

UC Irvine

UC Irvine Previously Published Works

Title

Analysis of stochastic stem cell models with control

Permalink

<https://escholarship.org/uc/item/8r46p2pf>

Journal

Mathematical Biosciences, 266

ISSN

0025-5564

Authors

Yang, Jienian

Sun, Zheng

Komarova, Natalia L

Publication Date

2015-08-01

DOI

10.1016/j.mbs.2015.06.001

Peer reviewed



Published in final edited form as:

Math Biosci. 2015 August ; 266: 93–107. doi:10.1016/j.mbs.2015.06.001.

Analysis of stochastic stem cell models with control

Jienian Yang, Zheng Sun, and Natalia L. Komarova*

Department of Mathematics, University of California Irvine, Irvine, CA 92617

Abstract

Understanding the dynamics of stem cell lineages is of central importance both for healthy and cancerous tissues. We study stochastic population dynamics of stem cells and differentiated cells, where cell decisions, such as proliferation vs differentiation decisions, or division and death decisions, are under regulation from surrounding cells. The goal is to understand how different types of control mechanisms affect the means and variances of cell numbers. We use the assumption of weak dependencies of the regulatory functions (the controls) on the cell populations near the equilibrium to formulate moment equations. We then study three different methods of closure, showing that they all lead to the same results for the highest order terms in the expressions for the moments. We derive simple explicit expressions for the means and the variances of stem cell and differentiated cell numbers. It turns out that the variance is expressed as an algebraic function of partial derivatives of the controls with respect to the population sizes at the equilibrium. We demonstrate that these findings are consistent with the results previously obtained in the context of particular systems, and also present two novel examples with negative and positive control of division and differentiation decisions. This methodology is formulated without any specific assumptions on the functional form of the controls, and thus can be used for any biological system.

1 Introduction

Tissue turnover dynamics, especially in the context of stem cell regulation, have attracted the attention of many researchers. Cell populations are assumed to possess a hierarchical structure, where different classes of cells can interact in intricate ways. In the simplest case, there are stem cells capable of self-renewing and regenerating the tissue, and differentiated cells which can perform the tissue's specific functions.

Differentiated cells are subject to relatively frequent cell death and need to be replenished by stem cell divisions. These divisions can be of several types. Specifically, a stem cell can differentiate by dividing into two differentiated cells, or it can proliferate, by dividing into two stem cells. Differentiation/proliferation decisions are thought to be under regulation coming from surrounding cells in the tissue. Various control loops help maintain a roughly

*Corresponding author: komarova@uci.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

constant overall tissue size, and keep variations in the numbers of stem and differentiated cells to a minimum.

There is significant theoretical literature exploring various aspects of stem cell dynamics. Conceptual theoretical issues for the studies of stem cells have been identified in [1, 2, 3]. Discrete and continuous models relevant for carcinogenesis have been studied [4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15]. Evolutionary modeling of stem cells in systems other than cancer was introduced in [16]. Modeling of stem cells in the hematopoietic system was proposed by several authors [17, 18, 19, 20, 21]. In these and other papers, both deterministic and stochastic models have been introduced and studied (see a great review of many modeling approaches provided in [22]). The deterministic (ODE) approach provides useful analytical insights into the dynamics and long-term behavior of cell lineages. Two- and multi-compartment models with several types of the regulation function have been studied in [23, 24], where the authors discuss important conceptual issues about stem cell regulation from the engineering perspective. A systematic linear stability analysis of two- and three-compartment models with regulation of self-renewal fractions or regulation of proliferation rates was performed in [25]. Another type of regulation was studied in two-compartment models by [26]. Analysis of the structure of stationary solutions in the n -compartment version of the model was presented in [27].

The stochastic approach allows to quantify the role of fluctuations in the behavior of the system of interest [28, 29, 30, 31]. Apart from several exceptions [32, 33], most of the literature is devoted to numerical explorations of stochastic stem cell systems. Recently, we performed analytical studies of two stochastic stem cell systems involving nonlinear control [34, 35] and found how the strength of control determines the amount of stochastic fluctuations in the numbers of stem and differentiated cells. This was done for several particular types of control functions. Unfortunately, the methods used in those papers cannot be extended to study other types of control loops.

In the present paper we develop a general, analytical methodology for studying the behavior of hierarchical, two-compartment (stem and differentiated cells) systems with nonlinear control. We assume that division, death, and differentiation/proliferation decisions are given by some (unspecified) functions of the numbers of stem and differentiated cells, and provide tools to calculate the moments of the cell numbers, and importantly, the means and the variances of the numbers of cells.

It turns out that under some general assumptions, the amount of variation in the system is a function of the local behavior of the control functions near the equilibrium. For example, in the simplest case of the constant total population systems, the variance in the number of stem cells is inversely proportional to the derivative of the control function with respect to the number of stem cells, evaluated at the equilibrium. For non-constant populations, we develop similarly general methods and provide explicit formulas approximating cell number means and variances.

The method developed here is algorithmically different, and simpler, than the linear noise approximation [36]. We studied the connection between the two methods and proved that

they give the same result to all orders of accuracy. Therefore, our method could be considered a short-cut compared with the Van Kampen power series expansion. We developed a computer program (written for *Mathematica* and presented in a supplement) which allows to apply our method to any system of stem and differentiated cells with given control functions. In other words, if we assign the rates of divisions, differentiation/proliferation, and death to be some functions of the numbers of stem and differentiated cells, our tools allow to calculate analytically the means and the variances of the stem and differentiated cell numbers as functions of the system parameters, and to study stability and robustness of the system.

The rest of this paper is organized as follows. In Section 2 we discuss systems with constant total populations, where only differentiation/proliferation decisions are under nonlinear regulation. In Section 3 we generalize this methodology to non-constant populations, where three types of processes (divisions, deaths, and differentiation/proliferation decisions) are under nonlinear regulation. In section 4, the results are illustrated by using previously solved regulation problems as well as two novel examples. In the first example, both division and differentiation decisions are under negative control from the population sizes. In the second example, divisions are negatively regulated while differentiation decisions are under a positive control loop. Section 5 compares and contrasts our new method with the power series expansion method of Van Kampen. Discussion is provided in section 6.

2 Modeling constant total cell populations

In the first set of models we will assume that the population consists of I stem cells and J differentiated cells, and that the total population size remains constant, $I + J = N$. This corresponds to a generalization of the well-known Moran process [37] in the presence of two sub-populations of different properties. In the classical Moran process, each update consists of a division event followed by a death event. All cells have an equal probability to die, and any cell has a chance to divide. A division event is a replacement of the dividing parent cell with two cells, which in the absence of mutations are both identical to the parent cell.

In the processes considered here, only differentiated cells die (with equal probabilities), and only stem cells divide (also with equal probabilities), see figure 1. Moreover, there are two types of stem cell divisions. A *proliferation event* results in two daughter cells which are both stem cells. A *differentiation event* leads to the creation of two differentiated cells. The probability of differentiation, p , is assumed to be under some regulatory loops from the stem and/or differentiated cell populations. Since $J = N - I$, we can simply say that $p = p_I$ a function of the number of stem cells, I .

The above model gives rise to a 1D Markov process with $\text{Prob}(I \rightarrow I - 1) = p_I$ and $\text{Prob}(I \rightarrow I + 1) = 1 - p_I$. Denoting by $\varphi_I(t)$ the probability to find the system at state I at time t , we can write down the following Kolmogorov forward equation:

$$\dot{\varphi}_I = \varphi_{I-1} (1 - p_{I-1}) + \varphi_{I+1} p_{I+1} - \varphi_I. \quad (1)$$

Depending on the functional form of the differentiation probability p_I , the system can exhibit different types of behavior, from oscillating around an equilibrium, to an unstable behavior resulting in extinction/overflow.

2.1 Previous results for specific cases

In [34], several types of the differentiation probability p_I have been studied.

- **No control.** It was shown that for $p_I = p = \text{const}$, the system rapidly drifts to one of the two extinction states: either the $I = 0$ state with no stem cells, or the $I = N$ state with no differentiated cells. This case corresponds to the absence of stem cell regulation.
- **A hyperbolic law.** In this case, we assume the following functional dependence:

$$p_I = \frac{\beta}{1 + hI}, \quad (2)$$

where β and h are parameters. The magnitude of h defines the degree of control, and the case $h = 0$ corresponds to the constant probability model. We obtained the following results for the mean and the variance of the stem cell numbers in this case:

$$E[I] = \frac{1 - 2\beta}{h} - \frac{1}{2}, \quad \text{Var}[I] = \frac{\beta}{h} + \frac{1}{4}. \quad (3)$$

- **A Hill-type law.** Consider the following functional form of differentiation probability:

$$p_I = \frac{I^\alpha}{k^\alpha + I^\alpha}, \quad (4)$$

with $0 < k < N$ and $\alpha \geq 0$. Here, $\alpha = 0$ is the constant- p model, and $\alpha \rightarrow \infty$ corresponds to the Heaviside function. If we assume that $k \gg 1$, then the following approximations for the mean and the variance of the stem cell number have been obtained:

$$E[I] = k + \frac{1}{2\alpha} + O(1/k), \quad \text{Var}[I] = \frac{k}{\alpha} + \frac{2\alpha - 1}{4\alpha^2} + O(1/k). \quad (5)$$

The methodology used to obtain the above formulas relies on formulating the moment equations from the Kolmogorov forward equation, and decoupling them by means of a truncation procedure (a cumulant and a central moment closure method). Because probability functions p_I in the hyperbolic and the Hill-type laws are not polynomials of I , obtaining the equations for the moments is a complicated procedure, which requires using some auxiliary probability functions, and cannot be generalized to different functional forms

of p_I . Thus the analytical results of [34] have very limited applicability. We would like to design a method that would allow to calculate means and variances for the stem cell populations under a general, nonlinear control function. This method is presented below.

2.2 The linearization method

Suppose that we have a 1D Markov process with $Prob(I \rightarrow I-1) = p_I$ and $Prob(I \rightarrow I+1) = 1 - p_I$, where p_I is a nondecreasing function of I . The mean value for I can be obtained by simply solving the equation

$$p_I = \frac{1}{2}$$

for the value I . Let us denote this value of I as i_0 : $p_{i_0} = 1/2$.

The variance is obtained in the following way. Let us denote by $i = I - i_0$, the difference between the current number of stem cells and the equilibrium number, i_0 . The variable i satisfies $|i| \ll i_0$, and can take positive or negative values. Let us assume that the dependence of the differentiation probability p_I on its variable i is slow. In the examples above this would mean that

$$\beta/h \gg 1, \quad k \gg 1.$$

Then we can expand the function p_i in terms of small i/i_0 to obtain

$$p_i \approx \frac{1}{2} + p_x i,$$

where p_x is the derivative of p_i with respect to i evaluated at $i = 0$. For the hyperbolic law described above (equation (2)), we have $i_0 = (1 - 2\beta)/h$ and

$$p_x = \frac{h}{4\beta}.$$

For the Hill-type law, equation (4), we obtain $i_0 = k$ and

$$p_x = \frac{\alpha}{4k}.$$

We have the following Kolmogorov forward equation for the probability function $\varphi_i(t)$:

$$\dot{\varphi}_i = \varphi_{i-1} (1 - p_{i-1}) + \varphi_{i+1} p_{i+1} - \varphi_i, \quad (6)$$

or

$$\dot{\varphi}_i = \varphi_{i-1} (1/2 - p_x (i-1)) + \varphi_{i+1} (1/2 + p_x (i+1)) - \varphi_i.$$

Let us use the following notation for the moments:

$$\sum_i \varphi_i i^j = E [i^j] = x_j.$$

Forming the equations for the first two moments, we obtain in steady state,

$$\frac{d}{dt} x_1 = 2p_x x_1 = 0, \quad (7)$$

$$\frac{d}{dt} x_2 = 1 - 4p_x x_2 = 0, \quad (8)$$

which corresponds to

$$E [i] = 0, \quad E [i^2] = Var [i] = \frac{1}{4p_x}. \quad (9)$$

This reproduces the first approximation to the mean and variance values.

To improve the precision of the method, let us consider higher terms in the expansion of the function p_i in terms of i . Let us present the expansion of p_i around the steady state as

$$p_i = \frac{1}{2} + \sum_{k=1}^N \frac{1}{k!} \frac{d^k p_i}{di^k} i^k,$$

where the derivatives are evaluated at $i = 0$ and the cut-off N is defined by our desired precision. In the master equation, equation (6), we can use this expansion to replace the function p_i and obtain equations for the unknown moments x_j by multiplying equation (6) by i^j and summing over i . We obtain in the j th order,

$$\sum_{k=1}^{[j/2]} \binom{j}{2k} x_{j-2k} - 2 \sum_{m=1}^N \frac{1}{m!} \frac{d^m p_i}{di^m} \sum_{k=1}^{[(j+1)/2]} \binom{j}{2k-1} x_{j+1-2k+m} - x_j = 0, \quad (10)$$

where $[.]$ stands for the integer part. For a particular choice of truncation N , equations (10) do not comprise a closed system, and an approximate method must be used to decouple the first N equations from the rest. There are at least 3 different truncation methods that can be readily implemented to solve this system.

- **Simple truncation method** assumes that all the higher moments are 0, that is, $x_k = 0$, for $k > N$, where N is defined by the desired precision.
- **Central moment truncation method** assumes that the central moments are equal to 0 for higher orders, that is, $E[(i - E[i])^k] = 0$, where $k > N$.
- **Cumulant truncation method** assumes that the higher order cumulants are equal to 0, that is, $\kappa_k = 0$, where $k > N$.

Below we demonstrate these methods by taking into account the second order terms in the Taylor expansion. We will use the notation

$$p_{xx} = \frac{d^2 p_i}{di^2},$$

evaluated at $i = 0$. We start by writing the Komogorov forward equation with the probability functions expanded up to the second order terms:

$$\dot{\varphi}_i = -\varphi_i + \varphi_{i-1} \left(\frac{1}{2} - p_x(i-1) - \frac{p_{xx}}{2}(i-1)^2 \right) + \varphi_{i+1} \left(\frac{1}{2} + p_x(i+1) + \frac{p_{xx}}{2}(i+1)^2 \right).$$

Multiplying this equation by i and i^2 on both sides, we obtain in steady state:

$$p_x x_1 + \frac{p_{xx}}{2} x_2 = 0, \quad (12)$$

$$p_x x_2 + \frac{p_{xx}}{2} x_3 = \frac{1}{4}. \quad (13)$$

In the simple truncation method, we simply assume that $x_3 = 0$. In the central moment and cumulant truncation methods, we assume that $x_3 - 3x_1x_2 + 2x_1^3 = 0$. Therefore, x_i , $i = 1, 2, 3$ can be found by solving system (12–13). The results for the mean and variance are as follows:

The hyperbolic model—In this case, we have $p_i = \frac{\beta}{1 - h(i_0 + i)}$, which corresponds to the expansion coefficients

$$p_x = \frac{h}{4\beta}, \quad p_{xx} = \frac{h^2}{4\beta^2}.$$

The simple truncation method yields:

$$x_1 = -\frac{1}{2}, \quad x_2 = \frac{\beta}{h}.$$

To express the mean and the variance of the variable i in terms of these quantities, we note that

$$E[I] = i_0 + x_1,$$

and that

$$\text{Var}[I] = E[(I - E[I])^2] = E[(I - i_0 - x_1)^2] = E[(i - x_1)^2] = x_2 - x_1^2.$$

This yields

$$E[I] = \frac{1 - 2\beta}{h} - \frac{1}{2}, \quad \text{Var}[I] = \frac{\beta}{h} - \frac{1}{4}.$$

The central moment and cumulant truncation methods give the following solution:

$$x_1 = -\frac{1}{2}, \quad x_2 = \frac{\beta}{h} + \frac{3}{4},$$

which corresponds to

$$E[I] = \frac{1 - 2\beta}{h} - \frac{1}{2}, \quad \text{Var}[I] = \frac{\beta}{h} + \frac{1}{2}.$$

The Hill-type model—In this case, we have $p_i = \frac{i^\alpha}{k^\alpha + (i_0 + i)^\alpha}$, and the expansion coefficients are:

$$p_x = \frac{\alpha}{4k}, \quad p_{xx} = \frac{\alpha}{2k^2}.$$

The simple truncation method gives:

$$x_1 = \frac{1}{\alpha}, \quad x_2 = \frac{k}{\alpha}.$$

The results for the mean and the variance are:

$$E[I] = k + \frac{1}{\alpha}, \quad \text{Var}[I] = \frac{k}{\alpha} - \frac{1}{\alpha^2}.$$

The central moment and cumulant truncation methods yield:

$$x_1 = \frac{1}{\alpha}, \quad x_2 = \frac{k}{\alpha} + \frac{3}{\alpha^2},$$

and the formulas for the mean and the variance are

$$E[I] = k + \frac{1}{\alpha}, \quad \text{Var}[I] = \frac{k}{\alpha} + \frac{1}{\alpha^2}.$$

3 General approach: modeling non-constant total cell populations

In a more general setting, the total number of cells in the system is not a constant number. The number of stem cells, I , and the number of differentiated cells, J , vary independently, giving rise to a 2D Markov process. Let us suppose that in an infinitesimal time-interval, Δt , the following events can occur:

- With probability $L_{I,J}\Delta t$ a stem cell divides. Two types of division are possible.
 - With probability $L_{I,J}P_{I,J}\Delta t$ a stem cell differentiation takes place resulting in a creation of two differentiated cells, $(I, J) \rightarrow (I-1, J+2)$.
 - With probability $L_{I,J}(1 - P_{I,J})\Delta t$ a stem cell proliferation takes place resulting in a creation of a stem cell, $(I, J) \rightarrow (I+1, J)$.
- With probability $D_{I,J}\Delta t$, a differentiated cell dies, $(I, J) \rightarrow (I, J-1)$.

All other events are assumed to happen with zero probability. The processes described above are illustrated schematically in figure 2. Let us denote by $\varphi_I(t)$ the probability to have I stem cells and J differentiated cells at time t . The Kolmogorov forward equation corresponding to the above processes is given by:

$$\dot{\varphi}_{I,J} = \varphi_{I,J+1}D_{I,J+1} + \varphi_{I-1,J}L_{I-1,J}(1 - P_{I-1,J}) + \varphi_{I+1,J-2}L_{I+1,J-2}P_{I+1,J-2} - \varphi_{I,J}(L_{I,J} + D_{I,J}). \quad (14)$$

3.1 Previous results for a specific case

A specific form of this process was studied in [35], where we assumed

$$L_{I,J} = \frac{b}{1+h(I+J)}, \quad P_{I,J} = \frac{r}{1+gJ}, \quad D_{I,J} = 1 - L_{I,J}, \quad (15)$$

and parameters h and g are small. In this case, the probabilities of divisions and deaths are functions of the variable $N = I + J$, and the probability of differentiation is a function of the number of differentiated cells, J , only. The resulting means and variances are listed below:

$$E[N] = \frac{1}{2} + \frac{2b-1}{h}, \quad (16)$$

$$Var[N] = \frac{1}{4} + \frac{b}{h}, \quad (17)$$

$$E[J] = \frac{2bg - 8h + 7gh + 20hr - \sqrt{4b^2g^2 + h^2(g-4r)^2 + 4bgh(4r-5g)}}{8gh},$$

$$Var[J] = \frac{23}{32} - \frac{b^2}{8h^2} - \frac{b}{8h} - \frac{r}{4g} + \frac{br}{2gh} + \frac{3r^2}{2g^2} + \frac{2gb + 7gh - 12hr}{32g^2h^2} \sqrt{4b^2g^2 + h^2(g-4r)^2 + 4bgh(4r-5g)}.$$

It is easier to interpret the results for the differentiated cells if we consider the behavior in the limit of weak control, that is when $h \rightarrow 0$ or $g \rightarrow 0$. We have

$$\text{If } h \rightarrow 0: E[J] = \frac{2r-1}{g} + 1 + O(h), \quad Var[J] = \frac{b}{h} + \frac{r}{g} + \frac{1}{2} + O(h), \quad (18)$$

$$\text{If } g \rightarrow 0: E[J] = \frac{2r-1}{g} + \frac{3}{2} + O(g), \quad Var[J] = \frac{3r}{g} - \frac{3}{4} + O(g). \quad (19)$$

In the following we will develop a general method of calculating the means and the variances of non-constant population systems. It will be demonstrated how the above results can be obtained in a way much simpler than that of [35].

3.2 The general 2D model of control

Let us define the steady state of the system, (i_0, j_0) , by the the following equations:

$$L_{i_0, j_0} = D_{i_0, j_0} \equiv L_0, \quad P_{i_0, j_0} = \frac{1}{2}. \quad (20)$$

Similar to the previous section, we will use lower-case letters to measure the difference between the current cell numbers and their equilibrium numbers: $i = I - i_0$, $j = J - j_0$.

Define $\tilde{\varphi}$ such that $\tilde{\varphi}_{i,j} = \varphi_{I,j}$, and $\tilde{Z}_{i,j} = Z_{I,j}$, where $Z_{I,j}$ denotes any of the functions $L_{I,j}$, $P_{I,j}$, $D_{I,j}$. Then equation (14) can be expressed as:

$$\dot{\tilde{\varphi}}_{i,j} = \tilde{\varphi}_{i,j+1} \tilde{D}_{i,j+1} + \tilde{\varphi}_{i-1,j} \tilde{L}_{i-1,j} \left(1 - \tilde{P}_{i-1,j}\right) + \tilde{\varphi}_{i+1,j-2} \tilde{L}_{i+1,j-2} \tilde{P}_{i+1,j-2} - \tilde{\varphi}_{i,j} \left(\tilde{L}_{i,j} + \tilde{D}_{i,j}\right). \quad (21)$$

Let us use the following short-hand notation for the moments:

$$x_{\alpha\beta} \equiv E \left[i^\alpha j^\beta \right] := \sum_{i,j} \tilde{\varphi}_{i,j} i^\alpha j^\beta. \quad (22)$$

Then, we obtain:

$$E[I] \equiv \sum_{I,J} I \cdot \varphi_{I,J} = \sum_{i,j} (i+i_0) \tilde{\varphi}_{i,j} = x_{10} + i_0; \quad (23)$$

$$E[I^2] \equiv \sum_{I,J} I^2 \cdot \varphi_{I,J} = \sum_{i,j} (i+i_0)^2 \tilde{\varphi}_{i,j} = x_{20} + 2i_0 x_{10} + i_0^2; \quad (24)$$

$$Var[I] = E[I^2] - E[I]^2 = x_{20} + x_{10}^2. \quad (25)$$

Similarly, we have:

$$E[J] = x_{01} + j_0; \quad Var[J] = x_{02} - x_{01}^2. \quad (26)$$

Our goal now is to find the quantities x_{01} , x_{10} , x_{02} , x_{20} , which are essential for calculating the expectation and variance for I and J . In order to derive equations for these quantities, we multiply equation (21) by $i^\alpha j^\beta$ with $\alpha + \beta \leq 2$, and perform the summation in i and j in the quasi-stationary state. Each of the resulting 5 moment equations involve higher moments, which means that the number of the unknowns is larger than the number of equations. More precisely, these 5 equations involve 20 unknown variables, $x_{\alpha\beta}$, with $1 \leq \alpha + \beta \leq 5$. In order to proceed, we need to implement a truncation methods to close the system. In other words, we need to derive the missing equations for the higher moments, $x_{\alpha\beta}$, $3 \leq \alpha + \beta \leq 5$.

As in the 1D case, there are at least 3 different truncation methods that can be readily implemented to solve this system:

- **Simple truncation method** assumes that all the higher moments are 0, that is, $x_{\alpha\beta} = 0$, for $\alpha + \beta \geq 3$. This method requires the least amount of calculations,

because by assuming that all the higher moments are 0, we get a 5×5 linear system of equations.

- **Central moment truncation method** assumes that the central moments are equal to 0 for higher orders, that is, $E[(i-E[i])^\alpha(j-E[j])^\beta] = 0$, where $\alpha + \beta \geq 3$.
- **Cumulant truncation method** assumes that the higher order multivariate cumulants are equal to 0, that is, $\kappa_{\alpha,\beta} = 0$, where $\alpha + \beta \geq 3$.

The three methods are compared and contrasted in Appendix C.

In order to solve the resulting system of algebraic equations, we use the approximation of weak dependencies of the control functions on the cell numbers. Let us suppose that we can represent the functions $L_{I,J}$, $D_{I,J}$ and $P_{I,J}$ near the equilibrium as $L_{I,J} = L(\varepsilon I, \varepsilon J)$, $P_{I,J} = P(\varepsilon I, \varepsilon J)$, and $D_{I,J} = D(\varepsilon I, \varepsilon J)$, where the parameter $\varepsilon \ll 1$ defines the weakness of the dependence. It is convenient to denote $x = \varepsilon I$, $y = \varepsilon J$. Then we can expand the functions $L_{I,J}$, $P_{I,J}$ and $D_{I,J}$ around the steady state in Taylor series:

$$\tilde{L}_{i,j} \equiv L_{I,J} = L(\varepsilon i_0 + \varepsilon(I - i_0), \varepsilon j_0 + \varepsilon(J - j_0)) = L_0 + l_x i + l_y j + \frac{1}{2} (l_{xx} i^2 + l_{yy} j^2 + 2l_{xy} ij) + \dots, \quad (27)$$

$$\tilde{D}_{i,j} \equiv D_{I,J} = L_0 + d_x i + d_y j + \frac{1}{2} (d_{xx} i^2 + d_{yy} j^2 + 2d_{xy} ij) + \dots, \quad (28)$$

$$\tilde{P}_{i,j} \equiv P_{I,J} = \frac{1}{2} + p_x i + p_y j + \frac{1}{2} (p_{xx} i^2 + p_{yy} j^2 + 2p_{xy} ij) + \dots, \quad (29)$$

where the subscripts denote the partial derivatives of the functions with respect to its argument, evaluated at the equilibrium, $(I, J) = (i_0, j_0)$, and $I = i_0 + i$, $J = j_0 + j$. We further adopt the following convention: $L_x = L_x \varepsilon$, $L_{xx} = L_{xx} \varepsilon^2$, etc. In this description, the upper case constants $L_x = O(1)$, $L_{xx} = O(1)$, etc are all of order one, and all the derivatives expressed by lower-case letters contain a power of ε . In particular, the first derivatives l_x , l_y , p_x , etc contain a factor ε , and all the second derivatives l_{xx} , l_{xy} , l_{yy} , p_{xx} , etc contain a factor ε^2 . In Appendix C we demonstrate that all three truncation methods give the same result in the highest order of expansion in terms of ε .

3.3 Results for the cell number means and variances

Here we present the results for the means and the variances of the cell numbers. Let us define the pair (i_0, j_0) by equation (20), and derive the equations for the moments (the summation equations) by expanding the probability functions around this point, see

Appendix A. By the simple truncation method we set $x_{\alpha\beta} = 0$ for all $\alpha + \beta > 2$ and obtain the following five equations for the first and second moments:

$$(1, 0) \quad L_0 P_x x_{10} + L_0 P_y x_{01} + \varepsilon L_x P_x x_{20} + \varepsilon (L_y P_x + L_x P_y) x_{11} + \varepsilon L_y P_y x_{02} = 0, \quad (30)$$

$$(0, 1) \quad (-D_x + L_x + 2L_0 P_x) x_{10} + (-D_y + L_y + 2L_0 P_y) x_{01} + 2\varepsilon L_x P_x x_{20} + 2\varepsilon (L_y P_x + L_x P_y) x_{11} + 2\varepsilon L_y P_y x_{02} = 0, \quad (31)$$

$$(2, 0) \quad L_0 + \varepsilon L_x x_{10} + \varepsilon L_y x_{01} - 4\varepsilon L_0 P_x x_{20} - 4\varepsilon L_0 P_y x_{11} = 0, \quad (32)$$

$$(1, 1) \quad \begin{aligned} &L_0 + \varepsilon (L_x + 2L_0 P_x) x_{10} \\ &+ \varepsilon (L_y + 2L_0 P_y) x_{01} \\ &+ \varepsilon (D_x - L_x - 2L_0 P_x + 2\varepsilon L_x P_x) x_{20} \\ &+ \varepsilon (D_y - L_y - 2L_0 (P_x - P_y) + 2\varepsilon (L_y P_x + L_x P_y)) x_{11} \\ &+ 2\varepsilon (L_0 P_y + \varepsilon L_y P_y) x_{02} = 0, \end{aligned} \quad (33)$$

$$(0, 2) \quad \begin{aligned} &3L_0 + \varepsilon (D_x + 2L_x + 4L_0 P_x) x_{10} \\ &+ \varepsilon (D_y + 2L_y + 4L_0 P_y) x_{01} \\ &+ 4\varepsilon^2 L_x P_x x_{20} \\ &+ 2\varepsilon (-D_x + L_x + 2L_0 P_x + 2\varepsilon (L_y P_x + L_x P_y)) x_{11} \\ &+ 2\varepsilon (-D_y + L_y + 2L_0 P_y + 2\varepsilon L_y P_y) x_{02} = 0. \end{aligned} \quad (34)$$

Solving these to the highest order terms in ε and using (23), (25), and (26), we obtain the following result:

$$E[I] = i_0 + O(\varepsilon^0), \quad (35)$$

$$E[J] = j_0 + O(\varepsilon^0), \quad (36)$$

$$\text{Var}[I] = \frac{q_y^2 + 8L_0^2 p_y^2 + 2L_0 \Delta}{4B\Delta} + O(\varepsilon^0), \quad (37)$$

$$\text{Var}[J] = \frac{q_x^2 + 8L_0^2 p_x^2 + 6L_0 \Delta}{4B\Delta} + O(\varepsilon^0), \quad (38)$$

where we used the following notations:

$$q_x = l_x - d_x, \quad q_y = l_y - d_y,$$

$$\Delta = p_y q_x - p_x q_y, \quad B = 2L_0 (p_x = p_y) - q_y. \quad (39)$$

Appendix B demonstrates the application of our methods to system (15). As discussed, all three methods yield the same result in the highest order of expansion, and the results coincide with the ones previously obtained. The correction terms are different in different methods.

Quantities Δ and B , equation (39), are key for determining the stability properties of the stem cell lineage. It was shown in [38] that conditions $\Delta > 0$, $B > 0$ are necessary and sufficient for stability. Furthermore, equations (37) and (38) relate these quantities with the size of variance experienced by the cells in the stem and differentiated compartment. Expressions (39) define a subset in the four-dimensional parameter space, (q_x, q_y, p_x, p_y) , that is compatible with stability. Minimizing the variance in expressions (37, 38) restricts this subset further to identify the most general parameter region that is consistent with stable homeostasis. Note that only local properties (the derivatives at the steady state) of the control functions are needed to characterize homeostasis.

4 Numerical Simulations

In this section, we will demonstrate that the formulas given by equations (35–38) agree with the results from numerical simulations via two examples on two different types of control.

Throughout this section, let us denote $x = \varepsilon I$, $y = \varepsilon J$, $\hat{q}_x = L_x - D_x$, $\hat{q}_y = L_y - D_y$. Thus, $\hat{q}_x$ and $\hat{q}_y$ are the partial derivatives of the net growth rate, $L - D$, with respect to x and y . To clarify the biological meaning of these parameters, consider the quantity L_y . If it is nonzero, it means that the probability of stem cell division is controlled by the differentiated cell population. Moreover, if $L_y < 0$, this means that the control is negative (the more differentiated cells in the system, the less likely the stem cells are to divide); $L_y > 0$ means the existence of a positive control loop. The other three quantities can be interpreted in a similar manner. Below are two examples.

Negative control of differentiation and division

Consider the following functional forms of negatively controlled rates of division and differentiation:

$$L_{I,J}=L(\varepsilon I, \varepsilon J)=e^{-\varepsilon I}, \quad P_{I,J}=P(\varepsilon I, \varepsilon J)=e^{-\varepsilon J},$$

$$D_{I,J}=D(\varepsilon I, \varepsilon J)=1-L_{I,J}. \quad (40)$$

We therefore have $P_x = 0$, $P_y = -e^{-\varepsilon J} < 0$, $\hat{q}_x = -2e^{-\varepsilon I} < 0$, $\hat{q}_y = 0$. The steady state of the system can be obtained by solving $P(x, y) = \frac{1}{2}$, and $L(x, y) = D(x, y)$:

$$i_0 = j_0 = \frac{\log 2}{\varepsilon}.$$

By equations (35–38), we can obtain the means and the variances of the system:

$$E[I] = i_0, \quad (41)$$

$$E[J] = j_0, \quad (42)$$

$$Var[I] = \frac{8L_0^2 P_y^2 + 2L_0 P_y \hat{q}_x}{-8L_0 P_y^2 \hat{q}_x} \cdot \frac{1}{\varepsilon}, \quad (43)$$

$$Var[J] = \frac{\hat{q}_x^2 + 6L_0 P_y \hat{q}_x}{-8L_0 P_y^2 \hat{q}_x} \cdot \frac{1}{\varepsilon}, \quad (44)$$

where all the partial derivatives are evaluated at (i_0, j_0) , and $L_0 = L(\varepsilon i_0, \varepsilon j_0) = 1/2$.

For each value of ε , we ran numerical simulations starting at the expected values of the cell population given above, and finishing either when the number of time-steps reached $2 \cdot 10^5$, or if any of the cell types went extinct. We then computed the means and the variances of the cell population over the time-course of each simulation. A typical run for a particular value of ε is presented in figure 3.

From figure 4, we observe that the theoretical results for the means and the variances show a good agreement with the numerical results for smaller values of ε , which is what we expect. Also, the means and the variances of the cell population decrease as the value of ε increases, which is already predicted by the formulas given by equations (41–44).

Positive control of differentiation and negative regulation of division

The second example is given by equations:

$$L_{I,J} = L(\varepsilon I, \varepsilon J) = \frac{1}{1 + \varepsilon J}, \quad P_{I,J} = P(\varepsilon I, \varepsilon J) = 0.7 \cdot \tanh(\varepsilon I),$$

$$D_{I,J} = D(\varepsilon I, \varepsilon J) = 1 - L(\varepsilon I, \varepsilon J). \quad (45)$$

A typical stochastic simulation of system (45) for a particular value of ε is presented in figure 5.

To calculate the variances, we calculate $P_x = 0.7 \cdot \text{sech}^2(\varepsilon I) > 0$, $P_y = 0$, $\hat{q}_x = 0$, $\hat{q}_y = -2(1 + \varepsilon J)^{-2} < 0$. The steady state of the system can be obtained by solving

$$P(x, y) = \frac{1}{2}, \text{ and } L(x, y) = D(x, y):$$

$$i_0 = \frac{\log 6}{2\varepsilon}, \quad j_0 = \frac{1}{\varepsilon}.$$

By equations (35–38), we can obtain the means and the variances of the system:

$$E[I] = i_0, \quad (46)$$

$$E[J] = j_0, \quad (47)$$

$$\text{Var}[I] = \frac{\hat{q}_y^2 + 2L_0\hat{\Delta}}{4\hat{B}\hat{\Delta}} \cdot \frac{1}{\varepsilon}, \quad (48)$$

$$\text{Var}[J] = \frac{8L_0^2P_x^2 + 6L_0\hat{\Delta}}{4\hat{B}\hat{\Delta}} \cdot \frac{1}{\varepsilon}, \quad (49)$$

where all the partial derivatives are evaluated at (i_0, j_0) , and $L_0 = 1/2$,

$$\hat{\Delta} = P_y \hat{q}_x - P_x \hat{q}_y = -P_x \hat{q}_y, \quad \hat{B} = 2L_0(P_x - P_y) - \hat{q}_y = 2L_0P_x - \hat{q}_y.$$

We used the same numerical scheme as in the previous example with $2 \cdot 10^6$ time steps. As observed in figure 6, the theoretical results are in good agreement with the numerical results for smaller values of ε , which is consistent with the previous example. The means and the

variances of the cell population decrease as the value of ε increases, which is foretold by equations (46–49). From figure 7, we can see the overall pattern of the relative error: the smaller the value of ε , the smaller the relative error, which is what we expect.

5 Connection with the power series expansion method of Van Kampen

In this section, we will show that the simple truncation method described here and the well-known power series expansion method of Van Kampen [36] give exactly the same results, up to any order of expansion. First, we demonstrate how the Van Kampen method can be used for our system of stem and differentiated cells (as was done in [38]), and then argue that the two methods give the same results.

5.1 The method of Van Kampen: review and notations

Let us introduce the operators $\mathbf{E}_I^k$ and $\mathbf{E}_J^k$, such that

$$\mathbf{E}_I^k[f_{I,J}] = f_{I+k,J}, \quad \mathbf{E}_J^k[f_{I,J}] = f_{I,J+k}.$$

Then equation (14) can be rewritten more conveniently,

$$\dot{\varphi}_{I,J} = (\mathbf{E}_J^{+1} - 1)[\varphi_{I,J} D_{I,J}] + (\mathbf{E}_I^{-1} - 1)[\varphi_{I,J} L_{I,J} (1 - P_{I,J})] + (\mathbf{E}_I^{+1} \mathbf{E}_J^{-2} - 1)[\varphi_{I,J} L_{I,J} P_{I,J}]. \quad (50)$$

Equation (50) is nonlinear, and a general solution cannot be found. Therefore, we will use approximate methods to solve it. Let us assume that the functions $L_{I,J}$, $P_{I,J}$, and $D_{I,J}$ depend weakly on their arguments:

$$L_{I,J} = \mathcal{L}(\varepsilon I, \varepsilon J), \quad P_{I,J} = \mathcal{P}(\varepsilon I, \varepsilon J), \quad D_{I,J} = \mathcal{D}(\varepsilon I, \varepsilon J),$$

where $\varepsilon \ll 1$. We will use this parameter to perform the Van Kampen master equation expansion, in order to formulate the linear noise approximation [39]. We expect that in the long run, the probability distribution, $\varphi_{I,J}$, will have a peak somewhere around the (large) values

$$i_0 = \frac{\phi_I}{\varepsilon}, \quad j_0 = \frac{\phi_J}{\varepsilon},$$

with $\phi_I \sim \varepsilon^0$, $\phi_J \sim \varepsilon^0$. Let us suppose that the width of those peaks scales with $1/\varepsilon^{1/2}$. This is expressed in the following change of variables,

$$I = \frac{\phi_I(t)}{\varepsilon} + \frac{\xi(t)}{\varepsilon^{1/2}}, \quad J = \frac{\phi_J(t)}{\varepsilon} + \frac{\eta(t)}{\varepsilon^{1/2}}. \quad (51)$$

This change of variables will be used in the master equation (50). First of all, the probability function $\varphi_{I,J}(t)$ is now a function of ξ and η :

$$\varphi_{I,J}(t) = \Pi(\xi, \eta; t).$$

Its time-derivative can be written as follows,

$$\frac{d\varphi_{I,J}(t)}{dt} = \frac{\partial \Pi}{\partial t} + \frac{\partial \Pi}{\partial \xi} \dot{\xi} + \frac{\partial \Pi}{\partial \eta} \dot{\eta}.$$

Because the left hand sides of expressions (51) are time-independent, we have $\dot{\xi} = -\dot{\phi}_I/\varepsilon^{1/2}$, $\dot{\eta} = -\dot{\phi}_J/\varepsilon^{1/2}$. Also, we will introduce a slow time-scale,

$$\tau = \varepsilon t$$

(the necessity for this rescaling will become apparent once all the terms at different orders of ε are collected in the master equation). Therefore, we have for the time-derivative of $\varphi_{I,J}$:

$$\dot{\varphi}_{I,J} = \varepsilon \frac{\partial \Pi}{\partial \tau} - \varepsilon^{1/2} \left(\frac{\partial \Pi}{\partial \xi} \frac{\partial \phi_I}{\partial \tau} + \frac{\partial \Pi}{\partial \eta} \frac{\partial \phi_J}{\partial \tau} \right). \quad (52)$$

Next, we evaluate the shift operators. A jump of size k in the value of I is reflected by the jump of size $k\varepsilon^{1/2}$ in the value of ξ :

$$I+k = \frac{\phi_I(t)}{\varepsilon} + \frac{\xi(t)}{\varepsilon^{1/2}} + k = \frac{1}{\varepsilon} \left(\phi_I + \varepsilon^{1/2} \left(\xi + k\varepsilon^{1/2} \right) \right).$$

Similar arguments hold for the values of J . This allows us to express the shift operators $\mathbf{E}_I^k$ and $\mathbf{E}_J^k$ in terms of a (Taylor) series of differential operators,

$$\mathbf{E}_I^k = 1 + k\varepsilon^{1/2} \frac{\partial}{\partial \xi} + \frac{k^2 \varepsilon}{2} \frac{\partial^2}{\partial \xi^2} + \cdots, \quad (53)$$

and similarly for the shift in the J -direction.

Finally, we use ansatz (51) to expand the functions $L_{I,J}$, $P_{I,J}$ and $D_{I,J}$. We have

$$\mathcal{L}(\varepsilon I, \varepsilon J) = \mathcal{L}(\phi_I + \varepsilon^{1/2} \xi, \phi_J + \varepsilon^{1/2} \eta).$$

It is convenient to denote $x = \varepsilon I$, $y = \varepsilon J$, such that $L_{I,J} = \mathcal{L}(x, y)$, and denote by the subscripts the derivatives of this function with respect to its argument, evaluated at (ϕ_I, ϕ_J) : $L_x = \partial \mathcal{L} / \partial x$, $L_y = \partial \mathcal{L} / \partial y$, etc. We have

$$\mathcal{L}(\phi_I + \varepsilon^{1/2}\xi, \phi_J + \varepsilon^{1/2}\eta) = \mathcal{L}(\phi_I, \phi_J) + \varepsilon^{1/2}\xi L_x + \varepsilon^{1/2}\eta L_y + \frac{\varepsilon}{2}\xi^2 L_{xx} + \frac{\varepsilon}{2}\eta^2 L_{yy} + \varepsilon\xi\eta L_{xy} + \dots$$

Similarly, we expand the functions $\mathcal{P}$ and $\mathcal{D}$. These expressions, together with the operator expansions (53) and the time-derivative (52), are substituted into the master equation (50).

Then the terms in different orders of ε are equated. At order $\varepsilon^{1/2}$ we have

$$\frac{\partial \Pi}{\partial \xi} \frac{d\phi_I}{d\tau} + \frac{\partial \Pi}{\partial \eta} \frac{d\phi_J}{d\tau} = \frac{\partial \Pi}{\partial \xi} \mathcal{L}(\phi_I, \phi_J) (1 - 2\mathcal{P}(\phi_I, \phi_J)) + \frac{\partial \Pi}{\partial \eta} (2\mathcal{L}(\phi_I, \phi_J) \mathcal{P}(\phi_I, \phi_J) - \mathcal{D}(\phi_I, \phi_J)).$$

This equation gives rise to two “macroscopic laws”,

$$\frac{d\phi_I}{d\tau} = \mathcal{L}(\phi_I, \phi_J) (1 - 2\mathcal{P}(\phi_I, \phi_J)), \quad \frac{d\phi_J}{d\tau} = 2\mathcal{L}(\phi_I, \phi_J) \mathcal{P}(\phi_I, \phi_J) - \mathcal{D}(\phi_I, \phi_J). \quad (54)$$

or in steady state simply

$$\mathcal{P}(\phi_I, \phi_J) = 1/2, \quad \mathcal{D}(\phi_I, \phi_J) = \mathcal{L}(\phi_I, \phi_J) \equiv L_0. \quad (55)$$

Let us introduce the notations

$$\tilde{q}_x = \frac{L_x - D_x}{L_0} = \frac{q_x}{\varepsilon L_0}, \quad \tilde{q}_y = \frac{L_y - D_y}{L_0} = \frac{q_y}{\varepsilon L_0};$$

where q_x, q_y are defined in section 3.3. At order ε of the master equation expansion, after rescaling time once more by

$$T = L_0 \tau = L_0 \varepsilon t,$$

we obtain the following linear Fokker-Planck equation:

$$\frac{\partial \Pi}{\partial T} = -(\tilde{q}_y + 2P_y)(\eta \Pi)_\eta - (\tilde{q}_x + 2P_x)(\xi \Pi)_\xi + 2P_x(\xi \Pi)_\xi + 2P_y(\eta \Pi)_\eta + \frac{1}{2}(3\Pi_{\eta\eta} - 2\Pi_{\xi\eta} + \Pi_{\xi\xi}). \quad (56)$$

This is the linear noise approximation of Van Kampen [39]. The validity of this approximation has been studied extensively, see e.g. [40, 41]. Here we mention that the relative size of typical fluctuations scales with $\varepsilon^{1/2}$, and thus for sufficiently small values of ε , the system will remain near the equilibrium and stochastic extinction is an unlikely event,

at least for a time-duration which grows with $1/\varepsilon$. For a rigorous study of extinction times of birth-death processes see e.g. [42, 43].

From equation (56) we can obtain the equations for the first and second moments in a standard way:

$$\frac{d\langle\xi\rangle}{dT} = -2(P_x\langle\xi\rangle + P_y\langle\eta\rangle), \quad (57)$$

$$\frac{d\langle\eta\rangle}{dT} = (\tilde{q}_y + 2P_y)\langle\eta\rangle + (\tilde{q}_x + 2P_x)\langle\xi\rangle, \quad (58)$$

$$\frac{d\langle\xi^2\rangle}{dT} = -4(P_x\langle\xi^2\rangle + P_y\langle\xi\eta\rangle) + 1, \quad (59)$$

$$\frac{d\langle\eta^2\rangle}{dT} = 2(\tilde{q}_y + 2P_y)\langle\eta^2\rangle + 2(\tilde{q}_x + 2P_x)\langle\xi\eta\rangle + 3, \quad (60)$$

$$\frac{d\langle\xi\eta\rangle}{dT} = (\tilde{q}_y + 2P_y)\langle\xi\eta\rangle + (\tilde{q}_x + 2P_x)\langle\xi^2\rangle - 2(P_x\langle\xi\eta\rangle + P_y\langle\eta^2\rangle) - 1. \quad (61)$$

As we will show in the next section, the above moment equations are exactly the same as the summation equations (30)–(34) if we only keep the leading order terms, and hence Van Kampen method and simple truncation method give the same results to the leading order, see (35)–(38).

5.2 Comparison of the simple truncation method and the Van Kampen method

Before we illustrate the equivalence of the two methods, we state for convenience some fundamental facts that we will use later:

- Taylor series expansion. We know the Taylor expansion of $f(j) = j^n$ center at $j = a$ is: $j^n = a^n + na^{n-1}(j-a) + \dots + na(j-a)^{n-1} + (j-a)^n$. Notice that the second to the last term is a product of the jump size a and the derivative of the last term $(j-a)^{n-1}$.
- Integration by parts. When we compute $\int \eta^n(\cdot)_{\eta} d\eta$, integration by parts will give $\eta^n(\cdot) - \int n \eta^{n-1}(\cdot) d\eta$.

The leading order—By using the ansatz (51) in the previous section, we have in steady states:

$$x_{\alpha\beta} = \sum_{I,J} \varphi_{I,J} (I - i_0)^\alpha (J - j_0)^\beta = \frac{\langle \xi^\alpha \eta^\beta \rangle}{\varepsilon^{(\alpha+\beta)/2}}; \quad (62)$$

where $x_{\alpha\beta}$ is defined in (22).

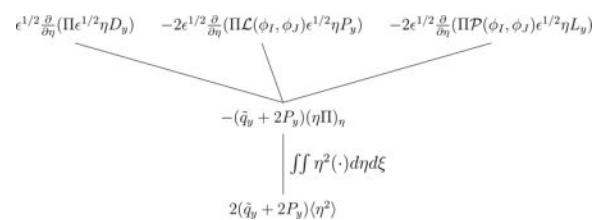
By using (62), we can rewrite the summation equations from section 3.3 in terms of $\langle \xi^\alpha \eta^\beta \rangle$. It turns out that they are the same as moment equations (57–61) if we only keep the leading order terms. Since the moment equations agree to the leading order, the two methods give the same results to the leading order.

The next order—Next, we investigate if the two methods provide the same result in the next order of accuracy. To this end, we will investigate the structure of a moment equation in depth from both methods by looking at a particular term. For illustration, we will analyze the first term in equation (60) and its counterpart in equation (34) from section 3.3:

$$\frac{d\langle \eta^2 \rangle}{dT} = 2(\tilde{q}_y + 2P_y) \langle \eta^2 \rangle + \dots \quad (63)$$

$$\frac{dx_{02}}{dt} = 2\varepsilon (-D_y + L_y + 2L_0 P_y) x_{02} + \dots \quad (64)$$

If we trace back the terms which contribute to $2(\tilde{q}_y + 2P_y) \langle \eta^2 \rangle$ in (63), we will have the following diagram:



Let us focus on the third term at the first level, we have the following observation: $-2\varepsilon^{1/2} \frac{\partial}{\partial \eta}$ is the second term of Taylor series of E_J^{-2} ; $\mathcal{P}(\phi_I, \phi_J)$ is the first term of $P_{I,J}$ in Taylor expansion; $\varepsilon^{1/2} \eta L_y$ is the third term of Taylor series of $L_{I,J}$.

To draw comparison, we will also trace back the terms which contribute to $2\varepsilon(-D_y + L_y + 2L_0 P_y)x_{02}$ in (64). We obtain the following picture:

$$\begin{array}{c}
 \sum_{i,j} \tilde{\varphi}_{i+1,j} D_{\eta}(j+1)[-2(j+1)] \quad \sum_{i,j} \tilde{\varphi}_{i+1,j-2} L_{\eta} P_{\eta}(j-2)4(j-2) \quad \sum_{i,j} \tilde{\varphi}_{i+1,j-2} \frac{1}{2} L_{\eta}(j-2)4(j-2) \\
 \swarrow \quad \downarrow \quad \searrow \\
 2(-D_{\eta} + L_{\eta} + 2L_{\eta} P_{\eta})x_{02}
 \end{array}$$

Let us look at the corresponding counterpart at the first level. Observe that $1/2$ is the first term of Taylor expansion of $P_{i+1,j-2}$; $L_{\eta}e(j-2)$ is the third term of $L_{i+1,j-2}$ in Taylor expansion; $4(j-2)$ is the second to the last term of Taylor expansion of j^2 centered at $j=2$.

It is not hard to see that the term $4(j-2)$ “captures” the jump size of E_j^{-2} ; namely 2, and the derivative of η^2 (obtained from integration by parts in the first diagram); moreover, the

differential operator $-\frac{\partial}{\partial \eta}$ will be offset after integration by parts. Hence, the two terms are exactly the same. In fact, it can be shown that any two corresponding terms at the first level (from both diagrams) are the same by similar analysis, which are due to the fundamental properties mentioned at the beginning of this section. This methodology is essentially carried over in any two corresponding terms in a homologous pair of moment equations.

To see whether the two methods agree to the next order corrections, we need to assume γ is sufficiently large, where γ is the constant for which $x_{a\beta} = 0$ for $a + \beta \geq \gamma$ in the simple truncation method. It turns out that the moment equations that are used to compute the next order correction are the same for the two methods if we set $\gamma = 5$, see Appendix E for details.

Generalization to higher orders of accuracy—By the methodology presented in the previous section, we can deduce that the two methods will produce the same moment equations for computing any order corrections if we set γ sufficiently large, hence the two methods give exactly the same results (up to any order) to the general two-step model as stated at the beginning. However, as the value of γ increases, the computation of moment equations will become more and more tedious. So, there is a trade off between efficiency and accuracy.

Under the same value of γ , cumulant truncation method is the most accurate among the three truncation methods presented in section 3.2. To see this, set $\gamma = 3$ and look at the leading order of the terms $x_{a\beta}$ for $a + \beta = 4$. From Appendix D, we see that the leading

order is $O\left(\frac{1}{\varepsilon}\right)$ and $O\left(\frac{1}{\varepsilon^2}\right)$ from central moment truncation and cumulant closure method, respectively. On the other hand, we can obtain the leading order of $\langle \xi^a \eta^\beta \rangle$ is $O(1)$ for $a + \beta = 4$ from Van Kampen method, as shown in E of appendix. By (62), the leading order of $x_{a\beta}$ is $O\left(\frac{1}{\varepsilon^2}\right)$ for $a + \beta = 4$. Therefore, cumulant truncation method is the most accurate though it's the most expensive in terms of computation, which verdicts the trade off between efficiency and accuracy.

In conclusion, simple truncation method produces the same results as Van Kampen method up to any order by setting γ sufficiently large. The advantage of simple truncation method is the straightforward calculations that it involves. To see this, we can compare the steps used

to obtain the moment/summation equations in each method. For simple truncation method, we only use Taylor expansion on the probability functions of the master equation. On the other hand, Van Kampen method uses integration besides the master equation expansion in Taylor series. Our method could be regarded as a short-cut compared to the Van Kampen derivation. To see this, recall that we multiplied the equation (56) by $\xi^a \eta^b$, and then integrated to obtain the moment equations given by (57–61), for $a + b \leq 2$. This extra step requires more computational work. Clearly, there will be more terms to integrate in order to compute the moment equations for the next order corrections, since we extend the equation (56) to order $O(e^{1/2})$, see appendix E.1. In the Supplementary Materials, we provide a *Mathematica* program that implements our methodology to compute the summation equations.

6 Discussion and conclusions

In this paper we presented a very general approach to the studies of stem cell/differentiated cell dynamics. We assumed that cell divisions and deaths happen according to a Markov process, where the probability rates for different events are some functions of the current populations. Our methodology allows us to compute the means and the variances of the stem and differentiated cell populations. It is based on the linearization of the control functions near the steady state and truncating the moment equations by using three different techniques. We have shown that all three techniques yield identical results for the highest order term in the mean and the variance. Therefore, it is to one's advantage to use the easiest of the three methods, namely, what we called the simple truncation method. We have shown that while the result of this method coincides with a variation of the Van Kampen power series expansion, our method is easier to implement in practice, as it requires only a Taylor series expansion, while the Van Kampen method also includes integration.

For the system with constant populations, the method yields 2 linear algebraic equations for the moments, which can be solved to give simple expressions for the cell number variances. For non-constant populations, we obtain 5 linear algebraic equations for the moments.

In the case of constant overall populations, where only differentiation/proliferation decisions lead to fluctuations of the stem cell numbers, we demonstrated that the variance of the stem cell numbers is given by 1/4 times the inverse of the derivative of the control function with respect to the number of stem cells, taken at the equilibrium, formula (9).

In the case where the overall population varies, we have derived an equivalent formula for the variances, expressed as a function of the partial derivatives of the controls with respect to the two population sizes, taken at the equilibrium, equations (37–38). The analytical expressions for the means and variances of the stem and differentiated cell populations are compared with numerical simulations for two different examples.

Applications

There are several general areas of application of this kind of modeling. Firstly, one can study the questions about the tissue architecture of different organs. Why does mouse epithelium at different locations have different structure and different division symmetry? We are

currently using a modification of the models presented here to explain this. Secondly, one could implement a parameterized model of stem cell lineage control and ask questions about the dynamics of cell renewal both in healthy tissues, and in cancer. What type of mutations can lead to rapid expansion? What is the effect of a given mutation on cell population dynamics, given the underlying control system? This is a continuation of the work started in [44], but with a more general theoretical basis for the description of control networks. Finally, one can attempt to solve the “inverse problem”: suppose we know the cell numbers in different compartments. Can we then reconstruct the underlying control network that governs the dynamics?

Our present study differs from previous theoretical literature on the subject because we do not make any prior assumptions on the type and direction of control loops, apart from the fact that a stable equilibrium exists, which biologically corresponds to the existence of homeostatic control. Instead of looking at particular functional forms of regulatory mechanisms, we investigate the population dynamics of cell lineages in the most general setting. The particular functional forms of control loops can be inferred from careful measurements of biological systems of interest.

Model parameterization

What defines the shape of the control loops, that is, the functional forms of the functions $L(I, J)$, $D(I, J)$, and $P(I, J)$? Stem cell regulation is often described in the context of the so-called stem cell niche, an anatomic location that regulates how stem cells participate in tissue generation, maintenance and repair [45]. The niche includes both cellular and non-cellular components that interact in order to control the adult stem cell [46]. Within a niche, the stem cell fate - that is, its division and differentiation decisions - are under the regulation of many different factors, including structural and physical forces, paracrine and endocrine signaling from neighboring and distant cells, metabolic factors and neural signaling [47].

The regulatory mechanisms that have been discussed in the literature include growth factors, cell-cell contacts, and cell-matrix adhesions [48], regulation by microRNAs [49, 50], signaling from mesenchymal cells, as well as differentiated cells [51]. In [46], both physical contact with the niche, and diffusible factors that regulate stem cell behavior, have been catalogued for neural, epidermal, haematopoietic, and intestinal stem cells. Many more mechanisms exist that are responsible for controlling cell decisions of both stem cells and other cell types, see e.g. [52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68].

For example, there is evidence in the literature that the differentiation probability, $P(I, J)$, is a decaying function of I , because of cell crowding and contact inhibition effects that take place inside the stem cell niche [62]. In our example (15) the function P was assumed a decaying function of the stem cell number.

In other circumstances, $P(I, J)$ could be a growing function of its variables [69], because in some systems, mechanical strain has been shown to increase cell differentiation [62, 53, 61]. In our example (45), P increases as a function of the stem cell population.

Negative control of differentiation from downstream (that is, by differentiated cells) has been reported in the context of the adult neurogenesis [52], in colon [51], in the haematopoietic system [64] and in the olfactory epithelium [23]. Our example (40) assumes that P decays with J , the number of differentiated cells.

It has been observed that the rate of cell divisions, $L(I, J)$, like the division type, is also under regulation of several types of control loops [67]. A “crowd-control” model is described in [58] consistent with $L(I, J) - D(I, J)$ being a decreasing function of cell numbers. In examples (40) and (45), the function $L - D$ is assumed to be a decreasing function of the number of stem cells and differentiated cells, respectively, while in example (15) it is a decreasing function of the total population $I + J$.

Outlook

In the present paper we show how the biological information of this kind can be incorporated in a rather transparent way to inform us about the efficiency of control in maintaining the homeostasis by keeping the variance of cell populations low. This is a first step in the direction of understanding the dynamics of different types of control loops. Future directions include generalizing this model to multiple cell compartments, and including a variety of other cellular processes such as de-differentiation and asymmetric stem cell divisions. Furthermore, the present model is non-spatial. Incorporation of spatial information about the geometry of the stem cell niche would be very useful. Cell signaling is mediated by soluble compounds and its effectiveness is intrinsically a function of distance. Cellular compartments that are separated in space will by necessity have weaker influence on each other's cell decisions compared to neighboring compartments. Therefore, developing effective analysis methods, which take explicit account of spatial interactions is an important future direction.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

NK gratefully acknowledges support of grant 1R01CA129286-01A1.

A The summation equations

Here we consider the patterns for the 5 summation equations which are derived from equation (21) by multiplying by $i^a j^b$ and summing over i and j , for $a + b \leq 2$. Let us denote the summation equation derived from multiplying by $i^a j^b$ by the pair of numbers (a, b) .

We present the example of equation (1, 0). Multiplying equation (21) by i and summing over i and j , we obtain

$$\begin{aligned}
0 = & \sum_i \sum_j \tilde{\varphi}_{ij} \left(-\frac{1}{2} \varepsilon (4iL_0P_x + 4jL_0P_y) - \frac{1}{2} \varepsilon^2 [2i^2(2L_xP_x + L_0P_{xx}) + 4ij(L_yP_x + L_0P_{xy} + L_xP_y) + 2j^2(2L_yP_y + L_0P_{yy})] - \frac{1}{2} \varepsilon^3 [i^3 \right. \\
& + ij^2(4L_{xy}P_y + 2L_xP_{yy} + 4L_yP_{xy} + 2L_{yy}P_x) + j^3(2L_{yy}P_y + 2L_yP_{yy})] \\
& - \frac{1}{2} \varepsilon^4 [i^4L_{xx}P_{xx} \\
& + 2i^3j(L_{xy}P_{xx} + L_{xx}P_{xy}) \\
& \left. + i^2j^2(L_{yy}P_{xx} + 4L_{xy}P_{xy} + L_{xx}P_{yy}) + j^2L_{yy}P_{yy}] \right)
\end{aligned}
\tag{65}$$

From the above equation, we notice that the power of ε is consistent with the power of $i^\alpha j^\beta$. Indeed, every term containing ε^k term multiplies a term of the form $i^\alpha j^\beta$, where $\alpha + \beta = k$. The reason for this can be seen by examining the structure of equation (21). In this equation, every term containing the power k of ε (as follows from the number of derivatives of the probability functions), is multiplied by $i^\alpha j^\beta$ with $\alpha + \beta = k$.

The same property holds for the other four summation equations, which we do not present here. To derive all the summation equations, we need to perform the summations in i and j and use definition (22) for the moments. Each equation will be coupled to other equations containing higher order moments. The easiest way to close the system is to use the simple truncation method. The resulting system of 5 equations is given by (30–34). The other two truncation methods are worked out below.

B A case study

In this section, we will consider a special case of the general model equation (14). We will see that the results for the general case coincide with the results found previously by a different methodology in [35] and reported in Section 3.1. In particular, we will demonstrate that all three truncation methods have the same result for $a_{\alpha\beta}^1$ but different results for $a_{\alpha\beta}^2$,

where $x_{\alpha\beta} = \frac{a_{\alpha\beta}^1}{\varepsilon^n} + \frac{a_{\alpha\beta}^2}{\varepsilon^{n-1}}$, $\alpha + \beta \leq 2$.

In this example, we assume that

$$Q(N) = \frac{b}{1+hN}, \quad P(J) = \frac{r}{1+gJ},$$

where $h, g \ll 1$ and $b, r = O(1)$ are two constants. The corresponding Kolmogorov forward equation is given by:

$$\dot{\varphi}_{J,N} = \varphi_{J+1,N+1} (1 - Q(N+1)) + \varphi_{J,N-1} Q(N-1) (1 - P(J)) + \varphi_{J-2,N-1} Q(N-1) P(J-2) - \varphi_{J,N}. \quad (66)$$

Let n_0 and j_0 be the steady state, and $n = N - n_0$, $j = J - j_0$. n_0 and j_0 are defined to satisfy the deterministic equations:

1. $\dot{N} = Q(N) - (1 - Q(N)),$
2. $\dot{J} = 2Q(N)P(J) - (1 - Q(N)).$

We can easily see that $n_0 = \frac{2b-1}{h}$.

Define $\tilde{\varphi}_{j,n}$ such that $\tilde{\varphi}_{j,n} = \varphi_{J,N}$, and $\tilde{Q}(n) = Q(N)$, $\tilde{P}(j) = P(J)$, then (66) can be rewritten as:

$$\dot{\tilde{\varphi}}_{j,n} = \tilde{\varphi}_{j+1,n+1} (1 - \tilde{Q}(n+1)) + \tilde{\varphi}_{j,n-1} \tilde{Q}(n-1) (1 - \tilde{P}(j)) + \tilde{\varphi}_{j-2,n-1} \tilde{Q}(n-1) \tilde{P}(j-2) - \tilde{\varphi}_{j,n}. \quad (67)$$

Expanding $Q(N)$ and $P(J)$ in Taylor series, we obtain

$$Q(N) = \frac{1}{2} + r_1 n + r_2 n^2 + \dots,$$

$$P(J) = \frac{1}{2} + s_1 j + s_2 j^2 + \dots,$$

Let us multiply both sides of equation (67) by $j^\alpha n^\beta$ and sum over j, n in the quasi-stationary state. We obtain the following 5 summation equations, where $x_{\alpha\beta} = \sum_j \sum_n j^\alpha n^\beta \tilde{\varphi}_{j,n}$:

$$r_1 x_{01} + r_2 x_{02} = 0$$

$$2r_1 x_{01} + s_1 x_{10} + 2r_2 x_{02} + 2r_1 s_1 x_{11} + 2r_2 s_1 x_{12} + s_2 x_{20} + 2r_1 s_2 x_{21} + 2r_2 s_2 x_{22} = 0$$

$$1 + 4r_1 x_{02} + 4r_2 x_{03} = 0$$

$$\begin{aligned}
&1 + s_1 x_{10} + 2r_1 x_{02} \\
&+ (2r_1 + s_1 + 2r_1 s_1) x_{11} + s_2 x_{20} + 2r_2 x_{03} \\
&+ 2(r_2 + r_1 s_1 + r_2 s_1) x_{12} \\
&+ (s_2 + 2r_1 s_2) x_{21} + 2r_2 s_1 x_{13} \\
&+ 2s_2(r_1 + r_2) x_{22} \\
&+ 2r_2 s_2 x_{23} = 0
\end{aligned}$$

$$\begin{aligned}
&\frac{3}{2} + r_1 x_{01} + 2s_1 x_{10} \\
&+ r_2 x_{02} + 4r_1(1 + s_1) x_{11} \\
&+ 2(s_1 + s_2) x_{20} \\
&+ 4r_2(1 + s_1) x_{12} \\
&+ 4r_1(s_1 + s_2) x_{21} \\
&+ 2s_2 x_{30} \\
&+ 4r_2(s_1 + s_2) x_{22} \\
&+ 4r_1 s_2 x_{31} \\
&+ 4r_2 s_2 x_{32} = 0
\end{aligned}$$

As in the general case, we expand every term in the Taylor series:

$$Q(N) = \frac{1}{2} - \frac{h_0}{4}n + \frac{h_0^2}{8}n^2,$$

$$P(J) = \frac{1}{2} - \frac{g_0}{4}j + \frac{g_0^2}{8}j^2,$$

where we introduce the following short-hand notations:

$$h_0 = \frac{h}{b}, \quad g_0 = \frac{g}{r}, \quad \eta = \frac{g_0}{h_0}.$$

We will use the truncation equations of Appendix D for central moment and cumulant

closure method to solve the system for $x_{\alpha\beta} = \frac{a_{\alpha\beta}^1}{h_0^n} + \frac{a_{\alpha\beta}^2}{h_0^{n+1}}$, $\alpha + \beta \leq 2$. The solutions are presented below.

1. Simple truncation method

$$E[j] = x_{10} = \frac{4+3\eta}{2(2+\eta)} + O(h_0^2)$$

$$E[n] = x_{01} = \frac{1}{2} + O(h_0^2)$$

$$E[j^2] = x_{20} = \frac{2+3\eta}{\eta(2+\eta)} * \frac{1}{h_0} + O(h_0)$$

$$E[jn] = x_{11} = \frac{2}{2+\eta} * \frac{1}{h_0} + O(h_0)$$

$$E[n^2] = x_{02} = \frac{1}{h_0} + O(h_0)$$

Comparing with the old results in formulas (16)–(19), here we have

$$E[N] = \frac{2b-1}{h} + \frac{1}{2}, \quad Var[N] = \frac{b}{h} - \frac{1}{4}$$

2. Central moment truncation method

$$x_{10} = \frac{4+3\eta}{2(2+\eta)} + \frac{44\eta+108\eta^2+93\eta^3+27\eta^4}{8(2+\eta)^3} h_0 + O(h_0^2)$$

$$x_{01} = \frac{1}{2} + \frac{3}{8} h_0 + O(h_0^2)$$

$$x_{20} = \frac{2+3\eta}{\eta(2+\eta)} * \frac{1}{h_0} + \frac{8+92\eta+138\eta^2+93\eta^3+27\eta^4}{4\eta(2+\eta)^3} + O(h_0)$$

$$x_{11} = \frac{2}{2+\eta} * \frac{1}{h_0} + \frac{16+24\eta+22\eta^2+9\eta^3}{2(2+\eta)^3} + O(h_0)$$

$$x_{02} = \frac{1}{h_0} + \frac{3}{4} + O(h_0) \quad (68)$$

3. Cumulant truncation method

$$x_{10} = \frac{4+3\eta}{2(2+\eta)} + \frac{\eta(228+556\eta+483\eta^2+135\eta^3)}{32(2+\eta)^3} h_0 + O(h_0^2)$$

$$x_{01} = \frac{1}{2} + \frac{3}{8} h_0 + O(h_0^2)$$

$$x_{20} = \frac{2+3\eta}{\eta(2+\eta)} * \frac{1}{h_0} + \frac{328+580\eta+438\eta^2+135\eta^3}{16(2+\eta)^3} + O(h_0)$$

$$x_{11} = \frac{2}{2+\eta} * \frac{1}{h_0} + \frac{128+140\eta+184\eta^2+105\eta^3}{16(2+\eta)^3} + O(h_0)$$

$$x_{02} = \frac{1}{h_0} + \frac{3}{4} + O(h_0)$$

From the above results, all three truncation methods have the same solution for $a_{\alpha\beta}^1$, but

different solutions for $a_{\alpha\beta}^2$, where $x_{\alpha\beta} = \frac{a_{\alpha\beta}^1}{h_0^n} + \frac{a_{\alpha\beta}^2}{h_0^{n+1}}$, $\alpha + \beta \leq 2$.

As we can see, the simple truncation, central moment truncation and cumulant truncation methods yield the same result for the highest order terms in $x_{\alpha\beta}$, where $\alpha + \beta \leq 2$.

Therefore, all three methods have the same result for the mean and variance of N and J , if we only keep the highest order term, as summarized below:

$$E[N] = \frac{2b-1}{h} + \frac{1}{2} + O(h), \quad (69)$$

$$\text{Var}[N] = \frac{b}{h} + O(1), \quad (70)$$

$$E[J] = \frac{2r-1}{g} + \frac{4+3\eta}{2(2+\eta)} + O(h), \quad (71)$$

$$\text{Var}[J] = \frac{2+3\eta}{\eta(2+\eta)} * \frac{1}{h_0} + O(1). \quad (72)$$

To compare the above results with formulas (16)–(19), derived from direct calculations in Section 3.1, we expand formulas (16)–(19) with respect to h_0 , and then only keep the highest order terms. The results are identical to equations (69–72).

C Comparison of the three truncation methods

Each of the three truncation methods has its own advantages and disadvantages. The advantage of the simple truncation method is its simplicity. Compared to the simple truncation method, both the central moment and cumulant truncation method require more extensive calculations. However, these methods can give more accurate results, as shown below.

The order of magnitude for the lower moments

Let us expand the moments $x_{a\beta}$ in a power series in terms of ε , $x_{a\beta} = \sum_{n=-\infty}^{\infty} X_{a\beta}^{(n)} \varepsilon^n$. We will consider only the two highest order terms in this expansion, with the corresponding coefficients denoted as $a_{a\beta}^1$ and $a_{a\beta}^2$. That is, we write

$$x_{a\beta} = \frac{a_{a\beta}^1}{\varepsilon^n} + \frac{a_{a\beta}^2}{\varepsilon^{n-1}} + O\left(1/\varepsilon^{n-2}\right),$$

where $a_{a\beta}^1, a_{a\beta}^2 = O(1)$ are unknown constants that we need to find. Next, we prove that for $x_{a\beta}$, where $\alpha + \beta \leq 2$, all three truncation methods yield the same result for $a_{a\beta}^1$, but the results of the three methods differ for $a_{a\beta}^2$.

To determine the largest contributions to the expansions for $x_{a\beta}$, we consider the five moment equations. At order (m, k) , we multiply equation (21) by $i^m j^k$ and perform a double-summations in i and j . We call the resulting equations the *summation equations*. Because of expansion (27–29), coefficients in front of different variables $x_{a\beta}$ will have a different order in terms of ε . In general, such an equation will contain terms multiplying $x_{m+s, k+r}$ with $s = 0, 1, \dots$ and $r = 0, 1, \dots$. The coefficient in front of the term $x_{m+s, k+r}$ is of the order ε^{s+r} . The summation equations may also contain a nonhomogeneous (constant) term of order $O(1)$.

Because $x_{01} = E[j]$, and $j = J - j_0$ is a small perturbation around the steady state, it is reasonable to assume that $x_{01} = a_{01}^1 + a_{01}^2 \varepsilon$ and $x_{10} = a_{10}^1 + a_{10}^2 \varepsilon$. Because $x_{02} = E[j^2]$, x_{02} should be at least of order $O\left(\frac{1}{\varepsilon}\right)$. But if x_{02} is of order $O\left(\frac{1}{\varepsilon^2}\right)$ or higher, then from the 5 summation equations, the coefficient of the terms with power $O\left(\frac{1}{\varepsilon^k}\right)$ with $k \geq 2$ should be

0, which is not the case. Therefore, $x_{02} = \frac{a_{02}^1}{\varepsilon} + a_{02}^2$. Similarly, we have $x_{11} = \frac{a_{11}^1}{\varepsilon} + a_{11}^2$ and $x_{20} = \frac{a_{20}^1}{\varepsilon} + a_{20}^2$. Calculations presented in Appendix A demonstrate these arguments in detail. As it will be discussed in E.3 of Appendix, these assumptions are indeed valid.

The order of magnitude of the higher moments

For the 3rd order truncation equations, the central moment and the cumulant closure methods have the same truncation equations, because for $\alpha + \beta = 3$ we have $E[(i - E[i])^\alpha (j - E[j])^\beta] = \kappa_{\alpha,\beta}$. Setting these moments to zero yields 4 equations (equations (73)–(76) in

Appendix D). Because $x_{\alpha\beta} = O(1)$ for $\alpha + \beta = 1$ and $x_{\alpha\beta} = O\left(\frac{1}{\varepsilon}\right)$ for $\alpha + \beta = 2$, in order to balance equations (73–76), $x_{\alpha\beta}$ with $\alpha + \beta = 3$ has to be of the order of $O\left(\frac{1}{\varepsilon}\right)$.

For $\alpha + \beta \geq 4$, the central moment and the cumulant closure methods have different truncation equations (see Appendix D.2). For the central moment truncation method, we have

$x_{\alpha\beta} = O\left(\frac{1}{\varepsilon}\right)$ with $\alpha + \beta = 4, 5$. For the cumulant truncation method, we have $x_{\alpha\beta} = O\left(\frac{1}{\varepsilon^2}\right)$, where $\alpha + \beta = 4, 5$ (see Appendix D).

All three methods coincide for $a_{\alpha\beta}^1$, but differ for $a_{\alpha\beta}^2$

Next, we will show that all three truncation methods give the same result for $a_{\alpha\beta}^1$, but

different results for $a_{\alpha\beta}^2$, where $x_{\alpha\beta} = \frac{a_{\alpha\beta}^1}{\varepsilon^n} + \frac{a_{\alpha\beta}^2}{\varepsilon^{n-1}} + O\left(\frac{1}{\varepsilon^{n-2}}\right)$ and $\alpha + \beta \leq 2$.

First, we look for the highest order term of ε in the 5 summation equations. By the above and equation (65) in Appendix A, the highest order term in the 5 summation equations is

$$\varepsilon f_1(x_{01}, x_{10}) + \varepsilon^2 f_2(x_{02}, x_{11}, x_{20}),$$

where f_1 is a function of x_{01}, x_{10} and f_2 is a function of x_{02}, x_{11}, x_{20} . Presenting $x_{\alpha\beta}$ with $\alpha + \beta \leq 2$ as a series, and keeping only the highest order terms in the summation equations, we obtain a linear system for $a_{\alpha\beta}^1$. Because the linear system derived from the moment equations is independent of the truncation methods, we can see that all three truncation methods yield the same result for $a_{\alpha\beta}^1$, where $\alpha + \beta \leq 2$. This is also the reason why the highest order contributions to the expectation and variance for i and j are the same for all the three methods, and so are the leading order to the expectation and variance for I and J by equations (23), (25), and (26).

Next, we will show that $a_{\alpha\beta}^2$ with $\alpha + \beta \leq 2$ are different in the 3 truncation methods by looking for the second highest order terms of ε in the 5 summation equations.

In the simple truncation method, the second highest order terms are $a_{\alpha\beta}^2 \varepsilon^2$, $\alpha + \beta \leq 2$, because we simply assume that all the higher moments are 0. Thus, we get a linear system for $a_{\alpha\beta}^2$, $\alpha + \beta \leq 2$.

In the central moment truncation method, the second highest order terms are $a_{\alpha\beta}^2 \varepsilon^2$ for $\alpha + \beta \leq 2$, and $a_{\alpha\beta}^1 \varepsilon^2$ for $\alpha + \beta = 3$, because when $\alpha + \beta \geq 4$, the power is at least $O(\varepsilon^3)$, by equation (65). Therefore, we obtain a linear system for $a_{\alpha\beta}^2$, $\alpha + \beta \leq 2$, which contains terms $a_{\alpha\beta}^1$ with $\alpha + \beta = 3$. We can solve the system for coefficients $a_{\alpha\beta}^1$ with $\alpha + \beta = 3$ from the third order central moment truncation equations.

Finally, for the cumulant truncation method, the second highest order terms are $a_{\alpha\beta}^2 \varepsilon^2$ for $\alpha + \beta \leq 2$, and $a_{\alpha\beta}^1 \varepsilon^2$ for $\alpha + \beta = 3, 4$ (the argument is similar to the one presented above). The difference from the central moment closure method is that the linear system for $a_{\alpha\beta}^2$ with $\alpha + \beta \leq 2$ does not only contain terms $a_{\alpha\beta}^1$ with $\alpha + \beta = 3$, but also contains terms $a_{\alpha\beta}^1$ with $\alpha + \beta = 4$. Similarly, $a_{\alpha\beta}^1$ with $\alpha + \beta = 4$ can be obtained from the cumulant truncation equations.

From the above considerations, we can see that the equations for $a_{\alpha\beta}^2$ with $\alpha + \beta \leq 2$ for the simple truncation method contain no information about the higher order terms $x_{\alpha\beta}$, $\alpha + \beta \geq 3$. For the central moment truncation method, these equations contain some information about the higher order terms, which is $x_{\alpha\beta}$ with $\alpha + \beta = 3$. Finally, for the cumulant truncation method, these equations contain information about $x_{\alpha\beta}$ with $\alpha + \beta = 3, 4$.

In conclusion, the three truncation methods produce the same result to the leading order of the mean and variance for the cell population. While we expect the central moment and cumulant truncation methods to give more accurate results to the next order correction, they require more extensive computations. In section 5.2, we show that the cumulant truncation is the most accurate among the three methods. The advantage of the simple truncation method is the straightforward calculations that it involves.

D Truncation equations

For more sophisticated truncation techniques employed here, we need to use truncation equations that express the higher moments in terms of the lower moments. Here we present these truncation equations for the central moment truncation method and the cumulant truncation method.

D.1 Truncation equations for $x_{\alpha\beta}$ where $\alpha + \beta = 3$

These two methods have the same truncation equations for $x_{\alpha\beta}$, where $\alpha + \beta = 3$:

$$2x_{10}^3 - 3x_{10}x_{20} + x_{30} = 0 \quad (73)$$

$$2x_{01}x_{10}^2 - 2x_{10}x_{11} - x_{01}x_{20} + x_{21} = 0 \quad (74)$$

$$2x_{10}x_{01}^2 - 2x_{01}x_{11} - x_{10}x_{02} + x_{12} = 0 \quad (75)$$

$$2x_{01}^3 - 3x_{02}x_{01} + x_{03} = 0 \quad (76)$$

Because $x_{10}, x_{01} = O(1)$ and $x_{20}, x_{11}, x_{02} = O\left(\frac{1}{\varepsilon}\right)$, we can see that $x_{\alpha\beta} = O\left(\frac{1}{\varepsilon}\right)$ with $\alpha + \beta = 3$.

D.2 Truncation equations for $x_{\alpha\beta}$ where $\alpha + \beta = 4, 5$

When $\alpha + \beta = 4, 5$, central moment closure method has different truncation equations compared to cumulant closure method.

(a). Central moment closure method

$$-3x_{01}x_{10}^3 + 3x_{10}^2x_{11} + 3x_{01}x_{10}x_{20} - 3x_{10}x_{21} - x_{01}x_{30} + x_{31} = 0$$

$$-3x_{01}^2x_{10}^2 + x_{02}x_{10}^2 + 4x_{01}x_{10}x_{11} - 2x_{10}x_{12} + x_{01}^2x_{20} - 2x_{01}x_{21} + x_{22} = 0$$

$$-3x_{10}x_{01}^3 + 3x_{01}^2x_{11} + 3x_{01}x_{10}x_{02} - 3x_{01}x_{12} - x_{10}x_{03} + x_{13} = 0$$

$$4x_{01}^2x_{10}^3 - x_{02}x_{10}^3 - 6x_{01}x_{10}^2x_{11} + 3x_{10}^2x_{12} - 3x_{01}^2x_{10}x_{20} + 6x_{01}x_{10}x_{21} - 3x_{10}x_{22} + x_{01}^2x_{30} - 2x_{01}x_{31} + x_{32} = 0$$

$$4x_{10}^2x_{01}^3 - x_{20}x_{01}^3 - 6x_{10}x_{01}^2x_{11} + 3x_{01}^2x_{21} - 3x_{10}^2x_{01}x_{02} + 6x_{01}x_{10}x_{12} - 3x_{01}x_{22} + x_{10}^2x_{03} - 2x_{10}x_{13} + x_{23} = 0$$

Because $x_{10}, x_{01} = O(1)$, $x_{20}, x_{11}, x_{02} = O\left(\frac{1}{\varepsilon}\right)$ and $x_{30}, x_{21}, x_{12}, x_{03} = O\left(\frac{1}{\varepsilon}\right)$, by the above truncation equations, $x_{\alpha\beta} = O\left(\frac{1}{\varepsilon}\right)$, for $\alpha + \beta = 4, 5$.

(b). Cumulant closure method

The truncation equations for cumulant closure method for $x_{\alpha\beta}$ where $\alpha + \beta = 4, 5$ are:

$$x_{31} - 3x_{10}x_{21} - x_{01}x_{30} - 3x_{20}x_{11} + 6x_{20}x_{01}x_{10} + 6x_{11}x_{10}^2 - 6x_{10}^3x_{01} = 0$$

$$x_{22} - 2x_{10}x_{12} - 2x_{01}x_{21} - x_{20}x_{02} - 2x_{11}^2 + 2x_{20}x_{01}^2 + 2x_{02}x_{10}^2 + 8x_{11}x_{10}x_{01} - 6x_{10}^2x_{01}^2 = 0$$

$$x_{13} - 3x_{01}x_{12} - x_{10}x_{03} - 3x_{02}x_{11} + 6x_{02}x_{01}x_{10} + 6x_{11}x_{01}^2 - 6x_{01}^3x_{10} = 0$$

$$\begin{aligned} & 24x_{01}^2x_{10}^3 - 6x_{02}x_{10}^3 \\ & - 36x_{01}x_{10}^2x_{11} \\ & + 12x_{10}x_{11}^2 \\ & + 6x_{10}^2x_{12} \\ & - 18x_{01}^2x_{10}x_{20} \\ & + 6x_{02}x_{10}x_{20} \\ & + 12x_{01}x_{11}x_{20} \\ & - 3x_{12}x_{20} \\ & + 12x_{01}x_{10}x_{21} \\ & - 6x_{11}x_{21} \\ & - 3x_{10}x_{22} \\ & + 2x_{01}^2x_{30} \\ & - x_{02}x_{30} - 2x_{01}x_{31} + x_{32} = 0 \end{aligned}$$

$$\begin{aligned} & 24x_{10}^2x_{01}^3 - 6x_{20}x_{01}^3 \\ & - 36x_{10}x_{01}^2x_{11} \\ & + 12x_{01}x_{11}^2 \\ & + 6x_{01}^2x_{21} \\ & - 18x_{10}^2x_{01}x_{02} \\ & + 6x_{20}x_{01}x_{02} \\ & + 12x_{10}x_{11}x_{02} \\ & - 3x_{21}x_{02} \\ & + 12x_{01}x_{10}x_{12} \\ & - 6x_{11}x_{12} \\ & - 3x_{01}x_{22} \\ & + 2x_{10}^2x_{03} \\ & - x_{20}x_{03} - 2x_{10}x_{13} + x_{23} = 0 \end{aligned}$$

By these truncation equations, we get $x_{\alpha\beta} = O\left(\frac{1}{\varepsilon^2}\right)$, for $\alpha + \beta = 4, 5$.

E Moment Equations

E.1 Linear Noise Approximation

To find the next order correction, we will collect terms up to $\mathcal{O}(\varepsilon^{3/2})$ in the master equation expansion. After rescaling time by $T = L_0 \tau = L_0 \varepsilon t$, we will extend the Fokker-Planck equation (56) to $\mathcal{O}(\varepsilon^{1/2})$. Then, we can obtain the moment equations by integrations. Here we will only illustrate several of them:

$$\frac{d\langle \xi \rangle}{dT} = -2(P_x \langle \xi \rangle + P_y \langle \eta \rangle) + \varepsilon^{1/2} \left[-2\left(P_{xx} + \frac{2L_x P_x}{L_0}\right) \langle \xi^2 \rangle - \left(P_{yy} + \frac{2L_y P_y}{L_0}\right) \langle \eta^2 \rangle - 2\left(P_{xy} + \frac{L_x P_y + L_y P_x}{L_0}\right) \langle \xi \eta \rangle \right], \quad (77)$$

$$\begin{aligned} \frac{d\langle \xi^2 \rangle}{dT} = & -4(P_x \langle \xi^2 \rangle + P_y \langle \xi \eta \rangle) \\ & + 1 + \varepsilon^{1/2} \left[-2\left(P_{xx} + \frac{2L_x P_x}{L_0}\right) \langle \xi^3 \rangle - 2\left(P_{yy} + \frac{2L_y P_y}{L_0}\right) \langle \xi \eta^2 \rangle - 4\left(P_{xy} + \frac{L_x P_y + L_y P_x}{L_0}\right) \langle \xi^2 \eta \rangle + \frac{L_x \langle \xi \rangle}{L_0} + \frac{L_y \langle \eta \rangle}{L_0} \right], \end{aligned} \quad (78)$$

$$\begin{aligned} \frac{d\langle \xi^3 \rangle}{dT} = & -6P_x \langle \xi^3 \rangle \\ & -6P_y \langle \xi^2 \eta \rangle \\ & + 3\langle \xi \rangle + \varepsilon^{1/2} \left[-3\langle \xi^4 \rangle \left(P_{xx} + \frac{2L_x P_x}{L_0}\right) - 3\langle \xi^2 \eta^2 \rangle \left(P_{yy} + \frac{2L_y P_y}{L_0}\right) - 6\langle \xi^3 \eta \rangle \left(P_{xy} + \frac{L_x P_y + L_y P_x}{L_0}\right) + 3\langle \xi^2 \rangle \frac{L_x}{L_0} + 3\langle \xi \eta \rangle \frac{L_y}{L_0} \right], \end{aligned} \quad (79)$$

$$\frac{d\langle \xi^4 \rangle}{dT} = -8P_x \langle \xi^4 \rangle - 8P_y \langle \xi^3 \eta \rangle + 6\langle \xi^2 \rangle. \quad (80)$$

E.2 Simple Truncation

Correspond with (77) – (80), we have:

$$\begin{aligned}
 & \mathbf{(1, 0)} \\
 & -2\varepsilon [L_0 (P_y x_{01} + P_x x_{10})] + [-2L_y (P_y x_{02} + P_x x_{11}) - 2L_x (P_y x_{11} + P_x x_{20}) - L_0 (P_{yy} x_{02} + 2P_{xy} x_{11} + P_{xx} x_{20})] \varepsilon^2 = 0,
 \end{aligned}
 \tag{81}$$

$$\begin{aligned}
 & \mathbf{(2, 0)} \\
 & L_0 - 4L_0 (P_y x_{11} + P_x x_{20}) \varepsilon \\
 & + (L_y x_{01} + L_x x_{10}) \varepsilon + \frac{1}{2} \varepsilon^2 [-4L_0 (P_{xx} + \frac{2L_x P_x}{L_0}) x_{30} - 4L_0 (P_{yy} + \frac{2L_y P_y}{L_0}) x_{12} - 8L_0 (P_{xy} + \frac{L_y P_x + L_x P_y}{L_0}) x_{21} + L_{yy} x_{02} + L_{xx} x_{20} \\
 & + L_{xy} x_{11}] = 0,
 \end{aligned}
 \tag{82}$$

$$\begin{aligned}
 & \mathbf{(3, 0)} \\
 & 3L_0 x_{10} + \varepsilon (3L_y x_{11} + 3L_x x_{20} - 2L_0 P_y x_{01} - 6L_0 P_y x_{21} - 2L_0 P_x x_{10} - 6L_0 P_x x_{30}) \\
 & + \varepsilon^2 (\\
 & \quad - 2L_x P_y x_{11} \\
 & \quad + \frac{3}{2} L_{yy} x_{12} \\
 & \quad - 2L_x P_x x_{20} \\
 & \quad + 3L_{xy} x_{21} \\
 & \quad + \frac{3}{2} L_{xx} x_{30} \\
 & \quad - 6L_x P_y x_{31} - 2L_y P_y x_{02} - 6L_y P_y x_{22} - 2L_y P_x x_{11} - 6L_y P_x x_{31} - 6L_x P_x x_{40} - L_0 P_{yy} x_{02} - 3L_0 P_{yy} x_{22} - 2L_0 P_{xy} x_{11}) = 0,
 \end{aligned}
 \tag{83}$$

$$\begin{aligned}
 & \mathbf{(4, 0)} \\
 & L_0 + 6L_0 x_{20} + \varepsilon (L_y x_{01} + 6L_y x_{21} + L_x x_{10} + 6L_x x_{30} - 8L_0 P_y x_{11} - 8L_0 P_y x_{31} - 8L_0 P_x x_{20} - 8L_0 P_x x_{40}) = 0.
 \end{aligned}
 \tag{84}$$

By using (62), equations (77)–(79) coincide with equations (81)–(83) to $\mathcal{O}(\varepsilon^{1/2})$, and equations (80) agrees with (84) to $\mathcal{O}(1)$. In fact, all the first, second, and third order moment equations from the two methods coincide to $\mathcal{O}(\varepsilon^{1/2})$, and so are the fourth order moment equations to $\mathcal{O}(1)$.

E.3 Methodology

To find $\mathcal{O}(\varepsilon^{1/2})$ terms of the first order moments $\langle \xi \rangle$, $\langle \eta \rangle$, (77) shows we only need $\mathcal{O}(1)$ terms of $\langle \xi^2 \rangle$, $\langle \eta^2 \rangle$, $\langle \xi \eta \rangle$. Since the two methods give the same results to the leading order of

the second order moments, $O(\epsilon^{1/2})$ terms of the first order moments (which turn out to be non-zero) are the same for the two methods. It follows that the next order correction to the first order moments are the same for the two methods.

To find $O(\epsilon^{1/2})$ terms of the second order moments $\langle \xi^2 \rangle, \langle \xi \eta \rangle, \langle \eta^2 \rangle$, (78) shows we only need $O(1)$ terms of $\langle \xi^a \eta^b \rangle$ for $a + b = 3$, and $O(1)$ terms of $\langle \xi \rangle, \langle \eta \rangle$. Through (79) and the other third order moment equations, we have $O(1)$ terms of $\langle \xi^a \eta^b \rangle$ for $a + b = 3$ are all zero since $O(1)$ terms of $\langle \xi \rangle, \langle \eta \rangle$ are zero. It follows that $O(\epsilon^{1/2})$ terms of the second order moments $\langle \xi^2 \rangle, \langle \xi \eta \rangle, \langle \eta^2 \rangle$ are all zero.

To $O(\epsilon)$ terms of the second order moments $\langle \xi^2 \rangle, \langle \xi \eta \rangle, \langle \eta^2 \rangle$, (78) shows we need $O(\epsilon^{1/2})$ terms of $\langle \xi^a \eta^b \rangle$ for $a + b = 3$, and $O(\epsilon^{1/2})$ terms of $\langle \xi \rangle, \langle \eta \rangle$. We know $O(\epsilon^{1/2})$ terms of $\langle \xi \rangle, \langle \eta \rangle$ are the same for the two methods. To find $O(\epsilon^{1/2})$ terms of $\langle \xi^a \eta^b \rangle$ for $a + b = 4$, and $O(1)$ terms of $\langle \xi^a \eta^b \rangle$ for $a + b = 2$ other (79). Now, (80) and the other fourth order moment equations show that $O(1)$ terms of $\langle \xi^a \eta^b \rangle$ coincide for the two methods, for $a + b = 4$, since both methods give the same results to the leading order of the second order moments. It follows that the next order corrections to the second order moments (which turn out to be non-zero) are the same for the two methods.

By using (62), we have just shown that Van Kampen method and simple truncation give the same results to the next order correction of the mean and variance of the cell population. In particular, let I and i_0 denote the number of stem cells and its steady state; respectively, then we have the following results:

$$\langle \xi \rangle = a_{10}^1 \epsilon^{1/2} + \dots; \quad (85)$$

$$\langle \xi^2 \rangle = a_{20}^1 + a_{20}^2 \epsilon + \dots; \quad (86)$$

$$E[I] = i_0 + a_{10}^1 + \dots; \quad (87)$$

$$Var[I] = \frac{a_{20}^1}{\epsilon} + a_{20}^2 - \left(a_{10}^1\right)^2 + \dots. \quad (88)$$

References

1. Loeffler M, Roeder I. Tissue stem cells: definition, plasticity, heterogeneity, self-organization and models—a conceptual approach. *Cells Tissues Organs* (Print). 2002; 171:8–26. [PubMed: 12021488]
2. d’Inverno, M., Theise, ND., Prophet, J. *Tissue Stem Cells: Biology and Applications*. 2nd. Marcell Dekker Inc; New York, NY: 2006. Mathematical modeling of stem cells: a complexity primer for the stem cell biologist.

3. Roeder, I., Galle, J., Loeffler, M. Theoretical concepts of tissue stem cell organization. Taylor and Francis; New York, London: 2006.
4. Tomlinson IP, Bodmer WF. Failure of programmed cell death and differentiation as causes of tumors: some simple mathematical models. *Proc Natl Acad Sci USA*. Nov.1995 92:11130–11134. [PubMed: 7479951]
5. d’Onofrio A, Tomlinson IP. A nonlinear mathematical model of cell turnover, differentiation and tumorigenesis in the intestinal crypt. *J Theor Biol*. Feb.2007 244:367–374. [PubMed: 17049944]
6. Johnston MD, Edwards CM, Bodmer WF, Maini PK, Chapman SJ. Mathematical modeling of cell population dynamics in the colonic crypt and in colorectal cancer. *Proc Natl Acad Sci USA*. Mar. 2007 104:4008–4013. [PubMed: 17360468]
7. Boman BM, Fields JZ, Cavanaugh KL, Guetter A, Runquist OA. How dysregulated colonic crypt dynamics cause stem cell overpopulation and initiate colon cancer. *Cancer Res*. May.2008 68:3304–3313. [PubMed: 18451157]
8. Hardy K, Stark J. Mathematical models of the balance between apoptosis and proliferation. *Apoptosis*. Aug.2002 7:373–381. [PubMed: 12101397]
9. Yatabe Y, Tavare S, Shibata D. Investigating stem cells in human colon by using methylation patterns. *Proc Natl Acad Sci USA*. Sep.2001 98:10839–10844. [PubMed: 11517339]
10. Ganguly R, Puri IK. Mathematical model for the cancer stem cell hypothesis. *Cell proliferation*. 2006; 39(1):3–14. [PubMed: 16426418]
11. Ganguly R, Puri IK. Mathematical model for chemotherapeutic drug efficacy in arresting tumour growth based on the cancer stem cell hypothesis. *Cell proliferation*. 2007; 40(3):338–354. [PubMed: 17531079]
12. Ashkenazi, Rina, Gentry, Sara N., Jackson, Trachette L. Pathways to tumorigenesis modeling mutation acquisition in stem cells and their progeny. *Neoplasia* (New York, NY). 2008; 10(11): 1170.
13. Michor F. Mathematical models of cancer stem cells. *Journal of Clinical Oncology*. 2008; 26(17): 2854–2861. [PubMed: 18539964]
14. Tomasetti C, Levy D. Role of symmetric and asymmetric division of stem cells in developing drug resistance. *Proceedings of the National Academy of Sciences*. 2010; 107(39):16766–16771.
15. Enderling H, Hahnfeldt P. Cancer stem cells in solid tumors: Is ‘evading apoptosis’ a hallmark of cancer? *Progress in Biophysics and Molecular Biology*. 2011
16. Mangel M, Bonsall MB. Phenotypic evolutionary models in stem cell biology: replacement, quiescence, and variability. *PLoS ONE*. 2008; 3:e1591. [PubMed: 18270578]
17. Glauche I, Cross M, Loeffler M, Roeder I. Lineage specification of hematopoietic stem cells: mathematical modeling and biological implications. *Stem Cells*. Jul.2007 25:1791–1799. [PubMed: 17412891]
18. Michor, Franziska, Hughes, Timothy P., Iwasa, Yoh, Branford, Susan, Shah, Neil P., Sawyers, Charles L., Nowak, Martin A. Dynamics of chronic myeloid leukaemia. *Nature*. 2005; 435(7046): 1267–1270. [PubMed: 15988530]
19. Colijn C, Mackey MC. A mathematical model of hematopoiesis–I. Periodic chronic myelogenous leukemia. *J Theor Biol*. Nov.2005 237:117–132. [PubMed: 15975596]
20. Adimy M, Crauste F, Ruan S. Modelling hematopoiesis mediated by growth factors with applications to periodic hematological diseases. *Bull Math Biol*. Nov.2006 68:2321–2351. [PubMed: 17086497]
21. Marciniak-Czochra, Anna, Stiehl, Thomas, Ho, Anthony D., Jäger, Willi, Wagner, Wolfgang. Modeling of asymmetric cell division in hematopoietic stem cells-regulation of self-renewal is essential for efficient repopulation. *Stem cells and development*. 2009; 18(3):377–386. [PubMed: 18752377]
22. Piotrowska, MJ., Enderling, H., van der Heiden, U., Mackey, MC. Mathematical modeling of stem cells related to cancer. In: Dittmar, T., Zanker, Kurt, editors. *Cancer and stem cells*. Nova Science Publishers, Inc; 2008.
23. Lander AD, Gokoffski KK, Wan FY, Nie Q, Calof AL. Cell lineages and the logic of proliferative control. *PLoS Biol*. Jan.2009 7:e15. [PubMed: 19166268]

24. Lo WC, Chou CS, Gokoffski KK, Wan FY, Lander AD, Calof AL, Nie Q. Feedback regulation in multistage cell lineages. *Math Biosci Eng.* Jan.2009 6:59–82. [PubMed: 19292508]
25. Nakata, Yukihiko, Getto, Philipp, Marciniak-Czochra, Anna, Alarcón, Tomás. Stability analysis of multi-compartment models for cell production systems. *Journal of Biological Dynamics.* 2012; 6(sup1):2–18.
26. Arino, Ovide, Kimmel, Marek. Stability analysis of models of cell production systems. *Mathematical Modelling.* 1986; 7(9):1269–1300.
27. Stiehl, Thomas, Marciniak-Czochra, Anna. Characterization of stem cells using mathematical models of multistage cell lineages. *Mathematical and Computer Modelling.* 2011; 53(7):1505–1517.
28. Enderling, Heiko, Chaplain, Mark AJ., Anderson, Alexander RA., Vaidya, Jayant S. A mathematical model of breast cancer development, local treatment and recurrence. *Journal of theoretical biology.* 2007; 246(2):245–259. [PubMed: 17289081]
29. Enderling, Heiko, Anderson, Alexander RA., Chaplain, Mark AJ., Beheshti, Afshin, Hlatky, Lynn, Hahnfeldt, Philip. Paradoxical dependencies of tumor dormancy and progression on basic cell kinetics. *Cancer research.* 2009; 69(22):8814–8821. [PubMed: 19887613]
30. Enderling H, Park D, Hlatky L, Hahnfeldt P. The importance of spatial distribution of stemness and proliferation state in determining tumor radioresponse. *Math Model Nat Phenom.* 2009; 4(3):117–133.
31. Enderling H, Hlatky L, Hahnfeldt P. Migration rules: tumours are conglomerates of self-metastases. *British journal of cancer.* 2009; 100(12):1917–1925. [PubMed: 19455139]
32. Agur Z, Daniel Y, Ginosar Y. The universal properties of stem cells as pinpointed by a simple discrete model. *Journal of mathematical biology.* 2002; 44(1):79–86. [PubMed: 11942526]
33. Dingli, David, Traulsen, Arne, Pacheco, Jorge M. Stochastic dynamics of hematopoietic tumor stem cells. *Cell Cycle.* 2007; 6(4):461–466. [PubMed: 17329969]
34. Sun Z, Komarova NL. Stochastic modeling of stem-cell dynamics with control. *Mathematical Biosciences.* 2012; 240:231–240. [PubMed: 22960597]
35. Sun Z, Komarova NL. Stochastic control of proliferation and differentiation in stem cell dynamics. *Jour Math Biol.* 2012
36. van Kampen NG. A power series expansion of the master equation. *Canadian Journal of Physics.* 1961; 39(4):551–567.
37. Moran, Patrick Alfred Pierce. The statistical processes of evolutionary theory. Clarendon Press; Oxford University Press; New York: 1962.
38. Komarova, Natalia L. Principles of regulation of self-renewing cell lineages. *PloS one.* 2013; 8(9):e72847. [PubMed: 24019882]
39. Van Kampen, Nicolaas Godfried. Stochastic processes in physics and chemistry. Elsevier; Amsterdam, Boston, London, Tokyo: 1992.
40. Gardiner, Crispin W., et al. Handbook of stochastic methods. Vol. 3. Springer; Berlin: 1985.
41. Wallace EWJ, Gillespie DT, Sanft KR, Petzold LR. Linear noise approximation is valid over limited times for any chemical system that is sufficiently large. *Systems Biology, IET.* 2012; 6(4): 102–115.
42. Doering, Charles R., Sargsyan, Khachik V., Sander, Leonard M. Extinction times for birth-death processes: Exact results, continuum asymptotics, and the failure of the fokker–planck approximation. *Multiscale Modeling & Simulation.* 2005; 3(2):283–299.
43. Assaf, Michael, Meerson, Baruch. Spectral theory of metastability and extinction in birth-death systems. *Physical review letters.* 2006; 97(20):200602. [PubMed: 17155670]
44. Rodriguez-Brenes IA, Komarova NL, Wodarz D. Evolutionary dynamics of feedback escape and the development of stem-cell–driven cancers. *Proceedings of the National Academy of Sciences.* 2011; 108(47):18983–18988.
45. Scadden DT. The stem-cell niche as an entity of action. *Nature.* 2006; 441(7097):1075–1079. [PubMed: 16810242]
46. Walker MR, Patel KK, Stappenbeck TS. The stem cell niche. *The Journal of pathology.* 2009; 217(2):169–180. [PubMed: 19089901]

47. Spiegel A, Kalinkovich A, Shivtiel S, Kollet O, Lapidot T. Stem cell regulation via dynamic interactions of the nervous and immune systems with the microenvironment. *Cell Stem Cell*. 2008; 3(5):484–492. [PubMed: 18983964]
48. Discher DE, Mooney DJ, Zandstra PW. Growth factors, matrices, and forces combine and control stem cells. *Science*. 2009; 324(5935):1673–1677. [PubMed: 19556500]
49. Gangaraju VK, Lin H. Micronas: key regulators of stem cells. *Nature Reviews Molecular Cell Biology*. 2009; 10(2):116–125. [PubMed: 19165214]
50. Heinrich EM, Dimmeler S. Micronas and stem cells control of pluripotency, reprogramming, and lineage commitment. *Circulation research*. 2012; 110(7):1014–1022. [PubMed: 22461365]
51. Takashima S, Hartenstein V. Genetic control of intestinal stem cell specification and development: a comparative view. *Stem Cell Reviews and Reports*. 2012;1–12. [PubMed: 21537995]
52. Liu M, Pleasure SJ, Collins AE, Noebels JL, Naya FJ, Tsai MJ, Lowenstein DH. Loss of beta2/neurod leads to malformation of the dentate gyrus and epilepsy. *Proceedings of the National Academy of Sciences*. 2000; 97(2):865–870.
53. Simmons CA, Matlis S, Thornton AJ, Chen S, Wang CY, Mooney DJ. Cyclic strain enhances matrix mineralization by adult human mesenchymal stem cells via the extracellular signal-regulated kinase (erk1/2) signaling pathway. *Journal of biomechanics*. 2003; 36(8):1087–1096. [PubMed: 12831733]
54. Alvarez-Buylla A, Lim DA. For the long run: maintaining germinal niches in the adult brain. *Neuron*. 2004; 41(5):683–686. [PubMed: 15003168]
55. Saha S, Ji L, de Pablo JJ, Palecek SP. Inhibition of human embryonic stem cell differentiation by mechanical strain. *Journal of cellular physiology*. 2006; 206(1):126–137. [PubMed: 15965964]
56. Lien WH, Klezovitch O, Fernandez TE, Delrow J, Vasioukhin V. {alpha} e-catenin controls cerebral cortical size by regulating the hedgehog signaling pathway. *Science's STKE*. 2006; 311(5767):1609.
57. Adams GB, Scadden DT. A niche opportunity for stem cell therapeutics. *Gene therapy*. 2007; 15(2):96–99. [PubMed: 18004404]
58. Dehay C, Kennedy H. Cell-cycle control and cortical development. *Nature Reviews Neuroscience*. 2007; 8(6):438–450. [PubMed: 17514197]
59. Orford KW, Scadden DT. Deconstructing stem cell self-renewal: genetic insights into cell-cycle regulation. *Nature Reviews Genetics*. 2008; 9(2):115–128.
60. Saha S, Ji L, De Pablo JJ, Palecek SP. Tgf [beta]/activin/nodal pathway in inhibition of human embryonic stem cell differentiation by mechanical strain. *Biophysical journal*. 2008; 94(10):4123–4133. [PubMed: 18234825]
61. Sen B, Xie Z, Case N, Ma M, Rubin C, Rubin J. Mechanical strain inhibits adipogenesis in mesenchymal stem cells by stimulating a durable β -catenin signal. *Endocrinology*. 2008; 149(12):6065–6075. [PubMed: 18687779]
62. Guilak F, Cohen DM, Estes BT, Gimple JM, Liedtke W, Chen CS. Control of stem cell fate by physical interactions with the extracellular matrix. *Cell Stem Cell*. 2009; 5(1):17–26. [PubMed: 19570510]
63. Lavado A, Lagutin OV, Chow LML, Baker SJ, Oliver G. Prox1 is required for granule cell maturation and intermediate progenitor maintenance during brain neurogenesis. *PLoS biology*. 2010; 8(8):e1000460. [PubMed: 20808958]
64. de Graaf CA, Kauppi M, Baldwin T, Hyland CD, Metcalf D, Willson TA, Carpinelli MR, Smyth GK, Alexander WS, Hilton DJ. Regulation of hematopoietic stem cells by their mature progeny. *Proceedings of the National Academy of Sciences*. 2010; 107(50):21689–21694.
65. Li L, Clevers H. Coexistence of quiescent and active adult stem cells in mammals. *Science*. 2010; 327(5965):542–545. [PubMed: 20110496]
66. Salomoni P, Calegari F. Cell cycle control of mammalian neural stem cells: putting a speed limit on g1. *Trends in cell biology*. 2010; 20(5):233–243. [PubMed: 20153966]
67. Hsieh J. Orchestrating transcriptional control of adult neurogenesis. *Genes & Development*. 2012; 26(10):1010–1021. [PubMed: 22588716]
68. Ordóñez-Morán P, Huelsken J. Lrig1: a new master regulator of epithelial stem cells. *The EMBO Journal*. 2012

69. Nusse R. Wnt signaling and stem cell control. *Cell research*. 2008; 18(5):523–527. [PubMed: 18392048]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Highlights

- We study stochastic population dynamics of stem cells and differentiated cells
- We ask how different types of control affect means and variances of cell numbers
- Simple explicit expressions for the means and variances are derived
- The method is general and works for any functional form of the controls

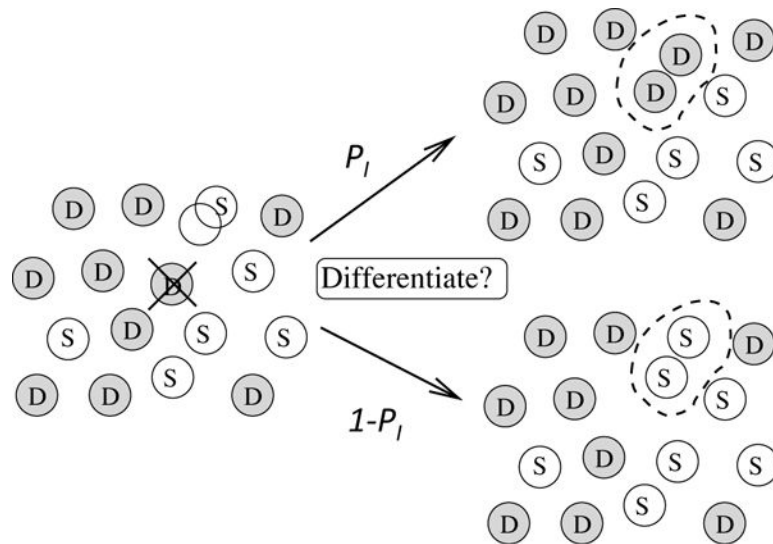


Figure 1.

A schematic showing one step of the update for the constant total population model. Circles represent stem cells (“S”) and differentiated cells (“D”). Following a death of a randomly chosen differentiated cells, one of the stem cells is chosen for division. With probability p_I (where I is the current number of stem cells in the system) it will differentiate, that is, divide into two daughter differentiated cells. With probability $1 - p_b$ it will proliferate, that is, divide into two stem cells.

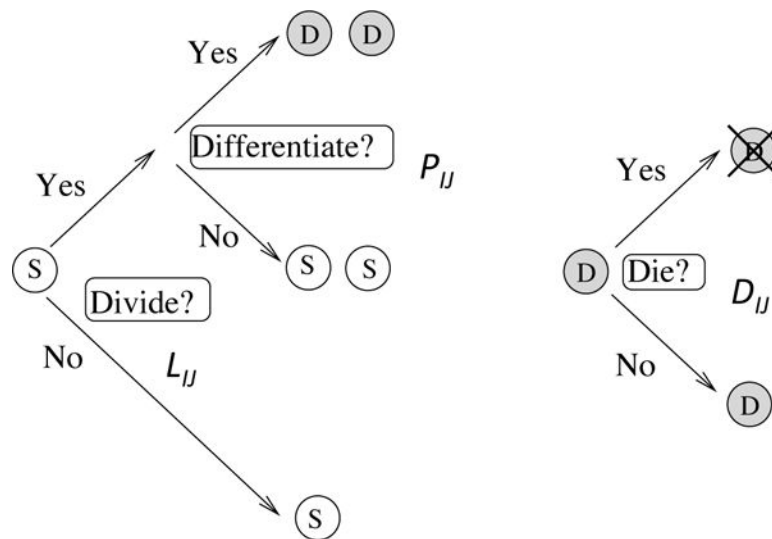


Figure 2.

A non-constant total population process. A schematics showing the cellular processes and their probabilities. Circles represent stem cells (“S”) and differentiated cells (“D”). A stem cell divides with probability $L_{I,J}$ where I and J are the current populations of stem and differentiated cells respectively. The division can be a differentiation event (with probability $P_{I,J}$) or a proliferation event (with probability $1 - P_{I,J}$). A differentiated cell dies with the rate $D_{I,J}$.

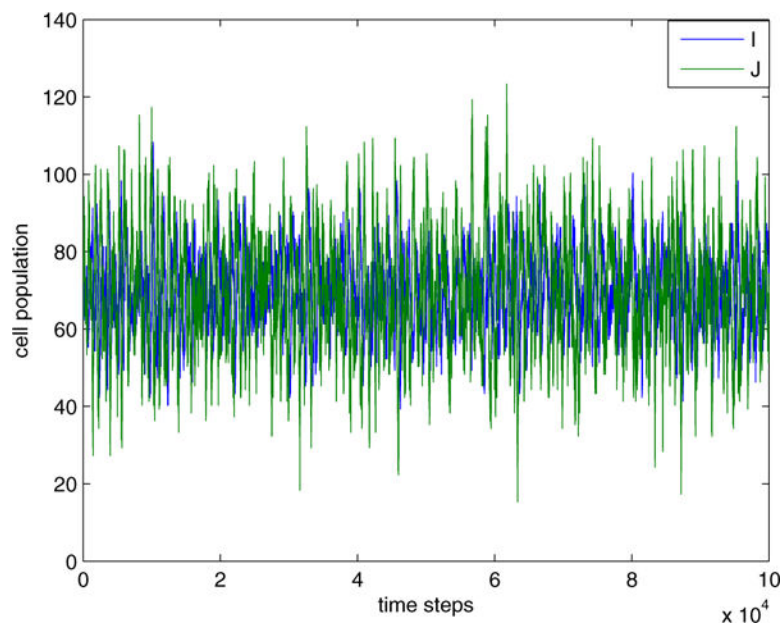


Figure 3. A numerical simulation of the system in (40) with $\varepsilon = 0.01$ ran for 10^5 time steps. ('I') stands for the stem cell population, and ('J') stands for the differentiated cell population.

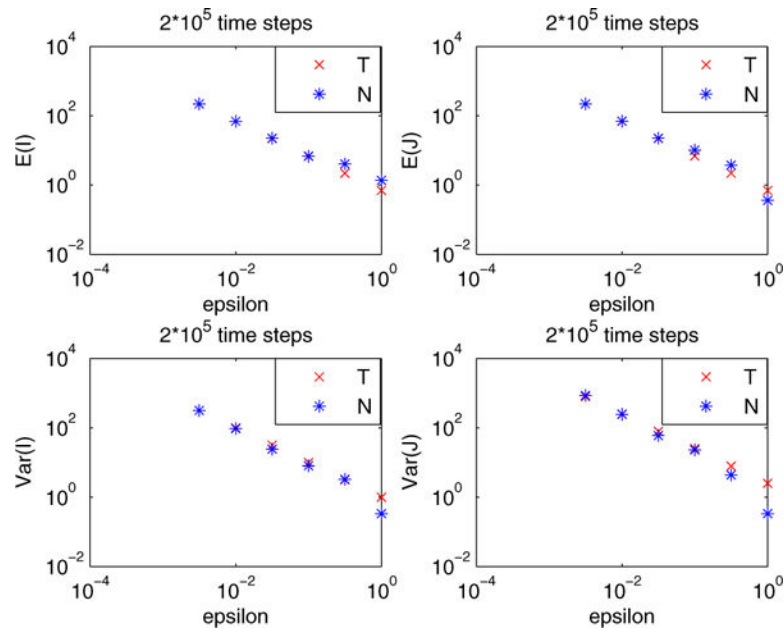


Figure 4.

The behavior of the means and the variances of the cell population described by equation (40). The analytical results given by equations (41–44) (X's) are compared with the values obtained by numerical simulations (stars), for different values of ε . ('T') stands for the theoretical results, and ('N') stands for the numerical results.

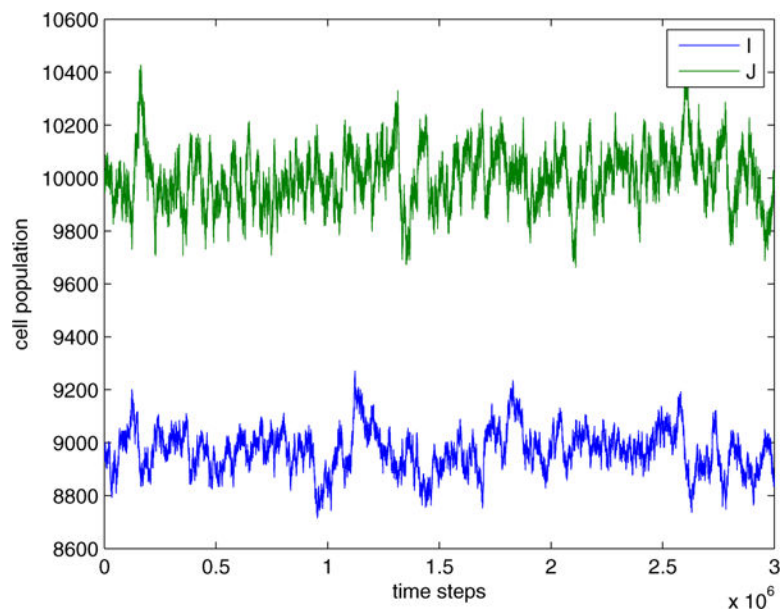


Figure 5.

A typical numerical simulation of example (45) with $\varepsilon = 10^{-4}$ and $3 \cdot 10^6$ time steps. ('I') stands for the stem cell population, and ('J') stands for the differentiated cell population.

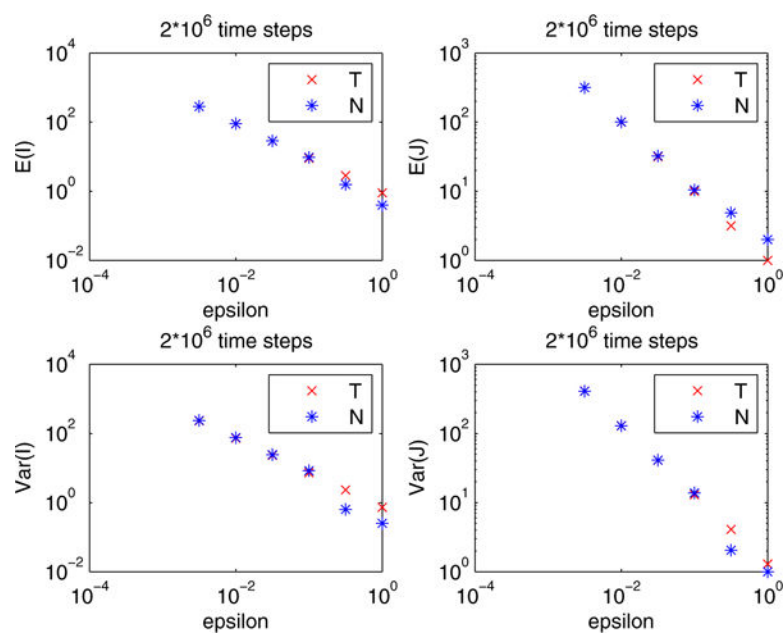


Figure 6. Same as in figure 4, except the means and variances are calculated of system (45).

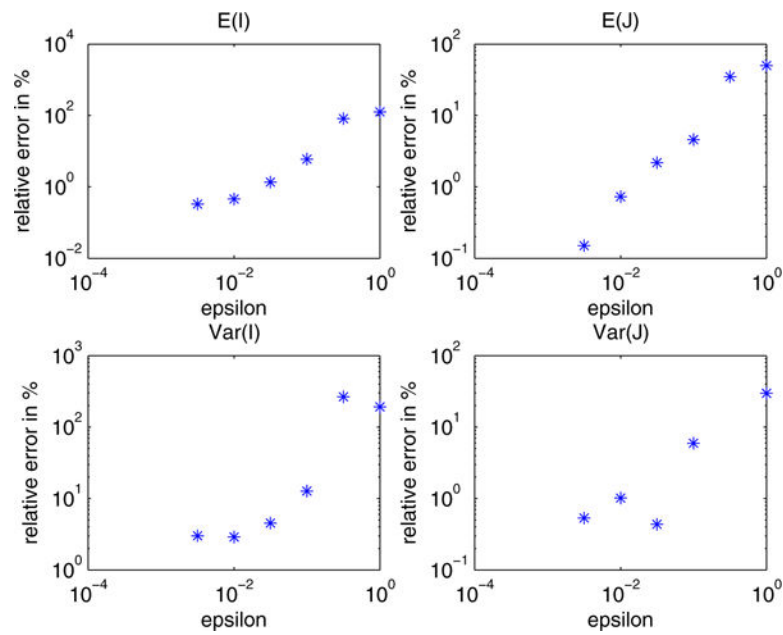


Figure 7.

The behavior of the relative error of the means and the variances for different values of ε .

We used the relative error = $\frac{|theoretical\ result - numerical\ result|}{numerical\ result} \times 100\%$.