SOFIA UNIVERSITY “ST. KLIMENT OHRIDSKI”

FACULTY OF MATHEMATICS AND INFORMATICS
DEPARTMENT OF PROBABILITY, OPERATIONS RESEARCH AND STATISTICS

ABSTRACT

OF DISSERTATION FOR PH. D. DEGREE IN MATHEMATICS

Professional direction: 4.5. Mathematics
Doctoral programme: Probability Theory and Mathematical Statistics

Branching processes - optimization and applications

Kaloyan Nikolaev Vitanov, M. Sc.

Ph. D. supervisor: prof. Maroussia Slavtchova-Bojkova

SOFIA, BULGARIA
2022

Contents

1 Introduction **5**
1.1 Short history of branching processes 5
1.2 A descriptive definition of a Sequential Decision Problem (SDP) 6
1.3 Conceptual organization of the dissertation 6

2 Multi-type continuous-time branching processes through probabilities of mutation between types **9**
2.1 Chapter overview and organization 9
2.2 Multi–type Sevastyanov Branching Processes through probabilities of Mutation between types (MSBPM) 12
2.2.1 Notation and definition of the MSBPM 12
2.2.2 Probability generating functions for the MSBPM 14
2.2.3 Probabilities of extinction for the MSBPM 14
2.2.4 Number of particles produced from $\mathbb{W}_e$ towards $\mathbb{W}$ within the MSBPM 16
2.2.5 Time until occurrence of the first "successful" particle produced from $\mathbb{W}_e$ towards $\mathbb{W}$ within the MSBPM 19
2.2.6 Immediate risk of producing a "successful" particle from $\mathbb{W}_e$ towards $\mathbb{W}$ within the MSBPM 20
2.2.7 Numerical schemes for computing obtained systems of integral equations for the MSBPM 21
2.3 Particular cases of the MSBPM 23
2.3.1 Decomposable Multi-type Sevastyanov Branching Process through probabilities of Mutation between types (DMSBPM) 23
2.3.2 Decomposable Multi-type Bellman-Harris Branching Process through probabilities of Mutation between types (DMBHBPM) 27

3 Sequential decision problems with branching process based dynamics **29**
3.1 Chapter overview and organization 29
3.2 Modeling of Sequential Decision Problems (SDPs) 31
3.3 Bellman's optimality equation for SDPs within "Universal Modeling Framework" 34
3.4 SDPs with underlying BGW branching process dynamics 35

3.5 SDPs with underlying exponential lifespan MBHBPM dynamics 37
3.6 SDPs with underlying MSBPM dynamics 38

Conclusion **41**
Approbation 41
Scientific contributions 41
Note regarding used software 42
Declaration for originality of obtained results 43
Gratitude 43

Bibliography **43**

CHAPTER 1

Introduction

1.1 Short history of branching processes

Branching processes are a class of stochastic processes that capture some of the fundamental aspects of division and propagation observed in nature. Branching processes model the evolution of a population of objects (these objects can correspond to real-world elementary particles, photons, electrons, atoms, molecules, cells, viruses, bacteria, animals, people, information, finances, and other entities) through time and study various characteristics of this evolution. It comes as no surprise that the areas of application of branching processes are diverse and numerous - physics, chemistry, biology, demography, ecology, economy, etc. This diversity of contexts stimulates the development of various kinds of branching processes, adapted towards answering particular questions of interest found within these contexts. Indeed, there are branching processes in discrete-time and continuous-time, with one or multiple types of objects/particles/cells, branching processes with random immigration, branching diffusion processes, spatial branching processes, controlled branching processes, and others.

The study of branching processes begins around the middle of the 19th century with the question of explaining the disappearance of aristocratic family lines in Europe. In 1845, the French mathematician and statistician Bienaymé first studied the process of extinction of the French noble families, [83], and created the first branching process model. Unfortunately, as Bienaymé left no students, his name, as well as his results, gradually faded into obscurity. A few decades later, in 1873, concerns regarding the extinction of noble family names reappeared among the scientific community with the famous work of Galton and Watson, [85]. Despite the fallacy in the final conclusion of Galton and Watson, [85] is considered by many as the beginning of branching process theory. It took more that 50 years for the correct solution to be published by the Danish mathematician J. F. Steffensen within [88] in 1930. It took 40 more years for Heyde and Seneta to first note in 1972 that Bienaymé already had the correct statement of the Criticality Theorem back in 1845 (see [84]).

After World War II, branching processes and their applications in physics became intensely researched, leading to the rapid development of the field. The term *branching process* itself is considered to have been coined by A. N. Kolmogorov and N. A. Dmitriev in their work [89] from 1947. Among the most influential monographs in the field are those of Harris [58] (1963), Sevastyanov [8] (1971), and [57] from Athreya and Ney (1972). In more contemporary times the focus of questions and applications explored through branching processes has shifted away from physics and has orientated towards biological contexts - see Kimmel and Axelrod [14] (2002), Haccou, Jagers, and Vatutin [10] (2007), Durrett [104] (2015), Pardoux [36] (2016).

1.2 A descriptive definition of a Sequential Decision Problem (SDP)

Within Chapter 3 of the dissertation, we investigate stochastic sequential decision problems in the context of systems with underlying branching process dynamics. Here, we give a brief informal description of what a "sequential decision problem" is.

Assume that there is a dynamic system with which we can interact. We observe the system at specified moments in time, called *decision epochs*. We will denote the set of decision epochs with $\hat{t}_T$, $\hat{t}_T = \{0 = t_0, t_1, t_2, \ldots, t_T = \hat{T}\}$, the distance between two neighboring decision epochs can vary but cannot be 0 or ∞. At each decision epoch $t_i \in \hat{t}_T$, except at $t_T = \hat{T}$, we make a decision (interaction with the system) that affects how the system evolves from t onward. Upon making a decision, we collect rewards (or incur costs), with the exception of $t_T = \hat{T}$ where we only collect predefined rewards (or incur predefined costs). A *sequential decision problem* is a problem of choosing such decisions so that the cumulative expected reward, after collecting the rewards at $t_T = \hat{T}$, is maximized (or the cumulative expected cost is minimized). The problem can be deterministic or stochastic.

1.3 Conceptual organization of the dissertation

This dissertation is conceptually divided into two topics explored in Chapter 2 and Chapter 3 respectively. Chapter 2 defines the novel Multi–type Sevastyanov Branching Processes through probabilities of Mutation between types (MSBPM) and obtains results of interest in the context of populations escaping extinction. Chapter 3 is dedicated towards the incorporation of branching processes, including the MSBPM, into optimization problems known as Sequential Decision Problems (SDPs).

The novel MSBPM from Chapter 2 is connected to the classic multi-type Sevastyanov branching process, however, within the MSBPM, probabilities for mutation are used for writing down expressions of interest. This makes the novel MSBPM well adapted towards modeling biological populations under stress that escape extinction. Within Chapter 2, we obtain various systems of equations for the MSBPM as well as for quantities relevant in the context of populations escaping extinction. To the best of our knowledge, such an in-depth investigation of the topic has not been done previously (excluding our earlier work in [7] as well as preceding papers [1] - [6]) for multi-type, continuous-time branching processes. We explore the case of the MSBPM starting with one particle of age 0 and the case of the MSBPM starting with one particle of age a, $a \neq 0$. The latter case, to the best of our knowledge, has not been explored previously in a systematic manner within the context of branching processes. Numerical schemes for calculating all obtained systems of equations are also developed within Chapter 2.

In Chapter 3, we begin with an introduction of the "Universal Modeling Framework" developed by Warren B. Powell in [82] (2022). The choice of modeling framework within which we specify our Sequential Decision Problems (SDPs) is of paramount importance for our perspective on the systems we attempt to model as well as for the ease of possible future extensions of our results. Our choice of framework allows us to utilize Bellman's optimality equation for finding solutions of SDPs, provided that our models conform to the assumptions of the framework. We proceed with our novel considerations and results as follows. We recast the multi-type Bienaymé-Galton-Watson (BGW) branching process optimization problem, considered in [77], as a SDP within the "Universal Modeling Framework". In Theorem 3.2 from Subsection 3.4.3 within the dissertation, we provide a novel proof for Theorem 3.1 from [77] that is based on Bellman's optimality equation. Theorem 3.2 enables us to efficiently find the solution of SDPs with underlying BGW dynamics. Next, we incorporate the multi-type Bellman-Harris branching process with exponential lifespan distributions as well as the Multi-type Bellman-Harris Branching Process through probabilities of Mutation between types (MBHBPM; a particular case of the MSBPM) with exponential lifespan distributions into a SDP and prove that a result, similar to Theorem 3.2, holds. We then shown that, with respect to a novel state space, the MSBPM and the multi-type Sevastyanov branching process can also be incorporated into SDPs. Unfortunately, an analogue of Theorem 3.2 is not available for these processes. Regardless, Bellman's optimality equation allows us to consider the Approximate Dynamic Programming (ADP) approach for finding the solution of obtained SDPs within future research. We conclude our investigations by outlining a general ADP algorithm based on post-decision state variables that may serve as a starting point for the future development of a specialized ADP algorithm for SDPs with branching process based dynamics.

Within the Appendix, we have provided some standard results regarding the Perron-Frobenius theorem. We reference these results in some of our MSBPM related discussions.

More detailed description of the structure of Chapter 2 and Chapter 3, as well as relevant remarks and discussions, can be seen in the corresponding "Chapter overview and organization" sections within these chapters.

CHAPTER 2

Multi-type continuous-time branching processes through probabilities of mutation between types

2.1 Chapter overview and organization

In this Chapter, we define the novel continuous time branching process model that plays the central role within this dissertation - the Multi–type Sevastyanov Branching Processes through probabilities of Mutation between types (MSBPM). The MSBPM can be considered as a relative of the classical multi-type Sevastyanov branching process as defined in Chapter VIII in [8]. The novel characteristic of the MSBPM, with respect to the classical formulation in [8], is the use of probabilities of mutation (a particle is a "mutant" if it is of type that is different from the type of its mother particle) between types. More specifically, through the use of probabilities of mutation, effectively, we decompose the classical probabilities $p^i_{\boldsymbol{\alpha}}(u)$ for a particle of type i of age u to transform into $\boldsymbol{\alpha}$ particles at the end of its lifespan (see page 229 in [8] or Subsection 1.3.1 from the dissertation) into two components: 1) Probabilities $p_{ik}(u)$ for the total k number of offspring, regardless of offspring type, of a type i particle of age u; 2) Probabilities for mutation of an offspring particle of a type i particle towards type j, u_{ij}.

The use of probabilities for mutation opens the way for applications of the MSBPM into many biological contexts. Most notably, the MSBPM is well suited for modeling biological populations under stress that face certain extinction unless a "beneficial" mutation occurs (or a combination of mutations occur), leading to supercritical behavior. Such situations are of interest in the areas of cancer modeling and treatment, spread of viruses, vaccination campaigns, control over agricultural pests and others (see, e.g., [61], [62], [63], [64], [65], [66], [104], [1] - [7]). In biological contexts it is easier to estimate the probabilities for the total number of offspring, $p_{ik}(u)$, and the probabilities of mutation, u_{ij}, found within the MSBPM, than the more abstract $p^i_{\boldsymbol{\alpha}}(u)$ used in the multi-type Sevastyanov branching process. The use of $p_{ik}(u)$ and u_{ij} often provides us with a model with more clear and straightforward interpretations.

Within this Chapter, we concentrate our efforts towards obtaining results for the MSBPM regarding quantities that are of interest in the context of populations escaping extinction. Other authors, see [64], [65], [66], discuss similar topics to the ones explored within the dissertation, however, their discussions is based on multi-type Bienaymé-Galton-Watson (BGW) processes. The BGW is a discrete time branching process while the MSBPM is in continuous time - a more difficult theoretical setting. [61], [62], [64], [65], explore populations escaping extinction under the assumption that probabilities of mutation are small quantities. Such an assumption is not made for obtaining the results for the MSBPM, making the model more general in terms of possible applications. [61] and [62] assume Poisson and geometric offspring distributions for obtaining their results. The results obtained for the MSBPM do not rely on particular assumptions of offspring distributions. [64], [65] consider only one supercritical type, in contrast the MSBPM can accommodate an arbitrary number of supercritical types, in addition to that almost all of our results do not depend on type criticality. Further, almost all results obtained for the MSBPM do not rely on an assumption about the process being non-decomposable, this assumption being central for many results valid for the classical multi-type Sevastyanov branching process. We also note that our results for the MSBPM do not rely on particular assumptions about the lifespan distributions among types. All mentioned features of the MSBPM and the results obtained within the current Chapter, highlight the flexibility and wide area of applicability of the MSBPM in modeling populations escaping extinction. However, the MSBPM is not to be understood as exclusively tied to biology - the model can be applied in other areas as well, provided a proper interpretation of u_{ij}.

This Chapter is a continuation and generalization of our previous work in [1] - [7] where the focus is on cancer modeling as well as modeling escape from extinction. More specifically, with respect to our latest work in Vitanov & Slavtchova-Bojkova [7] (2022), the MSBPM provides an extension in the following directions: 1) The MSBPM can be non-decomposable; 2) The “emitting” class and the target class can intersect. The process discussed in [7] is a particular case of the MSBPM, that is, the decomposable MSBPM (DMSBPM) explored in Subsection 2.3.1 of this dissertation. The DMSBPM is of particular interest for modeling mutation as it describes an irreversible path in the evolution of a population of particles. We note that the development of cancer resistance towards medical treatment in many situations can be attributed to biological mutations. Thus, the MSBPM and its decomposable variants rise as well suited candidates for modeling the risk of cancer reemerging even when an apparently successful treatment is applied. We note that the work in the following Chapter 3 can be considered as a further continuation of [1] - [7]. A real-world example of a sequential decision problem (we discuss these problems in Chapter 3) is the planning of cancer treatment administration throughout time with respect to cost and benefit considerations. While the results from Chapter 3 are not yet ready for handling the nuances of this particular example, Chapter 3 is to be understood as a step towards solving such problems.

Within this Chapter, we obtain systems of integral equations for the probability generating functions (p.g.f.s) of the MSBPM as well as for the probabilities of extinction within the MSBPM. We obtain p.g.f.s for the production of particles from one class of particle types within the process to another. For the general case the particles produced need not necessarily be mutants, however, for particular cases of interest of the MSBPM, such as the decomposable MSBPM (DMSBPM), we investigate only the production of mutants. The DMSBPM can be used to model, for example, a “beneficial” mutation that is reachable only after certain preceding mutations have occurred (see Figure 2.12 and

Figure 2 within [61]) or other relevant mutation schemes. We obtain the distribution of the random variable “time until first ’successful’ particle”/“time until first ’successful’ mutant” within the MSBPM, i.e., the time until the occurrence of a particle/mutant that initiates a non-extincting process. We also obtain expressions for the hazard function that is with respect to the occurrence of the first “successful” particle/mutant. We stress that the proofs of the results within this Chapter do not depend on assumptions whether a process is decomposable or not (although some statements are valid only for a decomposable process), decomposability is treated as a particular case where a particular set of probabilities of mutation contains only zeros.

In addition to obtaining results for the case where the MSBPM starts with one particle of age 0, we also obtain novel variants of our results for the case where the process starts with one particle of age a, $a \neq 0$. We have so far not detected other authors that consider initial particles with non-zero age. As can be seen from the various figures within the dissertation, significant difference in the behavior of the investigated quantities can be observed when we look at those t that are close to the beginning of the process. This observation has the potential to be very useful with respect future research stemming from Chapter 3, where optimization problems related to decision making are investigated in the context of branching processes.

All results obtained within the Chapter can be computed with the help of the novel Numerical Scheme 1 and Numerical Scheme 2 constructed in Subsection 2.2.7 of the dissertation.

This Chapter is organized as follows. In Section 2.2, we define the Multi–type Sevastyanov Branching Processes through probabilities of Mutation between types (MSBPM) and obtain various systems of equations for quantities of interest in the context of populations escaping extinction. More specifically, in Subsection 2.2.1 we define the MSBPM. We then obtain the system of integral equations for the probability generating functions (p.g.f.s) of the process in Subsection 2.2.2 as well as results the probabilities for extinction within Subsection 2.2.3. Next, in Subsection 2.2.4 we investigate the p.g.f.s for the number of particles produced within the process from a class of particle types towards all types within the process. We then continue with results concerning the occurrence of the first “successful” particle produced from any type within a class of types within the MSBPM in Subsection 2.2.5. In Subsection 2.2.6 we obtain expressions for the hazard function defined with respect to the occurrence of the first “successful” particle. In Subsection 2.2.7 we provide two numerical schemes that can be used for computing obtained systems of equations throughout Chapter 2. We finish this Section with specifications of the example MSBPM that we use, in conjunction with the constructed numerical schemes, for demonstrating results obtained within the Chapter. In Section 2.3, we investigate two particular cases of the MSBPM. In Subsection 2.3.1, we consider the decomposable MSBPM (DMSBPM). Within the Subsection, we obtain variants of the novel results for the MSBPM that are valid for the DMSBPM and also explore some additional results that stem from the enforced decomposability. In Subsection 2.3.2, we consider the particular case of the DMSBPM where there is no dependence of particle reproduction from particle age.

2.2 Multi–type Sevastyanov Branching Processes through probabilities of Mutation between types (MSBPM)

The current Section contains novel, yet unpublished, results that extend our recent publication Vitanov & Slavtchova-Bojkova [7] (March 2022).

2.2.1 Notation and definition of the MSBPM

We begin with the introduction of some of the notation and prerequisites that we extensively use throughout the dissertation:

1. Let $\mathbb{W} = \{1, 2, \dots, n\}$. $\mathbb{W}$ denotes the set of possible particle types.

2. Denote $\boldsymbol{\delta}^i = \left(\delta_1^i, \dots, \delta_n^i\right)^\top$, where $\delta_j^i = 0$ if $i \neq j$ and $\delta_j^i = 1$ if $i = j$. We will use $\boldsymbol{\delta}^i$ to specify a single initial particle of type i that is of age 0. For a single initial particle of type i that is of age a, $a \neq 0$, we will use $\boldsymbol{\delta}_a^i$. Again, we set $\boldsymbol{\delta}_a^i = \left(\delta_1^i, \dots, \delta_n^i\right)^\top$ with $\delta_j^i = 0$ if $i \neq j$ and $\delta_j^i = 1$ if $i = j$, however, the subscript "a" in $\boldsymbol{\delta}_a^i$ now specifies the age of the initial particle.

3. We will be denoting the lifespan cumulative distribution function (c.d.f.) at t for type i particles, of age 0, with $G_i(t)$. If a type i particle is of age a, we will denote the corresponding c.d.f., conditioned on the age of the particle, with $G_{i,a}(t)$.

4. If X is some random variable (r.v.), we denote with $\widetilde{X}$ an identical and independent copy of X. Also, if $\boldsymbol{X} = \left(X_1, \dots, X_n\right)^\top$ is a random vector, then $\widetilde{\boldsymbol{X}}$ is an identical and independent copy of $\boldsymbol{X}$.

5. The probability generating function (p.g.f.) of a discrete r.v. X is given by $\mathbb{E}\left[s^X\right] = \sum_{x=0}^{\infty} p_x s^x$, where $|s| \leq 1$. The p.g.f. of a random vector $\boldsymbol{X} = \left(X_1, \dots, X_n\right)^\top$, comprised of discrete r.v.s, is given by

$$\mathbb{E}\left[\prod_{i=1}^{n} s_i^{X_i}\right] = \sum_{x_1,\dots,x_n=0}^{\infty} \left[p(x_1, \dots, x_n) \prod_{i=1}^{n} s_i^{x_i}\right],$$

   where $max\{|s_1|, \dots, |s_n|\} \leq 1$. The last requirement can be written as $|\boldsymbol{s}| \leq 1$, where $\boldsymbol{s} = (s_1, \dots, s_n)^\top$.

We define the novel branching process of main interest within the dissertation:

Definition 2.1. *Define the Multi-type Sevastyanov Branching Process through probabilities of Mutation between types (MSBPM) as the multi-type branching process satisfying:*

1. *Each particle type is uniquely associated with an integer from $\mathbb{W}$ and conforms to:*

(a) *The lifespan of particles of type* i, $i \in \mathbb{W}$, *is modeled by a (continuous) r.v.* τ_i. *The corresponding cumulative distribution function (c.d.f.) is denoted by* $G_i(t) = \mathbb{P}(\tau_i \leq t)$, *also* $G_i(0^+) = 0$.

(b) *The number of particles in the offspring of a type* i, $i \in \mathbb{W}$, *particle of age* a *is modeled by a (discrete) r.v.* $\nu_i(a)$. *We denote with* $p_{ik}(a)$ *the probability that a type* i *particle of age* a *has* k, $k \in \mathbb{N}_0$, *offspring particles (regardless of their type). Thus,* $\nu_i(a)$ *is specified by given* $\left\{p_{ik}(a)\right\}_{k=0}^{\infty}$, $\sum_{k=0}^{\infty} p_{ik}(a) = 1$. *We denote the corresponding p.g.f. of* $\nu_i(a)$ *with* $f_i(a;s) = \mathbb{E}\left[s^{\nu_i(a)}\right] = \sum_{k=0}^{\infty} p_{ik}(a)s^k$, $|s| \leq 1$.

2. *Each daughter particle of a type* i *particle can be of any type* $j \in \mathbb{W}$. *The type of a daughter particle is determined at birth. If* $i \neq j$ *we say that a "mutation" occurs. The probability that a daughter particle of a type* i *particle is a type* j *particle is denoted by* u_{ij}, $u_{ij} \geq 0$, $\sum_{j=1}^{n} u_{ij} = 1$. *Further:*

   (a) *If type* i *cannot have daughters of type* j *we consider the corresponding* u_{ij} *as* $u_{ij} = 0$.

   (b) *Particles are not allowed to change their type within their lifespan.*

3. *All particles from all particle types evolve independently from one another, irrespective of generation.*

4. *Formally* $\left\{\boldsymbol{Z}(t) = \left(Z_1(t), Z_2(t), \ldots, Z_n(t)\right)^\top\right\}_{t\geq 0}$, *where* $\boldsymbol{Z}(t)$ *stands for the MSBPM at* t *and* $Z_i(t)$ *is the number of particles of type* i *that exist at* t.

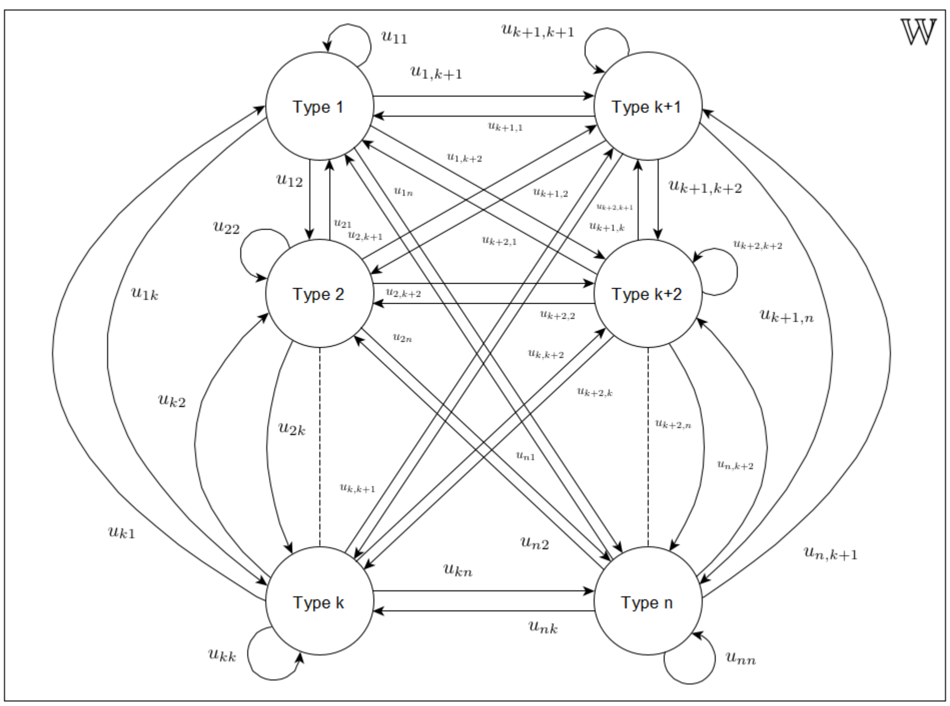

Figure 2.1: A diagram of the MSBPM depicting all possible paths of mutation within the process. Note that some of the u_{ij} may be equal to 0 depending on context.

From Definition 2.1, we can see the connection between the MSBPM and the multi-type Sevastyanov branching process defined in Chapter VIII of [8]. Through $p_{ik}(a)$ and u_{ij}, as specified in Definition 2.1, we can construct an analogue of $p^i_{\boldsymbol{\alpha}}(a)$ (see Chapter VIII of [8] page 229 or Subsection 1.3.1 from the dissertation) that has the same interpretation. This is done by setting $\sum_{j=1}^n \alpha_j = k$ and $p^i_{\boldsymbol{\alpha}}(a) := p_{ik}(a)\frac{k!}{\alpha_1!\ldots\alpha_n!}u_{i1}^{\alpha_1}\ldots u_{in}^{\alpha_n}$. For more details about the nature of the connection between the two models, the reader is referred to the discussion following Definition 2.1 in the dissertation.

2.2.2 Probability generating functions for the MSBPM

Definition 2.2. *We denote the p.g.f. of a MSBPM, starting with one particle of type i, $i \in \mathbb{W}$, that is of age 0, with:*

$$F_i(t;\boldsymbol{s}) = \mathbb{E}\Big(\prod_{j\in\mathbb{W}} s_j^{Z_j(t)} \mid \boldsymbol{Z}(0) = \boldsymbol{\delta}^i\Big),$$

where $|\boldsymbol{s}| \le 1$. We denote the p.g.f. of a MSBPM, starting with one particle of type i, $i \in \mathbb{W}$, that is of age a, $a \neq 0$, with

$$F_{i,a}(t;\boldsymbol{s}) = \mathbb{E}\Big(\prod_{j\in\mathbb{W}} s_j^{Z_j(t)} \mid \boldsymbol{Z}(0) = \boldsymbol{\delta}^i_a\Big),$$

where $|\boldsymbol{s}| \le 1$.

Theorem 2.1. *The following system of integral equations holds for the MSBPM, $i \in \mathbb{W}$:*

$$F_i(t;\,\boldsymbol{s}) = s_i\big(1 - G_i(t)\big) + \int_0^t f_i\Big(y; \sum_{r\in\mathbb{W}} u_{ir}F_r(t-y;\,\boldsymbol{s})\Big)dG_i(y). \tag{2.1}$$

Corollary 2.1. *Let a MSBPM start with one particle of type i, $i \in \mathbb{W}$, that is of age a, $a \neq 0$. Then*

$$F_{i,a}(t;\,\boldsymbol{s}) = s_i\big(1 - G_{i,a}(t)\big) + \int_0^t f_i\Big(a+y; \sum_{r\in\mathbb{W}} u_{ir}F_r(t-y;\,\boldsymbol{s})\Big)dG_{i,a}(y). \tag{2.2}$$

2.2.3 Probabilities of extinction for the MSBPM

Definition 2.3. *We define the probability of extinction until time t of a MSBPM beginning with one particle of type i, $i \in \mathbb{W}$, that is of age 0, as:*

$$q_i(t) = \mathbb{P}\left(\sum_{j\in\mathbb{W}} Z_j(t) = 0 \mid \boldsymbol{Z}(0) = \boldsymbol{\delta}^i\right).$$

Similarly, if the initial particle is of age a, $a \neq 0$:

$$q_{i,a}(t) = \mathbb{P}\left(\sum_{j\in\mathbb{W}} Z_j(t) = 0 \mid \boldsymbol{Z}(0) = \boldsymbol{\delta}^i_a\right).$$

Theorem 2.2. *The following system of integral equations holds for the MSBPM, $i \in \mathbb{W}$:*

$$q_i(t) = \int_0^t f_i\Big(y; \sum_{r\in\mathbb{W}} u_{ir} q_r(t-y)\Big) dG_i(y). \tag{2.5}$$

Corollary 2.2. *The following system of integral equations holds for the MSBPM, $i \in \mathbb{W}$:*

$$q_{i,a}(t) = \int_0^t f_i\Big(a + y; \sum_{r\in\mathbb{W}} u_{ir} q_r(t-y)\Big) dG_{i,a}(y). \tag{2.6}$$

Definition 2.4. *We denote the probability of extinction of a MSBPM beginning with one particle of type i, $i \in \mathbb{W}$, that is of age 0, as:*

$$q_i = \mathbb{P}\Bigg(\sum_{j\in\mathbb{W}} Z_j(t) = 0 \text{ for some } t > 0 \mid \boldsymbol{Z}(0) = \boldsymbol{\delta}^i\Bigg).$$

Similarly, if the initial particle is of age a, $a \neq 0$:

$$q_{i,a} = \mathbb{P}\Bigg(\sum_{j\in\mathbb{W}} Z_j(t) = 0 \text{ for some } t > 0 \mid \boldsymbol{Z}(0) = \boldsymbol{\delta}^i_a\Bigg).$$

Definition 2.5. *We define $\lim\limits_{y\to\infty}\lim\limits_{t\to\infty} q_i(t-y) = q_i$ and $\lim\limits_{y\to\infty}\lim\limits_{t\to\infty} q_{i,a}(t-y) = q_{i,a}$.*

Definition 2.5 is necessary for eliminating the ambiguity of the expression $\lim\limits_{y\to\infty}\lim\limits_{t\to\infty} q_i(t-y)$. The alternative to Definition 2.5 is to set $\lim\limits_{y\to\infty}\lim\limits_{t\to\infty} q_i(t-y) = q_i(b)$, where the choice of b can be arbitrary.

Theorem 2.3. *The following system of integral equations holds for the MSBPM, $i \in \mathbb{W}$:*

$$q_i = \int_0^\infty f_i\Big(y; \sum_{r\in\mathbb{W}} u_{ir} q_r\Big) dG_i(y). \tag{2.7}$$

Corollary 2.3. *The following system of integral equations holds for the MSBPM, $i \in \mathbb{W}$:*

$$q_{i,a} = \int_0^\infty f_i\Big(a + y; \sum_{r\in\mathbb{W}} u_{ir} q_r\Big) dG_{i,a}(y). \tag{2.8}$$

Corollary 2.4. *In the particular case where there is no dependence of particle reproduction from particle age, i.e., $f_i(y; s) = f_i(s)$, $i \in \mathbb{W}$, the systems of integral equations (2.7) and (2.8) become the following system of equations, $i \in \mathbb{W}$:*

$$q_i = q_{i,a} = f_i\Big(\sum_{r\in\mathbb{W}} u_{ir} q_r\Big). \tag{2.9}$$

2.2.4 Number of particles produced from $\mathbb{W}_e$ towards $\mathbb{W}$ within the MSBPM

Let $\mathbb{W}_e \subseteq \mathbb{W}$ be a subset of types within the MSBPM (the subscript "e" stands for "emit"). Note that we allow $\mathbb{W}_e = \mathbb{W}$. The number of occurred mutations from $\mathbb{W}_e$ towards types in $\mathbb{W} \setminus \mathbb{W}_e$ is a crucial quantity in the context of populations escaping extinction as the types that have supercritical reproduction are usually modeled to be outside of $\mathbb{W}_e$. We investigate the production of mutants from $\mathbb{W}_e \subset \mathbb{W}$ towards $\mathbb{W}_0 = \mathbb{W} \setminus \mathbb{W}_e$ in Subsection 2.3.1 and Subsection 2.3.2. In the current Subsection, we derive more general results. These general results concern general particle production, that is, the particles produced from $\mathbb{W}_e$ can be of any type within $\mathbb{W}$ and are not necessarily mutants. Particular cases of these results that correspond to particle production from $\mathbb{W}_e$ towards any subclass of $\mathbb{W}$ within the MSBPM, can be straightforwardly obtained by appropriately setting $u_{ij} = 0$, appropriately setting coordinates of $\boldsymbol{s}$ ot 1, and realizing in some occasions that $s^X = 1$ due to a relevant random variable X being always 0.

Definition 2.6. *Denote with $I_j^{\mathbb{W}_e}(t)$ the number of particles (mutants or not) of type j, $j \in \mathbb{W}$, produced from particles with types from $\mathbb{W}_e$ until t within a MSBPM. We do not count the initial particle within any of the $I_j^{\mathbb{W}_e}(t)$. For a MSBPM starting with one particle of type i, $i \in \mathbb{W}$, that is of age 0, denote with $h_i^{\mathbb{W}_e}(t;\boldsymbol{s})$ the following p.g.f.*

$$h_i^{\mathbb{W}_e}(t;\boldsymbol{s}) = \mathbb{E}\Big(\prod_{j\in\mathbb{W}} s_j^{I_j^{\mathbb{W}_e}(t)} \mid \boldsymbol{Z}(0) = \boldsymbol{\delta}^i\Big),$$

where $|\boldsymbol{s}| \leq 1$. We denote the corresponding p.g.f., when the MSBPM starts with one particle of type i, $i \in \mathbb{W}$, that is of age a, $a \neq 0$, with

$$h_{i,a}^{\mathbb{W}_e}(t;\boldsymbol{s}) = \mathbb{E}\Big(\prod_{j\in\mathbb{W}} s_j^{I_j^{\mathbb{W}_e}(t)} \mid \boldsymbol{Z}(0) = \boldsymbol{\delta}_a^i\Big),$$

where $|\boldsymbol{s}| \leq 1$.

We note that unlike $F_i(t;\boldsymbol{s})$, which compactly contain information about the number of particles, per type, that exist *at* t, $h_i^{\mathbb{W}_e}(t;\boldsymbol{s})$ contain information about the number of particles that have been produced *until* t (with respect to t some of the produced particles may no longer exist).

Theorem 2.4. *The following system of integral equations holds within the MSBPM:*

1. *For $i \in \mathbb{W}_e$*

$$h_i^{\mathbb{W}_e}(t;\boldsymbol{s}) = \big(1 - G_i(t)\big) + \int_0^t f_i\Big(y; \sum_{r\in\mathbb{W}} u_{ir} s_r h_r^{\mathbb{W}_e}(t-y;\boldsymbol{s})\Big) dG_i(y). \tag{2.10}$$

2. *For $i \notin \mathbb{W}_e$*

$$h_i^{\mathbb{W}_e}(t;\boldsymbol{s}) = \big(1 - G_i(t)\big) + \int_0^t f_i\Big(y; \sum_{r\in\mathbb{W}} u_{ir} h_r^{\mathbb{W}_e}(t-y;\boldsymbol{s})\Big) dG_i(y). \tag{2.11}$$

Corollary 2.5. *Let $\mathbb{W}_e = \mathbb{W}$. The following system of integral equations hold within the MSBPM, $i \in \mathbb{W}$:*

$$h_i^{\mathbb{W}}(t;\boldsymbol{s}) = \big(1 - G_i(t)\big) + \int_0^t f_i\Big(y; \sum_{r\in\mathbb{W}} u_{ir} s_r h_r^{\mathbb{W}}(t-y;\boldsymbol{s})\Big) dG_i(y). \tag{2.12}$$

Corollary 2.6. *Let a MSBPM start with one particle of type i, $i \in \mathbb{W}$, that is of age a, $a \neq 0$. Then the following system of integral equations holds:*

1. *For $i \in \mathbb{W}_e$*

$$h_{i,a}^{\mathbb{W}_e}(t;\boldsymbol{s}) = \big(1 - G_{i,a}(t)\big) + \int_0^t f_i\Big(a+y; \sum_{r\in\mathbb{W}} u_{ir} s_r h_r^{\mathbb{W}_e}(t-y;\boldsymbol{s})\Big) dG_{i,a}(y). \tag{2.13}$$

2. *For $i \notin \mathbb{W}_e$*

$$h_{i,a}^{\mathbb{W}_e}(t;\boldsymbol{s}) = \big(1 - G_{i,a}(t)\big) + \int_0^t f_i\Big(a+y; \sum_{r\in\mathbb{W}} u_{ir} h_r^{\mathbb{W}_e}(t-y;\boldsymbol{s})\Big) dG_{i,a}(y). \tag{2.14}$$

Corollary 2.7. *Let $\mathbb{W}_e = \mathbb{W}$. The following system of integral equations hold within the MSBPM, $i \in \mathbb{W}$:*

$$h_{i,a}^{\mathbb{W}}(t;\boldsymbol{s}) = \big(1 - G_{i,a}(t)\big) + \int_0^t f_i\Big(a+y; \sum_{r\in\mathbb{W}} u_{ir} s_r h_r^{\mathbb{W}}(t-y;\boldsymbol{s})\Big) dG_{i,a}(y). \tag{2.15}$$

Next, we investigate $I_j^{\mathbb{W}_e}(t)$ and $h_i^{\mathbb{W}_e}(t;\boldsymbol{s})$ as $t \to \infty$.

Definition 2.7. *Denote with $I_j^{\mathbb{W}_e}$ the number of particles (mutants or not) of type j, $j \in \mathbb{W}$, produced from particles with types from $\mathbb{W}_e$ during the whole MSBPM. We do not count the initial particle within any of the $I_j^{\mathbb{W}_e}$. For a MSBPM starting with one particle of type i, $i \in \mathbb{W}$, that is of age 0, denote with $h_i^{\mathbb{W}_e}(\boldsymbol{s})$ the following p.g.f.*

$$h_i^{\mathbb{W}_e}(\boldsymbol{s}) = \mathbb{E}\Big(\prod_{j\in\mathbb{W}} s_j^{I_j^{\mathbb{W}_e}} \mid \boldsymbol{Z}(0) = \boldsymbol{\delta}^i\Big),$$

where $|\boldsymbol{s}| \leq 1$. We denote the corresponding p.g.f., when the MSBPM starts with one particle of type i, $i \in \mathbb{W}$, that is of age a, $a \neq 0$, with

$$h_{i,a}^{\mathbb{W}_e}(\boldsymbol{s}) = \mathbb{E}\Big(\prod_{j\in\mathbb{W}} s_j^{I_j^{\mathbb{W}_e}} \mid \boldsymbol{Z}(0) = \boldsymbol{\delta}_a^i\Big),$$

where $|\boldsymbol{s}| \leq 1$.

Remark 2.9. *From Definition 2.7 it is evident that $I_j^{\mathbb{W}_e} := \lim_{t\to\infty} I_j^{\mathbb{W}_e}(t)$ almost surely. Considering this and the fact that there is a one-to-one correspondence between r.v.s and p.g.f.s, it follows that $h_i^{\mathbb{W}_e}(\boldsymbol{s}) = \lim_{t\to\infty} h_i^{\mathbb{W}_e}(t;\boldsymbol{s})$ and $h_{i,a}^{\mathbb{W}_e}(\boldsymbol{s}) = \lim_{t\to\infty} h_{i,a}^{\mathbb{W}_e}(t;\boldsymbol{s})$.*

Definition 2.8. *We define* $\lim_{y\to\infty}\lim_{t\to\infty} I_j^{\mathbb{W}_e}(t-y) = I_j^{\mathbb{W}_e}$. *Consequently* $\lim_{y\to\infty}\lim_{t\to\infty} h_i^{\mathbb{W}_e}(t-y;\boldsymbol{s}) = h_i^{\mathbb{W}_e}(\boldsymbol{s})$ *and* $\lim_{y\to\infty}\lim_{t\to\infty} h_{i,a}^{\mathbb{W}_e}(t-y;\boldsymbol{s}) = h_{i,a}^{\mathbb{W}_e}(\boldsymbol{s})$.

Definition 2.8 is necessary for eliminating the ambiguity of the expression $\lim_{y\to\infty}\lim_{t\to\infty} I_j^{\mathbb{W}_e}(t-y)$. The alternative to Definition 2.8 is to set $\lim_{y\to\infty}\lim_{t\to\infty} I_j^{\mathbb{W}_e}(t-y) = I_j^{\mathbb{W}_e}(b)$, where the choice of b can be arbitrary.

Theorem 2.5. *The following system of equations holds within the MSBPM,* $i \in \mathbb{W}$*:*

1. *Let* $i \in \mathbb{W}_e$. *Then*

$$h_i^{\mathbb{W}_e}(\boldsymbol{s}) = \int_0^\infty f_i\Big(y; \sum_{r\in\mathbb{W}} u_{ir} s_r h_r^{\mathbb{W}_e}(\boldsymbol{s})\Big) dG_i(y). \tag{2.18}$$

2. *Let* $i \notin \mathbb{W}_e$. *Then*

$$h_i^{\mathbb{W}_e}(\boldsymbol{s}) = \int_0^\infty f_i\Big(y; \sum_{r\in\mathbb{W}} u_{ir} h_r^{\mathbb{W}_e}(\boldsymbol{s})\Big) dG_i(y). \tag{2.19}$$

Corollary 2.8. *Let* $\mathbb{W}_e = \mathbb{W}$. *The following system of integral equations holds within the MSBPM,* $i \in \mathbb{W}$*:*

$$h_i^{\mathbb{W}}(\boldsymbol{s}) = \int_0^\infty f_i\Big(y; \sum_{r\in\mathbb{W}} u_{ir} s_r h_r^{\mathbb{W}}(\boldsymbol{s})\Big) dG_i(y). \tag{2.20}$$

Corollary 2.9. *Let a MSBPM start with one particle of type* i, $i \in \mathbb{W}$, *that is of age* a, $a \neq 0$. *Then the following system of integral equations holds:*

1. *Let* $i \in \mathbb{W}_e$. *Then*

$$h_{i,a}^{\mathbb{W}_e}(\boldsymbol{s}) = \int_0^\infty f_i\Big(a+y; \sum_{r\in\mathbb{W}} u_{ir} s_r h_r^{\mathbb{W}_e}(\boldsymbol{s})\Big) dG_{i,a}(y). \tag{2.21}$$

2. *Let* $i \notin \mathbb{W}_e$. *Then*

$$h_{i,a}^{\mathbb{W}_e}(\boldsymbol{s}) = \int_0^\infty f_i\Big(a+y; \sum_{r\in\mathbb{W}} u_{ir} h_r^{\mathbb{W}_e}(\boldsymbol{s})\Big) dG_{i,a}(y). \tag{2.22}$$

Corollary 2.10. *Let* $\mathbb{W}_e = \mathbb{W}$. *The following system of integral equations holds within the MSBPM,* $i \in \mathbb{W}$*:*

$$h_{i,a}^{\mathbb{W}}(\boldsymbol{s}) = \int_0^\infty f_i\Big(a+y; \sum_{r\in\mathbb{W}} u_{ir} s_r h_r^{\mathbb{W}}(\boldsymbol{s})\Big) dG_{i,a}(y). \tag{2.23}$$

Corollary 2.11. *In the particular case where there is no dependence of particle reproduction from particle age, i.e., $f_i(y;s) = f_i(s)$, $i \in \mathbb{W}$, the system of integral equations* (2.18), (2.19) *from Theorem 2.5, and the system of integral equations* (2.21), (2.22) *from Corollary 2.9, become the following system of equations:*

1. *Let $i \in \mathbb{W}_e$. Then*

$$h_i^{\mathbb{W}_e}(\boldsymbol{s}) = h_{i,a}^{\mathbb{W}_e}(\boldsymbol{s}) = f_i\Big(\sum_{r\in\mathbb{W}} u_{ir} s_r h_r^{\mathbb{W}_e}(\boldsymbol{s})\Big). \tag{2.24}$$

2. *Let $i \notin \mathbb{W}_e$. Then*

$$h_i^{\mathbb{W}_e}(\boldsymbol{s}) = h_{i,a}^{\mathbb{W}_e}(\boldsymbol{s}) = f_i\Big(\sum_{r\in\mathbb{W}} u_{ir} h_r^{\mathbb{W}_e}(\boldsymbol{s})\Big). \tag{2.25}$$

2.2.5 Time until occurrence of the first "successful" particle produced from $\mathbb{W}_e$ towards $\mathbb{W}$ within the MSBPM

We call a particle produced from $\mathbb{W}_e$ towards $\mathbb{W}$ "successful" if it initiates a non-extincting MSBPM.

Definition 2.9. *Denote with $T_{\mathbb{W}}^{\mathbb{W}_e}$ the r.v. that is the time until occurrence of the first "successful" particle produced from a type within $\mathbb{W}_e$ towards a type within $\mathbb{W}$ in a MSBPM starting with some combination of particles with types within $\mathbb{W}_e$. Without loss of generality, we set the starting number of particles per type $r \in \mathbb{W}_e$ to be k_r and the starting number of particles per type $r \in \mathbb{W} \setminus \mathbb{W}_e$ to be 0. We denote the so specified initial state of the process as $\boldsymbol{Z}(0) = \boldsymbol{\alpha}^*$. We define $T_{\mathbb{W}}^{\mathbb{W}_e} = \infty$ as the event that no "successful" particles have been produced from $\mathbb{W}_e$ towards $\mathbb{W}$ in a MSBPM beginning with an initial state $\boldsymbol{\alpha}^*$. Thus, we may write $T_{\mathbb{W}}^{\mathbb{W}_e} \in (0,\infty]$. If the MSBPM starts with a single particle of type i, $i \in \mathbb{W}_e$, of age 0, we use $T_{\mathbb{W},i}^{\mathbb{W}_e}$ as a shortcut notation. If the initial particle is of age a, $a \neq 0$, we use $T_{\mathbb{W},i,a}^{\mathbb{W}_e}$.*

Theorem 2.6. *Let the MSBPM start with k_r particles per type r, $r \in \mathbb{W}_e$. Let all particles form $\boldsymbol{\alpha}^*$ have age 0. The distribution of $T_{\mathbb{W}}^{\mathbb{W}_e}$ has the following properties:*

(i) $\mathbb{P}\Big(T_{\mathbb{W}}^{\mathbb{W}_e} > t \mid \boldsymbol{Z}(0) = \boldsymbol{\alpha}^*\Big) = \prod_{r\in\mathbb{W}_e} \Big[h_r^{\mathbb{W}_e}(t;\boldsymbol{q})\Big]^{k_r}.$

(ii) $\mathbb{P}\Big(T_{\mathbb{W}}^{\mathbb{W}_e} = \infty \mid \boldsymbol{Z}(0) = \boldsymbol{\alpha}^*\Big) = \prod_{r\in\mathbb{W}_e} \Big[h_r^{\mathbb{W}_e}(\boldsymbol{q})\Big]^{k_r}.$

(iii) If at least one particle type within $\mathbb{W}$ is supercritical, we have

$$\mathbb{E}\Big[T_{\mathbb{W}}^{\mathbb{W}_e} \mid T_{\mathbb{W}}^{\mathbb{W}_e} < \infty,\ \boldsymbol{Z}(0) = \boldsymbol{\alpha}^*\Big] =$$
$$= \frac{1}{1 - \prod_{r\in\mathbb{W}_e}\Big[h_r^{\mathbb{W}_e}(\boldsymbol{q})\Big]^{k_r}} \int_0^\infty \Bigg[\prod_{r\in\mathbb{W}_e}\Big[h_r^{\mathbb{W}_e}(t;\boldsymbol{q})\Big]^{k_r} - \prod_{r\in\mathbb{W}_e}\Big[h_r^{\mathbb{W}_e}(\boldsymbol{q})\Big]^{k_r}\Bigg] dt,$$

if not, then the expectation does not exist.

Theorem 2.7. *Let the MSBPM start with k_r particles per type r, $r \in \mathbb{W}_e$, let the starting particles in $\boldsymbol{\alpha}^*$ have ages $a_{r,c}$, $c \in \{1, 2, \dots, k_r\}$, where $a_{r,c}$ is the age of the c-th particle of type r. We allow $a_{r,c}$ to be 0. The distribution of $T_{\mathbb{W}}^{\mathbb{W}_e}$ has the following properties:*

(i) $\mathbb{P}\Big(T_{\mathbb{W}}^{\mathbb{W}_e} > t \mid \boldsymbol{Z}(0) = \boldsymbol{\alpha}^*\Big) = \prod_{r\in\mathbb{W}_e}\Big[\prod_{c=1}^{k_r} h_{r,a_{r,c}}^{\mathbb{W}_e}(t;\boldsymbol{q})\Big]$.

(ii) $\mathbb{P}\Big(T_{\mathbb{W}}^{\mathbb{W}_e} = \infty \mid \boldsymbol{Z}(0) = \boldsymbol{\alpha}^*\Big) = \prod_{r\in\mathbb{W}_e}\Big[\prod_{c=1}^{k_r} h_{r,a_{r,c}}^{\mathbb{W}_e}(\boldsymbol{q})\Big]$.

(iii) If at least one particle type within $\mathbb{W}$ is supercritical, we have

$$\mathbb{E}\Big[T_{\mathbb{W}}^{\mathbb{W}_e} \mid T_{\mathbb{W}}^{\mathbb{W}_e} < \infty,\ \boldsymbol{Z}(0) = \boldsymbol{\alpha}^*\Big] =$$
$$= \frac{1}{1 - \prod_{r\in\mathbb{W}_e}\Big[\prod_{c=1}^{k_r} h_{r,a_{r,c}}^{\mathbb{W}_e}(\boldsymbol{q})\Big]}\int_0^\infty \Bigg[\prod_{r\in\mathbb{W}_e}\Big[\prod_{c=1}^{k_r} h_{r,a_{r,c}}^{\mathbb{W}_e}(t;\boldsymbol{q})\Big] - \prod_{r\in\mathbb{W}_e}\Big[\prod_{c=1}^{k_r} h_{r,a_{r,c}}^{\mathbb{W}_e}(\boldsymbol{q})\Big]\Bigg]dt,$$

if not, then the expectation does not exist.

2.2.6 Immediate risk of producing a "successful" particle from $\mathbb{W}_e$ towards $\mathbb{W}$ within the MSBPM

We will study the immediate risk of escape facilitated by the particles with types within $\mathbb{W}_e$ via the following hazard function:

Definition 2.10. *Define for an initial particle of type i, $i \in \mathbb{W}_e$, the following hazard function:*

1. *If the initial particle is of age 0*

$$g_{\mathbb{W},i}^{\mathbb{W}_e}(t)dt = \mathbb{P}\Big(T_{\mathbb{W},i}^{\mathbb{W}_e} \in (t, t+dt] \mid T_{\mathbb{W},i}^{\mathbb{W}_e} > t\Big). \tag{2.26}$$

2. *If the initial particle if of age a, $a \neq 0$*

$$g_{\mathbb{W},i,a}^{\mathbb{W}_e}(t)dt = \mathbb{P}\Big(T_{\mathbb{W},i,a}^{\mathbb{W}_e} \in (t, t+dt] \mid T_{\mathbb{W},i,a}^{\mathbb{W}_e} > t\Big). \tag{2.27}$$

It is clear that for $i \in \mathbb{W}_e$

$$g_{\mathbb{W},i}^{\mathbb{W}_e}(t)dt = \frac{\mathbb{P}\Big(T_{\mathbb{W},i}^{\mathbb{W}_e} \in (t, t+dt],\, T_{\mathbb{W},i}^{\mathbb{W}_e} > t\Big)}{\mathbb{P}\Big(T_{\mathbb{W},i}^{\mathbb{W}_e} > t\Big)}.$$

Thus,

$$g_{\mathbb{W},i}^{\mathbb{W}_e}(t) = \frac{F^{(1)}_{T_{\mathbb{W},i}^{\mathbb{W}_e}}(t)}{\mathbb{P}\Big(T_{\mathbb{W},i}^{\mathbb{W}_e} > t\Big)}, \tag{2.28}$$

where $F^{(1)}_{T^{\mathbb{W}_e}_{\mathbb{W},i}}(t)$ is the probability density function of $T^{\mathbb{W}_e}_{\mathbb{W},i}$. Analogously, we can obtain

$$g^{\mathbb{W}_e}_{\mathbb{W},i,a}(t) = \frac{F^{(1)}_{T^{\mathbb{W}_e}_{\mathbb{W},i.a}}(t)}{\mathbb{P}\Big(T^{\mathbb{W}_e}_{\mathbb{W},i,a} > t\Big)}. \tag{2.29}$$

2.2.7 Numerical schemes for computing obtained systems of integral equations for the MSBPM

We organize the integral equations obtained so far into two tables. Table 2.1 contains all integral equations that are for a MSBPM starting with a particle of age 0. We also put into Table 2.1 the result of Theorem 2.8, i.e., equation (2.53), as it conforms to the same pattern. Let us denote $B_i(t;\boldsymbol{s}) = \int_0^t f_i\Big(y;\, C_i(t-y;\boldsymbol{s})\Big)dG_i(y)$, where $C_i(t-y;\boldsymbol{s})$ is the corresponding second argument of f_i with respect to the entry of interest in Table 2.1. Through Numerical Scheme 1, outlined below, we provide a general numerical method immediately applicable to the integral equations listed. We note that Numerical Scheme 1 can trace its origin to [2], where a model consisting of two particle types is discussed.

Eq.	$L_i(t;\boldsymbol{s})$		$A_i(t;\boldsymbol{s})$		$B_i(t;\boldsymbol{s})$	
(2.1)	$F_i(t;\,\boldsymbol{s})$	$=$	$s_i\big(1-G_i(t)\big)$	$+$	$\int_0^t f_i\Big(y;\sum_{r\in\mathbb{W}} u_{ir}F_r(t-y;\,\boldsymbol{s})\Big)dG_i(y)$	$i\in\mathbb{W}$
(2.5)	$q_i(t)$	$=$	0	$+$	$\int_0^t f_i\Big(y;\sum_{r\in\mathbb{W}} u_{ir}q_r(t-y)\Big)dG_i(y)$	$i\in\mathbb{W}$
(2.7)	q_i	$=$	0	$+$	$\int_0^\infty f_i\Big(y;\sum_{r\in\mathbb{W}} u_{ir}q_r\Big)dG_i(y)$	$i\in\mathbb{W}$
(2.10)	$h_i^{\mathbb{W}_e}(t;\boldsymbol{s})$	$=$	$\big(1-G_i(t)\big)$	$+$	$\int_0^t f_i\Big(y;\sum_{r\in\mathbb{W}} u_{ir}s_rh_r^{\mathbb{W}_e}(t-y;\boldsymbol{s})\Big)dG_i(y)$	$i\in\mathbb{W}_e$
(2.11)	$h_i^{\mathbb{W}_e}(t;\boldsymbol{s})$	$=$	$\big(1-G_i(t)\big)$	$+$	$\int_0^t f_i\Big(y;\sum_{r\in\mathbb{W}} u_{ir}h_r^{\mathbb{W}_e}(t-y;\boldsymbol{s})\Big)dG_i(y)$	$i\notin\mathbb{W}_e$
(2.12)	$h_i^{\mathbb{W}}(t;\boldsymbol{s})$	$=$	$\big(1-G_i(t)\big)$	$+$	$\int_0^t f_i\Big(y;\sum_{r\in\mathbb{W}} u_{ir}s_rh_r^{\mathbb{W}}(t-y;\boldsymbol{s})\Big)dG_i(y)$	$i\in\mathbb{W}$
(2.18)	$h_i^{\mathbb{W}_e}(\boldsymbol{s})$	$=$	0	$+$	$\int_0^\infty f_i\Big(y;\sum_{r\in\mathbb{W}} u_{ir}s_rh_r^{\mathbb{W}_e}(\boldsymbol{s})\Big)dG_i(y)$	$i\in\mathbb{W}_e$
(2.19)	$h_i^{\mathbb{W}_e}(\boldsymbol{s})$	$=$	0	$+$	$\int_0^\infty f_i\Big(y;\sum_{r\in\mathbb{W}} u_{ir}h_r^{\mathbb{W}_e}(\boldsymbol{s})\Big)dG_i(y)$	$i\notin\mathbb{W}_e$
(2.20)	$h_i^{\mathbb{W}}(\boldsymbol{s})$	$=$	0	$+$	$\int_0^\infty f_i\Big(y;\sum_{r\in\mathbb{W}} u_{ir}s_rh_r^{\mathbb{W}}(\boldsymbol{s})\Big)dG_i(y)$	$i\in\mathbb{W}$
(2.53)	$V_i(t)$	$=$	0	$+$	$\int_0^t f_i\Bigg(y;\Big[\sum_{m\in\mathbb{W}_e} u_{im}V_m(t-y)\Big]+\Big[\sum_{r\in\mathbb{W}_0} u_{ir}q_r\Big]\Bigg)dG_i(y)$	$i\in\mathbb{W}_e$

Table 2.1: Systems of integral equations for the case of a MSBPM starting with a particle of age 0.

Numerical Scheme 1. *Let $L_i(t;\boldsymbol{s})$ be from Table 2.1. The corresponding system of integral equations can be numerically computed via the following steps:*

1. *Let $t=0$. For every i that participates in the corresponding system of integral equations, compute the initial point $L_i(0;\boldsymbol{s}) = A_i(0;\boldsymbol{s})$.*

2. *Let $t = kh$, $k = 1, 2, \dots,$ where h is the chosen step size. For every i that participates in the corresponding system of integral equations compute*

$$L_i(kh; \boldsymbol{s}) \approx A_i(kh; \boldsymbol{s}) + \sum_{j=1}^{k} f_i\Big(jh;\, C_i\big((k-j)h; \boldsymbol{s}\big)\Big) \cdot \Big(G_i\big(jh\big) - G_i\big((j-1)h\big)\Big).$$

Our second table, Table 2.2, contains all integral equations that are obtained for a MSBPM starting with a particle of age a, $a \neq 0$. We also put into Table 2.2 the result of Corollary 2.22, i.e., equation (2.54), as it conforms to the same pattern. We denote $B_{i,a}(t; \boldsymbol{s}) = \int_0^t f_i\Big(a+y;\, C_i(t-y; \boldsymbol{s})\Big) dG_{i,a}(y)$, however, we stress that all $C_i(t-y; \boldsymbol{s})$ remain as in Table 2.1. Numerical Scheme 2, outlined below, is applicable to all integral equations listed within Table 2.2.

Eq.	$L_{i,a}(t; \boldsymbol{s})$		$A_{i,a}(t; \boldsymbol{s})$		$B_{i,a}(t; \boldsymbol{s})$	
(2.2)	$F_{i,a}(t; \boldsymbol{s})$	$=$	$s_i\big(1 - G_{i,a}(t)\big)$	$+$	$\int_0^t f_i\Big(a+y; \sum_{r\in\mathbb{W}} u_{ir} F_r(t-y; \boldsymbol{s})\Big) dG_{i,a}(y)$	$i \in \mathbb{W}$
(2.6)	$q_{i,a}(t)$	$=$	0	$+$	$\int_0^t f_i\Big(a+y; \sum_{r\in\mathbb{W}} u_{ir} q_r(t-y)\Big) dG_{i,a}(y)$	$i \in \mathbb{W}$
(2.8)	$q_{i,a}$	$=$	0	$+$	$\int_0^\infty f_i\Big(a+y; \sum_{r\in\mathbb{W}} u_{ir} q_r\Big) dG_{i,a}(y)$	$i \in \mathbb{W}$
(2.13)	$h_{i,a}^{\mathbb{W}_e}(t; \boldsymbol{s})$	$=$	$\big(1 - G_{i,a}(t)\big)$	$+$	$\int_0^t f_i\Big(a+y; \sum_{r\in\mathbb{W}} u_{ir} s_r h_r^{\mathbb{W}_e}(t-y; \boldsymbol{s})\Big) dG_{i,a}(y)$	$i \in \mathbb{W}_e$
(2.14)	$h_{i,a}^{\mathbb{W}_e}(t; \boldsymbol{s})$	$=$	$\big(1 - G_{i,a}(t)\big)$	$+$	$\int_0^t f_i\Big(a+y; \sum_{r\in\mathbb{W}} u_{ir} h_r^{\mathbb{W}_e}(t-y; \boldsymbol{s})\Big) dG_{i,a}(y)$	$i \notin \mathbb{W}_e$
(2.15)	$h_{i,a}^{\mathbb{W}}(t; \boldsymbol{s})$	$=$	$\big(1 - G_{i,a}(t)\big)$	$+$	$\int_0^t f_i\Big(a+y; \sum_{r\in\mathbb{W}} u_{ir} s_r h_r^{\mathbb{W}}(t-y; \boldsymbol{s})\Big) dG_{i,a}(y)$	$i \in \mathbb{W}$
(2.21)	$h_{i,a}^{\mathbb{W}_e}(\boldsymbol{s})$	$=$	0	$+$	$\int_0^\infty f_i\Big(a+y; \sum_{r\in\mathbb{W}} u_{ir} s_r h_r^{\mathbb{W}_e}(\boldsymbol{s})\Big) dG_{i,a}(y)$	$i \in \mathbb{W}_e$
(2.22)	$h_{i,a}^{\mathbb{W}_e}(\boldsymbol{s})$	$=$	0	$+$	$\int_0^\infty f_i\Big(a+y; \sum_{r\in\mathbb{W}} u_{ir} h_r^{\mathbb{W}_e}(\boldsymbol{s})\Big) dG_{i,a}(y)$	$i \notin \mathbb{W}_e$
(2.23)	$h_{i,a}^{\mathbb{W}}(\boldsymbol{s})$	$=$	0	$+$	$\int_0^\infty f_i\Big(a+y; \sum_{r\in\mathbb{W}} u_{ir} s_r h_r^{\mathbb{W}}(\boldsymbol{s})\Big) dG_{i,a}(y)$	$i \in \mathbb{W}$
(2.54)	$V_{i,a}(t)$	$=$	0	$+$	$\int_0^t f_i\left(a+y; \Big[\sum_{m\in\mathbb{W}_e} u_{im} V_m(t-y)\Big] + \Big[\sum_{r\in\mathbb{W}_0} u_{ir} q_r\Big]\right) dG_{i,a}(y)$	$i \in \mathbb{W}_e$

Table 2.2: Systems of integral equations for the case of a MSBPM starting with a particle of age a, $a \neq 0$.

Numerical Scheme 2. *Let $L_{i,a}(t; \boldsymbol{s})$ be from Table 2.2. The corresponding system of integral equations can be numerically computed via the following steps:*

1. *Let $t = 0$. For every i that participates in the corresponding system of integral equations, compute the initial point $L_{i,a}(0; \boldsymbol{s}) = A_{i,a}(0; \boldsymbol{s})$.*

2. *Let $t = kh$, $k = 1, 2, \dots,$ where h is the chosen step size. For every i that participates in the corresponding system of integral equations compute*

$$L_{i,a}(kh; \boldsymbol{s}) \approx$$
$$\approx A_{i,a}(kh; \boldsymbol{s}) + \sum_{j=1}^{k} f_i\Big(a + jh;\, C_i\big((k-j)h; \boldsymbol{s}\big)\Big) \cdot \Big(G_{i,a}\big(jh\big) - G_{i,a}\big((j-1)h\big)\Big).$$

2.3 Particular cases of the MSBPM

Within this Section, we explore the Decomposable Multi-type Sevastyanov Branching Process through probabilities of Mutation between types (DMSBPM) and the Decomposable Multi-type Bellman-Harris Branching Process through probabilities of Mutation between types (DMBHBPM). The DMSBPM is the process considered within our previous work Vitanov & Slavtchova-Bojkova [7] (2022), while the DMBHBPM is an extension of the process in our work Slavtchova-Bojkova & Vitanov [5] (2019). Both the DMSBPM and the DMBHBPM are particular cases of the MSBPM where decomposability holds.

2.3.1 Decomposable Multi-type Sevastyanov Branching Process through probabilities of Mutation between types (DMSBPM)

Within the current Abstract, we limit ourselves with presenting only core definitions and results. The definition of the DMSBPM is similar to the definition of the MSBPM, the difference being that within the DMSBPM, we consider two classes of types - class $\mathbb{W}_0 \subset \mathbb{W}$ and class $\mathbb{W}_e = \mathbb{W} \setminus \mathbb{W}_0$. Particles with types from $\mathbb{W}_e$ can produce particles with types from $\mathbb{W}$, while particles with types from $\mathbb{W}_0$ can only produce particles with types from $\mathbb{W}_0$. The full formal formulation of the DMSBPM can be seen in Definition 2.11 in Subsection 2.3.1.1 of the dissertation. The DMSBPM can be used for modeling irreversible paths in the evolution of a population escaping extinction.

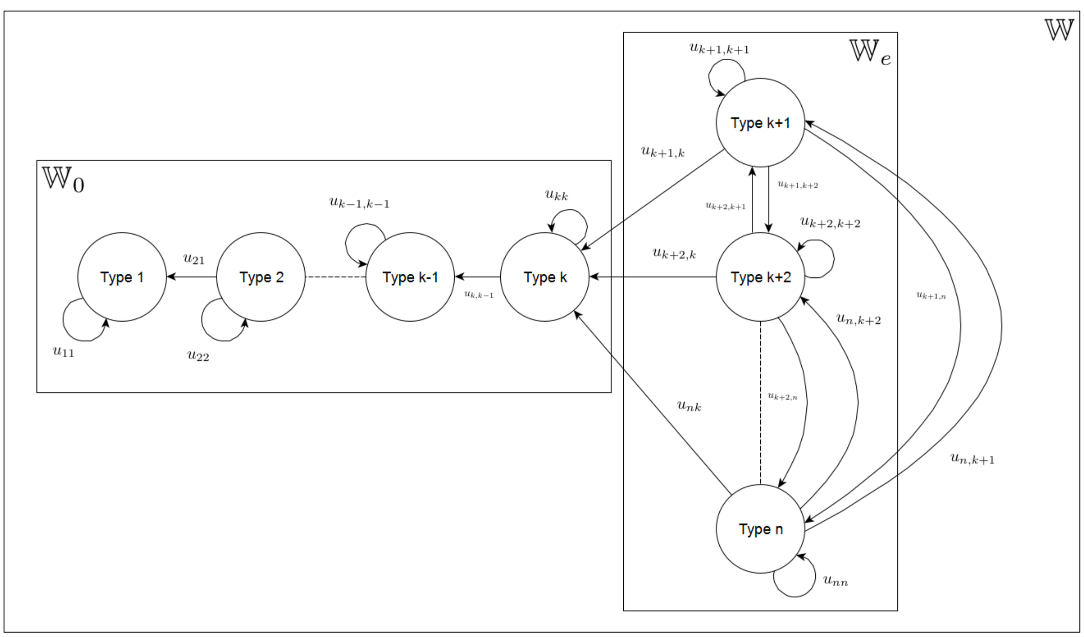

Figure 2.12: A particular case of the DMSBPM where type 1 is reachable only after mutations leading to type k, type $k-1$, ..., type 2 occur (assuming the process begins with particles with types from $\mathbb{W}_e$). A subcase of special interest arises when type 1 is the only supercritical type - types within $\mathbb{W}_e$ can be used to model an existing badly adapted population approaching extinction, while the mutation path towards type 1, provided by the types within $\mathbb{W}_0$, leads to possible escape from extinction.

In order to avoid ambiguity, without loss of generality, we impose the following ordering - if $|\mathbb{W}_0| = b$, then $\mathbb{W}_0 = \{1, 2, \ldots, b\}$ and $\mathbb{W}_e = \{b+1, b+2, \ldots, n\}$. We denote $\boldsymbol{s}_{\mathbb{W}_0} = (s_1, \ldots, s_b, 1, \ldots, 1)^\top$ and $\boldsymbol{q}_{\mathbb{W}_0} = (q_1, \ldots, q_b, 1, \ldots, 1)^\top$. In our considerations below, we assume that the i-th coordinate of $\boldsymbol{s}$ is always equal to the i-th coordinate of $\boldsymbol{s}_{\mathbb{W}_0}$, $i \in \mathbb{W}_0$. Analogously, the i-th coordinate of $\boldsymbol{q}$ is always equal to the i-th coordinate of $\boldsymbol{q}_{\mathbb{W}_0}$, $i \in \mathbb{W}_0$.

In the context of the DMSBPM, for clarity and convenience, we give the following definition.

Definition 2.12. *Given Definition 2.2, denote the p.g.f. of a DMSBPM starting with one particle of type $i \in \mathbb{W}$ as*

1. *For $i \in \mathbb{W}_e$*

$$F_i(t; \boldsymbol{s}) = \mathbb{E}\Big(\prod_{j \in \mathbb{W}} s_j^{Z_j(t)} |\ \boldsymbol{Z}(0) = \boldsymbol{\delta}^i \Big),$$

$$F_{i,a}(t; \boldsymbol{s}) = \mathbb{E}\Big(\prod_{j \in \mathbb{W}} s_j^{Z_j(t)} |\ \boldsymbol{Z}(0) = \boldsymbol{\delta}^i_a \Big),$$

where $|\boldsymbol{s}| \leq 1$.

2. *For $i \in \mathbb{W}_0$, due to the fact that there can be no mutations from $\mathbb{W}_0$ towards $\mathbb{W}_e$*

$$F_i(t; \boldsymbol{s}_{\mathbb{W}_0}) = \mathbb{E}\Big(\prod_{j \in \mathbb{W}_0} s_j^{Z_j(t)} |\ \boldsymbol{Z}(0) = \boldsymbol{\delta}^i \Big),$$

$$F_{i,a}(t; \boldsymbol{s}_{\mathbb{W}_0}) = \mathbb{E}\Big(\prod_{j \in \mathbb{W}_0} s_j^{Z_j(t)} |\ \boldsymbol{Z}(0) = \boldsymbol{\delta}^i_a \Big),$$

where $|\boldsymbol{s}_{\mathbb{W}_0}| \leq 1$.

Corollary 2.12. *For the DMSBPM, the following system of integral equations holds:*

1. *For $i \in \mathbb{W}_e$*

$$F_i(t;\, \boldsymbol{s}) = s_i\big(1 - G_i(t)\big) + \int_0^t f_i\Bigg(y;\, \Big[\sum_{m \in \mathbb{W}_e} u_{im} F_m(t-y;\, \boldsymbol{s})\Big] + \Big[\sum_{r \in \mathbb{W}_0} u_{ir} F_r(t-y;\, \boldsymbol{s}_{\mathbb{W}_0})\Big]\Bigg) dG_i(y). \tag{2.30}$$

2. *For $i \in \mathbb{W}_0$*

$$F_i(t;\, \boldsymbol{s}_{\mathbb{W}_0}) = s_i\big(1 - G_i(t)\big) + \int_0^t f_i\Bigg(y;\, \sum_{r \in \mathbb{W}_0} u_{ir} F_r(t-y;\, \boldsymbol{s}_{\mathbb{W}_0})\Bigg) dG_i(y). \tag{2.31}$$

Definition 2.13. *Given Definition 2.6, for a DMSBPM starting with one particle of type i, $i \in \mathbb{W}$, denote the p.g.f.s for the numbers of mutants produced from $\mathbb{W}_e$ towards $\mathbb{W}_0$, until t, as*

1. *For $i \in \mathbb{W}_e$*

$$h_i^{\mathbb{W}_e}(t;\boldsymbol{s}_{\mathbb{W}_0}) = \mathbb{E}\Big(\prod_{j\in\mathbb{W}_0} s_j^{I_j^{\mathbb{W}_e}(t)} \mid \boldsymbol{Z}(0)=\boldsymbol{\delta}^i\Big),$$

$$h_{i,a}^{\mathbb{W}_e}(t;\boldsymbol{s}_{\mathbb{W}_0}) = \mathbb{E}\Big(\prod_{j\in\mathbb{W}_0} s_j^{I_j^{\mathbb{W}_e}(t)} \mid \boldsymbol{Z}(0)=\boldsymbol{\delta}_a^i\Big),$$

where $|\boldsymbol{s}| \le 1$.

2. *For $i \in \mathbb{W}_0$, due to the fact that there can be no mutations from $\mathbb{W}_0$ towards $\mathbb{W}_e$*

$$h_i^{\mathbb{W}_e}(t;\boldsymbol{s}_{\mathbb{W}_0}) = h_{i,a}^{\mathbb{W}_e}(t;\boldsymbol{s}_{\mathbb{W}_0}) = 1.$$

Corollary 2.18. *For the DMSBPM the following system of integral equations holds:*

1. *For $i \in \mathbb{W}_e$*

(2.40)
$$h_i^{\mathbb{W}_e}(t;\boldsymbol{s}_{\mathbb{W}_0}) = \big(1-G_i(t)\big) + \int_0^t f_i\Bigg(y;\Big[\sum_{m\in\mathbb{W}_e} u_{im}h_m^{\mathbb{W}_e}(t-y;\boldsymbol{s}_{\mathbb{W}_0})\Big] + \Big[\sum_{r\in\mathbb{W}_0} u_{ir}s_r\Big]\Bigg)dG_i(y).$$

2. *For $i \in \mathbb{W}_0$*

$$h_i^{\mathbb{W}_e}(t;\boldsymbol{s}_{\mathbb{W}_0}) = 1.$$

Definition 2.14. *Given Definition 2.7, for a DMSBPM starting with one particle of type i, $i \in \mathbb{W}$, denote the p.g.f.s for the numbers of mutants produced from $\mathbb{W}_e$ towards $\mathbb{W}_0$, during the entire process, as*

1. *For $i \in \mathbb{W}_e$*

$$h_i^{\mathbb{W}_e}(\boldsymbol{s}_{\mathbb{W}_0}) = \mathbb{E}\Big(\prod_{j\in\mathbb{W}_0} s_j^{I_j^{\mathbb{W}_e}} \mid \boldsymbol{Z}(0)=\boldsymbol{\delta}^i\Big),$$

$$h_{i,a}^{\mathbb{W}_e}(\boldsymbol{s}_{\mathbb{W}_0}) = \mathbb{E}\Big(\prod_{j\in\mathbb{W}_0} s_j^{I_j^{\mathbb{W}_e}} \mid \boldsymbol{Z}(0)=\boldsymbol{\delta}_a^i\Big),$$

where $|\boldsymbol{s}| \le 1$.

2. *For $i \in \mathbb{W}_0$, due to the fact that there can be no mutations from $\mathbb{W}_0$ towards $\mathbb{W}_e$*

$$h_i^{\mathbb{W}_e}(\boldsymbol{s}_{\mathbb{W}_0}) = h_{i,a}^{\mathbb{W}_e}(\boldsymbol{s}_{\mathbb{W}_0}) = 1.$$

Corollary 2.20. *The following system of equations holds within the DMSBPM, $i \in \mathbb{W}$:*

1. *For $i \in \mathbb{W}_e$*

$$h_i^{\mathbb{W}_e}(\boldsymbol{s}_{\mathbb{W}_0}) = \int_0^\infty f_i\Bigg(y;\Big[\sum_{m\in\mathbb{W}_e} u_{im}h_m^{\mathbb{W}_e}(\boldsymbol{s}_{\mathbb{W}_0})\Big] + \Big[\sum_{r\in\mathbb{W}_0} u_{ir}s_r\Big]\Bigg)dG_i(y). \qquad (2.42)$$

2. For $i \in \mathbb{W}_0$

$$h_i^{\mathbb{W}_e}(\boldsymbol{s}_{\mathbb{W}_0}) = 1. \tag{2.43}$$

Proposition 2.1. *Let each particle type from* $\mathbb{W}_e$*, within the DMSBPM, be either subcritical or critical. Then for* $i \in \mathbb{W}_e$

$$q_i = h_i^{\mathbb{W}_e}(\boldsymbol{q}_{\mathbb{W}_0}) = \int_0^\infty f_i\bigg(y; \Big[\sum_{m\in\mathbb{W}_e} u_{im} h_m^{\mathbb{W}_e}(\boldsymbol{q}_{\mathbb{W}_0})\Big] + \Big[\sum_{r\in\mathbb{W}_0} u_{ir} q_r\Big]\bigg) dG_i(y) \tag{2.45}$$

and

$$q_{i,a} = h_{i,a}^{\mathbb{W}_e}(\boldsymbol{q}_{\mathbb{W}_0}) = \int_0^\infty f_i\bigg(a + y; \Big[\sum_{m\in\mathbb{W}_e} u_{im} h_m^{\mathbb{W}_e}(\boldsymbol{q}_{\mathbb{W}_0})\Big] + \Big[\sum_{r\in\mathbb{W}_0} u_{ir} q_r\Big]\bigg) dG_{i,a}(y). \tag{2.46}$$

Definition 2.15. *Denote with* $T_{\mathbb{W}_0}^{\mathbb{W}_e}$ *the r.v. that is the time until occurrence of the first "successful" mutant produced from a type within* $\mathbb{W}_e$ *towards a type within* $\mathbb{W}_0$ *in a DMSBPM starting with some combination of particles with types within* $\mathbb{W}_e$*. Without loss of generality, we set the starting number of particles per type* $r \in \mathbb{W}_e$ *to be* k_r *and denote the initial state of the process as* $\boldsymbol{Z}(0) = \boldsymbol{\alpha}^* = \big(0, \ldots, 0, Z_{b+1}(0) = k_{b+1}, \ldots, Z_n(0) = k_n\big)^\top$*. In* $\boldsymbol{\alpha}^*$*, without loss of generality, we have set* $|\mathbb{W}_0| = b$ *and we have arranged for the types from* $\mathbb{W}_0$ *to correspond to the first* b *coordinates. We define* $T_{\mathbb{W}_0}^{\mathbb{W}_e} = \infty$ *as the event that no "successful" mutants occur during a DMSBPM beginning with an initial state* $\boldsymbol{\alpha}^*$*. Thus, we may write* $T_{\mathbb{W}_0}^{\mathbb{W}_e} \in (0, \infty]$*. If the DMSBPM starts with a single particle of type* i*,* $i \in \mathbb{W}_e$*, of age* 0*, we use* $T_{\mathbb{W}_0,i}^{\mathbb{W}_e}$ *as a shortcut notation. If the initial particle is of age* a*,* $a \neq 0$*, we use* $T_{\mathbb{W}_0,i,a}^{\mathbb{W}_e}$*.*

Definition 2.17. *Define for an initial particle of type* i*,* $i \in \mathbb{W}_e$*, the following modified hazard function:*

1. If the initial particle is of age 0

$$\hat{g}_{\mathbb{W}_0,i}^{\mathbb{W}_e}(t)dt = \mathbb{P}\Big(T_{\mathbb{W}_0,i}^{\mathbb{W}_e} \in (t, t+dt] \mid T_{\mathbb{W}_0,i}^{\mathbb{W}_e} > t,\ \sum_{c\in\mathbb{W}_e} Z_c(t) > 0\Big).$$

2. If the initial particle if of age a*,* $a \neq 0$

$$\hat{g}_{\mathbb{W}_0,i,a}^{\mathbb{W}_e}(t)dt = \mathbb{P}\Big(T_{\mathbb{W}_0,i,a}^{\mathbb{W}_e} \in (t, t+dt] \mid T_{\mathbb{W}_0,i,a}^{\mathbb{W}_e} > t,\ \sum_{c\in\mathbb{W}_e} Z_c(t) > 0\Big).$$

We obtain

$$\hat{g}_{\mathbb{W}_0,i}^{\mathbb{W}_e}(t) = \frac{F^{(1)}_{T_{\mathbb{W}_0,i}^{\mathbb{W}_e}}(t)}{\mathbb{P}\Big(T_{\mathbb{W}_0,i}^{\mathbb{W}_e} > t\Big) - \mathbb{P}\Big(T_{\mathbb{W}_0,i}^{\mathbb{W}_e} > t,\ \sum_{c\in\mathbb{W}_e} Z_c(t) = 0\Big)}.$$

Definition 2.18. *For $i \in \mathbb{W}_e$, denote*

$$V_i(t) = \mathbb{P}\Big(T^{\mathbb{W}_e}_{\mathbb{W}_0,i} > t,\ \sum_{c\in\mathbb{W}_e} Z_c(t) = 0\Big),$$

$$V_{i,a}(t) = \mathbb{P}\Big(T^{\mathbb{W}_e}_{\mathbb{W}_0,i,a} > t,\ \sum_{c\in\mathbb{W}_e} Z_c(t) = 0\Big).$$

Theorem 2.8. *The probability $V_i(t)$ of the event that jointly the first "successful" mutant does not occur before or at t and there are no particles from $\mathbb{W}_e$ left at t, for a DMSBPM starting with a particle of age 0, satisfies the following system of integral equations:*

$$V_i(t) = \int_0^t f_i\Bigg(y; \Big[\sum_{m\in\mathbb{W}_e} u_{im}V_m(t-y)\Big] + \Big[\sum_{r\in\mathbb{W}_0} u_{ir}q_r\Big]\Bigg)dG_i(y), \quad i \in \mathbb{W}_e. \tag{2.53}$$

Corollary 2.22. *The probability $V_{i,a}(t)$ of the event that jointly the first "successful" mutant does not occur before or at t and there are no particles from $\mathbb{W}_e$ left at t, for a DMSBPM starting with a particle of age a, $a \neq 0$, satisfies the following system of integral equations:*

$$V_{i,a}(t) = \int_0^t f_i\Bigg(a + y; \Big[\sum_{m\in\mathbb{W}_e} u_{im}V_m(t-y)\Big] + \Big[\sum_{r\in\mathbb{W}_0} u_{ir}q_r\Big]\Bigg)dG_{i,a}(y), \quad i \in \mathbb{W}_e. \tag{2.54}$$

2.3.2 Decomposable Multi-type Bellman-Harris Branching Process through probabilities of Mutation between types (DMBHBPM)

The DMBHBPM is a particular case of the DMSBPM where there is no dependence of particle reproduction capabilities from particle age. The full formal definition of the DMBHBPM can be seen in Definition 2.19 in Subsection 2.3.2.1 of the dissertation. The results for the DMBHBPM are analogous to the results for the DMSBPM, however the dependence from particle age is dropped within f_i.

CHAPTER 3

Sequential decision problems with branching process based dynamics

3.1 Chapter overview and organization

In this Chapter, we incorporate branching processes into optimization problems known as Sequential Decision Problems (SDPs). For our modeling of SDPs, we use Warren B. Powell's "Universal Modeling Framework" developed in [82] (2022). Our motivation for this choice of framework can be summarized as follows: 1) The "Universal Modeling Framework" is an attempt to unify the 15 communities discussed in the Introduction of the dissertation (Section 1.5, pages 23 - 24). This may prove beneficial in future research, where we may consider complicating the branching process based SDPs that we investigate; 2) The "Universal Modeling Framework" is straightforwardly connected to Approximate Dynamic Programming (ADP; see [74], [76], [78]) and Reinforcement Learning (RL; see [203], [82]). ADP and RL rely on simulations in order to produce solutions (of varying quality) for complex SDPs. We envision future research stemming from this dissertation, regarding SDPs with branching process based dynamics, as simulations based; 3) For our purposes, the "Universal Modeling Framework" is conceptually and notationally close to the discussions within the Markov decision processes community (see Puterman [70] (2005)). This is a good starting point for considering SDPs with Bienaymé-Galton-Watson (BGW) branching process dynamics as the BGW process is Markovian within standard definitions.

Our modeling of SDPs within this dissertation is heavily based on the ideas developed in [82] and [78] and as such shares the strengths and weaknesses of the "Universal Modeling Framework". We note that, with respect to our purposes, we have made some minor contributions to the presentation in [82], those contributions being Proposition 3.1 and Proposition 3.2 from Subsection 3.2.7, as well as the inclusion of the discount factor γ in some equations and statements.

In Section 3.2 and Section 3.3, we present the "Universal Modeling Framework" by adapting parts of the presentation in [82]. In Section 3.4, Section 3.5, and Section 3.6, we obtain our novel results that incorporate certain branching processes into SDPs within the "Universal Modeling Framework". The results from Section 3.4, Section 3.5, and

Section 3.6, have not been published yet. In Section 3.7, we outline, but do not apply or investigate the properties of, a general ADP algorithm that can be used as a starting point for developing a specialized ADP algorithm for finding the solution of the SDP discussed within Section 3.6. We stress that in the dissertation, we do not consider stochastic differential equations within the optimization problems investigated.

The standard notion of “state” within the branching processes community postulates that the state of a branching process at t is to be understood as the number of particles, per type, that exist at t. With respect to a so defined “state” the BGW branching process as well as the Bellman-Harris branching process with exponential lifespans are Markovian. The scarce literature dedicated to combining specifically branching processes and SDPs, see [77], [199], [200], [201], [202], concentrates its efforts on branching processes that are Markovian under the standard notion of state. The papers listed, effectively, discuss multi-type Bienaymé-Galton-Watson branching processes. Our novel idea within Section 3.6 is to consider a novel definition of the state of the MSBPM (the MSBPM is generally non-Markovian under the standard notion of “state” since particle reproduction can depend on particle age). Under the newly defined “state”, we prove that the MSBPM is Markovian. This and the following considerations within Section 3.6 formally allow us to apply ADP and RL for the purpose of finding solutions of the corresponding SDP.

We note that [202] develops a model-free RL algorithm for a SDP with BGW branching process based dynamics. Contrary to [202], our agenda is to exploit a specified branching process model (such as the MSBPM) as much as possible. We note that while RL algorithms are usually model-free, ADP algorithms are usually model-based. This is why we outline a general ADP algorithm in Section 3.7 of the dissertation as an illustration of possible future research - although we successfully incorporate the (generally non-Markovian) MSBPM into SDPs within the “Universal Modeling Framework”, devising practical computational algorithms requires substantial further research. We hope that the computational tractability of the MSBPM, via Numerical Scheme 1 and Numerical Scheme 2, together with the theoretical foundation laid within the dissertation, will facilitate future success.

Within Section 3.4, we consider the paper of S. R. Pliska from 1976, [77]. [77] discusses a SDP with BGW branching process based dynamics and provides a theorem that allows us to efficiently obtain the solution of the (finite-horizon) SDP, described within Section 2 and Section 3 of the paper, via a Dynamic Programming algorithm (see, [67], [68], [69], [70]). Although [77] acknowledges that the algorithm obtained is a Dynamic Programming algorithm, the proof of Theorem 3.1. from [77] uses conditional expectations and does not use Bellman's optimality equation (see Section 3.3). Within Section 3.4, we recast the discussion in [77] into the more contemporary “Universal Modeling Framework” and provide a novel proof of Theorem 3.1 from [77] that is based on Bellman's optimality equation.

In Section 3.5, we consider the Multi-type Bellman-Harris Branching Process through probabilities of Mutation between types (MBHBPM; a special case of the MSBPM) with exponential lifespan distributions. The MBHBPM with exponential lifespan distributions is Markovian under the classical definition of state. Our novel contribution for this case is that we formally incorporate the process in SDPs within the “Universal Modeling Framework” and show that a result analogous to the Theorem 3.1. from [77] holds.

We stress that the algorithms obtained in Section 3.4 and Section 3.5, that allow for efficiently finding the solutions of the corresponding SDPs, discussed within these sections, have a limited scope of application. More precisely, these algorithms easily become non-applicable upon introducing further (appropriately modeled with respect to the "Universal Modeling Framework") dependencies within the discussed SDPs. However, for such complex SDPs, we can still consider the ADP and/or RL approach.

We note that the field of Controlled Branching Processes (CBPs; see [115], [124], [181], [182], [120], [121]) contains ideas that are close to the ideas found within the discussion of sequential decision problems. The relationship between CBPs and SDPs is that a CBP is a branching process and as such can be used as a model of uncertainty within a SDP.

This Chapter can also be viewed as a continuation of the efforts within [1]-[7], as well as Chapter 2, to model cancer evolution and populations escaping extinction. Indeed, this context can benefit much from an introduction of SDPs as SDPs provide a way for planning appropriate actions in advance. This can be very beneficial, for example, within the case of administering cancer therapies as the different costs and expected results associated with different therapies have to be considered by the recipient and medical personnel. SDPs model the outcomes of the choices available to us, thus, depending on our objectives, they have the potential to become a useful tool for finding the best way available for forcing a population into extinction or for maximizing its chance of survival.

The Chapter is organized as follows. In Section 3.2, we introduce relevant concepts from the "Universal Modeling Framework" proposed by Warren B. Powell in [82]. In Section 3.3, we discuss Bellman's optimality equation. In Section 3.4, we recast the model from [77] into the "Universal Modeling Framework" and utilize Bellman's optimality equation to provide a novel proof for Theorem 3.1 from [77]. In Section 3.5, we consider the MBHBPM with exponential lifespan distributions, construct a corresponding SDP and prove a result similar to Theorem 3.1 from [77] for this case. In Section 3.6, we consider the MSBPM and construct a SDP with MSBPM based dynamics. In Section 3.7 within the dissertation, we outline an ADP approach for solving the constructed SDP with MSBPM based dynamics. We finish with Section 3.8, where we give illustrative examples of SDPs with branching process based dynamics.

3.2 Modeling of Sequential Decision Problems (SDPs)

Within this Section, we introduce the "Universal Modeling Framework" developed by Warren B. Powell in [82] (2022) and [80]. We follow primarily Chapter 9 (page 467) from [82], however, we have streamlined the presentation in accordance with our purposes.

Recall the 15 mathematical communities, discussed in Section 1.5 within the dissertation, that consider Sequential Decision Problems (SDPs). Each of these communities has its distinctive toolset and perspective when working with deterministic and/or stochastic optimization. The choice of community/communities within which to develop and associate our work is thus of paramount importance. The "Universal Modeling Framework" aims at covering the particularities encountered within all of the 15 communities that work with SDPs in a unified way. This may facilitate a more easy incorporation of ideas from the aforementioned communities into subsequent future research that continues the work within this dissertation. The framework (as well as the presentation in [82] and [78]) is also oriented towards the use of simulations and computer resources - we believe this approach towards solving complicated SDPs to hold great promise, considering that closed-form solutions are rare for SDPs.

Having in mind our discussion in Section 1.2 of the Abstract (that corresponds to Description 1.1 in Section 1.5 from the dissertation), in what follows, we make the following arrangement:

Notational Choice 1. We index the decision epochs with t, $t = 0, 1, 2, \ldots$. Any variable indexed with t is understood as a variable corresponding to decision epoch with index t. When we talk about intervals, e.g. interval (t, t+1), we understand the interval between epochs with index t and index $t+1$. We assume that the distance between two decision epochs can vary between any two neighboring epochs but cannot be 0 or ∞. If there is a final epoch, its index is T.

Following [82] (page 470), there are 5 components when modeling any sequential decision problem: 1) Exogenous information variables; 2) Decision/action/control variables; 3) Transition function; 4) State variables; 5) Objective function. For the purposes of the Abstract, we give here a very quick overview of these concepts. For mode details and relevant discussions, see Section 3.2 within the dissertation or Chapter 9 from [82].

1. The *Exogenous information variables* capture all of the stochasticity of the system we are trying to model, with the possible exception of stochasticity related to the initial state. The exogenous information is generically denoted with W_t.
2. *Decision/action/control* variables model our possible interactions with the system. Decisions made at t are denoted as x_t and the space of all possible decisions is denoted with $\mathcal{X}$ (or $\mathcal{X}_t$). Decisions can affect the future evolution of the system. A policy π, $\pi \in \Pi$, where Π is the set of all policies, defines a *decision function* $X^\pi(\cdot)$ (or $X^\pi_t(\cdot)$) that outputs decision x_t for a given state of the system S_t.
3. The *Transition function* describes how the systems evolves between neighboring decision epochs.
4. The *State variables* contain all necessary information so that the evolution of the system is Markovian. We denote the state at t as S_t and the space of all possible states as $\mathcal{S}$ (or $\mathcal{S}_t$).
5. The *Objective function* specifies a relevant performance metric. We write the objective function via the *Contribution function* $C(S_t, x_t)$ (or $C_t(S_t, x_t)$) which outputs the result, with respect to our metric, of applying decision x_t onto state S_t. Let γ be a discount factor ($\gamma \leq 1$). If the initial state S_0 is probabilistic, the objective function can be written as

$$F^\pi(S_0) = \mathbb{E}_{S_0}\mathbb{E}_{W_t,\ldots,W_T|S_0}\left\{\sum_{t=0}^{T}\gamma^t C_t\big(S_t, X^\pi_t(S_t)\big)|S_0\right\},$$

if not then we may write

$$F^\pi(S_0) = \mathbb{E}_{W_t,\ldots,W_T|S_0}\left\{\sum_{t=0}^{T}\gamma^t C_t\big(S_t, X^\pi_t(S_t)\big)|S_0\right\},$$

or more compactly

$$F^{\pi}(S_0) = \mathbb{E}\left\{ \sum_{t=0}^{T} \gamma^t C_t\big(S_t, X_t^{\pi}(S_t)\big) | S_0 \right\}. \tag{3.1}$$

We are now ready to formalize our definition of a SDP. Definition 3.2, that we give below, constitutes an aggregation of the discussion within Chapter 9 from [82]. When writing Definition 3.2, we keep in mind Notational Choice 1.

Definition 3.2. *A finite-horizon Sequential Decision Problem (finite-horizon SDP) within the "Universal Modeling Framework", with final decision epoch T, is characterized by the sequence*

$$(S_0, x_0, W_1, S_1, \dots, S_t, x_t, W_{t+1}, S_{t+1}, \dots, S_T).$$

The objective of a finite-horizon SDP is to find a policy that satisfies

$$\max_{\pi \in \Pi} \mathbb{E}_{S_0} \mathbb{E}_{W_1,\dots,W_T|S_0} \left\{ \sum_{t=0}^{T} \gamma^t C_t\big(S_t, X_t^{\pi}(S_t)\big) | S_0 \right\}.$$

An infinite-horizon Sequential Decision Problem (infinite-horizon SDP) within the "Universal Modeling Framework" is characterized by the sequence

$$(S_0, x_0, W_1, S_1, \dots, S_t, x_t, W_{t+1}, S_{t+1}, \dots).$$

The objective of an infinite-horizon SDP is to find a policy that satisfies

$$\max_{\pi \in \Pi} \mathbb{E}_{S_0} \mathbb{E}_{W_1,W_2,\dots|S_0} \left\{ \sum_{t=0}^{\infty} \gamma^t C_t\big(S_t, X_t^{\pi}(S_t)\big) | S_0 \right\}.$$

If the set of all policies, Π, is infinite, we write "sup" instead of "max".

Within the dissertation, we will be interested only in finite-horizon SDPs. We leave the topic of incorporating branching processes into infinite-horizon SDPs for future considerations.

We stress that, as stated in [82], page 482, (vi), Definition 3.1 for state variables within the dissertation (i.e., the policy-dependent version of Definition 9.4.1 from [82]) implies that all properly modeled, with respect to the "Universal Modeling Framework", dynamic systems are Markovian by construction.

Next, we provide Proposition 3.1 and Proposition 3.2 independently of the discussion within [82]. We actively use these propositions, as well as the considerations following them, in subsequent Sections within the dissertation.

Proposition 3.1. *For any fixed policy π, a SDP within the "Universal Modeling Framework" constitutes a discrete-time (possibly non-stationary) Markov chain with respect to $t = 0, 1, \dots, T$.*

The following is also true

Proposition 3.2. *A discrete-time, with respect to* $t = 0, 1, \ldots, T$*, possibly non-stationary, Markov chain can be viewed as a SDP within the "Universal Modeling Framework".*

Perhaps the most difficult moment when modeling a dynamic system as a SDP within the "Universal Modeling Framework" is verifying that the defined state variables, decision variables, exogenous information variables, contribution and objective functions, and transition function, all satisfy the assumptions of a SDP within the framework. More specifically, proving the conditional independence of the transition function between t and $t+1$ from states and decisions prior to t can be especially problematic in the general case. Proposition 3.1 and Proposition 3.2 provide us with a way of checking that our model is indeed a SDP within the "Universal Modeling Framework" - what we need to do is verify that for every fixed policy π the resulting process is a discrete time Markov chain with respect to $t = 0, 1, \ldots, T$.

3.3 Bellman's optimality equation for SDPs within "Universal Modeling Framework"

Within this Section, we adapt the discussions in Chapter 14 from [82] and Chapter 3 from [78]. Similarly to the presentation for Markov Decision Processes (MDP) within Section 4.3 in [70], we will first define Bellman's optimality equation within the "Universal Modeling Framework" and then, we will show its relevance with respect to solving SPDs. In the current Section, we focus only on aspects of the discussion that we will need when incorporating branching processes into SDPs.

Definition 3.3. *In the context of a SDP within the "Universal Modeling Framework", as given by Definition 3.2, define (the expectation form of) Bellman's optimality equation at* t *as*

$$V_t(S_t) = \max_{x_t \in \mathcal{X}_t} \Big(C_t(S_t, x_t) + \gamma \mathbb{E}\big\{ V_{t+1}(S_{t+1}) | S_t, x_t \big\} \Big). \tag{3.2}$$

$V_t(S_t)$ is also known as the *value function* as it gives us the value of being in state S_t at t. If we write equation (3.2) for consecutive t it is evident that the Markov property is implied, which is to be expected knowing that all dynamic systems considered within the "Universal Modeling Framework" are Markovian. When using (3.2), we keep in mind that the exogenous information, if any, may or may not depend on S_t and x_t (and does not depend on states and decisions prior to t). If there is exogenous information affecting our system, its influence is captured by the expectation in (3.2).

Denote

$$F_t^\pi(S_t) = \mathbb{E}\left\{ \sum_{t'=t}^{T-1} \gamma^{t'-t} C_{t'}\big(S_{t'}, X_{t'}^\pi(S_{t'})\big) + \gamma^{T-t} C_T(S_T) | S_t \right\}$$

and

$$F_t^*(S_t) = \max_{\pi \in \Pi} F_t^\pi(S_t).$$

The following theorem, without γ, can be found in Subsection 14.12.1 of [82]. Within the dissertation, we incorporate γ into the proof given within [82].

Theorem 3.1. *Let $V_t(S_t)$ be a solution to Bellman's optimality equation* (3.2)

$$V_t(S_t) = \max_{x_t \in \mathcal{X}_t} \Big(C_t(S_t, x_t) + \gamma \mathbb{E}\big\{ V_{t+1}(S_{t+1}) | S_t, x_t \big\} \Big).$$

Then, for a finite-horizon SDP

$$F_t^*(S_t) = V_t(S_t).$$

Algorithms based on Bellman's optimality equation often (but not always) suffer from the so-called curses of dimensionality. More specifically, depending on the setting at hand, non-specialized algorithms may require iterating over all possible states and/or decisions. This quickly becomes impractical for discrete multi-dimensional state and decision spaces and is not an option when considering continuous state and decision spaces. Incorporation of branching process into SDPs, thus, needs to be done with care as the usual state space associated with a branching process, i.e., the number of particles, per type, that exist at t, is countably infinite and possibly multi-dimensional.

3.4 SDPs with underlying BGW branching process dynamics

To the best of our knowledge, stochastic sequential decision problems, where the dynamics is generated specifically by a branching process (the case of the Bienaymé-Galton-Watson branching process is investigated), are considered for the first time in [77] (1976). Our contributions within this Section are: 1) We recast the Markov decision process from [77] into SDP Model 1 within "Universal Modeling Framework"; 2) We provide a novel proof for Theorem 3.1 from [77] that is based on Bellman's optimality equation. These contributions have not been published yet.

Definition of SDP Model 1. Define SDP Model 1 as the finite-horizon Sequential Decision Problem that satisfies:

1. We observe a BGW branching process at successive times indexed with $t = 0, 1, 2, \ldots, T$.
2. Let k be the number of particles types within the BGW branching process. The state space $\mathcal{S}_t$ consists of all k-dimensional vectors whose coordinates are non-negative integers. The t index in $\mathcal{S}_t$ indicates that there are probability distributions and p.g.f.s associated with each type that may change with t (after a decision has been made). The state of the process at t is also called the "generation" or the "population" at t. The state of the process at t is given by $\boldsymbol{S}_t = (S_{1,t}, S_{2,t}, \ldots, S_{k,t})^\top$, where all $S_{i,t}$ are with values in $\mathbb{N}_0$ and $S_{i,t}$ is the (non-negative) number of particles of type i that exist at t. The initial state $\boldsymbol{S}_0$ is deterministic.
3. Each particle type i has a specific finite set of possible decisions (actions) $\widetilde{\mathcal{X}}_i$ associated with it. Hence, the decision space is given by $\mathcal{X} = \widetilde{\mathcal{X}}_1 \times \widetilde{\mathcal{X}}_2 \times \cdots \times \widetilde{\mathcal{X}}_k$. We denote the decisions made at t with $\boldsymbol{x}_t = (x_{1,t}, x_{2,t}, \ldots, x_{k,t})^\top$.

4. Let $c_i(x_{i,t})$ be the individual contribution (reward) received for a type i particle after making decision $x_{i,t}$ for all particles of type i at t. We assume that $-\infty < c_i(x_{i,t}) < \infty$ for all i and that $c_i(\cdot)$ do not depend on t. If we let $\boldsymbol{c}(\boldsymbol{x}_t) = \big(c_1(x_{1,t}), \ldots, c_k(x_{k,t})\big)^\top$, then the generation contribution at t is $C_t\big(\boldsymbol{S}_t, \boldsymbol{x}_t\big) = \sum_{i=1}^{k} S_{i,t} \cdot c_i(x_{i,t}) = \boldsymbol{S}_t^\top \boldsymbol{c}(\boldsymbol{x}_t)$. At $t = T$ no decisions are made, instead a terminal $\boldsymbol{c}_T = \big(c_1, \ldots, c_k\big)^\top$ is collected, hence the generation contribution at $t = T$ is $C_T(\boldsymbol{S}_T) = \boldsymbol{S}_T^\top \boldsymbol{c}_T$.
5. The decision selected for a particle affects the number of offspring, per type, that the particle has in the next generation.
   (a) For each k-dimensional vector $\boldsymbol{q} = (q_1, \ldots, q_k)^\top$ of non-negative integers, let $p_i(\boldsymbol{q}, x_{i,t})$ be the probability that a type i particle, whose corresponding decision is $x_{i,t}$, will produce exactly q_1 type 1 offspring, $\ldots$, q_k type k offspring, $\sum_{\boldsymbol{q}} p_t(\boldsymbol{q}, x_{i,t}) = 1$.
   (b) Corresponding to each $p_i(\cdot, x_{i,t})$ is the row vector $\boldsymbol{m}_i(x_{i,t}) = \big(m_{i1}(x_{i,t}), \ldots, m_{ik}(x_{i,t})\big)$, where $m_{ij}(x_{i,t})$ equals the expected number of type j offspring produced by a single particle of type i under decision $x_{i,t}$. We assume that $m_{ij}(x_{i,t}) < \infty$ for all $x_{i,t} \in \widetilde{\mathcal{X}}_i$ and $i, j = 1, \ldots, k$. Given $\boldsymbol{x}_t \in \mathcal{X}$, we organize the expectations into matrix $M(\boldsymbol{x}_t) = (\boldsymbol{m}_1(x_{1,t}), \ldots, \boldsymbol{m}_k(x_{k,t}))^\top$.
6. Denote a policy by π. The set of possible policies, with respect to $\mathcal{X}$, is Π. Denote the decision function at t, corresponding to policy π, with $X_t^\pi(\cdot)$.
7. A discount factor γ is given.
8. We want to obtain the maximum expected T-period discounted reward given by

$$\max_{\pi \in \Pi} \mathbb{E}\left\{ \sum_{t=0}^{T-1} \gamma^t C_t(\boldsymbol{S}_t, X_t^\pi(\boldsymbol{S}_t)) + \gamma^T C_T(\boldsymbol{S}_T) | \boldsymbol{S}_0 \right\}.$$

Definition 3.4. *Let $\boldsymbol{X}$ be a k-dimensional vector. The maximum return operator $\mathcal{R}$ is given by*

$$\mathcal{R}\boldsymbol{X} = \max_{\boldsymbol{x} \in \mathcal{X}} \big\{ \boldsymbol{c}(\boldsymbol{x}) + \gamma M(\boldsymbol{x}) \boldsymbol{X} \big\}.$$

We denote the n-fold composition of $\mathcal{R}$ as $\mathcal{R}^n$ and $\mathcal{R}^0$ is understood as the identity operator.

Within Theorem 3.2 below, we provide a novel proof of Theorem 3.1 from [77] using Bellman's optimality equation (3.2). We note that the original proof of Theorem 3.1 from [77] is based on conditional expectations.

Theorem 3.2. *For SDP Model 1, the value function $V_t(\boldsymbol{S}_t)$ satisfies*

$$V_t(\boldsymbol{S}_t) = \boldsymbol{S}_t^\top \mathcal{R}^{T-t} \boldsymbol{c}_T, \qquad t = 0, 1, \ldots, T-1. \tag{3.9}$$

Policy π, with corresponding $X_t^\pi(\cdot)$, that computes decisions $\boldsymbol{x}_t$ satisfying

$$\boldsymbol{c}(\boldsymbol{x}_t) + \gamma M(\boldsymbol{x}_t) \mathcal{R}^{T-t-1} \boldsymbol{c}_T = \mathcal{R}^{T-t} \boldsymbol{c}_T, \qquad t = 0, 1, \ldots, T-1, \tag{3.10}$$

is optimal.

3.5 SDPs with underlying exponential lifespan MB-HBPM dynamics

Our contributions within this Section are: 1) We provide a proof that the multi-type Bellman-Harris branching process with exponential lifespans, as well as the MBHBPM with exponential lifespans, are discrete-time Markov chains with respect to $t = 0, 1, \ldots, T$; 2) For these processes, we construct SDPs within the "Universal Modeling Framework"; 3) We show that a theorem similar to Theorem 3.2 holds for the newly constructed SDPs. These contributions have not been published yet.

Proposition 3.4. *A multi-type Bellman-Harris branching process with exponential lifespan distributions for all particle types, with states defined as the number of particles, per type, that exist at moment t, is a discrete-time Markov chain with respect to the moments in time indexed by $t = 0, 1, \ldots, T$.*

Definition of SDP Model 2. Define SDP Model 2 as the finite-horizon SDP that corresponds to the definition of SDP Model 1 upon which the following modifications are applied:

1. We observe a MBHBPM at epochs indexed with t, $t = 0, 1, 2, \ldots, T$. Regardless of t, lifespan distributions for particles of each type must be exponential.
2. Let k be the number of types of particles within the MBHBPM. The state space $\mathcal{S}_t$ consists of all k-dimensional vectors whose coordinates are non-negative integers. The t index in $\mathcal{S}_t$ indicates that there are probability distributions and p.g.f.s associated with each type that may change with t (after a decision has been made). However, although lifespan distribution may change their parameters they must continue to be exponential. The state of the process at t is also called the "generation" or the "population" at t. The state of the process at t is given by $\boldsymbol{S}_t = (S_{1,t}, S_{2,t}, \ldots, S_{k,t})^\top$, where all $S_{i,t}$ are with values in $\mathbb{N}_0$ and $S_{i,t}$ is the (non-negative) number of particles of type i that exist at t. The initial state $\boldsymbol{S}_0$ is deterministic.

5. The chosen decision $\boldsymbol{x}_t$ affects the lifespan distributions (however the distributions remain exponential), the distributions for the number of particles in the offspring, and the probabilities for mutation within the offspring, of all particles that exist at t. Thus, the types of all particles that exist at t are modified as a consequence of $\boldsymbol{x}_t$. Particles that exist at t can create only particles that are of the modified types, hence only the modified types are being propagated until $t + 1$.

   (a) Corresponding to the i-th coordinate of $\boldsymbol{x}_t$ is the row vector $\boldsymbol{m}_i(x_{i,t}) = \big(m_{i1}(x_{i,t}), \ldots, m_{ik}(x_{i,t})\big)$, where $m_{ij}(x_{i,t})$ denotes the expected number of type j particles at $t + 1$ within a MBHBPM with exponential lifespan distributions starting at t with a single particle of type i under decision $x_{i,t}$. We assume that $m_{ij}(x_{i,t}) < \infty$ for all $x_{i,t} \in \widetilde{\mathcal{X}}_i$ and $i, j = 1, \ldots, k$. Given $\boldsymbol{x}_t \in \mathcal{X}$, we organize the expectations into matrix $M(\boldsymbol{x}_t) = (\boldsymbol{m}_1(x_{1,t}), \ldots, \boldsymbol{m}_k(x_{k,t}))^\top$.

Proposition 3.5. *SDP Model 2 is a SDP within the "Universal Modeling Framework".*

Theorem 3.3. *For SDP Model 2, the value function $V_t(\boldsymbol{S}_t)$ satisfies*

$$V_t(\boldsymbol{S}_t) = \boldsymbol{S}_t^\top \mathcal{R}^{T-t}\boldsymbol{c}_T, \qquad t = 0, 1, \ldots, T-1. \tag{3.16}$$

Policy π, with corresponding $X_t^\pi(\cdot)$, that computes decisions $\boldsymbol{x}_t$ satisfying

$$\boldsymbol{c}(\boldsymbol{x}_t) + \gamma M(\boldsymbol{x}_t)\mathcal{R}^{T-t-1}\boldsymbol{c}_T = \mathcal{R}^{T-t}\boldsymbol{c}_T, \qquad t = 0, 1, \ldots, T-1, \tag{3.17}$$

is optimal.

3.6 SDPs with underlying MSBPM dynamics

Our contributions within this Section are: 1) We construct a novel state space and show that, with respect to it, the multi-type Sevastyanov branching process, as well as the MSBPM, constitute discrete-time Markov chains with respect to $t = 0, 1, \ldots, T$; 2) For these processes, we construct SDPs within the "Universal Modeling Framework". The contributions of this Section have not been published yet.

Definition 3.5. *Let there be k types. For each type denote with $\mathcal{D}_i$ the set of all 2-tuples of the following form:*

1. *The first element of the tuple is an integer. We denote this integer with r, $r \in \mathbb{N}_0$.*
2. *The second element of the 2-tuple is a r-tuple. We denote this r-tuple with $\boldsymbol{l}$. Each element l_i of $\boldsymbol{l}$ is a non-negative real number, i.e., $l_i \in \mathbb{R}_+$. The numbers within $\boldsymbol{l}$ are ordered from smallest to largest. Duplication is allowed in which case duplicating numbers are written next to each other.*

At t, associate with each $\mathcal{D}_i$ probability distributions. We will not write these distributions explicitly, but will consider them implicitly known. Denote $\mathcal{D}_i$ with associated distributions at t as $\mathcal{D}_{i,t}$. For the collection of k types at t, denote $\mathcal{D}_t^k = \mathcal{D}_{1,t} \times \mathcal{D}_{2,t} \times \cdots \times \mathcal{D}_{k,t}$.

Proposition 3.6. *A multi-type Sevastyanov branching process, with states defined as the elements of $\mathcal{D}_t^k$, is a discrete-time Markov chain with respect to $t = 0, 1, \ldots, T$.*

Definition of SDP Model 3. Define SDP Model 3 as the finite-horizon SDP that corresponds to the definition of SDP Model 1 upon which the following modifications are applied:

1. We observe a MSBPM, as defined in Definition 2.1, at epochs indexed with t, $t = 0, 1, 2, \ldots, T$.
2. Let k be the number of types of particles within the MSBPM. The state space is $\mathcal{D}_t^k$. The t index in $\mathcal{D}_t^k$ indicates that there are probability distributions and p.g.f.s associated with each type that may change with t (after a decision has been made). The state of the process at t is also called the "generation" or the "population" at t. The state of the process at t is given by $\boldsymbol{S}_t = (S_{1,t}, S_{2,t}, \ldots, S_{k,t})^\top$, where $S_{i,t} \in \mathcal{D}_{i,t}$ with the interpretation of the first component of $S_{i,t}$ being the number of particles of type i that exist at t and the interpretation of the second component of $S_{i,t}$ being the ages of each particle of type i that exists at t. The initial state $\boldsymbol{S}_0$ is deterministic.

5. The chosen decision $\boldsymbol{x}_t$ affects the lifespan distributions, the distributions for the number of particles in the offspring, and the probabilities for mutation within the offspring, of all particles that exist at t. Thus, the types of all particles that exist at t are modified as a consequence of $\boldsymbol{x}_t$. Particles that exist at t can create only particles that are of the modified types, hence only the modified types are being propagated until $t+1$. $\boldsymbol{x}_t$ does not affect the age of particles that exist at t.

Proposition 3.7. *SDP Model 3 is a SDP within the "Universal Modeling Framework".*

Unfortunately, for SDP Model 3 there is no analogue of Theorem 3.2 that allows for efficiently finding the solution. Currently the only algorithms available to us are generic dynamic programming algorithms that require iterating the entirety of the state and decision space. As the state space associated with Definition 3.5 requires information about particle age, iterating over all possible states is practically impossible.

We highlight that stochastic SDPs (and generally stochastic problems) are one of the most difficult optimization problems within the field of optimization. Nice and compact results are seldom available and different problems may require specialized algorithms solely designed for them. The mere existence of the 15 fragmented communities, discussed in Section 1.5 of the dissertation, that deal with sequential decision problems, testifies to the lack of an overarching approach that can handle a sufficiently large class of problems. In this context, the successful incorporation of the MSBPM (and other branching processes) into SDPs within the "Universal Modeling Framework" is a significant development, as it provides us with Bellman's optimality equation as a possible tool to be used when searching for solutions.

The validity of Bellman's optimality equation within the "Universal Modeling Framework" allows us to consider approaches such as Approximate Dynamic Programming (ADP; see [78]). Within the dissertation, we outline a general ADP algorithm based around the "post-decision state". This algorithm can serve as a starting point for the development of a more specialized ADP algorithm targeted at SDPs with branching process based dynamics.

Conclusion

Approbation

Results from the dissertation have been presented at: FMI Spring Scientific Session (March 2019, 2021, Sofia, Bulgaria), National Seminar on Probability and Statistics (June 2019, Sofia, Bulgaria), 21st European Young Statisticians Meeting (29 July - 02 August 2019, Belgrade, Serbia), Sofia University Young Researchers Conference (February 2020, Sofia, Bulgaria), The 19th Conference of the Applied Stochastic Models and Data Analysis International Society ASMDA2021 and DEMOGRAPHICS2021 WORKSHOP (June 2021, Athens, Greece), The 5th International Workshop on Branching Processes and their Applications - IWBPA 2021 (April 2021, Badajoz, Spain).

The following publications were written during the writing of the dissertation:

1. M. Slavtchova-Bojkova, K. Vitanov. Modelling cancer evolution by multi-type age-dependent branching processes. Comptes rendus de l'Acade'mie bulgare des Sciences, **71**, 10, 1297-1305, (2018).
2. M. Slavtchova-Bojkova, K. Vitanov. Multi-type age-dependent branching processes as models of metastasis evolution. Stochastic Models, **35**, 284-299, (2019), https://doi.org/10.1080/15326349.2019.1600410.
3. K. Vitanov, M. Slavtchova-Bojkova. On decomposable multi-type Bellman-Harris branching process for modeling cancer cell populations with mutations. 21st European Young Statisticians meeting - Proceedings, 113-118, (2019).
4. K. Vitanov, M. Slavtchova-Bojkova. Modeling escape from extinction with decomposable multi-type Sevastyanov branching processes. Stochastic Models, (2022), https://doi.org/10.1080/15326349.2022.2041037.

Scientific contributions

Within this dissertation the novel Multi–type Sevastyanov Branching Processes through probabilities of Mutation between types (MSBPM) is developed and explored in the context of populations escaping extinction. Unlike previous works in the field, the MSBPM and the results obtained do not depend on assumptions about mutations being small quantities or particular lifespan distributions nor on assumptions of non-decomposability or particular reproduction rates. As such, the novel MSBPM and the associated novel results constitute a continuous-time extension and/or generalization of previously obtained results by other authors concerning populations escaping extinction (see, e.g., [61], [62], [64], [65]) as well as a continuation of our previous results in the same field within Vitanov & Slavtchova-Bojkova [7] (2022) as well as preceding papers [1] - [6]. Various systems of equations have been obtained - systems of equations for the probability generating functions (p.g.f.s) of the process, for the probabilities of extinction, for the p.g.f.s of particle production from one class of particle types to another. Results concerning the time until occurrence of the first "successful" particle as well as the immediate risk of a "successful" particle emerging have also been obtained. To the best of our knowledge, such an in-depth investigation of the topic has not been done previously for multi-type, continuous-time branching processes (excluding our earlier work in [7] as well as preceding papers [1] - [6]). Aforementioned results have been obtained for the case of the MSBPM starting with one particle of age 0 and for the case of the MSBPM starting with one particle of age a, $a \neq 0$. The latter case, to the best of our knowledge, has not been explored previously in a systematic manner within the context of branching processes. Particular cases of decomposable MSBPMs have also been considered in the manner described above. Numerical schemes for calculating all obtained systems of equations have been developed.

Multi-type Bienaymé-Galton-Watson (BGW) branching processes, multi-type Bellman-Harris branching process with exponential lifespan distributions, multi-type Sevastyanov branching process, as well the MSBPM and its variants, have been successfully incorporated into Sequential Decision Problems (SDPs) within the "Universal Modeling Framework" developed in [82]. To the best of our knowledge, with the exception of the BGW branching process, branching processes have not been considered in the context of SDPs (within the "Universal Modeling Framework" or in other modeling frameworks). This incorporation formally opens the gate for techniques such as Approximate Dynamic Programming (ADP) and Reinforcement Learning (RL) to be applied onto SDPs with underlying branching process based dynamics. A novel proof of Theorem 3.1 from [77], concerning an efficient algorithm for finding the solution of SDPs with multi-type BGW branching processes dynamics, that uses Bellman's optimality equation, has been obtained. An analogous novel result for the case of the multi-type Bellman-Harris branching process with exponential lifespan distributions, as well as for the case of the Multi-type Bellman-Harris Branching Process through probabilities of Mutation between types (MBHBPM) with exponential lifespan distributions, has been identified. A novel state space has been constructed for the purpose of successfully incorporating the MSBPM and the multi-type Sevastyanov branching process into SDPs within the "Universal Modeling Framework".

Note regarding used software

All computations within the dissertation are done via code written in Python 3.8.13 [209]. The code uses the NumPy 1.20.3 [210] and SciPy 1.6.2 [211] libraries. Figures, that do not contain nodes, are created with Matplotlib 3.5.1 [212]. Figures that contain nodes are created with yEd 3.20.1 [213].

Declaration for originality of obtained results

I declare that the current dissertation "Branching processes - optimization and applications" contains original results obtained as a product of my research (supported by my supervisor). Results that have been obtained, published, or described by other authors are appropriately cited within the Bibliography.

The dissertation has not been applied for the purpose of obtaining a scientific degree from another school, university, or scientific institute.

Gratitude

I would like to thank my supervisor, prof. Maroussia Slavtchova-Bojkova, for our discussions, fruitful exchange of ideas, and support throughout the years! I am also grateful to my colleagues at the Faculty of Mathematics and Informatics of Sofia University "St. Kliment Ohridski" for always being kind and helpful!

Bibliography

[1] M. Slavtchova-Bojkova. On two-type decomposable branching processes in continuous time and time to escape extinction. In Branching Processes and their Applications, Lecture Notes in Statistics, **219**, Puerto, I.M., et al. Eds., Springer International Publishing: Switzerland, 319-329, (2016).

[2] M. Slavtchova-Bojkova, P. Trayanov, S. Dimitrov. Branching processes in continuous time as models of mutations: Computational approaches and algorithms. Computational Statistics and Data Analysis, **113**, 111–124, (2017).

[3] K. Vitanov, M. Slavtchova-Bojkova. Multitype branching processes in continuous time as models of cancer. Annuaire de l'Universite de Sofia "St. Kl. Ohridski", Fac. Math and Inf., **104**, 193-200, (2017).

[4] M. Slavtchova-Bojkova, K. Vitanov. Modelling cancer evolution by multi-type age-dependent branching processes. Comptes rendus de l'Acade'mie bulgare des Sciences, **71**, 10, 1297-1305, (2018).

[5] M. Slavtchova-Bojkova, K. Vitanov. Multi-type age-dependent branching processes as models of metastasis evolution. Stochastic Models, **35**, 284-299, (2019), https://doi.org/10.1080/15326349.2019.1600410.

[6] K. Vitanov, M. Slavtchova-Bojkova. On decomposable multi-type Bellman-Harris branching process for modeling cancer cell populations with mutations. 21st European Young Statisticians meeting - Proceedings, 113-118, (2019).

[7] K. Vitanov, M. Slavtchova-Bojkova. Modeling escape from extinction with decomposable multi-type Sevastyanov branching processes. Stochastic Models, (2022), https://doi.org/10.1080/15326349.2022.2041037.

[8] B. A. Sevastyanov. Branching processes. Mir, Moscow, (1971) (in Russian).

[10] P. Haccou, P. Jagers, V. Vatutin. Branching processes: Variation, Growth and Extinction of Populations. Cambridge University Press, (2007).

[14] M. Kimmel, D. Axelrod. Branching Processes in Biology. Springer: New York, (2002).

[36] E. Pardoux. Probabilistic Models of Population Evolution. Springer International Publishing: Switzerland, (2016).

[57] K. Athreya, P. Ney. Branching Processes. Springer: New York, (1972).

[58] T. Harris. The Theory of Branching Processes. Springer: Berlin, (1963).

[61] Y. Iwasa, F. Michor, M. Nowak. Evolutionary dynamics of escape from biomedical intervention. Proc. Biol. Sci. B, **270**, 2573-2578, (2003).

[62] Y. Iwasa, F. Michor, M. Nowak. Evolutionary dynamics of invasion and escape. J. Theor. Biol., **226**, 2, 205-214, (2004).

[63] P. Haccou, V. Vatutin. Establishment success and extinction risk in autocorrelated environments. Theoretical Population Biology, **64**, 303–314, (2003).

[64] M. Serra, P. Haccou. Dynamics of escape mutants. Theor. Popul. Biol., **72**, 167-178, (2007).

[65] S. Sagitov, M. Serra. Multitype Bienaymé–Galton–Watson processes escaping extinction. Advances in Applied Probability, 41(1), 225-246 (2009).

[66] A. Ghosh, M. Serra, P. Haccou. Quantifying stochastic introgression processes in random environments with hazard rates. Theoretical Population Biology, **100**, 1-5 (2015).

[67] R. Bellman. On the theory of dynamic programming. Proceedings of the National Academy of Sciences **38** (8): 716–719 (1952).

[68] R. Bellman. Dynamic Programming. Princeton University Press, Princeton, NJ (1957).

[69] R. Howard. Dynamic programming and Markov processes. Cambridge, MA: MIT Press (1960).

[70] M. L. Puterman. Markov Decision Processes, 2nd ed. John Wiley & Sons, Hoboken, NJ, (2005).

[74] K. Judd. Numerical Methods in Economics. MIT Press (1998).

[76] J. Si, A. Barto, W. Powell, and D. Wunsch. Learning and Approximate Dynamic Programming: Scaling up to the Real World. New York: John Wiley and Sons (2004).

[77] S. R. Pliska. Optimization of multitype branching processes. Management Science, **23**, 2, 117-124, (1976), http://www.jstor.org/stable/2629818.

[78] W. B. Powell. Approximate Dynamic Programming: Solving the Curses of Dimensionality, Second Edition. John Wiley & Sons, (2011), DOI:10.1002/9781118029176.

[80] W. B. Powell. A unified framework for stochastic optimization. European Journal of Operational Research, **275**, 3, 795-821, (2019), https://doi.org/10.1016/j.ejor.2018.07.014.

[81] W. B. Powell. From Reinforcement Learning to Optimal Control: A unified framework for sequential decisions. In: Handbook on Reinforcement Learning and Optimal Control, Studies in Systems, Decision and Control, 29–74 (2021).

[82] W. B. Powell. Reinforcement Learning and Stochastic Optimization. Wiley (2022).

[83] I. J. Bienayme. De la loi de multiplication et de la duree des familles. Soc. Philomath. Paris, Extraits, Ser. 5, 37 – 39 (1845).

[84] C. C. Heyde, E. Seneta. I. J. Bienaymé: Statistical theory anticipated. Springer, New York (1977).

[85] H. W. Watson, F. Galton. On the probability of the extinction of families. The Journal of the Anthropological Institute of Great Britain and Ireland **4**, 138 - 144 (1875).

[88] J. F. Steffenson. On Sandsynligheden for at Afkommet uddor. Matem. Tiddskr. B, 19-23 (1930).

[89] A. Kolmogorov, N. Dmitriev. Branching random processes. Dokl. Acad. Nauk SSSR, **56**, No. 1, 7 – 10 (See, English transl., Selected works of A. N. Kolmogorov, vol. II: Probability theory and Mathematical Statistics, Kluwer, Dordrecht, 1992), 1947.

[104] R. Durrett. Branching process models of cancer. Springer, New York (2015).

[115] B. A. Sevastyanov, A. M. Zubkov. Controlled branching processes. Theory Probab. Appl. **19**, 14-24 (1974).

[120] M. G. Velasco, I. García, G. P. Yanev. Controlled Branching Processes. John Wiley & Sons (2018).

[121] M. González, I. Del Puerto, N. Yanev, G. Yanev. Controlled branching processes with continuous time. Journal of Applied Probability, 58(3), 830-848 (2021). doi:10.1017/jpr.2021.8

[124] N. M. Yanev, K. V. Mitov. Controlled branching processes with non- homogeneous migration. Pliska Stud. Math. Bulg. **7**, 90 – 96 (1984) (in Russian).

[181] I. M. Del Puerto, N. M. Yanev. Branching processes with multi-type random control functions. C. R. Acad. Bulg. Sci. **57**, No 6 , 29 - 36 (2004).

[182] I. M. Del Puerto, N. M. Yanev. Stationary distributions for branching processes with multi-type random control functions. J. Appl. Stat. Sci. **16**, 91 - 102 (2008).

[199] K. Etessami, M. Yannakakis. Recursive markov decision processes and recursive stochastic games. J. ACM **62**(2), 11:1-11:69 (2015), https://doi.org/10.1145/2699431

[200] K. Etessami, A. Stewart, M. Yannakakis. Polynomial time algorithms for branching markov decision processes and probabilistic min(max) polynomial bellman equations. Math. Oper. Res. **45**(1), 34-62 (2020). https://doi.org/10.1287/moor.2018.0970

[201] U. Rothblum, P. Whittle. Growth optimality for branching Markov decision chains. Math. Oper. Res. **7**(4):582–601 (1982).

[202] E. M. Hahn, M. Perez, S. Schewe, F. Somenzi, A. Trivedi, D. Wojtczak. Model-Free Reinforcement Learning for Branching Markov Decision Processes. In: Silva, A., Leino, K.R.M. (eds) Computer Aided Verification. CAV 2021. Lecture Notes in Computer Science, **12760**, Springer, Cham. (2021) https://doi.org/10.1007/978-3-030-81688-9_30

[203] R. Sutton, A. Barto. Reinforcement Learning: An Introduction, Second Edition. MIT Press, Cambridge, MA, (2018).

[209] Python Software Foundation. Python Language Reference, version 3.8.13, https://docs.python.org/3.8/reference/index.html

[210] C. R. Harris, K. J. Millman, S. J. van der Walt, et al. Array programming with NumPy. Nature, **585**, 357–362 (2020), https://doi.org/10.1038/s41586-020-2649-2

[211] P. Virtanen, R. Gommers, T. E. Oliphant, et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nature Methods, **17**, 261-272 (2020), https://doi.org/10.1038/s41592-019-0686-2

[212] J. D. Hunter. Matplotlib: A 2D Graphics Environment. Computing in Science & Engineering, **9**, 90-95 (2007).

[213] yWorks GmbH. (2019). yEd. Retrieved from https://www.yworks.com/products/yed