Hindawi Publishing Corporation
Journal of Probability and Statistics
Volume 2011, Article ID 867493, 39 pages
doi:10.1155/2011/867493

Research Article

Simulating the Emergence and Survival of Mutations Using a Self Regulating Multitype Branching Processes

Charles J. Mode,[1] Towfique Raj,[2] and Candace K. Sleeman[3]

[1] *Department of Mathematics, Drexel University, Philadelphia, PA 19104, USA*
[2] *Division of Genetics, Department of Medicine, Brigham and Women's Hospital/Harvard Medical School, Boston, MA 02115, USA*
[3] *NAVTEQ Corporation, Malvern, PA 19335, USA*

Correspondence should be addressed to Charles J. Mode, cjmode@comcast.net

Received 18 May 2011; Revised 19 August 2011; Accepted 24 August 2011

Academic Editor: Shein-chung Chow

It is difficult for an experimenter to study the emergence and survival of mutations, because mutations are rare events so that large experimental population must be maintained to ensure a reasonable chance that a mutation will be observed. In his famous book, *The Genetical Theory of Natural Selection*, Sir R. A. Fisher introduced branching processes into evolutionary genetics as a framework for studying the emergence and survival of mutations in an evolving population. During the lifespan of Fisher, computer technology had not advanced to a point at which it became an effective tool for simulating the phenomenon of the emergence and survival of mutations, but given the wide availability of personal desktop and laptop computers, it is now possible and financially feasible for investigators to perform Monte Carlo Simulation experiments. In this paper all computer simulation experiments were carried out within a framework of self regulating multitype branching processes, which are part of a stochastic working paradigm. Emergence and survival of mutations could also be studied within a deterministic paradigm, which raises the issue as to what sense are predictions based on the stochastic and deterministic models are consistent. To come to grips with this issue, a technique was used such that a deterministic model could be embedded in a branching process so that the predictions of both the stochastic and deterministic compared based on the same assigned values of parameters.

1. Introduction

Branching processes are a class of stochastic processes that deal with the dynamics of evolving populations and have an extensive literature dating back one hundred years or more. Some references to this early literature may be found in the book of Harris [1]. Other books

on branching processes include those of Mode [2], Athreya and Ney [3], Jagers [4], Asmussen and Hering [5], Kimmel and Axelrod [6], and Haccou et al. [7], which contain discussions of these classes of stochastic processes and their applications to biology and other fields. Branching processes continue to be an ongoing field of research as exemplified by the recent master's degree thesis of Alexander [8] as well as elsewhere. For more details, it is suggested that an interested reader consult the Internet. The working paradigm used in these books and in even more recent works have consisted, for the most part, of the applications of various methods of classical mathematical analysis and probability theory to deduce properties of a class of branching process under consideration. But with the development of user friendly computers in more recent years, it has become feasible to use some numerical and graphical methods to help elucidate some properties of a branching process, but the greater part of the working paradigm for most investigators, working within this class of stochastic processes, has been based on mathematics. In this paper, however, due to the nonlinear properties of the class of branching process under consideration, after the basics of a branching process have been formalized in terms of mathematics, the working paradigm used to analyze their properties and deduce the evolutionary implications of these properties will be based on Monte Carlo simulation methods. As will be shown by examples, even though branching processes with nonlinearities are difficult to analyze, using classical methods of mathematical methods, nevertheless, the application of Monte Carlo simulation methods yields results that have interesting evolutionary implications.

Fisher [9], in Dover reprint of an edition of book first published in the late 1920s, was the first to apply what is now known as a one-type Galton-Watson process to the study of the survival of mutations in an evolving biological population. Mode and Gallop [10], in a review paper on the applications of Monte Carlo simulation methods in the study and analysis of the widely used Wright-Fisher process of evolutionary genetics, suggested that multitype branching processes could be used to eliminate the assumption of constant population size which characterizes most applications of this process in the study of biological evolution. In this paper, an overview of multitype branching processes with discrete generations, as described in the books Mode [2] and Harris [1], will be given. In most classical branching processes, whether one type or multitype, under some conditions on the parameters which are known as the super critical case, a population evolving according to a branching process will increase geometrically so that total population size grows without bound, which is contrary to what is frequently observed in nature where environmental and other conditions limit total population size. To correct this undesirable property of classical branching process, a class of self-regulating process such that population size is constrained within limits will also be considered in this paper.

Another methodological issue will also be addressed in this paper. There are at least two schools of thought concerning the applications of mathematical models to biology. In one school, the view that predominates is that deterministic models are sufficient to describe and analyze biological evolutionary phenomena mathematically, and that the introduction of stochastic models leads to unnecessary complications. A second school holds that if an investigator restricts attention to solely deterministic models, random effects, that are thought by many to play a significant role in biological evolution, will be missed so that predictions based on deterministic models may be misleading. In order to come to grips with this methodological issue, methods will be described and implemented whereby it is possible to embed deterministic models in a stochastic process. In such formulations, the embedded deterministic model and the stochastic process have the same parameters so that for each choice of numerical values of the parameters, the predictions based on the deterministic

model can be compared with those of a statistical summary of a sample of Monte Carlo realizations of the process. As will be shown in illustrative computer experiments, when the emergence of new beneficial mutations in an evolving population are under consideration, the predictions based on the embedded deterministic model can be misleading when compared with of the results of a Monte Carlo simulation experiment. The technique of embedding deterministic models in stochastic processes has been used extensively in other work, see for example, Mode and Sleeman [11, 12].

A question that naturally arises is to what class of biological organisms will the multitype branching processes described in this paper apply? A short answer to this question is any class biological organisms that do not reproduce by sexual reproduction, that is, those classes of organisms in which offspring arise only from the matings of females and males such as in humans and other vertebrates. Another question that arises is when during the process biological evolution on planet earth did such classes of organisms exits and flourish? According to the interesting and compelling lectures by Martin and Hawks [13], on major transitions in evolution, after the formation of our Earth and Solar System about 4.5 to 4.6 billions years ago and the earth cooled for about another 6 hundred million years or so, single celled organisms, prokaryotes, arose and persisted for about 2 billion years. Interestingly, some of these prokaryotes had evolved a capability for photosynthesis, which resulted in the evolution of the earth's atmospheric oxygen that in turn made it possible for plants and animals of present-day earth to live and breathe.

Many single-cell organisms, such as bacteria, reproduce by a process called binary fission, that is, a cell divides resulting in two daughter cells which in turn divide and so on. The class of multitype branching processes, which will be the focus of attention in this paper, will not only accommodate reproduction by binary fission but also other types of reproduction in which a multicellular organism may reproduce by releasing spores which subsequently germinate and grow into bodies of cells resembling their parents. But, applications of the class of branching processes described in this paper are not necessarily confined to single or multicellular organisms. For they may also be applied to large chemical molecules that have the capability of self-reproduction or replication. In this connection, it would be of interest to consult lectures on the origin of life by Hazen [14].

Up until now, evolution in deep time has been the focus of attention, but when the present is considered there is an abundance of examples for which the class of multitype branching processes with mutations may be applied. An immediate example is the human microbiome, which consists of a large number of species of microorganisms with very small cells when compared to cells of humans. Indeed, it has been estimated that there are more nonhuman cells than human cells in our bodies. Ordinarily, these cells live with us in a state of symbiosis, but when mutations occur, a disease condition may result. Populations of microorganisms also inhabit the bodies of plants and animals that we depend on for food and when mutations in these organisms occur and cause disease, which in turn may have devastating consequences for our food supply. These examples seem to be an adequate biological justification for the class of branching processes that will be the focus of attention in this paper, but in the discussion mention will be made of some further caveats. When attempting to study the occurrence and survival of mutations experimentally, large populations must be maintained by an experimenter to ensure that the event of a rare mutation that will occur will be likely. It seems reasonable, therefore, to use Monte Carlo simulation methods to study the emergence and survival of mutations, because the evolutionary dynamics of a large population can be simulated on computers with relative ease.

Although it may not be clear from the computer simulation experiments reported in this paper that the embedded deterministic model plays an essential role in exploring for and finding points in high-dimensional parameter spaces such that the evolutionary trajectory computed using the embedded deterministic model suggests potentially interesting biological implications, it is actually a necessary part of the working paradigm. For if one were to carry out such exploratory experiments by computing samples of realizations of the stochastic process, it would become clear that this approach would be prohibitive, because of the long time periods required to complete Monte Carlo simulation experiments when computing the number of replications of the process of a 100 or more takes hours and sometimes days of computer time to complete. But, when using the embedded deterministic model, the time needed to compute a trajectory corresponding to some point in a high-dimensional parameter space may require only a few minutes or at most an hour or two.

This experimental process of exploring and finding points in a high-dimensional parameter space that suggests potentially interesting biological applications using the embedded deterministic model also has the potential for connecting the stochastic behavior of the class of branching processes under consideration with fields of deterministic dynamics that have been studied extensively. For at some points in the high-dimensional parameter space, the computed trajectories of embedded nonlinear deterministic model exhibit periodicity and chaotic behavior, which provides a connection of branching processes to a field of nonlinear deterministic dynamics called chaos with an extensive literature. If a reader is interested in exploring the literature of this field, the book for the general reader by Gleick [15] may be consulted. In this connection, it is interesting to observe the essential role that computer-intensive methods played in the work of physicist Mitchell Feigenbaum whose numerical work led to a more general development of the field of chaos and its connections with Benoit Mandelbrot's concept of fractals. Another more technical book on chaos is the that of Gulick [16], and if a reader is interested in a series of informative lectures on this subject, the course by Strogatz [17] may be viewed.

In this paper the implications of chaos in the embedded deterministic model on the emergence of beneficial mutations in a population were not studied, but attention was focused on only a few points in the parameter space that led to lead to "regular" behavior of the trajectories of the embedded deterministic model. By considering these relatively few points in the parameter space of the model, it was possible to simulate the emergence of beneficial mutations in cases such that the driving forces of evolution were the reproductive success of a mutant type or it ability to compete for limiting resources within a background of lower probabilities of mutations per generation than those that had been studied heretofore as well as a case of neutral evolution where rare mutations may occur. However, a reader should be aware that before these few points in the parameter space were chosen, a rather large series of preliminary experiments with the embedded deterministic model were carried out.

2. Overview of Multitype Branching Processes with Discrete Generations

In this section, the formulation of a multitype branching process evolving on a time scale of discrete generations will be given along with a description of algorithms to compute a sample of Monte Carlo realizations of such a multitype branching processes. To illustrate the algorithms underlying a Monte Carlo simulation procedures to simulate a sample of realizations of this stochastic process, for the sake of simplicity, attention will be focused

on the case of $m = 3$ types, which will be referred as genotypes. Let $\mathbb{G} = \{1, 2, 3\}$ denote the set of three genotypes. Initially, it will be assumed that two components, reproductive success per individual and mutations among the three genotypes per generation, are the forces driving the evolution of a population. Reproductive success will be characterized in terms of random variables N_ν for $\nu \in \mathbb{G}$, taking values in the set $\mathbb{N} = \{n \mid n = 0, 1, 2, 3, \ldots\}$ of nonnegative integers and representing the number of offspring produced by each genotype ν per generation. The probability density functions of these random variables will be denoted by

$$P[N_\nu] = g_\nu(n), \quad \text{for } n \in \mathbb{N}, \tag{2.1}$$

and $\nu \in \mathbb{G}$. A useful measure of reproductive success for each individual per generation is the expected value

$$E[N_\nu] = \lambda_\nu \geq 0, \tag{2.2}$$

of the number of offspring per generation for each genotype $\nu \in \mathbb{G}$.

An experimenter is free to choose any parametric or nonparametric form of the probability density functions for the random variables $N_\nu, \nu \in \mathbb{G}$, but for the sake of simplicity and ease of computation and interpretation, it will be assumed in this paper that these densities have the simple Poisson form

$$g_\nu(n) = \exp[-\lambda_\nu] \frac{\lambda_\nu^n}{n!}, \tag{2.3}$$

for $n \in \mathbb{N}$ and $\nu \in \mathbb{G}$. It is easy to show that for the density in (2.3), $E[N_\nu] = \lambda_\nu > 0$ so that the measure of reproductive success λ_ν is the parameter for a Poisson density for each genotype. In the experiments reported in this paper, a decision was made to consider only the special case of Poisson distributions for the random variables N_ν for $\nu = 1, 2, 3$. It is well known for this simple distribution that the expectation and variance both equal the parameter λ. Thus, the use of the Poisson distribution with equal expectation and variance provides a useful set of computer experiments for distributions with this property that may form a basis for comparisons with other experiments depending on other values of parameters or even offspring distributions belonging to other families of distributions. It also seems to be a reasonable assumption for populations that reproduce by a process of binary fission, cell division, that the Poisson distribution may be a reasonable choices for offspring distributions for assignments of the parameters near the number 2. It should be mentioned that it would be straightforward to replace the Poisson offspring distribution that is currently in the software with some other distribution of interest to an investigator that would be appropriate model for the mechanism of reproduction for a particular species under consideration.

Another family of parametric distribution that could be chosen for offspring distributions is the two-parameter negative-binomial family. For this family, the expectation and variance are not equal, and it seems reasonable to expect that if the variances of the offspring distributions were greater than the expectation, then the variation among the realizations of a process would be greater than that for the Poisson case. Given any distribution with a finite variance, however, one would expect, a priori, that if one of the three parameters $(\lambda_1, \lambda_2, \lambda_3)$ is greater than the other two, then the genotype corresponding to

this number would become predominate in a population as evolutionary time progresses, and, in this sense, the process may be robust to the choices of offspring distributions. It should also be mentioned that if two parameter offspring distributions were used, then the dimension of the parameter space of the model would be increased, and, as will be mentioned subsequently, high-dimensional parameter spaces complicate the design and execution of computer simulation experiments.

A basic tenet of Darwinian evolution is that natural selection acts on variants, genotypes, in a population so that those genotypes, which arose as a result of beneficial mutations, become predominant in the long-run evolution of a population. Thus, to take into account novel variations of genotypes that may occur in an evolving population, mutations need to be taken into account. As a first step in formalizing the notion of mutation, let $\mu_{ij} \geq 0$ denote the conditional probability per generation that genotype i produces an offspring of genotype j for $i, j = 1, 2, 3$. For a more complete interpretation of these conditional probabilities, it is useful to represent them in the matrix form

$$\mathbb{M} = \begin{pmatrix} \mu_{11} & \mu_{12} & \mu_{13} \\ \mu_{21} & \mu_{22} & \mu_{23} \\ \mu_{31} & \mu_{32} & \mu_{33} \end{pmatrix}. \tag{2.4}$$

In this matrix, μ_{11}, for example, is the conditional probability that an individual of genotype 1 produces an offspring of genotype 1, but μ_{12} and μ_{13} are, respectively, the conditional probabilities that an individual of genotype 1 produces an offspring of genotype 2 or 3. For every $i \in \mathbb{G}$,

$$\sum_{j \in \mathbb{G}} \mu_{ij} = 1, \tag{2.5}$$

so that each row of the matrix $\mathbb{M}$ may be thought of as the vector of probabilities for a three dimensional multinomial distribution. Let the row vector $\mathbf{p}_i = (\mu_{i1}, \mu_{i2}, \mu_{i3})$ denote row $i \in \mathbb{G}$ of the matrix $\mathbb{M}$ in (2.4). Because of condition (2.5), it can be seen that one is free to choose 6 parameters of the model. Then, given these choices, constraint (2.5) will determine the other three parameters.

At this point, enough formal notation has been defined to begin a description of a set of algorithms for computing Monte Carlo realizations of the multitype branching process under consideration. Consider any individual of genotype $i \in \mathbb{G}$ in any generation. By assumption, the number of offspring produced by any individual of genotype $i \in \mathbb{G}$ in any generation is a realization $\mathbf{n}_i$ of a Poisson random variable N_i with parameter λ_i. The set of possible values of $\mathbf{n}_i$ is $\mathbb{N} = \{n \mid n = 0, 1, 2, 3, \ldots\}$, the set of nonnegative integers. To take the possibility of mutation into account, among the $\mathbf{n}_i$ offspring, let $\mathbf{x}_{ij} \geq 0$ denote the number of offspring of genotype $j \in \mathbb{G}$ produced by an individual of genotype $i \in \mathbb{G}$ such that these nonnegative integers satisfy the constraint

$$\mathbf{n}_i = \sum_{j \in \mathbb{G}} \mathbf{x}_{ij}. \tag{2.6}$$

Let $\mathbf{x}_i = (\mathbf{x}_{i1}, \mathbf{x}_{i2}, \mathbf{x}_{i3})$ denote a vector whose elements are in the sum (2.6). Then for $\mathbf{n}_i \neq 0$, a realization of the vector $\mathbf{x}_i$ is computed as a realization from a multinomial distribution with index, sample size, $\mathbf{n}_i$ and probability vector $\mathbf{p}_i = (\mu_{i1}, \mu_{i2}, \mu_{i3})$. If $\mathbf{n}_i = 0$, then $\mathbf{x}_i = \mathbf{0}$, the zero vector. If a reader is interested in more details, Mode and Gallop [10] may be consulted for a detailed description of an algorithm for computing realizations of a random vector from a multinomial distribution, given an index n and a multinomial probability vector $\mathbf{p}$.

In generation t, where $t = 0, 1, 2, 3, \ldots$, let the random function $X_i(t)$, taking values in the set $\mathbb{N}$, denote the number of individuals of genotype $i \in \mathbb{G}$ in generation t, and let

$$\mathbf{X}(t) = (X_1(t)), X_2(t), X_3(t) \tag{2.7}$$

denote a vector of these random functions. For $t = 0$, an experimenter needs to assign initial values in the set $\mathbb{N}$ to each of the elements in the vector (2.7). In generation $t > 0$, let the random function $Y_j(t)$ denote the total number of offspring of genotype $j \in \mathbb{G}$ produced by all three genotypes in generation t, and let the random function $Y_{ij}(t)$ demote the number of offspring of genotype j produced by the $X_i(t)$ individuals of genotype $i \in \mathbb{G}$ in generation t. If $X_i(t) = 0$, then $Y_{ij}(t) = 0$. But if $X_i(t) \neq 0$, then, given $X_i(t)$, let $(\mathbf{x}_{ij}^{(k)})$, for $k = 1, 2, \ldots, X_i(t)$, denote a sequence of conditionally independent random variables defined as in (2.6). Then, for every $j \in \mathbb{G}$, the random function $Y_j(t)$ is given by the formula

$$Y_j(t) = \sum_{i=1}^{3} Y_{ij} = \sum_{i=1}^{3} \sum_{k=1}^{X_i(t)} \mathbf{x}_{ij}^{(k)}. \tag{2.8}$$

Not all the offspring in generation t will survive and produce offspring in generation $t + 1$; consequently, in order to take into account those offspring who survive to produce the next generation, a parametric from of a survival function will be introduced. Among the canonical forms of widely used survival functions in applied probability is the Weibull-type which has the parametric form

$$\exp\left[-(\beta t)^{\alpha}\right], \tag{2.9}$$

where α and β are positive parameters. Usually, this function has the following interpretation. Let $\mathbb{R}$ denote the set of points on the real number line and consider a random variable T taking values in the set

$$[0, \infty) = \{x \in \mathbb{R} \mid x \geq 0\} \tag{2.10}$$

of nonnegative real numbers. Next suppose that the random variable T denotes the random lifespan of some piece of equipment or individual. Then, the probability that this individual survives beyond time $t > 0$ is assigned the probability

$$P[T > t] = \exp\left[-(\beta t)^{\alpha}\right], \tag{2.11}$$

where, in most applications, the values of the parameters α and β must either be assigned or estimated from data.

In the class of self-regulating branching under consideration, however, the formula in (2.11) will be used to formalize the conditional probability that an offspring of genotype $j \in \mathbb{G}$ in some generation t survives to produce offspring in generation $t+1$. In the formulation under consideration, it will be assumed that this conditional probability depends on total population size in generation t. By definition, total population size in generation t is given by the random function

$$Z(t) = \sum_{j \in \mathbb{G}} X_j(t), \tag{2.12}$$

see vector (2.7). Given a value of this random function in generation t, let $S_j(t \mid Z(t))$ denote the conditional probability that an individual offspring of genotype $j \in \mathbb{G}$ in generation t survives to produce offspring of generation $t+1$. By assumption, this conditional probability has the parametric form

$$S_j(t \mid Z(t)) = \exp\left[-\left(\beta_j Z(t)\right)^{\alpha_j}\right], \tag{2.13}$$

where the pair of parameters (α_j, β_j) may vary among genotypes $j \in \mathbb{G}$. For the sake of simplicity, in all the computer experiments reported in this paper, it was assumed that $\alpha_j = 2$ for all $j \in \mathbb{G}$. Observe that the survival function in (2.13) decreases as $Z(t)$ increases and that this decrease is more rapid for those t such that $\beta_j Z(t) > 1$. For example, if $\beta_j = 10^{-8}$, then the condition $10^{-8} Z(t) > 1$ implies that $Z(t) > 10^8$ so that when total population size reached this level, an offspring of genotype $j \in \mathbb{G}$ is less likely to survive and produce offspring in the next generation. The rationale for choosing the alpha a constant 2 will be discussed briefly in the next section.

Given the parametric survival function in (2.13), a Monte Carlo realization of the random function $X_j(t+1)$, the number of offspring of genotype $j \in \mathbb{G}$ in generation t who survive to produce the offspring of generation $t+1$ is easy to compute. Let $p(t) = S_j(t \mid Z(t))$ denote the conditional probability in (2.13) and let $Y_j(t)$ denote the random function in (2.8). Then, by assumption, $X_j(t+1)$ is a realization of a binomial random variable with index $Y_j(t)$ and probability $p(t)$. In subsequent sections in which computer experiments on quantifying selection and mutation are under consideration, rationales for choosing values of the β parameters by genotype will be discussed.

Before proceeding to the next section, it will be of interest to mention an alternative algorithm for computing realizations of the multitype branching process under consideration. Of all the Monte Carlo simulation procedures discussed in this section, that entailed in (2.8) will be the most time consuming, particularly if the random function $X_i(t)$ has a large value for some genotype. When $X_i(t)$ is large, a realization of the random variable $Y_j(t)$ may be computed more efficiently by using a central limit theorem approximation. Let N_i denote a random variable representing the number of offspring produced by an individual of genotype $i \in \mathbb{G}$ in any generation, and let η_i and σ_i^2 denote, respectively, the expectation and variance of the random variable N_i. Then, let

$$\left(N_i^{(k)} \mid k = 1, 2, \ldots, X_i(t)\right) \tag{2.14}$$

be a collection of independent and identically distributed random variables whose common distribution is that of the random variable N_i. Then, the random variable

$$H_i = \sum_{k=1}^{X_i(t)} N_i^{(k)} \tag{2.15}$$

is approximately normally distributed with a conditional expectation of $X_i(t)\eta_i$ and variance $X_i(t)\sigma_i^2$. Let Z be a realization of a standard normal random variable with expectation 0 and variance 1. Then,

$$\widehat{H_i} = \left[X_i(t)\eta_i + \sqrt{X_i(t)}\sigma_i Z\right], \tag{2.16}$$

where the function $[x]$ stands for the greatest nonnegative integer in the real number x, is a central limit theorem approximation to the integer valued random variable H_i in (2.15). By definition, if $x < 1$, then $[x] = 0$ and $1 < x \leq 2$, then $[x] = 1$ and so on. Let

$$\mathbf{Y}(t) = (Y_1(t), Y_2(t), Y_3(t)) \tag{2.17}$$

denote the vector valued random function to be approximated, see (2.8). Then, in a Monte Carlo simulation experiment such that $X_i(t)$ is large, let the random function $Y_{ij}(t)$ denote the total number of offspring of genotype j produced by the $X_i(t)$ individuals of genotype i in generation t. Then, $Y_{ij}(t)$ is component j in a vector realization from a multinomial distribution with index $\widehat{H_i}$ and probability vector $\mathbf{p}_i = (\mu_{i1}, \mu_{i2}, \mu_{i3})$. Thus, the total number of offspring of genotype j produced by the three genotypes in generation t is, by definition, given by the random function

$$Y_j = Y_{\circ j} = \sum_{i=1}^{3} Y_{ij}, \tag{2.18}$$

for $j = 1, 2, 3$.

In the software used to carry out the simulation experiments reported in this paper, the central limit theorem just outlined was used when $X_i(t)$ was large and the random variable in the collection in (2.14) had independent Poisson distributions. But, when $X_i(t)$ was small, the central limit approximation was not applied. It is easy to see that the procedures described for the case of $m = 3$ genotypes may be easily extended to the case of $m > 3$ genotypes. If one would like to see further details on computing Monte Carlo realizations of the class of multitype branching processes under consideration, which have been omitted in this overview, the book by Mode and Sleeman [12] may be consulted.

3. On Embedding a Deterministic Model in a Stochastic Process

Let $(X(t) \mid t = 0, 1, 2, \ldots)$ be stochastic process evolving in discrete time that takes values in the set $\mathbb{R}$ of real numbers. Furthermore, suppose that $E[X^2(t)]$ is finite for all $t = 0, 1, 2, \ldots$.

Then, it is well known that for $t+1$, if one wishes to find the best estimator $\widehat{X}(t+1)$ of the random function $X(t+1)$ such that

$$E\left[\left(X(t+1)-\widehat{X}(t+1)\right)^2\right] \tag{3.1}$$

is a minimum, $\widehat{X}(t+1)$ must be chosen as the conditional expectation

$$\widehat{X}(t+1)=E[X(t+1)\mid X(t)]. \tag{3.2}$$

For some stochastic processes, it is difficult to find a usable form of the conditional expectation on the right, but for the sake of simplicity, before considering a self-regulating multitype branching process, it will be helpful to consider the case of a one-type branching process, which is also known as a one-type Galton-Watson process. In what follows the idea of a conditional expectation will be used extensively.

For the case of a Galton-Watson process, let $(X(t)\mid t=0,1,2,\ldots)$ denote a sequence of random functions taking values in the set $\mathbb{N}$ of nonnegative integers and interpret $X(t)$ as the number of individuals in a population in generation $t\in\mathbb{N}$. Suppose that for each of these individuals, the number of offspring contributed to the generation $t+1$ is a realization of a random variable N taking values in the set of nonnegative integers $\mathbb{N}$ that has a Poisson distribution with parameter $\lambda>0$. Let the random function $X(t)$ denote the number of individuals in the population in generation t, and given $X(t)$, let N_k for $k=1,2,\ldots,X(t)$ denote a collection of conditionally independent and identically distributed random variables whose common distribution is that on the random variable N. Then, if $X(t)=0$, $X(t+1)=0$, but if $X(t)>0$, then $X(t+1)$ is the random sum

$$X(t+1)=\sum_{k=1}^{X(t)}N_k. \tag{3.3}$$

For each $k=1,2,\ldots,X(t)$, $E[N_k]=\lambda$. Therefore, for this process,

$$E[X(t+1)\mid X(t)]=E\left[\sum_{k=1}^{X(t)}N_k\mid X(t)\right]=X(t)\lambda. \tag{3.4}$$

At generation t, let $\eta(t)=E[X(t)]$ denote the unconditional expectation of the random function $X(t)$. Then, from (3.4), it follows that

$$\eta(t+1)=E[E[X(t+1)\mid X(t)]]=E[X(t)]\lambda=\eta(t)\lambda. \tag{3.5}$$

Therefore, if $X(0)=x_0$ is some assigned positive integer, which is interpreted as the number of individuals in the initial population, then (3.5) implies that

$$\eta(t)=x_0\lambda^t, \tag{3.6}$$

for every generation $t \in \mathbb{N}$. Thus, if λ is such that $0 < \lambda < 1$, then $\eta(t) \to 0$ as $t \to \infty$. And if $\lambda = 1$, then $\eta(t) = x_0$ for all $t \in \mathbb{N}$. But, if $\lambda > 1$, then $\eta(t) \to \infty$ as $t \to \infty$ so that the growth of the population is unbounded. As indicated in the introduction, such a model for an evolving population is not acceptable, because the size of any real biological population is limited by the available environmental resources.

To remove this limitation, it will be assumed that the process is self-regulating. For any realization $X(t) = x(t)$ of a one-type branching process in generation t, let

$$S(t \mid X(t)) = \exp\left[-\left(\beta X(t)\right)^{\alpha}\right] \tag{3.7}$$

denote the conditional probability that any offspring in generation t survives to produce offspring for generation $t+1$, where α and β are positive parameters which are to be assigned numerical values. Then, let $Y(t)$ denote the total number of offspring produced in generation t and suppose $Y(t)$ given by the random sum

$$Y(t) = \sum_{k=1}^{X(t)} N_k. \tag{3.8}$$

Next suppose, as in Section 2, that $X(t+1)$ is a realization of a binomial random variable with index $Y(t)$ and probability $S(t \mid X(t))$. Then, the conditional expectation of $X(t+1)$, given $Y(t)$, is

$$\begin{aligned} E[X(t+1) \mid Y(t)] &= Y(t)S(t \mid X(t)) \\ &= \left(\sum_{k=1}^{X(t)} N_k\right) S(t \mid X(t)), \end{aligned} \tag{3.9}$$

see (3.8). Therefore,

$$\begin{aligned} E[X(t+1) \mid X(t)] &= E\left[\left(\sum_{k=1}^{X(t)} N_k\right) S(t \mid X(t)) \mid X(t)\right] \\ &= X(t)S(t \mid X(t))\lambda. \end{aligned} \tag{3.10}$$

The random function on the right is the best predictor of the random function $X(t+1)$ for the self-regulating branching process under consideration. But, this value is completely known from a computational point of view only for the case $t = 0$. For in this case, it is assumed that $X(0) = x_0$, an assigned positive integer. Hence,

$$\widehat{X}(1) = x_0 S(t \mid x_0)\lambda. \tag{3.11}$$

But, this result for $t = 1$ suggests that, because $\widehat{X}(1)$ is known, it would be of interest to consider

$$\widehat{X}(2) = \widehat{X}(1)S\left(t \mid \widehat{X}(1)\right)\lambda \tag{3.12}$$

as an estimate of the random function $X(2)$. Of course, this result is not the best estimate of the random function $X(2)$ in the sense of (3.1) and (3.2), but it is useful from a computational point of view, because it reduces to a single value that may be easily computed. In general, we could continue this estimation procedure so that for $t \geq 0$, the random function $X(t+1)$ may be estimated according to the recursive formula

$$\widehat{X}(t+1) = \widehat{X}(t) \times S\left(t \mid \widehat{X}(t)\right) \times \lambda. \tag{3.13}$$

By definition, (3.13) may be viewed as a recursive deterministic model embedded in the self-regulating Galton-Watson process. Just as for the process, the trajectories based on the recursive relation in (3.13) depend only the three positive parameters α, β, and λ. Moreover, as will be shown in subsequent sections of this paper, by computing a sample of realizations of a multitype Galton-Watson process, it will be possible to compare such trajectories of the stochastic process such as quantiles, mean, and standard deviation trajectories to those based on the embedded recursive deterministic model in (3.13).

In (3.13) for the sake of simplifying the notation, let $x(t)$ denote the estimate $\widehat{X}(t)$ and let $f(x)$ denote the function of the right in (3.13). Note that the function $f(x)$ is a continuous function of x for $x \geq 0$. Then, if a sequence $(x(t) \mid t = 0, 1, 2, \ldots)$ that is computed recursively, where $x(0)$ is an assigned positive number, converges to a number x as $t \uparrow \infty$, then $x = f(x)$. For the case when the offspring distribution is a Poisson with parameter λ and the survival function is Weibull with parameters α and β, it is possible to show that the fixed point x as a simple function of the parameters α, β, and λ. Explicit forms of this function are given in the book Mode and Sleeman [12] for two choices of survivals functions in chapter 9 but these formulas will not be listed here for the sake of brevity. Furthermore, for these simple cases, it is possible to derive conditions, expressed in terms of the parameters, when the fixed point x is attracting (stable) or not attracting. For the case the fixed point x is attracting, the formula for x represents the total population size that a population evolving according the embedded deterministic model would attain in the limit.

From this result, for the Poisson case one may do simple numerical experiments by considering a set of values of the three parameters α, β, and λ as they affect asymptotic total population size in the stable case. Alternatively, by inspecting the form of this simple function for the fixed point x, it is possible to develop some intuition as what parameters are most important in determining population size in the limit. It should also be mentioned in passing that in chapter 9 referenced above, a different parametric form of the survival form was also used, and in this case the formula for the fixed point differs significantly from the Weibull case. As it turns out, for an embedded deterministic model in the multitype model that will be described below, it does not seem possible to derive a simple formula for the vector-valued fixed point in a three- or higher-dimensional space. Thus, from the biological and evolutionary point of view a one-type model is not very interesting, because it does not accommodate mutation, even though it may provide useful insights for the dynamic behavior of multitype models in the limit.

For the multitype branching process considered in section, a vector valued recursive deterministic model may be embedded in the stochastic process by extending the ideas outlined for the one-type case. Let the random row vector $\mathbf{X}(t) = (X_1(t), X_2(t), X_3(t))$ denote

the number of individuals of each of the three genotypes in generation $t \in \mathbb{N}$. Then, let $\mathbf{M}(t;\mathbf{X}(t))$ denote a 3×3 random matrix such that row i has the form

$$S_i(t \mid Z(t))\lambda_i\left(\mu_{i1}, \mu_{i2}, \mu_{i3}\right), \tag{3.14}$$

where $S_i(t \mid Z(t))$ is the survival function for individuals of genotype i, λ_i is the expected number of offspring produced by an individual of genotype i per generation and $\mathbf{Z}(t)$ is total population size in generation t. Observe that the matrix $\mathbf{M}(t;\mathbf{X}(t))$ contains the conditional expectations of the number of offspring of each genotypes contributed by individuals in generation t to the next generation, given the random vector $\mathbf{X}(t)$. For more details see in Section 2 see (2.4), (2.12), and (2.13). Then if $\widehat{\mathbf{X}}(0) = \mathbf{X}(0) = \mathbf{x}(0)$, an assigned vector of nonnegative integers, the embedded deterministic model for self-regulating multitype branching process is the recursive vector-matrix equation

$$\widehat{\mathbf{X}}(t+1) = \widehat{\mathbf{X}}(t)\mathbf{M}\left(t;\widehat{\mathbf{X}}(t)\right), \tag{3.15}$$

for $t \geq 0$. In subsequent sections of this paper, the results of Monte Carlo simulation experiments will be reported in which the evolutionary trajectories based on (3.15) will be compared with stochastic trajectories estimated from a Monte Carlo sample of realizations of the stochastic process.

At this point in the discussion, it is helpful to mention the rational underlying the decision to consider only cases in which the three alpha parameters for the three genotypes were set equal to two in all subsequent computer experiments reported in the paper. The dimension of the parameter space for the three-type model is six with respect to the survival component of the process, and when such large dimensional spaces are under consideration, it is helpful in planning computer experiments to reduce the dimension of the parameter space as much as feasible. To lend more clarity to the discussion, it will be useful to have access to the fixed point for the one-type case for the case of a Poisson offspring distribution and a Weibull-type survival function. In this simple case, the formula for the fixed point has the simple form $x = (\ln\lambda)^{1/\alpha}/\beta$, where $\alpha > 0$, $\beta > 0$ and $\lambda > 1$. From this simple formula, it can be seen that of the fixed values of α and λ, x is large when β is small and large values of β will result in small values of x. For this simple case, for fixed values of β and λ, let $x(\alpha)$ denote the fixed point of a function of α. Then it is easy to see that $x(1) > x(2)$. Thus, for this case, if alpha is assigned the value $\alpha = 1$, the limiting population size would be larger than if alpha were assigned the value $\alpha = 2$. Although there is no proof that this result will also hold in a multidimensional case, it is thought that for a case in which all parameters were chosen as alpha equal to one, the limiting population size for each type would be larger than for alpha equal to two.

From an inspection of the fixed point formula shown above for the one type case, it was clear that the values of the beta parameter would be most important in determining final population size for the case the fixed was attracting or stable. It seems plausible that values of the beta parameters in the multitype case would be the most important for determining limiting population size for all genotypes in multitype models, but this conclusion was based on the results of exploratory computer experiments that are not reported in this paper. A possible exception to this statement is the case when the parameter α is small, but such cases were not considered. The lambda parameters also have clear biological interpretations as

a measure of reproductive success for individuals of each of the three genotypes. Consequently, from the biological point of view, it seemed to be of significant interest to allow a measure of reproductive success for each of the three genotypes under consideration. The assumption that the alpha parameters were constant reduced the dimension of parameter space with respect to the survival component of the model from 6 to 3 and thus simplified the problem of choosing interesting values of the parameters in computer simulation experiment. However, it should also be kept in mind that in the mutation matrix we are free to choose two values in each row of the matrix of mutation probabilities so that the dimension of the parameter space for the full model is 9, which will force an investigator to reduce the dimension of the parameter by making simplifying assumption whenever they seem reasonable. The decision of set all alpha parameters equal to 2 was subjective and was done primarily for the sake of simplicity, even at the time it was thought that if all alpha parameters were set equal to 1, the results would differ only quantitatively with respect to the limiting sizes of the subpopulations for each of the three genotypes.

To some readers it may seem strange, because little mention was made that it is well known that most mutations are deleterious and are either fatal or render the mutant genotype less fit than the parental genotype. But, the study of a two-type process accommodating a wild-type and deleterious mutant are not very interesting, because in such cases the wild type will predominate in the long-run evolution of the process, even though the mutant type may persist in the population in the long run because of recurrent mutations. However, in three-type models, it is possible to consider cases in which it is supposed that one genotype is the wild type of the founding population and the two mutant types may be such that one is deleterious and the other is beneficial so that, in principle, both types of mutations may be studied in a computer simulation experiment. In the computer simulation experiments that are reported in subsequent sections of this paper, these two types of mutations were tacitly assumed to be in force, and these ideas were quantified by choosing what seemed to be interesting assignments of parameter values that had not been considered heretofore.

It can be shown that for some choices of parameter values in the parameter space, the iterates of (3.13) and (3.15) can become chaotic, but such examples will not be presented in this paper. If a reader is interested in pursuing the subject of chaos, it is suggested that the book by Gulick [16] be consulted. Among the models considered by Gulick was the logistic equation, which has the form

$$f(x) = \mu x(1 - x), \tag{3.16}$$

where $\mu > 0$ and $0 \leq x \leq 1$. The properties of the iterates of (3.16) depend on the value of the parameter μ, and for some values of μ the iterates of the function in (3.16) will become chaotic. Actually, the function in (3.16) belongs to a class of functions whose iterates may become chaotic. Evidently, the function on the right in (3.13) also belongs to this class of functions. Moreover, the function on the right in (3.15) belongs to a class of vector-valued functions whose iterates may become chaotic at some points in the parameter space, but a discussion of the technical details characterizing these classes will not be given in this paper. If a reader is interested in more details regarding the embedding of a deterministic model in a self-regulating multitype branching process in cases in which the iterates of the embedded deterministic become chaotic, chapters 9 and 10 in the book by Mode and Sleeman [12] may be consulted.

To some, perhaps many, the computer experiments reported in this paper do not provide insights to a path toward a more general theoretical understanding of the class of stochastic process under consideration with embedded deterministic models that may have predictive value regarding realizations of a stochastic process. But, in the present state of the development of the art, well thought out computer simulation experiments are the best way, at the moment, for exploring what the implications of a formulation may have for useful interpretations for the study of biological evolution and in some cases chemical evolution. Of course, for one-type model, it is easier to do the analytics required for a deeper theoretical understanding of the system, but one-type models are not of significant biological interest when mutations are considered.

Given the present state of the art, computer simulation experiments can contribute information that will be useful in a more theoretical analysis of the implications of a system for biological evolution. For example, through computer experimentation it is known that large values for the lambda parameters will result in periodic and chaotic behavior of the iterates of the embedded deterministic model and such behavior is also reflected in the properties of the sample functions of the stochastic process. It is also known that smaller values of the lambda parameters will lead to a more stable behavior of the embedded deterministic model that is also reflected in the behavior of the sample functions of the stochastic process. Indeed, from a historical point of view, many subfields of probability and statistics, as well as mathematics in general, are rooted in practical problems from experience and the use of numerical experimentation to gain insights into solutions of these problems.

In the best of all worlds, a team of investigators would be working on the class of processes with embedded deterministic models such as that under consideration. Among the members of such a team would be an expert probabilistic analyst who would be potentially capable of handling the high dimensionality of parameter spaces that often arises when considering biological evolution and, at the same time, be acquainted with some of the literature on equations of deterministic dynamical systems with nonlinearities, which may have chaotic solutions. At times such a person may wish to call in a symbolic computation engine to do the messy algebra and numerical computations that often arise in dealing with problems of coping with multidimensional parameter spaces. For example, when the mutation matrix has relatively many rows and columns and is reducible, such a model would be of significant interest in the study of speciation in evolution but was not considered in this paper. Other members of such a team would include someone with extensive experience with modelling and a person with computer expertise. Another person of such a team would include a biologist who would assist in motivating the research problems to be considered and would also play role constructive criticism. Should any readers of this paper have the expertise of a probabilistic analyst, the existing team would welcome a cooperative arrangement with such an individual or individuals.

4. Beneficial Mutations and Differential Reproductive Success as Driving Forces of Evolution

The objective in this section is to report the results of a computer simulation experiment in which beneficial mutations and differential reproductive success were the driving forces of evolution. Within this framework, a beneficial mutation will be characterized in terms of whether one of the three genotypes under consideration has a reproductive advantage over the other two. To quantify this idea, a genotype will have a reproductive advantage, if,

on average, individuals of this genotype contribute more offspring to the next generation than individuals of the other two genotypes. In the experiment reported in this section, as well as those in subsequent sections, the evolutionary timespan considered was 6,000 generations, and in each Monte Carlo simulation experiment, 6,000 generations of evolution were replicated 100 times, which provided a sample of realizations of the stochastic process that could subsequently be used to compute statistically informative summarizations of the simulated data. Because the embedded deterministic model was, by definition, not stochastic, it sufficed to compute only one trajectory of this model for 6,000 generations of evolution.

A question that naturally arises at this point in the description of the experiment under consideration is why was the number 100 chosen as the number of replications? Whenever an investigator is in the process of designing a Monte Carlo simulation experiments, a question as to what sample size is needed to produce an informative statistical summarization of the simulated data will always arise. Answers to this question will depend on theoretical as well as practical considerations. When one is contemplating a Monte Carlo simulation experiment with a large number of replications, it will be necessary to make a large number of calls to the random number generator used in the experiment. All random number generators will have a large but finite period and when the number of calls to the generator exceeds the period, questions will arise as to whether the random numbers used in the experiment had a sufficiently long period to assure credibility of the simulated data as to its randomness. This issue was discussed in some detail in Mode and Gallop [10] in Section 2 of that paper. From an example given in that paper, if an investigator wished to do a large Monte Carlo simulation experiment based on the Wright-Fisher process, the default random number generator in many software packages did not have a sufficiently long period to assure the randomness of the simulated data. Consequently, an alternative random number generator with a very long period was used in the paper published in 2008, and if a reader is interested in the details Section 2 of the paper by Mode and Gallop [10] may be consulted as well as the references cited therein. This random number generator with a very long period was also used in all computer simulation experiments reported in this paper to help assure the "randomness" of the simulated data.

With regard to practical considerations, the choice of sample size for a Monte Carlo simulation experiment will depend on the computer platform available to an experimenter. All the computer simulation experiments reported in this paper, as well as those reported in the book by Mode and Sleeman [12], were done on personal desktop or laptop computers. Before a Monte Carlo simulation experiment was attempted, a number of exploratory experiments were conducted using the embedded deterministic model with various combinations of parameter values. When an interesting result, which was judged subjectively, was found, a decision was made to run a preliminary Monte Carlo simulation experiment with sample size as small as 10. Such a small sample size is often sufficient to give an experimenter some idea as to how well the embedded deterministic model will predict the behavior of the process. If a judgement was made that it would be worthwhile to conduct a confirmatory Monte Carlo simulation experiment with a larger number of replications, then such an experiment was executed.

On personal desktop or laptop computers, preliminary experiments based on the embedded deterministic model may be executed within minutes or about an hour. But to finish a Monte Carlo simulation with 100 replications of 6,000 generations of evolution may require a timespan of a few hours or, in some cases, over 24 or more hours were needed to finish a Monte Carlo simulation experiment. Indeed, it is interesting to note that the times needed to complete a Monte Carlo simulation experiment, increase beyond 24 hours with

increasing complexity of the model under consideration. If an investigator was confined to using only one or two personal computers and the accomplishment of multitude tacks and there were required to be accomplished each day, then having a computer tied up for a day or more can be very disruptive for a research schedule. Based on this experience over a timespan of a decade or more, the idea that 100 replications were sufficient to provide an informative statistical summarization of the simulated data was made by trial and error. Of course, if an investigator has access to a network of computers, which may run for days without disruption of a research schedule, it would be practical to entertain larger numbers of replications from which judgements could be made as to whether such sample sizes would be sufficient to have confidence that statistical summarizations of the simulated data were sufficiently informative. In this and subsequent sections of the paper, for each experiment considered, remarks regarding the sufficiency of the number of replications will be made.

There is another case that is worthy of mention when a decision is being made as to whether an increase in the number of replications after a preliminary confirmatory experiments has been completed would be justified. If there is evidence that the process had converged to a quasistationary distribution based on the observation that the statistical quantiles appear to relatively flat after some number of generations of evolution, then increasing the number of replications in a follow-up experiment may not be sufficient to justify the computer time needed to complete a simulation experiment with a larger number of replications.

In the experiment reported in this section, as well as those in subsequent sections, only one set of numerical values of the parameters were under consideration. For example, $\boldsymbol{\lambda}$, the components of the 1×3 vector denoting the expected number of offspring contributed to the next generation by each of the three genotypes were assigned the values

$$\boldsymbol{\lambda} = (1.02, 1.02, 1.05). \tag{4.1}$$

If the three genotypes are denoted by the symbols A_1, A_2, and A_3, then the mutation $A_1 \rightarrow A_2$ would not be beneficial but the mutation $A_1 \rightarrow A_3$ would, by definition, be beneficial. In the experiment under consideration, the probabilities of mutations among the three genotypes were assigned the values indicated in the matrix

$$\mathbb{M} = \begin{pmatrix} \mu_{11} & 10^{-7} & 0 \\ 10^{-12} & \mu_{22} & 10^{-14} \\ 10^{-15} & 10^{-17} & \mu_{33} \end{pmatrix}, \tag{4.2}$$

where the principal diagonal elements are chosen such that the sum of each row of the matrix is 1. By way of interpretation of this matrix, in the first row the assigned probability for the mutation $A_1 \rightarrow A_2$ was $\mu_{12} = 10^{-7}$ per generation, but it was assumed that the mutation $A_1 \rightarrow A_3$ would not occur in this experiment so that, by assumption, $\mu_{13} = 0$. As can be seen from the second row of the matrix, it was assumed that the probability of the favorable mutation $A_2 \rightarrow A_3$ was $\mu_{23} = 10^{-14}$ so in this computer experiment this mutation was rare, which raised the question as to whether it would actually appear in a simulated population. In conversations with biologists, who study mutations in bacteria, the probabilities of a mutational event were somewhere in the interval $(10^{-10}, 10^{-6})$ per generation. Thus, based on this anecdotal information, the value 10^{-7} was within range of observed values but experiments

suggest that the value 10^{-14} would be a probability for a rare event. It should also be mentioned in passing, that if the expected number of offspring produced by mutant genotype 2 were assigned the value $\lambda_2 = 1$, then from the point of view of reproductive success, this mutation would be deleterious in the sense that individual of this genotype would be at a selective disadvantage, because $\lambda_2 = 1$, but the lambda parameters for the other two genotype were greater than 1. Thus, in such an experiment, the evolution of deleterious and beneficial mutation could be considered simultaneously in a computer simulation experiment.

The next step in the assignment of parameter values consisted of assigning values for the regulation of total population size. As was mentioned in a previous section, the vector of alpha parameters in the Weibull survival function were assigned the values $\boldsymbol{\alpha} = (2, 2, 2)$, and, moreover, these values will be used in all experiments reported in subsequent sections. Recall that the rationale underlying these parameter assignment was discussed in a previous section. The second set of parameters for the regulation of population size was the vector of beta parameters for each genotype. For the experiment under consideration, the parameters in this vector were assigned the values $\boldsymbol{\beta} = (10^{-6}, 10^{-6}, 10^{-6})$, indicating that the maximum total number of individuals for each of the three genotypes was about one million individuals. Because all the alpha and beta parameters for the three genotype were assumed to be equal in this experiment, the regulation of population size was not a component of natural selection. The last numerical input in the experiment was that of assigning values to the three component vector $\mathbf{X}(0)$, the number of individuals of each genotype in the initial population. In the experiment reported in this section this vector had the components

$$\mathbf{X}(0) = (10,000, 0, 0). \tag{4.3}$$

Thus, it was assumed that the initial population consisted only of $10,000$ individuals of genotype A_1 and the appearance of genotypes A_2 and A_3 as the population evolved would be due to mutations. Subsequently, in this and following sections this experiment will be referred to as experiment 1.

Because the model under consideration is stochastic, the structure to be predicted in a computer experiment is the dynamic distribution of the numbers of the three genotypes in any generation of the evolutionary process. From the analytic point of view, the problem of deriving a useful form for this dynamic distribution would be very difficult. However, given a sample of Monte Carlo realizations of the process, informative properties of this distribution can be estimated and inferred. As discussed in Mode and Gallop [10], it is feasible to estimate such summary statistics of this dynamic distribution, including the extreme value trajectories denoted by MIN and MAX as well as quantile trajectories denoted by $Q25, Q50$, and $Q75$. Two other trajectories of interest would be mean and standard deviation denoted by MEAN and SD. Unlike the stochastic process, whose predictive structure is a dynamic distribution, the predictive structure of the embedded deterministic model consists of estimates of the numbers of each of the three genotypes as the population evolves on a time scale of generations. There is also another property differentiating the stochastic process from the embedded deterministic model. The sample functions of the stochastic model always take values in the set $\mathbb{N}$ of nonnegative integers whereas the numerical values of the trajectories of the embedded deterministic model take values in the set $\mathbb{R}^+ = [0, \infty)$ of nonnegative real numbers.

It was observed in a preliminary experiment that in the embedded deterministic model, estimates of the number of individuals of genotype A_3 converged to a constant within

Table 1: Comparisons of numerical values of stochastic and deterministic trajectories for experiment 1.

(a) Stochastic trajectories for genotype 1

GEN	MIN	$Q25$	$Q50$	$Q75$	MAX	MEAN	SD
2200	129749	137196	138686	140035	142339	138556.82	2116.18
2201	130070	137337	138512	140068	142463	138581.65	2099.05
2202	129457	137403	138583	140117	143120	138632.99	2123.60
2203	129224	137541	138738	140075	142707	138632.62	2135.92
2204	129302	137223	138918	140126	143430	138667.74	2208.46

(b) Stochastic trajectories for genotype 2

GEN	MIN	$Q25$	$Q50$	$Q75$	MAX	MEAN	SD
2200	65	1154	1932	2742	11321	2181.62	1623.22
2201	62	1163	1912	2757	11209	2180.71	1623.55
2202	68	1174	1893	2743	11126	2178.24	1624.77
2203	69	1163	1901	2769	11133	2173.69	1623.14
2204	66	1155	1909	2726	11082	2178.11	1627.84

(c) Stochastic trajectories for genotype 3

GEN	MIN	$Q25$	$Q50$	$Q75$	MAX	MEAN	SD
2200	0	0	0	0	0	0	0
2201	0	0	0	0	0	0	0
2202	0	0	0	0	0	0	0
2203	0	0	0	0	0	0	0
2204	0	0	0	0	0	0	0

(d) Deterministic trajectories of genotypes 1, 2, and 3

GEN	genotype 1	genotype 2	genotype 3
2200	3.96×10^{-5}	8.78×10^{-9}	220884.96
2201	3.85×10^{-5}	8.54×10^{-9}	220884.96
2202	3.74×10^{-5}	8.30×10^{-9}	220884.96
2203	3.63×10^{-5}	8.07×10^{-9}	220884.96
2204	3.53×10^{-5}	7.85×10^{-9}	220884.96

2,200 generations. Therefore, a decision was made to examine the estimated trajectories of the stochastic process for generation 2,200 and four generations thereafter. Presented in Table 1 are estimated trajectories for each of the three genotypes for the stochastic process as well as those for the deterministic model for five generations 2,200 through 2,204.

Observe that in the first columns of this table, the symbol GEN denotes generations. By viewing the numerical values in this table, it is straight forward to compare the predicted values of the numbers of each of genotype based on the embedded deterministic model with the estimated trajectories of the stochastic process. When rounded up to the nearest integer, the embedded deterministic predicts that the number of individuals of genotypes 1 and 2 would be 0 at the generations displayed. However, as can be seen from the estimated trajectories for the numbers of genotypes 1 and 2 for the stochastic model in the upper two subtables of Table 1, all these numbers differ significantly from 0. But, there is a greater

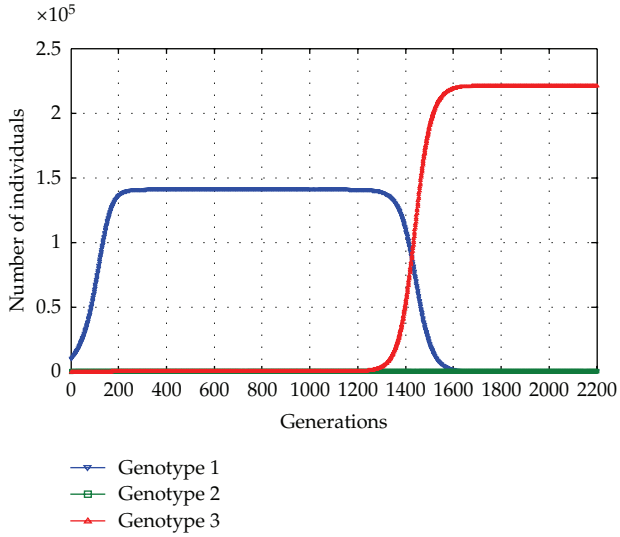

Figure 1: The evolutionary trajectories of the three genotypes as predicted by the embedded deterministic model.

difference in the predictions of the embedded deterministic model and the stochastic process for genotype 3. For according to the predictions of the embedded deterministic model in the subtable at the bottom of Table 1, the number of individuals in the population of genotype 3 would be about 220,884.96. But, in the subtable immediately above that for the predictions of the deterministic model, the numbers for the stochastic model are uniformly 0, indicating that in the sample of Monte Carlo simulations of the process, the mutant genotype A_3 did not appear in 100 replications of 6,000 generations of evolution.

To get a more informative overview of the predictions of the embedded deterministic model, the trajectories of the three genotypes were graphed for the first 2,200 generations of the experiment. Presented in Figure 1 are the graphs of the trajectories of the three genotypes for the embedded deterministic model.

From an inspection of the graphs of the trajectories for the three genotypes, it can be seen that the estimated number of individuals of genotype 1 began a steep decline somewhere between 1,300 and 1,400 generations into the simulated evolutionary time period. Accompanied by this steep decline in the estimated number of individuals of genotype 1 was a steep increase in the number of individuals of genotype 3. By generation 1,800 the number of individuals of genotype 3 had reached a level value at about a midpoint between the numbers 2×10^5 and 2.5×10^5 whereas the number of individuals of genotype 1 has declined to a value near zero. Throughout the 2,200 generations displayed in Figure 1, the number of individuals of mutant genotype 2 remained at relatively low numbers. Because it was assumed that individuals of genotype 3 had a reproductive advantage, the forms of the graphs of the trajectories in Figure 1 were expected, but at the outset of the experiment it was not clear as to the number of generations required before the number of individuals of genotype 3 rose to predominance in the population.

It is of methodological interest to ask the question: in what sense are the trajectories computed from the embedded deterministic model measures of central tendencies of realized sample functions of the process. Because the median trajectories, $Q50$, for the three genotypes were estimated from the simulated Monte Carlo data for each of the three genotypes and is

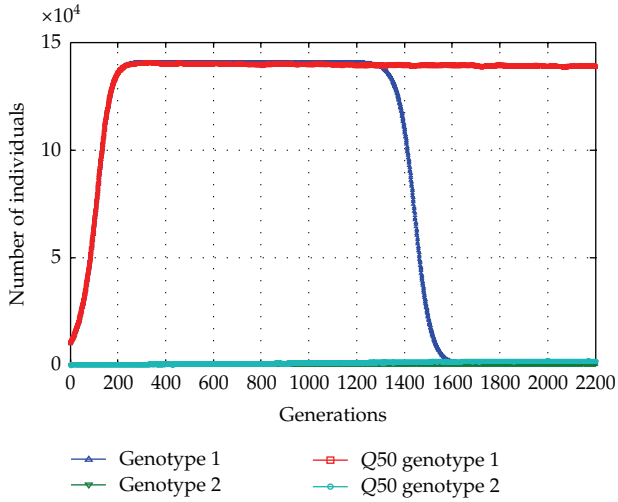

Figure 2: A comparison of deterministic and stochastic $Q50$ trajectories for genotypes 1 and 2.

a measure of central tendency for the sample functions of the process, it is of interest to plot the deterministic and $Q50$ trajectories for the three genotypes. But, in the experiment under consideration, no individuals of mutant genotype 3 appeared in 100 replications of 6,000 generations of evolution. Consequently, for the experiment under consideration, it is possible to focus attention only on the trajectories of genotypes 1 and 2. Presented in Figure 2 are graphs of the deterministic and $Q50$ trajectories for the numbers of individuals of genotypes 1 and 2 for the first 2,200 generations of simulated evolution.

As can be seen from Figure 2, the deterministic trajectory for genotype, denoted by genotype 1, and the $Q50$ trajectory, which was estimated from a sample of Monte Carlo realizations of the process, are quite close in relative terms for nearly 1,400 generations of the 2,200 generations of evolution presented in the figure. But, after 1,400 generations, genotype 1, the deterministic trajectory for genotype 1 declines steeply until somewhere between 1,600 and 1,800 generations it reaches a value near zero, but the $Q50$ trajectory for genotype 1 remains nearly constant at a value less than 15×10^4 throughout the remaining generations shown in the figure. But, the deterministic trajectory, genotype 2, and the stochastic trajectory $Q50$ for genotype 2 remain at low values throughout a period of 2,200 generations of simulated evolution. According to the predictions of the embedded deterministic model as displayed in Figure 1, genotype 3 rises to predominance in the population after about 1,500 generations of evolution. But, in the Monte Carlo simulation experiment under consideration, individuals of mutant genotype 3 did not appear in the simulated population for 100 replications of 6,000 generations of evolution. Hence, because individuals of genotype 3 did not rise to predominance in the Monte Carlo simulation experiment, the expected decline in the $Q50$ trajectory for genotype 2 did not occur in the simulated realizations of the stochastic process but remained relatively constant during generations in the timespan of 1,600 to 2,200 as shown in the figure.

It has been observed in a number of computer experiments not reported here that even if the probability that a beneficial mutation occurs is small, when the trajectories for the embedded deterministic model are computed, the rare mutation will always occur and become predominant in the population in the long run. However, the event that the rare mutation occurs and becomes predominate in a Monte Carlo simulation experiment may not

be observed in a simulated sample of realizations of the stochastic process. As mentioned previously, in a deterministic projection, the range of the set of values for the number of each of the three genotypes is the set $\mathbb{R}^+ = [0, \infty)$ of nonnegative real numbers. A subset of this set is the interval $[0, 1) = \{x \in \mathbb{R} \mid 0 \leq x < 1\}$ and at the beginning of every projection based on the deterministic model small positive numbers in the interval $[0, 1)$ will be observed, but eventually these numbers increase until they exceed the number 1. Apart from the number 0, all numbers in this interval represent fictional individuals, because the actual count of the number of individuals in a real population must be some whole number in the set $\mathbb{N}$ of nonnegative integers. Because there is ample evidence that many events that have occurred during the evolution of any species were, in some sense, due to chance or stochastic effects, a stochastic model seems to be a better approximation to the evolution of a real population, but, nevertheless, as will be shown in subsequent section, the computed trajectories based on an embedded deterministic model are quite often reasonable measures of central tendency for the sample functions of a stochastic process.

In retrospect, it may of been have interest to test whether the beneficial mutation A_3 would eventually appear in a simulated population if larger number of generations and replications of the simulated process were considered. Given the results of the experiment with 6,000 generations of evolution replicated 100 times, it seems unlikely that an experiment with more generations of evolution and a greater number of replications would justify the computer time needed to complete such an experiment. However, this was a value judgement and other investigators may entertain another judgement and do an experiment with a larger number of generations and replications of the stochastic process.

It should be mentioned that the results of this experiment were not arrived on a de novo basis. Rather they were motivated by an experiment with a more complex two-sex model in which genotypes resulting from beneficial mutation did not appear in a simulated population in a simulation experiment consisting of 6,000 generations of evolution replicated 100 times. Unlike the experiment reported in this section, the result with the two-sex model was found serendipitously and was not preplanned. An interested reader may consult Mode et al. [18] for details.

5. Smaller Probabilities of Beneficial Mutations and Differential Reproductive Success

This section will be devoted to the question: if the parameter assignments used in the experiments reported in Section 4 were held constant except for the mutation probability μ_{23}, then what range of values of this probability would be needed to insure that beneficial mutant genotype A_3 would eventually appear and become predominant in an evolving population? The goal of this section is to report on the results of two computer experiments in which the probability of a beneficial mutation, $A_2 \rightarrow A_3$, was assigned larger probabilities than those in Section 4. In the first of these experiments, the probability of the beneficial mutation was assigned the value $\mu_{23} = 10^{-10}$ in contrast to the value $\mu_{23} = 10^{-14}$, that was used in experiment 1 reported in Section 4. Before this values was chosen, several preliminary experiments were conducted. For this set of assigned values for the parameters, individuals of mutant genotype A_3 did eventually become predominant in the population as one might expect, because, on average, individuals of genotype 3 contributed more offspring to the next generation than those of genotypes A_1 and A_2. In this computer experiment, however, there was greater variation among the numbers of individuals of genotypes 1 and 2 than in the experiment reported in Section 4. One approach to having an idea as to the amount of this

variation is to compute the fraction or the frequency of the realizations among the 100 Monte Carlo replications of 6,000 generations of evolution considered in the experiment that had no individuals of each of the three genotypes in generation 6,000. For the experiment under consideration, these frequencies were 0.23, 0.21, and 0, respectively, for genotypes 1, 2, and 3. By way of contrast for the experiment reported in Section 4, these three frequencies were 0, 0, and 1, indicating that in all the 100 Monte Carlo replications of 6,000 generations of evolution, the numbers of genotypes 1 and 2 were always positive and those for genotype 3 were always 0 in generation 6,000.

Rather than reporting the results of this experiment in more detail, in this section attention will be focused on a similar experiment in which the probability of the mutation $A_1 \rightarrow A_2$ was assigned a lower value than the value $\mu_{12} = 10^{-7}$ used in the experiment reported in Section 4 as well as the one briefly described in this section. From now on in this section the values assigned for the mutations $A_1 \rightarrow A_2$ and $A_2 \rightarrow A_3$ were $\mu_{12} = \mu_{23} = 10^{-10}$. With this assignment, it was expected that the waiting time until an appearance of the mutant genotype A_2 and the subsequent appearance of the beneficial mutant A_3 would, on average, be longer. All other parameters of the model were assigned the same values as those used in the experiment reported in Section 4. In the discussion that follows, this experiment will be referred to as experiment 2. Shown in Table 2 are statistical summaries of simulated data for five generations of evolution for genotypes 1, 2 and 3 as well as the deterministic trajectories after the embedded deterministic model had reached a point at which the trajectory for genotype 3 was constant in experiment 2.

As can be seen from the lower subtable of Table 2, by generation 2,105 the deterministic trajectory for genotype 3 had converged to a constant of about 220,884 individuals and that the numbers of individuals of genotypes 1 and 2 were essentially zero. On the other hand, the estimated stochastic trajectories for genotypes 1, 2, and 3 present a different story. For example, in the subtable for genotype 1, in which the estimated predictive trajectories of the stochastic process for this genotype are displayed, the estimated trajectories are signatures of high levels of variation, stochasticity, among the 100 Monte Carlo realizations of the process for generations 2,105 through 2,109. For example, the value of MIN trajectory is 0 for all these generations, which indicated that among some of the 100 Monte Carlo realizations of the process, no individuals of genotype 1 were present in the simulated population during these generations. Interestingly, apart from two exceptions, the values of the $Q25$, $Q50$, and $Q75$ trajectories were 1 during these generations. It is also interesting to note that the median $Q50 = 1$ trajectory is near that predicted by the embedded deterministic model for these generations. The high level of stochasticity displayed among the 100 Monte Carlo realizations of the process is displayed very clearly by the MAX, MEAN, and SD trajectories in the right-most three columns of the subtable for genotype 1. For in these columns, the MAX exceeds 141,000 individuals, the MEAN trajectory exceeds 6,800, and the standard deviation trajectory, SD, exceeds the value 29,000 in all generations. Observe that values of SD are 4- to 5-times grater than the values of the mean in these generations, which is also a signature of high levels of stochasticity.

The estimated stochastic trajectories for the stochastic process in the generations under consideration in the subtables for genotypes 2 and 3, however, display signatures of more moderate levels of stochasticity among the 100 Monte Carlo realizations of the process. In particular, it is interesting to note that values of the $Q50$ trajectory in the subtable for genotype 2, which exceed the value 220,000 in all five generations, are much greater than that predicted by the embedded deterministic model, but rather closely match those predicted by the embedded deterministic model for genotype 3 which were also greater than 220,000 in

Table 2: Comparisons of numerical values of stochastic and deterministic trajectories for experiment 2.

(a) Stochastic trajectories for genotype 1

GEN	MIN	$Q25$	$Q50$	$Q75$	MAX	MEAN	SD
2105	0	1	1	1	141357	6915.43	30084.44
2106	0	0	1	1	141866	6892.56	30002.01
2107	0	1	1	1	141905	6879.54	29960.46
2108	0	0	1	1	142067	6874.84	29949.86
2109	0	1	1	1	142875	6860.15	29895.47

(b) Stochastic trajectories for genotype 2

GEN	MIN	$Q25$	$Q50$	$Q75$	MAX	MEAN	SD
2105	0	219892	220873	221502	223817	210124.57	46984.89
2106	0	219777	220764	221457	223816	210044.14	46937.52
2107	0	220039	220542	221381	223952	210087.86	46921.99
2108	0	219893	220751	221191	224191	210111.26	46892.89
2109	0	219958	220799	221469	224603	210195.93	46865.36

(c) Stochastic trajectories for genotype 3

GEN	MIN	$Q25$	$Q50$	$Q75$	MAX	MEAN	SD
2105	0	219892	220873	221502	223817	210124.57	46984.89
2106	0	219777	220764	221457	223816	210044.14	46937.52
2107	0	220039	220542	221381	223952	210087.86	46921.99
2108	0	219893	220751	221191	224191	210111.26	46892.89
2109	0	219958	220799	221469	224603	210195.93	46865.36

(d) Deterministic trajectories for the Three genotypes

GEN	genotype 1	genotype 2	genotype 3
2105	6.04×10^{-5}	9.00×10^{-11}	220884.96
2106	5.87×10^{-5}	8.97×10^{-11}	220884.96
2107	5.70×10^{-5}	8.93×10^{-11}	220884.96
2108	5.54×10^{-5}	8.90×10^{-11}	220884.96
2109	5.38×10^{-5}	8.87×10^{-11}	220884.96

all generations. From the subtable for genotype 3 it can be seen that the $Q50$ trajectory is also close to that predicted by the embedded deterministic model, but the MIN trajectory for genotype 3 has the constant value 0 in all generations and this together with the SD, which exceeds 46,000 in all generations, are indicative of significant levels of variation among the 100 Monte Carlo realizations of the process.

The reported values in Table 2 provide detailed insights for the evolution of the stochastic process for the five generations considered, but to obtain broader overview of the simulated data three graphs will be presented. Presented in Table 2 are the estimated $Q50$ and deterministic, DET, trajectories for genotype 1 for the first 2200 generations of 6,000 simulated generations of evolution.

From this figure a clearer picture of the evolution of the numbers of individuals of genotype 1 during the first 2,200 generations of the experiment emerges as predicted by

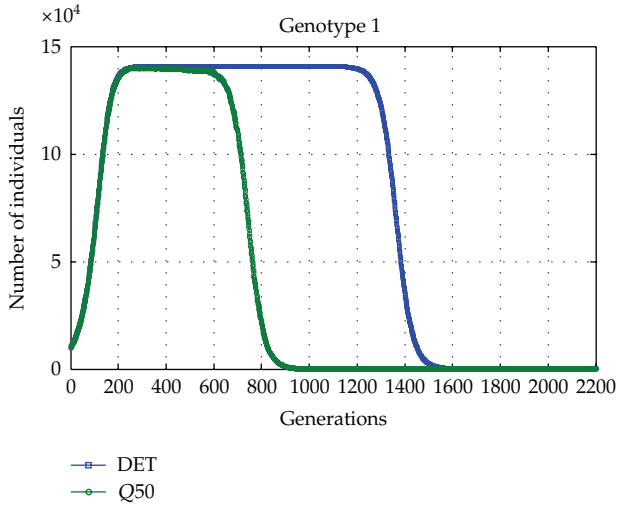

Figure 3: Estimated $Q50$ and DET trajectories for genotype 1 in experiment 2.

the stochastic process and the embedded deterministic model. Thus, it can be seen from this figure that the $Q50$ trajectory of the process and that for the embedded deterministic model essentially agree for nearly 800 generations of the simulation experiment. But, at about 800 generations into the experiment, the $Q50$ trajectory of the stochastic process declines to values near 0 by generation 1,000. On the other hand, the DET trajectory remains essentially constant until about 1,300 generations into the experiment and then declines steeply to values near 0 by generation 1,600. The observation that the generation times needed to reach near zero values for the $Q50$ and DET trajectories differ by about 600 generations confirms that in experiment 2, the predictions for genotype 1 according to the embedded deterministic model were not good measures of central tendencies of the process throughout the first 2,200 generations of the experiment. Note, however, that the two trajectories are both essentially 0 after 1,600 generations of simulated evolution.

For the sake of brevity, a graph showing the $Q50$ and DET trajectories for genotype 2 will not be displayed in this section, but suffice to mention that these trajectories did not coincide throughout the 2,200 generations considered in Figure 3. In the following two figures attention will be focused on genotype 3, which was expected to become predominant in the population as it evolved after a sufficiently long period of time. Presented in Figure 4 are the estimated $Q50$ and DET trajectories for individuals of genotype 3 for the first 2,200 generations of the experiment.

When this figure is viewed, it becomes clear that also for genotype 3, the trajectory computed using the embedded deterministic model was not a good measure of central tendency for the stochastic process throughout the first 2,200 simulated generations of evolution. For in this case, the $Q50$ and DET trajectories essentially agreed for the first 600 generations of the experiment but thereafter differed significantly until about generation 1,600. Thus, it can be seen that by about 700 generations into the experiment, the $Q50$ trajectory had reached an essentially constant value between 2×10^5 and 2.5×10^5 individuals but the DET trajectory did not rise to this nearly constant value until about 1,400 generations into the experiment. Thus, in this experiment the number of generations in which there was not good agreement between the stochastic and deterministic trajectories was about

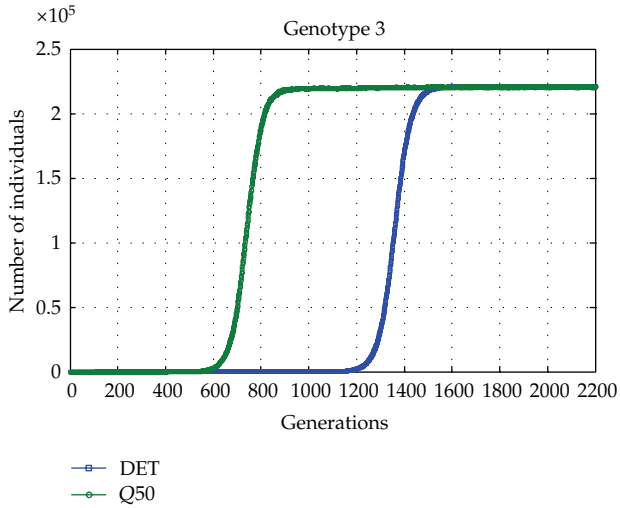

Figure 4: Estimated $Q50$ and DET Trajectories for Genotype 3 in Experiment 2.

700 generations. But after 1,600 generations of evolution, these two trajectories essentially coincided.

Figures 3 and 4 represent transitory evolutionary periods in which the numbers of individuals of genotype 1 were decreasing and those of genotype 3 were increasing, due to its selective advantage. It is during this transitory period that the trajectory of the embedded deterministic model is not a good measure of central tendency for the sample functions of the stochastic process, but, outside this transitory period, the $Q50$ trajectory of the process and that of the embedded deterministic model, DET, are quite close. A question that naturally arises is how close are these two trajectories for each genotype near the end of 6,000 generations of simulated evolution. Presented in Figure 5 are graphs of the trajectories, $Q50$ and DET, for individuals of genotype 3 in the last 200 generations of the experiment.

As can be seen from this figure, the trajectory of embedded deterministic model is flat, as one would expect after viewing the graph of this trajectory in Figure 4 for genotype 3. In Figure 5, however, it can be seen that $Q50$ trajectory of the process varies around the DET trajectory, but in terms of actual numbers of individuals in the population of genotype 3 this variation is relatively small. It is known that for some classes of branching processes, the process converges in distribution to a quasistationary distribution. There is a rather extensive literature on quasistationary distributions that may be found by consulting the Internet, but the technical ideas underlying such distributions here will be confined to some brief remarks.

Consider a vector-valued stochastic Markov process, $\mathbf{X}(t)$ for $t = 0, 1, 2, 3, \ldots$, taking values in the set $\mathbb{N}^{(3)}$ of three-dimensional nonnegative integers. Let $\mathbb{S}_1$ denote the set of absorbing states and let $\mathbb{S}_2$ denote the set of transient states. To illustrate these ideas, for the branching process under consideration, $\mathbb{N}^{(3)}$ is the state space and $\mathbf{0} = (0, 0, 0)$ is called an absorbing state, because in the absence of immigration, if the population enters this state, it is extinct and evolution ends for this species. Thus, in this example, the set of absorbing states is $\mathbb{S}_1 = \{\mathbf{0}\}$, the set consisting of only the zero vector $\mathbf{0}$. And the set $\mathbb{S}_1 = \{\mathbf{x} \in \mathbb{N}^{(3)} \mid \mathbf{x} \neq \mathbf{0}\}$ is the set of transient states in which the population evolves. Next, let $E(t)$ denote the event that process is in some transient state in generation $t \in \mathbb{N}$. Then, the conditional probability

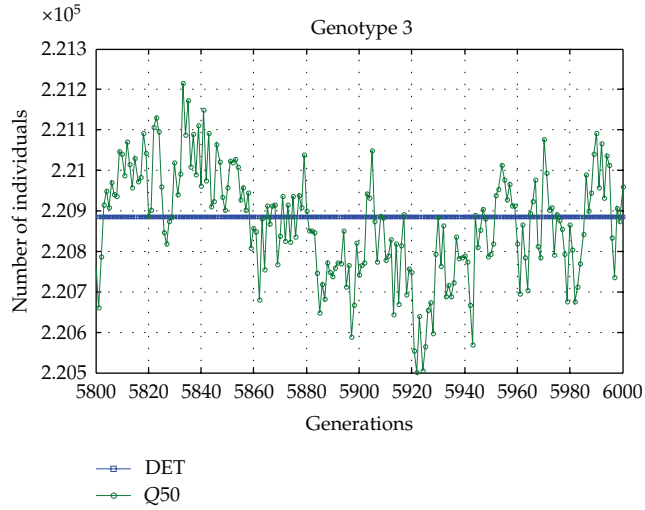

Figure 5: Graphs of the trajectories $Q50$ and DET of genotype 3 for generations 5,800 to 6,000 in experiment 2.

that the process is in some transient state $\mathbf{x} \in \mathbb{S}_2$ in generation t, given that it started in state $\mathbf{x}_0 \in \mathbb{S}_2$ at time $t = 0$, is

$$P[E(t) \mid \mathbf{X}(0) = \mathbf{x}_0] = \sum_{\mathbf{x}} P[\mathbf{X}(t) = \mathbf{x} \mid \mathbf{X}(0) = \mathbf{x}_0], \tag{5.1}$$

where the sum ranges over all states $\mathbf{x} \in \mathbb{S}_2$. Therefore, the conditional probability that the population is in some transient state $\mathbf{y} \in \mathbb{S}_2$ in generation t, given the event $E(t)$, is

$$P[\mathbf{X}(t) = \mathbf{y} \in \mathbb{S}_2 \mid E(t)] = \frac{P[\mathbf{X}(t) = \mathbf{y} \mid \mathbf{X}(0) = \mathbf{x}_0]}{P[E(t) \mid \mathbf{X}(0) = \mathbf{x}_0]}, \tag{5.2}$$

for every $\mathbf{y} \in \mathbb{S}_2$. For some cases, it has been shown that

$$\lim_{t \uparrow \infty} P[\mathbf{X}(t) = \mathbf{y} \in \mathbb{S}_2 \mid E(t)] = \pi(\mathbf{y}), \tag{5.3}$$

for $\mathbf{y} \in \mathbb{S}_2$ and the limit does not depend on the initial condition $\mathbf{X}(0) = \mathbf{x}_0 \in \mathbb{S}_2$. When it can be proven that this limit exists, it is called the quasistationary distribution on the set $\mathbb{S}_2$ of transient states. No attempt will be made in this paper to prove that this limit exists, but the experimental evidence, such as that displayed in Figure 5 and in subsequent figures, suggests that the process under consideration did converge to a quasistationary distribution as $t \uparrow \infty$, because in none of these experiments extinction of the population within 6,000 generations of evolution has not been observed. It will be noted from Figure 5 that even though the embedded deterministic model has converged to a constant for genotype 3, the $Q50$ trajectory for this genotype fluctuates around this constant, which is indicative of the concept of convergence in distribution. For if a process converges in distribution, the sample functions of the process continue to fluctuate around some measure of central tendency, and,

in many cases, the stationary distribution will have a constant and finite expectation and variance but the variation among the realizations of the process continues indefinitely as the population evolves in time. The graphs for genotypes 1 and 2 were not displayed for the last 200 generations of experiment 2, because they were noninformative. In particular, the $Q50$ trajectory was a constant 1 and the DET was essentially 0 throughout the last 200 generations for both genotypes 1 and 2, indicating that both these genotypes were present in the simulated population only in small numbers.

6. Differential Capacities to Compete as Driving Forces of Evolution

In this section, the results of two computer simulation experiments will also be reported. In both of these experiments, the vector denoting the expected number of offspring produced per generation for each of the genotype was assigned as

$$\boldsymbol{\lambda} = (2, 2, 2), \tag{6.1}$$

so that a population would grow at a faster rate than those considered in Sections 4 and 5. The assignment of expected number of 2 offsprings of each genotype was considered, because it was assumed that each individual in the population was a cell that reproduced by binary cell division. According to these assignments, reproductive success as a component of natural selection in these experiments was, by definition, neutral. It was also supposed that the probability of the mutation $A_2 \rightarrow A_3$ was $\mu_{23} = 10^{-14}$, which was the same values as that for experiment 1 reported in Section 4. To expedite the reading of this section, it will be helpful to present all the probabilities of mutation among the three genotypes used in these experiments as indicated in the 3×3 matrix

$$\mathbb{M} = \begin{pmatrix} \mu_{11} & 10^{-10} & 0 \\ 10^{-14} & \mu_{22} & 10^{-14} \\ 10^{-15} & 10^{-17} & \mu_{33} \end{pmatrix}, \tag{6.2}$$

where elements on the principal diagonal are chosen such that all rows of the matrix add up to one. Differential capacities to compete for food and other resources among the three genotypes were quantified in terms of the vector

$$\boldsymbol{\beta} = \left(10^{-14}, 10^{-14}, 10^{-16}\right). \tag{6.3}$$

According to these assignments, the total size of a population attained in a computer simulation experiment would be much larger than those in experiments 1 and 2 reported in Sections 4 and 5, and genotype 3 would have a selective advantage over the other two genotypes, because the parameter assignment $\beta_3 = 10^{-16}$ allowed individuals of genotype 3 to compete more successfully for resources in large populations than those of genotypes 1 and 2. This assignment of parameter values was motivated by the desire to let the population grow to a sufficient size so as to increase the likelihood that a mutational event of low probability

would actually occur. The initial vector, denoting the numbers of each of the three genotypes in the population at generation 0 was chosen as

$$\mathbf{X}(0) = (10,000, 0, 0), \tag{6.4}$$

indicating, by assumption, that there are 10,000 individuals of genotype 1 in the initial population, but the number of individuals of each of the genotypes 2 and 3 was 0. All other parameters of the model were assigned the same values as used in experiments 1 and 2 reported in Sections 4 and 5.

Given the small values for the β displayed in (6.3), large population sizes that were permitted to occur in a computer experiment, so it was expected that the rare beneficial $A_2 \rightarrow A_3$, which occurred with probability $\mu_{23} = 10^{-14}$ per generation, would be more likely to arise in this experiment and eventually become predominant in the population. This expectation was realized as indicated the frequencies of the events that number of individuals of each of the three genotypes in generation 6,000 of the experiment was 0. The observed frequencies of these event were 0, 0.1 and 0. Thus, for genotypes 1 and 3, among all the 100 Monte Carlo replications of 6,000 generations of simulated evolution, the number of individuals of each of the genotype 1 and 3 was at least 1. But, for the case of genotype 2, in generation 6,000 of this experiment, in 10 of the Monte Carlo replications, there were no individuals of genotype 2 but in the other 90 replications there was at least one individual of genotype 2. Given the very small probabilities of the back mutations, $A_3 \rightarrow A_2$ and $A_3 \rightarrow A_1$, see row 3 of the matrix in (6.2), it seems plausible that in the 10 Monte Carlo replications in which the number of individuals of genotype 2 was 0 in generation 6,000, this rare back mutation did not occur.

Instead of discussing the results of this experiment more thoroughly in the remainder of this section, attention will be focused on an experiment in which all the parameters were chosen as above but the initial vector for the population was chosen as

$$\mathbf{X}(0) = (1, 0, 0), \tag{6.5}$$

indicating that in the initial population there was only one individual of genotype 1. To motivate this choice of initial condition, think of one individual, a spore for example, being dispersed to a new environment and the evolution of this very small founder population was subsequently followed for a long period of time. In such a situation, it seems natural to think about whether the offspring of the single initial individual would survive and evolve into a large population or would the population become extinct. From now on, this computer run will be referred to as experiment 4.

As expected in this experiment, the evolution of the population occurred at a more rapid pace than in the experiments reported in the preceding sections. Presented in Table 3 are five generations of statistically summarized simulated data for each of the three genotypes after the embedded deterministic model had converged.

All values in the subtable for the stochastic trajectories of genotype 3 should be multiplied by 10^{15}, indicating that in experiment 4 the number of individuals of genotype 3 had become very large after a short period of 150 generations of evolution. A feature common to all statistically summarized Monte Carlo simulation data in Table 3 is that for genotypes 1, 2, and 3, the trajectories MIN and $Q25$ are all zero. After this observation, these trajectories were inspected for 6,000 generations of evolution, and it was observed that in all these generations the MIN and $Q25$ trajectories were 0. This observation indicates that in at

Table 3: Comparisons of numerical values of stochastic and deterministic trajectories for experiment 4.

(a) Stochastic trajectories for genotype 1

GEN	MIN	$Q25$	$Q50$	$Q75$	MAX	MEAN	SD
150	0	0	5	8	16	5.18	3.99
151	0	0	6	8	18	5.45	4.14
152	0	0	5	8	17	5.17	3.91
153	0	0	5	9	14	5.48	4.07
154	0	0	6	8	15	5.38	3.94

(b) Stochastic trajectories for genotype 2

GEN	MIN	$Q25$	$Q50$	$Q75$	MAX	MEAN	SD
150	0	0	1	1	1	0.70	0.46
151	0	0	1	1	1	0.68	0.47
152	0	0	1	1	1	0.65	0.48
153	0	0	1	1	1	0.67	0.47
154	0	0	1	1	1	0.65	0.48

(c) Stochastic trajectories for genotype 3

GEN	MIN	$Q25$	$Q50$	$Q75$	MAX	MEAN	SD
150	0	0	8.33	8.33	8.33	6.16	3.67
151	0	0	8.33	8.33	8.33	6.16	3.67
152	0	0	8.33	8.33	8.33	6.16	3.67
153	0	0	8.33	8.33	8.33	6.16	3.67
154	0	0	8.33	8.33	8.33	6.16	3.67

(d) Deterministic trajectories for the three genotypes

GEN	genotype 1	genotype 2	genotype 3
150	8.33	0.08	8.33×10^{15}
151	8.33	0.08	8.33×10^{15}
152	8.33	0.08	8.33×10^{15}
153	8.33	0.08	8.33×10^{15}
154	8.33	0.08	8.33×10^{15}

least 25 of the 100 Monte Carlo replications considered, the population had become extinct. Thus, this experiment suggests that with probability of at least 0.25, a population evolving from a single individual of genotype 1 would become extinct and in the contrary case the population of individuals of genotype 3 would continue to grow until the carrying capacity of the environment limited total population size, but the number of individuals of genotypes 1 and 2 would remain relatively small. Interestingly, in generation 6,000 of this experiment, the fractions of the 100 realizations of the process that contained zero individuals for each of the three genotype were 0.27, 0.3, and 0.26 for genotypes 1, 2, and 3, respectively.

From the point of view of evolutionary dynamics, the most interesting period of simulated evolution is that in which the stochastic process is in a transient phase before converging in distribution to a quasistationary distribution. Presented in Figure 6 are the graphs of $Q50$ and DET trajectories for the first 160 generations of evolution.

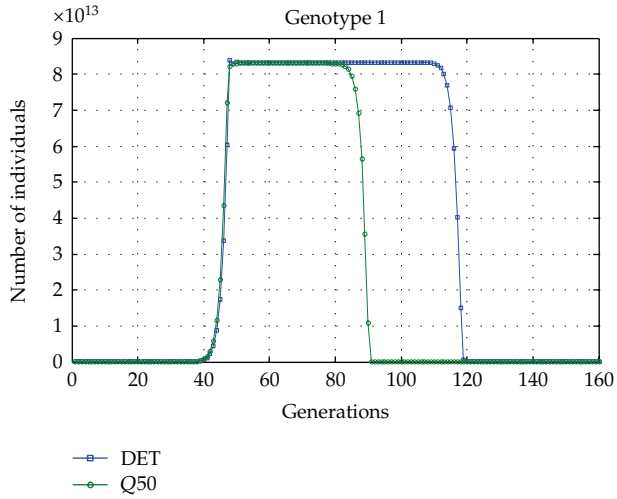

Figure 6: Estimated $Q50$ and DET trajectories for genotype 1 in experiment 4.

At the outset, it should be realized that the estimated trajectories in Figure 6, as well as in the two figures that follow, are based on those realizations of the process in which the population did not become extinct. As can be seen from this figure, in the rapidly evolving population of experiment 4, the median number of individuals of genotype 1 evolved from a single individual of genotype 1 in the initial generation to a median population size of over 8×10^{13} in a little over 40 generations. During this transitory period of evolution, the $Q50$ and DET trajectories nearly coincide until about generation 90, when the $Q50$ trajectory declines to small values. But, the DET trajectory does not undergo this steep decline until about 120 generations into the simulation experiment, and after 120 generations the two trajectories nearly coincide. This pattern is similar to those reported in previous section, but in experiment 4 evolution progressed much more rapidly.

In this experiment, the evolutionary pattern in the transitory phase of the evolving stochastic process for genotype 2 was also very interesting. Presented in Figure 7 are graphs of the $Q50$ and DET trajectories for genotype 2.

Throughout the 160 generations of evolution displayed in Figure 7, the DET trajectory for genotype 2 remains essentially near zero, but the $Q50$ trajectory begins to rise steeply after about 40 generations and then reaches a plateau of less than 4.5×10^{10} individuals and remains there until about generation 90, when it declines to nearly zero. This steep increase to about 4.5×10^{10} individuals of this genotype was a crucial event in the evolution of the population, because evidently the population of genotype 2 had reached a sufficiently high number to make it probable that the rare beneficial mutation $A_2 \rightarrow A_3$ with probability $\mu_{23} = 10^{-14}$ per generation occurs. After individuals of genotype 3 appeared in the evolving population, they soon rose to predominance by outcompeting individuals of genotypes 1 and 2. To complete the picture of the transitory period of evolution of the process, presented in Figure 8 are graphs of the $Q50$ and DET and for genotype 3 during this period.

Just as one would expect after viewing Figure 7 for genotype 2, the steep rise of the $Q50$ trajectory for genotype 3 to predominance in the population begins after about 90 generations of evolution and soon reaches a number of over 8×10^{15} individuals, at which the carrying capacity of the environment limits further growth of the population of individuals

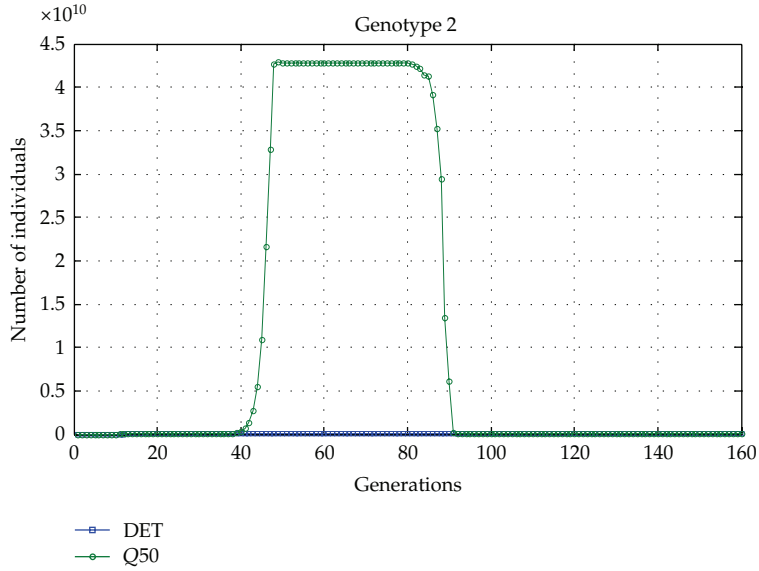

Figure 7: Estimated $Q50$ and DET trajectories for genotype 2 in experiment 4.

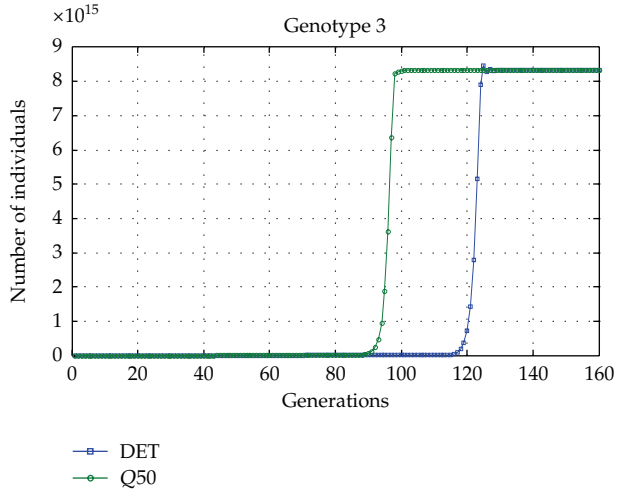

Figure 8: Estimated $Q50$ and DET trajectories for genotype 3 in experiment 4.

of genotype 3. The steep rise of the DET trajectory, however, does not begin until about 120 generations of evolution and thereafter coincides with the stochastic $Q50$ trajectory.

After inspecting the stochastic trajectories for genotype 3 in Table 3, one may get the impression that, given that the population did not become extinct and had converged to a quasistationary distribution, at stationarity there was essentially no variation around the $Q50$ trajectory. But, after viewing the trajectories MEAN and SD for genotype 3 in Table 3, it can be seen that such an impression was not entirely correct. From this table it can be seen that the $Q50$ trajectory of the quasistationary distribution is about 8.33×10^{15} for this genotype, which is very dependent on those realizations of the process in which the population did not

become extinct. But the MEAN and SD trajectories also take into account those replications of the experiment in which the population became extinct. According to Table 3, the values of these trajectories for the mean and standard deviation were about MEAN $= 6.16 \times 10^{15}$ and SD $= 3.67 \times 10^{15}$. Thus, if one examined the variations among the sample functions of the process after it had converged to a quasistationary distribution and examined them for many decimal places, variation would be present but its range would be small in relative terms.

7. Neutral Evolution—Mutation but No Selection

Within the class of branching processes under consideration, evolution is said to be neutral if all components of natural selection for each of the three genotypes are assigned equal values. In the model under consideration, there are two components of natural selection, namely, differences in reproductive success, expressed in terms of the expected number of offspring contributed by members of each genotype to the next generation, and competitive abilities among members of the three genotypes to compete for environmental resources expressed in terms of parameters that regulate the total population size for each genotype. Thus, in the experiment reported in this section, the components of the 1×3 vector $\boldsymbol{\lambda}$ were assigned the same values as those in (6.1) so that on average each member of the three genotypes in any generation would contribute two offspring to the next generation. Given these assigned parameters values, it was expected that total population size would increase rapidly and thus increase the likelihood that mutant genotypes would actually appear in a simulation experiment.

To expedite comparisons among the experiments presented in the preceding sections, the 3×3 matrix of mutation probabilities $\mathbb{M}$, containing probabilities of mutations among the three genotypes per generation, were assigned the same values as in (6.2). Unlike the $\boldsymbol{\beta}$ vector displayed in (6.3), however, in the experiment under consideration, the values in this vector were assigned as

$$\boldsymbol{\beta} = \left(10^{-14}, 10^{-14}, 10^{-14}\right), \tag{7.1}$$

so that there were no differences in the abilities to compete among the three genotypes. To initiate the simulated evolution of the population, the elements in the initial vector $\mathbf{X}(0)$ were assigned the same values as in (6.4). Underlying the choice of this initial vector was the desire to make it possible to study the long-term effects an initial population of consisting only 10,000 individuals of genotype 1 under the assumption of neutral evolution. In subsequent evolution of such a population, individuals of genotypes 2 and 3 could appear only from the process of mutation among the three genotypes. To further simplify the description of the experiment, all other assignments of parameter values were chosen just as in the first experiment described in Section 6.

In the three figures that follow, the DET and $Q50$ trajectories are plotted for the first 200 generations of evolution for each of the three genotypes. Then, in order to arrive at an understanding as to how these trajectories will appear after a long period of evolution, these trajectories were also plotted for the last 200 generations of the 6,000 generations considered in the experiment.

As can be seen from Figure 9(a), the graphs of the DET and $Q50$ trajectories for genotype 1 virtually coincide for the first 200 generations of the experiment. Therefore, for

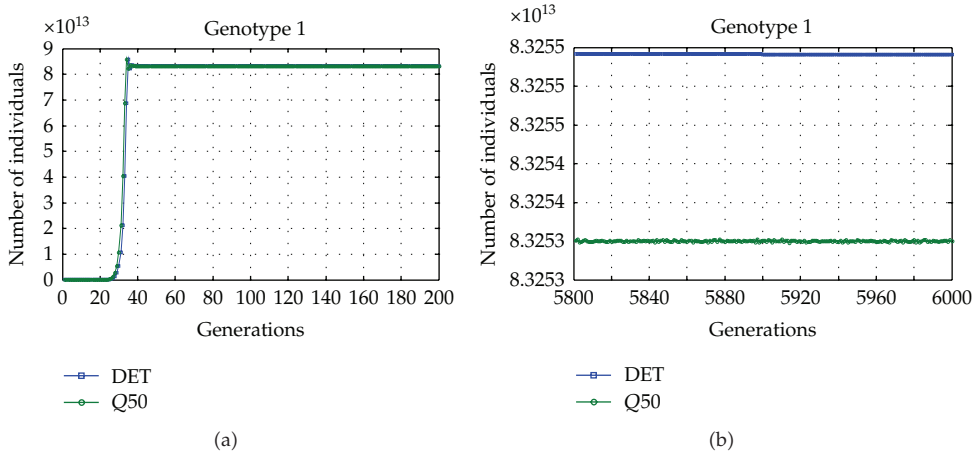

Figure 9: Graphs of the DET and $Q50$ trajectories for genotype 1.

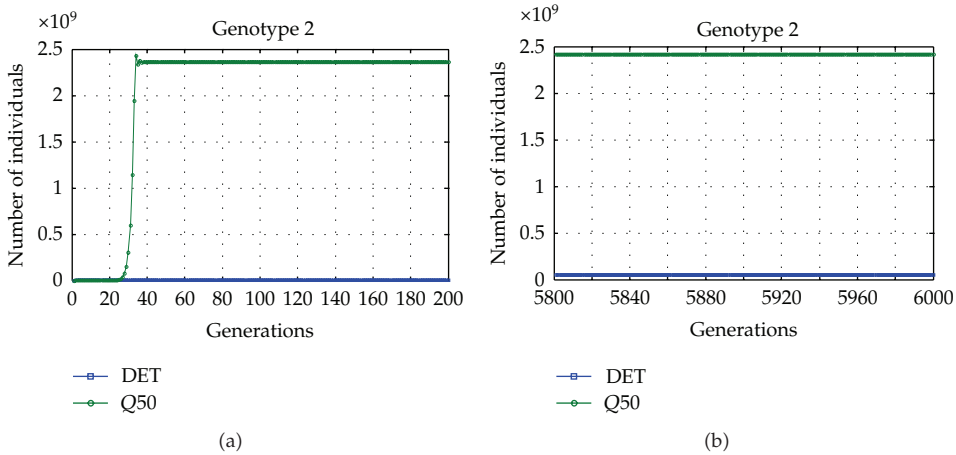

Figure 10: Graphs of the DET and $Q50$ trajectories for genotype 2.

genotype 1, the DET trajectory was a good predictor of central tendency for the stochastic process during the first 200 generations of the experiment. As can be seen, however, from Figure 9(b) for the last 200 generations, at first site it appears that the flat $Q50$ trajectory is uniformly below that of DET trajectory, but when one takes a closer look of the vertical scale of the graph, where the range of points is from 8.3253×10^{13} to 8.3255×10^{13}, it can be seen that the distance between the two trajectories is small. Thus, the *DET* trajectory is also a good predictor of the central tendency of the stochastic process for the last 200 generations of the experiment, but these trajectories do not coincide as they did in Figure 9(a).

By design, in the experiment under consideration, individuals of genotypes 2 and 3 could appear in the evolving population only if one or more mutations had been realized, since, by assumption, the initial population was composed of only individuals of genotype 1. In Figure 10(a) depicting the first 200 generations of evolution for individuals of genotype 2, it can be seen that DET trajectory rises steeply between 20 and 40 generations into the experiment and thereafter levels off at a value near 2.5×10^9. But, in this experiment, the $Q50$

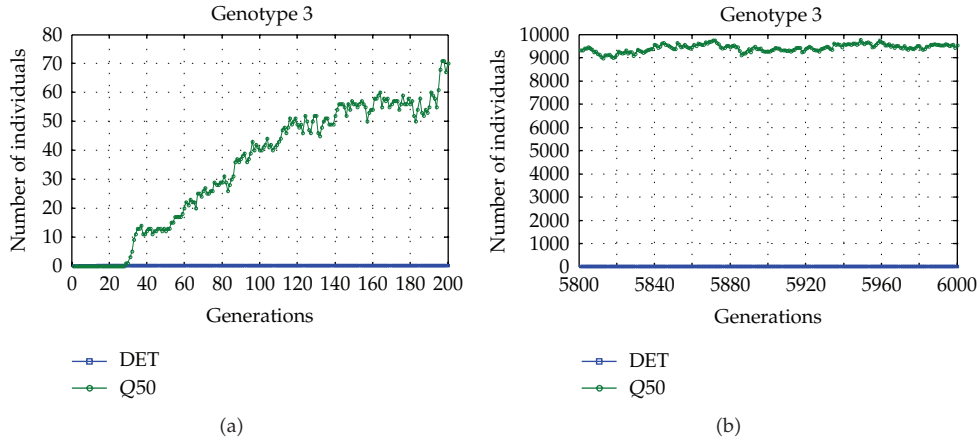

Figure 11: Graphs of the DET and $Q50$ trajectories for genotype 3.

trajectory remained near 0 throughout the first 200 generations of evolution. Therefore, it appears that, under an assumption of neutral evolution and individuals of genotype 2 could arise only through the process of mutation, the DET trajectory was not a good predictor for the central tendency of the process. As can be seen from Figure 10(a) depicting the last 200 generations of evolution for individuals of mutant genotype 2, the distance between the DET and $Q50$ trajectories remained apart for rather large distances so that, even in the long run, the DET trajectory was not a good predictor of central tendency for the stochastic process.

From Figure 11(a), it can be seen that the $Q50$ trajectory for the number of individuals of mutant genotype 3 in the simulated population had increased steadily from 0 in the initial generation to a level of about 70 individuals by generation 200. This result indicated that among the 100 Monte Carlo replications of the stochastic process, 50 contained more than 70 individuals in generation 200 and 50 of the other replications of the process in this generation contained less than 70 individuals. From Figure 11(b), it can be seen that by generation 5,800, the $Q50$ trajectory for individuals of genotype 3 had reached a number between 9,000 and 10,000 and remained in this range thereafter, indicating the number of individuals of genotype 3 had risen to significant numbers in the population but had remained relatively small when compared with the numbers of individuals of genotypes 1 and 2 in generation 6,000. This result was expected, because before individuals of genotype 3 could appear in the population, a mutation from genotype 1 to genotype 2, $A_1 \rightarrow A_2$, would need to appear and grow to a sufficiently large number before the mutation $A_2 \rightarrow A_3$ could occur. On the other hand, the DET trajectory for mutant genotype 3 had remained near 0 for the first and last 200 generations of the 6,000 generations of simulated evolution. Thus, like the situation for mutant genotype 2, the trajectory computed using the embedded deterministic model was not a good predictor for the central tendency of the stochastic process with respect to genotype 3 as it evolved in time. It should be mentioned in passing that if the simulated population were a real population observed after a long period of evolution, in the absence of data on reproductive success of each genotype as well as their competitive abilities, it would be difficult for an observer to decide whether the observed population had arisen as a result of mutation and selection or it was the result of neutral evolution with mutation.

When thinking about a population in which genotype 1 was predominate, it would be tempting to conclude that this genotype had a selective advantage over the other two,

Table 4: Deterministic Estimates of the Number Individuals of Each of the Three Genotypes in the Last Five Generations of the Experiment.

Generation	genotype 1	genotype 2	genotype 3
5996	$8.325541120 \times 10^{13}$	49,919,959.52	0.001496350936
5997	$8.325541119 \times 10^{13}$	49,928,285.06	0.001496850136
5998	$8.32551118 \times 10^{13}$	49,936,610.60	0.001497349419
5999	$8.32554117 \times 10^{13}$	49,944,936.14	0.001497848785
6000	$8.325441116 \times 10^{13}$	49,953,261.68	0.001498348234

but, by design, in this experiment, the predominance of genotype 1 in the population was due to the long-term effects of the founding population consisting of 10,000 individuals of genotype 1, for in this case, the momentum of population growth coupled with no selection, let the population of individuals of genotype 1 to continue to grow until it had reached the carrying capacity of the environment. Evidently, the observation that at 6,000 generations into the experiment, the number of individuals of genotype 2 was greater than that of genotype 3 was due primarily to the fact that in the simulation experiment, mutant genotype 2 appeared earlier than genotype 3.

Rather than relying totally on the above plots to get an impression of the magnitude of estimates of the number of individuals of each of the three genotypes as computed using the embedded deterministic model, it is of interest to actually view these numbers as displayed in Table 4 for the last 5 generations of the experiment as predicted by the embedded deterministic model.

As one can see, the estimates for the number of individuals of genotype 1 are indeed large, but the deterministic model had not converged during the last 5 generations of the experiment, because there was agreement in all five generations at only three decimal places. Interestingly, the estimate of the number of individuals of genotype 2 was nearly 50 million, but as can be seen from Figure 10(b), this number was relatively small when compared with 2.5×10^9. The small estimates of the number of individuals of genotype 3 in the last 5 generations are small as actually shown in the lower figure of Figure 11. The rather noisy numbers presented in Table 4 are representative of the behavior of the embedded deterministic model under the assumption of neutral evolution. In those experiments reported in this paper, in which the driving force of evolution was some component of selection, it was observed that, in the long run, the deterministic trajectories were less noisy than for the case of neutral evolution.

8. Discussion

The class of self-regulating branching processes described and applied in this paper may be extended in several ways that would expedite a wider range of applications. An assumption underlying the class of branching processes applied in the foregoing sections had the property that generations are distinct and nonoverlapping, which implies, for the case reproduction by division of a mother cell into two daughter cells, that division of all cells in the population must be synchronized. In many cell populations, however, this is not the case, because, in general, divisions in these populations are not synchronized. There is a class of branching process that evolves in continuous time and has the property that reproduction is not synchronized among individuals. This class of branching processes is sometimes referred

to as general branching processes. For a treatment of the one-type case see Jagers [4] and for a treatment of the multitype case Mode [2] may be consulted. In both these books, the branching processes considered are not self-regulating and would thus need to be generalized to render them more realistic from the biological point of view. However, in chapter 12 of Mode and Sleeman [12] a formal treatment is given to a class of self-regulating age-dependent two-sex branching processes with partnership formation and dissolution, which from the genetic point of view, accommodates a single autosomal locus with two alleles. Moreover, a simplification of this class of processes could be accomplished mathematically and used to conduct computer simulation experiments by writing the appropriate software in a programming language of a developer's choosing. Given such software, computer experiments, which would be extensions of those reported in this paper, could be executed.

There is another and perhaps more important field of research that has a rich potential for applications of self-regulating branching processes. For nearly a decade or more, there has been an extensive research effort involving many investigators to find signals of the impact of recent natural selection in genome-wide sweeps of the human genome and those of other species. The recent review paper by Stranger et al. [19] may be consulted for details along with a long list of references. Another reference with a long list of cited papers on regions of the human genome that have been implicated as sites of positive selection is that of Sabeti et al. [20] in the supporting online material. Among the many statistical techniques used in these research efforts are those that apply computer simulation methods which simulate the evolution of model genomes with up to one million base pairs under the influence of mutation, linkage and natural selection. Chapter 14 in the work by Mode and Sleeman [12] contains an overview of these simulation methods, which have been proposed and applied by several teams of investigators. Also contained in this chapter is a critique of these methods, because of the lack of transparency in the description of the mathematical algorithms underlying their methods so that, in general, it is not clear to a reader literate in mathematics as to the fundamental mathematical nature of these algorithms.

Under these circumstances, many readers of such papers will view the reported results of simulation experiments using various methods with considerable skepticism, because of a lack of mathematical transparency of the algorithms implemented in the simulation software. As a first step in overcoming this lack of transparency, in chapter 14 of the work by Mode and Sleeman [12], an attempt has been made to write a transparent account of a set of preliminary algorithms designed to simulate the evolution of model genomes that accommodate various types of mutations, linkage, and natural selection. Included in the types of mutations are nucleotide substitutions, deletions, inversions, and gene conversions. A goal of such a research program would be that of incorporating such software within a branching process framework so that the components of natural selection could be quantified and the combined effects of mutation and selection could be simulated at the genomic level. In contrast with the abstract level at which mutations have been treated in this paper with no formal connection to changes in the genome, linking the effects of mutations and selection to a model genome would be more informative and satisfying.

In carrying out such a program, a detailed account of the mathematics underlying the algorithms would increase the transparency of the methods used to obtain the results of computer simulation experiments and would thus increase their credibility to a reader. It should be mentioned, in passing, that in this paper an attempt has been made to present the mathematics underlying the software in a transparent fashion so that a reader can view the reported results of simulated experiments with confidence. For, in principle, given

the mathematics outlined in this paper, software could be written by a reader to check the validity and reproducibility of the reported experimental results.

Whenever an investigator is considering the computer implementation of some mathematical structure of interest, an issue that always arises is that of deciding what programming language or combination of programming languages and available software packages to use in conducting and reporting the results of computer simulation experiments. In the computer experiments reported in this paper, a version of a programming language called APL 2000 was used to write the software that implemented the underlying mathematics as well as the procedures to produce statistical summarizations of simulated data produced by using Monte Carlo methods. A well-known software package called MATLAB was used to produce the graphs and the widely used spreadsheet software called excel, produced by Microsoft, was used to transfer simulated data among the computers used by the research team. These choices of software were made, primarily, because the principal investigator had many years of experience with the APL programming language and the software packages. But, given the mathematics underlying our software, any team of investigators would be free to write software, using programming languages and software packages that worked best for them.

After the required software packages had been developed, they could be used not only in cutting-edge research but also in beginning and advanced courses on biological evolution to graphical illustrate the processes of mutation and various components of selection as driving forces of the evolutionary process. In advanced courses, populated by people using statistical software packages or students with a grasp of the principles underlying stochastic processes and statistical inference, the technical details could be presented and discussed. But, for beginning courses on evolution, such software could be used in connection with the development of computer animation of difficult concepts underlying the science of evolution that would aid beginning students to get a firmer grasp of the subject.

It was decided at the outset to use Monte Carlo simulation methods to study the emergence and survival of mutations in an evolving population, because it is difficult to maintain sufficiently large experimental populations to ensure that a rare mutation will appear in an experiment with reasonably high probability. But, nevertheless, it would be of interest to investigate whether the class of branching processes under consideration, could be used as dynamic models for statistical inference in the experimental study of an evolving population. Such an investigation is beyond the scope of this paper, but it may be of interest for an interested reader to consider a Bayesian-Monte Carlo integration strategy proposed in work by Mode [21] for dealing with interfacing stochastic models of HIV/AIDS epidemics with data. At the moment, it appears to be plausible that the strategy used in that paper could also be used in the experimental study of the class of evolving populations considered in this paper, but the details regarding the actual implementation of such a strategy will be left as a research project for the future.

References

[1] T. E. Harris, *The Theory of Branching Processes*, Springer-Verlag, New York, NY, USA, 1963.
[2] C. J. Mode, *Multitype Branching Processes—Theory and Applications*, American Elsevier, New York, NY, USA, 1971.
[3] K. B. Athreya and P. Ney, *Branching Processes*, Springer-Verlag, New York, NY, USA, 1972.
[4] P. Jagers, *Branching Processes with Biological Applications*, John Wiley & Sons, New York, NY, USA, 1975.
[5] S. Asmussen and H. Hering, *Branching Processes*, vol. 3, Birkhauser, Boston, MA, USA, 1983.
[6] M. Kimmel and D. E. Axelrod, *Branching Processes in Biology*, Springer-Verlag, New York, NY, USA, 2002.

[7] P. Haccou, P. Jagers, and V. A. Vatutin, *Branching Processes—Variation, Growth, and Extinction of Populations*, Cambridge University Press, Cambridge, UK, 2007.
[8] H. K. Alexander, *Modelling Pathogen Evolution with Branching Processes*, M.S. thesis, Queen's University, Ontario, Canada, 2010.
[9] R. A. Fisher, *The Genetical Theory of Natural Selection—Second Revised Edition*, Dover Publications, New York, NY, USA, 1958.
[10] C. J. Mode and R. J. Gallop, "A review on Monte Carlo simulation methods as they apply to mutation and selection as formulated in Wright-Fisher models of evolutionary genetics," *Mathematical Biosciences*, vol. 211, no. 2, pp. 205–225, 2008.
[11] C. J. Mode and C. K. Sleeman, *Stochastic Processes in Epidemiology—HIV/AIDS, Other Infectious Diseases and Computers*, World Scientific, London, UK, 2000.
[12] C. J. Mode and C. K. Sleeman, *Stochastic Processes in Genetics and Evolution—Computer Experiments in the Quantification of Mutation and Selection*, World Scientific, London, UK, 2011.
[13] A. Martin and J. Hawks, *Major Transitions in Evolution*, The Teaching Company, Chantilly, Va, USA, 2010.
[14] R. M. Hazen, *Origins of Life*, The Teaching Company, Chantilly, Va, USA, 2005.
[15] J. Gleick, *Chaos—Making a New Science*, Penguin Group, New York, NY, USA, 1987.
[16] D. Gulick, *Encounters With Chaos*, McGraw-Hill, New York, NY, USA, 1992.
[17] S. Strogatz, *Chaos*, The Teaching Company, Chantilly, Va, USA, 2008.
[18] C. J. Mode, T. Raj, and C. K. Sleeman, "Monte Carlo implementations of two sex density dependent branching processes and their applications in evolutionary genetics," in *Applications of Monte Carlo Methods in Biology, Medicine and Other Fields of Science*, C. J. Mode, Ed., pp. 273–296, INTECH, 2011.
[19] B. E. Stranger, E. A. Stahl, and T. Raj, "Progress and promise of genome-wide association studies for human complex trait genetics," *Genetics*, vol. 187, no. 2, pp. 367–383, 2011.
[20] P. C. Sabeti, S. F. Schaffner, B. Fry et al., "Positive natural selection in the human lineage," *Science*, vol. 312, no. 5780, pp. 1614–1620, 2006.
[21] C. J. Mode, "A bayesian Monte Carlo integration strategy for connecting stochastic models of HIV/AIDS with data," in *Deterministic and Stochastic Models of AIDS Epidemics and HIV Infections with Intervention*, W. Y. Tan and H. Wu, Eds., chapter 3, pp. 61–76, World Scientific, 2005.