

Numerical Solution of an Inverse Problem in Size-Structured Population Dynamics

Marie Doumic-Jauffret, Benoît Perthame, Jorge Zubelli

▶ **To cite this version:**

Marie Doumic-Jauffret, Benoît Perthame, Jorge Zubelli. Numerical Solution of an Inverse Problem in Size-Structured Population Dynamics. Inverse Problems, 2009, 25 (4), pp.045008. 10.1088/0266-5611/25/4/045008 . hal-00327151

HAL Id: hal-00327151
https://hal.science/hal-00327151

Submitted on 7 Oct 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Numerical Solution of an Inverse Problem in Size-Structured Population Dynamics

Marie Doumic [*‡] Benoît Perthame[†‡] Jorge Zubelli[§]

October 8, 2008

Abstract

We consider a size-structured model for cell division and address the question of determining the division (birth) rate from the measured stable size distribution of the population. We propose a new regularization technique based on a filtering approach. We prove convergence of the algorithm and validate the theoretical results by implementing numerical simulations, based on classical techniques. We compare the results for direct and inverse problems, for the filtering method and for the quasi-reversibility method proposed in [1].

1 Introduction

The use of size-structured models to describe biological systems has attracted the interest of many authors and has a long standing tradition. In particular, the use of size structures was very well documented and compared to experiments in the 70's. This led to the survey book [2] and subsequent mathematical analysis (see also the references in [3]). Needless to say, in such models it is crucial for the analysis, computer simulation and prediction to calibrate the corresponding model parameters so as to obtain good quantitative results. Indeed, in the inverse problem literature, a number of authors have addressed the calibration of certain structured population models. See for example [4, 5, 6, 7] and references therein.

In this article, we consider theoretical and numerical aspects of the inverse problem of determining the division rate coefficient $B = B(x)$ in the following specific size-structured model for cell division:

$$\begin{cases} \frac{\partial}{\partial t} n(t,x) + \frac{\partial}{\partial x} n(t,x) + B(x) n(t,x) = 4B(2x) n(t,2x), & x \geqslant 0,\ t \geqslant 0, \\ n(t, x=0) = 0, t > 0, \\ n(0,x) = n^0(x) \geq 0. \end{cases} \tag{1}$$

[*] Département de Mathématiques et Applications, École Normale Supérieure, INRIA projet BANG, 45 rue d'Ulm, F 75230 Paris cedex 05, France; email: doumic@dma.ens.fr

[†] Université Pierre et Marie Curie-Paris 6, UMR 7598 LJLL, BC187, 4, place Jussieu, F-75252 Paris cedex 5, and Institut Universitaire de France; email: perthame@ann.jussieu.fr

[‡]INRIA Rocquencourt, projet BANG, Domaine de Voluceau, BP 105, 781153 Rocquencourt, France; emails: marie.doumic@inria.fr, benoit.perthame@inria.fr

[§]IMPA, Est. D. Castorina 110, Rio de Janeiro, RJ 22460-320 Brazil; email: zubelli@impa.br

Here, the cell density is represented by $n(t,x)$ at time t and size x. The division rate B expresses the division of cells of size $2x$ into two cells of size x.

By making use of flux cytometry technologies for instance, it is possible to determine cell populations with certain properties as protein content on a large scale of tenths of thousands of cells. In other applications, like coagulation fragmentation equation [8, 9, 10, 11, 12], or prion aggregation and fragmentation [13, 14, 15], similar equations arise, and much less is known on aggregate size repartition. The division rate $B(x)$, on the contrary, is not directly measurable.

The long time behavior of solutions is well known. Indeed, it was proved in [16, 17] that under fairly general conditions on the coefficients, there is a unique solution (N, λ_0) to the following eigenvalue problem

$$\begin{cases} \frac{\partial}{\partial x}N + (\lambda_0 + B(x))N = 4B(2x)N(2x), \qquad x \geqslant 0, \\ \\ N(x=0) = 0, \\ \\ N(x) > 0 \ \text{ for } x > 0, \qquad \int_0^\infty N(x)dx = 1, \end{cases} \tag{2}$$

where $\lambda_0 > 0$ and $Ne^{\mu x} \in L^\infty \cap L^1$ for all $\mu < \lambda_0$.

It was shown in [18, 16]

$$n(t,x)e^{-\lambda_0 t} \xrightarrow[t\to\infty]{} m_0 N(x), \quad \text{ in } \ L^1(\mathbb{R}_+, \phi(x)dx),$$

where the weight ϕ is the unique solution to the adjoint problem

$$\begin{cases} -\frac{\partial}{\partial x}\phi + (\lambda_0 + B(x))\phi = 2B(x)\phi(\frac{x}{2}), \qquad x \geqslant 0, \\ \\ \phi(x) > 0, \qquad \int_0^\infty \phi(x)N(x)dx = 1. \end{cases} \tag{3}$$

In other words, λ_0 is the growth rate of such a system and is usually called "Malthus parameter" in population biology. From [18, 16, 3] we also know that λ_0 is related to N by the relation

$$\lambda_0 = \frac{\int_0^\infty N dx}{\int_0^\infty xN dx}. \tag{4}$$

The question we address here is the following: How can we estimate the division rate B from the knowledge of the steady dynamics N and λ_0 ? The inverse problem thus consists of finding B a solution to

$$4B(2x)N(2x) - B(x)N(x) = L(x) := \frac{\partial}{\partial x}N(x) + \lambda_0 N(x), \quad x \geqslant 0, \tag{5}$$

assuming that (N, λ_0) is known, or thanks to (4) that N is known. As seen in [1], this problem is well-posed if N satisfies strong regularity properties such as $\frac{\partial}{\partial x}N(x) \in L^p(\mathbb{R}_+)$ for some $p \geqslant 1$.

However, in practical applications we have only an *approximate* knowledge of (N, λ_0), given by noisy data $(N_\varepsilon, \lambda_\varepsilon)$, with $N_\varepsilon \in L^2_+(\mathbb{R}_+)$ for instance. [1] This means that we have no way of controlling $\frac{\partial}{\partial x}N_\varepsilon$, so we cannot control the precision of a solution B_ε to problem (5) when a perturbed N_ε replaces N. Furthermore, it is not even clear whether such a B_ε exists.

[1] Actually, our knowledge of λ_0 is presumably an order of precision higher than that of N, since the rate λ_0 can be estimated independently by means of time information.

The question we focus on is then: How to approximate the problem (5) in order to get a solution B_ε as close as possible to the exact division rate B?

We remark that, in the context of noisy data, the inverse problem under consideration is ill-posed [1] and thus regularization would be required. A natural tool to be invoked from the inverse problem literature would be some kind of Tikhonov regularization method [19, 20]. However, this would lead to computationally intensive problems. Indeed, for each forward problem evaluation a dilation-differential equation of the form (2) would have to be solved.

In [1], two of the present authors proposed a method of regularization consisting in the solution of the following approximate problem:

$$\begin{cases} \alpha\frac{\partial}{\partial y}(B_{\varepsilon,\alpha}N_\varepsilon) + 4B_{\varepsilon,\alpha}(y)N_\varepsilon(y) = B_{\varepsilon,\alpha}\left(\frac{y}{2}\right)N_\varepsilon\left(\frac{y}{2}\right) + \lambda_0 N_\varepsilon\left(\frac{y}{2}\right) + 2\frac{\partial}{\partial y}\left(N_\varepsilon\left(\frac{y}{2}\right)\right), & y > 0, \\ (B_{\varepsilon,\alpha}N_\varepsilon)(0) = 0, \end{cases}$$

where α is a regularizing parameter. It was shown that a convergence rate of order $\sqrt{\varepsilon}$ could be obtained, for $\alpha = O(\sqrt{\varepsilon})$, where ε is the error on the data N in an appropriate norm.

The above method of the solution to the inverse problem will be called *quasi reversibility* in accordance with the general spirit of the terminology of [21, 22]. The main goal of this work is to investigate the numerics of such approach, to consider an alternative technique based on filtering ideas and to compare the performance of the different methods. The alternative technique is also analyzed from the theoretical point of view and estimates are presented.

In this work, we have modified slightly the original regularization equation by writing $\lambda_{\varepsilon,\alpha}$ instead of λ_0 for the reasons we shall explain in the sequel. Thus, we work with

$$\begin{cases} \alpha\frac{\partial}{\partial y}(B_{\varepsilon,\alpha}N_\varepsilon) + 4B_{\varepsilon,\alpha}(y)N_\varepsilon(y) = B_{\varepsilon,\alpha}\left(\frac{y}{2}\right)N_\varepsilon\left(\frac{y}{2}\right) + \lambda_{\varepsilon,\alpha} N_\varepsilon\left(\frac{y}{2}\right) + 2\frac{\partial}{\partial y}\left(N_\varepsilon\left(\frac{y}{2}\right)\right), & y > 0, \\ (B_{\varepsilon,\alpha}N_\varepsilon)(0) = 0. \end{cases} \tag{6}$$

Indeed, in order to conserve regularity properties of the solution $H = BN$ to the inverse problem, we want it to be both in $L^1(\mathbb{R}_+)$ and in $L^1(\mathbb{R}_+, xdx)$ in order to express that both the total number of cells and the total biomass are finite. Hence, formal integration of Equation (6) gives

$$\lambda_{\varepsilon,\alpha}\int_0^\infty N_\varepsilon dx = \int_0^\infty B_{\varepsilon,\alpha}N_\varepsilon dx, \tag{7}$$

and integration against the weight x gives

$$-\alpha\int_0^\infty B_{\varepsilon,\alpha}N_\varepsilon dx = 4\lambda_{\varepsilon,\alpha}\int_0^\infty xN_\varepsilon dx - 4\int_0^\infty N_\varepsilon dx. \tag{8}$$

Hence, we have to choose, according to the eigenvalue theory:

$$\lambda_{\varepsilon,\alpha} = \frac{\int_0^\infty N_\varepsilon dx}{\int_0^\infty xN_\varepsilon dx + \frac{\alpha}{4}\int_0^\infty N_\varepsilon dx}. \tag{9}$$

The choice of $\lambda_{\varepsilon,\alpha}$ can be understood as a compatibility condition when $\alpha > 0$ and for $\alpha = 0$ it tells us that (N, λ_0) is overdetermined data for the inverse problem. Therefore, if we have *a priori* knowledge on λ_0, we could verify its distance to $\lambda_{\epsilon,\alpha}$ as a way of checking the error of the inverse problem solution.

The plan of this work is the following: In Section 2, we propose yet another method to regularize the inverse problem, and obtain a convergence rate. The convergence rate turns out to be as good as the one in [1]. In Section 3 we give a numerical method to solve it, and in Section 4 we show some numerical simulations so as to compare the accuracy of the different methods.

2 Regularization by Filtering

2.1 Filtering approach

Taking a closer look at Equation (5), we see that all the difficulties come from the differential term $\frac{\partial}{\partial x}N$. In [1], the choice was to add an equivalent derivative $\alpha\frac{\partial}{\partial x}(B\,N)$ to the equation; here on the contrary, we choose to regularize it by a convolution method.

For $\alpha > 0$, we use the notation

$$\rho_\alpha(x) = \frac{1}{\alpha}\rho(\frac{x}{\alpha}), \qquad \rho \in \mathcal{C}_c^\infty(\mathbb{R}), \quad \int_0^\infty \rho(x)\,dx = 1, \quad \rho \geqslant 0, \quad \mathrm{Supp}(\rho) \subset [0,1], \tag{10}$$

and we replace in (5) the term $\frac{\partial}{\partial x}N_\varepsilon + \lambda_0 N_\varepsilon$ by

$$\big(\frac{\partial}{\partial x}N_\varepsilon + \lambda_{\varepsilon,\alpha}N_\varepsilon\big) * \rho_\alpha(x) = N_\varepsilon * \big(\frac{\partial}{\partial x}\rho_\alpha + \lambda_{\varepsilon,\alpha}\rho_\alpha\big)(x) = \int_0^\infty N_\varepsilon(x')\big(\frac{\partial}{\partial x}\rho_\alpha + \lambda_{\varepsilon,\alpha}\rho_\alpha\big)(x-x')dx'.$$

We now use the notation

$$N_{\varepsilon,\alpha} = N_\varepsilon * \rho_\alpha.$$

In this way, we obtain a smooth term in $L^2(\mathbb{R}_+)$. Furthermore, $N_{\varepsilon,\alpha}$ converges to N_ε in $L^2(\mathbb{R}_+)$ when α tends to zero. We now have to consider the following problem:
Find $B_{\varepsilon,\alpha}$ solution of

$$4B_{\varepsilon,\alpha}(2x)N_{\varepsilon,\alpha}(2x) + B_{\varepsilon,\alpha}(x)N_{\varepsilon,\alpha}(x) = \frac{\partial}{\partial x}N_{\varepsilon,\alpha} + \lambda_{\varepsilon,\alpha}N_{\varepsilon,\alpha}(x), \quad x \geqslant 0\ . \tag{11}$$

As in Equation (6), for the quasi-reversibility method, we need to choose $\lambda_{\varepsilon,\alpha}$ appropriately. Indeed, we perform the same manipulations leading to Equation (9) to get

$$\lambda_{\varepsilon,\alpha} = \frac{\int_0^\infty N_{\varepsilon,\alpha}(x)dx}{\int_0^\infty xN_{\varepsilon,\alpha}(x)dx}. \tag{12}$$

By Theorem A.3 (see the Appendix), we know that the problem in Equation (11) has a unique solution $B_{\varepsilon,\alpha} \in L^2(\mathbb{R}_+, N_{\varepsilon,\alpha}^2 dx)$.

2.2 Estimates for the filtering approach

The main result of this section establishes an estimate for the regularization of the inverse problem by means of the filtering method described above.

Theorem 2.1 *Suppose that $N \in H^2(\mathbb{R}_+)$ and $B \in L^\infty(\mathbb{R}_+)$, $B \geqslant 0$ verify (2). Let $\varepsilon > 0$ and $N_\varepsilon \in L^2(\mathbb{R}_+)$, $N_\varepsilon(x) > 0$ for $x > 0$, such that*

$$||N_\varepsilon - N||_{L^2(\mathbb{R}_+)} \leqslant \varepsilon||N||_{L^2(\mathbb{R}+)}.$$

Let $B_{\varepsilon,\alpha} \in L^2(\mathbb{R}_+, N^2_{\varepsilon,\alpha}dx)$ be the unique solution of (10) and (11). We have the following estimate:

$$||B_{\varepsilon,\alpha} - B||_{L^2(N^2_\varepsilon dx)} \leqslant C(\alpha + |\lambda_{\varepsilon,\alpha} - \lambda_0|)||N||_{H^2(\mathbb{R}_+)} + \frac{C}{\alpha}||N_{\varepsilon,\alpha} - N||_{L^2(\mathbb{R}_+)}, \tag{13}$$

where C is a constant depending only on $||B||_{L^\infty}$, $||B_{\varepsilon,\alpha}||_{L^\infty}$ and the regularizing function ρ.

This theorem relies on a first estimate.

Proposition 2.2 *Using the same notations as in Theorem 2.1, we have*

$$||B_{\varepsilon,\alpha}N_{\varepsilon,\alpha} - B\,N||^2_{L^2(dx)} \leqslant C\left(1+\lambda_0^2\right)\left(1+\frac{1}{\alpha^2}\right)||N_\varepsilon - N||^2_{L^2(dx)} + C(\alpha^2 + |\lambda_{\varepsilon,\alpha} - \lambda_0|^2)||N||^2_{H^2(\mathbb{R}_+)}\ , \tag{14}$$

where C depends only on the regularizing function ρ.

Proof of Prop. 2.2: Denote by $Q = B_{\varepsilon,\alpha}N_{\varepsilon,\alpha} - BN$, $R = N_{\varepsilon,\alpha} - N$ and $\delta = \lambda_{\varepsilon,\alpha} - \lambda_0$. From Equations (2) and (11), Q verifies:

$$\begin{cases} \frac{\partial}{\partial x}R(x) + \lambda_0 R(x) + \delta N_{\varepsilon,\alpha}(x) + Q(x) = 4Q(2x), \qquad x \geqslant 0, \\ \\ Q(x=0) = 0. \end{cases} \tag{15}$$

(Since $N_{\varepsilon,\alpha} \in H^1(\mathbb{R}_+)$, the definition of $Q(x=0)$ is not ambiguous.) Multiplying this equation by $Q(2x)$ and integrating on the interval $(0, y)$ yields

$$4\int_0^y Q(2x)^2dx = \int_0^y Q(2x)\frac{\partial}{\partial x}R(x)dx + \lambda_0\int_0^y Q(2x)R(x)dx$$

$$+\delta\int_0^y Q(2x)N_{\varepsilon,\alpha}(x)dx + \int_0^y Q(2x)Q(x)dx.$$

From the Cauchy-Schwarz inequality, after the change of variables $x \to 2x$, we have

$$4\int_0^y Q(2x)^2dx \leqslant \frac{1}{2}\int_0^y\left(\frac{\partial}{\partial x}R\right)^2(x)dx + \frac{1}{2}\int_0^y Q(2x)^2dx + \frac{\lambda_0}{2}\int_0^y CR(x)^2dx + \frac{\lambda_0}{2}\int_0^y\frac{Q(2x)^2}{C}dx$$

$$+\frac{|\delta|^2}{2}\int_0^y N_{\varepsilon,\alpha}(x)^2dx + \frac{1}{2}\int_0^y Q(2x)^2dx + \frac{1}{2}\int_0^y Q(2x)^2dx + \int_0^{\frac{y}{2}} Q(2x)^2dx.$$

We take, for instance, $C = \lambda_0$. We obtain

$$||B_{\varepsilon,\alpha}N_{\varepsilon,\alpha} - BN||^2_{L^2} \leqslant ||N_\varepsilon * \frac{\partial}{\partial x}\rho_\alpha - \frac{\partial}{\partial x}N||^2_{L^2} + \lambda_0^2||N_\varepsilon * \rho_\alpha - N||^2_{L^2} + |\lambda_{\varepsilon,\alpha} - \lambda_0|^2||N_\varepsilon * \rho_\alpha||^2_{L^2}. \tag{16}$$

The last two terms of this inequality are easy to estimate, writing

$$||N_\varepsilon * \rho_\alpha - N||_{L^2} \leqslant ||N_\varepsilon * \rho_\alpha - N * \rho_\alpha||_{L^2} + ||N * \rho_\alpha - N||_{L^2} \leqslant C\left(||N_\varepsilon - N||_{L^2} + \alpha||N||_{H^1}\right),$$

and

$$||N_\varepsilon * \rho_\alpha||_{L^2} \leqslant C||N||_{L^2}.$$

It remains to evaluate the first term on the right-hand side of inequality (16). We write

$$||N_\varepsilon * \frac{\partial}{\partial x}\rho_\alpha - \frac{\partial}{\partial x}N||^2_{L^2} \leqslant 2||N_\varepsilon * \frac{\partial}{\partial x}\rho_\alpha - N * \frac{\partial}{\partial x}\rho_\alpha||^2_{L^2} + 2||N * \frac{\partial}{\partial x}\rho_\alpha - \frac{\partial}{\partial x}N||^2_{L^2}.$$

By a convolution estimate we evaluate the first term as

$$||N_\varepsilon * \frac{\partial}{\partial x}\rho_\alpha - N * \frac{\partial}{\partial x}\rho_\alpha||^2_{L^2(\mathbb{R}_+,dx)} \leqslant ||N_\varepsilon - N||^2_{L^2(\mathbb{R}_+,dx)}||\frac{\partial}{\partial x}\rho_\alpha||^2_{L^1} \ .$$

Since $\int_0^\infty |\frac{\partial}{\partial x}\rho_\alpha(x)|dx = \frac{1}{\alpha}\int_0^\infty |\frac{\partial}{\partial x}\rho(y)|dy$, we have

$$||N_\varepsilon * \frac{\partial}{\partial x}\rho_\alpha - \frac{\partial}{\partial x}N||^2_{L^2(\mathbb{R}_+,dx)} \leqslant \frac{C(\rho)}{\alpha^2}||N_\varepsilon - N||^2_{L^2(\mathbb{R}_+,dx)} + 2||N * \frac{\partial}{\partial x}\rho_\alpha - \frac{\partial}{\partial x}N||^2_{L^2(\mathbb{R}_+,dx)} \ .$$

To evaluate the last term $||N * \frac{\partial}{\partial x}\rho_\alpha - \frac{\partial}{\partial x}N||^2_{L^2}$, we extend to $\mathbb{R}$ the functions N and $\frac{\partial}{\partial x}N$ by zero and consider their Fourier transforms. We denote $\hat{f}(\xi)$ the Fourier transform of $f \in L^2(\mathbb{R}_+)$ at ξ, where f is extended as zero on $\mathbb{R}_-$. We obtain by Fourier analysis

$$||N * \frac{\partial}{\partial x}\rho_\alpha - \frac{\partial}{\partial x}N||^2_{L^2(\mathbb{R}_+,dx)} = ||i\xi\hat{N}\hat{\rho}_\alpha - i\xi\hat{N}||^2_{L^2(\mathbb{R}_+,dx)} \leqslant \int_{-\infty}^{\infty} |\hat{N}(\xi)|^2|\xi|^4\frac{|\hat{\rho}_\alpha(\xi)-1|^2}{|\xi|^2}d\xi.$$

Using that

$$|\frac{|\hat{\rho}_\alpha(\xi)-1|^2}{\xi^2}| \leqslant C(\rho)\alpha^2, \tag{17}$$

where $C(\rho)$ only depends on the regularization function ρ, we have that

$$||N * \frac{\partial}{\partial x}\rho_\alpha - \frac{\partial}{\partial x}N||^2_{L^2(\mathbb{R}_+,dx)} \leqslant C(\rho)\alpha^2||N||^2_{H^2(\mathbb{R}_+)} \ .$$

Going back to (16), this concludes the proof of Proposition 2.2. □

We can now deduce the proof of Theorem 2.1. We write:

$$||B_{\varepsilon,\alpha} - B||_{L^2(N_\varepsilon^2 dx)} \leqslant ||B_{\varepsilon,\alpha}N_\varepsilon - B_{\varepsilon,\alpha}\,N_{\varepsilon,\alpha}||_{L^2(\mathbb{R}_+)} + ||B_{\varepsilon,\alpha}\,N_{\varepsilon,\alpha} - B\,N||_{L^2(\mathbb{R}_+)} + ||B\,N - B\,N_\varepsilon||_{L^2(\mathbb{R}_+)}.$$

Using Proposition 2.2, and the fact that

$$||N_\varepsilon - N_{\varepsilon,\alpha}||_{L^2} \leqslant 2||N_\varepsilon - N||_{L^2} + \alpha||N||_{H^1},$$

this inequality gives the result. □

3 Numerical Solution of the Inverse Problem

This section is concerned with the numerical aspects of the solution of the inverse problem. In order to do that we start with a description of the solution to the direct one in Subsection 3.1.

3.1 Direct Problem

In the direct problem, we assume we know the proliferation rate B, we look for N and $\lambda_0 > 0$ solutions of (2). For this purpose, we solve the time-dependent problem (1) and look for a steady dynamics. As already said, this problem is well-posed (see for instance [3]) and it was proved in [18] that solutions grow at an exponential rate towards $\rho N(x)e^{\lambda_0 t}$ with $\rho = \int_0^\infty n(0,x)\phi(x)dx$, recalling the notation in (3). Furthermore, under more restrictive conditions it was shown in [16] that there exists constants $\mu > 0$ and $C(n^0) > 0$, such that

$$||n(t,x)e^{-\lambda_0 t} - \rho N(x)||_{L^1(\mathbb{R}_+,\phi(x)dx)} \leqslant Ce^{-\mu t}.$$

To solve it numerically, we discretize the problem (1) along a regular grid, denote by Δt the time step and by $\Delta x = L/I$ the spatial step, where I denotes the number of points and L the computational domain length: $x_i = i\Delta x$, $0 \leqslant i \leqslant I$.

We use an upwind finite volume method (cf. [23, 24, 25])

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

For the time discretization, we use a marching technique. We choose the time step Δt so as to satisfy the largest possible CFL stability criteria $\theta := \frac{\Delta t}{\Delta x} = 1$.

The numerical scheme is given, for $i = 1, ..., I$, by $n_0^k = 0$ and

$$\frac{n_i^{k+1} - n_i^k}{\Delta t} + \frac{n_i^k - n_{i-1}^k}{\Delta x} + B_i n_i^{k+1} = B_{2i-1}n_{2i-1}^k + 2B_{2i}n_{2i}^k + B_{2i+1}n_{2i+1}^k \ , \tag{18}$$

with the convention that $n_j = 0$ for $j > I$. For stability reasons, we have used an implicit method for the division term in the left hand side and explicit for the right hand side of the equation. The specific form for the right hand side is simply motivated by the need of also dividing cells of odd labels.

According to the power algorithm, we do not keep n^{k+1} from (18) but rather renormalize it as

$$\widetilde{n}^{k+1} = \frac{n^k}{\Delta x \sum_{j=1}^{I} n_j^k}.$$

It is standard, for these positive matrices arising in (18), that

$$\widetilde{n}^{k+1} \underset{k\to\infty}{\longrightarrow} N, \qquad \sum_{i=1}^{I} N_i = 1, \qquad N_i > 0,$$

where N is the dominant eigenvector for the problem

$$\frac{N_i - N_{i-1}}{\Delta x} + (\lambda_0 + B_i)N_i = B_{2i-1}N_{2i-1} + 2B_{2i}N_{2i} + B_{2i+1}N_{2i+1}.$$

One can also find the dominant eigenvalue as

$$\lambda_0 = \lim_{k\to\infty} \frac{1}{\Delta t}\log\left(\frac{\sum_{i=1}^{I} n_i^{k+1}}{\sum_{i=1}^{I} n_i^k}\right).$$

For matrices with one dominant eigenvalue and a corresponding one-dimensional eigenspace, it is known that the power algorithm is fast and in fact converges with exponential rate [26]. In practice we can stop the iterations when the relative error on the normalized quantity

$$\frac{1}{\Delta t}\left(\sum_{i=1}^{I}\widetilde{n}_i^{k+1}-\sum_{i=1}^{I}\widetilde{n}_i^{k}\right)$$

is small enough, say of the order of 10^{-10}.

3.2 Inverse Problem: General Strategy

In the sequel, we denote by H the product $B.N$ and its approximations. Indeed, from Equations (6) or (11), we have to search for the product $H = B_{\varepsilon,\alpha}N_\varepsilon$ or $H = B_{\varepsilon,\alpha}N_{\varepsilon,\alpha}$ before computing $B_{\varepsilon,\alpha}$. In particular, we cannot avoid a loss of information where N_ε is small, i.e., for $x \approx 0$ or $x \gg 1$.

The inverse problem (5), as well as (11), can be written as

$$4H(2x) - H(x) = L(x), \tag{19}$$

with different expressions for H and L. We may think of two possible numerical approaches.

Strategy 1. Compute $H(2x)$ from $H(x)$: This means that we re-write Equation (19) with the new variable $y = 2x$, and arrive at

$$4H(y) - H(\frac{y}{2}) = L(\frac{y}{2}). \tag{20}$$

The scheme departs from zero, and one deduces the values of H_i step by step, from the knowledge of H_j for $j \leqslant i-1$.

Strategy 2. Compute $H(x)$ from $H(2x)$: The scheme departs from the largest point $x = L$ of our simulation domain. We suppose that for $x \geqslant L$ we have $H(x) = H(L) = 0$ (it is relevant since we suppose that N vanishes for x large: see below), and then deduce the smaller values H_i step by step, from the knowledge of H_j for $j \geqslant i+1$.

The two approaches do not necessarily lead to the same result because the continuous equation

$$4H(2x) - H(x) = 0 \tag{21}$$

has infinitely many solutions. This issue is interesting on its own and is related to the construction of wavelets, see [27]. It is discussed in Proposition A.1 of the Appendix.

By imposing $H \in L^2(\mathbb{R}_+)$, we select a unique solution, as shown in Theorem A.3. The question is then: Which numerical strategy should we use to select the *correct* solution, *i.e.* the one in $L^2(\mathbb{R}_+)$?

Among the solutions of Equation (19), we single out two, defined by the power series:

$$H^{(1)}(x) = \sum_{n=1}^{+\infty} 2^{-2n}L(2^{-n}x) \quad \text{and } H^{(2)}(x) = -\sum_{n=0}^{+\infty} 2^{2n}L(2^n x) \quad , \quad \forall\, x > 0.$$

Proposition A.1 shows that for $L \in L^2(\mathbb{R}_+, x^p dx)$, there is a unique solution in $L^2(\mathbb{R}_+, x^p dx)$, given by $H^{(1)}$ if $p < 3$ and by $H^{(2)}$ if $p > 3$ (and the power series converge in the corresponding spaces).

For $B > 0$ smooth and bounded from above and from below, we know that N is smooth and vanishes at $x \approx 0$ and $x \approx \infty$, and BN inherits these properties. For instance, we know that

$H \in L^2(dx) \cap L^2(x^4 dx)$. By uniqueness of a solution in each space, Proposition A.1 implies that $H^{(2)} = H^{(1)}$, or equivalently:

$$\sum_{n=-\infty}^{+\infty} 2^{2n} L(2^n x) = 0, \quad \forall\, x \geqslant 0.$$

This very particular property cannot be verified at the discrete level. Hence, the two strategies generally give two different approximations of the same solution of (19). The first strategy selects an approximation of the solution $H^{(1)}$ whereas the second selects an approximation of the solution $H^{(2)}$. In the case of a very regular data N, then $H^{(2)}$ will perform better around infinity, whereas $H^{(1)}$ will be better around zero. However, if N is a solution of Equation (2), when we increase the number of points, the two approaches converge to the same solution since $H^{(2)} = H^{(1)}$.

Since our simulation domain $[0, L]$ is bounded and contains zero, we prefer the first strategy. This choice is confirmed by all the numerical tests we have performed: the second approach has always lead to a solution exploding around zero. However, for the sake of completeness, we also describe the scheme we used for the second approach.

3.3 Inverse Problem: Filtering Approach

According to strategies 1 and 2, we now present two approaches to handle the numerical solution of the inverse problem regularized with the *filtering approach*. Both need to first compute the convolution terms arising in (11). To do so we first take the Fast Fourier Transform F of N_ε, multiply it by $i\xi\hat{\rho}_\alpha(\xi)$, and then take the inverse Fast Fourier Transform F^*. We choose and define the regularization function ρ_α by its Fourier transform:

$$\hat{\rho}_\alpha(\xi) = \frac{1}{\sqrt{1 + \alpha^2 \xi^2}}.$$

This leads us to the numerical approximation

$$\frac{\partial}{\partial x} N_{\varepsilon,\alpha} \approx dN_\alpha = \mathsf{F}^*\Big(i\xi \hat{\rho}_\alpha(\xi) \mathsf{F}(N_\varepsilon)(\xi) \Big). \tag{22}$$

We also impose $dN_{\alpha,0} = 0$ for compatibility with the continuous equation and further use.

As mentioned earlier, there are two alternatives, either starting from zero or coming from infinity.

The Filtering Approach Starting from Zero (strategy 1). We solve Equation (11) considered as an equation in the variable $y = 2x$, that is to say (20), in order to compute its solution $H^{(1)}(x)$. At the discrete level, we use the notations

$$H_i^f \approx B_i N_i, \qquad L_i^f = dN_{\alpha,i} + \lambda_{\varepsilon,\alpha} N_i, \quad L_0^f = 0.$$

The discrete version of (20) reads

$$4H_i^f = H_{\frac{i}{2}}^f + L_{\frac{i}{2}}^f, \qquad \forall\, 0 \leqslant i \leqslant I, \tag{23}$$

and we need to define the quantities $G_{\frac{i}{2}}$. We choose

$$G_{\frac{i}{2}} = \begin{cases} G_{\frac{i}{2}} & \text{when } i \text{ is even,} \\ \frac{1}{2}\big(G_{\frac{i-1}{2}} + G_{\frac{i+1}{2}}\big) & \text{when } i \text{ is odd.} \end{cases} \tag{24}$$

In particular, we have $H_0^f = 0$.

Summing up all the terms in (23) for $1 \leqslant i \leqslant I$, we find (with I even to simplify):

$$4\sum_{i=0}^{I} H_i^f = 2\sum_{i=0}^{\frac{I}{2}}(H_i^f + L_i^f) - \frac{1}{2}(H_{I/2}^f + L_{I/2}^f).$$

Since we have assumed that N has exponential decay for $x \gg 1$, it follows that

$$\sum_{i=0}^{I} H_i^f = \sum_{i=1}^{I} L_i^f + E_I, \qquad \text{with} \qquad |E_I| \leqslant 2\sum_{i=\frac{I}{2}}^{I} |H_i|. \tag{25}$$

Multiplying (23) by x_i and summing up again, we find

$$\sum_{i=1}^{\frac{I}{2}} x_i L_i^f = F_I, \qquad \text{with} \qquad |F_I| \leqslant \sum_{i=\frac{I}{2}}^{I} x_i |H_i|. \tag{26}$$

As a consequence, we can choose:

$$\lambda_{\varepsilon,\alpha} = -\frac{\sum x_i dN_{\alpha,i}}{\sum x_i N_i}, \tag{27}$$

as the discrete version of the relations (4) or (12).

The Filtering Approach Starting from Infinity (strategy 2). Another method is to discretize the formulation (19) in order to compute its solution $H^{(2)}(x)$. We define the extension $H_i^f = 0$ for $i \geqslant I+1$, and for $2 \leq i \leq I$, we define by backward iterations

$$H_i^{f\infty} = 2H_{2i}^{f\infty} + H_{2i+1}^{f\infty} + H_{2i-1}^{f\infty} - L_i^f. \tag{28}$$

This however does not apply to the indices $i = 0,\ 1$ and we set $H_0^f = \frac{L_0^f}{3} = 0$ and $H_1^f = 4H_2^f - L_1^f$. By summing up all the terms in (28), we find balance properties equivalent to (25)–(26), but with remainders E_I and F_I depending on H_1 and H_2 instead of $H_{i\geqslant\frac{I}{2}}$. One has to check *a posteriori* that these last quantities are very small ; it is not the case in a standard calculation, but becomes true when the precision of the direct problem scheme increases.

3.4 Inverse problem: Quasi-Reversibility Approach

In this section, we present a numerical scheme for the regularized inverse problem proposed in [1]. This problem leads to solving (6) taken at $y = 2x$, that is

$$\begin{cases} \alpha\frac{\partial}{\partial y}(B_{\varepsilon,\alpha}N_\varepsilon) + 4B_{\varepsilon,\alpha}(y)N_\varepsilon(y) = B_{\varepsilon,\alpha}\left(\frac{y}{2}\right)N_\varepsilon\left(\frac{y}{2}\right) + \lambda_{\varepsilon,\alpha}N_\varepsilon\left(\frac{y}{2}\right) + 2\frac{\partial}{\partial y}\left(N_\varepsilon\left(\frac{y}{2}\right)\right), \qquad y > 0, \\ \\ (B_{\varepsilon,\alpha}N_\varepsilon)(0) = 0, \end{cases}$$

where $\alpha > 0$ is the regularizing parameter and $\lambda_{\varepsilon,\alpha}$ is defined by (9). This gives, in a discretized version, after dropping the index ε,

$$\lambda_{\varepsilon,\alpha} = \frac{\sum N_i}{\sum x_i N_i + \frac{\alpha}{4}\sum N_i}. \tag{29}$$

For the numerical discretization we set $H^Q_{-1} = 0$ and also recall that $N_0 = 0$ and assume that the data satisfies $N_{I+1} = 0$. We use a standard upwind scheme for the differential term:

$$\frac{\alpha}{\Delta x}(H^Q_i - H^Q_{i-1}) + 4H^Q_i = H^Q_{\frac{i}{2}} + L^Q_{\frac{i}{2}}, \tag{30}$$

where we have defined the fractional indices as in the filtering approach by (24), and here

$$L^Q_i = \lambda_\varepsilon N_i + \frac{N_{i+1} - N_i}{\Delta x}.$$

If we neglect the terms $H^Q_{i \geqslant \frac{I}{2}+1}$, we can easily verify a discrete version of the balance laws (7) and (9), equivalent to (25)–(26).

4 Numerical Tests

As input data, we take the values of the function N obtained by the numerical solution of the direct problem in Section 3.1, we add a random noise uniformly distributed in $[-\frac{\varepsilon}{2}, \frac{\varepsilon}{2}]$, and we enforce nonnegativity of the data

$$N_\varepsilon = \max(N + \varepsilon\, r, 0).$$

We solve the direct problem on a regular grid of $I + 1$ points, on an interval $[0, 2L]$. We need L large enough, such that it is possible to assume that $N(x \geqslant L) \approx 0$ and we have checked it *a posteriori.* Indeed, we have seen that this property is essential when we use the inverse schemes on a domain $[0, L]$ in order to verify the balance laws (7)–(9). In other words, we solve the direct problem on a domain twice larger than for the inverse problem. In the numerical tests we take $L = 4$, and we show the numerical solution N only on the interval $[0, L]$ since it is uniformly small on $[L, 2L]$.

We solve the inverse problem by the different methods on a regular grid of $I_1 + 1$ points on $[0, L]$, with $\Delta x_1 = L/I_1$. This grid is taken ten times finer than the grid used for the direct problem, *i.e.* we take $I_1 = 10I$. Since we have chosen L large enough so that $N(x \geqslant L) \approx 0$, we have always obtained that indeed $H(x \geqslant L) \approx 0$.

As before, we denote by H^Q and H^f the solution data H obtained respectively by the quasi-reversibility method of Section 3.4 and by the first filtering approach (from zero) of Section 3.3. We also define a solution H^{fQ} by mixing both methods, *i.e.* by solving the following equation:

$$\begin{cases} \alpha \frac{\partial}{\partial x}(B_{\varepsilon,\alpha} N_\varepsilon)(y) + 4B_{\varepsilon,\alpha}(y) N_\varepsilon(y) - B_{\varepsilon,\alpha}(x) N_\varepsilon(x) = \left(\frac{\partial}{\partial x} N_\varepsilon + \lambda_{\varepsilon,\alpha} N_\varepsilon \right) * \rho_\alpha(x), \quad x \geqslant 0, \\ \\ B_{\varepsilon,\alpha}(x = 0) N_\varepsilon(x = 0) = 0, \end{cases} \tag{31}$$

where $\lambda_{\varepsilon,\alpha}$ is defined by

$$\lambda_{\varepsilon,\alpha} = \frac{\int_0^\infty N_\varepsilon * \rho_\alpha dx}{\int_0^\infty x N_\varepsilon * \rho_\alpha dx + \frac{\alpha}{4} \int_0^\infty N_\varepsilon * \rho_\alpha dx}. \tag{32}$$

The relative error is measured, as seen in Theorem 2.1 and in Theorem 5.1 of [1], by

$$\delta^Q = \frac{||BN_\varepsilon - H^Q||^2_{l^2}}{||N_\varepsilon||_{l^2}}, \qquad \delta^f = \frac{||BN_\varepsilon - H^f||^2_{l^2}}{||N_\varepsilon||_{l^2}}, \qquad \delta^{fQ} = \frac{||BN_\varepsilon - H^{fQ}||^2_{l^2}}{||N_\varepsilon||_{l^2}}.$$

We have divided by $||N_\varepsilon||_{L^2}$ and not by $||N||_{H^2}$ because in practice we only know the entry data with noise.

In order to illustrate the accuracy of our method, we also compare it to a naive way (brute force) of considering the equation. Namely, we approximate $\frac{\partial}{\partial x}N(x)$ by a second-order Euler scheme without regularization. It gives a solution H^b by the same formula (30), where we simply take $\alpha = 0$.

The Direct Problem. We have first tested the direct problem for various division rates B. Three different solutions N for three given division rates B are depicted in Figure 1 with 800 grid points.

In the particular case when B is constant, we can go further and evaluate the computational error. Then, we know that $\lambda = B$ and the exact solution N_{exact} can be explicitly calculated, as shown in [16, 3], by the formula:

$$N_{\text{exact}}(x) = \bar{N} \sum_{n=0}^{\infty} \alpha_n e^{-2^n Bx}, \tag{33}$$

where the coefficients are defined recursively by $\alpha_0 = 1$ and $\alpha_n = (-1)^n \frac{2\alpha_{n-1}}{2^n - 1}$, and $\bar{N}$ is chosen to ensure the mass one normalization. We take $B = 1$ and obtain the continuous curve of Figure 1. We can measure here the relative error by

$$\delta^D = \frac{\|N - N_{\text{exact}}\|_{l^1}}{\|N_{\text{exact}}\|_{l^1}},$$

where N represents the numerical solution of Section 3.1. We choose this norm because for B constant, the solution of the adjoint problem is $\phi = 1$ and the General Relative Entropy Principle ([18, 3]) gives us that this quantity decreases along the time iterations. Still for 800 points, we obtain $\delta^D = 7.7.10^{-3}$.

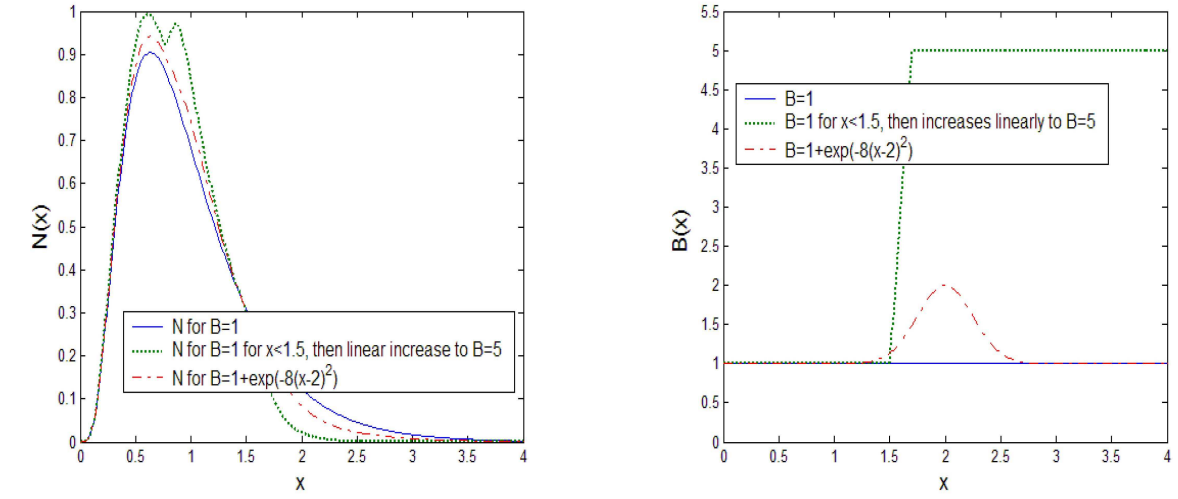


Figure 1: Solutions N (left) obtained by the numerical resolution of Section 3.1 for the direct problem with three different division rates B (right).

The noiseless case ($\varepsilon = 0$). In the simplest case where the data is perfectly known, *i.e.* for $\varepsilon = 0$, we verify that the different schemes allow us to recover B. Since the precision of the data is directly

linked to the number of points used in the scheme, we run the codes with 1.000 points for the direct problem (below, we will take only 100 points).

We test several values of α and we use the three functions B of Figure 1 for each method for the inverse problem. The error estimate is found to depend on the method used but not significantly on the division rate B. Therefore we have drawn in Figure 2 the average error estimates for the three division rates B. In Figure 3 we have depicted the products $B.N$ in the case $B = 1$ and $\alpha = 0.01$ (other cases are similar): it shows that the precision obtained is satisfactory. In Figures 4, 5 and 6 we have drawn the approximations of B in each of the three cases, calculated only for $N > 0.01$ (indeed, for N too small the division leads to insignificant results on B).

Not surprisingly, the brute force method reveals to be satisfactory, with an error estimate of $\delta^b = 1.3.10^{-2}$, since we are in the case where N is very regular. The filtering method can reach this level of error for $\alpha = 10^{-2}$ but cannot go further. However, both the quasi reversibility method and the mixed method given by Equation (31) improve it with minimum values $\delta^Q = 6.9.10^{-3}$ and $\delta^{fQ} = 6.5.10^{-3}$ reached for $\alpha = 10^{-2}$.

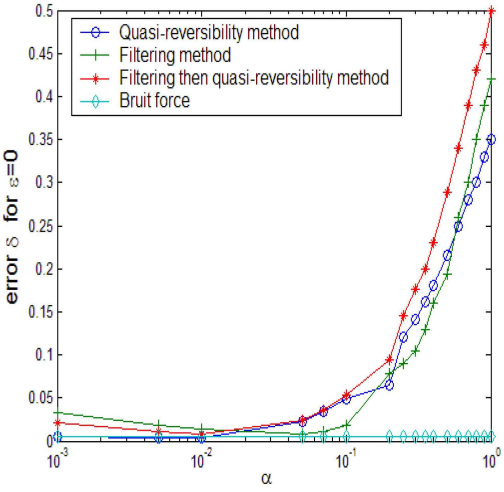


Figure 2: For $\varepsilon = 0$, numerical errors obtained with the different methods for the inverse problem.

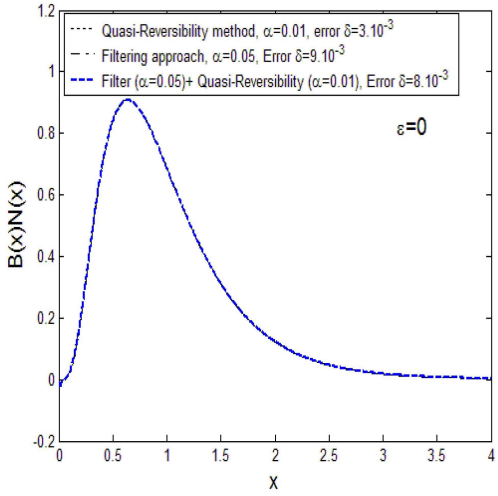


Figure 3: Numerical reconstruction of $B.N$ obtained by each method for the inverse problem when $B = 1$, $\varepsilon = 0$ and $\alpha = 0.01$.

Link between the noise level ε and the regularization parameter α. For noise levels $\varepsilon = 0.01$, $\varepsilon = 0.05$ and $\varepsilon = 0.1$ respectively, the Figures 7, 8 and 9 give the curves ε as a function of α for the three inverse methods. We compare the reconstructed division rates B in Figures 10 and 11.

Each of the error curves presents a minimum for an optimal value of α, as expressed by estimate (13) for instance. In Figures 12, 13, 14 and 15, we have compared three curves, drawn in a log-log scale: $\sqrt{\varepsilon}$ to serve as a reference curve, $f(\varepsilon) = \min_{\alpha} \delta(\alpha, \varepsilon)$, and $g(\varepsilon) = \arg\min_{\alpha} \delta(\alpha, \varepsilon)$. One can see that for each method, these three curves have comparable slopes ($\frac{1}{2}$ on a log-log scale): they show that even though the combination of filtering and quasi-reversibility method improves the optimal errors in absolute value, it does not change the order of convergence of the approximation, which remains of order $O(\sqrt{\varepsilon})$. Figure 15 gives also the convergence of the filtering method for much smaller values

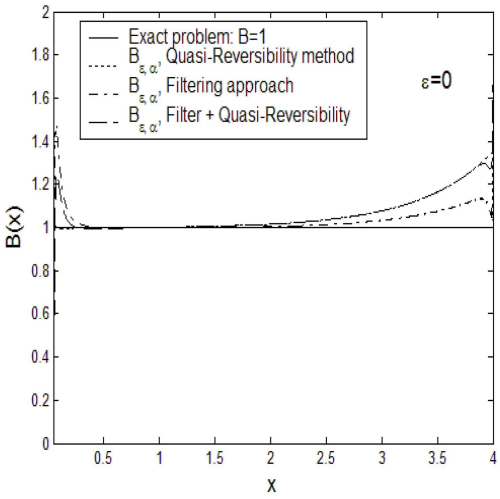


Figure 4: Reconstructed division rate B using the three inverse methods, for $\varepsilon = 0$, $\alpha = 0.01$ with N computed from $B = 1$.

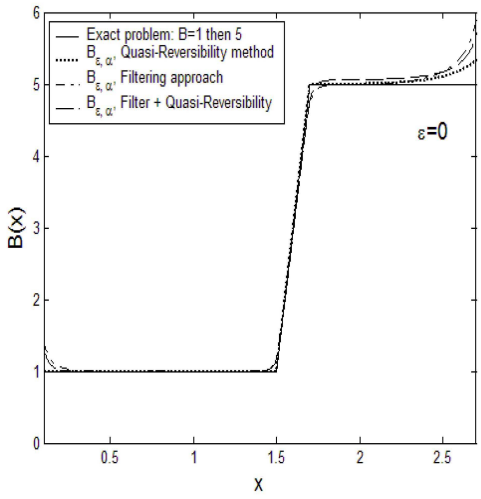


Figure 5: Reconstructed division rate B, for $\varepsilon = 0$, $\alpha = 0.01$ and a jump $B = 1$ to 5 as in Figure 1

of ε (for which an increased number of 500 points has been taken, in order to avoid numerical bias): the comparison with $\sqrt{\epsilon}$ is there particularly evident, and we have obtained similar curves for the two other methods. The speed of convergence is though in complete accordance with the theoretical estimate (13).

Influence of the choice of λ_0 instead of $\lambda_{\varepsilon,\alpha}$. To evaluate the influence of the error term due to the distance $|\lambda_{\varepsilon,\alpha} - \lambda_0|$, we compare the curves obtained respectively by taking on the inverse code the exact λ_0 or the value $\lambda_{\varepsilon,\alpha}$ expressed by the balance laws. They are drawn in Figure 13 for the quasi-reversibility method. They show that even though the *a priori* knowledge of λ_0 improves the error in absolute value, it does not change the order of convergence of the scheme. Thus it is in complete accordance with Estimate (13).

5 Conclusions

We have considered size-structured equations connected to several areas of biology from cell division to prion proliferation by aggregation and fragmentation. We have addressed the numerical efficiency of some inverse problem solution methods to tackle the problem of recovering the division rate from the size distribution of cells. The latter involves a dilation equation with a singular right-hand side that needs regularization for actual implementation. For that purpose, we have introduced a filtering method and proved its convergence for noisy data. This method brings in an operator that has a non-trivial kernel and we have selected a numerical approximation that is able to recover the natural solution we want to reach.

The implementation of the inverse algorithm, based on the filtering method, confirms the convergence analysis. In particular, there is an optimal regularization parameter as can be seen in the graphs of Figures 7, 8 or 9 for instance. Comparison with a quasi-reversibility method introduced

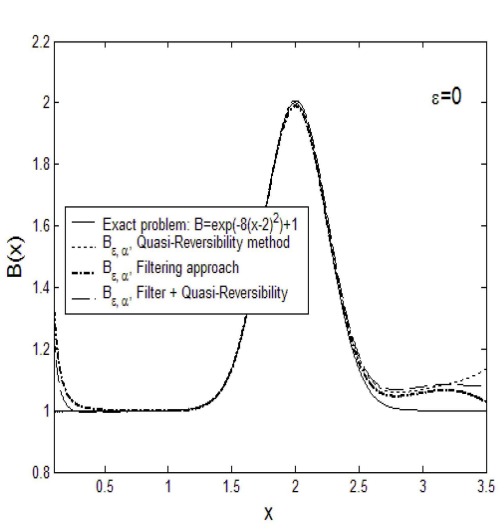


Figure 6: Reconstructed division rate B, for $\varepsilon = 0$, $\alpha = 0.01$ and $B = 1 + exp(-8(x-2)^2)$.

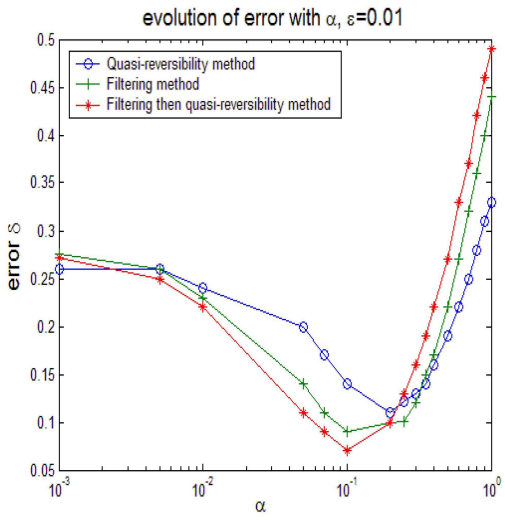


Figure 7: Numerical error when $\varepsilon = 0.01$ for the different methods.

earlier leads to the conclusion that a combination of filtering and quasi-reversibility methods seems to be more efficient because the oscillations are reduced, but without improving the rate of convergence.

We also analyzed the impact of using the exact value of λ_0 or the λ_ε on the different solutions of the inverse problem. In our simulations, the difference between using λ_0 or λ_ε seemed to be immaterial as far as the accuracy of method is concerned. This is in perfect accordance with the theoretical estimate (13).

The above remarks open several directions for continuation and extension of the present work. On the practical side, the present work sets the stage for the use of experimental data either from the existing literature or from more recent biological experiments. On the theoretical side, the possibility of improving the convergence by combining the filtering and quasi-reversibility methods should be investigated further.

Finally, we point out that although the Tikhonov method is more standard, we did not study it so far because it seems more time consuming. Indeed, iterations are needed to solve both the direct problem and the inverse one. To overcome such difficulty a completely new theory has to be developed so as to suit the particular structure of our model. This provides yet another direction for future work.

Acknowledgments

The authors were supported by the CNPq-INRIA agreement INVEBIO. JPZ was supported by CNPq under grants 302161/2003-1 and 474085/2003-1. JPZ and BP are thankful to the RICAM special semester and to the International Cooperation Agreement Brazil-France. Part of this work was conducted during the Special Semester on Quantitative Biology Analyzed by Mathematical Methods, October 1st, 2007 -January 27th, 2008, organized by RICAM, Austrian Academy of Sciences.

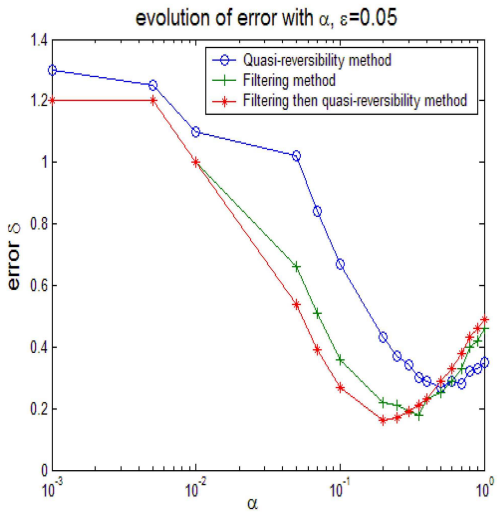


Figure 8: Numerical error when $\varepsilon = 0.05$.

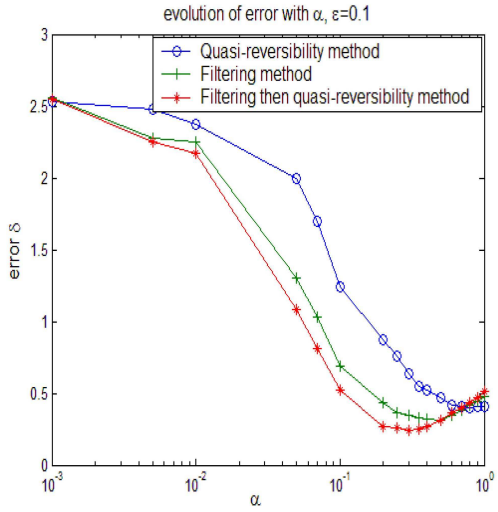


Figure 9: Numerical error when $\varepsilon = 0.1$.

A Well-Posedness of Functional Equation Associated to the Inverse Problem.

We have seen that the regularization method for the inverse problem relies mostly on solving the equation

$$4H(2x) - H(x) = L(x), \qquad x \geq 0. \tag{34}$$

Eventhough this equation is formally very simple, its analysis reveals some complexity. It may admit several solutions in general. Among them, we can mention two with simple representation formulas (we leave to the reader to check they are indeed formally solutions)

$$H^{(1)}(x) = \sum_{n=1}^{+\infty} 2^{-2n} L(2^{-n}x). \tag{35}$$

$$H^{(2)}(x) = -\sum_{n=0}^{+\infty} 2^{2n} L(2^{n}x), \tag{36}$$

To clarify this issue and motivate our choice of a solution, we first state general results concerning solutions to (34) and then come back to our original problem (5).

We first mention the following

Proposition A.1 *Let $L \in L^2(\mathbb{R}_+, x^p dx)$, with $p \neq 3$, then there exists a unique solution $H \in L^2(\mathbb{R}_+, x^p dx)$ to (34) and*

- *for $p < 3$, this solution is given explicitly by the formula (35). Furthermore, for $1 \leq q \leq \infty$, if $L \in L^q(\mathbb{R}_+)$ then $H^{(1)} \in L^q(\mathbb{R}_+)$.*
- *for $p > 3$, this solution is given explicitly by the formula (36).*

Because we look for an integrable function H (the number of cells is supposedly finite), the function $H^{(1)}$ is preferable (take $q = 1$). It also behaves better near $x \approx 0$ because the weight $p < 3$ imposes that $H^{(1)}$ vanishes at 0 as we expect.

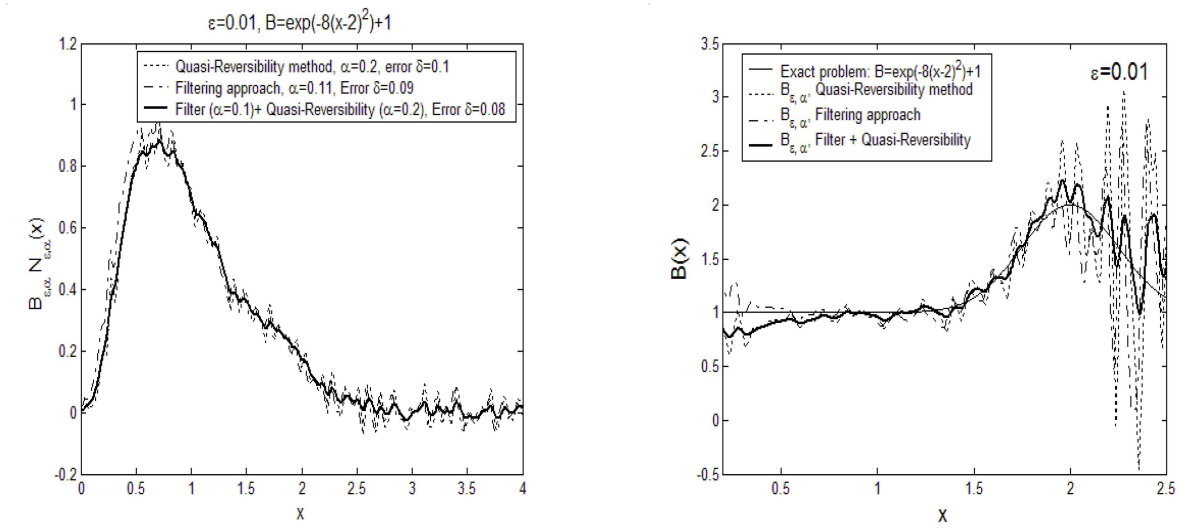


Figure 10: In the case $\varepsilon = 0.01$, $\alpha = 0.05$, $B = 1 + e^{-8(x-2)^2}$, numerical solution $B.N$ (left) and B (right) by the different methods.

From the point of view of exact solutions of the direct problem, we find that $H = BN$ and L belong to all spaces $L^2(\mathbb{R}_+, x^p dx)$ for all $p \in \mathbb{R}$. Therefore, the two solutions coincide and in principle we could choose any of them. In practice, errors on the data L are better handled by $H^{(1)}$ than by $H^{(2)}$ for the afore mentioned reason. Notice indeed that these two solutions are different in general. One can check for instance that for $L = 0$, there is a singular distributional solution $\delta'_{x=0}$. Furthermore,

Lemma A.2 *The solutions to (34) with $L = 0$ in $\mathcal{D}'(0,\infty)$ have the form $\frac{f(\log(x))}{x^2}$ with $f \in \mathcal{D}'(\mathbb{R})$ a $\log(2)-$ periodic distribution.*

Proof of Proposition A.1: We consider the Hilbert space $X = L^2(\mathbb{R}_+, x^p dx)$ and we simply apply the Lax-Milgram theorem to a properly chosen bilinear form.

Case 1, $p < 3$. We solve the equation in the variable $y = 2x$ that is (20). and consider the bilinear form $a(u, v)$ on $X \times X$ defined by

$$a(u,v) = 4\int_0^\infty u(y)v(y)y^p dy - \int_0^\infty u(\frac{y}{2})v(y)y^p dy.$$

This form is obviously continuous and it remains to prove that it is coercive. We have

$$a(u,u) = 4\int_0^\infty u(y)^2 y^p dy - \int_0^\infty u(\frac{y}{2})u(y)y^p dy \geqslant (4 - 2^{\frac{p+1}{2}}) \int_0^\infty u(y)^2 y^p dy,$$

and it is indeed coercive as long as $\alpha = 4 - 2^{\frac{p+1}{2}}$ is positive which holds true for $p < 3$. The Lax-Milgram Theorem asserts that there is a unique $H \in X$ such that $a(H, \cdot) = (L, \cdot)$, where $(\cdot, \cdot)$ denotes the inner product in X, that is a solution of (20).

Case 2, $p > 3$. We work in the variable x and consider the continuous bilinear form $b(u, v)$ on $X \times X$ defined by

$$b(u,v) = -4\int_0^\infty u(2x)v(x)x^p dx + \int_0^\infty u(x)v(x)x^p dx.$$

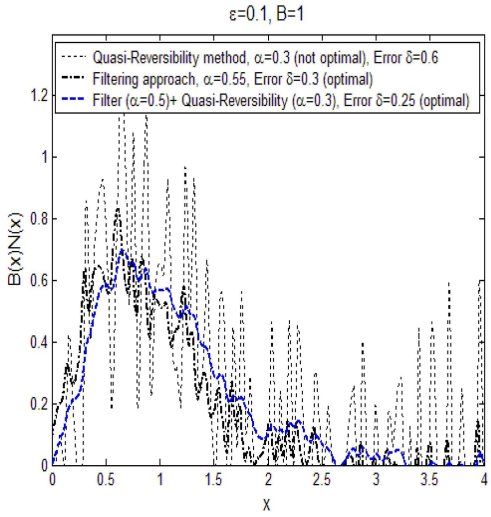


Figure 11: In the case $\varepsilon = 0.1$ and $B = 1$, the numerical solution $B.N$ by the different methods.

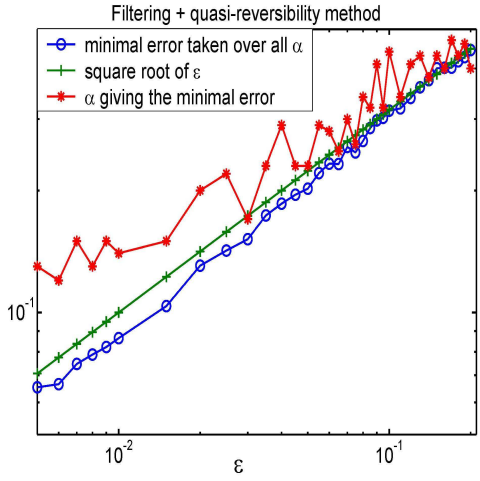


Figure 12: Filtering and quasi-reversibility method: minimal error as a function of the noise level ε for the optimal value α, with a comparison to the theoretical curve $\sqrt{\varepsilon}$.

The same calculation leads us to:

$$b(u,u) \geqslant \alpha \int_0^\infty u(x)^2 x^p dx, \qquad \text{with} \quad \alpha = 1 - 2^{\frac{3-p}{2}} > 0,$$

and the same conclusion holds.

To check formulae (36) and (35), it remains to prove that these solutions belong to the corresponding spaces:

$$\|H^{(1)}\|_{L^2(\mathbb{R}_+,x^p dx)} \leqslant \sum_{n=1}^{\infty} 2^{-2n} ||L(2^{-n}x)||_{L^2(\mathbb{R}_+,x^p dx)} = \sum_{n=1}^{\infty} 2^{\frac{n}{2}(p-3)} ||L(x)||_{L^2(\mathbb{R}_+,x^p dx)}.$$

This sum converges *iff* $p > 3$. In the same way, we write:

$$||H^{(2)}||_{L^2(x^p dx)} \leqslant \sum_{n=0}^{\infty} 2^{2n} ||L(2^n x)||_{L^2(\mathbb{R}_+,x^p dx)} = \sum_{n=0}^{\infty} 2^{\frac{n}{2}(3-p)} ||L(x)||_{L^2(\mathbb{R}_+,x^p dx)},$$

which converges *iff* $p > 3$. □

Proof of Lemma A.2: When $L = 0$, we first define $\mathcal{H} \in \mathcal{D}'(0,\infty)$ as the second antiderivative of H, and notice that it should verify

$$\mathcal{H}(2x) = \mathcal{H}(x).$$

We perform the change of variables $y = \log(x)$ and notice that, if $\mathcal{H} \in \mathcal{D}'(0,\infty)$, it is equivalent to look for solutions $f \in \mathcal{D}'(\mathbb{R})$ of

$$f\big(y + \log(2)\big) = f(y). \tag{37}$$

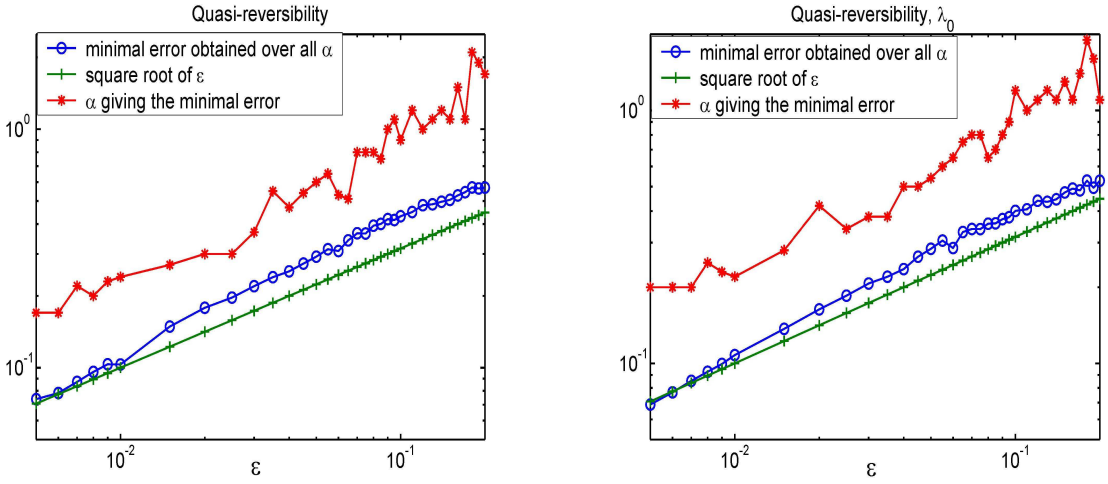


Figure 13: Quasi-reversibility method, with (left) $\lambda_{\varepsilon,\alpha}$ given by relation (9) or (right) *a priori* knowledge of λ_0 : minimal error and optimal regularization parameter α as functions of the noise level ε, with a comparison to the theoretical curve $\sqrt{\varepsilon}$. We see that the *a priori* knowledge of λ_0 does not improve the speed of convergence of the scheme.

Hence, all the solutions in $\mathcal{D}'(0,\infty)$ are given by $\frac{f\big(\log(x)\big)}{x^2}$, where $f \in \mathcal{D}'(\mathbb{R})$. □

To conclude this Appendix, we come back to our original problem (5) and draw the consequences in terms of B, not H.

Theorem A.3 *Let $N \in L^2(\mathbb{R}_+)$, with $N(x) > 0$ for $x > 0$. Let $L \in L^2(\mathbb{R}_+)$. There exists a unique $B \in L^2(\mathbb{R}_+, N^2dx)$ solution of*

$$4B(2x)N(2x) - B(x)N(x) = L(x). \tag{38}$$

Proof: The theorem follows directly from Proposition A.1 for $p = 0$, and since $N > 0$, we can define $B = H/N$ for $B \in L^2(\mathbb{R}_+, N^2dx)$. □

This theorem shows that we can find a solution B of (5) for all N and all λ, this is the basis of our algorithm. However, if we want that the solution B belongs to the space $L^1(\mathbb{R}_+; xN(x)dx)$, integration of (38) multiplied by x shows that L has to satisfy the condition

$$\int_0^\infty xL(x)dx = 0.$$

Applying this to Equation (5), we recover that $\lambda_0 = \int_0^\infty N(x)dx / \int_0^\infty xN(x)dx$. In the case of Equations (6) and (11) respectively, we get formulae (9) and (12), which discrete versions are expressed by (29) and (27).

In view of these considerations, it is better to use a discrete scheme defined by a matrix A that preserves a similar discrete property. Namely, for all $H = (H_i)$, we should have $\sum_i i(AH)_i = 0$, in other words the vector of components i belongs to the kernel of the adjoint of A. Indeed, this property yields the (discrete) regularity $H \in L^1(\mathbb{R}_+; xdx)$.

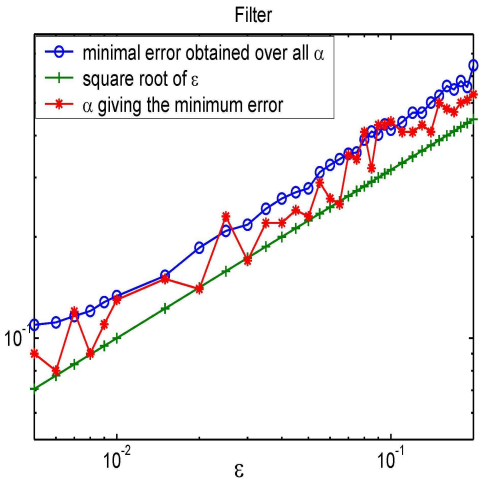


Figure 14: Filtering method for standard levels of noise.

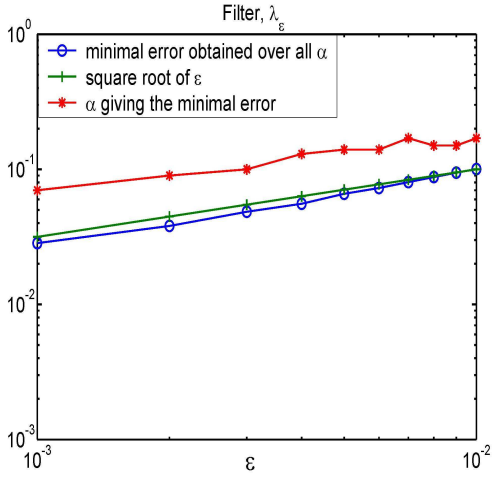


Figure 15: Filtering method for smaller values of ε and increased number of points.

References

[1] Benoît Perthame and Jorge P. Zubelli. On the inverse problem for a size-structured population model. *Inverse Problems*, 23(3):1037–1052, 2007.

[2] J. A. J. Metz and O. Diekmann. Formulating models for structured populations. In *The dynamics of physiologically structured populations (Amsterdam, 1983)*, volume 68 of *Lecture Notes in Biomath.*, pages 78–135. Springer, Berlin, 1986.

[3] Benoît Perthame. Transport equations arising in biology. In *Frontiers in Mathematics*, Frontiers in Mathematics. Birkhauser, 2007.

[4] Heinz W. Engl, William Rundell, and Otmar Scherzer. A regularization scheme for an inverse problem in age-structured populations. *J. Math. Anal. Appl.*, 182(3):658–679, 1994.

[5] Mats Gyllenberg, Andrei Osipov, and Lassi Päivärinta. The inverse problem of linear age-structured population dynamics. *J. Evol. Equ.*, 2(2):223–239, 2002.

[6] William Rundell. Determining the birth function for an age structured population. *Math. Population Stud.*, 1(4):377–395, 397, 1989.

[7] Michael Pilant and William Rundell. Determining a coefficient in a first-order hyperbolic equation. *SIAM J. Appl. Math.*, 51(2):494–506, 1991.

[8] J.A. Carrillo and T. Goudon. A numerical study on large-time asymptotics of the Lifshitz-Slyozov system. *J. Sci. Comput.*, 20(1):69–113, 2004.

[9] Jean-François Collet, Thierry Goudon, Frédéric Poupaud, and Alexis Vasseur. The Beker-Döring system and its Lifshitz-Slyozov limit. *SIAM J. Appl. Math.*, 62(5):1488–1500, 2002.

[10] Jean-François Collet, Thierry Goudon, and Alexis Vasseur. Some remarks on large-time asymptotic of the Lifshitz-Slyozov equations. *J. Statist. Phys.*, 77(1-2):139–152, 1999.

[11] Philippe Laurençot. Convergence to self-similar solutions for a coagulation equation. *Z. Angew. Math. Phys.*, 56(3):398–411, 2005.

[12] Philippe Laurençot and Stéphane Mischler. Liapunov functionals for Smoluchowski's coagulation equation and convergence to self-similarity. *Monatsh. Math.*, 146(2):127–142, 2005.

[13] V. Calvez, N. Lenuzza, D. Oelz, J.-P. Deslys, F. Mouthon, P. Laurent, and B. Perthame. Bimodality, prion aggregates infectivity and prediction of strain phenomenon, arXiv : 0802.2024 (2008).

[14] J. Masel, V.A.A. Jansen, and M.A. Nowak. Quantifying the kinetic parameters of prion replication. *Biophys. Chem*, 77:139–152, 1999.

[15] Meredith L. Greer, Laurent Pujo-Menjouet, and Glenn F. Webb. A mathematical analysis of the dynamics of prion proliferation. *J. Theoret. Biol.*, 242(3):598–606, 2006.

[16] Benoît Perthame and Lenya Ryzhik. Exponential decay for the fragmentation or cell-division equation. *J. Differential Equations*, 210(1):155–177, 2005.

[17] Philippe Michel. Existence of a solution to the cell division eigenproblem. *Model. Math. Meth. Appl. Sci.*, 16(suppl. issue 1):1125–1153, 2006.

[18] Philippe Michel, Stéphane Mischler, and Benoît Perthame. General relative entropy inequality: an illustration on growth models. *J. Math. Pures Appl. (9)*, 84(9):1235–1260, 2005.

[19] Johann Baumeister and Antonio Leitão. *Topics in inverse problems*. Publicações Matemáticas do IMPA. [IMPA Mathematical Publications]. Instituto Nacional de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, 2005. 25º Colóquio Brasileiro de Matemática. [25th Brazilian Mathematics Colloquium].

[20] Heinz W. Engl, Martin Hanke, and Andreas Neubauer. *Regularization of inverse problems*, volume 375 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1996.

[21] Robert Lattès. Non-well-set problems and the method of quasi reversibility. In *Functional Analysis and Optimization*, pages 99–113. Academic Press, New York, 1966.

[22] R. Lattès and J.-L. Lions. *Méthode de quasi-réversibilité et applications*. Travaux et Recherches Mathématiques, No. 15. Dunod, Paris, 1967.

[23] François Bouchut. *Nonlinear stability of finite volume methods for hyperbolic conservation laws, and well-balanced schemes for sources*. Frontiers in Mathematics. Birkhäuser, 2004.

[24] Randall J. LeVeque. *Finite Volume Methods for Hyperbolic Problems*. Frontiers in Mathematics. Cambridge University Press, 2002.

[25] Edwige Godlewski and Pierre-Arnaud Raviart. *Numerical approximation of hyperbolic systems of conservation laws*. Applied Mathematical Sciences, vol. 118. Springer, 1996.

[26] Denis Serre. *Matrices: Theory and Applications*. TELOS, 2002.

[27] Gilbert Strang. Wavelets and dilation equations: a brief introduction. *SIAM Review*, 31(4):614–627, 1989.