\infty)}(y)$ for $y \in \mathcal{Z}_+$. Assume that $0 < \theta^{(3)}(m) < 1$ for $m = 1, 2, \ldots$ and such that the sequence $\{\theta^{(3)}(m)\}_{m=1}^{\infty}$ is non-decreasing and further $\lim_{m\to\infty}\theta^{(3)}(m) = \theta^{(3)}(\infty) < 1$. Also assume that $\{m(\theta^{(3)}(\infty) - \theta^{(3)}(m))\}_{m=1}^{\infty}$ is a non-increasing sequence and further $\sum_{m=1}^{\infty}(\theta^{(3)}(\infty) - \theta^{(3)}(m)) < \infty$.

With $S_{n,j} = (X_{n,j}, Y_{n,j})$, the number of daughters and sons of the jth female in the $(n-1)$th generation, for $j = 1, 2, \ldots, F_{n-1}$, we determine the distribution of $S_{n,j} = (X_{n,j}, Y_{n,j})$ in two steps. On the first step we allow the jth female in $(n-1)$th generation to get married and reproduce independent of the other females and with probability $\theta^{(3)}(M_{n-1})$ or stay single and not reproduce with probability $1-\theta^{(3)}(M_{n-1})$. We show this by the outcome of a Bernolli random variable $R_{n,j}$, i.e. for given n, we have

$$\begin{aligned} &\{R_{n,j}\}_{j=1}^{F_{n-1}} \mid \mathcal{F}_{n-1} \text{ are i.i.d.} \\ &\text{and } R_{n,j} \mid \mathcal{F}_{n-1} \sim \text{Bernolli}(\theta^{(3)}(M_{n-1})), \end{aligned} \tag{4.3.6}$$

for $n = 1, 2, \ldots$. Given $\mathcal{F}_{n-1}$, let $U_{n,j}$ and $U'_{n,j}$ denote the hypothetical numbers of daughters and sons which the jth female in the nth generation would like to have. These numbers are conditionally independent of her marital status and they have the following probability density functions:

$$\text{Given } \mathcal{F}_{n-1} \; U_{n,j} \text{ and } U'_{n,j} \text{ are i.i.d., and}$$

$$P(U_{n,j} = u \mid \mathcal{F}_{n-1}) = \begin{cases} \eta(u), & \text{if } M_{n-1} > 0, \\ \mathbf{1}_{\{0\}}(u), & \text{if } M_{n-1} = 0, \end{cases} \tag{4.3.7}$$

for $u \in \mathcal{Z}_+$. Hence, the conditional distribution of the number of children of the jth female in the $(n-1)$th generation, i.e. $S_{n,j} = (X_{n,j}, Y_{n,j}) = R_{n,j}(U_{n,j}, U'_{n,j})$, for $j = 1, 2, \ldots, F_{n-1}$, are i.i.d. with the common distribution

$$P((X_{n,j}, Y_{n,j}) = (x, y) \mid \mathcal{F}_{n-1})$$

$$= \begin{cases} \eta(x)\eta(y)\theta^{(3)}(M_{n-1}), & \text{if } \max(x,y) > 0, \text{ and } M_{n-1} > 0, \\ 1 - \theta^{(3)}(M_{n-1}) + \eta^2(0)\theta^{(3)}(M_{n-1}), & \text{if } \max(x,y) = 0, \text{ and } M_{n-1} > 0, \\ \mathbf{1}_{\{0\}}(\max(x,y)), & \text{if } M_{n-1} = 0, \end{cases} \tag{4.3.8}$$

for $y \in \mathcal{Z}_+$.

Hence, in this example, the reproduction law of each female can be broken into two stages. These stages are similar to those of the 2S-GWBP studied in Chapter 2. If we let $K_0 = 0$, one can easily see that the distribution of $S_{n,j}$ satisfies the assumptions of Theorem 4.7. However, since a random variable with density function η does not have higher moments than the second, the methods developed for the 2S-GWBP (Model 2.1) can not be used to study its asymptotic properties.

□

Chapter 5

γ-Coupling On $\mathcal{Z}_+$

In Chapter 4, we studied the asymptotic behaviour of a bisexual population with the help of a coupling argument. To prepare the necessary tools for studying Model 4.1 in (4.1.8), we referred to Strassen's theorem (see Lindvall [31]) which gives the existence of a coupling between two random vectors on $\mathcal{Z}_+^2$. To be more clear, let $\xi^* = (\xi_1^*, \xi_2^*)$ and $\xi^{**} = (\xi_1^{**}, \xi_2^{**})$,be two random vectors on $\mathcal{Z}_+^2$ with probability measures $\mathbf{P}^*$ and $\mathbf{P}^{**}$, respectively, such that $\mathbf{P}^* \overset{\mathrm{D}}{\leq} \mathbf{P}^{**}$. Strassen's theorem gives the existence of a pair of random vectors, say $(\hat{\xi}^*, \hat{\xi}^{**})$, with joint probability measure $\mathbf{P}$ such that $\hat{\xi}^* = (\hat{\xi}_1^*, \hat{\xi}_2^*)$ and $\hat{\xi}^{**} = (\hat{\xi}_1^{**}, \hat{\xi}_2^{**})$ have marginal distributions as ξ^* and ξ^{**}, respectively. And further, $\mathbf{P}(\hat{\xi}_i^* \leq \hat{\xi}_i^{**}$, for $i = 1, 2)$=1. A special case occurs when components of random vectors $\xi^* = (\xi_1^*, \xi_2^*)$ and $\xi^{**} = (\xi_1^{**}, \xi_2^{**})$, are independent of each other. In this chapter, we study the construction of coupling which can be used in this situation.

This chapter is devoted to γ-coupling, where we present the construction of γ-coupling between two stochastically ordered, $\mathcal{Z}_+$-valued r.v.s such that the coupled r.v.s maintain the order almost surely. This coupling can be used for the case where the numbers of daughters and sons of a female are independent. The construction of a γ-coupling for the case where there is dependence between the number of daughters and sons is still an open question.

Let $E = \mathcal{Z}_+$, and Z and Z^* be two r.v.s on $(E, \mathcal{E})$ (where $\mathcal{E}$ is the set of all subsets of E), with corresponding probability measures $\mathbf{P}$ and $\mathbf{P}^*$, respectively. For convenience, denote the density functions of r.v.s Z and Z^* by $\mathbf{p}$ and $\mathbf{p}^*$, respectively. We say $Z \overset{\mathrm{D}}{\leq} Z^*$ or equivalently, $\mathbf{P} \overset{\mathrm{D}}{\leq} \mathbf{P}^*$, if

$$\mathbf{P}([x, \infty)) \leq \mathbf{P}^*([x, \infty)) \text{ for all } x \in E. \tag{5.0.1}$$

Let (U, U^*) be any coupling of r.v.s Z and Z^*, i.e. (U, U^*) is a pair of r.v.s defined on the space $(E^2, \mathcal{E}^2)$ such that U and U^* have the marginal laws $\mathbf{P}$ and $\mathbf{P}^*$, respectively. Then by the (basic) coupling inequality (see [31], page 12), we have

$$\begin{aligned} \| \mathbf{P} - \mathbf{P}^* \| \quad &:= \sup_{|f|\leq 1}\{| \textstyle\int_E f d(\mathbf{P} - \mathbf{P}^*) |: f \text{ measurable}\} \\ &\leq 2P(U \neq U^*), \end{aligned} \tag{5.0.2}$$

i.e. for any coupling (U, U^*) of Z and Z^*, the total variation norm of $\mathbf{P} - \mathbf{P}^*$ is bounded from above by $2P(U \neq U^*)$. A γ-coupling (or a maximal-coupling) is a coupling which gives equality in the coupling inequality (5.0.2).

The objective of this section is to show that for the r.v.s Z and Z^* on $(E, \mathcal{E})$ such that $Z \overset{\mathrm{D}}{\leq} Z^*$, we can construct a coupling, say $(\hat{Z}, \hat{Z}^*)$, such that

$$\begin{aligned} &(i) \quad 2P(\hat{Z} \neq \hat{Z}^*) = \| \mathbf{P} - \mathbf{P}^* \| \quad \text{and} \\ &(ii) \quad \hat{Z} \overset{\text{a.s.}}{\leq} \hat{Z}^*. \end{aligned} \tag{5.0.3}$$

The following general lemma gives a value for the total variation distance between two probability measures.

Lemma 5.1. *Let U and U^* be two random elements on the probability space $(F, \mathcal{F})$ with the corresponding probability measures Q and Q^*. Let λ be a measure on $(F, \mathcal{F})$ such that both Q and Q^* are absolutely continuous with respect to it and with the Radon-Nikodym derivatives g and g^*, respectively,* i.e.

$$g = \frac{dQ}{d\lambda} \quad \text{and} \quad g^* = \frac{dQ^*}{d\lambda}. \tag{5.0.4}$$

If we let $L = \{g > g^*\}$, then the total variation distance between U and U^* is equal to

$$\| Q - Q^* \| = 2(Q(L) - Q^*(L)). \tag{5.0.5}$$

Proof of Lemma 5.1: We write $f \in \mathcal{F}$ in order to say that f is a real-valued measurable function on $(F, \mathcal{F})$. Then, one can write the left hand side of (5.0.5) as

$$\begin{aligned} &\| Q - Q^* \| = \sup\{| \textstyle\int f(g - g^*)d\lambda |:| f |\leq 1 \text{ and } f \in \mathcal{F}\} \\ &= \textstyle\int_L (g - g^*)d\lambda + \int_{L^c}(-1)(g - g^*)d\lambda \\ &= \textstyle\int_L g d\lambda - \int_{L^c} g d\lambda - \int_L g^* d\lambda + \int_{L^c} g^* d\lambda \\ &= Q(L) - Q(L^c) - Q^*(L) + Q^*(L^c) = 2(Q(L) - Q^*(L)). \end{aligned} \tag{5.0.6}$$

■

Define the set H as

$$H := \{x \in E : \mathbf{p}(x) < \mathbf{p}^*(x)\}, \tag{5.0.7}$$

where $\mathbf{p}$ and $\mathbf{p}^*$ are the density functions of the r.v.s Z and Z^*, respectively.

The following easy lemma highlights some simple properties of the set H.

Lemma 5.2. *With the set H as in (5.0.7) we have*

$$\begin{aligned} &(i) \quad H^c \neq \emptyset \quad \text{and} \\ &(ii) \quad \text{if} \;\; H = \emptyset \;\; \text{then} \;\; \mathbf{P} = \mathbf{P}^*. \end{aligned} \tag{5.0.8}$$

Proof of Lemma 5.2: From (5.0.1) we have

$$\mathbf{P}([1,\infty)) \leq \mathbf{P}^*([1,\infty)) \Longrightarrow \mathbf{p}^*(0) \leq \mathbf{p}(0). \tag{5.0.9}$$

Thus $0 \in H^c \neq \emptyset$. This proves (5.0.8)(i).

In order to prove (5.0.8)(ii), notice that if $H = \emptyset$ and there exists a $z \in H^c = E$ such that $\mathbf{p}(z) > \mathbf{p}^*(z)$, then

$$1 = \mathbf{P}^*(H^c) = \textstyle\sum_{x \in H^c} \mathbf{p}^*(x) < \sum_{x \in H^c} \mathbf{p}(x) = \mathbf{P}(H^c) = 1, \tag{5.0.10}$$

provides a contradiction. Thus (5.0.8)(ii) is true. ■

Now define recursively, the real-valued function $\mathbf{q}$ on $E \times E$ as follows

- For $x \in H$, let

$$\mathbf{q}(x,y) = \begin{cases} \mathbf{p}(x), & \text{if} \;\; y = x, \\ 0, & \text{otherwise.} \end{cases} \tag{5.0.11}$$

- If $x \in H^c$, then

$$\mathbf{q}(x,y) = \begin{cases} \mathbf{p}^*(x), & \text{if} \;\; y = x, \\ \min\{\mathbf{p}^*(y) - \mathbf{p}(y) - \sum_{l=0}^{x} \mathbf{q}(l,y), & \\ \mathbf{p}(x) - \sum_{j=x}^{y-1} \mathbf{q}(x,l)\}, & \text{if} \;\; x < y \\ & \text{and} \;\; y \in H, \\ 0, & \text{otherwise.} \end{cases} \tag{5.0.12}$$

In the following example, we derive the function $\mathbf{q}$ for a pair of stochastically ordered random elements on $E = \{0, 1, 2, 3, 4\}$.

Example. Let r.v.s Z and Z^* have the following distributions.

x	$\mathbf{p}(x)$	$\mathbf{p}^*(x)$
0	0.4	0.1
1	0.2	0.1
2	0.2	0.3
3	0.1	0.2
4	0.1	0.3

(5.0.13)

One can easily see that $Z \overset{\mathrm{D}}{\leq} Z^*$. From Table 5.0.13, one can see that the set H defined in (5.0.7) is equal to

$$H = \{2, 3, 4\}. \tag{5.0.14}$$

To calculate the value of the function $\mathbf{q}$ for this example (the result shown in Table 5.0.15), we proceed as follows:

First from (5.0.11)-(5.0.12) we have $\mathbf{p}(x, y) = 0$ if $y < x$, and $\mathbf{q}(x, x) = \mathbf{p}(x)$ if $x \in H$. Now, start from the smallest element of the set H^c, i.e. zero. Since $0 \in H^c$, (5.0.12) gives us $\mathbf{q}(0, 0) = \mathbf{p}^*(0) = 0.1$. And $\mathbf{q}(0, y) = 0$ for all $y \in H^c \setminus \{0\}$. $2 \in H$ is the smallest element of H. Hence $\mathbf{q}(0, 2)$ will be to the amount of mass to be added to the mass $\mathbf{p}(2)$ from the mass difference $\mathbf{p}(0) - \mathbf{p}^*(0)$ in order to obtain at most the mass $\mathbf{p}^*(2)$. Repeat this for the second smallest element of H, i.e. 3. Then, $\mathbf{q}(0, 3)$ will be equal to the amount of mass to be added to the mass $\mathbf{p}(3)$ from the reminder of the mass $\mathbf{p}(0) - \mathbf{p}^*(0) - \mathbf{q}(0, 2)$, to obtain at most the mass $p^*(3)$, and so on we continue until we distribute all the mass difference $\mathbf{p}(0) - \mathbf{p}^*(0)$ among the points of H.

The second smallest element of H^c is one. Let $\mathbf{q}(1, 1) = \mathbf{p}^*(1)$ and for $y \in H$, $\mathbf{q}(1, y)$ (again as it given in (5.0.12)), can be determined in the similar manner to that of $\mathbf{q}(0, y)$ except that this time we add that amount of the mass difference $\mathbf{p}(1) - \mathbf{p}^*(1)$ to each of the elements of H which are greater than 1 (in an order mentioned above), such that we obtain at most the mass $\mathbf{p}^*(y) - \mathbf{p}(y)$, for $y \in H$. Table 5.0.15 presents the results of this calculation.

$\mathbf{q}(x,y)$	$y=0$	1	2	3	4	$\mathbf{p}(x)$
$x=0$	0.1	0	0.1	0.1	0.1	0.4
1		0.1	0	0	0.1	0.2
2			0.2	0	0	0.2
3				0.1	0	0.1
4					0.1	0.1
$\mathbf{p}^*(y)$	0.1	0.1	0.3	0.2	0.3	

(5.0.15)

As one can see, the function $\mathbf{q}$ presented in Table (5.0.15), is the joined density function of a coupling between r.v.s Z and Z^*. □

The following lemma gives some of the properties of the function $\mathbf{q}$.

Lemma 5.3. *For the function* $\mathbf{q}$ *given in (5.0.11)-(5.0.12) we have*

$$\begin{aligned} &(i) \quad \mathbf{q}(x,y) \geq 0, \text{ for all } x,y \in E, \\ &(ii) \quad \mathbf{q}(x,y) = 0, \text{ if } x > y, \\ &(iii) \quad \textstyle\sum_{0 \leq l < y} \mathbf{q}(l,y) = \mathbf{p}^*(y) - \mathbf{p}(y), \text{ for all } y \in H, \\ &(iv) \quad \textstyle\sum_{l > x} \mathbf{q}(x,l) = \mathbf{p}(x) - \mathbf{p}^*(x), \text{ for all } x \in H^c, \end{aligned} \tag{5.0.16}$$

where the set H *is as in (5.0.7).*

Proof of Lemma 5.3: (5.0.16)(i) is obvious from (5.0.11)-(5.0.12). To see (5.0.16)(ii) notice that by the construction of $\mathbf{q}$ in (5.0.11)-(5.0.12), $\mathbf{q}(x,y)$ can be non-zero only if $x \leq y$.

Let $\pi : H \to H^c$ such that $\pi(y) := \max\{x \in H^c : x < y\}$. Since by (5.0.12) for $y \in H$, $\mathbf{q}(z,y) = 0$ for $\pi(y) < z < y$. Then we have

$$\mathbf{q}(\pi(y),y) \leq \mathbf{p}^*(y) - \mathbf{p}(y) - \sum_{l=0}^{\pi(y)-1} \mathbf{q}(l,y). \tag{5.0.17}$$

Since $\mathbf{q}(l,y) = 0$ for $\pi(y) < l < y$, (5.0.17) implies that

$$\mathbf{p}^*(y) - \mathbf{p}(y) \geq \sum_{l=0}^{\pi(y)} \mathbf{q}(l,y) = \sum_{l=0}^{y-1} \mathbf{q}(l,y). \tag{5.0.18}$$

To show (5.0.16)(iii), assume that for some $y_0 \in H$, we have

$$\mathbf{p}^*(y_0) - \mathbf{p}(y_0) > \sum_{l=0}^{\pi(y_0)} \mathbf{q}(l,y_0) \tag{5.0.19}$$

and we show that this leads to a contradiction. If in (5.0.18) we sum over all $y \in H$ and $y \leq y_0$, we obtain

$$\sum_{y=1}^{y_0} (\mathbf{p}^*(y) - \mathbf{p}(y))\mathbf{1}_H(y) \geq \sum_{y=1}^{y_0} \sum_{l=0}^{y-1} \mathbf{q}(l, y)\mathbf{1}_H(y). \tag{5.0.20}$$

Now, from the construction of the function $\mathbf{q}$, one can see that for each point $x \in H^c$, we distribute the mass difference $\mathbf{p}(x) - \mathbf{p}^*(x)$ to the points in H which are greater than x. We do this in an increasing order, i.e. if y is a point in H greater than $x \in H^c$, then first we do not allow any mass to be moved from the mass $\mathbf{p}(x) - \mathbf{p}^*(x)$ to the point y, unless we have distributed enough mass difference to all the points in H which are smaller than y. So, if (5.0.19) holds, it means that $\sum_{y=1}^{y_0} \sum_{l=0}^{y-1} \mathbf{q}(l, y)\mathbf{1}_H(y)$ is equal to all the possible mass that we could distribute from H^c to the points $\{y \in H : y \leq y_0\}$, i.e. $\sum_{y=0}^{y_0}(\mathbf{p}(y) - \mathbf{p}^*(y))\mathbf{1}_{H^c}(y)$. Hence we have

$$\begin{aligned}
\textstyle\sum_{y=1}^{y_0}(\mathbf{p}^*(y) - \mathbf{p}(y))\mathbf{1}_H(y) \quad & > \textstyle\sum_{y=1}^{y_0} \sum_{l=0}^{y-1} \mathbf{q}(l, y)\mathbf{1}_H(y) \\
& = \textstyle\sum_{y=0}^{y_0}(\mathbf{p}(y) - \mathbf{p}^*(y))\mathbf{1}_{H^c}(y) \\
\Longrightarrow \quad & \textstyle\sum_{y=1}^{y_0} \mathbf{p}^*(y)\mathbf{1}_H(y) + \sum_{y=0}^{y_0} \mathbf{p}^*(y)\mathbf{1}_{H^c}(y) \\
& > \textstyle\sum_{y=1}^{y_0} \mathbf{p}(y)\mathbf{1}_H(y) + \sum_{y=0}^{y_0} \mathbf{p}(y)\mathbf{1}_{H^c}(y) \\
\Longrightarrow \quad & \mathbf{P}([0, y_0]) < \mathbf{P}^*([0, y_0]) \\
\Longrightarrow \quad & \mathbf{P}([y_0 + 1, \infty)) > \mathbf{P}^*([y_0 + 1, \infty)).
\end{aligned} \tag{5.0.21}$$

A contradiction with (5.0.1), thus for all $y \in H$ we have

$$\begin{aligned}
\mathbf{p}^*(y) - \mathbf{p}(y) &= \textstyle\sum_{l=0}^{\pi(y)} \mathbf{q}(l, y) \\
\text{or equivalently} & \\
\mathbf{p}^*(y) - \mathbf{p}(y) &= \textstyle\sum_{l=0}^{y-1} \mathbf{q}(l, y).
\end{aligned} \tag{5.0.22}$$

This proves (5.0.16)(iii)

To show (5.0.16)(iv) we proceed similarly. Again from (5.0.11)-(5.0.12) we can see that for $x \in H^c$ we have

$$\mathbf{p}(x) - \mathbf{p}^*(x) \geq \sum_{l>x} \mathbf{q}(x, l)\mathbf{1}_H(l). \tag{5.0.23}$$

Now, assume that for some $x_0 \in H^c$ we have

$$\mathbf{p}(x_0) - \mathbf{p}^*(x_0) > \sum_{l>x_0} \mathbf{q}(x_0, l)\mathbf{1}_H(l), \tag{5.0.24}$$

and we show that this leads to a contradiction. Since $\mathbf{p}$ and $\mathbf{p}^*$ are probability densities on E and $x_0 \in H^c$. We have

$$\begin{aligned} &\textstyle\sum_{k=0}^{\infty} \mathbf{p}^*(k) = \sum_{k=0}^{\infty} \mathbf{p}(k) = 1 \implies \\ &\implies \textstyle\sum_{k>0}(\mathbf{p}^*(k) - \mathbf{p}(k))\mathbf{1}_H(k) = \sum_{k\geq 0}(\mathbf{p}(k) - \mathbf{p}^*(k))\mathbf{1}_{H^c}(k). \end{aligned} \tag{5.0.25}$$

Now (5.0.23), (5.0.24) and (5.0.25) imply that

$$\begin{aligned} &\textstyle\sum_{k>0}(\mathbf{p}^*(k) - \mathbf{p}(k))\mathbf{1}_H(k) \\ &> \textstyle\sum_{k\geq 0}\sum_{l>k} \mathbf{q}(k,l)\mathbf{1}_H(l)\mathbf{1}_{H^c}(k) \\ &= \textstyle\sum_{l>0}\sum_{k<l} \mathbf{q}(k,l)\mathbf{1}_{H^c}(k)\mathbf{1}_H(l). \end{aligned} \tag{5.0.26}$$

On the other hand, from (5.0.11)-(5.0.12) for $k < l$ we have $\mathbf{q}(k,l) = 0$ when $k \in H$. Also from (5.0.16)(iii) we have $\sum_{k<l} \mathbf{q}(k,l)\mathbf{1}_H(l) = (\mathbf{p}^*(l) - \mathbf{p}(l))\mathbf{1}_H(l)$. Hence we can write (5.0.26) as

$$\textstyle\sum_{k>0}(\mathbf{p}^*(k) - \mathbf{p}(k))\mathbf{1}_H(k) > \sum_{l>0}(\mathbf{p}^*(l) - \mathbf{p}(l))\mathbf{1}_H(l) \tag{5.0.27}$$

A contradiction. Thus (5.0.16)(iv) is true. ■

Now we are ready for the main result of this Chapter which is presented in the following theorem.

Theorem 5.4. *Let Z and Z^* be* r.v.s *on $E = \mathcal{Z}_+$, such that $Z \overset{\mathrm{D}}{\leq} Z^*$. Then there is a coupling $(\hat{Z}, \hat{Z}^*)$ of Z and Z^* such that*

$$\begin{aligned} &(i) \quad P(\hat{Z} = x, \hat{Z}^* = y) = \mathbf{q}(x,y), \ \text{ for } \ x, y \in E, \\ &(ii) \quad \hat{Z} \leq \hat{Z}^* \ \text{ a.s., and} \\ &(iii) \quad (\hat{Z}, \hat{Z}^*) \ \text{ is a } \gamma-\text{coupling of } Z \text{ and } Z^*, \end{aligned} \tag{5.0.28}$$

where the function $\mathbf{q}$ is given in (5.0.11)-(5.0.12).

Proof of Theorem 5.4: From Lemma 5.3 we have $\mathbf{q}(x,y) \geq 0$ for all $x, y \in E$. To show that $\mathbf{q}$ is a probability density function of a coupling between the r.v.s Z and Z^*, it is enough to show that $\sum_{z\in E}\mathbf{q}(x,z)=\mathbf{p}(x)$ and $\sum_{z\in E}\mathbf{q}(z,x)=\mathbf{p}^*(x)$, for all $x\in E$.

If $x\in H$ then from (5.0.11) we have $\sum_{z\in E}\mathbf{q}(x,z)=\mathbf{p}(x)$. If $x\in H^c$, then from (5.0.12) and Lemma 5.3(iv) we have

$$\begin{aligned}&\textstyle\sum_{z\in E}\mathbf{q}(x,z)=\mathbf{q}(x,x)+\sum_{l>x}\mathbf{q}(x,l)\\ &=\mathbf{p}^*(x)+(\mathbf{p}(x)-\mathbf{p}^*(x))=\mathbf{p}(x).\end{aligned}\tag{5.0.29}$$

Similarly, if $x\in H$ then from (5.0.11) and Lemma 5.3(iii) we have

$$\begin{aligned}&\textstyle\sum_{z\in E}\mathbf{q}(z,x)=\mathbf{q}(x,x)+\sum_{l<x}\mathbf{q}(l,x)\\ &=\mathbf{p}(x)+(\mathbf{p}^*(x)-\mathbf{p}(x))=\mathbf{p}^*(x).\end{aligned}\tag{5.0.30}$$

If $x\in H^c$, from (5.0.12) we have $\sum_{z\in E}\mathbf{q}(z,x)=\mathbf{p}^*(x)$. Hence we have proved (5.0.28)(i).

Lemma 5.3(ii) gives (5.0.28)(ii), i.e. for $x,y\in E$ and $y<x$ we have $\mathbf{q}(x,y)=P(\hat{Z}=x,\hat{Z}^*=y)=0$.

To show (5.0.28)(iii), by Lemma 5.1 it is enough to show that $P(\hat{Z}\le\hat{Z}^*)=\mathbf{P}^*(H)-\mathbf{P}(H)$. But by (5.0.28)(ii), we can write $P(\hat{Z}\le\hat{Z}^*)$ as

$$\begin{aligned}&P(\hat{Z}<\hat{Z}^*)=1-P(\hat{Z}=\hat{Z}^*)\\ &=1-P(\hat{Z}=\hat{Z}^*\in H)-P(\hat{Z}=\hat{Z}^*\in H^c)\\ &\textstyle=1-\sum_{z\in E}\mathbf{q}(z,z)\mathbf{1}_H(z)-\sum_{z\in E}\mathbf{q}(z,z)\mathbf{1}_{H^c}(z).\end{aligned}\tag{5.0.31}$$

Now from (5.0.11)-(5.0.12) we can see that $\mathbf{q}(z,z)\mathbf{1}_H(z)=\mathbf{p}(z)\mathbf{1}_H(z)$ and $\mathbf{q}(z,z)\mathbf{1}_{H^c}(z)=\mathbf{p}^*(z)\mathbf{1}_{H^c}(z)$. Thus we can write (5.0.31) as

$$\begin{aligned}&\textstyle P(\hat{Z}<\hat{Z}^*)=1-\sum_{z\in E}\mathbf{p}(z)\mathbf{1}_H(z)-\sum_{z\in E}\mathbf{p}^*(z)\mathbf{1}_{H^c}(z)\\ &=1-\mathbf{P}(H)-\mathbf{P}^*(H^c)=\mathbf{P}^*(H)-\mathbf{P}(H).\end{aligned}\tag{5.0.32}$$

Hence $(\hat{Z},\hat{Z}^*)$ is a γ-coupling of Z and Z^*.

■

Bibliography

[1] Alsmeyer, M. and Rösler, U., (1996). The Bisexual Galton-Watson Process Promiscuous Mating: Extinction Probabilities in the Supercritical Case. *Ann. Appl. Probab.* **6**, no. 3, 922-939.

[2] Asmussen, S. , (1980). On Some Two-Sex Population Models. *The Annals of Probability* **8**, No. 4, 727-744.

[3] Asmussen, S. and Hering, H., (1983). Branching Processes. Birkhäuser, Boston.

[4] Athreya, K. B. and Ney, P. E. , (1972). Branching Processes. Springer-Verlag, Berlin.

[5] Badalbaev I. S. and Muhitdinov A. A., (1990). Statističeskie zadači mnogotipnyh vetvyaščihsya protsessov. Fan, Tashkent.

[6] Chow, Yuan Shih and Teicher, Henry, (1978). Probability theory: independence, interchangeability and Martingales. Springer-Verlage, New York.

[7] Daley, D. J., (1968). Extinction Conditions for Certain Bisexual Galton-Watson Branching Processes. *Z. Wahrscheinlichkeitstheorie verw. Geb.* **9**, 315-322.

[8] Daley, D. J. , Hull, David M. and Taylor, James M., (1986). Bisexual Galton-Watson Branching Processes with Superadditive Mating Functions. *J. Appl. Prob.* **23**, 585-600.

[9] Fujimagari, Tetsuo, (1976). Controlled Galton-Watson Process and its Asymptotic Behavior. *Kodai Math. Sem. Rep.* **27**, 11-18.

[10] Goodman, L. A., (1953). Population growth of the sexes. *Biometrics* **9**, 212-225.

[11] Goodman, L. A., (1968). Stochastic models for the population growth of the sexes. *Biometrika* **55**, 469-487.

[12] González, M. and Molina, M., (1992). Some Theoretical Results for Bisexual Galton-Watson Branching Processes. *Extracta Math.* **7**, no. 2-3, 140-143.

[13] González, M. and Molina, M., (1996). On the Limit behaviour of a Supperadditive Bisexual Galton-Watson Branching Process. *J. Appl. Prob.* **33**, 960-967.

[14] González, M. and Molina, M., (1997). On the L^2-Convergence of a Supperadditive Bisexual Galton-Watson Branching Process. *J. Appl. Prob.* **34**, 575-582.

[15] Hull, D. M., (1982). A necessary condition for extinction in those bisexual Galton-Watson branching processes governed by superadditive mating functions. *J. Appl. Prob.* **19**, 847-850.

[16] Hull, D. M. , (1998). A reconsideration of Galton's Problem (Using a Two-Sex Population). *Theoretical Population Biology* **54**, 105-116.

[17] Jagers, Peter, (1975). Branching processes with biological applications. John Wiley & Sons, London.

[18] Jagers, Peter, (1997). Coupling and population dependence in branching processes. *The Annals of Applied Probability*, **7**, 281-298.

[19] Jagers, Peter, (1997). Towards dependence in general branching processes. *Classical and modern branching processes (Minneapolis, MN, 1994)*, 127-139, IMA Vol. Math. Appl., 84, *Springer, New York*, 1997.

[20] Keller, G., Kersting, G. and Rösler, U. (1987). On the Asymptotic behaviour of Discrete Time Stochastic Growth Processes. *The Annals of Probability* **15**, 305-343.

[21] Kendall, D. G., (1949). Stochastic Processes and Population Growth. *Journal of the Royal Statistical Society, Series B* **11**, 230-282.

[22] Kersting, G., (1986). On Recurrence and Transience of Growth models. *J. Appl. Prob.* **23**, 614-625.

[23] Kersting, G. (1992). Asymptotic Γ-distribution for stochastic difference equations. *Stochastic Processes and their Applications* **40**, 15-28.

[24] Klebaner, Fima C., (1984). Geometric rate of growth in population-size-dependent branching processes. *J. Appl. Prob.* **21**, 40-49.

[25] Klebaner, Fima C., (1985). A limit theorem for population-size-dependent branching processes. *J. Appl. Prob.* **22**, 48-57.

[26] Klebaner, Fima C., (1989). Linear growth in near-critical population-size-dependent multitype Galton-Watson processes. *J. Appl. Prob.* **26**, 431-445.

[27] Klebaner, Fima C., (1989). Stochastic Difference Equations and Generalized Gamma Distributions. *The Annals of Probability* **17**, 178-188.

[28] Klebaner, Fima C., (1989). Geometric Growth in Near-Supercritical Population Size Dependent Multitype Galton-Watson Processes. *The Annals of Probability* **17**, 1466-1477.

[29] Krebs, J. R. and Davies, N. B. , (1991). Behavioural ecology: an evolutionary approach, 3rd ed. Blackwell Scientific Publications, Oxford.

[30] Krebs, J. R. and Davies, N. B. , (1997). Behavioural ecology: an evolutionary approach, 4th ed. Blackwell Scientific Publications, Oxford.

[31] Lindvall, Torgny, (1992). Lectures on the coupling method. John Wiley & Sons, New York.

[32] Mode, C. J., (1974). A Discrete Time Stochastic Growth Process for Human Populations Accommodating Marriages. *Mathematical Biosciences* **19**, 201-219.

[33] Mode, C. J., (1974). On the Evaluation of a Determinant Arising in a Bisexual Stochastic Growth Process. *Journal of Mathematical Analysis and Applications* **46**, 190-211.

[34] Pečarić, J. E. , Proschan, F. and Tong, Y. L. , (1992). Convex Functions, Partial Orderings, and Statistical Applications. Academic Press, Inc.