Age-dependent stochastic models for understanding population fluctuations in continuously cultured cells

Evgeny B. Stukalin\textsuperscript{1,2}, Ivie Aifuwa\textsuperscript{2,3}, Jin Seob Kim\textsuperscript{2}, Denis Wirtz\textsuperscript{2,3} and Sean X. Sun\textsuperscript{1,2}

\textsuperscript{1}Department of Mechanical Engineering, \textsuperscript{2}Physical Sciences in Oncology Center, and \textsuperscript{3}Department of Chemical Engineering, The Johns Hopkins University, Baltimore, MD 21218, USA

For symmetrically dividing cells, large variations in the cell cycle time are typical, even among clonal cells. The consequence of this variation is important in stem cell differentiation, tissue and organ size control, and cancer development, where cell division rates ultimately determine the cell population. We explore the connection between cell cycle time variation and population-level fluctuations using simple stochastic models. We find that standard population models with constant division and death rates fail to predict the level of population fluctuation. Instead, variations in the cell division time contribute to population fluctuations. An age-dependent birth and death model allows us to compute the mean squared fluctuation or the population dispersion as a function of time. This dispersion grows exponentially with time, but scales with the population. We also find a relationship between the dispersion and the cell cycle time distribution for synchronized cell populations. The model can easily be generalized to study populations involving cell differentiation and competitive growth situations.

1. Introduction

All cell types reproduce by repeated cycles of division and growth. If a cell divides symmetrically, then both daughter cells are expected to divide roughly at the same rate when environmental factors and cell population densities are kept constant. In a recent experiment where individual bacterial cell division events were monitored, it was shown that the time between cell divisions (or generation time) is a highly stochastic variable, even for clonal cells \cite{1}. Furthermore, the generation time distribution is a non-monotonic function of chronological cell age, \(a\), defined as the time since its birth. Typical probability distributions of the time between successive cell divisions from experiments are shown in figure 1b,c. This behaviour is not restricted to prokaryotes, and is also true for eukaryotes. It is known that the probability for cell division depends both on age and on cell size \cite{1,2}. For instance, for cells of the same size (mouse lymphoblasts L1210) the probability of division increases with cell age and for cells of the same age, the probability of division increases with cell size \cite{2}. Cell division and cell growth cycles seem to be interrelated \cite{3}. Other factors are also involved in the division decision. For instance, in Escherichia coli cells the specific growth rate of a population is found to linearly correlate with RNA/protein ratio and ribosome production \cite{4}. The linear correlation holds for cell growth with doubling times varying from minutes to hours. These studies and others demonstrate that the cell cycle decision is an intrinsically noisy process where cell age and other factors are important, but biochemical noise also dominates.

Given the probabilistic nature of cell division events, we ask how cell division time variation ultimately translates to cell number fluctuations in a growing population. Understanding this connection is important for understanding tissue growth and homeostasis, where not just the average population but also population fluctuations must be carefully controlled \cite{5}. Quantitative models for this question must be stochastic in nature. The simplest model for studying
stochastic population dynamics is a Markovian master equation (or a birth–death process) with constant division and death probabilities per unit time \[6\]. The average population and population fluctuations for this model can be solved exactly (see the electronic supplementary material, section A). At long times, the relative population fluctuation is a function of time. We find that there is a direct link between division time distribution, \(\omega(a)\), and population fluctuations. In the case of synchronized populations, a simple relationship between \(\omega(a)\) and \(\sigma(t)\) is found. We also examine more realistic models of cell population homeostasis where the death rate (or the net growth rate) is controlled by the overall population.

Note that in the long time limit, population growth described by the age-dependent model exhibits some properties of the age-independent (stationary) stochastic model, e.g. the average number of cells grows exponentially. However, the predicted exponential population growth rate and population size fluctuations (dispersion) around the mean are different in these models, even when the average cell cycle times are kept constant. These differences are highlighted in this paper. There have been other quantitative studies of cell growth with time-dependent division rates \[5,7–10\]. For instance, Horowitz \textit{et al.} \[7\] consider a probabilistic model of microbial growth and mortality in which both cell division and cell death transition probabilities depend on time, e.g. in the form of logistic expressions. Two simulation algorithms are considered: one for tracking the fates of individual cells; another for simulation at the population level using a simplified model. Although the authors have succeeded in reproducing the experimental growth curves for a bacteria, the age of cells as an important characteristics of mitosis is not considered.

Another approach is applied for stochastic modelling of population growth by Pin & Baranyi \[8\] based on the assignment of processes in cell division, synchronous growth curves and/or age distributions of exponential cultures within deterministic models \[11–15\]. For instance, Bremer \[14\] has reduced the cell cycle variability of \textit{E. coli} to variability of times between

\begin{equation}
\lim_{t \to \infty} \sigma(t) = \frac{1}{\sqrt{N_0}} \left[ \frac{k_b + k_d}{k_b - k_d} \right]^{1/2}.
\end{equation}

where \(N(t)\) is the average population and \(\sigma(t)\) is the average population fluctuation (dispersion). \(k_b\) and \(k_d\) are the cell birth and death rates, respectively. \(N_0\) is the initial population at \(t = 0\). This scaling result is simple. However, since the model assumes constant division and death probabilities, it is in fact not applicable to typical cell division processes since the division probability per unit time is clearly not constant (figure 1). A constant division rate amounts to the assumption that a newly born cell has just as much likelihood to divide as a mature cell. But this is not the case, as shown in figure 1. In this paper, we develop a stochastic model where division and death probabilities are functions of cell age, and use experimentally measured division time distributions to predict population fluctuations. We consider a homogeneous cell population and with no correlation between interdivision times for cells in different generations. The age \(a\) of cells is considered explicitly as the variable determining the propensity of cell division. We examine growth dynamics when the mean division time is kept constant, but the spread of the division time distribution changes (figure 1d). We show that our model is equivalent to the solution of the von Foerster equation when considering the average population. However, we also obtain quantitative results regarding population fluctuations as a function of time.
the end of DNA replication and the next cell division. This is also a variation of the original idea by Smith & Martin [11] that there exists a highly variable period before DNA replication. Engelberg [16] has derived a simple model establishing the relationship between the loss of synchrony in cell divisions and the width of division time distribution in synchronized cell cultures. However, there is a lack of systematic studies analysing population size fluctuations of growing cell colonies. Our paper is aimed at providing such analysis and establishing the link between population size fluctuations and the stochastic nature of the cell division process at different growth conditions (both in the absence and in the presence of cell death, for conditions of restrictive growth, etc.).

Our age-dependent model is in principle applicable to any cell growth situation, e.g. in vitro in a reactor as well as in vivo where multiple cell types may influence each other. In conditions of saturating nutrient and low cell densities, cell division and cell death probabilities are known and the application of our model is straightforward. In conditions of limiting nutrient and/or space, competition between individuals cells results in death rate proportional to total number of cells. A simple example of this situation is also considered below. For multiple types of interacting cells, cell division and death probabilities will depend on signalling, cooperative as well as competitive interactions. Additional models and information are needed to quantify cell division and cell death probabilities. Our model is not applicable, however, to cell division probabilities that has memory of prior growth conditions. This is because our model is Markovian. In addition, our model currently does not consider spatial patterns of growing cell colonies, only the overall population in a well mixed situation.

2. Model

2.1. Cell population and division in the age-time representation

Before discussing the stochastic dynamics of cell population, it is useful to consider deterministic dynamics of the average population as a function of time. For this purpose, the von Foerster equation describes the cell number density as a function of age and time [17]. The age of a cell, \( a \), is defined as the time elapsed since its birth. The ongoing time, \( t \), can be chosen arbitrarily, but it is natural to measure this from the time of cell culture seeding. In the age-time representation the equation is written as

\[
\frac{\partial n}{\partial t} + \frac{\partial n}{\partial a} = -\lambda(a)n, \tag{2.2}
\]

where \( n(a,t) \) represents the number of cells in the age interval from \( a \) to \( a + da \) at time \( t \) and \( \lambda \) is a cell loss factor. The total number of cells at time \( t \) is therefore

\[
N(t) = \int_0^\infty n(a,t) \, da. \tag{2.3}
\]

The loss factor \( \lambda \) is assumed to be due to both cell division (mitosis) as well as cell death or disappearance. The von Foerster equation is a form of a conservation equation that describes the flow of population densities. From conservation of population, the number of cells \( n(a + \Delta a, t + \Delta t) \Delta a \) at the age \( a + \Delta a \) at time \( t + \Delta t \) equals the number of cells that have matured from an earlier age \( a \) at time \( t \), i.e. \( n(a, t) \Delta a \) less the number of cells that have been lost. The number of lost cells is assumed to be proportional to the number of cells \( n(a, t) \Delta a \) and the time interval \( \Delta t \). Cell loss due to mitosis should be understood as the mother cell of age \( a \) is replaced by two replicates of zero age \( (a = 0) \). If there are no other reasons for cell loss besides mitosis, then

\[
\lambda = \lambda(a), \text{ where } \lambda(a) \text{ is an age-dependent probability per unit time that a cell undergoes mitosis.}
\]

The solution and the properties of the von Foerster equation regardless the initial and boundary condition have been discussed by Trucco [18,19]. Rubinson [13] gave a solution for the specific initial and boundary conditions

\[
n(a,0) = N_0 \delta(a), \tag{2.4}
\]

and cell birth is described by the boundary condition

\[
n(0,t) = 2 \int_0^\infty k(a)n(a,t) \, da. \tag{2.5}
\]

Thus, the age-dependent function \( k(a) \) accounts for the variability in observed generation times. Here, we note that the total number of cells \( N(t) \) at time \( t \) is obtained by integrating \( n(a,t) \) for all ages and is found as the infinite sum of terms (\( j = 1, 2, \ldots \)), each represents the contribution of the \( j \)th generation (see the electronic supplementary material, section B).

The probability per unit time that a cell undergoes mitosis is related to probability density function \( \omega(a) \) for cell division times:

\[
k(a) = \frac{\omega(a)}{1 - \int_0^a \omega(a') \, da'}. \tag{2.6}
\]

This relation is derived as follows. Let us assume that we have \( N \gg 1 \) cells of zero age at time \( t = 0 \). In this case, time \( t \) and age \( a \) are synchronized and can be considered as a single variable, say \( a \). By time \( a > 0 \), the number of cells that remain undivided is \( N(1 - \int_0^a \omega(a') \, da') \). The number of cells that undergo mitosis during a time interval from \( a \) to \( a + \Delta a \) is \( N \omega(a) \Delta a \). Thus, the conditional probability that a cell divides between \( a \) and \( a + \Delta a \) (provided it had not divided earlier) is given by the ratio of the number of cells divided during this time interval to the total (remaining) number of cells at time \( a \), i.e.

\[
\Delta P(a) = \omega(a) \Delta a/(1 - \int_0^a \omega(a') \, da') = k(a) \Delta a.
\]

The cell division probability per unit time that \( k(a) \) is obtained from \( \Delta P(a) \) above resulting in equation (2.6).

Following Powell [12], in the limit of long times, the solution of the von Foerster equation reduces to exponential growth:

\[
N(t) = N_0^r \exp(\nu t) \tag{2.7}
\]

with the steady-state growth rate, \( \nu \), determined by generation time distribution \( \omega(a) \) only. The pre-factor \( N_0^r \) depends on the initial age distribution and is proportional to the initial number of live cells \( N_0 \). The following ansatz satisfies the von Foerster equation at long times:

\[
n(a,t) = \exp(\nu t) r(a), \tag{2.7}
\]

where \( r(a) \) is a function that only depends on the age \( a \). Substitution equation (2.7) into equation (2.2) gives

\[
\frac{dr}{da} = -[\nu + k(a)]r, \tag{2.8}
\]

and the solution is

\[
r(a) = r(0) \exp[-\nu a - \int_0^a k(a') \, da']. \tag{2.9}
\]

Using equation (2.9), we re-write the boundary condition equation (2.5) in the form (after cancelling the non-negative
Equation (2.10) then becomes
\[ 1 = 2 \int_0^\infty k(a) \exp \left( -va - \int_0^a k(a') \, da' \right) \, da. \]  
(2.10)

From equation (2.6), one may show that
\[ k(a) \exp \left( -\int_0^a k(a') \, da' \right) = \omega(a). \]  
(2.11)

This equation has a simple physical interpretation. If we consider mitosis as a one-step irreversible transition, the first term is the transition probability per unit of time, while the second term is the survival probability \( \Omega(a) \). Thus, the distribution of cell division times can be expressed through the survival probability as \( \omega(a) = -(\partial \Omega(a))/\partial a \).  
Equation (2.10) then becomes
\[ 2 \int_0^\infty \exp(-va) \omega(a) \, da = 1. \]  
(2.12)

This equation gives the fundamental relationship between division time distribution, \( \omega(a) \), and the long time growth rate of the population, \( v \).

Using the long time ansatz solution, we can also find the steady-state age distribution \( g(a) \), which is
\[ g(a) = \frac{n(a,t)}{\int n(a,t) \, da} = \frac{\exp(-r(a)/v)}{\int_0^\infty \exp(-r(a')/v) \, da'}. \]  
(2.13)

Since \( r(a)/r(0) = \exp[-\int_0^a k(a') \, da'] \), and in light of equation (2.10), we have
\[ \int_0^\infty \frac{r(a)}{r(0)} \, da = \int_0^\infty \exp(-a) \omega(a) \, da \exp(-a) \, da = \frac{1}{2v}. \]  
(2.14)

where we have used the fact that \( \exp(-\int_0^a k(a) \, da) = 0 \). Thus, for the age distribution, we obtain
\[ g(a) = 2v \exp(-va) \int_0^a \omega(a') \, da'. \]  
(2.15)

We now can examine the behaviour for different generation time distributions \( \omega(a) \). When the generation time is not dispersed and all cells divide exactly at the same age, then \( \omega(a) = \delta(a - \tau) \), where \( \delta(a - \tau) \) is the Dirac delta function satisfying the condition \( \int_0^\infty \delta(a - \tau) \, da = 1 \). The growth rate is \( v = \ln 2/\tau \) by equation (2.12) and the doubling time of the cell population equals the average generation time \( \tau^* = \ln 2/v = \tau \). The age distribution is
\[ g(a) = 2 \ln \frac{2}{\tau} \exp\left( -\frac{a \ln 2}{\tau} \right) \quad \text{for} \quad a < \tau. \]  
(2.16)

The age distribution described by equation (2.16) is an ideal age distribution assuming that the population has reached the steady state. In reality \( g(a) \) is smoothed-out. The average age of population with an ideal age distribution is obtained as \( \bar{a} = \int_0^\infty a g(a) \, da = (1/\ln 2 - 1) \tau \approx 0.44 \tau \). Let us consider another hypothetical case when the rate of cell division is age-independent \( k(u) = 1/\tau \) corresponding to exponential distribution for generation times \( \omega(a) = 1/\tau \exp(-a/\tau) \). In this case, the average growth rate is equal to \( v = 1/\tau \) by equation (2.12) and the average doubling time is equal to \( \tau^* = \ln 2/v = \tau \ln 2 \). The age distribution is the exponential function \( g(a) = 2/\tau \exp(-2a/\tau) \) with the average age of population equal to a half of the average generation time \( \bar{a} = \tau/2 \).

Powell has pointed out that in general (see equation (2.12)) the average doubling time \( \tau^* \) is not equal to the average generation time \( \tau \) but in practice this difference varies depending on the division time distribution.

For most cells, experimental data show that the cell division time distribution is neither exponential nor a delta function. Some typical results are given in figure 1. The realistic generation time distribution is often best described by the \( a_0 \)-shifted gamma distribution \([9,12,13,20–22] \)
\[ \omega(a) = (a - a_0)^{a-1} \exp(-(a - a_0)/\beta) / \beta \Gamma(a), \]  
(2.17)

where \( \alpha > 1 \), \( \beta > 0 \) and \( a_0 \geq 0 \) are parameters of shape, scale and shift of generation time distribution, respectively. Typically, data show that \( \omega(a) \) is not symmetric but skewed towards higher values (the skewness of gamma distribution is \( 2/\sqrt{\alpha} \)). The gamma distribution is widely used in chemical kinetics for dwell-time distributions (e.g. in analysing enzyme kinetics) \([23,24]\). The probability distribution of first passage times for a multi-step irreversible process 
\[ A \xrightarrow{k_1} X_1 \xrightarrow{k_2} X_2 \xrightarrow{k_3} \cdots \xrightarrow{k_{m-1}} X_m \xrightarrow{k_m} B \]  
with \( m \) rate-limiting steps can be shown to correspond to equation (2.17). The average growth rate of the population is the solution of equation (2.12) which reads
\[ a_0 v + a \ln(1 + v \beta) = \ln 2. \]  
(2.18)

For simplicity, take \( a_0 = 0 \). The average doubling time in this case is \( \tau^* = \ln(2/v) = \ln(2/\sqrt{\alpha} - 1) \), whereas \( \tau^* = a \beta \). For \( \alpha \gg 1 \), the average doubling time, \( \tau^* \approx a \beta = \tau \), equals the average generation time.

Interestingly, for \( E. \ coli \) the division time distribution (figure 1 and [1]) is not completely described by the shifted gamma distribution. Rather, it is better described by a combination of two exponential functions with a sharp peak. The mechanistic argument for this was discussed by Bremer \([14]\).

In addition to cell birth, cell death can also contribute to cell loss. In this paper, we consider constant cell death probability per unit time \( k_d \). In this case, \( \lambda = k(b) + k_d \) in equation (2.2) with the same boundary condition given by equation (2.5). If we are interested in the long time limit solution identical to equation (2.7), the net growth rate in the presence of cell death becomes \( v = v_0 - k_d \), where \( v_0 \) is the growth rate in the absence of cell death found in equation (2.10). Of course, if \( k_d \) depends on the cell age, this simple relation does not hold. The electronic supplementary material, figure S1 of SI shows simulation results of \( v \) as a function of constant \( k_d \) for age-dependent cell division. Note that cell death with age-independent probability does not perturb steady-state age distribution since cells of different ages die with the same rate. Constant cell death rescales both numerator and denominator of equation (2.13) by a factor \( e^{-k_d t} \).

The von Foerster equation describes the average age-dependent population change. In this sense, it does not contain information about population fluctuations. In order to obtain fluctuation information, we must incorporate age dependence in a stochastic model of cell population growth. In §2.2, we describe a stochastic simulation of age-dependent population growth and analyse the results. In general, the mathematical quantity we have to consider is the probability distribution of cell number density as a function of age \( P[n(a),t] \). This quantity is a functional and it is possible to write a stochastic equation to describe its evolution. The theoretical treatment of this model will be discussed elsewhere.
2.2. Stochastic simulations of age-dependent cell populations

Our stochastic algorithm takes explicitly into consideration the age-dependence of rates of ‘elementary’ events. Mitosis is considered in the present work as a one-step event with a probability per unit time \( k(a) \) that depends on the cell age \( a \), i.e. the time elapsed since its birth. Our simple simulation model is a version of the kinetic Monte Carlo algorithm in which the acceptance probability of any event for each cell depends on the exponential factor \( \exp[-(k(a) + k_d)\Delta t] \), where \( k(a) \) and \( k_d \) are cell division and cell death probabilities per unit time for the \( j \)-th cell, respectively. \( \Delta t \) is a time step. A random number \( 0 < r_1 < 1 \) is compared with \( \exp[-(k(a) + k_d)\Delta t] \) to determine whether an event occurs. At each time step, \( t = m\Delta t \), a particular cell remains alive but undivided if \( r_1 < f_1(a) = \exp[-(k(a) + k_d)\Delta t] \). Otherwise this cell dies or divides. If an event occurs for the \( j \)-th cell, the choice between mitosis and cell death is made by comparing another random number \( 0 < r_2 < 1 \) with \( f_2(a) = k(a)/k(a) + k_d \). Thus, cell division occurs if \( r_1 > f_1(a) \) and \( r_2 < f_2(a) \), giving rise to a pair of newborn cells with ages equal to 0. The \( j \)-th cell dies if \( r_1 > f_1(a) \) and \( r_2 > f_2(a) \). Note, that if no event occurs with a particular cell at a given time step, the cell becomes older by \( a = a + \Delta t \). As time is incremented by \( \Delta t \), a small portion of the cells undergoes mitosis or death. Thus, ages of cells are tracked explicitly. This procedure generates a histogram of cells with different ages \( n(h, a, t)\Delta a \), from which we can obtain the number density of cells as a function of cell age at time \( t \), where \( \Delta a \) is the bin size of the histogram. \( n(h) \) is the cell number density from the particular simulation trajectory. This algorithm amounts to a Markovian stochastic simulation of age-dependent cell division and growth.

Note that since the model is stochastic, for each simulation, the obtained cell number density is different. The proper interpretation is that the age distribution averaged over many simulations, \( \langle n(h, a, t) \rangle \), will reproduce the von Foerster equation result. The average is over many simulations. From the average population density, we can also compute the total number of cells by integrating over age. We also compute the population fluctuation, which is the average variation of a particular simulation from the mean, defined below in equation (3.3).

Figure 2 shows the average number of cells \( N(t) \) plus and minus standard deviation \( \sigma(t) \) of the mean obtained from our simulations. The population is grown from a single cell of age \( a = 0 \). Results of our simulations are in excellent agreement with analytical solution of the von Foerster equation shown also for comparison in figure 2. The contributions from different generations to \( N(t) \) are also shown. All contributions excluding the initial generation go through a maxima. In separate short runs, we analyse the distribution of waiting times to first cell division starting from a single cell of zero age. This histogram is also in the full agreement with theoretical generation time distribution (see the inset in figure 2). Thus, these two comparisons demonstrate the relevance and accuracy of our simulation algorithm.

3. Results

3.1. Average growth rates of cell colonies and doubling times

We first examine growing clonal populations with no cell death. In most of the simulations, cells are grown from a single clone with age \( a = 0 \). In some simulations, we start with more than one cell. We perform simulations for (a) initially synchronized and (b) non-synchronized cell populations. In the first case (a), the initial age of the cell (or cells) is taken to be zero. In the second case (b), the initial age of the cells is chosen randomly between 0 and \( T \) (typically, \( T = \tau \)). We find that in all cases at long times the average number of cells \( N(t) \), defined

$$N(t) = \left\langle \int_0^\infty n(h, a, t) \, da \right\rangle, \tag{3.1}$$

grows exponentially (see the electronic supplementary material, figures S2 and S3 of SI). For initially synchronized populations for the case of the distribution with the smallest spread (distribution I, see electronic supplementary material, figure S4), we observe the step-like pattern in the average growth trajectory.

Figure 2. Growth of a clonal cell colony using age-dependent stochastic model. (a) The average number of cells, \( N(t) \), grown from a single cell of age \( a = 0 \) (synchronized initial condition) for division time distribution II is shown from 1000 simulations (thick green line). Magenta lines represent \( N(t) \) plus and minus the standard deviation \( \sigma(t) \). The red line is the analytical solution of the von Foerster equation for synchronized populations. The thin curves (with maxima) represent the contributions from different generations (see the electronic supplementary material, section B). The results of our simulations are in excellent agreement with analytical solution. The inset depicts the probability density function of division times obtained from simulations (10^6 runs) for the theoretical division time distribution II (see caption to figure 1d). Our simulation recovers the original division time distribution \( \sigma(a) \) (red line). (b) The computed average doubling times compared to the theoretical result of equation (1.12) for division time distributions (I–IV) and the exponential distribution (V). There is complete agreement between the theoretical prediction and our simulation.
(‘mitotic waves’) [25–27]. However, eventually these ‘oscillations’ are damped at long times. The loss of synchrony results from the variation in division times. The growth rates $v$ are found from exponential fits to the average number of cells as a function of times. For comparison, for the exponential distribution of division times $V$, the cell division rate is constant, so synchronization does not matter. Note that for all these distributions I–V, the average generation times are the same ($\tau = 20$ arb. units) but their spreads are different. We characterize the width of generation time distributions by the dimensionless coefficient of variation of $\tau$ defined as

$$c_v = \frac{1}{\tau} \int_0^\infty (a - \tau)^2 \omega(a) \, da \right)^{1/2}. \quad (3.2)$$

For the exponential division time distribution $c_v = 1$. For division time distributions I–IV in figure 1, $0 < c_v < 1$. We run the simulation many times and obtain the average growth rate. The doubling time of the cell cultures is $\tau^* = \ln 2 / v$. Figure 2b shows the comparison between theory and simulation results for the doubling time. Excellent agreement is observed. For generation time distribution with larger spread (larger $c_v$), the difference between the average generation time $\tau$ and the average doubling time $\tau^*$ is larger than for more narrow distribution of $\tau$ in full accord with theory.

The effect of initial age distribution on the average growth rates is negligible (see the electronic supplementary material, figure S4 of SI). For instance, for distribution I the doubling times for initially synchronized and non-synchronized cell populations are $\tau^* = 19.972 \pm 0.004$ and $\tau^* = 19.934 \pm 0.001$. For wider distribution II, these values are $\tau^* = 19.713 \pm 0.002$ and $\tau^* = 19.695 \pm 0.001$, respectively.

### 3.2. Steady-state age distribution in the growing colony

At long times, the age distribution of cells in the colony should not depend on time, as indicated from the long time ansatz of the von Foerster equation. Indeed, we observe stationary age distributions in our simulations, regardless of the initial age distribution (figure 3). We also observe oscillations in the average age of populations because of sequential cell divisions. However, these oscillations eventually disappear at long times. The rate at which these oscillations disappear depends on $c_v$ of the division time distribution (figure 4). The age distribution at long times in all cases has a maximum at zero age $a = 0$ and decreases monotonically for $a > 0$. For the narrowest distribution (case I) with the smallest $c_v$, the age distribution is closer to the ‘ideal’ age distribution with two distinct regions: exponential decrease of probability density $g(a) \sim \exp(-av)$ for $a < \tau$ followed by much faster decay of $g(a)$ for $a > \tau$. Our simulation results for the steady-state age distributions are in excellent agreement with theoretical ones obtained by equation (2.15).

### 3.3. Population fluctuations in initially synchronized growing clones

Beyond the average clone population and the age distribution, we obtain results for population fluctuations defined as

$$\sigma^2(t) = \left\{ \int_0^\infty n_h(a, t) \, da - N(t) \right\}^2. \quad (3.3)$$

Simulations show that $\sigma(t)$ grows exponentially at long times with the same rates as the average numbers of cells $N(t)$. Thus, relative standard deviations defined as $\sigma(t)/N(t)$ converge to constant values at long times. The results of our simulations for different generation time distributions I–IV in the absence of the cell death are shown in figure 4. Distribution V is the exponential generation time distribution, which corresponds to a constant division probability per unit time (see above). Exact analytic results are available (see the electronic supplementary material, section A). Note that these populations have the same average division time. Thus, we see that the width (or the dispersion) of the generation time distribution has a dramatic effect on the population fluctuation; this dependence should have important consequences in real biological settings.

We will consider the asymptotic value of $\sigma(t)/N(t)$ at $t \to \infty$, which we denote as $\sigma(N)/N$lim. However, if we consider the time dependence of the relative fluctuation, we will keep $t$ in the notation. In general, the asymptotic value is less than 1. From our simulations, it appears that the following simple relation for $\sigma(N)/N$lim, for initially synchronized cell populations is valid:

$$\frac{\sigma_a}{N} = \frac{c_v}{\sqrt{N_0}}, \quad (3.4)$$

c_v being defined earlier by equation (3.2). It is easy to see that equation (3.4) is valid for delta function ($c_v = 0$) and exponential ($c_v = 1$) generation time distributions. However, the initial population must be synchronized to satisfy the simple relation $\sigma(N)/N$lim = $c_v/\sqrt{N_0}$. Note that for the constant cell division rate (exponential distribution), the asymptotic value of standard deviation remains the same regardless of the initial age distribution. This conjecture seems to be reasonable because the more dispersed distribution of division times results in larger variation of the growth rates and eventually the larger relative fluctuations of sizes of cell populations around mean values.
dependent cell division probability per unit time (cases I – IV) but does not matter for constant cell division rate (case V). The coefficients of variation for theoretical zero age 0.0025, 0.005 and 0.0075 (arb. units.)

The asymptotic values of the relative population fluctuations of synchronized cell populations \( \sigma(N)_{\text{lim}} \) at long times for theoretical distributions I – V taken from figure 4a are shown as a function of coefficient of variation of generation times \( c_v \) defined by equation (3.2). The solid line represents the linear fit \( \sigma(N)_{\text{lim}} = b s c_v \) with \( b_s = 0.99 \pm 0.01 \). The condition of synchronized populations at \( t = 0 \) is essential to observe this simple relation between \( \sigma(N)_{\text{lim}} \) and \( c_v \) for age-dependent cell division probability per unit time (cases I – IV) but does not matter for constant cell division rate (case V). The coefficients of variation for theoretical distributions I–V by equation (3.2) are \( c_v = 0.111, 0.222, 0.330, 0.408 \) and 1, respectively.

Figure 4. Population fluctuations in growing clonal colonies with initially synchronized population. (a) The average relative population fluctuations in clones grown from a single cell \( (N_0 = 1) \) with no cell death, \( \sigma(t) \) is defined in equation (3.3). The fluctuations depend on the division time distribution. Curves I – IV are for \( \omega(a) \) in figure 1d. Curve V corresponds to exponential distribution of division times (i.e. constant division rate). The relative fluctuations reach constant values for long times. (b) The asymptotic values of the relative population fluctuations of synchronized cell populations \( \sigma(N)_{\text{lim}} \) at long times for theoretical distributions I – V taken from figure 4a are shown as a function of coefficient of variation of generation times \( c_v \) defined by equation (3.2). The solid line represents the linear fit \( \sigma(N)_{\text{lim}} = b_s c_v \) with \( b_s = 0.99 \pm 0.01 \). The condition of synchronized populations at \( t = 0 \) is essential to observe this simple relation between \( \sigma(N)_{\text{lim}} \) and \( c_v \) for age-dependent cell division probability per unit time (cases I – IV) but does not matter for constant cell division rate (case V). The coefficients of variation for theoretical distributions I–V by equation (3.2) are \( c_v = 0.111, 0.222, 0.330, 0.408 \) and 1, respectively.

Figure 5. Effect of cell death on population fluctuations. (a) A constant cell death rate always increases the relative population fluctuation. The bottom line is the simulation result with no cell death taken from figure 4 (generation time distribution II, see caption to figure 1d). The other curves are for cell death rates \( k_d = 0.0025, 0.005 \) and 0.0075 (arb. units.)\(^{-1} \) respectively. \( \sigma(t) \) and \( N(t) \) are obtained at least for 200 runs for all cases. The populations are grown from a single cell of the zero age \( a = 0 \). The respective growth curves from simulations for some \( k_d \) are presented in electronic supplementary material, figure S5 of SI. (b) Asymptotic values of \( \sigma(N)_{\text{lim}} \) are shown as a function of cell death rate \( k_d \). The symbols are simulation data for age-dependent cell division \( k(a) \) (generation time distribution II) and constant cell death \( k_d \) obtained by analysing at least 200 runs for each case. The black line represents theoretical dependence for both contant rates \( k_b \) and \( k_d \) given by equation (59) in electronic supplementary material, section A.

Note that only the dispersion but not the general shape of \( \omega(a) \) matters. For other shapes such as the combination of two exponential functions proposed by Bremer [14], equation (3.4) is also valid.

3.4. Effect of cell death and initial conditions on population fluctuations

We now examine other aspects of typical cell population that effect population fluctuations. First, we model cell death with a constant death rate \( k_d \) that does not depend on the age of individuals. The results show that cell death has a significant effect on the average population \( N(t) \) and population fluctuations \( \sigma(t) \) (figure 5). The average growth curves for \( k_d > 0 \) are shown in electronic supplementary material, figure S5 of SI. Both \( N(t) \) and \( \sigma(t) \) grow exponentially as \( \sim e^{(\mu - k_d)t} \) with the net growth rate \( \nu = \nu_0 - k_d \), where \( \nu_0 \) is the mean growth rate in absence of cell death. The relative standard deviation \( \sigma(t)/N(t) \) converges to an asymptotic value that is larger than in absence of cell death. Figure 5a shows that the relative fluctuations grow with increasing cell death rate \( k_d \).

Apparently, a similar result is seen in models with constant (age-independent) cell division \( k(a) = k_b \) and cell death \( k_d \) rates. In this case, \( \sigma(t)/N(t) \) actually diverge when \( k_d \) approaches \( k_b \), i.e. equation (2.1) (see electronic supplementary material, section A). For \( k_d < k_b \), we have

\[
\left( \frac{\sigma(N)_{\text{lim}}}{N} \right) \approx \frac{1}{\sqrt{N_0}} \left( 1 + \frac{k_d}{k_b} \right).
\]  

(3.5)

However, in the age-dependent cell division case, the relative fluctuation does not agree with equation (3.5). Here, analytical results are not available, and the quantitative dependence is an open question. We compare the relative population fluctuations as a function of constant cell death rate \( k_d \) for age-dependent model with age-independent model in figure 5b. A significant difference between these models is seen. The effect of cell death on asymptotic relative fluctuations is stronger for the age-dependent model (cf. the initial slopes of dependences shown in figure 5b). The dependence is not linear for small death rates. Since cell division processes are generally age-dependent, this suggests that a more sophisticated understanding of population fluctuations is needed.
The red arrows show increasing the asymptotic values of relative dispersions material, figures S6 and S7 of SI). For asynchronous population fluctuation. However, the dependence is trivial since this model is well known as the Malthus–Verhulst problem.

For asynchronous growth the initial age of a cell is chosen randomly from 0 to \( \tau \). The red arrows show increasing the asymptotic values of relative dispersions with synchronized growth compared with synchronized condition. The effect of initial age distribution on \( \langle \sigma/N \rangle_{hm} \) is greater for less dispersed generation time distribution. In figure 6, we compare the relative population fluctuation at steady-state scales as \( \sigma(t)/N(t) \lim \) is larger for generation times with the smaller spread. Again, for constant division probability, the effect of initial age distribution on \( \sigma(t) \) and \( \sigma(t)/N(t) \) is absent.

In this paper, we consider the relationship between cell division time variability and the fluctuations in cell population using a stochastic modelling approach. We examine the simple case of symmetrical cell division, and explicitly consider the time-dependent probability of cell division. The average population in this model exactly coincides with the deterministic

\[
\frac{dN}{dt} = (k_b - k_d)N - \gamma_0 N^2, \tag{3.6}
\]

which has a stable steady-state population \( N_{ss} \):

\[
N_{ss} = \frac{k_b - k_d}{\gamma_0}. \tag{3.7}
\]

In our model, we consider the age-dependent cell division rate \( k(a) \) and all other rates are taken the same as in the Malthus–Verhulst problem, i.e. independent of the age of cells. We analyse the results for the growth of a competitive population from a single cell of the zero age \( a = 0 \) from our simulations. The ‘natural’ cell death rate is taken to be zero for simplicity \( k_d = 0 \), while \( \gamma_0 > 0 \). Indeed, we observe the steady state at long times (figure 7a) corresponding to a population of the macroscopic size equal to

\[
N_{ss} = \frac{v}{\gamma_0}, \tag{3.8}
\]

where \( v \) is the growth rate of a population in the absence of any cell death events. Recall that this growth rate appears as the solution of the long time ansatz of the von Foerster equation, see equation (2.12). Thus, for cells with the same average division time, but different shapes \( \omega(a) \), we expect a different steady-state population.

At long times, fluctuations around the mean population approach a steady value as in the Malthus–Verhulst problem. Interestingly, the standard deviation \( \sigma(t) \) of the mean is a non-monotonic function of time (see figure 7b). At early times when the cell number is small and death events are rare, \( \sigma(t) \) grows exponentially. In the crossover regime, \( \sigma(t) \) goes through the maximum and at last reaches its steady-state value. The steady-state value of \( \sigma(t) \), however, does depend on \( \gamma_0 \) and whether the model considers cell age. Figure 7d shows the comparison between the age-independent model (see the electronic supplementary material, section C) and the age-dependent model. The population fluctuation at steady-state scales as \( \sigma_{ids} \) for the age-dependent model and \( \sigma_{ids} = 0.5 \) (see the electronic supplementary material, section C). In addition, the behaviours of \( \sigma(t) \) before reaching steady state are also quite different (figure 7d).

In spite of the fact that the average cell death rate depends on time, the age distribution in this case is the same as in absence of cell death. We have analysed age distributions for the case of growth of competitive population at two times. One \( (t = 252 \text{ arb. units}) \) is chosen around the maximum of \( \sigma(t) \), and the other time \( (t = 450 \text{ arb. units}) \) is within the steady-state regime. For these times, age distributions turned out to be the same (the results are not shown) and identical to steady-state age distribution in absence of cell death (figure 3). Since the cell death rate is age-independent, cells of different ages have the same probability to die, so the age distribution remains intact.

3.5. Growth of competitive populations

In organized and biologically realistic cell colonies, cell division and cell death are carefully controlled. The control strategy has a strong effect on the total cellular population as well as population fluctuations. There are many possible models to describe these situations. Here, we examine a competitive population model where in addition to natural cell division \( k(a) \) and cell death rates \( k_d \) the competition gives rise to an additional death rate \( \gamma = \gamma_0 (t - 1) \) proportional to the number of other cells \( N \) in the colony at time \( t \). For the age-independent case, \( k(a) = k_a \) and constant \( k_d \) and \( \gamma_0 \) this model is well known as the Malthus–Verhulst problem.

\[
\frac{dN}{dt} = (k_b - k_d)N - \gamma(t)N^2, \tag{3.9}
\]

The equation for the average number of cells is

\[
\frac{dN}{dt} = (k_b - k_d)N - \gamma_0 N^2, \tag{3.10}
\]

In our model, we consider the age-dependent cell division rate \( k(a) \) and all other rates are taken the same as in the Malthus–Verhulst problem, i.e. independent of the age of cells. We analyse the results for the growth of a competitive population from a single cell of the zero age \( a = 0 \) from our simulations. The ‘natural’ cell death rate is taken to be zero for simplicity \( k_d = 0 \), while \( \gamma_0 > 0 \). Indeed, we observe the steady state at long times (figure 7a) corresponding to a population of the macroscopic size equal to

\[
N_{ss} = \frac{v}{\gamma_0}, \tag{3.8}
\]

where \( v \) is the growth rate of a population in the absence of any cell death events. Recall that this growth rate appears as the solution of the long time ansatz of the von Foerster equation, see equation (2.12). Thus, for cells with the same average division time, but different shapes \( \omega(a) \), we expect a different steady-state population.

At long times, fluctuations around the mean population approach a steady value as in the Malthus–Verhulst problem. Interestingly, the standard deviation \( \sigma(t) \) of the mean is a non-monotonic function of time (see figure 7b). At early times when the cell number is small and death events are rare, \( \sigma(t) \) grows exponentially. In the crossover regime, \( \sigma(t) \) goes through the maximum and at last reaches its steady-state value. The steady-state value of \( \sigma(t) \), however, does depend on \( \gamma_0 \) and whether the model considers cell age. Figure 7d shows the comparison between the age-independent model (see the electronic supplementary material, section C) and the age-dependent model. The population fluctuation at steady-state scales as \( \sigma_{ids} \) for the age-dependent model and \( \sigma_{ids} = 0.5 \) (see the electronic supplementary material, section C). In addition, the behaviours of \( \sigma(t) \) before reaching steady state are also quite different (figure 7d).

In spite of the fact that the average cell death rate depends on time, the age distribution in this case is the same as in absence of cell death. We have analysed age distributions for the case of growth of competitive population at two times. One \( (t = 252 \text{ arb. units}) \) is chosen around the maximum of \( \sigma(t) \), and the other time \( (t = 450 \text{ arb. units}) \) is within the steady-state regime. For these times, age distributions turned out to be the same (the results are not shown) and identical to steady-state age distribution in absence of cell death (figure 3). Since the cell death rate is age-independent, cells of different ages have the same probability to die, so the age distribution remains intact.

4. Discussion

In this paper, we consider the relationship between cell division time variability and the fluctuations in cell population using a stochastic modelling approach. We examine the simple case of symmetrical cell division, and explicitly consider the time-dependent probability of cell division. The average population in this model exactly coincides with the deterministic

\[
\frac{dN}{dt} = (k_b - k_d)N - \gamma_0 N^2, \tag{3.6}
\]

which has a stable steady-state population \( N_{ss} \):

\[
N_{ss} = \frac{k_b - k_d}{\gamma_0}. \tag{3.7}
\]

In our model, we consider the age-dependent cell division rate \( k(a) \) and all other rates are taken the same as in the Malthus–Verhulst problem, i.e. independent of the age of cells. We analyse the results for the growth of a competitive population from a single cell of the zero age \( a = 0 \) from our simulations. The ‘natural’ cell death rate is taken to be zero for simplicity \( k_d = 0 \), while \( \gamma_0 > 0 \). Indeed, we observe the steady state at long times (figure 7a) corresponding to a population of the macroscopic size equal to

\[
N_{ss} = \frac{v}{\gamma_0}, \tag{3.8}
\]

where \( v \) is the growth rate of a population in the absence of any cell death events. Recall that this growth rate appears as the solution of the long time ansatz of the von Foerster equation, see equation (2.12). Thus, for cells with the same average division time, but different shapes \( \omega(a) \), we expect a different steady-state population.

At long times, fluctuations around the mean population approach a steady value as in the Malthus–Verhulst problem. Interestingly, the standard deviation \( \sigma(t) \) of the mean is a non-monotonic function of time (see figure 7b). At early times when the cell number is small and death events are rare, \( \sigma(t) \) grows exponentially. In the crossover regime, \( \sigma(t) \) goes through the maximum and at last reaches its steady-state value. The steady-state value of \( \sigma(t) \), however, does depend on \( \gamma_0 \) and whether the model considers cell age. Figure 7d shows the comparison between the age-independent model (see the electronic supplementary material, section C) and the age-dependent model. The population fluctuation at steady-state scales as \( \sigma_{ids} \) for the age-dependent model and \( \sigma_{ids} = 0.5 \) (see the electronic supplementary material, section C). In addition, the behaviours of \( \sigma(t) \) before reaching steady state are also quite different (figure 7d).

In spite of the fact that the average cell death rate depends on time, the age distribution in this case is the same as in absence of cell death. We have analysed age distributions for the case of growth of competitive population at two times. One \( (t = 252 \text{ arb. units}) \) is chosen around the maximum of \( \sigma(t) \), and the other time \( (t = 450 \text{ arb. units}) \) is within the steady-state regime. For these times, age distributions turned out to be the same (the results are not shown) and identical to steady-state age distribution in absence of cell death (figure 3). Since the cell death rate is age-independent, cells of different ages have the same probability to die, so the age distribution remains intact.

4. Discussion

In this paper, we consider the relationship between cell division time variability and the fluctuations in cell population using a stochastic modelling approach. We examine the simple case of symmetrical cell division, and explicitly consider the time-dependent probability of cell division. The average population in this model exactly coincides with the deterministic
von Foerster equation for age-dependent cell populations. Indeed, our simulations confirm this, and accurately reproduce the growth curves of cell populations. As expected, the width of the division time distribution influences the overall cell colony growth rate, even when the average cell division time is identical. This is in contrast with models with constant cell division and death rates, which does not reproduce the correct population growth.

The main focus of our study, unlike the deterministic von Foerster model, is to delineate the underlying cause of cell population fluctuations. We find that the division time distribution, \( \alpha(t) \), has a strong influence on the population-level fluctuation, even for long times. For initially synchronized cells, we find a simple relation between the asymptotic values of relative fluctuations \( \sigma /N \) and the width of the division time distributions \( \alpha(t) \). We also find that the initial age distribution affects the long time population fluctuation. Asynchronous population has a larger population fluctuation. We also incorporate cell death in the model and compute the population fluctuation as a function of an age-independent cell division rate, \( k_d \). Results show that cell death generally increases cell population fluctuation, but the scaling with respect to \( k_d \) does not follow equation (2.1). Finally, since in most cell colonies both cell division and death are affected by the overall population, we study a simple model of a competitive growth where the death rate increases with the overall population. Once again, the age-dependent model predicts a different fluctuation behaviour. We expect that in general, age-dependent cell growth models will have qualitatively different results than the constant birth–death models. This fundamental result suggests a revision of the usual theoretical approach to model growth and evolution of cell populations.

To predict time evolution of cell populations in biologically realistic situations, it is important to obtain good knowledge of the cell division and death rates as functions of cell age. In our studies, we have incorporated realistic cell division rates but assumed that the death rate is a constant. Note that our approach is Markovian, i.e. there is no memory between successive division events. This assumption is justified by our data and others [1] which showed a lack of correlation between successive divisions. However, cell division and death rates are regulated by other factors such as population density and nutrient levels. Cells may also transform into a different state where cell division is inhibited. These factors suggest that rich and interesting population dynamics exist, even in the simplest colonies. For the living tissue, it is known that several cell types must coexist, and transformations from stem or progenitor cells will influence the final population. Our work here suggests that an age-dependent approach may be appropriate to understand those situations as well.

Equation (3.4) suggests that the distributions of cell colony sizes broaden with increasing spread of the generation time distribution. This result is for exponentially growing colonies, but similar effects may be present for homeostatic

**Figure 7.** Competitive growth. (a) The average numbers of cells as functions of time for competitive populations with age-dependent cell division rate \( k(a) \) (distribution II, see caption to figure 1d) and death rates proportional to the number of other individuals \( \gamma = \gamma_0(a - 1) \). The populations are grown from a single cell of the zero age \( a = 0 \). The numbers of cells reach the steady state \( N_{ss} \) at long times. The data for each \( \gamma_0 \) are obtained from simulations for 120 growth trajectories. (b) The standard deviations \( \sigma(t) \) of the mean sizes of competitive populations from data in (a) are non-monotonic functions of time. At long times, \( \sigma(t) \) approach the limiting values \( \sigma_{ss} \) in contrast to exponentially growing populations. The population size distributions at steady state are stationary Gaussian around mean sizes \( N_{ss} \) (see the electronic supplementary material, figure S9 of SI). (c) The average numbers of cell \( N_{ss} \) (circles) and its standard deviations \( \sigma_{ss} \) (diamonds) at steady state are shown as a function of ‘competitive’ cell death rate \( \gamma_0 \). The data are analysed in the time interval from \( t = 400 \) arb. units to \( t = 474 \) arb. units for all values of \( \gamma_0 \). The mean population is inversely proportional to \( \gamma_0 \). (d) Comparison of the population fluctuation for age-dependent (blue) and age-independent (constant division rate \( k_d \) (red) models. \( \gamma_0 \) for these models are identical. Parameters are chosen so that the average growth rates in the absence of cell death are identical. The age-dependent model shows a significantly different behaviour.
populations. During organ growth, for example, the final size of the organ and the variation in organ size may be related to cell population fluctuations. This fluctuation may be influenced by the age-dependence of cell events. Age-dependent dynamics are also expected to be important in colonies seeded by a few cells where population fluctuations may be significant, e.g. tumour growth, drug-resistant cell populations. Finally, it is important to note that by measuring not just the average population, but also the population fluctuation, important information can be obtained about the underlying growth and control mechanism. We have seen that the population fluctuation can be quite different, even when the average population growth rate is the same. Indeed, the full theory must consider all moments of the population distribution. Therefore, statistical fluctuations in population dynamics contain vital information about the behavioral mechanisms of growth and evolution.

This work was supported by NIH 1U54CA143868-01, 1R01GM075305 and NSF PHY-1205795.

References


\(N_0\) initial number of cells
\(N(t)\) average number of cells at time \(t\)
\(N_{\text{lost}}/N_{\text{alive}}\) ratio of numbers of lost to live cells
\(\tau^*\) (arb. units) average doubling time of cell population

\(\nu\) (arb. units\(^{-1}\)) average growth rate of cell population
\(g(a)\) (arb. units\(^{-1}\)) PDF of age distribution at steady state
\(\sigma(t)\) standard deviation of the mean \(N(t)\)
\((\sigma/N)_{\text{lim}}\) asymptotic value of \(\sigma(t)/N(t)\) at long \(t\)

\(\gamma_0\) (arb. units\(^{-1}\)) cell death rate for competitive population
\(N_{\text{ss}}\) steady-state size of competitive population
\(\sigma_{\text{ss}}\) standard deviation of \(N(t)\) at steady state

\(p(n; t)\) (dimensionless value) PDF of population size distribution at time \(t\)
\(p(t_n)\) (arb. units\(^{-1}\)) PDF of distribution of times \(t_n\) to reach \(n\) cells