Computers and Chemical Engineering 112 (2018) 27–36

Contents lists available at ScienceDirect

Computers and Chemical Engineering

journal homepage: www.elsevier.com/locate/compchemeng

Effect of cell heterogeneity on isogenic populations with the synthetic genetic toggle switch network: Bifurcation analysis of two-dimensional cell population balance models

Panagiotis Chrysinas, Michail E. Kavousanakis*, Andreas G. Boudouvis

National Technical University of Athens, School of Chemical Engineering, Athens 15780, Greece

ARTICLE INFO

Article history:
Received 22 August 2017
Revised 17 January 2018
Accepted 29 January 2018
Available online 31 January 2018

Keywords:
Cell heterogeneity
Toggle switch
Bistability
Broyden's algorithm

ABSTRACT

The dynamics of gene regulatory networks are often modeled with the assumption of cellular homogeneity. However, this assumption contradicts the plethora of experimental results in a variety of systems, which designates that cell populations are heterogeneous systems in the sense that properties such as size, shape, and DNA/RNA content are unevenly distributed amongst their individuals. In order to address the implications of heterogeneity, we utilize the so-called cell population balance (CPB) models. Here, we solve numerically multivariable CPB models to study the effect of heterogeneity on populations carrying the toggle switch network, which features nonlinear behavior at the single-cell level. In order to answer whether this nonlinear behavior is inherited to the heterogeneous population level, we perform bifurcation analysis on the steady-state solutions of the CPB model. We show that bistability is present at the population level with the pertinent bistability region shrinking when the impact of heterogeneity is enhanced.

1. Introduction

Advances that recently occurred in the fields of biotechnology, genomics and computational biology have supplied us with powerful techniques and methods that can shed light on the complex mechanisms taking place at the single-cell level. However, it is of equally great importance to understand the impact of intra and inter-cellular processes on the average population phenotype. In particular, we are interested in studying the effect of cell population heterogeneity, which has been observed in numerous biological systems. Indicatively, we report the burst variation of bacteriophages (Delbruck, 1945), and the existence of transcriptional states of heterogeneity in sporulating cultures of *Bacilus subtillis* (Chung and Stephanopoulos, 1995). Cellular heterogeneity has also been observed in various isogenic *Escherichia coli* systems (Elowitz et al., 2002), in endothelial cell surface markers (Oh et al., 2004), transcriptional states at single-cell-resolution (Tischler and Surani, 2013) and single-cell metabolomics (Rubakhin et al., 2013). Finally, we report recent studies showing the impact of heterogeneity on drug discovery and optimal design of therapeutic strategies (Gough et al., 2017, 2014).

Despite the experimental evidence for the importance of cell heterogeneity, a number of modeling approaches (e.g., Chung and Stephanopoulos, 1995; Fedoroff and Fontana, 2002; Sadeghpour et al., 2017) are based on the assumption that populations are homogeneous. Despite the fact, that this assumption leads to simple mathematical models (systems of ordinary differential equations), disregarding cell heterogeneity can lead to false quantitative predictions (Aviziotis et al., 2015a, 2015b; Kavousanakis et al., 2009; Mantzaris, 2005; McAdams and Arkin, 1998).

In this work, we adopt the modeling approach of cell population balance (CPB) to model the dynamics and compute the steady-state solution of heterogeneous isogenic populations, (all individuals carry the same genetic network). In such populations, cell heterogeneity originates from two main sources. The first source, the so-called *intrinsic* heterogeneity, is the result of stochastic fluctuations of regulatory molecules (Alberts et al., 1994), which exist in small concentrations and control a network of intracellular reactions. Gene regulatory molecules are a set of DNA segments inside the cell that interact with each other through their RNA and protein expression products as well as with other intracellular substances. The type and the number of genes expressed at each moment alongside with the intracellular reactions define the phenotype of each cell. Furthermore, gene expression is a stochastic process as shown in Blake et al. (2003); Elowitz et al. (2002) leading

* Corresponding author.
E-mail address: mihkavus@chemeng.ntua.gr (M.E. Kavousanakis).

https://doi.org/10.1016/j.compchemeng.2018.01.021
0098-1354/

also to phenotypic variability, which originates from intracellular processes.

The second source of heterogeneity is the so-called *extrinsic* heterogeneity, which is the result of the uneven distribution of the intracellular content -with the exception of DNA- from a mother cell to its daughter cells during cellular division. The uneven distribution of mother content to the offsprings results in different phenotypes as a result of the different rates of the intracellular reaction network. Furthermore, it is not only the intracellular content which is distributed unevenly; the regulatory molecules are also unevenly distributed, and the phenomenon repeats itself due to the process of cell cycle leading to further phenotypic variability. It has been shown by experimental studies (Elowitz et al., 2002), for *E. coli* populations, that *extrinsic* heterogeneity has a more significant quantitative impact; in this work we focus on the *extrinsic* heterogeneity impact on *E. coli* populations carrying the genetic toggle switch.

In order to quantify the heterogeneity and combine it with the pertinent genetic network, we introduce the CPB models, which were developed in mid-1960s (Eakman et al., 1966; Fredrickson et al., 1967; Tsuchiya et al., 1966). They are partial integro-differential equations and are characterized from high mathematical complexity (even with the application of model-reduction techniques (Stamatakis, 2013)). Analytical solutions cannot be obtained for the general case and the use of numerical methods is mandatory (Liou et al., 1997; Mantzaris et al., 2001a, 2001b, 2001c; Zhang et al., 2003, 2002; Zhu et al., 2000). However, a common feature of the applied numerical methods is the assumption that the physiological state space boundaries (e.g., the boundaries of the intracellular content) are known *a priori*. This assumption may be valid for the minimum intracellular content -which we can assume that is equal to zero- but this does not apply when it comes to the value of the maximum intracellular content. In order to bypass this impediment one can apply a free boundary formulation as presented in Kavousanakis et al. (2009), based on a valid assumption that the maximum intracellular content is a positive multiple of its average value.

The mathematical formulation of the applied free boundary CPB model is described in Section 2. In particular, we present a two-variable CPB model in order to describe the dynamics of *E. coli* cells carrying a synthetic toggle switch which has been presented in Gardner et al. (2000). A brief description of its design and mathematical formulation is provided in Section 3. A key feature of this synthetic network is its nonlinear behavior and the existence of a range of extracellular inducer concentration values - IPTG (isopropyl-β-D-thiogalactopyranoside)- with multiple co-existing steady-state phenotypes. In order to examine whether this nonlinear behavior is inherited also to the population level, we first study homogeneous populations using systems of ODEs which describe their dynamics, and the *pseudo* arc-length continuation algorithm (Keller, 1977) to track the entire steady-state solution space as a function of the [IPTG].

The study of heterogeneous populations is presented in Section 4, where we utilize the *pseudo* arc-length method in combination with CPBs, in order to determine and quantify the impact of heterogeneity on the range of bistability (the interval of [IPTG] values with multiple solutions). We need to stress at this point, that the steady-state solution of multivariable CPBs is not a trivial numerical task, with significantly large computational and memory requirements. In order to bypass these difficulties we resort to Newton-like algorithms, and in particular Broyden's algorithm, (Broyden, 1965), which requires only an approximation of the Jacobian matrix, and not the Jacobian matrix itself (as required in Newton–Raphson), thus saving significant computational effort as compared to Newton's method.

In Section 5, we present temporal and steady-state computations for the aforementioned CPB model, which is discretized with the finite element method. We also present the steady-state solution space of heterogeneous populations carrying the synthetic toggle switch as a function of the IPTG concentration. The pertinent bifurcation diagrams show that bistability is also present for heterogeneous cell populations, however the range is narrowed down as the impact of heterogeneity is enhanced. Furthermore, we also study the impact of other parameters on the range of bistability, including the parameters which quantify the asymmetry and sharpness of the division mechanism. Finally, in Section 6 we provide a brief summary of the main results of this study.

2. Cell population balance modeling

In this work, we study a two-dimensional CPB model, which describes the dynamics of a heterogeneous population carrying the synthetic toggle switch (Gardner et al., 2000). Each individual of the evolved distribution is characterized by the values of two intracellular variables, namely x and y. In the more general case of a k-variable CPB model, each cell is characterized by a vector of k intracellular content values, $\underline{x} \equiv (x_1, \ldots, x_k)$, the dynamics of the population are described by the following expression (Mantzaris, 2006):

$$\frac{\partial u(\underline{x},t)}{\partial t} + \nabla_{\underline{x}} \cdot [\underline{R}(\underline{x}) u(\underline{x},t)] + \Gamma(\underline{x}) u(\underline{x},t) = 2 \int_{\underline{x}}^{\underline{x}_{max}} \Gamma(\underline{x}') P(\underline{x},\underline{x}') u(\underline{x}',t) \mathrm{d}^k \underline{x}' - u(\underline{x},t) \int_{\Lambda} \Gamma(\underline{x}) u(\underline{x},t) \mathrm{d}^k \underline{x}, \quad (2.1)$$

where:

$$\Lambda = [0, x_{1,max}] \times \ \ldots \ \times [0, x_{k,max}] \subseteq \mathbb{R}^k, k \in \mathbb{N}, \quad (2.2)$$

and $\underline{x}_{\max} \equiv (\underline{x}_{1,\max}, \ldots, x_{k,\max})$ denotes the vector with the maximum intracellular content values. The number density function, $u(\underline{x},t)$, (Fredrickson et al., 1967), expresses the number of cells with content $\underline{x}$ at time t divided by the total number of cells at this time. The boundary conditions imposed to (2.1) require that the population cells do not grow outside the domain, Λ, i.e.:

$$u(\underline{0},t) = u(\underline{x}_{max},t) = 0. \quad (2.3)$$

The first term in (2.1) quantifies accumulation, and the second denotes the rate at which cells with intracellular content $\underline{x}$ change their content due to intracellular reactions, $\underline{R}(\underline{x})$. The third term represents division, which yields cells with lower content, when the cell division rate is $\Gamma(\underline{x})$. The first term at the right hand side describes the birth of cells with content, $\underline{x}$, by cells with larger intracellular content. The factor 2 multiplies the integral to model the birth of two cells at the end of each division. The function, $P(\underline{x},\underline{x}')$, models the mechanism of intracellular content distribution amongst the two daughter cells; in effect $P(\underline{x},\underline{x}')$ models the probability that a mother cell with content, x', produces a daughter cell with content, x, and one of content, x'-x. Finally, the last term of the right hand side (dilution term) is the one forcing the solution to reach a steady-state; at this state, the non-normalized distribution of cells reaches a time-invariant shape, while cells continue to proliferate.

Taking into account that $u(\underline{x},t)$ is the number density function (already normalized by the total number of cells), the following condition must apply:

$$\int_{\Lambda} u(\underline{x},t) \mathrm{d}^k \underline{x} = 1. \quad (2.4)$$

Eq. (2.1) incorporates single-cell operations through three key functions: $\Gamma(\underline{x})$, $\underline{R}(\underline{x})$ and $P(\underline{x},\underline{x}')$ known in the relative literature