Inverse problem for cell division rate in population dynamics

Léon Matar-Tine
Université de Lyon 1, UMR 5208, Institut Camille Jordan, INRIA Grenoble Rhône-Alpes-Dracula, France

Abstract. In population dynamics, more precisely in cell structured population, experimental measures of the temporal dynamics of some parameters are often difficult or even impossible because of the lack of appropriate technology. It is the case of the division rate parameter for a structured population of cells. So theoretical and numerical approaches are needed to recover such parameters by indirect observation of the solution of the modeling system. In this work we present how to estimate the division rate from a growth-fragmentation type equation as done in the work of Doumic and Tine (see [1]) and following the work done in [2] and [3]. The estimation strategy is based on the observation of the long-term solution of the model for a size structured population when the growth and division rates are fully general. The issue of the theoretical and numerical results are discussed here.

1. Introduction

The observation of a population of cells (proteins, bacteria, etc.) allows to understand the dynamics of the considered population such as its distribution, the behavior of its parameters (growth rate, death rate, division rate). This understanding depends on technologies that biologists have for experiences. For instance, one knows that from cytometric flow technology it is possible to measure at each time the distribution of the population what is not the case for instance for the measurement of the division rate because of the lack of technology able to follow a single cell from its birth to its division. So, to estimate this division rate which is a crucial parameter (in the case of pathogenic cells as E. Coli bacteria), mathematical and numerical approaches are needed. Then without lost of generality let consider the following growth-fragmentation type model:

\[
\begin{align*}
\hat{c}_1 n(t, x) + \hat{c}_x (g(x)n(t, x)) + B(x)n(t, x) &= 2 \int_0^\infty B(y)\kappa(x, y)n(t, y) \, dy, \\
n(0, x) &= n_0(x), \quad x \geq 0, \\
g(0)n(t, 0) &= 0, \quad t \geq 0,
\end{align*}
\]

where \(n(t, x)\) is the density repartition function of the cells structured by their size variable \(x\) at time \(t\). We assume the cell population to evolve only by growth and division process without death and birth events. The growth rate \(g\) is a given or measured function which depends on the size \(x\). The total rate
of division for cell of size $x$ is denoted by $B(x)$, so the product $B(y)\kappa(x, y)$ is the rate at which a single cell of size $y$ divides into cells of size respectively $x$ and $y - x$. The kernel division $\kappa(x, y)$ is required to satisfy:

$$\kappa(x, y) = 0 \quad \forall \ x > y, \quad \int_0^\infty \kappa(x, y) \, dx = 1, \quad \kappa(x, y) = \kappa(y - x, y), \quad \int_0^\infty x \kappa(x, y) \, dx = \frac{y}{2}. \quad (2)$$

With this model (1) we can formally deduce by simple integration that, in one hand the time variation of the total number of cells depends on the division rate $B$ what means

$$\frac{d}{dt} \int_0^\infty n(t, x) \, dx = \int_0^\infty B(x)n(t, x) \, dx.$$ 

In the other hand, one deduce that the time variation of the total mass of the population depends on the growth rate $g$ what means

$$\frac{d}{dt} \int_0^\infty x n(t, x) \, dx = \int_0^\infty g(x)n(t, x) \, dx.$$ 

The aim of this paper is to show how to estimate the division rate $B$ from the knowledge of measurements on the density repartition $n(t, x)$ and on the growth rate $g(x)$. Let us recall that these assumed known quantities can be experimentally measured by cytometric flows technology, while the division rate need theoretical and numerical approaches for its estimation. Estimating this division rate is crucial if one is interested of pathologic cells (as bacteria or misfolding proteins) which cause disease. Work in this field allows us to focus the estimation of the division rate on the time asymptotic behavior of the model (1).

The fact that, at any time we can have observation of the size distribution of the cell population allows us to focus the estimation of the division rate on the time asymptotic behavior of the model (1).

So the paper is organized as follow: in Sect. 1, we present the asymptotic behavior of the model and related results proved by generalized relative entropy principle (GRE). In Sect. 2 we present the inverse problem of estimation the division rate $B$ thanks to two regulation methods. The Sect. 3, is devoted to the numerical methods and results. The last section is for the conclusion.

2. Asymptotic behavior of the model

The asymptotic behavior of the model (1) is based on the GRE principle (see [8]) which need basic hypothesis on the parameters.

With the notation

$$\mathbb{P} := \left\{ f \in L^1_{\text{loc}}(\mathbb{R}^+_\nu), \ f \geq 0 : \ \exists \mu, \nu \geq 0, \ \lim_{x \to \infty} \sup_{x^{-\mu}} f(x) < \infty, \ \lim_{x \to \infty} \inf_{x^{\nu}} f(x) > 0 \right\},$$

we assume:

the kernel division fulfill in addition to (2), the relations

$$\exists \ c > 0, \ \forall p \geq 2, \ \int_0^y \frac{x^p}{y^p} \kappa(x, y) \, dx \leq c < \frac{1}{2}. \quad (3)$$

The growth rate and the division rate satisfy

$$B \in L^1_{\text{loc}}(\mathbb{R}^+_\nu) \cap \mathbb{P}, \ \exists \nu_0 \geq 0, \ \forall t \geq 0, \ g \in L^\infty_{\text{loc}}(\mathbb{R}^+_\nu, x^{\nu_0} \, dx) \cap \mathbb{P}. \quad (4)$$
and for all compact set $K \subset \mathbb{R}^+$,
\[ \exists m_K > 0 \text{ such that } g(x) \geq m_K \forall x \in K. \]
\[ \exists b \geq 0, \text{ supp} B = [b, \infty[. \] (5)

The cells of size zero forming is avoided thanks to the assumption
\[ \exists C > 0, \eta \geq 0, \text{ such that } \int_0^1 \kappa(z, y) dz \leq \min\left(1, C \left(\frac{x}{y}\right)^{\eta}\right) \text{ and } \frac{x^\eta}{g(x)} \in L^1_0. \] (6)

For the cells of small size, the growth is assumed to be more important than de division, that means
\[ \frac{B(x)}{g(x)} \in L^1_0 \text{ with } L^1_0 = L^1(0, a), \ a > 0. \] (7)

For the cells of big size, in order to avoid the formation of infinite size cells at finite time (gelation phenomenon), the growth and division rates satisfy
\[ \lim_{x \to +\infty} \frac{x}{g(x)} = +\infty, \ \forall t \geq 0. \] (8)

Thanks to the general assumptions (3)–(8), the existence, uniqueness and properties of eigenelements associated to (1) is proved in [4]. In particular it is shown by the GRE principle that
\[ \int_0^\infty |n(t, x)e^{-\lambda t} - \rho_0 N(x)|\phi(x) dx \to 0; \ \rho_0 = \int_0^\infty n^0(x)\phi(x)dx, \]
where $(\lambda, N, \phi)$ is the unique eigentriplet solution of the following stationary equation and its adjoint equation:
\[ \begin{cases} \ c\hat{\lambda}(g(x)N(x)) + (B(x) + \lambda)N(x) = 2 \int_0^{+\infty} B(y)\kappa(x, y)N(y) dy, \\ gN(x = 0) = 0; \ N(x) \geq 0; \int_0^\infty N(x) dx = 1, \ \lambda > 0. \end{cases} \] (9)

\[ \begin{cases} \ cg(x)\hat{\phi}(x) - (B(x) + \lambda)\phi(x) = -2B(x) \int_0^x \kappa(y, x)\phi(y) dy, \\ \phi(x) \geq 0; \int_0^{\infty} \phi(x)N(x) dx = 1. \end{cases} \] (10)

The estimation approaches of the division rate we develop here is based on the problem (9) and the assumption on the knowledge of $g, N$ or their experimental measurements. Of course there is two fundamental estimates for the eigenvalue $\hat{\lambda}$ from simple integration of (9) which are written as follow
\[ \hat{\lambda} = \int_0^\infty B(x)N(x) dx \text{ and/or } \int_0^\infty x N(x) dx = \frac{1}{\hat{\lambda}} \int_0^\infty g(x)N(x) dx. \]

3. Estimation of the division rate

In this part, we investigate the question of recovering the division rate $B$ when the growth rate $g$ is given and a sequence of measurements $N_\varepsilon$ such that, with an adequate norm they fulfill, $\|N - N_\varepsilon\| < \varepsilon$ and $|\hat{\lambda} - \hat{\lambda}_\varepsilon| < \varepsilon$ for $\hat{\lambda}$ and $N$ satisfying Eq. (9).
The possible noise on the measures can make the problem ill posed what is inherent to almost inverse problem. What’s happen in the problem (9) when \( N \) is replaced by \( N_\varepsilon \) and one wants to quantify the term \( \partial_\varepsilon (g(x)N_\varepsilon(x)) \). To be more clear, let’s consider the following sample.

**Example 3.1:** Let consider \( f \in \mathbb{C}^\infty(\mathbb{R}) \) and for arbitrary \( n \in \mathbb{N}^* \). We define for all \( \varepsilon > 0 \), the approximation \( f_\varepsilon(x) = f(x) + \varepsilon \sin \left( \frac{nx}{\varepsilon} \right) \). It is obvious to see that \( \| f - f_\varepsilon \|_\infty \leq \varepsilon \) means that \( f_\varepsilon \) as an excellent approximation of \( f \). But when one looks to the derivative \( f_\varepsilon'(x) = f'(x) + n \cos \left( \frac{nx}{\varepsilon} \right) \), it fulfill \( \| f' - f'_\varepsilon \|_\infty = n \) and we imagine when \( n \) is chosen as big as possible.

In order to come up this difficulty of controlling the term \( \partial_\varepsilon (gN_\varepsilon) \), we propose two methods of regularization named Filtering and Quasi-reversible methods that we present here below by introducing the new unknown \( H = BN \), solution of the approximated problem

\[
H_\varepsilon(x) - 2 \int_0^\infty \kappa(x, y)H_\varepsilon(y) \, dy = -c_\varepsilon \partial_\varepsilon \left( gN_\varepsilon(x) \right) - \lambda_\varepsilon N_\varepsilon(x).
\]

### 3.1 Filtering method

Here, we introduce the mollifier

\[
\rho_\alpha = \frac{1}{2} \rho \left( \frac{x}{\alpha} \right); \quad \rho \in C_\infty(\mathbb{R}_+); \quad \int_0^\infty \rho(x) \, dx = 1; \quad \alpha > 0,
\]

and define the following regularized problem

\[
H_{\varepsilon, \alpha}(x) - 2 \int_0^\infty \kappa(x, y)H_{\varepsilon, \alpha}(y) \, dy = \rho_\alpha \ast \left( -c_\varepsilon \partial_\varepsilon \left( gN_\varepsilon(x) \right) - \lambda_\varepsilon N_\varepsilon(x) \right).
\] (11)

For this regularized problem we prove the existence and uniqueness (see [1]) by using the Lax-Milgram theorem which is based on the bilinear form

\[
A(u, v) = \int_0^\infty u(x)v(x)x^p \, dx - 2 \int_0^\infty u(x) \int_x^\infty \kappa(x, y)u(y) \, dy \, dx = < u, v >_{L^2(x^p \, dx)} - B(u, v),
\]

with \( B(u, u) = 2 \int_0^\infty \int_x^\infty x^p u(x)u(y) \kappa(x, y) \, dy \, dx \), that we show the continuity and coercivity.

In addition we have by priori estimates and Minkowski inequality, the following approximations

\[
|c_{\varepsilon, \alpha} - c| \leq C(p, \rho, N) \varepsilon, \quad \| B_{\varepsilon, \alpha} - B \|_{L^2(x^p Na^p \, dx)} \leq C(p, \rho, N) \left( \alpha + \frac{\varepsilon}{\alpha} \right),
\]

where \( C(p, \rho, N) \) is a constant depending on \( p, \rho \) and \( N \).

### 3.2 Quasi-reversible method

The key point of this approach is to add another derivative term in the problem in order to balance the term source of ill posed term. So the regularized problem is casted as follow for \( k \in \mathbb{R}_+, k \geq 2 \) and \( \alpha > 0 \).

\[
\alpha x^{-k} \partial_\varepsilon \left( x^{k+1} H_{\varepsilon, \alpha} \right) + H_{\varepsilon, \alpha} - 2 \int_0^\infty \kappa(x, y)H_{\varepsilon, \alpha}(y) \, dy = -c_\varepsilon \partial_\varepsilon \left( gN_\varepsilon \right) - \lambda_\varepsilon N_\varepsilon.
\] (12)

Here, we prove the existence and uniqueness of the solution (see [1]) by the Banach fixed point theorem. We begin by writing the regularized problem with a general unknown \( u \) as:

\[
\alpha x \partial_\varepsilon u(x) + \left( \alpha(k + 1) + 1 \right) u = 2 \int_0^\infty \kappa(x, y)u(y) \, dy + F(x),
\]

01003-p.4
where $F(x) = -c_{e,x} \partial_x (g N_i(x)) - \dot{\lambda}_e N_i$. So we define the application in $L^1(\mathbb{R}_+, x \, dx)$ such that $\mathcal{T} : v \rightarrow u$ with $u = \mathcal{T}(v)$ is the explicit solution of

$$
\dot{x} \partial_x u(x) + \left( \chi(k + 1) + 1 \right) u = 2 \int_0^\infty \kappa(x, y) v(y) \, dy + F(x).
$$

We prove that $\mathcal{T}$ is a contracting Lipschitz function.

In addition we have by priori estimates and Minkowski inequality, the following approximations

$$
|c_{e,x} - c| \leq C(p, N) \varepsilon, \quad ||B_{e,x} - B||_{L^2(\mathbb{R}^2 \, dx)} \leq C(p, N) \left( \chi + \frac{\varepsilon}{\chi} \right),
$$

where $C(p, N)$ is a constant depending on $p$ and $N$.

### 4. Numerical simulation for the inverse problem of estimating $B$

This part is devoted to the numerical test of the estimated division rate $B$ from the two proposed regularization methods. Experimental data on the long time behavior of the cell repartition function and on the growth rate are needed. In [1], we have no data available, so we numerical generate synthetic data on $N$ by solving first the direct problem (1) with given $g, \kappa$ and $B$; what allows to obtain the asymptotic profile of the size distribution function of cells. Once this asymptotic profile is obtained, we keep the same parameters $g, \kappa$ and we forget the division rate $B$ that we want to recover. By adding noise on $N$, we numerically recover the division rate $B_e$ thanks to our two regularized methods and compare $B_e$ and $B$ which was taken into account in the resolution of the direct problem (1).

#### 4.1 Results on the direct problem

The simulation of the direct problem follows a finite volume method applied to (1). For the length domain $L$ and grid number points $n_{\text{max}}$ we have $x_i = i \Delta x$, $0 \leq i \leq n_{\text{max}}$, $\Delta x = \frac{L}{n_{\text{max}}}$. So, for a time step $\Delta t$ we define $t_k = k \Delta t$, $k \in \mathbb{N}$ and take the following approximations

$$
n_i^k = \frac{1}{\Delta x} \int_{x_{i-1/2}}^{x_{i+1/2}} n(k \Delta t, y) \, dy, \quad n_i^{k+1} = \frac{1}{\Delta t} \int_0^{\Delta t} n(k \Delta t + s, x_i) \, ds.
$$

We use time splitting and solve first the growth part of (1), $\partial_t n + \partial_x (g n) = 0$, as follows

$$
n_i^{k+1/2} = n_i^k - \frac{\Delta t}{\Delta x} ((gn)_i^{k+1/2} - (gn)_i^{k-1/2}),
$$

where $(gn)_i^{k+1/2}$ are incoming and outgoing flows. In second, we solve the gain and lost part by cell division as follows

$$
n_i^{k+1} = (1 - \Delta t B_i) n_i^{k+1/2} + 2 \Delta t \mathcal{F}_i^k \quad \text{with} \quad \mathcal{F}_i^k \approx \int_{x_i}^{x_{n_{\text{max}}}} B(y) n^{k+1/2}(y) \kappa(x_i, y) \, dy.
$$

In order to get the asymptotic profile, we renormalize the discrete solution:

$$
\bar{n}^{k+1} = \frac{n_i^{k+1}}{\sum_{j=1}^{n_{\text{max}}} n_j^{k+1} \Delta x},
$$

then we have $\bar{n}^{k+1} \rightarrow N$, $\sum_{i=1}^{n_{\text{max}}} N_i \Delta x = 1$, $N_i > 0$ for index $i$ and $N$ is the dominant eigenvector associated to the dominant eigenvalue $\lambda$ for the stationary problem (9).
For different initial conditions and parameter values on $g$, $\kappa$ and $B$ satisfying the general assumption of existence and uniqueness of solution our scheme converge to the same asymptotic behavior as one can see in figure (Fig. 1) below. Let recall that the time step is computed with respect the CFL (Courant-Friedrichs-Lewi) stability condition, $\Delta t < 1 / \max_{0 \leq i \leq n_{\text{max}}} \left( B_i + \frac{g_i}{\Delta x} \right)$, with the notation $B_i = B(i\Delta x)$ and $g_i = g(i\Delta x)$. Once results on the direct problem are obtained, means the asymptotic behavior, we numerically simulate the regularized problems in order to estimate the division rate.

4.2 Results on the regularized problems

For the quasi-reversibility method, the associated regularized problem (12) is discretized as follows

$$-\chi x_i^{-k} \left( x_i^{k+1} H_{i+1} - x_i^{k+1} H_i \right) + H_i - 2 \sum_{j=i}^{k\alpha} H_j \kappa(i, j) \Delta x = L_i, \quad i = 1, \ldots, n_{\text{max}}.$$
Figure 2. Comparison of $BN$ in the direct problem and $BN$ in the regularized problem for $\epsilon = 0$ and $\kappa(x, y) = \frac{1}{x} 1_{[x<3]}$. Top left: $g(x) = x$, $B(x) = \exp(-0.08(x-12)^2)$. Top right: $g(x) = x^{1/2}$, $B(x) = \min(1, \frac{x^2}{10})$. Down center: $g(x) = x^{1/3}$, $B$ is here a tray-function (see 13).

We then rewrite this discrete equation under matrix shape $A \times H = L$ where $A$ appears as a triangular matrix, so we deduce the solution $H$ by solving this simple iteration

$$
H_{n_{\text{max}}} = \frac{L_{n_{\text{max}}}}{A_{n_{\text{max}}, n_{\text{max}}}};
$$

$$
H_i = \frac{1}{A_{i,i}} \left( L_i - \sum_{j=i+1}^{n_{\text{max}}} A_{i,j} H_j \right), \quad \forall i = n_{\text{max}} - 1, ..., 1.
$$

The non-oscillatory of this scheme is ensured by the constraint $A_{i,i} = 1 + \alpha i - 2\kappa(i,i)\Delta x > 0$. For the filtering method, we apply the same discretization as in quasi-reversibility case to (11). The convolution term is computed by FFT (Fast Fourier Transform) followed by an IFFT (Inverse Fast Fourier Transform) what leads to the approximations

$$
\rho_x * N \approx F^* (\hat{\rho}_x(\xi) F(N)(\xi)) \quad \text{and} \quad \rho_x * \left( \frac{\hat{\xi}(\xi)}{\hat{\xi}(\xi)} \right) \approx F^* (i \xi \hat{\rho}_x(\xi) F(gN)(\xi)).
$$
We then write the discrete equation of (11) under matrix shape $A \times H = L$ with $A$ an upper triangular matrix and deduce the approximation of $H = BN$ as follows

$$
\begin{align*}
H_{n_{\text{max}}} &= L_{n_{\text{max}}} / A_{n_{\text{max}}, n_{\text{max}}}, \\
H_i &= \frac{1}{A_{i,i}} \left( L_i - \sum_{j=i+1}^{n_{\text{max}}} A_{i,j} H_j \right), \quad \forall i = n_{\text{max}} - 1, ..., 1.
\end{align*}
$$

The numerical results show a good accuracy of our regularization methods as depicted in the pictures below in the case where we assume that we have exact measures ($\varepsilon = 0$) that means $N_\varepsilon = N$. These pictures compare the exact value on $B$ taken in the direct problem and its estimated values obtained by...
solving the quasi-reversibility or filtering methods. We define a Tray-function for $B$ as follows

$$B(x) = \begin{cases} 
0 & \text{for } x < 2, \\
\left(\frac{x-2}{13}\right)^2 & \text{for } x \in [2, 15], \\
1 & \text{for } x > 15.
\end{cases} \quad (13)$$

In order to take into account the uncertainty in measurement, we add some multiplicative noise which is uniformly distributed in $N_\varepsilon = \max(N(1 + l\varepsilon), 0)$ with $l \in \left[\frac{-1}{2}, \frac{1}{2}\right]$ and $\varepsilon \in [0, 1]$.

So, with $N_\varepsilon$ considered as the measured data, we obtain the following results with various choice of $\varepsilon$ and the fixed parameters $g(x) = x$, $\kappa(x, y) = \frac{1}{2}1_{x < y}$, $B(x) = \exp(-0.08(x - 13)^2)$.

In the simulations we investigate for any $\varepsilon$, the optimal regularization parameter $\lambda$ which gives the best estimation $B_{\varepsilon, \lambda}$ of the division rate $B$ by minimizing the $L^2$-norm of the following relative error

$$err = \frac{\|B_{\varepsilon, \lambda} - B\|_{L^2}}{\|BN\|_{L^2}}.$$

### 5. Conclusion

This study presents how to recover the division rate for a general growth-fragmentation model for a size structured population where measurements on the time asymptotic profile of the density. Two method of regularization are highlighting here in order to overcome the difficulty of the ill-posed inverse problem. Theoretical estimates are built with strong analysis on the well-posedness of the regularized problem and numerical methods are implemented. As show on the above numerical illustrations, the estimation of the division rate is quite good. However, some issues remain to be addressed, specially for small size $x$, our algorithm need to be performed such as the appearance of oscillations.

This proceeding is a part of a work paper [1] done in collaboration with Marie Doumic, head of the MAMBA team of INRIA. The work was partially financed by the Agence Nationale de la Recherche, Grant No. ANR-09-BLAN-0218 TOPPAZ. I thank Abdelkader Lakmeche of Sidi Bel-Abbes university (Algeria) who gave me the opportunity to present this work during the Workshop WMLS’14 SBA.

### References

