PHYSICAL REVIEW E **105**, 054409 (2022)

Effect of dedifferentiation on noise propagation in cellular hierarchy

Yuman Wang,[1] Jintong Zhao,[1] Hye Jin Park,[2,3] and Da Zhou[1,*]
[1]*School of Mathematical Sciences, Xiamen University, Xiamen 361005, People's Republic of China*
[2]*Asia Pacific Center for Theoretical Physics, Pohang 37673, Republic of Korea*
[3]*Department of Physics, Inha University, Incheon 22212, Republic of Korea*

(Received 26 August 2021; revised 25 January 2022; accepted 25 April 2022; published 18 May 2022)

Many fast renewing tissues have a hierarchical structure. Tissue-specific stem cells are at the root of this cellular hierarchy, which give rive to a whole range of specialized cells via cellular differentiation. However, increasing evidence shows that the hierarchical structure can be broken due to cellular dedifferentiation in which cells at differentiated stages can revert to the stem cell stage. Dedifferentiation has significant impacts on many aspects of hierarchical tissues. Here we investigate the effect of dedifferentiation on noise propagation by developing a stochastic model composed of different cell types. The moment equations are derived, via which we systematically investigate how the noise in the cell number is changed by dedifferentiation. Our results suggest that dedifferentiation have different effects on the noises in the numbers of stem cells and nonstem cells. Specifically, the noise in the number of stem cells is significantly reduced by increasing dedifferentiation probability. Due to the dual effect of dedifferentiation on nonstem cells, however, more complex changes could happen to the noise in the number of nonstem cells by increasing dedifferentiation probability. Furthermore, it is found that even though dedifferentiation could turn part of the noise propagation process into a noise-amplifying step, it is very unlikely to turn the entire process into a noise-amplifying cascade.

DOI: 10.1103/PhysRevE.105.054409

I. INTRODUCTION

In multicellular organisms, many fast renewing tissues form a hierarchical structure [1,2]. Very few long lived tissue-specific stem cells are at the root of this cellular hierarchy [3,4]. They give rise to more specialized and short lived cells via cellular differentiation. It is argued that many cancers could also be characterized by this hierarchical structure, in which cancer stem cells possess very similar functions as tissue-specific stem cells in normal tissues [5,6]. The hierarchical architectures of both normal and cancerous tissues propose an irreversible transition relationship of cells at different stages of differentiation. That is, cells at less differentiated stages can generate more differentiated cells, but not vice versa.

There is growing evidence that the irreversible relationship can be broken due to cellular dedifferentiation [7–10]. In the process of dedifferentiation, cells at more differentiated stages can, under some circumstances, return to a less differentiated stage, or even the stem cell stage. In recent ten years, special attention has been paid to the effect of cellular dedifferentiation on different aspects of hierarchical architecture by using mathematical models [11]. Previous work has, e.g., considered the effects of dedifferentiation on the waiting time to mutation acquisition [12], the mutant fixation [13], the radiation sensitivity [14], the transient dynamics [15], and the phenotypic equilibrium [8].

Due to the stochastic nature of biochemical reactions and fluctuating environments, the cellular system is inherently noisy [16,17]. Increasing evidence shows that noise is not only a major source of phenotypic variation [18], but also associated with many important biological functions at various scales of biological organization [19–21]. In multicellular tissues, the cell population size of different cell types is the most important variable for quantifying the multicellular systems. Even though the cell population size in multicellular systems is normally large and deterministic models are more widely used, stochastic cell population models still receive much attention [22]. First of all, the number of tissue-specific stem cells could be very small, the fluctuation in the number of stem cells is thus significant and informative [23–25]. Moreover, the growth and regeneration of tissues are inevitably affected by various types of noises from biochemical reactions, cell-to-cell interactions, fluctuating microenvironments, etc. Therefore, stochastic cell population models provide powerful approaches that allow to capture and understand the role of various fluctuations in the multicellular systems of interest. In the hierarchical organization of cell populations, a significant problem is to understand how the noise propagates from tissue specific stem cells to terminally differentiated cells. Previous work has considered this issue via developing a noise decomposition method [26,27], in which the fluctuation of a given cell type is decomposed into intrinsic noise, transmitted noise, and conversion noise. Here we are interested in the role of dedifferentiation in multicellular tissues. Note that it has been reported that dedifferentiation influences many different aspects of hierarchical structure, so it is interesting to see how dedifferentiation affects the

---

[*] zhouda@xmu.edu.cn

fluctuations in the number of different types of cells in multicellular tissues.

To address this issue, we develop a multi-compartment model of a stage-structured cell population. Each compartment represents a certain level of cellular differentiation, ranging from stem cells to fully differentiated cells. Given the hierarchically structured cell population, we are concerned about how the noises in the numbers of cells at different stages of cellular differentiation are changed by taking dedifferentiation into account. For the number of each cell type, the square of the coefficient of variation [see Eq. (2)] is used for describing the total noise in the cell number. By using noise decomposition method [26], we can extract the propagated noise from the total noise. Our results show that the effect of dedifferentiation on the noise in the number of stem cells compartment is straightforward. As the dedifferentiation probability increases, the noise in the number of stem cells significantly decreases. For nonstem cells compartments, however, the change of the noise is not as predictable as that of stem cells compartment. Specifically, increasing dedifferentiation probability does not necessarily increase or decrease the noise in the number of nonstem cells. It heavily depends on another parameter which is called reshaping factor characterizing how dedifferentiation reshapes the division pattern of transient amplifying cells. Moreover, we explore the relation of the noises in the number of different cell types via calculating the noise ratios between compartments. We found that, due to the opposite effect of reshaping factor on the noises in the number of transient amplifying cells and fully differentiated cells, a tradeoff phenomenon arises: If the noise is amplified from stem cells to transient amplifying cells, then the noise is more likely to be reduced from transient amplifying cells to fully differentiated cells. In other words, even though dedifferentiation could significantly alter the propagation pattern of noise, it is very unlikely to make the noise progressively amplified along cellular hierarchy. We hope that our work could contribute to the theoretical understanding of the impact of dedifferentiation on hierarchical tissues.

The paper is organized as follows. The model is presented in Sec. II. In Sec. III the moment equations for different cell compartments are derived by using master equation [28] and moment closure approximation [29]. The main results are shown in Sec. IV. The conclusions are included in Sec. V.

II. MODEL

Consider a cell population composed of three compartments. Each compartment represents a certain level of cellular differentiation (see Fig. 1). Specifically, compartment 1 represents stem cells (SCs), compartment 2 represents transient amplifying cells (TACs), and compartment 3 represents terminally differentiated cells (TDCs). Even though the number of cell compartments could be diverse for different tissues, the three-compartment assumption has been widely accepted for abstracting the salient features of hierarchical tissues [27,30].

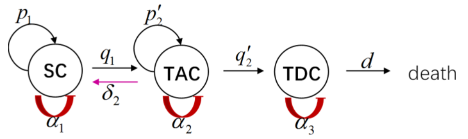

FIG. 1. Representation of our model. Each compartment represents a certain level of cellular differentiation. From left to right, it represents stem cell (SC), transient amplifying cell (TAC), and terminally differentiated cells (TDC), respectively. For each stem cell, it can either give birth to two identical stem cells with probability p_1 or two identical daughter transient amplifying cells with probability q_1. When cell competition happens, one of the two competing stem cells dies at equal chance, in which α_1 is the competition death rate controlling the competition strength. Similar cell division and competition pattern can also happen to transient amplifying cells. Due to dedifferentiation, each transient amplifying cell can also give birth to two daughter stem cells with probability δ_2. For each terminally differentiated cell, it cannot divide and is removed from the tissue at rate d.

For compartment 1, we assume that each stem cell divides at rate r_1, i.e., the waiting time for each cell division event follows exponential distribution with parameter r_1. It can either undergo self-renewal via symmetric division with probability p_1, or differentiate into two transient amplifying cells via symmetric differentiation with probability q_1 ($p_1 + q_1 = 1$). For compartment 2, similarly, each transient amplifying cell is assumed to divide at rate r_2. Either the symmetric division happens with probability p_2, or the symmetric differentiation (giving rise to two terminally differentiated cells) happens with probability q_2 ($p_2 + q_2 = 1$). For compartment 3, the terminally differentiated cells cannot divide and are removed from the population at rate d.

Given the above hierarchically structured cell population, let us now introduce cellular dedifferentiation. Since it is not very clear how a non dedifferentiating cell acquires the ability of dedifferentiation, we regard dedifferentiation as a perturbation to the hierarchical model as used in Ref. [31]. More specifically, besides the symmetric division and symmetric differentiation, each transient amplifying cell can undergo symmetric dedifferentiation with probability δ_2 via which two stem cells are generated. It should be pointed out that dedifferentiation is not allowed to happen to terminally differentiated cells as in Ref. [31]. Note that $p_2 + q_2 = 1$, i.e., the sum of the symmetric division and symmetric differentiation probabilities of each transient amplifying cell equals to 1 when dedifferentiation is not taken into account. Due to the introduction of dedifferentiation, however, the sum of these two probabilities is reduced from 1 to $1 - \delta_2$. Current knowledge regarding the effect of dedifferentiation on these two probabilities is still very lacking. Biologically it is poorly understood how much these two probabilities change individually. In view of this, we introduce a parameter $\kappa \in (0, 1)$ to represent how the dedifferentiation reshapes the division pattern of transient amplifying cells, which is thus called reshaping factor [31]. As κ increases, symmetric differentiation is preferred over symmetric division. In this way, the symmetric division probability of each transient amplifying cell is given by $p_2' = p_2 - \kappa\delta_2$, and its symmetric differentiation probability is given by $q_2' = q_2 - (1 - \kappa)\delta_2$. Both dedifferentiation probability δ_2 and reshaping factor κ are the key parameters relating to dedifferentiation. A major task in this work is to explore how the noises in the numbers of different cell types relate to both δ_2 and κ.

The schematic representation of cell processes summarize the above model assumptions as follows:

(1) SC $\xrightarrow{r_1 p_1}$ SC + SC,
(2) SC $\xrightarrow{r_1 q_1}$ TAC + TAC,
(3) TAC $\xrightarrow{r_2 p_2'}$ TAC + TAC,
(4) TAC $\xrightarrow{r_2 q_2'}$ TDC + TDC,
(5) TAC $\xrightarrow{r_2 \delta_2}$ SC + SC,
(6) TDC $\xrightarrow{d}$ $\varnothing$.

It should be noted that, asymmetric cell division is not accounted for in this model. In Appendix A we discuss the effect of asymmetric cell division on the model and derive the moment equations accordingly. Moreover, cell competitions are also taken into account in our model. Instead of using nonlinear feedback regulation to maintain the population equilibrium [27,32,33], we consider a very simple mode of intra-compartment cell competition as follows:

(7) SC + SC $\xrightarrow{\alpha_1}$ SC,
(8) TAC + TAC $\xrightarrow{\alpha_2}$ TAC,
(9) TDC + TDC $\xrightarrow{\alpha_3}$ TDC.

That is, each cell competition event causes one of the two competing cells in the same compartment to die at equal chance. α_i $(i = 1, 2, 3)$ is the competition death rate controlling the competition strength and thus regulating the population size [34].

Let X_t, Y_t, and Z_t be the cell numbers of SCs, TACs, and TDCs at time t, respectively. Their stochastic dynamics are captured by the following master equation:

$$\begin{aligned}\frac{\partial \varphi_{(i,j,k)}}{\partial t} &= \varphi_{(i-1,j,k)}(i-1)r_1p_1 + \varphi_{(i+1,j-2,k)}(i+1)r_1q_1 + \varphi_{(i+1,j,k)}(i+1)i\alpha_1\\ &\quad + \varphi_{(i-2,j+1,k)}(j+1)r_2\delta_2 + \varphi_{(i,j-1,k)}(j-1)r_2p_2' + \varphi_{(i,j+1,k-2)}(j+1)r_2q_2' + \varphi_{(i,j+1,k)}(j+1)j\alpha_2\\ &\quad + \varphi_{(i,j,k+1)}(k+1)d + \varphi_{(i,j,k+1)}(k+1)k\alpha_3\\ &\quad - \varphi_{(i,j,k)}[ir_1 + i(i-1)\alpha_1 + jr_2 + j(j-1)\alpha_2 + kd + k(k-1)\alpha_3], \end{aligned} \tag{1}$$

where $\varphi_{(i,j,k)} := P(X_t = i, Y_t = j, Z_t = k)$, representing the joint probability distribution of (X_t, Y_t, Z_t). In what follows we will use (X, Y, Z) short for (X_t, Y_t, Z_t). Let $\langle\cdot\rangle$ be the expectation or the moment of random variable. For example, $\langle X\rangle$ means the first moment of X, $\langle X^2\rangle$ means the second moment of X, and $\langle(X - \langle X\rangle)^2\rangle = \langle X^2\rangle - \langle X\rangle^2$ defines the variance characterizing the stochastic fluctuation of X around average value. To nondimensionalize variance, the coefficient of variation (CV) is often used to quantify the noise in various biological systems [17,26]. Here we define the square of CV, i.e.,

$$\mathrm{C}_X := \frac{\langle X^2\rangle - \langle X\rangle^2}{\langle X\rangle^2}, \tag{2}$$

to be the noise in the number of SCs. That is, C_X is the dimensionless quantity that describes the fluctuation in the number of stem cells. Similarly, C_Y and C_Z quantify the noises in the numbers of TACs and TDCs, respectively.

Note that C characterizes the total noise in the number of given cell type, which can be conceptually divided into two components. The first is purely generated from the the given cell type in the absence of the propagation from other cell types (intrinsic noise). The second is the component of propagated noise transmitted from other cell types. By using the noise decomposition method (see Appendix B), we can extract the propagated noise from the total noise [see Eqs. (B6), (B7), and (B8)]. In what follows, Adjusted C_X, Adjusted C_Y, and Adjusted C_Z represent the noises of X, Y, and Z by removing the effect of intrinsic components, respectively. In other words, compared to C, Adjusted C emphasizes more on propagated effect due to cell differentiation and dedifferentiation.

Stochastic simulation (e.g., Gillespie algorithm [35]) is frequently used in continuous-time stochastic processes, but it is time-consuming when investigating the statistical properties of the model. Instead, we will derive the moment equations, with which it is more convenient to calculate the noises in the numbers of different cell types.

III. THE MOMENT EQUATIONS

Note that the coefficient of variation depends on the first and second moments of X, Y, or Z. To explore the effect of cellular dedifferentiation on the noise in the cell number of each cell compartment, it is very useful to obtain the mathematical equations governing the dynamics of the first moments $\langle X\rangle$, $\langle Y\rangle$, $\langle Z\rangle$, and the second moments $\langle X^2\rangle$, $\langle Y^2\rangle$, $\langle Z^2\rangle$. In this section, we make use of the master equation Eq. (1) and moment closure approximation method [29] to obtain the moment equations for our model.

For the first moment, noticing that $\langle X\rangle = \sum_{i,j,k} i\varphi_{(i,j,k)}$, $\langle Y\rangle = \sum_{i,j,k} j\varphi_{(i,j,k)}$, $\langle Z\rangle = \sum_{i,j,k} k\varphi_{(i,j,k)}$, based on Eq. (1), we obtain the ordinary differential equations (ODEs) as follows (see Appendix C):

$$\begin{aligned}\frac{d\langle X\rangle}{dt} &= \langle X\rangle(r_1p_1 - r_1q_1 + \alpha_1) - \langle X^2\rangle\alpha_1 + 2\langle Y\rangle r_2\delta_2,\\ \frac{d\langle Y\rangle}{dt} &= 2\langle X\rangle r_1q_1 + \langle Y\rangle(r_2p_2' - r_2q_2' - r_2\delta_2 + \alpha_2) - \langle Y^2\rangle\alpha_2,\\ \frac{d\langle Z\rangle}{dt} &= 2\langle Y\rangle r_2q_2' + \langle Z\rangle(-d + \alpha_3) - \langle Z^2\rangle\alpha_3. \end{aligned} \tag{3}$$

We can see that the main challenge of analyzing the above ODEs lies in the presence of the second moments $\langle X^2\rangle$, $\langle Y^2\rangle$, and $\langle Z^2\rangle$. In other words, Eq. (3) is not closed. When the population size is sufficiently large, the stochastic fluctuation is negligible, and then approximately we can overcome this issue directly replacing $\langle X^2\rangle$, $\langle Y^2\rangle$, $\langle Z^2\rangle$ with $\langle X\rangle^2$, $\langle Y\rangle^2$, $\langle Z\rangle^2$. However, this method is invalid when the noise is nonnegligible. In our case, we should take into account both the first and second moments together. For the second moments $\langle X^2\rangle = \sum_{i,j,k} i^2\varphi_{(i,j,k)}$, $\langle Y^2\rangle = \sum_{i,j,k} j^2\varphi_{(i,j,k)}$, and $\langle Z^2\rangle = \sum_{i,j,k} k^2\varphi_{(i,j,k)}$, based on Eq. (1) we have the ODEs as follows (see Appendix C):

$$\begin{aligned}
\frac{d\langle X^2\rangle}{dt} &= \langle X^2\rangle(2r_1p_1 - 2r_1q_1 + 3\alpha_1) + \langle X\rangle(r_1p_1 + r_1q_1 - \alpha_1)\\
&\quad - 2\langle X^3\rangle\alpha_1 + 4\langle XY\rangle r_2\delta_2 + 4\langle Y\rangle r_2\delta_2,\\
\frac{d\langle Y^2\rangle}{dt} &= \langle Y^2\rangle(2r_2p_2' - 2r_2q_2' + 3\alpha_2 - 2r_2\delta_2)\\
&\quad + \langle Y\rangle(r_2p_2' + r_2q_2' + r_2\delta_2 - \alpha_2) - 2\langle Y^3\rangle\alpha_2\\
&\quad + 4\langle XY\rangle r_1q_1 + 4\langle X\rangle r_1q_1,\\
\frac{d\langle Z^2\rangle}{dt} &= \langle Z^2\rangle(3\alpha_3 - 2d) + \langle Z\rangle(d - \alpha_3) - 2\langle Z^3\rangle\alpha_3\\
&\quad + 4\langle YZ\rangle r_2q_2' + 4\langle Y\rangle r_2q_2'.
\end{aligned} \tag{4}$$

Here $\langle XY\rangle = \sum_{i,j,k} ij\varphi_{(i,j,k)}$, $\langle YZ\rangle = \sum_{i,j,k} jk\varphi_{(i,j,k)}$, and $\langle XZ\rangle = \sum_{i,j,k} ik\varphi_{(i,j,k)}$ are the mixed moments whose dynamics is captured by (see Appendix C)

$$\begin{aligned}
\frac{d\langle XY\rangle}{dt} &= \langle XY\rangle(r_1p_1 - r_1q_1 + \alpha_1 + r_2p_2' - r_2q_2' - r_2\delta_2 + \alpha_2)\\
&\quad + 2\langle X^2\rangle r_1q_1 - 2\langle X\rangle r_1q_1 - \langle X^2Y\rangle\alpha_1 - \langle XY^2\rangle\alpha_2\\
&\quad + 2\langle Y^2\rangle r_2\delta_2 - 2\langle Y\rangle r_2\delta_2,\\
\frac{d\langle YZ\rangle}{dt} &= \langle YZ\rangle(r_2p_2' - r_2q_2' - r_2\delta_2 + \alpha_2 - d + \alpha_3) + 2\langle XZ\rangle r_1q_1\\
&\quad + 2\langle Y^2\rangle r_2q_2' - 2\langle Y\rangle r_2q_2' - \langle Y^2Z\rangle\alpha_2 - \langle YZ^2\rangle\alpha_3,\\
\frac{d\langle XZ\rangle}{dt} &= \langle XZ\rangle(r_1p_1 - r_1q_1 + \alpha_1 - d + \alpha_3) + 2\langle YZ\rangle r_2\delta_2\\
&\quad - \langle X^2Z\rangle\alpha_1 - \langle XZ^2\rangle\alpha_3 + 2\langle XY\rangle r_2q_2'.
\end{aligned} \tag{5}$$

We can see that Eqs. (3), (4), and (5) are coupled together, and even the coupled equations are not closed due to the presence of the third moment items. To overcome the "not closed" issue, the moment closure approximation methods are often used to reduce the higher order moments into lower ones [29,36]. Here we adopt the cumulant-neglect moment closure approximation [29] to replace all the third moment items by the functions of the first and second moments as follows:

$$\begin{aligned}
\langle X^3\rangle &\approx 3\langle X^2\rangle\langle X\rangle - 2\langle X\rangle^3,\\
\langle Y^3\rangle &\approx 3\langle Y^2\rangle\langle Y\rangle - 2\langle Y\rangle^3,\\
\langle Z^3\rangle &\approx 3\langle Z^2\rangle\langle Z\rangle - 2\langle Z\rangle^3,\\
\langle X^2Y\rangle &\approx \langle X^2\rangle\langle Y\rangle + 2\langle XY\rangle\langle X\rangle - 2\langle X\rangle^2\langle Y\rangle,\\
\langle XY^2\rangle &\approx \langle Y^2\rangle\langle X\rangle + 2\langle XY\rangle\langle Y\rangle - 2\langle Y\rangle^2\langle X\rangle,\\
\langle Y^2Z\rangle &\approx \langle Y^2\rangle\langle Z\rangle + 2\langle YZ\rangle\langle Y\rangle - 2\langle Y\rangle^2\langle Z\rangle,\\
\langle YZ^2\rangle &\approx \langle Z^2\rangle\langle Y\rangle + 2\langle YZ\rangle\langle Z\rangle - 2\langle Z\rangle^2\langle Y\rangle,\\
\langle X^2Z\rangle &\approx \langle X^2\rangle\langle Z\rangle + 2\langle XZ\rangle\langle X\rangle - 2\langle X\rangle^2\langle Z\rangle,\\
\langle XZ^2\rangle &\approx \langle Z^2\rangle\langle X\rangle + 2\langle XZ\rangle\langle Z\rangle - 2\langle Z\rangle^2\langle X\rangle.
\end{aligned} \tag{6}$$

In this way, we can obtain the closed system of ordinary differential equations for all the second moments as follows:

$$\begin{aligned}
\frac{d\langle X^2\rangle}{dt} &= \langle X^2\rangle(2r_1p_1 - 2r_1q_1 + 3\alpha_1) + \langle X\rangle(r_1p_1 + r_1q_1 - \alpha_1)\\
&\quad - 2[3\langle X^2\rangle\langle X\rangle - 2(\langle X\rangle)^3]\alpha_1 + 4\langle XY\rangle r_2\delta_2 + 4\langle Y\rangle r_2\delta_2,\\
\frac{d\langle Y^2\rangle}{dt} &= \langle Y^2\rangle(2r_2p_2' - 2r_2q_2' + 3\alpha_2 - 2r_2\delta_2)\\
&\quad + \langle Y\rangle(r_2p_2' + r_2q_2' + r_2\delta_2 - \alpha_2) - 2[3\langle Y^2\rangle\langle Y\rangle\\
&\quad - 2(\langle Y\rangle)^3]\alpha_2 + 4\langle XY\rangle r_1q_1 + 4\langle X\rangle r_1q_1,\\
\frac{d\langle Z^2\rangle}{dt} &= \langle Z^2\rangle(3\alpha_3 - 2d) + \langle Z\rangle(d - \alpha_3) - 2[3\langle Z^2\rangle\langle Z\rangle\\
&\quad - 2(\langle Z\rangle)^3]\alpha_3 + 4\langle YZ\rangle r_2q_2' + 4\langle Y\rangle r_2q_2.
\end{aligned} \tag{7}$$

$$\begin{aligned}
\frac{d\langle XY\rangle}{dt} &= \langle XY\rangle(r_1p_1 - r_1q_1 + \alpha_1 + r_2p_2' - r_2q_2' - r_2\delta_2 + \alpha_2)\\
&\quad + 2\langle X^2\rangle r_1q_1 - [\langle X^2\rangle\langle Y\rangle + 2\langle XY\rangle\langle X\rangle\\
&\quad - 2(\langle X\rangle)^2\langle Y\rangle]\alpha_1 - [\langle Y^2\rangle\langle X\rangle + 2\langle XY\rangle\langle Y\rangle\\
&\quad - 2(\langle Y\rangle)^2\langle X\rangle]\alpha_2 + 2\langle Y^2\rangle r_2\delta_2\\
&\quad - 2\langle Y\rangle r_2\delta_2 - 2\langle X\rangle r_1q_1,\\
\frac{d\langle YZ\rangle}{dt} &= \langle YZ\rangle(r_2p_2' - r_2q_2' - r_2\delta_2 + \alpha_2 - d + \alpha_3) + 2\langle XZ\rangle r_1q_1\\
&\quad - [\langle Z^2\rangle\langle Y\rangle + 2\langle YZ\rangle\langle Z\rangle - 2(\langle Z\rangle)^2\langle Y\rangle]\alpha_3\\
&\quad - [\langle Y^2\rangle\langle Z\rangle + 2\langle YZ\rangle\langle Y\rangle - 2(\langle Y\rangle)^2\langle Z\rangle]\alpha_2\\
&\quad + 2\langle Y^2\rangle r_2q_2' - 2\langle Y\rangle r_2q_2',\\
\frac{d\langle XZ\rangle}{dt} &= \langle XZ\rangle(r_1p_1 - r_1q_1 + \alpha_1 - d + \alpha_3) + 2\langle YZ\rangle r_2\delta_2\\
&\quad - [\langle Z^2\rangle\langle X\rangle + 2\langle XZ\rangle\langle Z\rangle - 2(\langle Z\rangle)^2\langle X\rangle]\alpha_3\\
&\quad - [\langle X^2\rangle\langle Z\rangle + 2\langle XZ\rangle\langle X\rangle - 2(\langle X\rangle)^2\langle Z\rangle]\alpha_1\\
&\quad + 2\langle XY\rangle r_2q_2'.
\end{aligned} \tag{8}$$

It should be noted that, the cumulant-neglect moment closure approximation applies to Gaussian-like random variables. That is, the more the random variable is normal-distributed, the better the effectiveness of the approximation is. Figure 2 illustrates the effectiveness of the cumulant-neglect moment closure approximation in our model. Figures 2(a) and 2(b) show the comparison between the approximated moment equations and agent-based stochastic simulations using Gillespie algorithm. Specifically, the equilibrium states predicted by the approximated moment equations are in good agreement with stochastic simulations. Besides, Fig. 2(c) shows that X, Y, and Z all follow approximately normal distributions when they approach the steady states. Therefore, the moment Eqs. (3), (7), and (8) provide high-quality approximation to the original model in the equilibrium regime.

IV. RESULTS

In this section, we make use of the moment equations at equilibrium regime to investigate how the noises in the numbers of different cell types are changed by taking dediffer-

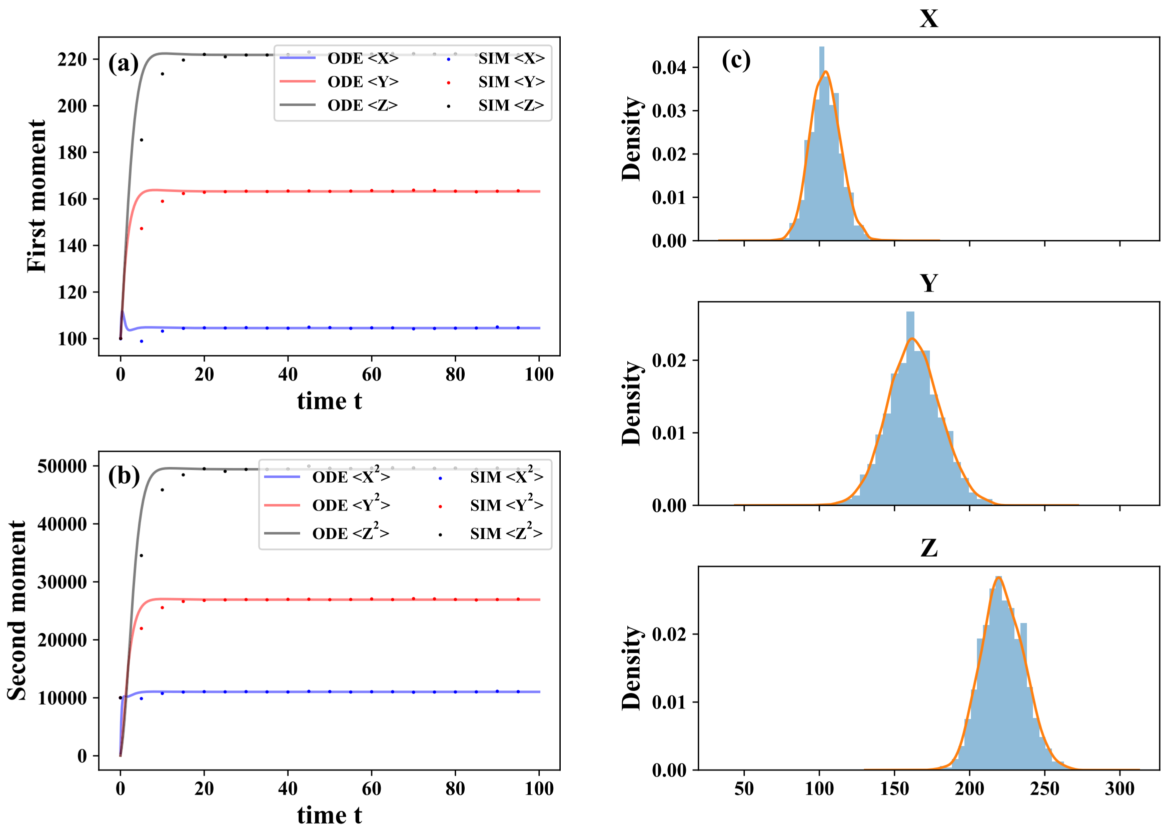

FIG. 2. Validation of the cumulant-neglect moment closure approximation. Panels (a) and (b) illustrate the comparison between the moment equations and the stochastic simulations. In both panels, the continuous lines represent the moment equations from Eqs. (3) and (4) by using the moment closure approximation method and the discrete dots represent stochastic simulations, which agree very well with each other in the equilibrium regime. Panel (c) illustrates the empirical distributions of X, Y, and Z in the equilibrium regime, and they all approximately follow normal distributions. Parameters: $(\delta_2, \kappa, p_1, q_1, p_2, q_2, r_1, r_2, \alpha_1, \alpha_2, \alpha_3, d) = (0.35, 0.5, 0.7, 0.3, 0.6, 0.4, 0.7, 0.7, 0.01, 0.001, 0.001, 0.01)$.

entiation into account. There are two fundamental parameters relating to dedifferentiation. One is δ_2 representing the probability that each transient amplifying cell undergoes symmetric dedifferentiation. It has been reported that the happening of dedifferentiation is generally not very often [8], we thus assume that δ_2 is less than a half. The other parameter is the reshaping factor κ, representing how the introduction of dedifferentiation redistributes the probabilities for self-renewal and differentiation of each transient amplifying cell, namely, $p_2 \rightarrow p_2 - \kappa\delta_2$, $q_2 \rightarrow q_2 - (1-\kappa)\delta_2$. In what follows, we focus our attention to δ_2 and κ, while regard all the other parameters as background and keep them fixed. We discuss the background parameters in Appendix D.

The main results are illustrated in Figs. 3 and 4. In Fig. 3, C and Adjusted C are shown as functions of dedifferentiation probability δ_2. From Figs. 3(a) to 3(d), the value of reshaping factor κ is increasing in turn. For the stem cells compartment, we can see that C_X is significantly decreasing as δ_2 increases. The decreasing trend is quite robust to the reshaping factor κ. This is actually in line with our intuition. Dedifferentiation contributes to the increase of X because the stem cells compartment receives the influx from transient amplifying cells by symmetric dedifferentiation. Note that X can be approximately regarded as a birth-and-death Markov process. As the population size increases, diffusion approximation (such as Ω expansion) shows that the noise is approximately proportional to the inverse of the mean population size [17,26,28] (see Fig. 7 in Appendix E). Thus, enlarging the dedifferentiation probability δ_2 increases the number of stem cells, and then reduce the noise in the number of stem cells. Furthermore, we can also see that the trend of Adjusted C_X is quite similar to C_X. That is, Adjusted C_X is also significantly decreasing as δ_2 increases. Noticing that Adjusted C_X equals to C_X minus the intrinsic noise generated by X, it implies that the decreasing trend of C_X is not only due to the increase of X but also to noise transmission.

However, the effect of dedifferentiation on the nonstem cells compartments is not as straightforward as on the stem cells compartment. From Fig. 3 we can see that dedifferentiation does not necessarily increase or decrease the noises in the numbers of transient amplifying cells and terminally differentiated cells. For example, when κ is small, C_Y is gradually decreasing with δ_2. When κ becomes larger, it is no longer monotonically decreasing but a convex function of δ_2. In other words, compared to stem cells, the noise in the number of nonstem cells is quite sensitive to the dedifferentiation parameters. This complexity actually arises from the dual effect of dedifferentiation on nonstem cells. Let us take transient amplifying cells as an example. The number of transient amplifying cells depends on both the

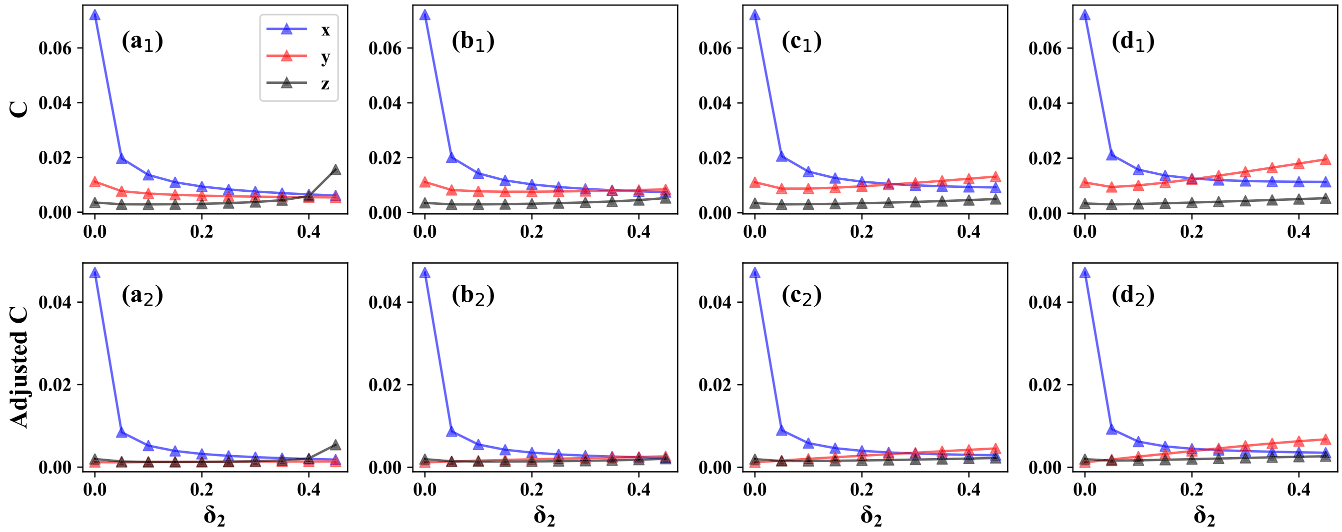

FIG. 3. The effect of dedifferentiation probability δ_2 on C and Adjusted C of different cell types at different levels of reshaping factor. From panels (a) to (d), the values of reshaping factor κ are 0.13, 0.33, 0.53, and 0.73, respectively. The noises (C and Adjusted C) in the numbers of stem cells, transient amplifying cells and terminally differentiated cells are represented by blue, red, and black curves, respectively. The values of the background parameters are the same as those in Fig. 2.

self-renewal and differentiation from stem cells. Even though the introduction of dedifferentiation slows down the self-renewal rate of each transient amplifying cell, dedifferentiation increases the number of stem cells whose differentiation would consequently contribute to the growth of transient amplifying cells. Therefore, the dual role of dedifferentiation in nonstem cells makes C_Y and C_Z much more unpredictable than C_X.

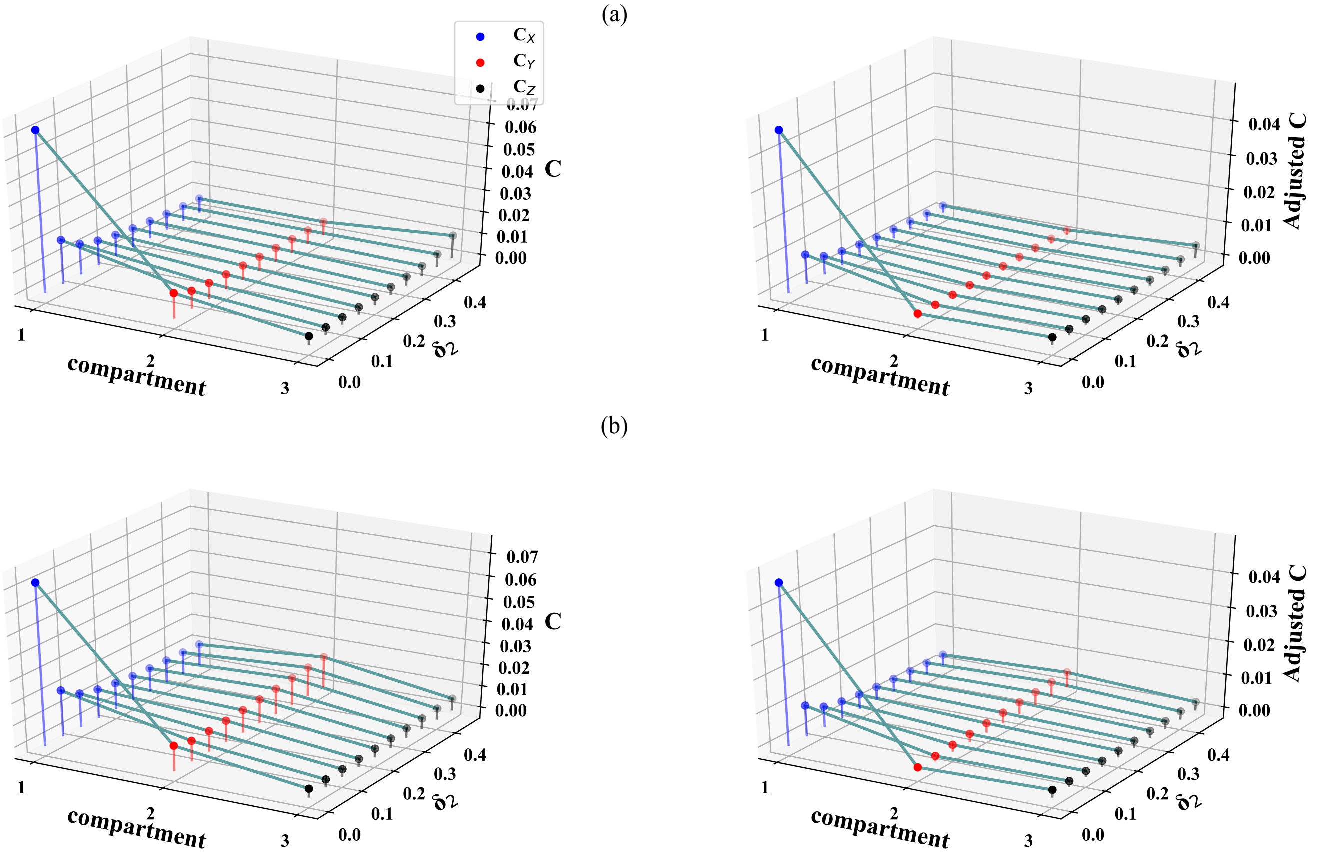

FIG. 4. The effect of dedifferentiation on the noise propagation in cellular hierarchy. Compartments 1, 2, and 3 represent stem cells, transient amplifying cells, and terminally differentiated cells, respectively. In panel (a) $\kappa = 0.155$, and in panel (b) $\kappa = 0.555$. The values of the background parameters are the same as those in Fig. 2. The 2D version of this figure is shown in Fig. 8.

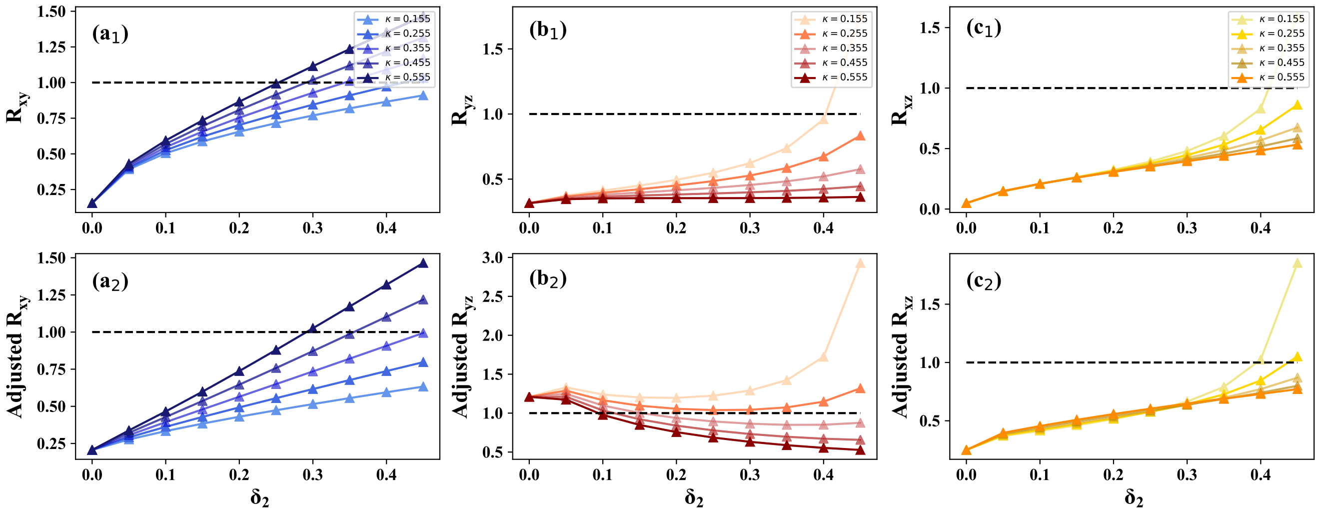

FIG. 5. The effect of dedifferentiation on the noise ratios. From panels (a) to (c), the noise ratios are shown as functions of dedifferentiation probability δ_2 at different levels of reshaping factor κ. The panels in the first row are about R_{XY}, R_{YZ}, and R_{XZ}, while the panels in the second row are about Adjusted R_{XY}, Adjusted R_{YZ}, and Adjusted R_{XZ}. The values of the background parameters are the same as those in Fig. 2.

Besides the effect of dedifferentiation on each cell compartment, an even more interesting question is how the noise propagates along the cellular hierarchy, i.e., the relation of C_X, C_Y, and C_Z, and the relation of Adjusted C_X, Adjusted C_Y, and Adjusted C_Z. In the absence of dedifferentiation, from stem cells to terminally differentiated cells, C is decreasing in sequence (see Fig. 4 for 3D plot and Fig. 8 for 2D plot),i.e., $C_X > C_Y > C_Z$. However, the order of Adjusted C is different. In other words, by removing the effect of intrinsic noise, Adjusted $C_X >$ Adjusted $C_Z >$ Adjusted C_Y. We are particularly interested in how the orders of C and Adjusted C would be changed by dedifferentiation. Figure 4 (see also Fig. 8) illustrates two different scenarios. When κ is small [panel (a)], as the increase of δ_2, C_Z will gradually become the largest among the three compartments. When κ is large [panel (b)], C_Y will gradually become the largest as δ_2 increases. Similar behavior happens to Adjusted C_Z and Adjusted C_Y. Note that reshaping factor κ is used for coordinating the self-renewal and differentiation probabilities of transient amplifying cells. Increasing κ reduces the self-renewal rate and meanwhile enhance the differentiation rate. It tends to produce the opposite effect on transient amplifying cells and terminally differentiated cells.

To further investigate this issue, let us consider the noise ratio defined as follows [27,37,38]:

$$R_{XY} := \frac{C_Y}{C_X}, \quad R_{YZ} := \frac{C_Z}{C_Y}, \quad R_{XZ} := \frac{C_Z}{C_X}. \tag{9}$$

$$\text{Adjusted } R_{XY} := \frac{\text{Adjusted } C_Y}{\text{Adjusted } C_X},$$

$$\text{Adjusted } R_{YZ} := \frac{\text{Adjusted } C_Z}{\text{Adjusted } C_Y},$$

$$\text{Adjusted } R_{XZ} := \frac{\text{Adjusted } C_Z}{\text{Adjusted } C_X}. \tag{10}$$

They are all the ratios of downstream compartments to upstream compartments. For example, R_{XY} characterizes the relation between transient amplifying cells and stem cells. When $R_{XY} > 1$, the noise is amplified from stem cells to transient amplifying cells, and vice versa. Hence, the comparison of the noises in different compartments is equivalent to the comparison between the noise ratios and 1.

Let us first consider the noise ratios R_{XY}, R_{YZ}, and R_{XZ}. Noticing that in the absence of dedifferentiation, all the noise ratios are smaller than 1, we are concerned about whether the noise ratios could become larger than 1 by changing the values of δ_2 and κ. The result is illustrated in Fig. 5. For R_{XY}, it is a monotonically increasing function of δ_2 for different levels of κ. Besides, it is found that R_{XY} is also an increasing function of κ. In other words, R_{XY} is more likely to exceed to 1 provided larger values of dedifferentiation probability and reshaping factor. For R_{YZ}, it is a monotonically increasing function of δ_2 only for small κ. R_{YZ} is very unlikely to exceed to 1 for large κ. For R_{XZ}, it is again monotonically increasing as δ_2 increases. However, the relation between R_{XZ} and κ is opposite to the relation between R_{XY} and κ. Namely, R_{XZ} is a decreasing function of κ, and then it is more likely to exceed to 1 provided larger dedifferentiation probability and smaller reshaping factor. The opposite effect of reshaping factor κ on R_{XY} and R_{XZ} leads to an interesting tradeoff phenomenon. For large κ, it is more likely for R_{XY} to exceed to 1, but less likely for R_{XZ} to exceed to 1. On the contrary, for small κ, it is less likely for R_{XY} to exceed to 1, but more likely for R_{XZ} to exceed to 1. The tradeoff phenomenon implies that even though dedifferentiation could significantly alter the order of C_X, C_Y, and C_Z (the original order is $C_X > C_Y > C_Z$ when dedifferentiation is absent), it is quite difficult to continuously amplify the noise along cellular hierarchy for a large range of parameter values. Similar argument can be used in the adjusted noise ratios between compartments. However, note that in the absence of dedifferentiation Adjusted R_{YZ} is already larger than 1, whereas R_{YZ} is smaller than 1. Even so, it is still very unlikely to realize Adjusted $C_Z >$ Adjusted $C_Y >$ Adjusted C_X.

V. CONCLUSIONS

In this study, we have explored the effect of cellular dedifferentiation on the noise propagation in cellular hierarchy using a stochastic model composed of different cell phenotypes. By performing noise decomposition and solving the moment equations of our model, we have systematically investigated how the noises in the numbers of different cell types are changed with dedifferentiation parameters.

According to our results, dedifferentiation works quite differently on the noises in the numbers of stem cells and nonstem cells. For stem cells, the result is very straightforward that the noise is significantly reduced by introducing dedifferentiation. The larger the dedifferentiation probability is, the more significant the noise is reduced. This relation is quite robust to other parameters such as reshaping factor. However, the effect of dedifferentiation on nonstem cells is not as straightforward as on stem cells. The complexity lies in two aspects. The first is the dual role of dedifferentiation, namely, dedifferentiation could decrease the number of transient amplifying cells via lowering their self-renewal rate as well as increase their number via enhancing the influx from stem cells. Hence, it is quite difficult to predict how the noise in the number of nonstem cells is changed with dedifferentiation. The second is the combined effect of dedifferentiation probability and reshaping factor on the propagation of noise in number of nonstem cells. Specifically, the reshaping factor tend to give the opposite effect to transient amplifying cells and terminally differentiated cells. Noticing that the non dedifferentiating cellular hierarchy generates a noise decreasing cascade from stem cells to terminally differentiated cells, our result suggests that it is very difficult for dedifferentiation to simultaneously amplify both the stem-cells-to-transient-amplifying-cells noise propagation and the transient-amplifying-cells-to-fully-differentiated-cells noise propagation. Even though dedifferentiation could turn part of the noise propagation process into noise-amplifying step, it is very unlikely to turn the entire process into a noise-amplifying cascade.

It should be noted that, in this study we have not related our theoretical results with real experimental data. To do this, both mean population size and the fluctuation around the mean population size should be measured in biological experiments. In future, it would be interesting to validate our model with experimental data and develop statistical methods for inferring dedifferentiation.

ACKNOWLEDGMENTS

We thank two reviewers for their constructive comments. D.Z. acknowledges the sponsorship by the NSFC (Grant No. 11971405). H.J.P. was supported by the National Research Foundation of Korea grant funded by the Korea government (MSIT), Grant No. 2020R1A2C1101894, and by an appointment to the JRG Program at the APCTP through the Science and Technology Promotion Fund and Lottery Fund of the Korean Government. This was also supported by the Korean Local Governments–Gyeongsangbuk-do Province and Pohang City.

APPENDIX A: ASYMMETRIC CELL DIVISION

Let us take asymmetric cell division into account. The schematic representation of cell processes summarize is shown as follows:

(1) $\mathrm{SC} \xrightarrow{r_1 p_1} \mathrm{SC} + \mathrm{SC}$,
(2) $\mathrm{SC} \xrightarrow{r_1 q_1} \mathrm{TAC} + \mathrm{TAC}$,
(3) $\mathrm{SC} \xrightarrow{r_1 o_1} \mathrm{SC} + \mathrm{TAC}$,
(4) $\mathrm{TAC} \xrightarrow{r_2 p_2'} \mathrm{TAC} + \mathrm{TAC}$,
(5) $\mathrm{TAC} \xrightarrow{r_2 q_2'} \mathrm{TDC} + \mathrm{TDC}$,
(6) $\mathrm{TAC} \xrightarrow{r_2 o_2'} \mathrm{TAC} + \mathrm{TDC}$,
(7) $\mathrm{TAC} \xrightarrow{r_2 \delta_2} \mathrm{SC} + \mathrm{SC}$,
(8) $\mathrm{TAC} \xrightarrow{r_2 \delta_1} \mathrm{TAC} + \mathrm{SC}$,
(9) $\mathrm{TDC} \xrightarrow{d} \varnothing$,
(10) $\mathrm{SC} + \mathrm{SC} \xrightarrow{\alpha_1} \mathrm{SC}$,
(11) $\mathrm{TAC} + \mathrm{TAC} \xrightarrow{\alpha_2} \mathrm{TAC}$,
(12) $\mathrm{TDC} + \mathrm{TDC} \xrightarrow{\alpha_3} \mathrm{TDC}$.

Here (3), (6), and (8) are asymmetric cell divisions. The master equation is obtained as follows:

$$\begin{aligned}\frac{\partial \varphi_{(i,j,k)}}{\partial t} &= \varphi_{(i-1,j,k)}(i-1)r_1p_1 + \varphi_{(i+1,j-2,k)}(i+1)r_1q_1 + \varphi_{(i,j-1,k)}ir_1o_1 + \varphi_{(i+1,j,k)}(i+1)i\alpha_1 \\ &\quad + \varphi_{(i-2,j+1,k)}(j+1)r_2\delta_2 + \varphi_{(i-1,j,k)}jr_2\delta_1 + \varphi_{(i,j-1,k)}(j-1)r_2p_2' + \varphi_{(i,j+1,k-2)}(j+1)r_2q_2' \\ &\quad + \varphi_{(i,j,k-1)}jr_2o_2' + \varphi_{(i,j+1,k)}(j+1)j\alpha_2 + \varphi_{(i,j,k+1)}(k+1)d + \varphi_{(i,j,k+1)}(k+1)k\alpha_3 \\ &\quad - \varphi_{(i,j,k)}[ir_1 + i(i-1)\alpha_1 + jr_2 + j(j-1)\alpha_2 + kd + k(k-1)\alpha_3]. \qquad \text{(A1)}\end{aligned}$$

Based on Eq. (A1), we can all the moment equations as follows:

$$\begin{aligned}
\frac{d\langle X\rangle}{dt} &= \langle X\rangle(r_1p_1 - r_1q_1 + \alpha_1) - \langle X^2\rangle\alpha_1 + \langle Y\rangle(r_2\delta_1 + 2r_2\delta_2),\\
\frac{d\langle Y\rangle}{dt} &= \langle X\rangle(2r_1q_1 + r_1o_1) + \langle Y\rangle(r_2p_2' - r_2q_2' - r_2\delta_2 + \alpha_2) - \langle Y^2\rangle\alpha_2,\\
\frac{d\langle Z\rangle}{dt} &= \langle Y\rangle(2r_2q_2' + r_2o_2') + \langle Z\rangle(-d + \alpha_3) - \langle Z^2\rangle\alpha_3,
\end{aligned} \tag{A2}$$

$$\begin{aligned}
\frac{d\langle X^2\rangle}{dt} &= \langle X^2\rangle(2r_1p_1 - 2r_1q_1 + 3\alpha_1) + \langle X\rangle(r_1p_1 + r_1q_1 - \alpha_1) - 2\langle X^3\rangle\alpha_1\\
&\quad + \langle XY\rangle(4r_2\delta_2 + 2r_2\delta_1) + \langle Y\rangle(4r_2\delta_2 + r_2\delta_1),\\
\frac{d\langle Y^2\rangle}{dt} &= \langle Y^2\rangle(2r_2p_2' - 2r_2q_2' + 3\alpha_2 - 2r_2\delta_2) + \langle XY\rangle(4r_1q_1 + 2r_1o_1)\\
&\quad + \langle X\rangle(4r_1q_1 + r_1o_1) + \langle Y\rangle(r_2p_2' + r_2q_2' + r_2\delta_2 - \alpha_2) - 2\langle Y^3\rangle\alpha_2,\\
\frac{d\langle Z^2\rangle}{dt} &= \langle Z^2\rangle(3\alpha_3 - 2d) + \langle Z\rangle(d - \alpha_3) - 2\langle Z^3\rangle\alpha_3 + \langle YZ\rangle(4r_2q_2' + 2r_2o_2') + \langle Y\rangle(4r_2q_2' + r_2o_2'),
\end{aligned} \tag{A3}$$

$$\begin{aligned}
\frac{d\langle XY\rangle}{dt} &= \langle XY\rangle(r_1p_1 - r_1q_1 + \alpha_1 + r_2p_2' - r_2q_2' - r_2\delta_2 + \alpha_2) + \langle X^2\rangle(2r_1q_1\\
&\quad + r_1o_1) - 2\langle X\rangle r_1q_1 - \langle X^2Y\rangle\alpha_1 - \langle XY^2\rangle\alpha_2\\
&\quad + \langle Y^2\rangle(2r_2\delta_2 + r_2\delta_1) - 2\langle Y\rangle r_2\delta_2,\\
\frac{d\langle YZ\rangle}{dt} &= \langle YZ\rangle(r_2p_2' - r_2q_2' - r_2\delta_2 + \alpha_2 - d + \alpha_3) + \langle XZ\rangle(2r_1q_1 + r_1o_1)\\
&\quad + \langle Y^2\rangle(2r_2q_2' + r_2o_2') - 2\langle Y\rangle r_2q_2' - \langle Y^2Z\rangle\alpha_2 - \langle YZ^2\rangle\alpha_3,\\
\frac{d\langle XZ\rangle}{dt} &= \langle XZ\rangle(r_1p_1 - r_1q_1 + \alpha_1 - d + \alpha_3) + \langle YZ\rangle(2r_2\delta_2 + r_2\delta_1)\\
&\quad + \langle XY\rangle(2r_2q_2' + r_2o_2') - \langle X^2Z\rangle\alpha_1 - \langle XZ^2\rangle\alpha_3.
\end{aligned} \tag{A4}$$

By comparing the moment Eqs. (3), (4), and (5) and Eqs. (A2), (A3), and (A4), we can see that the introduction of asymmetric division shifts the values of coefficients, without changing the structure of the moment equations.

APPENDIX B: NOISE DECOMPOSITION

The noise in the number of a given cell type can be conceptually divided into two parts. The first part is generated from the population of the given cell type in the absence of the propagation from other cell populations. This part is often called intrinsic noise [26]. The second part of noise is from other cell populations, i.e., the component of propagated noise from adjacent cell compartments. By applying the noise decomposition method proposed in Refs. [17,26], here we decompose C_X, C_Y, and C_Z as follows:

$$\mathrm{C}_X = \frac{1}{\langle X\rangle H_{XX}} + \frac{H_{YX}}{H_{XX}}\frac{H_{XY}}{\tau_2}\epsilon\frac{1}{\langle X\rangle H_{XX}} + \frac{H_{YX}}{H_{XX}}\frac{H_{YX}}{\tau_1}\epsilon\frac{1}{\langle Y\rangle H_{YY}} + \frac{H_{YX}}{H_{XX}}\epsilon\left(\frac{p_1}{\langle Y\rangle} + \frac{\delta_2}{\langle X\rangle}\right), \tag{B1}$$

$$\mathrm{C}_Y = \frac{1}{\langle Y\rangle H_{YY}} + \frac{H_{XY}}{H_{YY}}\frac{H_{YX}}{\tau_1}\epsilon\frac{1}{\langle Y\rangle H_{YY}} + \frac{H_{XY}}{H_{YY}}\frac{H_{XY}}{\tau_2}\epsilon\frac{1}{\langle X\rangle H_{XX}} + \frac{H_{XY}}{H_{YY}}\epsilon\left(\frac{q_1}{\langle Y\rangle} + \frac{\delta_2}{\langle X\rangle}\right), \tag{B2}$$

$$\begin{aligned}
\mathrm{C}_Z &= \frac{1}{\langle Z\rangle H_{ZZ}} + \frac{H_{YZ}}{H_{ZZ}}\frac{H_{XY}}{\tau_2}\frac{H_{YZ}}{\tau_3}\epsilon\eta\left(\frac{q_1}{\langle Y\rangle} + \frac{\delta_2}{\langle X\rangle}\right) + \frac{H_{YZ}}{H_{ZZ}}\eta\left(\frac{H_{YY}}{\tau_2} + \frac{H_{ZZ}}{\tau_3}\right)\frac{q_2'}{\langle Z\rangle}\\
&\quad + \frac{H_{YZ}}{H_{ZZ}}\epsilon\eta\left(\frac{H_{XY}}{\tau_2}\right)^2\frac{H_{ZZ}}{\tau_3}\frac{1}{\langle X\rangle H_{XX}} + \frac{H_{YZ}}{H_{ZZ}}\epsilon\eta\frac{H_{XY}}{\tau_2}\frac{H_{YZ}}{\tau_3}\frac{H_{YX}}{\tau_1}\frac{1}{\langle Y\rangle H_{YY}} + \frac{H_{YZ}}{H_{ZZ}}\eta\frac{H_{YZ}}{\tau_3}\left(\frac{H_{XX}}{\tau_1} + \frac{H_{ZZ}}{\tau_3}\right)\mathrm{C}_Y.
\end{aligned} \tag{B3}$$

Here

$$\begin{aligned}
\epsilon &= \left(\frac{H_{XX}}{\tau_1} + \frac{H_{YY}}{\tau_2} - \frac{H_{YX}}{\tau_1}\frac{H_{XY}}{H_{YY}} - \frac{H_{XY}}{\tau_2}\frac{H_{YX}}{H_{XX}}\right)^{-1},\\
\eta &= \left[\left(\frac{H_{YY}}{\tau_2} + \frac{H_{ZZ}}{\tau_3}\right)\left(\frac{H_{XX}}{\tau_1} + \frac{H_{ZZ}}{\tau_3}\right) - \frac{H_{YX}}{\tau_1}\right]^{-1},
\end{aligned} \tag{B4}$$

and H_{IJ} is the logarithmic gain [26] that measures how the change of cell type J in number is affected by cell type I. In particular,

$$\begin{aligned}
H_{XZ} &= H_{ZX} = H_{ZY} = 0, \quad H_{YZ} = -1,\\
H_{XX} &= \frac{r_1q_1 + 2(\langle X\rangle - 1)\alpha_1}{r_1q_1 + (\langle X\rangle - 1)\alpha_1} - \frac{\langle X\rangle r_1p_1}{\langle X\rangle r_1p_1 + 2\langle Y\rangle r_2\delta_2}\\
&= 1 + \frac{\langle X\rangle\alpha_1}{r_1q_1 + (\langle X\rangle - 1)\alpha_1} - \frac{\langle X\rangle r_1p_1}{\langle X\rangle r_1p_1 + 2\langle Y\rangle r_2\delta_2}\\
&= \frac{\langle X\rangle\alpha_1}{r_1q_1 + (\langle X\rangle - 1)\alpha_1} + \frac{2\langle Y\rangle r_2\delta_2}{\langle X\rangle r_1p_1 + 2\langle Y\rangle r_2\delta_2},\\
H_{XY} &= -\frac{2\langle X\rangle r_1q_1}{2\langle X\rangle r_1q_1 + \langle Y\rangle r_2p_2'}, \quad H_{YX} = -\frac{2\langle Y\rangle r_2\delta_2}{\langle X\rangle r_1p_1 + 2\langle Y\rangle r_2\delta_2},\\
H_{YY} &= \frac{r_2q_2' + r_2\delta_2 + 2(\langle Y\rangle - 1)\alpha_2}{r_2q_2' + r_2\delta_2 + (\langle Y\rangle - 1)\alpha_2} - \frac{\langle Y\rangle r_2p_2'}{2\langle X\rangle r_1q_1 + \langle Y\rangle r_2p_2'}\\
&= 1 + \frac{\langle Y\rangle\alpha_2}{r_2q_2' + r_2\delta_2 + (\langle Y\rangle - 1)\alpha_2} - \frac{\langle Y\rangle r_2p_2'}{2\langle X\rangle r_1q_1 + \langle Y\rangle r_2p_2'}\\
&= \frac{\langle Y\rangle\alpha_2}{r_2q_2' + r_2\delta_2 + (\langle Y\rangle - 1)\alpha_2} + \frac{2\langle X\rangle r_1q_1}{2\langle X\rangle r_1q_1 + \langle Y\rangle r_2p_2'}.\\
H_{ZZ} &= \frac{d + 2(\langle Z\rangle - 1)\alpha_3}{d + (\langle Z\rangle - 1)\alpha_3} = 1 + \frac{\langle Z\rangle\alpha_3}{d + (\langle Z\rangle - 1)\alpha_3}.
\end{aligned} \tag{B5}$$

Note that $\frac{1}{\langle X\rangle H_{XX}}$, $\frac{1}{\langle Y\rangle H_{YY}}$, and $\frac{1}{\langle Z\rangle H_{ZZ}}$ are the noises purely generated from X, Y, and Z, respectively [17]. To extract the component of propagated noise from the total noise, we define Adjusted C as follows:

$$\text{Adjusted C}_X = \text{C}_X - \frac{1}{\langle X\rangle H_{XX}}, \tag{B6}$$

$$\text{Adjusted C}_Y = \text{C}_Y - \frac{1}{\langle Y\rangle H_{YY}}, \tag{B7}$$

$$\text{Adjusted C}_Z = \text{C}_Z - \frac{1}{\langle Z\rangle H_{ZZ}}. \tag{B8}$$

APPENDIX C: DERIVATION OF MOMENT EQS. (3), (4), AND (5)

Let us consider the first moment $\langle X\rangle$. Noticing that $\langle X\rangle = \sum_{i,j,k} i\varphi_{(i,j,k)}$, we multiply the master equation Eq. (1) by i, and then sum over i, j, and k. We have

$$\begin{aligned}
\frac{d\langle X\rangle}{dt} &= \langle X^2\rangle r_1p_1 + \langle X\rangle r_1p_1 + \langle X^2\rangle r_1q_1 - \langle X\rangle r_1q_1 + \langle X^3\rangle\alpha_1 - 2\langle X^2\rangle\alpha_1 + \langle X\rangle\alpha_1 + \langle XY\rangle r_2p_2'\\
&\quad + \langle XY\rangle r_2q_2' + \langle XY^2\rangle\alpha_2 - \langle XY\rangle\alpha_2 + \langle XZ\rangle d + \langle XZ^2\rangle\alpha_3 - \langle XZ\rangle\alpha_3 + \langle XY\rangle r_2\delta_2 + 2\langle Y\rangle r_2\delta_2\\
&\quad - (\langle X^2\rangle r_1p_1 + \langle X^2\rangle r_1q_1 + \langle X^3\rangle\alpha_1 - \langle X^2\rangle\alpha_1 + \langle XY\rangle r_2p_2' + \langle XY\rangle r_2q_2' + \langle XY^2\rangle\alpha_2\\
&\quad - \langle XY\rangle\alpha_2 + \langle XY\rangle r_2\delta_2 + \langle XZ\rangle d + \langle XZ^2\rangle\alpha_3 - \langle XZ\rangle\alpha_3)\\
&= \langle X\rangle r_1p_1 - \langle X\rangle r_1q_1 + \langle X\rangle\alpha_1 - \langle X^2\rangle\alpha_1 + 2\langle Y\rangle r_2\delta_2.
\end{aligned} \tag{C1}$$

Similarly, for $\langle Y\rangle$ and $\langle Z\rangle$ we have

$$\frac{d\langle Y\rangle}{dt} = 2\langle X\rangle r_1q_1 + \langle Y\rangle(r_2p_2' - r_2q_2' - r_2\delta_2 + \alpha_2) - \langle Y^2\rangle\alpha_2 \tag{C2}$$

and

$$\frac{d\langle Z\rangle}{dt} = 2\langle Y\rangle r_2q_2' + \langle Z\rangle(-d + \alpha_3) - \langle Z^2\rangle\alpha_3. \tag{C3}$$

For the second moment $\langle X^2\rangle$, noticing that $\langle X^2\rangle = \sum_{i,j,k} i^2\varphi_{(i,j,k)}$, and based on Eq. (1) we have

$$\begin{aligned}\frac{d\langle X^2\rangle}{dt} = {} & \langle X^3\rangle r_1p_1 + 2\langle X^2\rangle r_1p_1 + \langle X\rangle r_1p_1 + \langle X^3\rangle r_1q_1 - 2\langle X^2\rangle r_1q_1 + \langle X\rangle r_1q_1 + \langle X^4\rangle\alpha_1\\ & - 3\langle X^3\rangle\alpha_1 + 3\langle X^2\rangle\alpha_1 - \langle X\rangle\alpha_1 + \langle X^2Y\rangle r_2p_2' + \langle XY^2\rangle r_2q_2' + \langle X^2Y^2\rangle\alpha_2 - \langle X^2Y\rangle\alpha_2\\ & + \langle X^2Z\rangle d + \langle X^2Z^2\rangle\alpha_3 - \langle X^2Z\rangle\alpha_3 + \langle X^2Y\rangle r_2\delta_2 + 4\langle XY\rangle r_2\delta_2 + 4\langle Y\rangle r_2\delta_2\\ & - (\langle X^3\rangle r_1p_1 + \langle X^3\rangle r_1q_1 + \langle X^4\rangle\alpha_1 - \langle X^3\rangle\alpha_1 + \langle X^2Y\rangle r_2p_2' + \langle X^2Y\rangle r_2q_2' + \langle X^2Y^2\rangle\alpha_2\\ & - \langle X^2Y\rangle\alpha_2 + \langle X^2Y\rangle r_2\delta_2 + \langle X^2Z\rangle d + \langle X^2Z^2\rangle\alpha_3 - \langle X^2Z\rangle\alpha_3) \qquad \text{(C4)}\\ = {} & \langle X^2\rangle(2r_1p_1 - 2r_1q_1 + 3\alpha_1) + \langle X\rangle(r_1q_1 + r_1q_1 - \alpha_1) - 2\langle X^3\rangle\alpha_1\\ & + 4\langle XY\rangle r_2\delta_2 + 4\langle Y\rangle r_2\delta_2. \qquad \text{(C5)}\end{aligned}$$

Similarly, for $\langle Y^2\rangle$ and $\langle Z^2\rangle$, we have

$$\frac{d\langle Y^2\rangle}{dt} = \langle Y^2\rangle(2r_2p_2' - 2r_2q_2' + 3\alpha_2 - 2r_2\delta_2) + \langle Y\rangle(r_2p_2' + r_2q_2' + r_2\delta_2 - \alpha_2) - 2\langle Y^3\rangle\alpha_2 + 4\langle XY\rangle r_1q_1 + 4\langle X\rangle r_1q_1 \tag{C6}$$

and

$$\frac{d\langle Z^2\rangle}{dt} = \langle Z^2\rangle(3\alpha_3 - 2d) + \langle Z\rangle(d - \alpha_3) - 2\langle Z^3\rangle\alpha_3 + 4\langle YZ\rangle r_2q_2' + 4\langle Y\rangle r_2q_2'. \tag{C7}$$

For the mixed moment $\langle XY\rangle$, noticing that $\langle XY\rangle = \sum_{i,j,k} ij\varphi_{(i,j,k)}$, and based on Eq. (1) we have

$$\begin{aligned}\frac{d\langle XY\rangle}{dt} = {} & \langle X^2Y\rangle r_1p_1 + \langle XY\rangle r_1p_1 + \langle X^2Y\rangle r_1q_1 + 2\langle X^2\rangle r_1q_1 - \langle XY\rangle r_1q_1 - 2\langle X\rangle r_1q_1 + \langle X^3Y\rangle\alpha_1\\ & - 2\langle X^2Y\rangle\alpha_1 + \langle XY\rangle\alpha_1 + \langle XY^2\rangle r_2p_2' + \langle XY\rangle r_2p_2' + \langle XY^2\rangle r_2q_2' - \langle XY\rangle r_2q_2' + \langle XY^3\rangle\alpha_2\\ & - 2\langle XY^2\rangle\alpha_2 + \langle XY\rangle\alpha_2 + \langle XYZ\rangle d + \langle XYZ^2\rangle\alpha_3 - \langle XYZ\rangle\alpha_3 + \langle XY^2\rangle r_2\delta_2 - \langle XY\rangle r_2\delta_2\\ & + 2\langle Y^2\rangle r_2\delta_2 - 2\langle Y\rangle r_2\delta_2\\ & - (\langle X^2Y\rangle r_1p_1 + \langle X^2Y\rangle r_1q_1 + \langle X^3Y\rangle\alpha_1 - \langle X^2Y\rangle\alpha_1 + \langle XY^2\rangle r_2p_2' + \langle XY^2\rangle r_2q_2' + \langle XY^3\rangle\alpha_2\\ & - \langle XY^2\rangle\alpha_2 + \langle XY^2\rangle r_2\delta_2 + \langle XYZ\rangle d + \langle XYZ^2\rangle\alpha_3 - \langle XYZ\rangle\alpha_3)\\ = {} & \langle XY\rangle(r_1p_1 - r_1q_1 + \alpha_1 + r_2p_2' - r_2q_2' - r_2\delta_2 + \alpha_2) + 2\langle X^2\rangle r_1q_1 - 2\langle X\rangle r_1q_1\\ & + 2\langle Y^2\rangle r_2\delta_2 - 2\langle Y\rangle r_2\delta_2 - \langle X^2Y\rangle\alpha_1 - \langle XY^2\rangle\alpha_2. \qquad \text{(C8)}\end{aligned}$$

Similarly, for $\langle YZ\rangle$ and $\langle XZ\rangle$ we have

$$\begin{aligned}\frac{d\langle YZ\rangle}{dt} = {} & \langle YZ\rangle(r_2p_2' - r_2q_2' - r_2\delta_2 + \alpha_2 - d + \alpha_3) + 2\langle XZ\rangle r_1q_1\\ & + 2\langle Y^2\rangle r_2q_2' - 2\langle Y\rangle r_2q_2' - \langle Y^2Z\rangle\alpha_2 - \langle YZ^2\rangle\alpha_3\end{aligned} \tag{C9}$$

and

$$\begin{aligned}\frac{d\langle XZ\rangle}{dt} = {} & \langle XZ\rangle(r_1p_1 - r_1q_1 + \alpha_1 - d + \alpha_3) + 2\langle YZ\rangle r_2\delta_2\\ & - \langle X^2Z\rangle\alpha_1 - \langle XZ^2\rangle\alpha_3 + 2\langle XY\rangle r_2q_2'.\end{aligned} \tag{C10}$$

APPENDIX D: ABOUT THE BACKGROUND PARAMETERS

Noticing that our interest is how dedifferentiation affects the noise in the numbers of different cell types, dedifferentiation probability δ_2 and reshaping factor κ are the focus of this study. We treat all the other parameters as background. These parameters can be classified as three different types. The first is about cell division mode (e.g., r_1p_1 represents symmetric division rate of each stem cell). The second is about cell death of terminally differentiated cells. The third is about cell competition (e.g., α_1 represents cell competition strength between stem cells).

Instead of investigating some specific multicellular systems, we are more interested in the conceptual models, so the setup of all the background parameters are based on a very simple but common biological fact that in the absence of dedifferentiation, the population size is amplifying from stem cells compartment to terminally differentiated cells compartment [39–41], i.e., it is required that $0 < \langle X\rangle < \langle Y\rangle < \langle Z\rangle$. Figure 2 illustrates a typical case that realizes the requirement. The main idea is that the carrying capacity of stem cells compartment is the smallest, while the carrying capacity of terminally differentiated cells compartment is the largest. In this way, the equilibrium population size of stem cells is smaller than that of transient amplifying

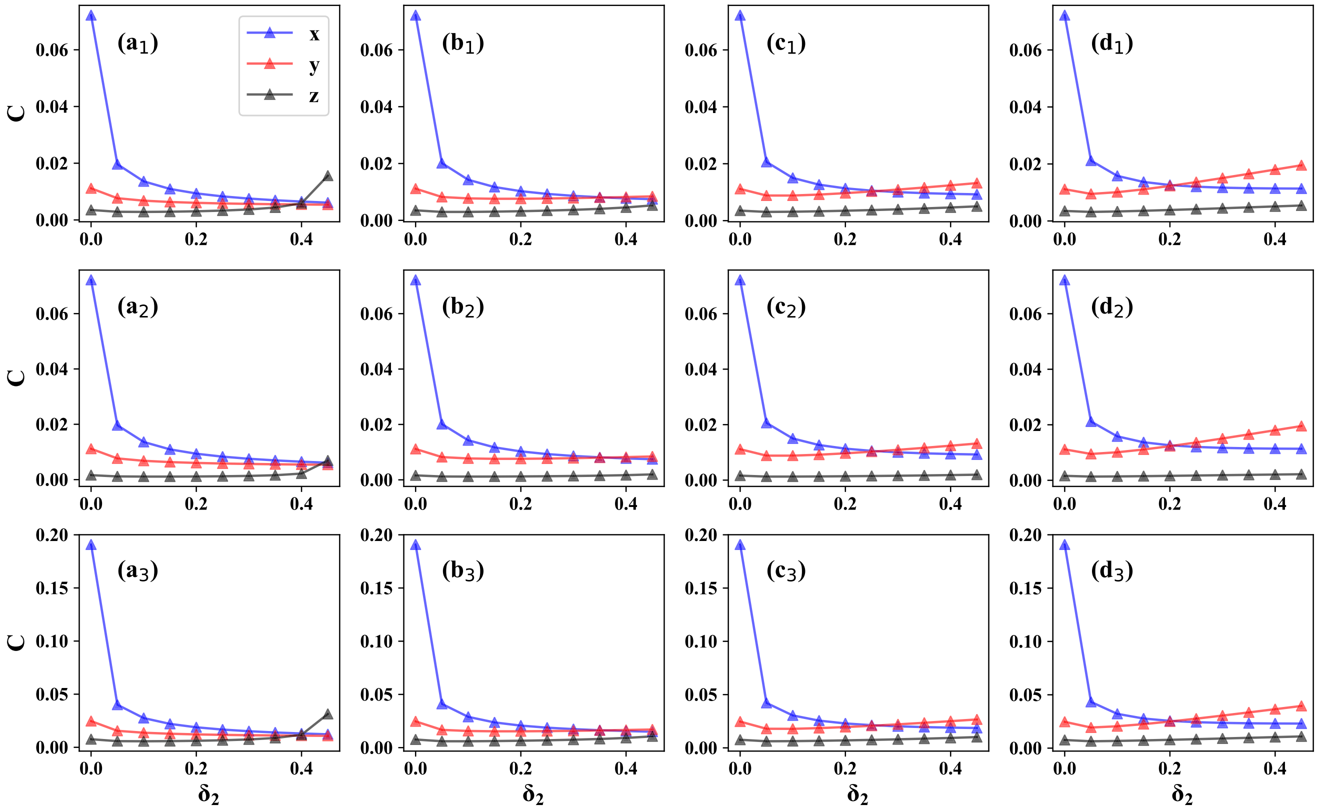

FIG. 6. Illustration of the robustness of the results in Fig. 3. In the fist row, $\alpha_1 = 0.01, \alpha_2 = 0.001, \alpha_3 = 0.001$; in the second row, $\alpha_1 = 0.015, \alpha_2 = 0.002, \alpha_3 = 0.002$; in the third row, $\alpha_1 = 0.01, \alpha_2 = 0.001, \alpha_3 = 0.001$; the joint parameters are $(p_1, q_1, p_2, q_2, r_1, r_2, d) = (0.7, 0.3, 0.6, 0.4, 0.7, 0.7, 0.01)$.

cells which is also smaller than that of terminally differentiated cells. However, noticing that there are numerous combinations of parameters that can realizes this requirement, Fig. 6 shows the robustness of our main result on the background parameters by changing the competition strengths α_1, α_2, and α_3.

APPENDIX E: RELATION BETWEEN C AND MEAN POPULATION SIZE

Figure 7 shows C_X (or Adjusted C_X) as a function of the inverse of the mean popoulation size of stem cells. We can see that C_X (or Adjusted C_X) is approximately proportional to the inverse of the mean population size.

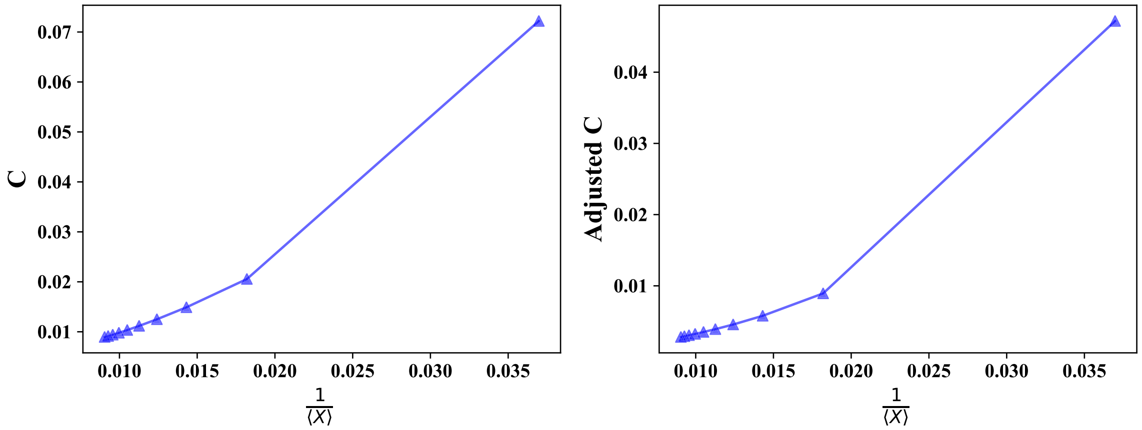

FIG. 7. Illustration of the relation between C_X (or Adjusted C_X) and the inverse of the mean population size of stem cells. The background parameters are $(\kappa, p_1, q_1, p_2, q_2, r_1, r_2, \alpha_1, \alpha_2, \alpha_3, d) = (0.5, 0.7, 0.3, 0.6, 0.4, 0.7, 0.7, 0.01, 0.001, 0.001, 0.01)$.

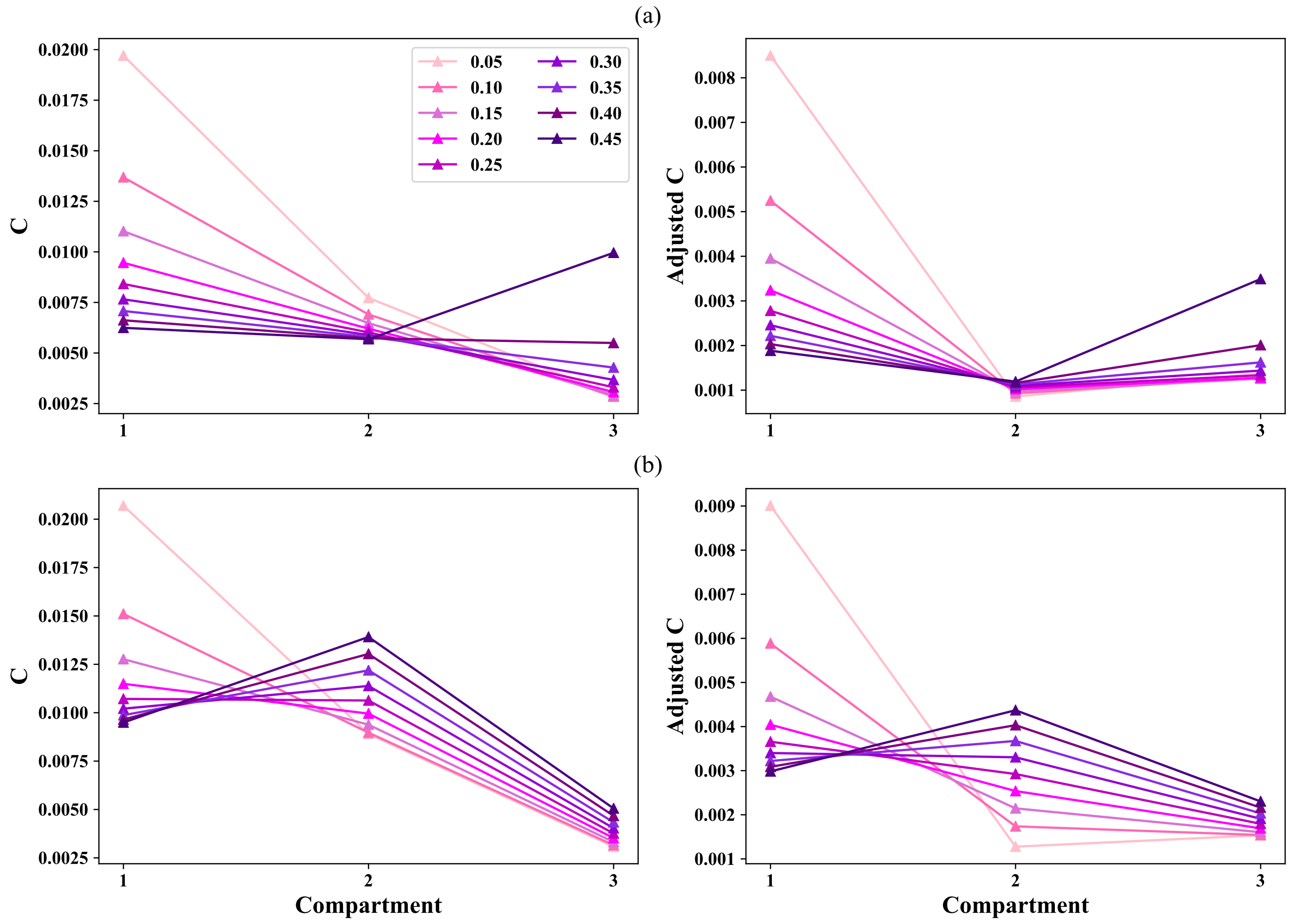

FIG. 8. The 2D version of Fig. 4. In panel (a) $\kappa = 0.155$, and in panel (b) $\kappa = 0.555$. Different colors correspond to different values of dedifferentiation probability δ_2.

APPENDIX F: TWO-DIMENSIONAL PLOT OF FIG. 4

Figure 8 shows the 2D version of Fig. 4.

---

[1] T. Tumbar, G. Guasch, V. Greco, C. Blanpain, W. E. Lowry, M. Rendl, and E. Fuchs, Science **303**, 359 (2004).
[2] B. Werner, D. Dingli, T. Lenaerts, J. M. Pacheco, and A. Traulsen, PLoS Comput. Biol. **7**, e1002290 (2011).
[3] M. C. Mackey, Cell Prolif. **34**, 71 (2001).
[4] E. A. McCulloch and J. E. Till, Nat. Med. **11**, 1026 (2005).
[5] C. Jordan, M. Guzman, and M. Noble, N. Engl. J. Med. **355**, 1253 (2006).
[6] L. Nguyen, R. Vanner, P. Dirks, and C. Eaves, Nat. Rev. Cancer **12**, 133 (2012).
[7] C. Chaffer, I. Brueckmann, C. Scheel, A. Kaestli, P. Wiggins, L. Rodrigues, M. Brooks, F. Reinhardt, Y. Su, K. Polyak *et al.*, Proc. Natl. Acad. Sci. USA **108**, 7950 (2011).
[8] P. Gupta, C. Fillmore, G. Jiang, S. Shapira, K. Tao, C. Kuperwasser, and E. Lander, Cell **146**, 633 (2011).
[9] P. R. Tata, H. Mou, A. Pardo-Saganta, R. Zhao, M. Prabhu, B. M. Law, V. Vinarsky, J. L. Cho, S. Breton, A. Sahay *et al.*, Nature (London) **503**, 218 (2013).
[10] C. E. Meacham and S. J. Morrison, Nature (London) **501**, 328 (2013).
[11] A. Jilkine, Curr. Stem Cell Rep. **5**, 66 (2019).
[12] A. Jilkine and R. N. Gutenkunst, PLoS Comput. Biol. **10**, e1003481 (2014).
[13] D. Wodarz, J. Theor. Biol. **448**, 86 (2018).
[14] K. Leder, K. Pitter, Q. LaPlant, D. Hambardzumyan, B. D. Ross, T. A. Chan, E. C. Holland, and F. Michor, Cell **156**, 603 (2014).
[15] X. Chen, Y. Wang, T. Feng, M. Yi, X. Zhang, and D. Zhou, J. Theor. Biol. **390**, 40 (2016).
[16] L. Cai, N. Friedman, and X. Xie, Nature (London) **440**, 358 (2006).
[17] J. Paulsson, Nature (London) **427**, 415 (2004).
[18] N. Eling, M. D. Morgan, and J. C. Marioni, Nat. Rev. Genet. **20**, 536 (2019).
[19] J. Paulsson, O. G. Berg, and M. Ehrenberg, Proc. Natl. Acad. Sci. USA **97**, 7148 (2000).

[20] M. Kittisopikul and G. M. Süel, Proc. Natl. Acad. Sci. USA **107**, 13300 (2010).
[21] G. Balázsi, A. van Oudenaarden, and J. J. Collins, Cell **144**, 910 (2011).
[22] L. J. Allen, *An Introduction to Stochastic Processes with Applications to Biology* (CRC Press, Boca Raton, FL, 2010).
[23] D. Dingli, A. Traulsen, and F. Michor, PLoS Comput. Biol. **3**, e53 (2007).
[24] Z. Wu, Y. Wang, K. Wang, and D. Zhou, Math. Biosci. **332**, 108541 (2021).
[25] Z. Sun and N. L. Komarova, J. Math. Biol. **71**, 883 (2015).
[26] Q. M. Pei, X. Zhan, L. J. Yang, J. Shen, L. F. Wang, K. Qui, T. Liu, J. B. Kirunda, A. A. M. Yousif, A. B. Li, and Y. Jia, Phys. Rev. E **92**, 012721 (2015).
[27] X. F. Hou, B. Q. Zhou, Y. F. Zhou, C. O. Apata, L. Jiang, and Q. M. Pei, Phys. Rev. E **102**, 062411 (2020).
[28] C. W. Gardiner, *A Handbook of Stochastic Methods: For Physics, Chemistry and the Natural Sciences*, 4th Printing (Springer, Berlin, 1991).
[29] D. Schnoerr, G. Sanguinetti, and R. Grima, J. Chem. Phys. **143**, 185101 (2015).
[30] M. D. Johnston, C. M. Edwards, W. F. Bodmer, P. K. Maini, and S. J. Chapman, Proc. Natl. Acad. Sci. USA **104**, 4008 (2007).
[31] D. Zhou, Y. Luo, D. Dingli, and A. Traulsen, PLoS Comput. Biol. **15**, e1007167 (2019).
[32] W. Lo, C. Chou, K. K. Gokoffski, F. Y.-M. Wan, A. D. Lander, A. L. Calof, and Q. Nie, Math. Biosci. Eng. **6**, 59 (2009).
[33] A. D. Lander, K. K. Gokoffski, F. Y. M. Wan, Q. Nie, and A. L. Calof, PLoS Biol. **7**, e1000015 (2009).
[34] H. J. Park and A. Traulsen, Phys. Rev. E **96**, 042412 (2017).
[35] D. T. Gillespie, J. Phys. Chem. **81**, 2340 (1977).
[36] C. S. Gillespie, IET. Syst. Biol. **3**, 52 (2009).
[37] T. Shibata and K. Fujimoto, Proc. Natl. Acad. Sci. USA **102**, 331 (2005).
[38] G. Hornung and N. Barkai, PLoS Comput. Biol. **4**, e8 (2008).
[39] D. Dingli, A. Traulsen, and J. M. Pacheco, PLoS ONE **2**, e345 (2007).
[40] D. Dingli, J. M. Pacheco, and A. Traulsen, Phys. Rev. E **77**, 021915 (2008).
[41] B. Werner, D. Dingli, and A. Traulsen, J. R. Soc. Interface **10**, 20130349 (2013).