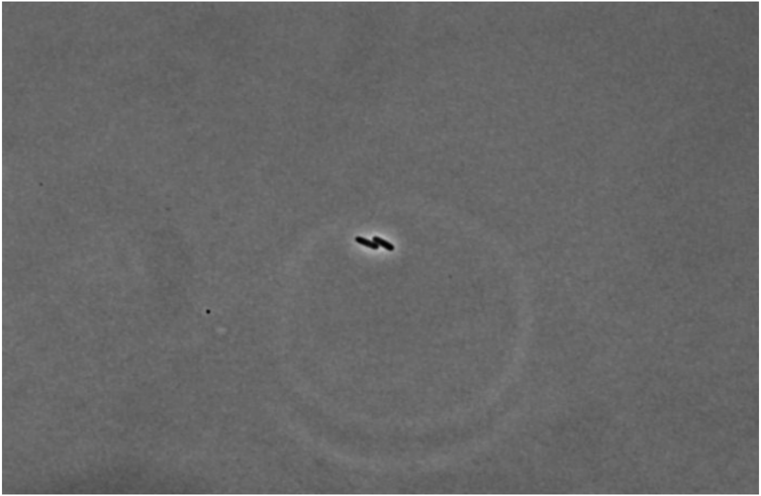
Comment estimer la division cellulaire ?
(et la fragmentation de polymères)

Marie Doumic

INRIA and Sorbonne Université, Paris

Aussois, chaire MMB, 14-17 juin 2022

Bacterial growth (E. coli here)



From E. J. Stewart, R. Madden, G. Paul, F. Taddei, Plos Biol, 2005

What triggers bacterial division?

OR

Different ways of investigation:

- details the intracellular mechanisms
  many studies
- Observe and understand the population dynamics

What triggers bacterial division?

OR

Different ways of investigation:

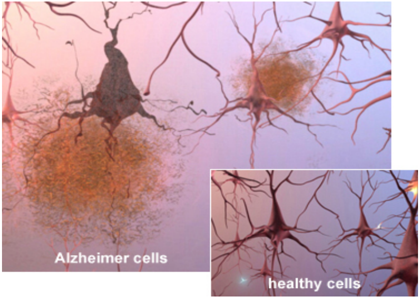
- details the intracellular mechanisms
  many studies
- Observe and understand the population dynamics

Question: Can we deduce laws from our observations?

Protein polymerization

Common point between:

- Alzheimer's (illustrated)
- Prion (mad cow)
- Huntington's
- and some others (Parkinson's, etc)?



(J. of Alzh.'s D., 2014)

Neurodegenerative diseases characterized by abnormal accumulation of protein aggregates called AMYLOIDS

Protein polymerization: main issues

- Understand what are the key polymerization mechanisms
- Identify transient species, and the "most infectious" sizes of polymers
- Study the models...
- How to select and calibrate the models; write new models...

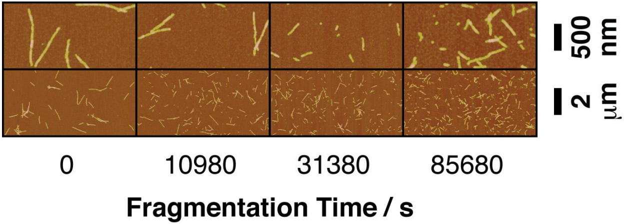
Of key importance: size distributions

1000 nm

Protein polymers fragmentation

Experimental device: Atomic Force Microscopy (AFM)
Performed at the University of Kent, UK, by W.F. Xue's team.

Several proteins: β_2m, αsynuclein, Lysozyme, βLactoglobuline
Fragmentation by agitation



(From W.F. Xue, S. Radford, Biophys. J., 2013)

Can we estimate the division features (rate, where the fibrils divide) from such images?

Steps towards "laws" of division

Steps towards "laws" of division

1. Make the most of direct observations
   Methods: statistical analysis, density estimation...

Steps towards "laws" of division

1. Make the most of direct observations
   Methods: statistical analysis, density estimation...

2. Make assumptions or simplifications

Steps towards "laws" of division

1. Make the most of direct observations
   Methods: statistical analysis, density estimation...

2. Make assumptions or simplifications

3. Build model(s)
   Methods: stochastic processes / ODE or PDE ...

Steps towards "laws" of division

1. Make the most of direct observations
   Methods: statistical analysis, density estimation...

2. Make assumptions or simplifications

3. Build model(s)
   Methods: stochastic processes / ODE or PDE ...

4. Analyse the model(s): long-time asymptotics, qualitative behaviours...

Steps towards "laws" of division

1. Make the most of direct observations
   Methods: statistical analysis, density estimation...

2. Make assumptions or simplifications

3. Build model(s)
   Methods: stochastic processes / ODE or PDE ...

4. Analyse the model(s): long-time asymptotics, qualitative behaviours...

5. Calibrate the model(s): estimation of unobserved parameters
   Methods: inverse problems, statistics

Steps towards "laws" of division

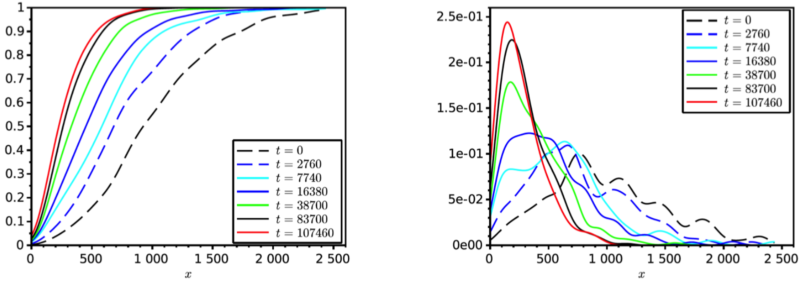
1. Make the most of direct observations
   Methods: statistical analysis, density estimation...
2. Make assumptions or simplifications
3. Build model(s)
   Methods: stochastic processes / ODE or PDE ...
4. Analyse the model(s): long-time asymptotics, qualitative behaviours...
5. Calibrate the model(s): estimation of unobserved parameters
   Methods: inverse problems, statistics
6. Back to the data to (in)validate the model(s)

First step: take the most of our data

(before writing down a math model)

1. Observations for the protein fragmentation case

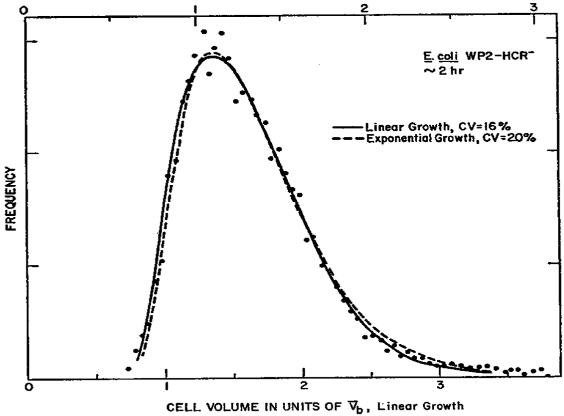
At different times, a sample of fibril sizes is measured $\rightsquigarrow \frac{n(t,x)}{\int n(t,x)dx}$.



Left: cumulative distribution functions, Right: density functions, at several time points.

2. Observations of the population for bacteria

1st historical observations, the simplest and often the only possible ones, and confirm the asymptotic behavior:



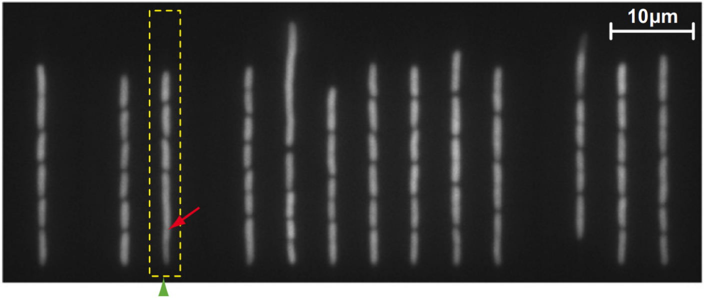
Observation (from Kubitschek, 1969): DOUBLING TIME and STEADY SIZE DISTRIBUTION.

3."Complete" observations for bacteria

Major advantage of in vitro bacterial growth: **EVERYTHING may be measured** to control/validate the assumptions.

2 types of data:

- initial video: all descendants till a certain time, several microcolonies (Stewart et al, Plos Biol, 2005)
- 1 daughter cell kept at each generation, till a certain time, several lineages (Wang, Robert et al, Current Biology, 2010)



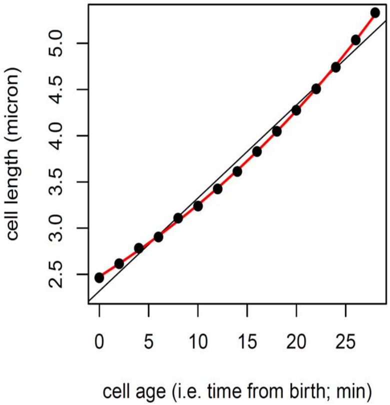
The way we observe the data influence the math modeling.

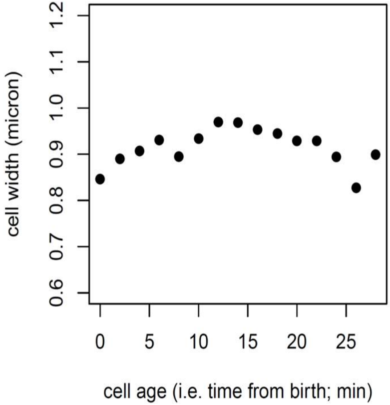
3. Complete observations: individual growth

commonly accepted after much debate: exponential growth:

$$\frac{dx}{dt} = \kappa x.$$

(Stewart et al, Plos Biol, 2005)





3. Complete observation: individual growth

variability of the exponential rate κ among cells

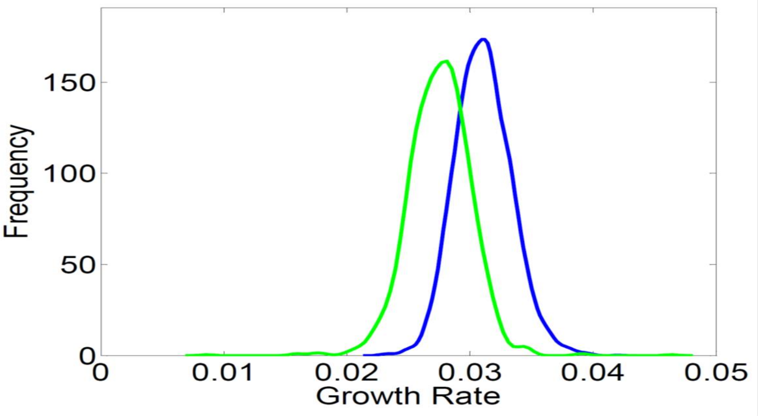


Figure: growth rate distrib. (min^{-1})

Heritability? See (Delyon, de Saporta, Krell, Robert, 2018)

3. Complete observations: population growth

Growth of the population: exponential with Malthus parameter λ (almost) equal to the (average) individual growth rate κ.
Doubling time ($= Log(2)/\kappa$) of approx. 20 min.

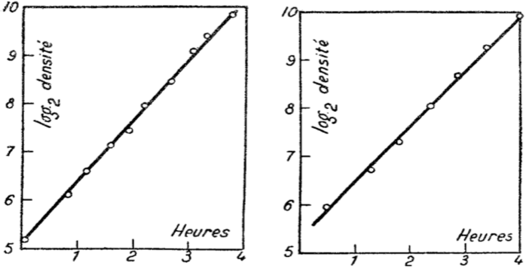


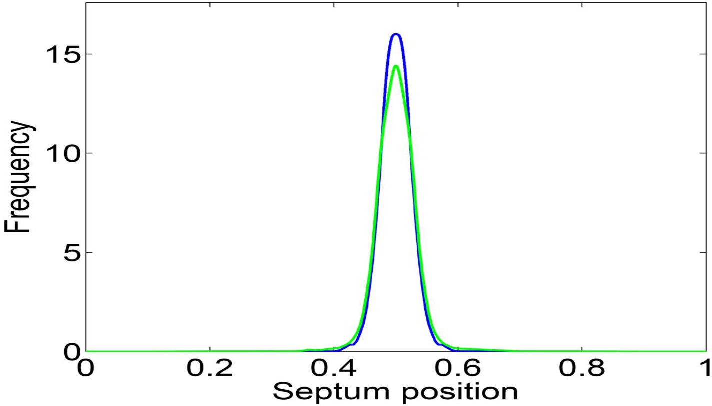
Fig. 10. — Phase exponentielle de la croissance d'une culture de *B. coli* en milieu synthétique, avec 300 mgr. par l. de glucose. Coordonnées semi-logarithmiques.

Fig. 11. — Phase exponentielle de la croissance d'une culture de *B. subtilis* en milieu synthétique, avec 500 mgr. par l. de saccharose. Coordonnées semi-logarithmiques.

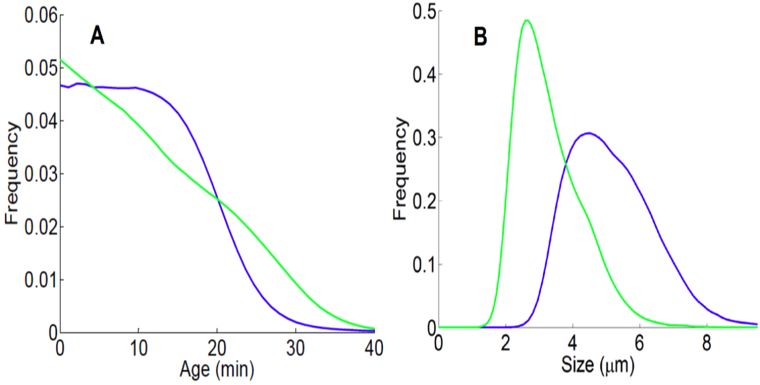
Figure: Monod's 1942 thesis on *E. Coli* culture cells.

3. Complete observation: division

Distribution of the ratio (size of daughter/size of mother)

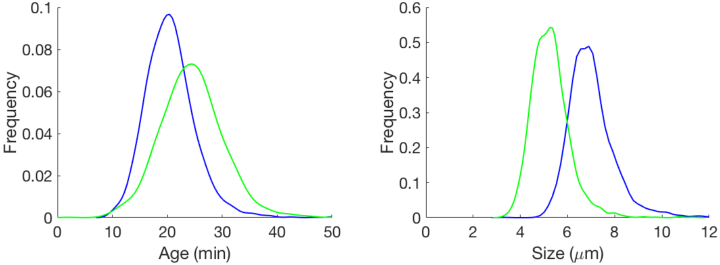


3. Complete observation: "all cells" distributions



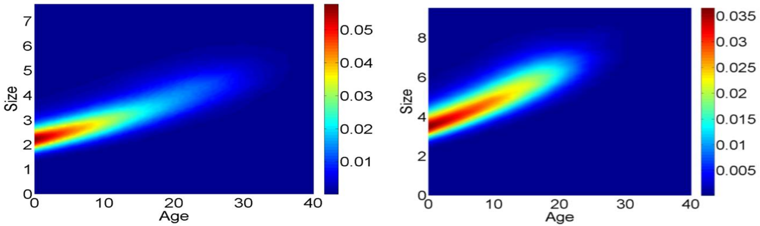
Blue: 1 branch/genealogical data
Green: whole tree data till a certain time

3. Complete observation: "at division" distributions



Blue: 1 branch/genealogical data
Green: whole tree data till a certain time

3. Complete observation: joint age-size distribution



Left: Age-Size Distribution for all cells - "petri dish" / whole population case
Right: Age-Size Distribution for microfluidic device - "1-branch data"

Second step: making assumptions
(before writing down a math model)

Assumptions: some simplification

based on direct observations:

- daughter cell size= half of mother cell size
- growth rate = constant among cells (neglect variability)

$$\frac{dx}{dt} = \kappa x$$

- infinite nutrient and space
- first cell selected at random

Assumptions: modeling

- no memory
- a particle of size x may divide with a division rate B depending on age
  OR
- a particle of size x may divide with a division rate B depending on size
  OR
- a particle of size x may divide with a division rate B depending on size AND age AND/OR something else...

Third step: models

(that we will analyse and calibrate)

Models

2 main ways of translating mathematically the previous assumptions:

1. probability: model each cell
2. PDE: model the population of cells, considered either as large or in expectation

Mathematical Modelling of the protein fragmentation experiment

Noise model:

At time t, we measure $x_1, \cdots x_n$ an *i.i.d.* sample of density $n(t,x)$

Model for $n(t,x)$: the fragmentation equation

$$\underbrace{\frac{\partial n}{\partial t}(t,x)}_{\text{Evol. of number of polymers}} = \underbrace{-B(x)n(t,x)}_{\text{Death}} + \underbrace{\int_{y=x}^{y=\infty} k(y,x)B(y)n(t,y)dy}_{\text{Creation}}$$

Measurement: at different times t_i, a (noisy) $n(t_i,x)$ provided by samples $x_1(t_i), \cdots x_{n(t_i)}(t_i)$

Unknowns: the non-parametric functions $B(x)$ (fragmentation rate) and $k(y,x)$ (fragmentation kernel)

The pure fragmentation equation: basic properties

"Fragmentation conserves the mass": $\forall\ B(\cdot)n(t,\cdot) \in L^1(xdx)$:

$$\int\limits_0^\infty xB(x)n(t,x)dx = \int\limits_0^\infty \int\limits_x^\infty xk(y,x)B(y)n(t,y)dydx$$

The fragmentation kernel $k(y,x)$ must satisfy

- $y \to k(y,\cdot)$ nonnegative measure with $Supp\big(k(y,\cdot)\big) \subset [0,y]$ (and $\forall\ \psi\ C^0$, $y \to \int \psi(x)k(y,dx)$ is Lebesgue-measurable)
- mass conservation

The pure fragmentation equation: basic properties

"Fragmentation conserves the mass": $\forall\ B(\cdot)n(t,\cdot) \in L^1(xdx)$:

$$\int_0^\infty xB(x)n(t,x)dx = \int_0^\infty \int_x^\infty xk(y,x)B(y)n(t,y)dydx$$

The fragmentation kernel $k(y,x)$ must satisfy

- $y \to k(y,\cdot)$ nonnegative measure with $Supp\big(k(y,\cdot)\big) \subset [0,y]$ (and $\forall\ \psi\ C^0$, $y \to \int \psi(x)k(y,dx)$ is Lebesgue-measurable)
- mass conservation $\implies \int_0^y xk(y,dx) = y$
- If binary fragmentation:

The pure fragmentation equation: basic properties

"Fragmentation conserves the mass": $\forall\ B(\cdot)n(t,\cdot) \in L^1(xdx)$:

$$\int_0^\infty xB(x)n(t,x)dx = \int_0^\infty \int_x^\infty xk(y,x)B(y)n(t,y)dydx$$

The fragmentation kernel $k(y,x)$ must satisfy

- $y \to k(y,\cdot)$ nonnegative measure with $Supp\big(k(y,\cdot)\big) \subset [0,y]$ (and $\forall\ \psi\ C^0$, $y \to \int \psi(x)k(y,dx)$ is Lebesgue-measurable)
- mass conservation $\implies \int\limits_0^y xk(y,dx) = y$
- If binary fragmentation: $\implies k(y,x) = k(y,y-x)$ (may be relaxed); with the mass conservation it implies $\int\limits_0^y k(y,dx) = 2$

Self-similar fragmentation: $k(y,x) := \frac{2}{y}k_0(\frac{x}{y})$, with $Supp(k_0) \subset [0,1]$.
2 main examples: uniform $k_0(z) \equiv 2$, equal mitosis $k_0(z) = 2\delta_{z=\frac{1}{2}}$.

Models: Branching processes modeling

See e.g. (Bansaye, Delmas, Marsalle, Tran, 2011); (Champagnat, Ferrière, Méléard, 2006 & 2008); (Bansaye, Méléard, 2015)

Piecewise Deterministic Markov Processes (PDMP):

- start: a singe cell of size x_0.
- cell's growth: deterministic.
- at each time, it has an instantaneous probabillity rate B to divide (jump); B depends on size x or age a of the cell.
- At division, two offspring of age 0 and initial size $x_1/2$, where x_1 is the size of the mother at division.
- The two offspring start independent growth (Markov property) according to the (deterministic) rate κ and divide according to the (probabilistic) rate B.

Stochastic models

Genealogical tree: infinite random marked tree

$$\mathcal{U} = \bigcup_{n=0}^{\infty} \{0,1\}^n \quad \text{with} \quad \{0,1\}^0 := \emptyset.$$

To each node $u \in \mathcal{U}$, we associate a cell with size at birth ξ_u and lifetime ζ_u.

If u^- denotes the parent of u then

$$\boxed{\xi_u = \frac{\xi_{u^-}}{2} \exp\left(\kappa \zeta_{u^-}\right).}$$

Stochastic models

Age model: the division depends on the age of the individual:

$$\mathbb{P}(\zeta_u \in (a, a+da) | \zeta_u \geq a) = B(a)da, \qquad \mathbb{P}(\zeta_u \geq a) = e^{-\int_0^a B(s)ds}$$

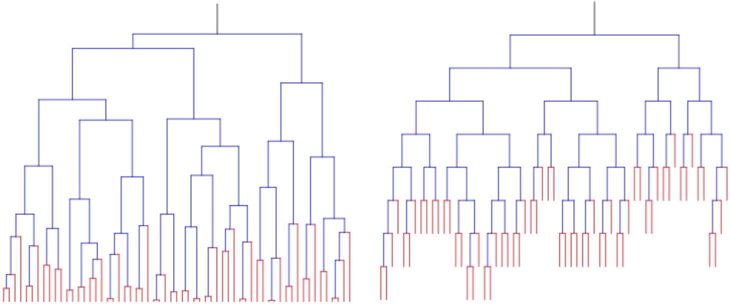


Figure: *Left: the size of each segment represents the lifetime of an individual. Individuals alive at time t are represented in red. Right: genealogical representation of the same realisation of the tree. Figure taken from (Hoffmann, Olivier, 2016).*

Models: From probability to PDE...

Equivalent view: random measures

$X(t) = \big(X_1(t), X_2(t), \dots\big)$ process of the sizes of the population at time t, or $A(t) = \big(A_1(t), A_2(t), \dots\big)$ of ages at time t.
$X(t)$ has values in the space of finite point random measures on $\mathbb{R}_+ \setminus \{0\}$ via

$$Z_t^{(x)} = \sum_{i=1}^{\sharp X(t)} \delta_{X_i(t)}, \qquad Z_t^{(a)} = \sum_{i=1}^{\sharp A(t)} \delta_{A_i(t)}$$

microfluidic / genealogical case: only 1 individual $\delta_{X_1(t)}$

Stochastic evolution equation for the age model

ask Bertrand, Chi, Sylvie, Vincent... or refer to (Bansaye, Méléard, 2015)

$$Z_t^{(k,a)} = \tau_t Z_0 + \int_0^t \sum_{i \leq \langle Z_{s-}^{(k,a)}, \mathbb{1} \rangle} \int_0^\infty \left(k\delta_{t-s} - \delta_{a_i(Z_{s-}^{(k,a)})+t-s}\right) \mathbb{1}_{\{\vartheta \leq B(a_i(Z_{s-}^{(k,a)}))\}} N_i(ds, d\vartheta),$$

$k = 1$: genealogical case / microfluidic device

$k = 2$: population case

Age model: renewal process and renewal equation

$$\mathbb{P}(\zeta_u \in (a, a+da) | \zeta_u \geq a) = B(a)da, \qquad \mathbb{P}(\zeta_u \geq a) = e^{-\int_0^a B(s)ds}$$

Set, for (regular compactly supported) f

$$\langle n(t,\cdot), f\rangle := \mathbb{E}[\langle Z_t^{(k,a)}, f\rangle] = \mathbb{E}\Big[\sum_{i=1}^{\infty} f\big(A_i(t)\big)\Big].$$

In a weak sense:

$$\partial_t n(t,a) + \partial_a n(t,a) = -B(a)n(t,a),$$

$$n(t,0) = 2\int_0^{\infty} B(a)n(t,a)da \quad OR \quad n(t,0) = \int_0^{\infty} B(a)n(t,a)da$$

So the mean empirical distribution of $A(t)$ satisfies the deterministic renewal equation.

Size model: growth-fragmentation process or equation

$$\mathbb{P}(\zeta_u \geq a | \xi_u = x) = e^{-\int_0^a B(xe^{\kappa s})ds}$$

Set, for (regular compactly supported) f

$$\langle n(t,\cdot), f\rangle := \mathbb{E}\left[\sum_{i=1}^{\infty} f\big(X_i(t)\big)\right].$$

Proof: tagged fragment approach (Bertoin, Haas, ...), many-to-one formula (Bansaye et al, 2009, Cloez, 2011, Bertoin & Watson, 2019...)

We have (in a weak sense) IF we keep the 2 daughters at each generation:

$$\boxed{\partial_t n(t,x) + \partial_x\big(\kappa x\, n(t,x)\big) + B(x)n(t,x) = 4B(2x)n(t,2x).}$$

So the mean empirical distribution of $X(t)$ satisfies the deterministic growth-fragmentation / size-structured / cell division equation (with binary fission and equal mitosis).

Size model: growth-fragmentation process or equation

$$\mathbb{P}(\zeta_u \geq a | \xi_u = x) = e^{-\int_0^a B(xe^{\kappa s})ds}$$

Set, for (regular compactly supported) f

$$\langle n(t,\cdot), f \rangle := \mathbb{E}\Big[\sum_{i=1}^{\infty} f\big(X_i(t)\big)\Big].$$

Proof: tagged fragment approach (Bertoin, Haas, ...), many-to-one formula (Bansaye et al, 2009, Cloez, 2011, Bertoin & Watson, 2019...)

We have (in a weak sense) IF we keep 1 daughter at each generation:

$$\boxed{\partial_t n(t,x) + \partial_x\big(\kappa x\, n(t,x)\big) + B(x)n(t,x) = 2B(2x)n(t,2x).}$$

So the mean empirical distribution of $X(t)$ satisfies a deterministic conservative growth-fragmentation equation (also encountered e.g. for TCP/IP protocol)

Age and Size model: PDE

$n(t, a, x)$ density of cells of size x and age a.
PDE obtained from the PDMP (as previously) or by a mass balance:

$$\frac{\partial}{\partial t} n + \frac{\partial}{\partial a} n + \frac{\partial}{\partial x}\big(\kappa x n\big) = -B(a, x) n(t, a, x),$$

$$n(t, a = 0, x) = 4 \int_0^\infty B(a, 2x) n(t, a, 2x) da$$

with $n(0, a, x) = n^{(0)}(a, x)$, $x \geq 0$.
IF $B = B(x)$: back to growth-fragmentation equation
IF $B = B(a)$: back to renewal equation
IF we keep only 1 daughter at each generation:

$$n(t, a = 0, x) = 2 \int_0^\infty B(a, 2x) n(t, a, 2x) da$$

A Specific Age and Size model: the "adder model"

(Taheri et al., Cell, 2015; A. Amir, PRL, 2014; Hall, Wake & Gandar, JMB, 1991)

$n(t, a, x)$ density of cells of size x and increment a.

Definition of an increment: difference between size and size at birth

A Specific Age and Size model: the "adder model"

(Taheri et al., Cell, 2015; A. Amir, PRL, 2014; Hall, Wake & Gandar, JMB, 1991)

$n(t, a, x)$ density of cells of size x and increment a.
Definition of an increment: difference between size and size at birth

PDE obtained from the PDMP (as previously): same as the age process:

$$\mathbb{P}(\zeta_u \geq a) = e^{-\int\limits_0^a B(s)ds},$$

A Specific Age and Size model: the "adder model"

(Taheri et al., Cell, 2015; A. Amir, PRL, 2014; Hall, Wake & Gandar, JMB, 1991)

$n(t, a, x)$ density of cells of size x and increment a.

Definition of an increment: difference between size and size at birth

PDE obtained from the PDMP (as previously): same as the age process:

$$\mathbb{P}(\zeta_u \geq a) = e^{-\int_0^a B(s)ds}, \qquad \frac{da}{dt} = \kappa x$$

A Specific Age and Size model: the "adder model"

(Taheri et al., Cell, 2015; A. Amir, PRL, 2014; Hall, Wake & Gandar, JMB, 1991)

$n(t,a,x)$ density of cells of size x and increment a.

Definition of an increment: difference between size and size at birth

PDE obtained from the PDMP (as previously): same as the age process:

$$\mathbb{P}(\zeta_u \geq a) = e^{-\int_0^a B(s)ds}, \qquad \frac{da}{dt} = \kappa x$$

$$\frac{\partial}{\partial t} n + \frac{\partial}{\partial a}(\kappa x n) + \frac{\partial}{\partial x}(\kappa x n) = -\kappa x B(a) n(t,a,x),$$

$$n(t, a=0, x) = 8 \int_0^\infty x B(a, 2x) n(t,a,2x) da$$

IF we keep only 1 daughter at each generation:

$$n(t, a=0, x) = 4 \int_0^\infty x B(a) n(t,a,2x) da$$

Fourth step: model analysis: long-time behaviour

The age model

A very pedagogical reference: B. Perthame, *Transport Equations in Biology*, 2007

historically the first structured-population model to be studied (Kermack and Mc Kendrick, 1927 ; Metz and Diekmann, 1981)
$n(t,a)e^{-\lambda t} \to N(a)$, with λ and N uniquely determined by

$$\frac{\partial}{\partial a}N + \lambda N = -B(a)N, \qquad N(0) = 2\int_0^\infty B(a)N(a)da.$$

Explicit solution: $N(a) = N(0)e^{-\lambda a - \int_0^a B(s)ds}$,
λ uniquely determined by the boundary condition:
either $\lambda = 0$ (1 branch case) or

$$2\int_0^\infty B(a)e^{-\lambda a - \int_0^s B(s)ds} da = 1$$

The fragmentation and growth-fragmentation equations
General form

From a stochastic viewpoint:

$$\frac{\partial}{\partial t} n(t,dx) + \frac{\partial}{\partial x}\big(\tau(x) n(t,dx)\big) =$$

$$-B(x) n(t,dx) + \sum_{j\geq 0} j p(j) \int\limits_{y=x}^{\infty} P^{(j)}(y,dx) B(y) n(t,dy),$$

in a weak sense (for measure solutions: see e.g. (Canizo, Carrillo, Cuadrado, 2013); (MD, Gwiazda, Wiedemann, 2018))

$P^{(j)}(y,dx)$: probability of an individual of size y to split in j parts, one of them of size in the interval dx. In a more compact way:

$$k(y,dx) := \sum_{j\geq 0} j p(j) P^{(j)}(y,dx), \qquad \text{with}$$

$$\int\limits_{x=0}^{y} x k(y,dx) = \sum_{j\geq 0} p(j) \int\limits_{0}^{y} j x P^{(j)}(y,dx) = y \sum_{j\geq 0} p(j) = y.$$

The fragmentation and growth-fragmentation equations General form

$$\frac{\partial}{\partial t} n(t,dx) + \frac{\partial}{\partial x}\big(\tau(x) n(t,dx)\big) =$$

$$-B(x)n(t,dx) + \int_{y=x}^{\infty} k(y,dx)B(y)n(t,dy),$$

with

$$\int_0^y x k(y,dx) = y, \qquad \int_0^y k(y,dx) = m > 1.$$

"One branch" process: $k_1(y,dx) := \sum_{j\geq 0} p(j) P^{(j)}(y,dx)$:

$$\frac{\partial}{\partial t} n_1(t,dx) + \frac{\partial}{\partial x}\big(\tau(x) n_1(t,dx)\big) =$$

$$-B(x)n_1(t,dx) + \int_{y=x}^{\infty} k_1(y,dx)B(y)n_1(t,dy).$$

The growth-fragmentation equation

Two fundamental relations
(and more generally: moments equations)

- First moment: mass balance only evolves by growth

$$\frac{d}{dt}\int x n(t,x)dx = \int \tau(x) n(t,x)dx.$$

- Zeroth moment: number of individuals only evolves by fragmentation:

$$\frac{d}{dt}\int n(t,x)dx = \int B(x)\Big(\int_0^x k(x,dy) - 1\Big) n(t,x)dx.$$

The growth-fragmentation equation

Two fundamental relations
(and more generally: moments equations)

- First moment: mass balance only evolves by growth

$$\frac{d}{dt}\int xn(t,x)dx = \int \tau(x)n(t,x)dx.$$

- Zeroth moment: number of individuals only evolves by fragmentation:

$$\frac{d}{dt}\int n(t,x)dx = \int B(x)\Big(\int_0^x k(x,dy) - 1\Big)n(t,x)dx.$$

- More generally: balance between growth & fragmentation

$$\tfrac{d}{dt}\int_0^\infty x^p n(t,x)dx = \int_0^\infty px^{p-1}\tau(x)n(t,x)dx$$

$$+\int_0^\infty B(x)x^p\Big(1-\int_0^x \tfrac{y^p}{x^p}k(x,dy)\Big)n(t,x)dx$$

Asymptotic behaviour 1: balance assumption on $\tau(x)$ and $B(x)$:
$\Rightarrow$ convergence to a steady profile + exponential growth
starts in the 1980s (Diekmann, Heijmans, Thieme and Gyllenberg & Webb)

$$n(t,x)e^{-\lambda t} \to N(x) \int n^0(x)dx$$

(N, λ) : dominant eigenpair of the semi-group generator $L^* + \mathcal{F}^*$.

For compact strictly positive operators: Krein-Rutman.

Stochastic approaches: for recent ref. see (Bertoin& Watson, 2018); (B. Cavalli, 2019); (Bansaye, Cloez, Gabriel, Marguet, 2021); (Champagnat, Villemonais, 2018)...

Long-time asymptotics 1: steady growth

Eigenvalue problem and adjoint problem:

$$\begin{cases} \frac{\partial}{\partial x}(\tau(x)N(x)) + \lambda N(x) = -B(x)N(x) + \int_x^\infty B(y)k(x,y)N(y)dy, \\ \tau N(x=0) = 0, \qquad N(x) \geq 0, \qquad \int_0^\infty N(x)dx = 1, \\ -\tau(x)\frac{\partial}{\partial x}(\phi(x)) + \lambda\phi(x) = B(x)(-\phi(x) + \int_0^x k(y,x)\phi(y)dy), \\ \phi(x) \geq 0, \qquad \int_0^\infty \phi(x)N(x)dx = 1. \end{cases} \tag{1}$$

If $\tau(x) = x^\nu$, $B(x) = x^\gamma$: if $1 + \gamma - \nu > 0$ (Michel, M3AS, 2004)

which optimal assumptions on (τ, k, B) ?

Long-time asymptotics

Theorem (MD, P. Gabriel, M3AS, 2010)

Under balance assumptions on τ, B *and* k, *there exists a unique triplet* (λ, N, ϕ) *with* $\lambda > 0$, *solution of the eigenproblem* (5) *and*

$$x^\alpha \tau N \in L^p(\mathbb{R}^+), \quad \forall \alpha \geq -\gamma, \quad \forall p \in [1, \infty], \quad x^\alpha \tau N \in W^{1,1}(\mathbb{R}^+),$$

$$\exists p > 0 \ s.t. \ \frac{\phi}{1 + x^p} \in L^\infty(\mathbb{R}^+), \quad \tau \frac{\partial}{\partial x} \phi \in L^\infty_{loc}(\mathbb{R}^+).$$

Generalizes previous results by Michel, M3AS, 2004.

$$\boxed{\int_{\mathbb{R}_+} \left| n(t,x) e^{-\lambda t} - \langle n^{(0)}, \phi \rangle N(x) \right| \phi(x) dx \to 0 \ \text{ as } \ t \to \infty}$$

Proof: General Relative Entropy (Michel, Mischler, Perthame, 2004)
See also many recent improvements...

Some ideas on the proof

2 opposite dynamics:

- Growth $\Rightarrow$ bigger and bigger $\Rightarrow$ mass goes to infinity ?
- Fragmentation $\Rightarrow$ smaller and smaller $\Rightarrow$ dust formation ?

Some ideas on the proof

2 opposite dynamics:

- Growth $\Rightarrow$ bigger and bigger $\Rightarrow$ mass goes to infinity ?
- Fragmentation $\Rightarrow$ smaller and smaller $\Rightarrow$ dust formation ?

Balance: asymptotic steady profile.

- Enough growth at zero: $\frac{B(x)}{\tau(x)} \in L^1_0$
- avoid *shattering* (0-size polymers)

$$\exists\, C > 0, \gamma \geq 0 \quad s.t. \qquad \int_0^x k(y, dz) \leq \min\Big(m, C\Big(\frac{x}{y}\Big)^\gamma\Big)$$

  and $\frac{x^\gamma}{\tau(x)} \in L^1_0$
- Enough fragmentation at infinity: $\frac{xB(x)}{\tau(x)} \to_{x\to\infty} \infty$

Some ideas on the proof

2 opposite dynamics:

- Growth $\Rightarrow$ bigger and bigger $\Rightarrow$ mass goes to infinity ?
- Fragmentation $\Rightarrow$ smaller and smaller $\Rightarrow$ dust formation ?

Balance: asymptotic steady profile.

- Enough growth at zero: $\frac{B(x)}{\tau(x)} \in L^1_0$
- avoid *shattering* (0-size polymers)

$$\exists\, C > 0, \gamma \geq 0 \quad s.t. \qquad \int_0^x k(y, dz) \leq \min\Big(m, C\Big(\frac{x}{y}\Big)^\gamma\Big)$$

  and $\frac{x^\gamma}{\tau(x)} \in L^1_0$
- Enough fragmentation at infinity: $\frac{xB(x)}{\tau(x)} \to_{x\to\infty} \infty$

Proof:

- regularized equation: Krein-Rutman/Perron-Frobenius
- balance assumptions$\Rightarrow$ compactness through successive moments estimates
- uniqueness and convergence by entropy method

Long-time asymptotics 1
Further comments on the "steady growth regime"

- Under extra assumptions, exponential convergence in some sense: (Laurençot, Perthame, 2009) (Balagué, Cañizo, Gabriel, 2012) (Bernard, Gabriel, 2019) (Càceres, Cañizo, Mischler, 2011)

- (Mischler, Scher, 2015): spectral gap for a large class
  for a more restrictive norm $L^1_\psi \subsetneq L^1_\phi$
  Based on semi-group spectral analysis & a generalization of Krein-Rutman theorem
  Proof of no spectral gap in L^1_ϕ (Bernard & Gabriel, 2017, & 2019)
  Measure solutions (MD, Gwiazda, Wiedemann, 2018; Bansaye, Cloez, Gabriel, Marguet, preprint, 2021)
- Age-size models: (MD, 2007), increment (Gabriel & Martin, 2019)

Other types of behaviours?

Growth-fragmentation eq., $\nu = 1$, $k_0(dz) = 2\delta_{z=\frac{1}{2}}$

$$
\begin{cases}
\dfrac{\partial}{\partial t} n(t,x) + \dfrac{\partial}{\partial x}\big(x n(t,x)\big) + B(x) n(t,x) = 4B(2x) n(t,2x), \ x > 0, \\[2ex]
n(0,x) = n_0(x).
\end{cases}
\tag{2}
$$

Growth-fragmentation eq., $\nu = 1$, $k_0(dz) = 2\delta_{z=\frac{1}{2}}$

$$\begin{cases} \frac{\partial}{\partial t} n(t,x) + \frac{\partial}{\partial x}\big(x n(t,x)\big) + B(x) n(t,x) = 4B(2x) n(t,2x), \ x > 0, \\ \\ n(0,x) = n_0(x). \end{cases} \tag{2}$$

Same case but $g(x) \equiv 1$: (Perthame, Ryzhik, 2004, +...)

$$n(t,x)e^{-\lambda t} \to N(x)$$

Growth-fragmentation eq., $\nu = 1$, $k_0(dz) = 2\delta_{z=\frac{1}{2}}$

$$\begin{cases} \frac{\partial}{\partial t} n(t,x) + \frac{\partial}{\partial x}\big(x n(t,x)\big) + B(x)n(t,x) = 4B(2x)n(t,2x), \ x > 0, \\ \\ n(0,x) = n_0(x). \end{cases} \tag{2}$$

Same case but $g(x) \equiv 1$: (Perthame, Ryzhik, 2004, +...)

$$n(t,x)e^{-\lambda t} \to N(x)$$ fails here

Growth-fragmentation eq., $\nu = 1$, $k_0(dz) = 2\delta_{z=\frac{1}{2}}$

$$\begin{cases} \frac{\partial}{\partial t} n(t,x) + \frac{\partial}{\partial x}\big(xn(t,x)\big) + B(x)n(t,x) = 4B(2x)n(t,2x), \ x > 0, \\ n(0,x) = n_0(x). \end{cases} \tag{2}$$

Same case but $g(x) \equiv 1$: (Perthame, Ryzhik, 2004, +...)

$$n(t,x)e^{-\lambda t} \to N(x) \qquad \text{fails here}$$

Intuition: stochastic process: if $X(t) = x_0$, all descendants live on the countable set of curves $x_0 e^t 2^{-n}$

Where usual proofs (eigenproblem, entropy) fail?
semi-groups on compact support: abstract result (Greiner, Nagel, 1988)

Eigenproblem

$$\lambda N(x) + \big(xN(x)\big)' + B(x)N(x) = 4B(2x)N(2x),$$
$$\lambda\phi(x) - x\phi'(x) + B(x)\phi(x) = 2B(x)\phi\left(\tfrac{x}{2}\right). \tag{3}$$

Eigenproblem

$$\lambda N(x) + \big(xN(x)\big)' + B(x)N(x) = 4B(2x)N(2x),$$
$$\lambda \phi(x) - x\phi'(x) + B(x)\phi(x) = 2B(x)\phi\Big(\frac{x}{2}\Big). \tag{3}$$

Assumption on B:

$$\begin{cases} B:\ (0,\infty) \to (0,\infty) \text{ is measurable}, B(x)/x \in L^1_{loc}(\mathbb{R}_+), \\ \exists \gamma_0, \gamma_1, K_0, K_1, x_0 > 0, \quad K_0 x^{\gamma_0} \le B(x \ge x_0) \le K_1 x^{\gamma_1}. \end{cases} \tag{4}$$

Eigenproblem

$$\lambda N(x) + \big(xN(x)\big)' + B(x)N(x) = 4B(2x)N(2x),$$
$$\lambda \phi(x) - x\phi'(x) + B(x)\phi(x) = 2B(x)\phi\Big(\frac{x}{2}\Big). \tag{3}$$

Assumption on B:

$$\begin{cases} B:\ (0,\infty) \to (0,\infty) \text{ is measurable}, B(x)/x \in L^1_{loc}(\mathbb{R}_+), \\ \exists \gamma_0, \gamma_1, K_0, K_1, x_0 > 0, \quad K_0 x^{\gamma_0} \le B(x \ge x_0) \le K_1 x^{\gamma_1}. \end{cases} \tag{4}$$

Theorem (from MD, P. Gabriel, M3AS, 2010)

Under standard assumptions, $\exists!$ *positive eigentriplet* $\lambda = 1$, $N \in L^1(\mathbb{R}_+)$, $\phi(x) = x$, *with* $\int_0^\infty xN(x)dx = 1$.

Eigenproblem

$$
\begin{aligned}
&\lambda N(x) + \big(xN(x)\big)' + B(x)N(x) = 4B(2x)N(2x), \\
&\lambda \phi(x) - x\phi'(x) + B(x)\phi(x) = 2B(x)\phi\Big(\frac{x}{2}\Big).
\end{aligned}
\tag{3}
$$

Assumption on B:

$$
\begin{cases}
B:\ (0,\infty) \to (0,\infty) \text{ is measurable}, B(x)/x \in L^1_{loc}(\mathbb{R}_+), \\
\exists \gamma_0, \gamma_1, K_0, K_1, x_0 > 0, \quad K_0 x^{\gamma_0} \le B(x \ge x_0) \le K_1 x^{\gamma_1}.
\end{cases}
\tag{4}
$$

Theorem (from MD, P. Gabriel, M3AS, 2010)

Under standard assumptions, $\exists!$ *positive eigentriplet* $\lambda = 1$, $N \in L^1(\mathbb{R}_+)$, $\phi(x) = x$, *with* $\int_0^\infty xN(x)dx = 1$.

$$
\lambda_k = 1 + \frac{2ik\pi}{\log 2}, \qquad N_k(x) = x^{-\frac{2ik\pi}{\log 2}} N(x), \qquad \phi_k(x) = x^{1+\frac{2ik\pi}{\log 2}},
$$

Balance laws and Entropy

$$\forall\ k \in \mathbb{Z}, \text{ and } \forall (k,l) \in \mathbb{Z}^2, \qquad \int_0^\infty N_k(x)\phi_l(x)dx = \delta_{kl}.$$

$$\forall k \in \mathbb{Z},\ \forall t \geq 0, \qquad \int_0^\infty n(t,x), \phi_k(x)dx e^{-\lambda_k t} = \int_0^\infty n_0(x)\phi_k(x)dx.$$

Lemma (General Relative Entropy Inequality)

$n(t,x)$ sol. of (2), $H : \mathbb{C} \to \mathbb{R}_+$ positive, differentiable & convex.

$$\frac{d}{dt}\int_0^\infty x\, N(x) H\Big(\frac{n(t,x)}{N(x)e^t}\Big)dx = -D^H[n(t)e^{-t}] \leq 0,$$

$$\text{with } D^H[n] := \int_0^\infty xB(x)\, N(x)\Big[H\Big(\frac{n(\frac{x}{2})}{N(\frac{x}{2})}\Big) - H\Big(\frac{n(x)}{N(x)}\Big) - \nabla H\Big(\frac{u(\frac{x}{2})}{N(\frac{x}{2})}\Big)\cdot\Big(\frac{n(\frac{x}{2})}{N(\frac{x}{2})} - \frac{n(x)}{N(x)}\Big)\Big]dx.$$

Dissipation of entropy

For H strictly convex, $n : \mathbb{R}_+ \to \mathbb{C}$ satisfies $D^H[u] = 0$ *iff*

$$\frac{n(x)}{N(x)} = \frac{n(2x)}{N(2x)}, \qquad \text{a.e.} \quad x > 0.$$

In particular, for all $k \in \mathbb{Z}$, $D^H[N_k] = 0$.
(Escobedo, Mischler, Rodriguez Ricard, 2004), lemma 3.5 fails.

Dissipation of entropy

For H strictly convex, $n : \mathbb{R}_+ \to \mathbb{C}$ satisfies $D^H[u] = 0$ *iff*

$$\frac{n(x)}{N(x)} = \frac{n(2x)}{N(2x)}, \qquad \text{a.e.} \quad x > 0.$$

In particular, for all $k \in \mathbb{Z}$, $D^H[N_k] = 0$.
(Escobedo, Mischler, Rodriguez Ricard, 2004), lemma 3.5 fails.

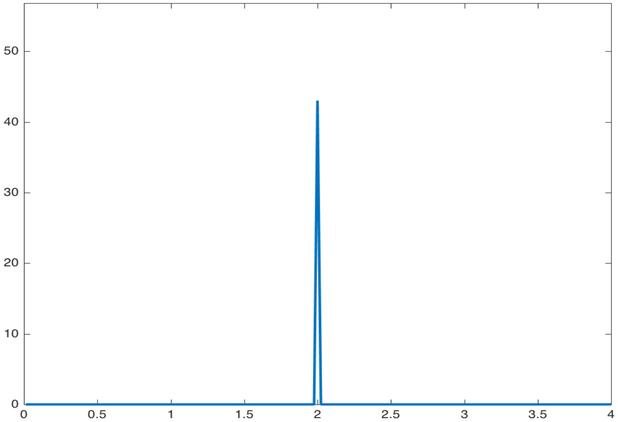
Theorem (E. Bernard, MD, P. Gabriel, Kin. Rel. Mod., accepted)

Under Hyp. (4), *for any* $n_0 \in L^2(\mathbb{R}_+, x/N(x)dx)$, *the unique solution* $n(t,x) \in C\big(\mathbb{R}_+, L^2(\mathbb{R}_+, x/N(x)dx)\big)$ *to* (2) *satisfies*

$$\int_0^\infty \left| n(t,x)e^{-t} - \sum_{k=-\infty}^{+\infty} (n_0, N_k)\, N_k(x) e^{\frac{2ik\pi}{\log 2} t} \right|^2 \frac{x\,dx}{N(x)} \xrightarrow[t\to+\infty]{} 0,$$

with $(n_0, N_k) = \int n_0 \phi_k(x) dx$

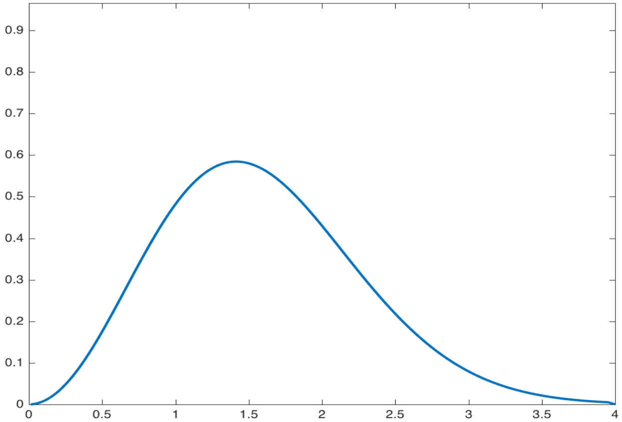
Numerical illustration



Non dissipative scheme:

- splitting transport & fragmentation
- grid $x_k = (1 + 2^{\frac{1}{n}})^{k-N}$

Numerical illustration



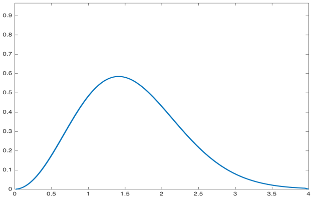
Non dissipative scheme:

- splitting transport & fragmentation
- grid $x_k = (1 + 2^{\frac{1}{n}})^{k-N}$

The case $\tau(x) = \kappa x$

If $B(x) = x^\gamma$:

- $\gamma > 0$: in general, convergence (at an exponential speed) given by $n(t,x)e^{-\lambda t} \to N(x)$
- $\gamma > 0$ and $k(y,x) = \delta_{x=\frac{y}{2}}$ (our "idealised" case!): convergence to an oscillatory profile (Bernard, MD, Gabriel, 2018), (Martin & Gabriel, 2021) remains true for any model where growth is exponential and division in two equally-sized daughters



Intuition: depart from a cell of size x_0, at time t all its descendants live on $x_0 e^{\kappa t} 2^{-\mathbb{N}}$

The pure fragmentation case: $\tau = 0$

Classical assumptions on the fragmentation equation

- $B(x) = \alpha x^\gamma$
- $k(y,x) = \frac{1}{y} k_0\left(\frac{x}{y}\right)$, where k_0 is a measure on $[0,1]$.

$$\frac{\partial}{\partial t} u(t,x) + x^\gamma u(t,x) = \int_0^1 (\frac{x}{z})^\gamma u(t,\frac{x}{z}) \frac{k_0(dz)}{z}$$

For $\gamma > 0$, at a power law speed, we have (Escobedo-Mischler-Ricard, 2005)

$$\lim_{t\to\infty} \int_0^\infty \left| u(t,y) - t^{-\frac{2}{\gamma}} g\left(t^{\frac{1}{\gamma}} y\right)\right| y dy = 0.$$

where g called the "self-similar profile" is the unique solution of

$$\frac{\partial}{\partial z}(zg(z)) + (1+\alpha\gamma z^\gamma) g(z) = \alpha\gamma \int_z^\infty \frac{1}{y} k_0(\frac{z}{y}) y^\gamma g(y) dy, \quad \int_0^\infty z g(z) dz = \rho.$$

The fragmentation equation

Focus: $\tau(x) \equiv 0$, $B(x) \equiv x^\gamma$: e.g. protein fibril fragmentation

$$\frac{\partial}{\partial t} u(t,x) + x^\gamma u(t,x) = \int_0^1 (\frac{x}{z})^\gamma u(t, \frac{x}{z}) \frac{k_0(dz)}{z}$$

- $\gamma > 0$: self-similar profile (Escobedo, Mischler, Ricard, 2004)

$$\lim_{t\to\infty} \int_0^\infty \left| u(t,y) - t^{-\frac{2}{\gamma}} g\left(t^{\frac{1}{\gamma}} y\right)\right| ydy = 0.$$

- $\gamma < 0$: shattering: loss of mass + self-similar profile or steady profile according to the initial condition (Haas, 2010, Bertoin & Watson 2017 & 2018, Escobedo 2017...)

The fragmentation equation

Focus: $\tau(x) \equiv 0$, $B(x) \equiv x^\gamma$: e.g. protein fibril fragmentation

$$\frac{\partial}{\partial t} u(t,x) + x^\gamma u(t,x) = \int_0^1 (\frac{x}{z})^\gamma u(t, \frac{x}{z}) \frac{k_0(dz)}{z}$$

- $\gamma > 0$: self-similar profile (Escobedo, Mischler, Ricard, 2004)

$$\lim_{t\to\infty} \int_0^\infty \left| u(t,y) - t^{-\frac{2}{\gamma}} g\left(t^{\frac{1}{\gamma}} y\right)\right| y dy = 0.$$

- $\gamma < 0$: shattering: loss of mass + self-similar profile or steady profile according to the initial condition (Haas, 2010, Bertoin & Watson 2017 & 2018, Escobedo 2017...)
- $\gamma = 0$: critical case. Close to a mutation model (G. Garnier's PhD) (Bertoin 2003, MD Escobedo 2016, Bertoin & Watson 2016)

Fifth step: model calibration

Model calibration for the bacteria case

Only unobserved parameter: the division rate B.
Estimation procedure:

- mathematical analysis: asymptotic regime (PDMP or PDE)
- estimation methods
- comparison of calibrated model results and data

Use of the long-time asymptotics

Example: PDE - Size model asymptotics

Recall: if $B(x) = x\beta(x)$ such that $\beta \in L^1_0$ and $\beta \rightarrow_{x\rightarrow\infty} \infty$,
$\exists!$ $(\lambda > 0, N \geq 0)$ solution of

$$\begin{cases} \frac{\partial}{\partial x}(\kappa x N(x)) + \lambda N(x) = -B(x)N(x) + 4B(2x)N(2x)dx, \\ \\ N(x) \geq 0, \qquad \int_0^\infty N(x)dx = 1. \end{cases} \tag{5}$$

Moreover here $\kappa = \lambda$ and

$$\boxed{\int_{\mathbb{R}_+} \left| n(t,x)e^{-\lambda t} - \langle n^{(0)}, x\rangle N(x) \right| x dx \rightarrow 0 \ \text{ as } \ t \rightarrow \infty}$$

false here (oscillations) but true in practice: experimental variability

Estimation methods

3 methods:

- use the "all cells" distributions: "indirect/inverse" approach, based on $N(x)$ or $N(a)$

- use the "at division" distributions: "direct" approach: PDMP or $B(x)N(x)/\int BNdx$

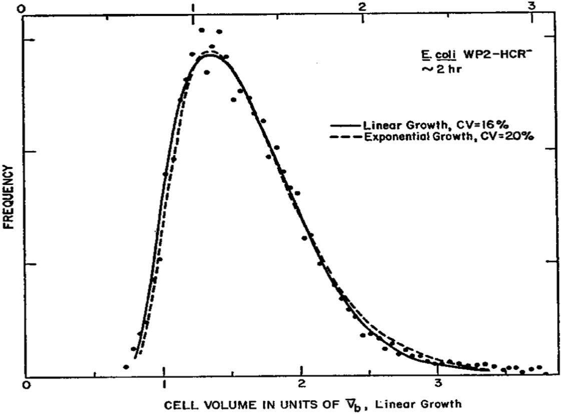
- use both ! "direct" approach: measure of both $B(x)N(x)/\int BNdx$, and $N(x)$

With E. coli: choose any of the 3 schemes and select the most accurate

Preliminaries: How to estimate these densities?

First method, preliminaries: estimation of $N(x)$

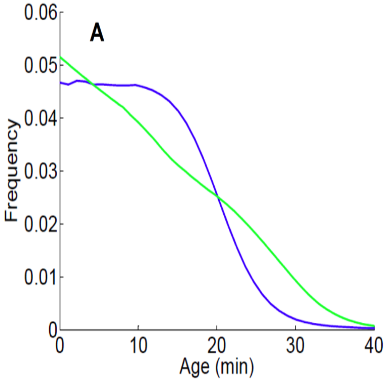
1st historical observations, the simplest and often the only possible ones, and confirm the asymptotic behavior:



Observation (from Kubitschek, 1969): doubling time and steady size distribution

First method: an indirect approach

Any cell at any time put together in this asymptotic distribution



cf. video at the beginning: around 30.000 to 60.000 observations
(Blue: 1 branch, Green: whole tree)

Inverse Problem for the age model

From a (noisy) measure of $N(a)$ and λ, we look for $B(a)$.
Since we have the explicit relation

$$N(a) = N(0)e^{-\lambda a - \int_0^a B(s)ds},$$

we get

$$B(a) = -\lambda - \frac{\partial_a N(a)}{N(a)}.$$

From a noisy version of N: regularization is needed:
"degree of ill-posedness"=1: if N is in H^s_{loc}, B is in H^{s-1}_{loc}

Inverse Problem for the size model

Inverse Problem: estimating the division rate $B(x)$

From: measurements of (κ, N) with

$$\frac{\partial}{\partial x}(\kappa x N(x)) + \lambda N(x) = -B(x)N(x) + 4B(2x)N(2x)dx.$$

Choice of a **Hilbert space:** $L^2(\mathbb{R}_+, x^p dx)$
(Engl, Hanke, Neubauer, *Regularization of Inverse Problems*, 1995)

Similar to the age problem: the equation implies a derivative for N

Inverse Problem for the Size Model

Estimate B through

$$L(N) = G(BN), \qquad \text{with}$$

$$G(f)(x) = 4f(2x) - f(x), \tag{6}$$

$$L(N)(x) = \kappa \partial_x \big(xN(x)\big) + \kappa N(x), \tag{7}$$

2 main steps:

- Solve $G(f) = L$ for f, L in suitable weighted L^2 spaces:
  PDE part. the problem $N \to f = BN$ is now linear.
- Find an estimate for $L(N)$ in this L^2 space:
  PDE or statistical part

Inverse Problem for the Size Model

Step 1: solve a dilation equation

Defining

$$G : f \to G(f) = 4f(2x) - f(x)$$

We want to invert G in a weighted L^2 space: knowing $L \in L^2$, find $f \in L^2$ solution of

$$L(x) = 4f(2x) - f(x) \tag{8}$$

Inverse Problem for the Size Model

Step 1: solve a dilation equation

Defining

$$G : f \to G(f) = 4f(2x) - f(x)$$

We want to invert G in a weighted L^2 space: knowing $L \in L^2$, find $f \in L^2$ solution of

$$L(x) = 4f(2x) - f(x) \qquad (8)$$

Proposition (MD, Perthame, Zubelli, 2009)

$\forall\ L \in L^2(x^p dx)$, $p \neq 3$, there exists a unique solution $f \in L^2(x^p dx)$ to (8). Moreover, defining

$$H_0 := \sum_{j=1}^{\infty} 2^{-2j} L(2^- jx), \qquad H_\infty := -\sum_{j=0}^{\infty} 2^{2j} L(2^j x),$$

we have $f = H_0$ if $p < 3$ and $f = H_\infty$ if $p > 3$. Moreover if $L \in L^q$ then $H_0 \in L^q$ for any $1 \leq q \leq \infty$. For $L = 0$, any distribution of the form $f(\frac{\log x}{x^2})$ with $f \in \mathcal{D}'(\mathbb{R}_+)$ $\log -2$ periodic is solution.

Inverse Problem for the Size Model

Step 1: solve a dilation equation for self-similar kernels
(Bourgeron, MD, Escobedo, Inv. Prob., 2014)

$G(f)$ becomes in the case of a self-similar fragmentation kernel:

$$G : g \to G(f), \quad G(f)(x) := \int_x^\infty k_0(\frac{x}{y}) f(y) \frac{dy}{y} - f(x),$$

Mellin transform: "Multiplicative Fourier transform on $\mathbb{R}_+$":
$\mathcal{M}$ isometry between $L^2(x^q dx)$ and $L^2(\frac{q+1}{2} + i\,\mathbb{R})$ defined by

$$\mathcal{M}[f](s) := \int_0^\infty x^{s-1} f(x) dx, \ \mathcal{M}_q^{-1}[F](x) := \int_{-\infty}^\infty x^{-\frac{q+1}{2} - iv} F(\frac{q+1}{2} + iv) dv$$

$$\mathcal{M}[\mathcal{G}(f)](s) = (\mathcal{M}[k_0](s) - 1)\mathcal{M}[f](s)$$

Zeros of $\mathcal{M}_{k_0}(s) - 1$: at least for $s = 2$, since $\int_0^1 x k_0(x) dx = 1$. So for $q \neq 3$:

$$H_q := \mathcal{M}_q^{-1}\left[\frac{\mathcal{M}[\mathcal{G}(f)](s)}{\mathcal{M}[k_0](s) - 1}\right]$$

Estimating B with the Mellin transform

(Bourgeron, MD, Escobedo, Inv. Prob. 2014)

We measure N with a noise:

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

Theory of linear inverse problems: by the optimal regularisation method of your choice, of parameter $\alpha > 0$, define an approximation $L(N_\varepsilon)_\alpha$ such that, for $N \in H^m(\mathbb{R}_+)$, and $q > 3$, we have

$$\|L(N_\varepsilon)_\alpha - L(N)\|_{L^2((1+x^q)dx)} \leq C(\frac{\varepsilon}{\alpha} + \alpha^m),$$

Estimating B with the Mellin transform

(Bourgeron, MD, Escobedo, Inv. Prob. 2014)

We measure N with a noise:

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

Theory of linear inverse problems: by the optimal regularisation method of your choice, of parameter $\alpha > 0$, define an approximation $L(N_\varepsilon)_\alpha$ such that, for $N \in H^m(\mathbb{R}_+)$, and $q > 3$, we have

$$\|L(N_\varepsilon)_\alpha - L(N)\|_{L^2((1+x^q)dx)} \le C(\frac{\varepsilon}{\alpha} + \alpha^m),$$

and since we want $H = BN$ in $L^2((1+x^q)dx)$ with $q > 3$ large, define for some $a > 0$

$$H_{\varepsilon,\alpha} := \mathcal{M}_{0}^{-1}\left[\frac{\mathcal{M}[L(N_\varepsilon)_\alpha](s)}{\mathcal{M}[k_0](s) - 1}\right]\mathbb{1}_{x\le a} + \mathcal{M}_{q}^{-1}\left[\frac{\mathcal{M}[L(N_\varepsilon)_\alpha](s)}{\mathcal{M}[k_0](s) - 1}\right]\mathbb{1}_{x>a}$$

Estimating B with the Mellin transform

(Bourgeron, MD, Escobedo, Inv. Prob. 2014)

Proposition

For $N \in H^s(\mathbb{R}_+)$ solution to the eigenequation we have

$$\|N - N_\varepsilon\|_{L^2(\mathbb{R}_+)} \leq \varepsilon \implies \|H_{\varepsilon,\alpha} - BN\|_{L^2((1+x^q)dx)} \leq C(\frac{\varepsilon}{\alpha} + \alpha^s \|N\|_{H^s})$$

optimal error in $O(\varepsilon^{\frac{s}{s+1}})$

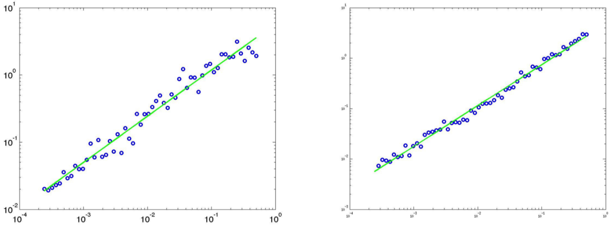


Figure : Rates of convergence for ρ_1 and ρ_2, loglog plot

Numerical slopes: 0.68 (for ρ_1) and 0.76 (for ρ_2)
Theoretical slopes: 0.50 (for ρ_1) and 0.66 (for ρ_2)

Indirect Observation Scheme
Step 2: regularization - **statistical setting**

Joint work with M. Hoffmann, P. Reynaud-Bouret & V. Rivoirard
we have supposed

$$||N - N_\varepsilon||_{L^2} \leq \varepsilon$$

But why an L^2 norm ? What about real data ?

Indirect Observation Scheme
Step 2: regularization - **statistical setting**

Joint work with M. Hoffmann, P. Reynaud-Bouret & V. Rivoirard
we have supposed

$$||N - N_\varepsilon||_{L^2} \leq \varepsilon$$

But why an L^2 norm ? What about real data ?

We observe a sample of n cells, of sizes $x_1, \cdots, x_n$ realizations of $X_1, \ldots, X_n$, *i.i.d.* random variables with density N

$$L_{\alpha,n}(x) := \rho_\alpha * L\left(\frac{1}{n}\sum_{i=1}^{n} \delta_{x=X_i}\right),$$

Inverse Problem for the age model: statistical treatment

We observe a sample of n cells, of ages $a_1, \cdots, a_n$ realizations of $A_1, \ldots, A_n$, *i.i.d.* random variables with density N,
(complete proof of this ansatz: M. Hoffmann, A. Olivier, 2016)
That is, your measure of $N(a)$ is

$$N_{\not{\varepsilon} n}(a) = \frac{1}{n} \sum_{i=1}^{n} \delta_{a=a_i}$$

Regularization: kernel method for instance: mollifier ρ_α

$$N_{n,\alpha}(a) = \rho_\alpha * \Big(\frac{1}{n} \sum_{i=1}^{n} \delta_{a=a_i} \Big)$$

with $\rho_\alpha = \frac{1}{\alpha}\rho(\frac{x}{\alpha})$ with $\rho \in \mathcal{C}_c^\infty(\mathbb{R})$ and $\int \rho(x)dx = 1$, and define

$$B_{n,\alpha}(a) = -\lambda - \frac{\partial_a N_{n,\alpha}(a)}{max(N_{n,\alpha}(a), threshold)}.$$

Indirect Observation Scheme
Step 2: regularization - **statistical setting**

How to adaptively select α ?

Goldenshluger & Lepski, Ann. Statist, 2009; Ann. Probab., 2010

We have a statistical estimator $L_{\alpha,n} = \rho_\alpha * L(\frac{1}{n}\sum \delta_{X_i})$,
we plug the first **PDE** step to inverse G and obtain

Theorem (MD, Hoffmann, Reynaud-Bouret, Rivoirard, 2012)

If $B \in H^s$ $(s > 1/2)$, then (under suitable assumptions)

$$\mathbb{E}\left[\left\|(B_\alpha^n - B)1_{[a,b]}\right\|_2\right] = O\left(n^{-\frac{s}{2s+3}}\right).$$

Indirect Observation Scheme Step 2: regularization - comparison of stat and **deterministic** settings

This optimal rate $n^{-\frac{s}{2s+3}}$ is to be compared with the **deterministic** rate $\varepsilon^{s/(s+1)}$.

see Engl, Hanke, Neubauer, 1995: for linear problems,
if a is the degree of ill-posedness, the optimal rate is $\varepsilon^{\frac{s}{s+a}}$
Here, by the Central Limit and Berry-Essen Theorems, heuristically:

$$\varepsilon \approx n^{-1/2}$$

Degree of ill-posedness: $a = 1$ for a noise in L^2, gives $\varepsilon^{\frac{s}{s+1}}$
Degree $a = 1 + 1/2$ for a noise in $H^{-1/2}$, gives $\varepsilon^{\frac{s}{s+3/2}} = n^{-\frac{s}{2s+3}}$

Indirect Observation Scheme
Step 2: regularization - comparison of stat and **deterministic** settings

This optimal rate $n^{-\frac{s}{2s+3}}$ is to be compared with the **deterministic** rate $\varepsilon^{s/(s+1)}$.

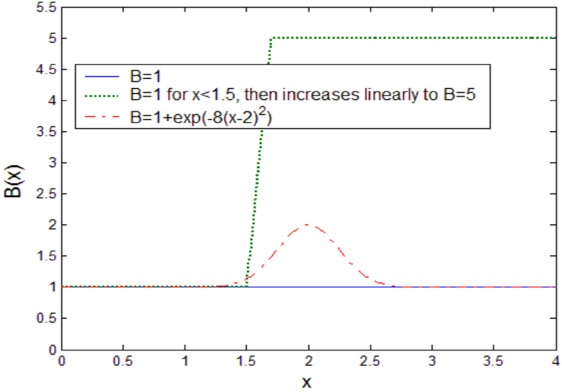
see Engl, Hanke, Neubauer, 1995: for linear problems,
if a is the degree of ill-posedness, the optimal rate is $\varepsilon^{\frac{s}{s+a}}$
Here, by the Central Limit and Berry-Essen Theorems, heuristically:

$$\varepsilon \approx n^{-1/2}$$

Degree of ill-posedness: $a = 1$ for a noise in L^2, gives $\varepsilon^{\frac{s}{s+1}}$
Degree $a = 1 + 1/2$ for a noise in $H^{-1/2}$, gives $\varepsilon^{\frac{s}{s+3/2}} = n^{-\frac{s}{2s+3}}$

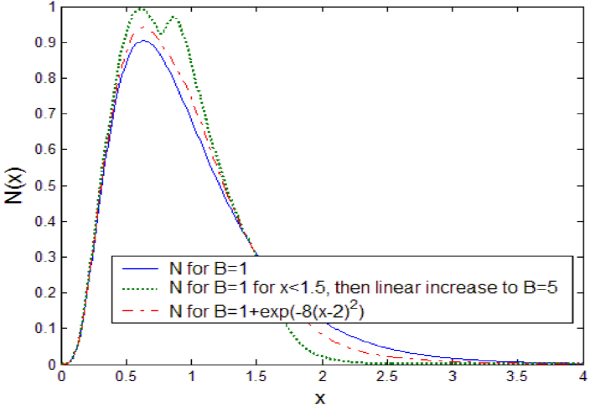
Coherence of the PDE and stat. settings

Numerical Results - Size Structured



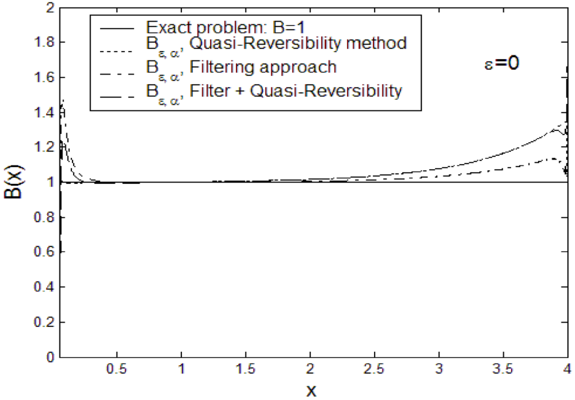
Three tested division rates B

Numerical Results - Size Structured



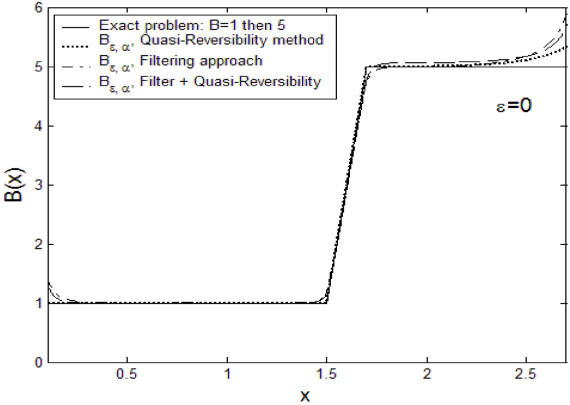
Three related asymptotic distributions N

Numerical Results - Size Structured



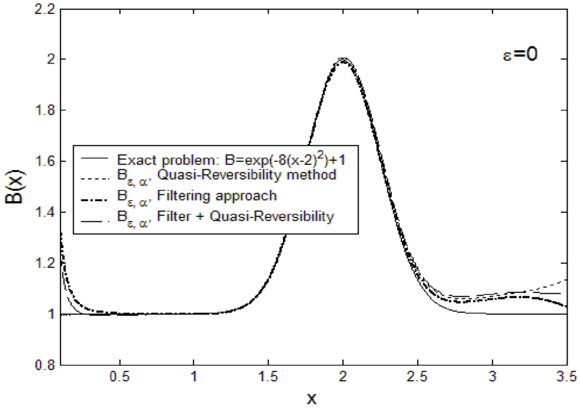
Results with no noise - constant B

Numerical Results - Size Structured



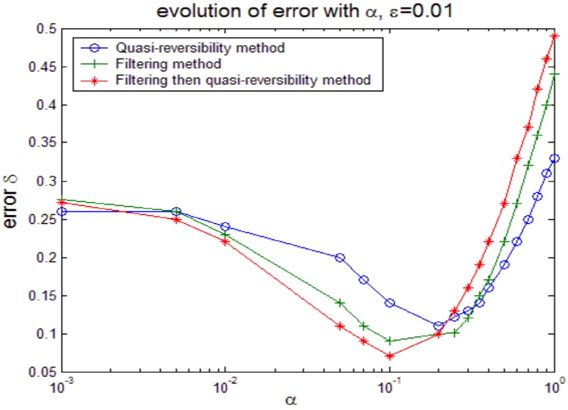
Results with no noise - step B

Numerical Results - Size Structured



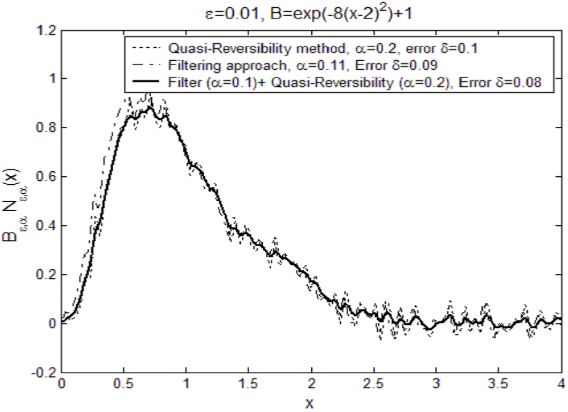
Results with no noise - varying B

Numerical Results - Size Structured



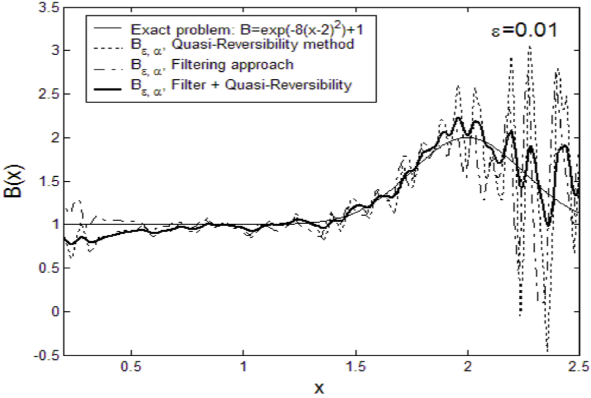
Results with noise $\varepsilon = 0.01$ - Error with respect to the regularization parameter α

Numerical Results - Size Structured



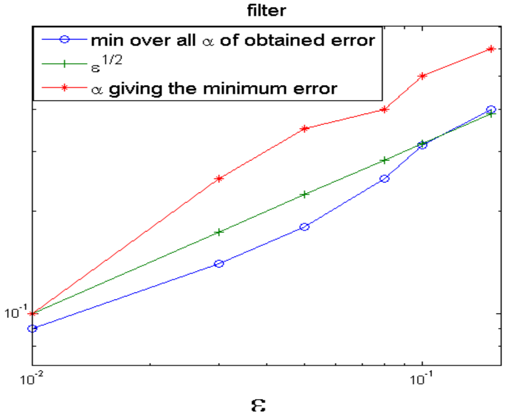
Results with noise $\varepsilon = 0.01$ - BN

Numerical Results - Size Structured



Results with noise $\varepsilon = 0.01$ - B

Numerical Results - Size Structured



Optimal α with respect to ε, compared to $\sqrt{\varepsilon}$ and the optimal error

Indirect measurement: the incremental model

With A. Olivier, L. Robert, DCDS-B, 2020

Recall: $n(t, a, x) \to e^{\kappa t} N(a, x)$ density of cells of size x and increment a.
Definition of an increment: difference between size and size at birth
PDE obtained from the PDMP :

$$\mathbb{P}(\zeta_u \geq a) = e^{-\int_0^a B(s)ds},$$

Indirect measurement: the incremental model

With A. Olivier, L. Robert, DCDS-B, 2020

Recall: $n(t,a,x) \to e^{\kappa t} N(a,x)$ density of cells of size x and increment a.
Definition of an increment: difference between size and size at birth
PDE obtained from the PDMP :

$$\mathbb{P}(\zeta_u \geq a) = e^{-\int\limits_0^a B(s)ds}, \qquad \frac{da}{dt} = \kappa x$$

$$\kappa N + \frac{\partial}{\partial a}\big(\kappa x N\big) + \frac{\partial}{\partial x}\big(\kappa x N\big) = -\kappa x B(a) N(a,x),$$

$$N(0,x) = 8 \int\limits_0^\infty B(a) N(a,2x) da$$

Inverse problem for the increment-structured equation / adder model

Reconstruction formula, deterministic setting - with A. Olivier and L. Robert

If we only measure $\mathcal{N}(x) = \int\limits_0^x N(a,x)da$, can we estimate $B(a)$?

Inverse problem for the increment-structured equation / adder model

Reconstruction formula, deterministic setting - with A. Olivier and L. Robert

If we only measure $\mathcal{N}(x) = \int_0^x N(a,x)da$, can we estimate $B(a)$?

Proposition (MD, A. Olivier, L. Robert, 2020, DCDS-B)

We have the following reconstruction formula:

$$B(a) = \frac{f(a)}{\int_a^\infty f(s)ds}, \qquad f(a) := \mathcal{F}^{-1}\left(1 + i\xi \frac{\mathcal{F}[\tau x^2 \mathcal{N}(x)](\xi)}{\mathcal{F}[4xH(2x)](\xi)}\right),$$

where $H(x)$ is the solution of the dilation equation:

$$\mathcal{L}(x) = \kappa\mathcal{N} + \frac{\partial}{\partial x}(\kappa x \mathcal{N}) = 4H(2x) - H(x).$$

severely ill-posed inverse problem: infinite ("+1"!) degree of ill-posedness...

Inverse problem for the increment-structured equation / adder model

Idea of the proof: solve the equation along the characteristics and integrate in a $\Longrightarrow$

$$\kappa x^2 \mathcal{N}(x) = 4 \int_0^x (y-a) H(2(y-a)) e^{-\int_0^a B(s)ds} da$$

$\Longrightarrow$ deconvolution problem, where $4xH(2x)$ plays the role of "noise".

Estimates would require a priori bounds on $\mathcal{F}[4xH(2x)]$, e.g.

- Ordinary smooth "noise" of order β:
  $c_1|t|^{-\beta} \leq \mathcal{F}[4xH(2x)](t) \leq c_2|t|^{-\beta}$ for $|t| \geq M$
- Super smooth "noise" of order β:
  $c_1|t|^{\gamma_1} e^{-c_0|t|^\beta} \leq |\mathcal{F}[4xH(2x)](t)| \leq c_2|t|^{\gamma_2} e^{-c_0|t|^\beta}$

Inverse problem for the increment-structured equation / adder model

Reconstruction formula, statistical setting - with A. Olivier and L. Robert, DCDS-B, 2020

We observe $X_1, \cdots X_n$ an *i.i.d.* sample of law $\mathcal{N}(x)$

$$\widehat{B}_{n,h}(a) = \frac{\widehat{f}_{n,h}(a)}{\widehat{S}_{n,h}(a) \vee \varpi_2} = \frac{\int\limits_{-1/h}^{1/h} \Big(1 + \mathrm{i}\xi \frac{\widehat{\mathcal{C}_n^*}(\xi)}{\widehat{\mathcal{G}_n^*}(\xi)} 1_{\Omega_n}(\xi)\Big) e^{-\mathrm{i}a\xi} d\xi}{\int\limits_{s}^{+\infty} \int\limits_{-1/h}^{1/h} \Big(1 + \mathrm{i}\xi \frac{\widehat{\mathcal{C}_n^*}(\xi)}{\widehat{\mathcal{G}_n^*}(\xi)} 1_{\Omega_n}(\xi)\Big) e^{-\mathrm{i}s\xi} d\xi ds \vee \varpi}$$

with

$$\widehat{\mathcal{C}_n^*}(\xi) = \frac{1}{n} \sum_{j=1}^{n} \tau X_j^2 e^{\mathrm{i}X_j\xi}, \qquad \widehat{\mathcal{G}}_n(y) = 4y \widehat{H}_n(2y)$$

Inverse problem for the increment-structured equation / adder model

Simulation protocols - with A. Olivier and L. Robert

To analyse separately each term of the formula, we tested 4 protocols:

1. Protocol 1: from all direct functions, FFT & IFFT
2. Protocol 2: from "exact" (simulated) $\mathcal{N}(x)$
3. Protocol 3: from $X_i \sim \mathcal{N}(x)$ and "exact" (simulated) $H(x)$
4. Protocol 4: from $X_i \sim \mathcal{N}(x)$.

Inverse problem for the increment-structured equation / adder model

Simulation results - with A. Olivier and L. Robert

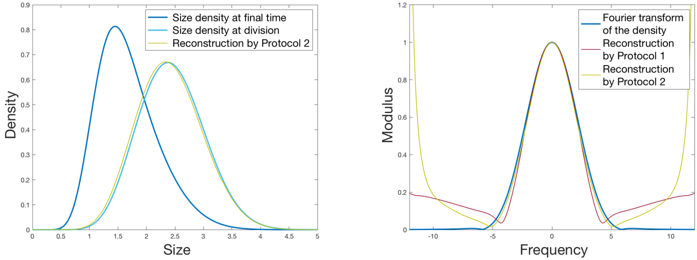


Figure: Left: $\mathcal{N}(x)$, $H(x)$ and $\widehat{H}(x)$ by protocol 2
Right: $|\mathcal{F}(f)(\xi)|$, $|\widehat{\mathcal{F}(f)}_1|$ (Protocol 1) and $|\widehat{\mathcal{F}(f)}_2|$ (Protocol 2)

Inverse problem for the increment-structured equation / adder model

Simulation results - with A. Olivier and L. Robert, DCDS-B, 2020

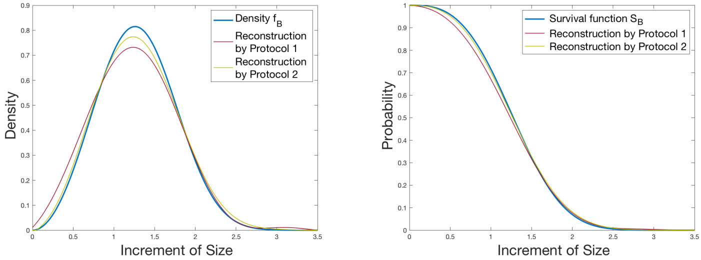


Figure: Left: $f(a)$, $\widehat{f_1}(a)$ (Protocol 1) and $\widehat{f_2}(a)$ (Protocol 2)
Right: $S(a) = \int_a^\infty f(s)ds$, $\widehat{S}_1(a)$ and $\widehat{S}_2(a)$

Inverse problem for the increment-structured equation / adder model

Simulation results - with A. Olivier and L. Robert, DCDS-B, 2020

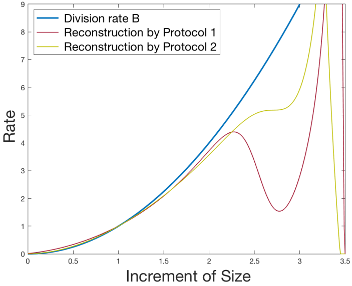


Figure: $B(a)$, $\widehat{B}_1(a)$ (Protocol 1) and $\widehat{B}_2$ (Protocol 2)

Inverse problem for the increment-structured equation / adder model

Simulation results - with A. Olivier and L. Robert, DCDS-B, 2020

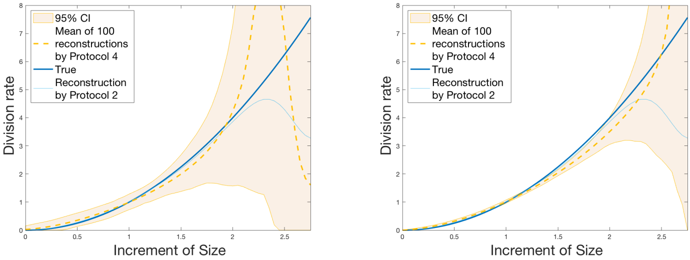


Figure: $\widehat{B}_n$ with $n = 500$ (left), $n = 50\ 000$ (right)

Inverse problem for the increment-structured equation / adder model

Simulation results - with A. Olivier and L. Robert

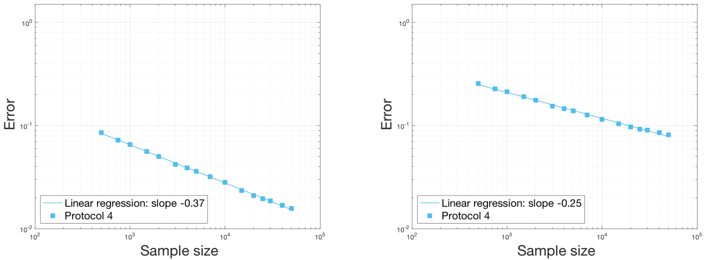


Figure: Estimation of $\mathcal{N}(x)$ (left) and $\frac{d\mathcal{N}}{dx}$ (right)

Inverse problem for the increment-structured equation / adder model

Simulation results - with A. Olivier and L. Robert

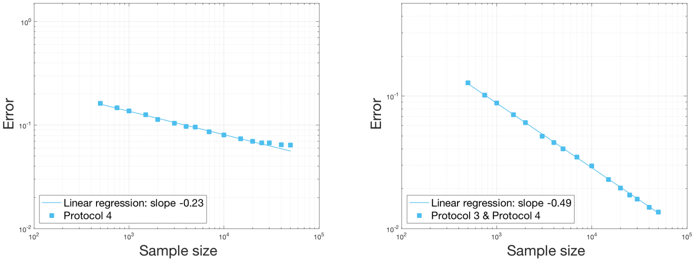


Figure: Estimation of $xH(x)$ (left) and $\widehat{\mathcal{C}_n^*}$ (right)

Inverse problem for the increment-structured equation / adder model

Simulation results - with A. Olivier and L. Robert

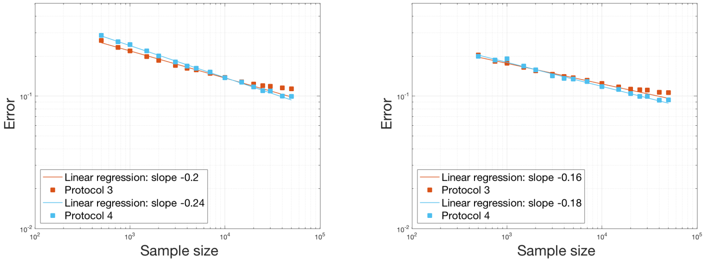
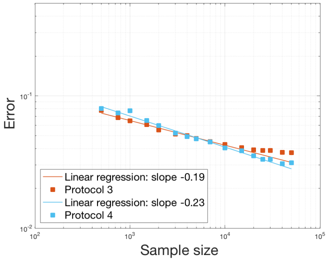


Figure: Estimation of $\mathcal{F}(f)$ (left) and f (right)

Inverse problem for the increment-structured equation / adder model

Simulation results - with A. Olivier and L. Robert



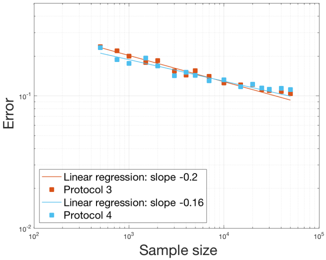
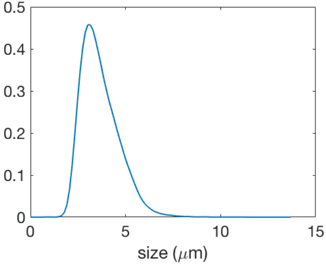


Figure: Estimation of $\int_a^\infty f(s)ds$ (left) and B (right)

Inverse problem for the increment-structured equation / adder model

Test on experimental data - with A. Olivier and L. Robert



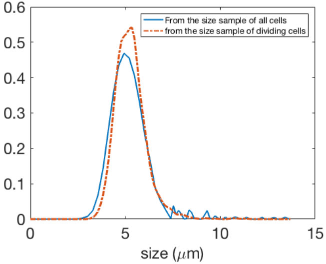


Figure: experimental size distribution (left), reconstructed "at division" size distribution (right)

Inverse problem for the increment-structured equation / adder model

Test on experimental data - with A. Olivier and L. Robert

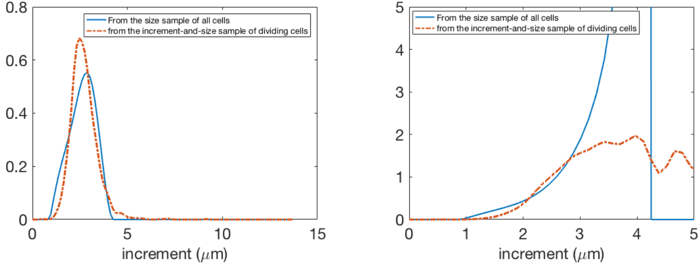
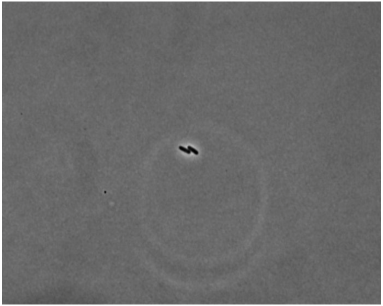


Figure: experimental size distribution (left), reconstructed "at division" size distribution (right)

What if we observe more ?



Second method: full observation

Second method: direct and full observation

Statistical reconstruction
(MD, M. Hoffmann, N. Krell, L. Robert, 2015)
Observation scheme

$$\{(\xi_u, \zeta_u), \quad u \in \mathcal{U}_n\},$$

with $\mathcal{U}_n \subset \mathcal{U}$ a set of n nodes having the property

$$\text{If } u \in \mathcal{U}_n \text{ then } u^- \in \mathcal{U}_n.$$

Asymptotics taken as $n \to \infty$.

We use the link between $f(t)$ the density of the lifetime and the division rate B.

Second method: full observation

We have for the age model

$$\mathbb{P}(\zeta_u \in [t, t+dt]\,|\zeta_u \geq t, \xi_u = x) = B(t)dt$$

or for the size model

$$\mathbb{P}(\zeta_u \in [t, t+dt]\,|\zeta_u \geq t, \xi_u = x) = B(xe^{\kappa t})dt$$

from which we obtain the density of the lifetime $\zeta_u = t$, for the age model:

$$\boxed{f(t) = B(t)\exp\Big(-\int_0^t B(s)ds\Big)}$$

For the size model it is conditional on the size at birth $= x$:

$$\boxed{f(t,x) = B(xe^{\kappa t})\exp\Big(-\int_0^t B(xe^{\kappa s})ds\Big)}$$

Second method: full observation
Age Model (Hoffmann, Olivier, 2015)

To make it short: survival analysis:

we observe a sample of n cells, of ages at division $a_1, \cdots, a_n$ realizations of $A_1, \ldots, A_n$, *i.i.d.* random variables with density $f_k^d(a) = B(a)N_k(a)/\int BN_k da$, and it is well-known that (branch tree)

$$B(a) = \frac{f_1^d(a)}{\int_a^\infty f_1^d(s)ds} = \frac{f_2^d(a)e^{\lambda a}}{\int_a^\infty f_2^d(s)e^{\lambda s}ds}.$$

For the whole tree data till a certain time: "bias" term: f_1^d is replaced by $f_2^d(a)e^{\lambda a} = cf_1^d$ for a normalisation constant c (Efromovich, Ann. Statis. 2004)

Second method: full observation
Size Model (M.D., M. Hoffmann, N. Krell, L. Robert, Bernoulli, 2015)

- explicit representation for the transition kernel $\mathcal{P}_B$ (which links the daughter size/age law to its mother size/age law) reminiscent of conditional survival function estimation.
- Under appropriate condition on B close to the conditions for the eigenvalue PDE problem, the Markov chain is geometrically ergodic (but not reversible).
- existence and uniqueness of an invariant measure $\nu_B(dx)$ such that

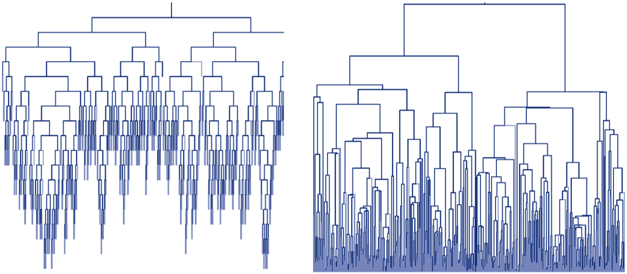
$$\nu_B \mathcal{P}_B = \nu_B.$$

  Convergence through a Lyapunov function.

Second method: direct and full observation

Influence of the observations on the estimator 3 fundamental cases:

- sparse tree case: a line of descendants $(\emptyset, u_1, \cdots, u_n)$
- full tree case: $n = 2^{k_n}$, k_n first generations
- measurements stop at a given time (independent of the number of generations)



The first two cases are equivalent, the third is different.

Second method: full observation

Size Model (M.D., M. Hoffmann, N. Krell, L. Robert, Bernoulli, 2015)

We prove

$$B(y) = \frac{\kappa y}{2} \frac{f_1^b(y/2)}{\mathbb{E}_{f_1^b}\left[1_{\{\xi_u^- \leq y,\, \xi_u \geq y/2\}}\right]}.$$

Stat. estimation: introduce a mollifier sequence to estimate f_1^b

Error estimate: if $B \in H^s$, for appropriate α, we find $B_{\alpha,n}$ such that

$$\boxed{\mathbb{E}\left[\|B_{\alpha,n} - B\|_{L^2}^2\right]^{1/2} \lesssim (\log n) n^{-s/(2s+1)}}$$

convergence rate to compare with the indirect approach: $n^{-\frac{s}{2s+3}}$.

Second method: full observation

From stat back to **PDE...**

Key representation:

$$B(y) = \frac{\kappa y}{2} \frac{f_1^b(y/2)}{\mathbb{E}_{f_1^b}\left[1_{\{\xi_u^- \leq y,\, \xi_u \geq y/2\}}\right]} = \frac{\kappa y}{2} \frac{f_1^b(y/2)}{\int\limits_{\frac{y}{2}}^{y} f_1^b(x)dx}.$$

1 branch data: steady state:

$$\partial_x\big(\kappa x\, N(x)\big) + B(x)N(x) = 2B(2x)N(2x).$$

we identify, up to a constant c, $f_1^b(x) = 2f_1^d(2x) = 2cB(2x)N(2x)$

$$B(y) = \frac{BN(y)}{N(y)} = \kappa y \frac{BN(y)}{\int\limits_{y}^{2y} BN(x)dx} = \frac{\kappa y}{2} \frac{f_1^b(\frac{y}{2})}{\int\limits_{\frac{y}{2}}^{y} f_1^b(x)dx}$$

Second method: full observation

Comparison of the convergence rates and conclusion

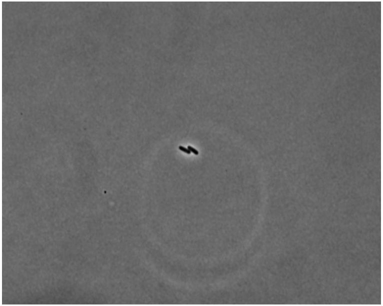
- Reference case: renewal: $B(a) = \frac{f_1^d(a)}{\int_a^\infty f_1^d(s)ds} = \frac{f_2^d(a)e^{\lambda a}}{\int_a^\infty f_2^d(s)e^{\lambda s}ds}$
- Deterministic problem: well-posed! Degree of ill-posedness $a = 0$ - estimate in $O(\varepsilon)$ -
- Statistical viewpoint: density estimate, $H^{-1/2}$ to L^2 so that $a = 1/2$

$$\varepsilon^{s/(s+1/2)} = n^{-s/(2s+1)}$$

- to be compared to the indirect method: error in the order of $\varepsilon^{s/(s+3/2)} = n^{-s/(2s+3)}$.
- Population case: formula to adapt (MD, Hoffmann, 2022)

$$B(x) \quad = \frac{\tau(x)f_2^d(x)}{\int_x^\infty \big(f_2^d(y)-2f_2^b(y)\big)e^{\lambda_2 \int_x^y \frac{ds}{\tau(s)}}dy} = \frac{\kappa x^2 f_2^d(x)}{\int_x^{2x} y f_2^d(y)dy}.$$

Step 6: Finally back to the data...



OR

Will we be able to select or reject our models ?

6. Back to the data

(M.D., M. Hoffmann, N. Krell, L. Robert, BMC Biology, 2014)

To test a model:

- calibrate it (previously seen methods and data)
- simulate the **age-size** PDE model:

$$\frac{\partial}{\partial t} n + \frac{\partial}{\partial a} n + \frac{\partial}{\partial x}\big(\kappa x n\big) = -B(a,x) n(t,a,x),$$

$$n(t, a=0, x) = 4 \int_0^\infty B(a, 2x) n(t, a, 2x) da$$

  till its asymptotic steady behaviour $n(t,a,x) = e^{\lambda t} N(a,x)$
- compare quantitatively data and simulations
- conclude !

6. Back to the data

(M.D., M. Hoffmann, N. Krell, L. Robert, BMC Biology, 2014)

To test a model:

- calibrate it (previously seen methods and data)
- simulate the **age-size** PDE model:

$$\frac{\partial}{\partial t} n + \frac{\partial}{\partial a} n + \frac{\partial}{\partial x}\left(\kappa x n\right) = -B(a,x) n(t,a,x),$$

$$n(t, a = 0, x) = 4 \int_0^\infty B(a, 2x) n(t, a, 2x) da$$

  till its asymptotic steady behaviour $n(t,a,x) = e^{\lambda t} N(a,x)$
- compare quantitatively data and simulations
- conclude ! If possible...

6. Back to the data
experimental age/size data - whole tree till a certain time

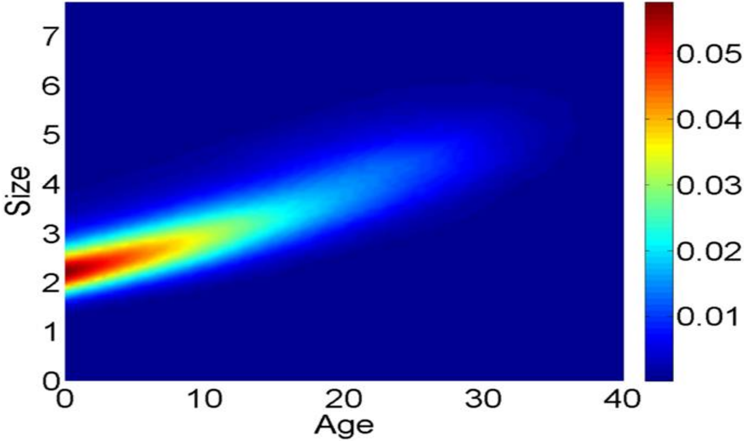


Figure: Age Size Distribution for all cells - whole tree data

6. Back to the data
experimental age/size data - 1 branch data

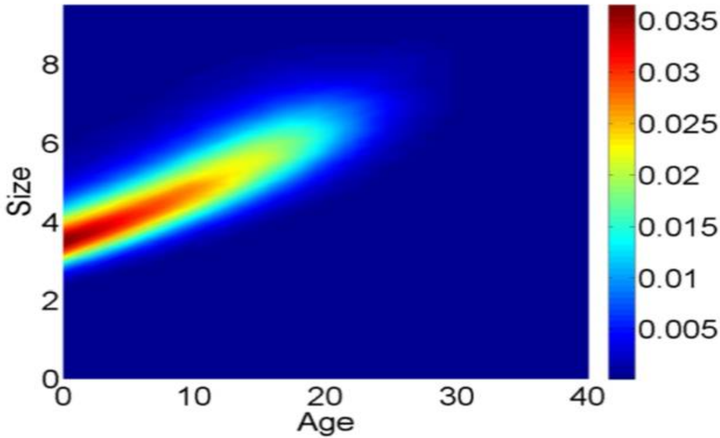


Figure: Age Size Distribution for all cells - tree branches data

Testing the Age Model

Back to the data: testing the Age Model

(M.D., M. Hoffmann, N. Krell, L. Robert, BMC Biology, 2014)

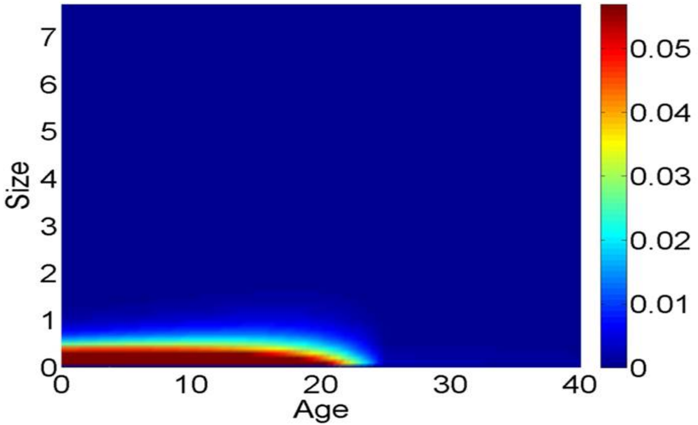


Figure: Age Size simulation for the Age Model - whole tree data

Back to the data: testing the Age Model with a corrected growth rate

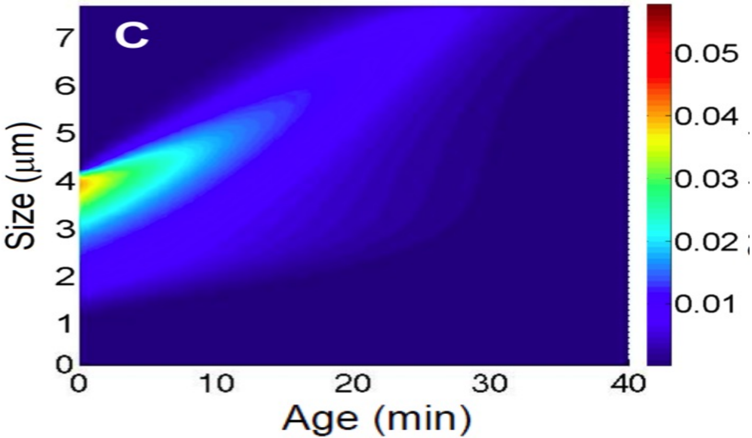


Figure: Age Size simulation for the Age Model - whole tree data

Back to the data: testing the Age Model

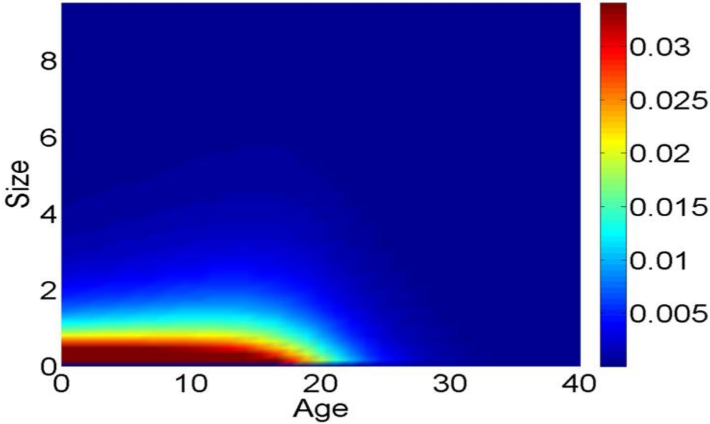


Figure: Age Size simulation for the Age Model - branch tree data

Back to the data: testing the Age Model with a corrected growth rate

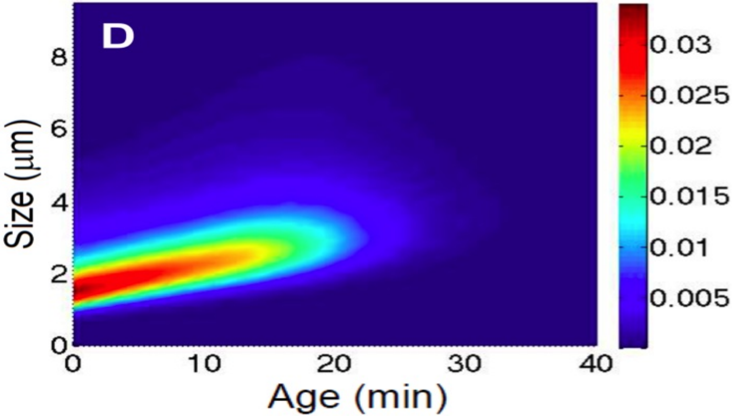


Figure: Age Size simulation for the Age Model - branch tree data

Age Model: conclusion

- As it is, this model is rejected
- Theoretical reason: exponential growth + age-dependent division rate lead to accumulation towards 0.
- Refer to theoretical results for the asymptotic regime: we need $\frac{B(x)}{x} \in L^1_0$ - false here
- This theory is not sufficient: corrected growth rate dependence on these corrections is too important

Testing the Size Model

Back to the data: testing the Size Model

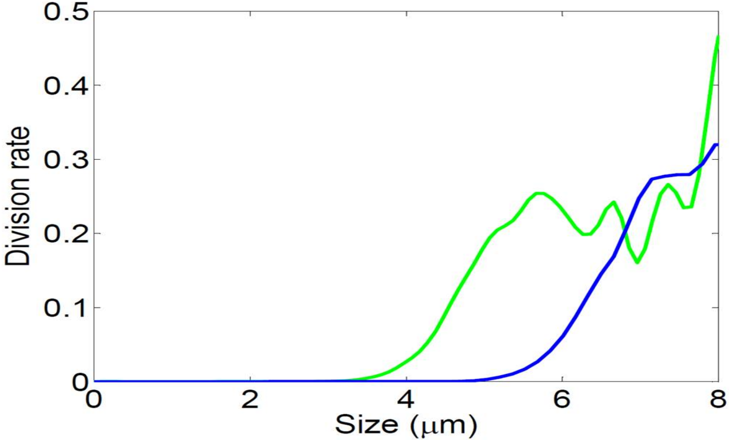
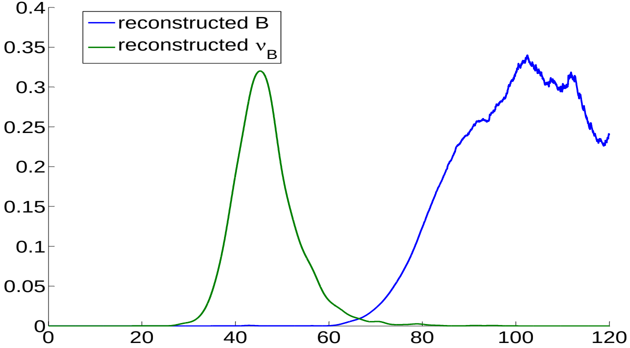


Figure: Reconstruction of the division rate - green: whole tree, blue: branches data

Size Model: reconstruction for B



Back to the data: testing the Size Model

Not too bad but...

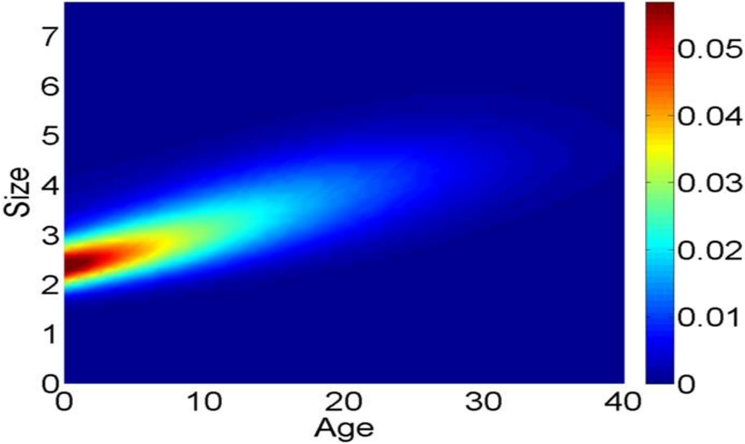


Figure: Age Size simulation for the Size Model - whole tree data

Back to the data: testing the Size Model

Not too bad but...

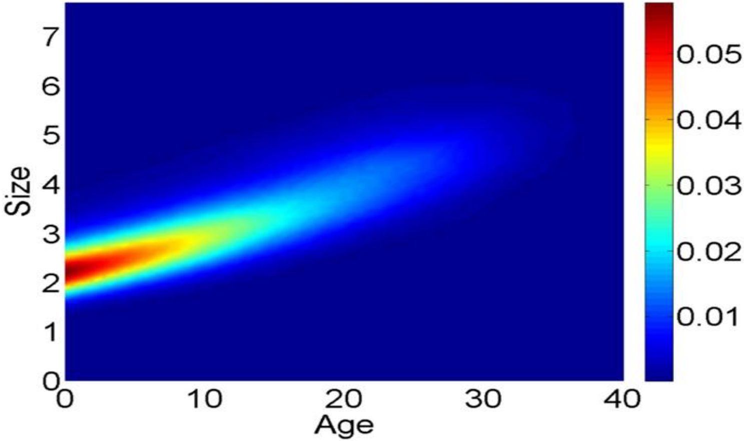


Figure: Age Size experimental data - whole tree data

Back to the data: testing the Size Model

Not too bad but...

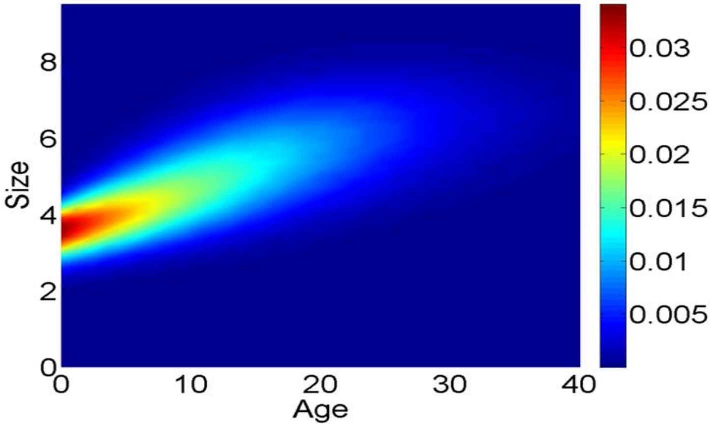


Figure: Age Size simulation for the Size Model - branch tree data

Back to the data: testing the Size Model

Not too bad but...

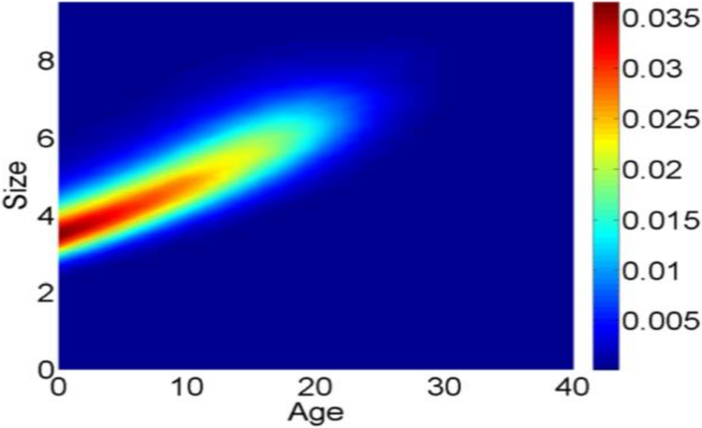


Figure: Age Size experimental data - branch tree data

The incremental/adder model

Rich data / "direct" approach: from "at division" distributions

The incremental model:

Increment= difference between size and size at birth

PDE obtained from the PDMP (as previously): asymptotically, for the 1-branch case:

$$\mathbb{P}(a \leq \zeta_u \leq a{+}da) = f(a) = B(a)e^{-\int\limits_0^a B(s)ds} = \frac{B(a)\int\limits_0^\infty xN(a,x)dx}{\iint xB(a)N(a,x)dadx}$$

$$\frac{\partial}{\partial a}\big(\kappa x N\big) + \frac{\partial}{\partial x}\big(\kappa x N\big) = -\kappa x B(a)N(a,x),$$

$$N(0,x) = 4\int\limits_0^\infty B(a)N(a,2x)da$$

The best argument to date: the correlation between size at birth and increment of size at division

(increment model: 0, size model: ~ -0.4, data: ~ -0.1)

What about an Age-Size Model ?

To test it, we would need an extra variable:

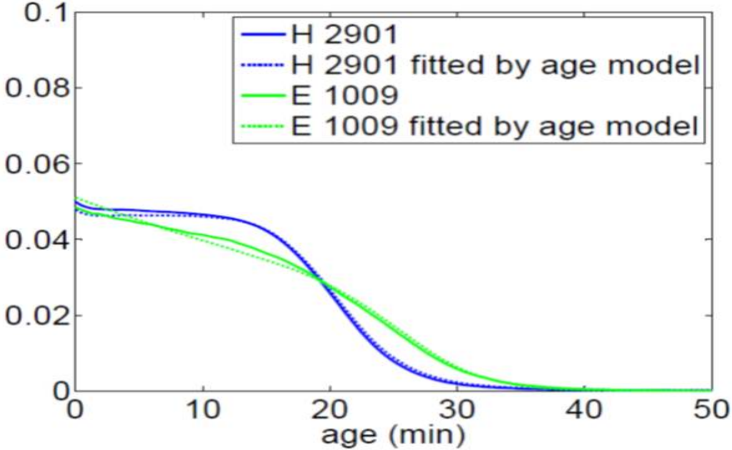


Figure: Age distribution: data and fit by the age model

What about an Age-Size Model ?

To test it, we would need an extra variable:

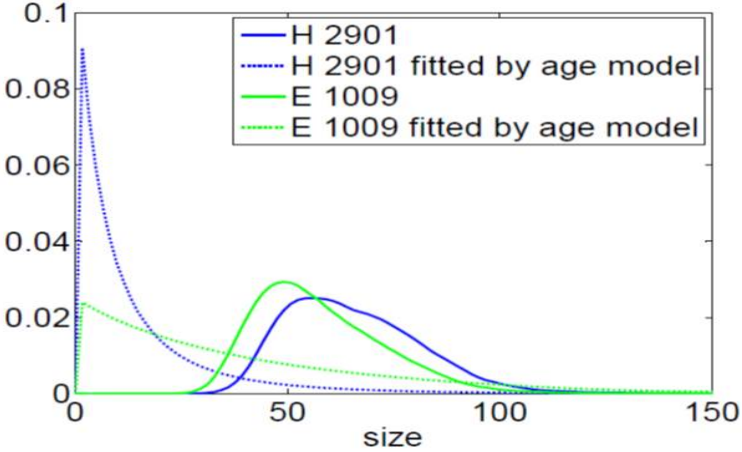


Figure: Size distribution: data and fit by the age model

The fragmentation case

application to fragmenting protein fibrils

(with Miguel Escobedo, Bilbao and Magali Tournus, Marseille, data from W.F. Xue's group in Canterbury)

Classical assumptions on the fragmentation equation

Also assumed by W.F. Xue and S. Radford, Biophys. J., 2013

- $B(x)$ = Fragmentation rate of particles of size x.
  $B(x) = \alpha x^\gamma$
- $k(x, y)$ = Fragmentation kernel.
  $k(x, y) = \frac{1}{y} k_0\left(\frac{x}{y}\right)$, where k_0 is a measure on $[0, 1]$.

Theorem (Escobedo-Mischler-Ricard – Ann. IHP 2005)

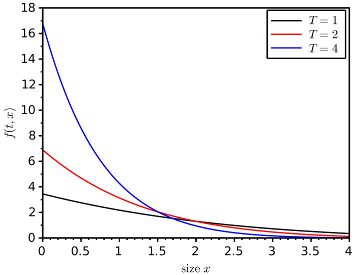
Under reasonable technical assumptions, for large time, the profile tends to a self-similar profile g :

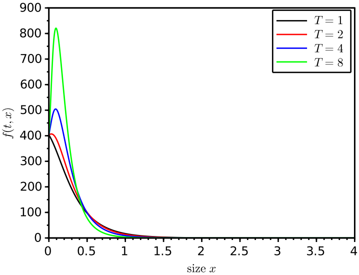
$$n(t, x) \to t^{\frac{2}{\gamma}} g\left(xt^{\frac{1}{\gamma}}\right), \qquad L^1(x\,dx) \tag{9}$$

where g is the unique solution of

$$\frac{\partial}{\partial z}(zg(z)) + (1 + \alpha\gamma z^\gamma)g(z) = \alpha\gamma \int_z^\infty \frac{1}{y} k_0\left(\frac{z}{y}\right) y^\gamma g(y)dy, \; \int_0^\infty zg(z)dz = \rho.$$

Two examples.





$k_0(x) = 2\mathbb{1}_{[0,1]}(x)$ $\qquad$ $k_0(x) = 2\delta_{x=1/2}(x)$

First reconstruction idea: use self-similar profile g to estimate α, γ and k_0

First reconstruction idea: use self-similar dynamics

- For fragmentation equations: Old problem
recover the transition probability of droplet breakage from experimental measurements of transient drop size distributions in a stirred liquid-liquid dispersion: using a fragmentation equation assuming self similarity.
Valentas, K. J., and N. R. Amundson, I.E.C. Fundls., 1966, 1968.
G. Narsimhan, D. Ramkrishna, J. P. Gupta, Chem. Ing. Sci , 1979

- Similar idea as seen above for growth fragmentation equations, where steady Malthusian behaviours replace self similarity.

Inverse problem observing g

Estimate all the fragmentation characteristics γ, α, and k_0

$$\frac{\partial}{\partial z}(zg(z)) + (1 + \alpha\gamma z^{\gamma})g(z) = \alpha\gamma \int\limits_z^{\infty} \frac{1}{y} k_0(\frac{z}{y}) y^{\gamma} g(y) dy$$

Mellin transform: $\mathcal{M}[g](s) = \int_0^{\infty} x^{s-1} g(x) dx$

$$(2 - s)\mathcal{M}[g](s) + \alpha\gamma\mathcal{M}[g](s + \gamma) = \alpha\gamma\mathcal{M}[k_0](s)\mathcal{M}[g](s + \gamma),$$

Theorem (MD,Escobedo, Tournus, Ann. IHP, 2018)

For any $g \in L^1(\mathbb{R}_+)$ such that for all $k \geq 0$ $\int x^k g(x) dx < \infty$, there exists at most one triplet $(\gamma, \alpha, k_0(x)) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathcal{M}^1([0, 1])$ such that g is the self-similar profile of the fragmentation equation.

Some ideas and comments on the proof

$(2-s)\mathcal{M}[g](s) + \alpha\gamma\mathcal{M}[g](s+\gamma) = \alpha\gamma\mathcal{M}[k_0](s)\mathcal{M}[g](s+\gamma)$

First step: determine γ

Proposition

Given any constant $R > 0$*:*

$$\lim_{s\to\infty,\, s\in\mathbb{R}^+} \frac{s\,\mathcal{M}[g](s)}{\mathcal{M}[g](s+R)} = \begin{cases} 0, & \forall R > \gamma \\ \alpha\gamma, & \text{if } R = \gamma \\ \infty, & \forall R \in (0,\gamma) \end{cases}$$

Use the asymptotic behaviour of $g(x)$ in 0 and $+\infty$ / of $\mathcal{M}[g](s)$ for $s \to +\infty$

[other result: direct estimates in (Balagué, Cañizo, Gabriel, 2013)]

Second step: determine α : Plug $s = 2$.

$$\alpha = \frac{\mathcal{M}[g](1)}{\gamma\mathcal{M}[g](1+\gamma)}.$$

Some ideas and comments on the proof

$(2-s)\mathcal{M}[g](s) + \alpha\gamma\mathcal{M}[g](s+\gamma) = \alpha\gamma\mathcal{M}[k_0](s)\mathcal{M}[g](s+\gamma)$

Third step: determine $\mathcal{M}[k_0]$. ($\rightsquigarrow k_0$)

$$(\mathcal{M}[k_0](s) - 1) = \frac{\mathcal{M}[g](s)(2-s)}{\alpha\gamma\mathcal{M}[g](s+\gamma)}, \qquad s \in \mathbb{C}.$$

Cauchy integral to solve this equation ; first prove that the denominator does not vanish by explicit solution.
(see also Hoang Ngoc Rivoirard Tran, 2020)

Existence: of a reconstruction formula $\Longrightarrow$ invert Mellin

Some ideas and comments on the proof

$(2-s)\mathcal{M}[g](s) + \alpha\gamma\mathcal{M}[g](s+\gamma) = \alpha\gamma\mathcal{M}[k_0](s)\mathcal{M}[g](s+\gamma)$

Third step: determine $\mathcal{M}[k_0]$. ($\rightsquigarrow k_0$)

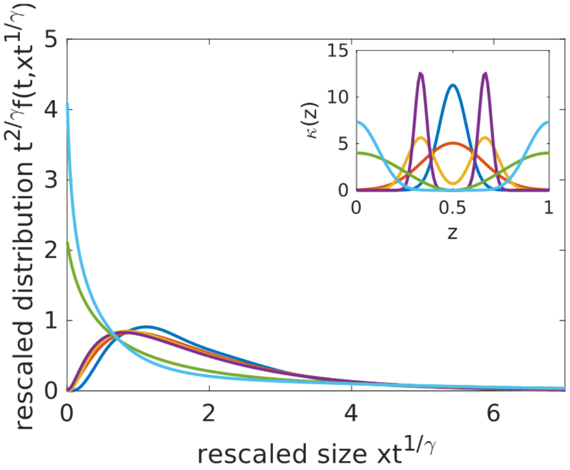
$$(\mathcal{M}[k_0](s) - 1) = \frac{\mathcal{M}[g](s)(2-s)}{\alpha\gamma\mathcal{M}[g](s+\gamma)}, \qquad s \in \mathbb{C}.$$

Cauchy integral to solve this equation ; first prove that the denominator does not vanish by explicit solution.
(see also Hoang Ngoc Rivoirard Tran, 2020)

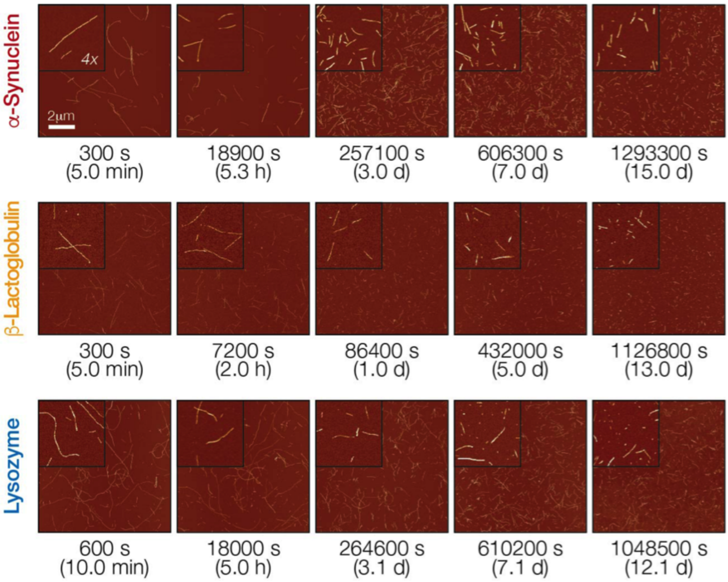
Existence: of a reconstruction formula $\Longrightarrow$ invert Mellin
Stability only in a very weak sense: severely ill-posed inverse problem
$+$ estimation for α and γ use $g(0)$ or $g(+\infty)$: impossible to observe

Some asymptotic profiles in practice...

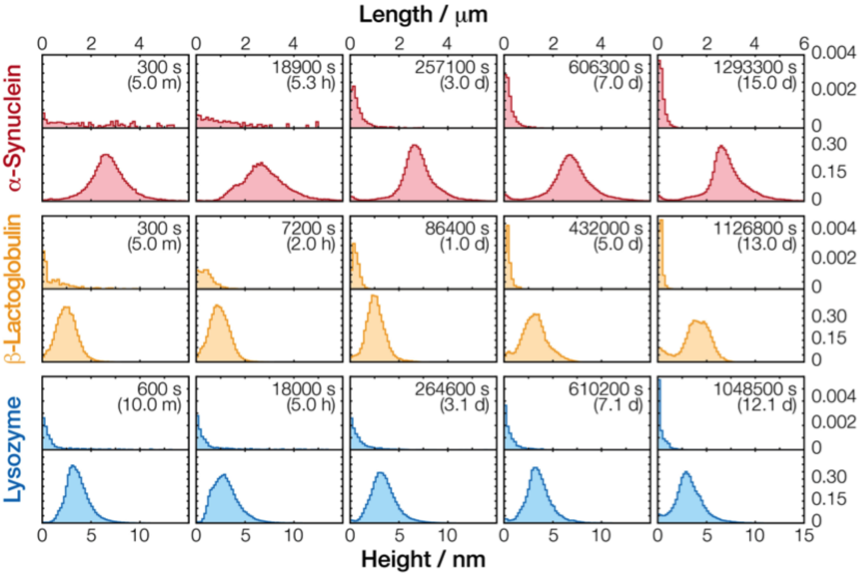
Tournus, Escobedo, Xue, MD, PLoS Comp Biol, 2021



Back to biologists... and to experimental data



Back to biologists... and to experimental data



What did experimentalists before they met us?

W.F. Xue, S. Radford, PNAS 2008 & Biophys. J., 2013

Question : Determine $\gamma \in \mathbb{R}, \alpha \in \mathbb{R}$ and k_0.

- Regularization of the data. Polynomial functions (instead of kernel regularization).
- Parametrization of the fragmentation kernel $k_0 \rightsquigarrow$ The problem becomes : Determine $\gamma, \alpha, k_1, k_2, k_3, k_4 \in \mathbb{R}^6$
- Solve the direct problem for the comprehensive set of admissible parameters $\gamma, \alpha, k_1, k_2, k_3, k_4 \in \mathbb{R}^6$.
- Total linear least square analysis to determine which set of parameters fits best.

... and it worked quite well in practice...

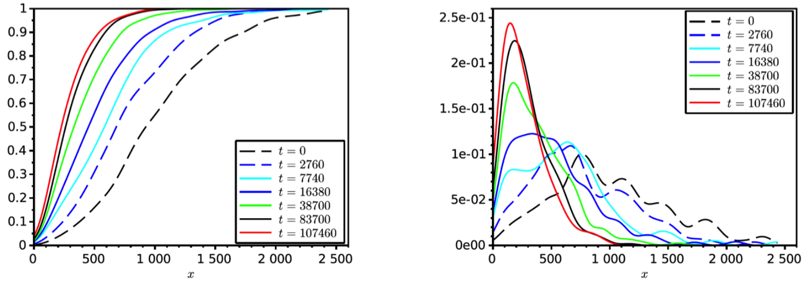
What we proposed them to do

D.M. Beal, M. Tournus,... M. Doumic, W-F. Xue, biorXiv

At different times, a sample of fibril sizes is measured: $f(t,x) := \frac{n(t,x)}{\int n(t,x)dx}$.

Average length: $\mu(t) = \int x f(t,x) dx \sim_{t\to\infty} C t^{-\frac{1}{\gamma}}$

$$\alpha \sim_{t\to\infty} \frac{1}{\gamma t} \frac{1}{\int x^\gamma f(t,x) dx}$$

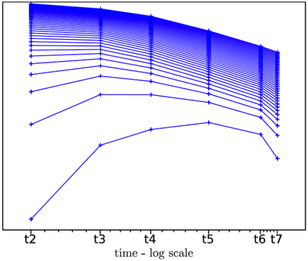


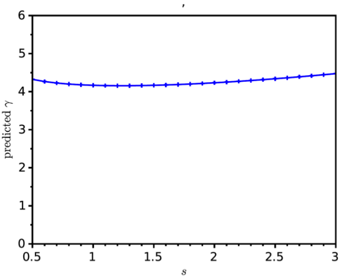
Left: cumulative distribution functions, Right: density functions, at several time points.

Estimate γ + First test on the model

For large times, $\log\left(\mathcal{M}[\frac{u}{\int u dx}](s+1,t)\right) = -\frac{s}{\gamma}\log(t)+\log(C_s).$

γ is the slope of $\log(t) \mapsto -\log\left(\mathcal{M}[\frac{u}{\int u dx}](s+1,t)\right)/s$, for $s \in [0,+\infty]$.

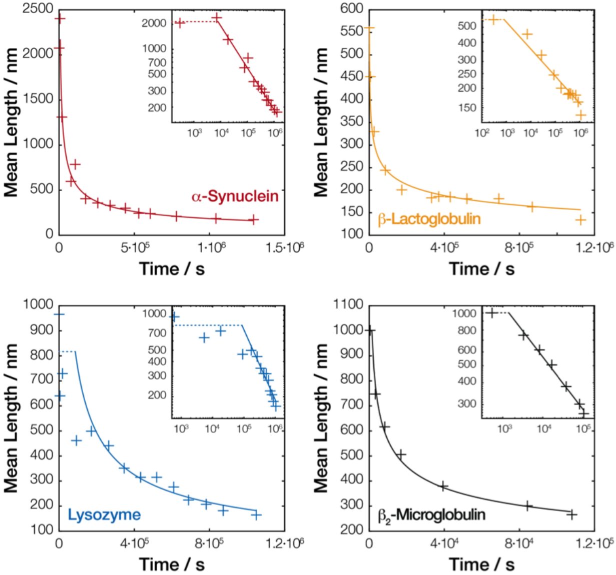




Here we predict $\gamma \approx 4.2$: small fibrils more unlikely to break up.

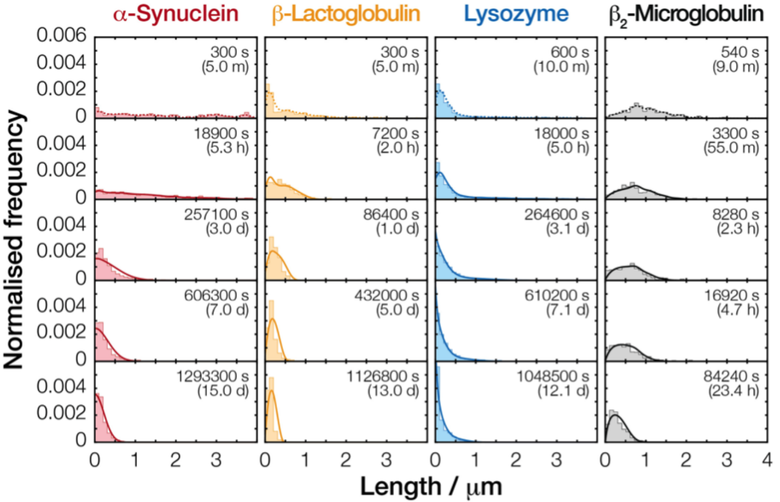
Estimate γ with $\mu(t) = \int x f(t,x)dx \sim_{t\to\infty} Ct^{-\frac{1}{\gamma}}$

D.M. Beal, M. Tournus,... M. Doumic, W-F. Xue, iScience, 2020

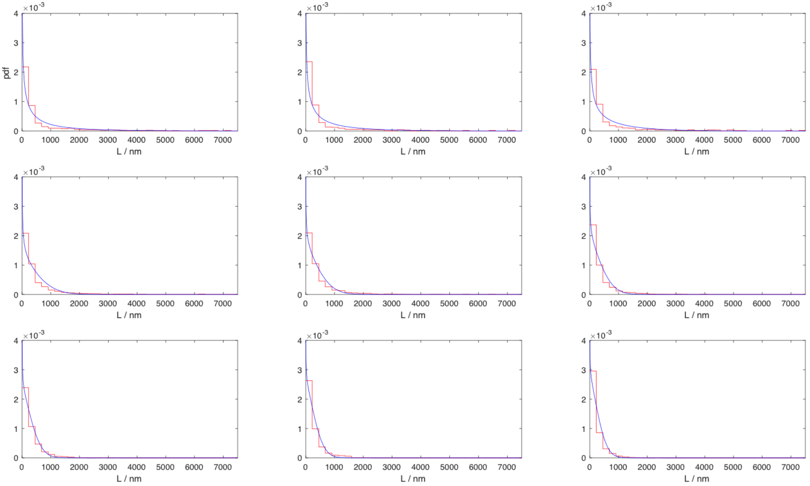


Back to the data: simulations with (α, γ) little influence of k_0

D.M. Beal, M. Tournus,... M. Doumic, W-F. Xue, biorXiv

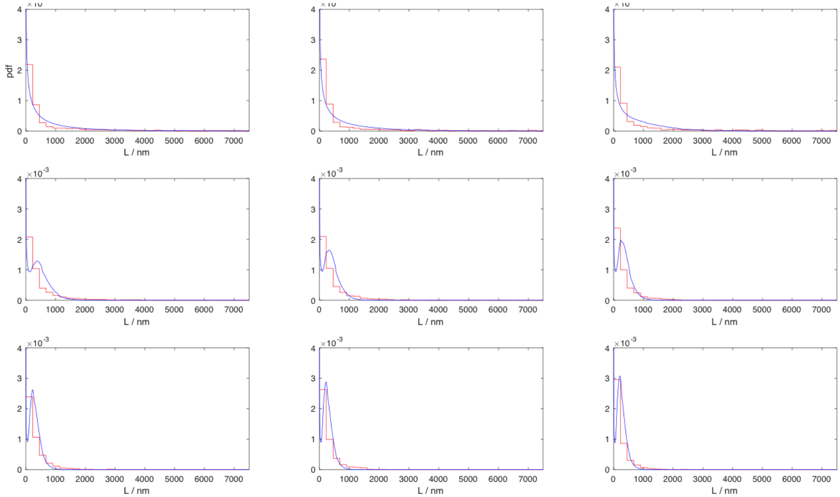


Results: influence of α and γ, small influence of k_0...



k_0 uniform - Lyzozyme c

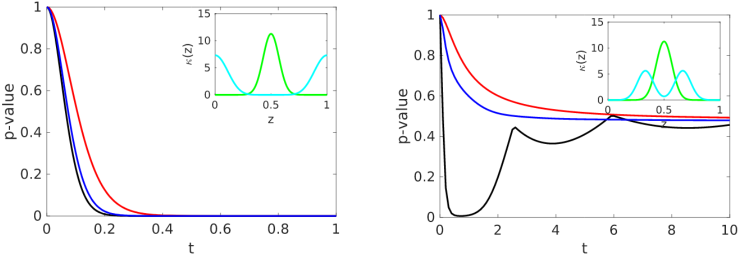
Results: influence of α and γ, small influence of k_0...



k_0 Delta Dirac in $1/2$ - Lyzozyme c

Then what to do? Some numerical investigation first

Tournus, Escobedo, Xue, MD, PLoS Comp Biol, 2021

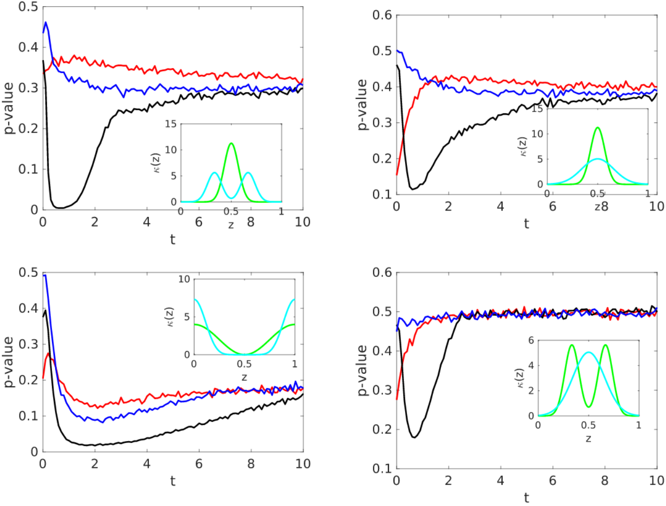


When can we distinguish 2 distributions?
Insets: 2 different kernels
Initial condition: Black: peaked gaussian -
Blue: spread gaussian - Red: decreasing exponential

Time evolution of the p-value of a Kolmogorov-Smirnov test

Then what to do? Some numerical investigation first

Tournus, Escobedo, Xue, MD, PLoS Comp Biol, 2021



$N = 200$

Use the short time behaviour

Some heuristics first

If $u(0,x) = \delta(x-1)$, and $0 < t < 1$,

$$\frac{\partial u}{\partial t}(t,x) + \alpha x^\gamma u(t,x) = \alpha \int_0^1 (\frac{x}{z})^\gamma u(t,\frac{x}{z}) \frac{k_0(dz)}{z}$$

$$\frac{u(t+\Delta t,x) - u(t,x)}{\Delta t} + \alpha x^\gamma u(t,x) \approx \alpha \int_0^1 (\frac{x}{z})^\gamma u(t,\frac{x}{z}) \frac{k_0(dz)}{z}$$

$$\frac{u(\Delta t,x) - \delta(x-1)}{\Delta t} + \alpha x^\gamma \delta(x-1) \approx \alpha \int_0^1 (\frac{x}{z})^\gamma \delta(\frac{x}{z}-1) \frac{k_0(dz)}{z}$$

$$\frac{u(\Delta t,x) - \delta(x-1)}{\Delta t} + \alpha \delta(x-1) \approx \alpha k_0(x)$$

$$\boxed{k_0(x) \approx k^{est}(x) = \frac{1}{\alpha \Delta t}\left(u(\Delta t,x) - (1-\alpha\Delta t)\delta(x-1)\right).}$$

Use the short time behaviour

A first Total Variation result

Theorem (MD, Escobedo, Tournus, preprint arXiv:2112.10423)

The unique fundamental solution U to the fragmentation equation with the initial condition $u_0 = \delta(x-1)$ *satisfies, for* $t \in [0, T]$ *and for some* $K > 0$ *depending on* T *and* α

$$\left\| \frac{U(t) - e^{-\alpha t}\delta(x-1)}{\alpha t} - k_0 \right\|_{TV} \leq Kt.$$

$$||\mu||_{TV} = \sup \left\{ \int_{[0,\infty)} \varphi(x) d\mu(x), \varphi \in L^1(d|\mu|) \cap L^\infty,\ ||\varphi||_\infty \leq 1 \right\}.$$

BUT: The situation for the experimentalists:
1.- $\delta(x-1)$ as initial data impossible $\rightarrow$ build something "close" $u_{q,0}$
2.- Do not measure $u_{q,0}$ and its solution $u_q(t)$, but $u_{q,0,\varepsilon_0}$ and $u_{q,\varepsilon}(t)$.

Use the short time behaviour

A stability result in a Bounded-Lipshitz norm

Theorem (MD, Escobedo, Tournus, preprint arXiv:2112.10423)

Let $u_{q,0} \in \mathcal{M}(\mathbb{R}^+)$ *such that* $Supp(u_{q,0}) \subset [m, M]$ *for* $m, M > 0$ *and*

$$\|u_{q,0} - \delta(x-1)\|_{BL} \ \leq q.$$

Let u_q *the unique solution to the frag eq. with* $u_q(0) = u_{q,0}$. *Let* $u_{q,0,\varepsilon_0}$ *and* $u_{q,\varepsilon}$ *noisy measurements:*

$$\|u_{q,0,\varepsilon_0} - u_{q,0}\|_{BL} \leq \varepsilon_0, \quad \|u_{q,\varepsilon}(t) - u_q(t)\|_{BL} \leq \varepsilon.$$

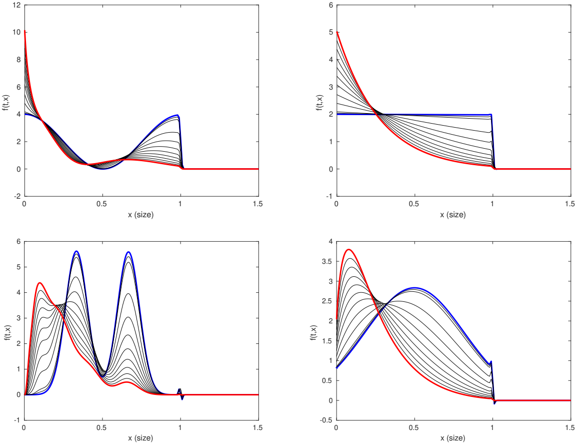
Then, for constants K_1 *and* K_2 *depending on* M *and* T,

$$\left\| \frac{u_{q,\varepsilon}(t) - e^{-\alpha t} u_{q,0,\varepsilon_0}}{\alpha t} - k_0 \right\|_{BL} \leq K_1 t + \frac{K_2 q + \varepsilon_0 + \varepsilon}{\alpha t},$$

where $\|\mu\|_{BL} = \sup \left\{ \int_{[0,\infty)} \varphi(x) d\mu(x), \varphi \in L^1(d|\mu|) \cap W^{1,\infty}, \ ||\varphi||_{W^{1,\infty}} \leq 1 \right\}$

Use the short time behaviour

Numerical illustration



Plot of $u(t,x) - e^{-\alpha t}\delta(x-1)$ for $\alpha = \gamma = 1$ and 4 different k_0.
Blue: $t = 10^{-3}$; Red: $t = 3$.
A good approximation of the kernel is seen on the curves in blue.

Conclusion and perspectives

- Method may be adapted to other cases and models
- Coherence and complementarity between PDE, stoch and stat
- a basis for new biological questions: coordination between growth and division, influence of variability...
- Short-time behaviour well-adapted to estimate the frag kernel; to test on real data ... and study from a stochastic viewpoint
- A new problem: estimate the mutation rate in bacteria - G. Garnier's Ph.D

Many have contributed...

Pierre Gabriel, Thibault Bourgeron, Miguel Escobedo, Magali Tournus, Benoit Perthame, Jorge Zubelli, Pedro Maia, Marc Hoffmann, Patricia Reynaud-Bouret, Lydia Robert, Vincent Rivoirard, Nathalie Krell, Adélaïde Olivier, Adeline Fermanian, Anaïs Rat, Wei-Feng Xue, Cédric Doucet...

to be continued!

The fragmentation and growth-fragmentation equations
General form

Recall of the probabilist view: "our" operator is "their" adjoint

$$\frac{\partial}{\partial t} n = L^* n + \mathcal{F}^* n,$$

where

- L^* is the adjoint of the infinitesimal generator L of the càdlàg strong Markov process $(X_t)_{t \geq 0}$. Here $\mathcal{X} = (0, \infty)$ and L^* is taken deterministic: $Lf = \tau(x) f'(x)$ so that $L^* n = \big(\tau n\big)'$.

- $\mathcal{F}^*$ is the adjoint of the fragmentation operator

  $$\mathcal{F} f(x) := B(x) \int_{\mathcal{X}} \sum_{j \geq 0} \big(j f(y) - f(x)\big) p(j) P^{(j)}(x, dy),$$

  where $P^{(j)}(x, dy)$ is the symetrized fragmentation kernel: probability of an individual of size x to split in j parts, one of them of size y.

Inverse problem for the increment-structured equation / adder model

Reconstruction formula, deterministic setting - with A. Olivier and L. Robert

If we only measure $\mathcal{N}(x) = \int_0^x N(a,x)da$, can we estimate $B(a)$?

Inverse problem for the increment-structured equation / adder model

Reconstruction formula, deterministic setting - with A. Olivier and L. Robert

If we only measure $\mathcal{N}(x) = \int_0^x N(a,x)da$, can we estimate $B(a)$?

Proposition (MD, A. Olivier, L. Robert, 2020, DCDS-B)

We have the following reconstruction formula:

$$B(a) = \frac{f(a)}{\int_a^\infty f(s)ds}, \qquad f(a) := \mathcal{F}^{-1}\left(1 + i\xi \frac{\mathcal{F}[\tau x^2 \mathcal{N}(x)](\xi)}{\mathcal{F}[4xH(2x)](\xi)}\right),$$

where $H(x)$ is the solution of the dilation equation:

$$\mathcal{L}(x) = \kappa \mathcal{N} + \frac{\partial}{\partial x}(\kappa x \mathcal{N}) = 4H(2x) - H(x).$$

severely ill-posed inverse problem: infinite ("+1"!) degree of ill-posedness...

Inverse problem solution with the Mellin transform

Problem.

Without any a priori knowledge on the fragmentation process, but measuring g identify the parameters γ, α, and k_0.

Supplementary hypothesis on k_0: no Dirac mass at $x = 0$ or $x = 1$,

$$\exists \varepsilon > 0,\ k_0 \in C[0,\varepsilon] \cap C[1-\varepsilon,\varepsilon],$$
$$\exists \varepsilon' > 0,\ \eta_2 > \eta_1 > 0;\ \ k_0(z) \geq \varepsilon',\ \forall z \in [\eta_1, \eta_2].$$

Theorem (MD,Escobedo, Tournus, Ann. IHP, 2018)

For any $g \in L^1(\mathbb{R}_+)$ such that for all $k \geq 0$ $\int x^k g(x)dx < \infty$, there exists at most one triplet $(\gamma, \alpha, k_0(x)) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathcal{M}^1([0,1])$ such that g is the self-similar profile of the fragmentation equation.

The fragmentation and growth-fragmentation equations

First focus: $\tau(x) \equiv x$

$$\frac{\partial}{\partial t} n(t,x) + \frac{\partial}{\partial x}\big(x n(t,x)\big) + B(x) n(t,x) = \int_0^1 B(\frac{x}{z}) n(t,\frac{x}{z}) \frac{k_0(dz)}{z}$$

Linked to the fragmentation equation

$$\frac{\partial}{\partial t} u(t,x) + B(x) u(t,x) = \int_0^1 B(\frac{x}{z}) u(t,\frac{x}{z}) \frac{k_0(dz)}{z}$$

by $u(t,x) = e^t n(t, x e^t)$

Critical fragmentation: first insight in the asymptotics

$$\begin{cases} \frac{\partial}{\partial t} u(t,x) + u(t,x) = \int\limits_0^1 u(t, \frac{x}{z}) \frac{k_0(dz)}{z}, \\ \\ u(0,x) = u^{in}(x) \in L^1\big((1+x)dx\big) \end{cases} \qquad (10)$$

Proposition

A solution $u \in C^1((0,\infty); L^1((1+x)dx))$ *of* (10) *satisfies*

$$xu(t,x) \rightharpoonup M\delta, \quad \text{as } t \to +\infty, \quad \text{in } \mathcal{D}'(\mathbb{R}^+), \quad M = \int xu^{in}(x)dx.$$

Mellin transform for the fragmentation equation

$$\mathcal{M}_f(s) = \int_0^\infty x^{s-1} f(x) dx$$

The Mellin transform is the Fourier transform in $y = \log x$

Denote $U(t,s) := \mathcal{M}_{u(t,\cdot)}(s)$, $U_0(s) = \mathcal{M}_{u_0}(s)$, $K(s) := \mathcal{M}_{k_0}(s)$.

$$\frac{\partial}{\partial t} U(t,s) + U(t,s) = K(s)\, U(t,s)$$

$$\Rightarrow U(t,s) = U_0(s)\, e^{(K(s)-1)t}$$

Mellin transform for the fragmentation equation

$$\mathcal{M}_f(s) = \int_0^\infty x^{s-1} f(x) dx$$

The Mellin transform is the Fourier transform in $y = \log x$

Denote $U(t,s) := \mathcal{M}_{u(t,\cdot)}(s),\ U_0(s) = \mathcal{M}_{u_0}(s),\ K(s) := \mathcal{M}_{k_0}(s).$

$$\frac{\partial}{\partial t} U(t,s) + U(t,s) = K(s)\, U(t,s)$$

$$\Rightarrow U(t,s) = U_0(s)\, e^{(K(s)-1)t}$$

Formally (assumptions on k_0, u_0 and $\nu \in \mathbb{R}$ required)

$$\boxed{u(t,x) = \frac{1}{2\pi i} \int_{\nu - i\infty}^{\nu + i\infty} U_0(s)\, e^{(K(s)-1)t} x^{-s} ds}$$

Mellin transform for the fragmentation equation

$$\mathcal{M}_f(s) = \int_0^\infty x^{s-1} f(x)dx$$

The Mellin transform is the Fourier transform in $y = \log x$

Denote $U(t,s) := \mathcal{M}_{u(t,\cdot)}(s),\ U_0(s) = \mathcal{M}_{u_0}(s),\ K(s) := \mathcal{M}_{k_0}(s).$

$$\frac{\partial}{\partial t} U(t,s) + U(t,s) = K(s)\,U(t,s)$$

$$\Rightarrow U(t,s) = U_0(s)\,e^{(K(s)-1)t}$$

Formally (assumptions on k_0, u_0 and $\nu \in \mathbb{R}$ required)

$$\boxed{u(t,x) = \frac{1}{2\pi i} \int_{\nu - i\infty}^{\nu + i\infty} U_0(s)\,e^{(K(s)-1)t} x^{-s} ds}$$

Nice formula... But asymptotically?...

Mellin transform and self-similar profiles

$$\boxed{u(t,x) = \frac{1}{2\pi i} \int_{\nu - i\infty}^{\nu + i\infty} U_0(s)\, e^{(K(s)-1)t} x^{-s} ds}$$

(as for the case $\gamma > 0$): does there exist Φ s.t.

$$f(t)\Phi\big(xg(t)\big)$$

is a solution to (10) and so that, for any u^{in},

$$u(t,x) \approx_{t\to\infty} f(t)\Phi\big(xg(t)\big) \quad ?$$

Mellin transform and self-similar profiles

$$u(t,x) = \frac{1}{2\pi i} \int_{\nu - i\infty}^{\nu + i\infty} U_0(s)\, e^{(K(s)-1)t} x^{-s} ds$$

(as for the case $\gamma > 0$): does there exist Φ s.t.

$$f(t)\Phi\big(xg(t)\big)$$

is a solution to (10) and so that, for any u^{in},

$$u(t,x) \approx_{t\to\infty} f(t)\Phi\big(xg(t)\big) \quad ?$$

Proposition

If we look for $\Phi \in L^1\big((1+x)dx\big)$, no such solution.

But for all $s > p_1$, pointwise self-similar solutions are given by

$$e^{(K(s)-1)t} x^{-s} = \exp((K(s)-1)t - s\log(x)) := \exp(\phi(s,t,x))$$

First step: integration domain for the Mellin transform

$$[1,2] \subset I(u_0) := \left\{ p \in \mathbb{R}; \;\; U_0(p) = \int_0^\infty u_0(x) x^{p-1} dx < \infty \right\} := (p_0, q_0).$$

$$p_0 := \inf I(u_0), \quad q_0 := \sup I(u_0), \quad p_1 := \inf I(k_0) < 2.$$

$$u_0 \approx_0 x^{-p_0}, \qquad u_0 \approx_{+\infty} x^{-q_0}$$

Proposition

For $p_1 := \inf I(k_0) < 2$*,* $\exists !$ *sol. to* (10)*,* $\forall \max(p_0, p_1) < \nu < q_0$*:*

$$u(t,x) = \frac{1}{2\pi i} \int_{\nu - i\infty}^{\nu + i\infty} U_0(s)\, e^{-s \log(x) + t(K(s)-1)} ds.$$

First step: integration domain for the Mellin transform

$$[1,2] \subset I(u_0) := \left\{ p \in \mathbb{R};\ \ U_0(p) = \int_0^\infty u_0(x) x^{p-1} dx < \infty \right\} := (p_0, q_0).$$

$$p_0 := \inf I(u_0), \quad q_0 := \sup I(u_0), \quad p_1 := \inf I(k_0) < 2.$$

$$u_0 \approx_0 x^{-p_0}, \qquad u_0 \approx_{+\infty} x^{-q_0}$$

Proposition

For $p_1 := \inf I(k_0) < 2$, $\exists !$ *sol. to* (10), $\forall \max(p_0, p_1) < \nu < q_0$*:*

$$u(t,x) = \frac{1}{2\pi i} \int_{\nu - i\infty}^{\nu + i\infty} U_0(s)\, e^{-s \log(x) + t(K(s)-1)} ds.$$

For $\nu > 2$ and x fixed: $K(\nu) - 1 < 0 \Rightarrow$ exponential decay $t \to \infty$.
But which exponential rate? And when $t \to \infty$ and $x \to 0$?

Main idea: study $\phi(s,t,x)$

$u(t,x) = \frac{1}{2\pi i} \int_{\nu - i\infty}^{\nu + i\infty} U_0(s)\, e^{\phi(s,t,x)} ds$ with $\phi(s,t,x) = -s\log(x) + t(K(s) - 1)$

$s \in \mathbb{R} \to \phi(s,t,x)$ is convex: define for $x < 1$

$$s_+(t,x) := \arg \min_{s \in (p_0, q_0)} \phi(s,t,x) = K'^{-1}\left(\frac{\log(x)}{t}\right)$$

In the zone $s_+(t,x) > q_0$: $\Rightarrow \phi(s_+,t,x) < \phi(q_0,t,x)$

$\Rightarrow$ Steps for $s_+ > q_0$ or $s_+ < p_0$:

- move to the residue q_0
- cross it: residue theorem (+ extra regularity assumptions)
- evaluate the rest as small since $\Re(\phi(s_+,t,x)) < \Re(\phi(q_0,t,x))$

The zones of convergence

Example: mitosis kernel

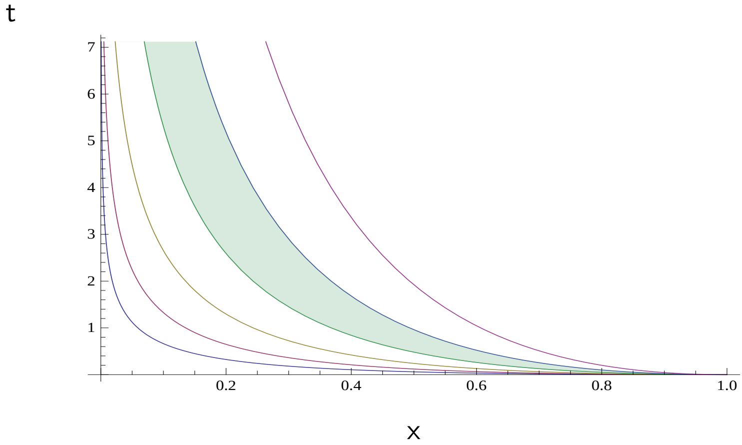


Figure: Different curves of the form $s_+ = \gamma$ for different values of $\gamma > 0$, so that $2t = -\gamma^2 \log x$. As $t \to \infty$, the function $xu(t,x)$ concentrates in the interval $x \in \left(e^{-\frac{2t}{\gamma_\ell^2}}, e^{-\frac{2t}{\gamma_r^2}} \right)$. .

Numerical Illustration

Example: mitosis kernel

$$\frac{\partial}{\partial t} n(t,y) + n(t,y) = 4n(t, y + \log 2), \quad n(0,y) = n^{in}(y).$$

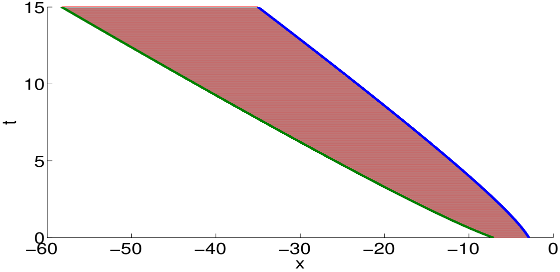


Figure: solution in a log-scale. Inside the blue and green curves, $u(t,x) \geq 10\% max_x u(t,\cdot)$.

Case $x > e^{tK'(q_0)}$

$$u(t,x) = \frac{1}{2\pi i} \int_{\nu - i\infty}^{\nu + i\infty} U_0(s)\, e^{(K(s)-1)t} x^{-s} ds.$$

Theorem

As $t \to \infty$ and $q_0 < s_+(t,x)$:

$$u(t,x) = a_0 x^{-q_0} e^{(K(q_0)-1)t} \left(1 + \mathcal{O}\left(x^{-\nu'+q_0} e^{(K(\nu')-K(q_0))t}\right)\right).$$

for a $\nu' > q_0$.

Case $x > e^{tK'(q_0)}$

$$u(t,x) = \frac{1}{2\pi i} \int_{\nu - i\infty}^{\nu + i\infty} U_0(s)\, e^{(K(s)-1)t} x^{-s} ds.$$

Theorem

As $t \to \infty$ and $q_0 < s_+(t,x)$:

$$u(t,x) = a_0 x^{-q_0} e^{(K(q_0)-1)t} \left(1 + \mathcal{O}\left(x^{-\nu'+q_0} e^{(K(\nu')-K(q_0))t}\right)\right).$$

for a $\nu' > q_0$.

$\Longrightarrow$ Rate of convergence: exponential.

case $e^{tK'(p_0)} < x < e^{tK'(q_0)}$

$$u(t,x) = \frac{1}{2\pi i} \int_{\nu - i\infty}^{\nu + i\infty} U_0(s)\, e^{\phi(s,t,x) - t} ds.$$

case $e^{tK'(p_0)} < x < e^{tK'(q_0)}$

$$u(t,x) = \frac{1}{2\pi i} \int_{\nu - i\infty}^{\nu + i\infty} U_0(s)\, e^{\phi(s,t,x) - t} ds.$$

$\Rightarrow$ steps:

- Choose $\nu = s_+(t,x)$
- Method of the stationary phase to localize the dominant contribution in the integral

case $e^{tK'(p_0)} < x < e^{tK'(q_0)}$

$$u(t,x) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} U_0(s_+ + iv)\, e^{\phi(s_+ + iv, t, x) - t} dv.$$

Lemma

$\Re e\big(\phi(s_+ + iv, t, x)\big)$ maximal iff

- *$v = 0$ if k_0 has an absolutely continuous part,*

case $e^{tK'(p_0)} < x < e^{tK'(q_0)}$

$$u(t,x) = \frac{1}{2\pi i}\int_{-\infty}^{\infty} U_0(s_+ + iv)\, e^{\phi(s_+ + iv,t,x) - t} dv.$$

Lemma

$\Re e\big(\phi(s_+ + iv, t, x)\big)$ maximal iff

- *$v = 0$ if k_0 has an absolutely continuous part,*
- *for $k_0(z) = 2\delta_{s=\frac{1}{2}}$:*

case $e^{tK'(p_0)} < x < e^{tK'(q_0)}$

$$u(t,x) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} U_0(s_+ + iv)\, e^{\phi(s_+ + iv,t,x) - t} dv.$$

Lemma

$\Re e\big(\phi(s_+ + iv, t, x)\big)$ maximal iff

- *$v = 0$ if k_0 has an absolutely continuous part,*
- *for $k_0(z) = 2\delta_{s=\frac{1}{2}}$: $v \in \frac{2\pi}{\log 2}\mathbb{Z}$*

(more complex probability measures also dealt with, but not all...)

Case $e^{tK'(p_0)} < x = e^{tK'(s_+)} < e^{tK'(q_0)}$

& $v = 0$ only max of $\Re e(\phi)$

Theorem

For any $\delta > 0$, for $p_0 + \delta < s_+(t, x) < q_0 - \delta$ and $t \to \infty$, we have

$$u(t, x) = \frac{U_0(s_+)x^{-s_+}e^{(K(s_+)-1)t}}{\sqrt{2\pi t K''(s_+)}} + O(t^{-\frac{1}{2}-\alpha}),$$

for $\alpha > 0$ well chosen.

$\Rightarrow$ Rate of convergence: at most polynomial.

Case $e^{tK'(p_0)} < x = e^{tK'(s_+)} < e^{tK'(q_0)}$ & $k_0 = 2\delta_{z=\frac{1}{2}}$

Same analysis around each $s_k = s_+ + \frac{2i\pi k}{\log 2}$.

Theorem (MD, M. Escobedo)

For any $\delta > 0$, for $p_0 + \delta < s_+(t,x) < q_0 - \delta$ and $t \to \infty$, we have

$$u(t,x) = x^{-s_+(t,x)} e^{(K(s_+(t,x))-1)t} \frac{\sum_{k\in\mathbb{Z}} U_0(s_k) x^{\frac{2i\pi k}{\log 2}}}{\sqrt{2\pi t K''(s_+)}} + \cdots ,$$

Poisson summation formula:

$$u(t,x) \sim \log 2 \frac{e^{(K(s_+)-1)t}}{\sqrt{2\pi t K''(s_+)}} \sum_{n\in\mathbb{Z}} u_0(2^{-n}x).$$

$\Rightarrow$ Rate of convergence: at most polynomial.

Comparison with (Bertoin, 2003)

see also (Bertoin, Watson, 2016)

Stochastic process $X = (X(t), t \geq 0)$, values in $\mathcal{S}^{\downarrow}(y)$ set of all sequences $Y = (y_i)_{i \in \mathbb{N}^*}$ such that

$$y_1 \geq ... \geq y_i \geq y_{i+1} \geq ... \geq 0 \quad \text{and} \quad y = \sum_{i=1}^{\infty} y_i \leq 1,$$

Random measure $\rho_t(dy)$ defined by

$$\rho_t(dy) = \sum_{i=1}^{\infty} X_i(t) \delta_{\frac{1}{t} \log X_i(t)}(dy)$$

converges to $\delta_{-\mu}$ in probability for some $\mu < \infty$.
$\tilde{\rho}_t$ image of ρ_t by $x \to \sqrt{t}(x + \mu)/\sigma$
converges in probability to the standard normal distribution $\mathcal{N}(0, 1)$.

Comparison with (Bertoin, 2003)

see also (Bertoin, Watson, 2016)

The laws of $\rho_t(dy)$ and $\tilde{\rho}_t(dy)$ correspond to rescalings of u:

$$r(t,y) := tye^{2ty}u(t,e^{ty}), \qquad \tilde{r}(t,z) := r(t, y_0 + \frac{\sigma z}{\sqrt{t}})\frac{\sigma}{\sqrt{t}},$$

with $y_0 := K'(2)$ and $\sigma^2 := K''(2)$.
Under previous assumptions we prove

$$r(t,\cdot) \rightharpoonup \delta_{K'(2)}U_0(2), \qquad \tilde{r}(t,\cdot) \rightharpoonup U_0(2)G,$$

with $G(z) = \frac{e^{-\frac{\cdot^2}{2}}}{\sqrt{2\pi}}$, in the weak sense of measures.

Fragmentation + binary fission: oscillations?

with Bruce van Brunt

Dirac kernel: an explicit formula

with B. van Brunt

Here

$$k_0 = 2\delta_{x=\frac{1}{2}} \implies K(s) = 2^{2-s},$$

For $x = e^{-tK'(s_+)}$ with $-K'(p_0) < K'(s_+) < -K'(q_0)$:

$$u(t,x) \sim \log 2 \frac{e^{(K(s_+)-1)t}}{\sqrt{2\pi t K''(s_+)}} \sum_{n\in\mathbb{Z}} u_0(2^n x).$$

Dirac kernel: an explicit formula

with B. van Brunt

Here

$$k_0 = 2\delta_{x=\frac{1}{2}} \implies K(s) = 2^{2-s},$$

For $x = e^{-tK'(s_+)}$ with $-K'(p_0) < K'(s_+) < -K'(q_0)$:

$$u(t,x) \sim \log 2 \frac{e^{(K(s_+)-1)t}}{\sqrt{2\pi t K''(s_+)}} \sum_{n\in\mathbb{Z}} u_0(2^n x).$$

Direct formula:

$$u(t,x) = e^{-t} \sum_{k=0}^{\infty} u_0(2^k x) \frac{(4t)^k}{k!}$$

"oscillations" in these formulae?

Dirac kernel: oscillations?

Recall (Bertoin, 2003):

$$r(t,y) := tye^{2ty}u(t,e^{ty}), \qquad \tilde{r}(t,z) := r(t, y_0 + \frac{\sigma z}{\sqrt{t}})\frac{\sigma}{\sqrt{t}},$$

with $y_0 := K'(2) = -\log 2$ and $\sigma^2 := K''(2) = (\log 2)^2$.
Under previous assumptions we prove

$$r(t,\cdot) \rightharpoonup \delta_{K'(2)} U_0(2), \qquad \tilde{r}(t,\cdot) \rightharpoonup U_0(2) G,$$

with $G(z) = \frac{e^{-\frac{\cdot^2}{2}}}{\sqrt{2\pi}}$, in the weak sense of measures.

not contradictory with oscillations: weak convergence

Dirac kernel: oscillations?

with B. van Brunt

$$\frac{r(t,y_0)}{-\log 2} = t2^{-2t}u(t,2^{-t}) = \sqrt{\frac{t}{2\pi}}\sum_{k\in\mathbb{Z}} U_0(2+\frac{2ik\pi}{\log 2})e^{-2i\pi kt}\left(1+o(t^{-\beta})\right)$$

$\Longrightarrow$ oscillations for $\frac{r}{\sqrt{t}}$ of period 1.

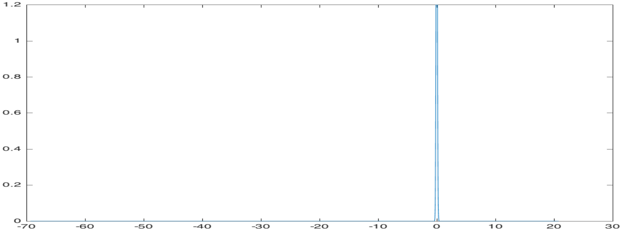


Illustration: $\sqrt{t}n(t,y)$ with $n(t,y) = e^{2y}u(t,e^y)$ solution to

$$\frac{\partial}{\partial t}n(t,y) + n(t,y) = n(t,y+\log 2), \qquad n(0,y) = n_0(y)$$

Series representation of the solution

The fundamental solution $U \in \mathcal{M}(\mathbb{R}^+)$ with initial data $u_0 = \delta(x-1)$:

$$U = e^{-\alpha t}\delta(x-1) + \sum_{n=0}^{\infty}(\alpha t)^n a_n; \qquad a_0(x) = 0,$$

$$a_{n+1}(x) = \frac{1}{n+1}\left(-x^\gamma a_n(x) + \int_0^1 (\frac{x}{z})^\gamma a_n(\frac{x}{z})\frac{k_0(dz)}{z} + k_0(x)\frac{(-1)^n}{n!}\right).$$

The series is convergent in the TV norm for measures: with

$$||a_n||_{TV} \leq u_n = \frac{1}{n!}\sum_{j=0}^{n-1} 3^{n-j}(-1)^j,\ \forall n \geq 1$$

Proof of the TV convergence result

We have

$$\frac{U - e^{\alpha t}\delta(x-1)}{\alpha t} - k_0 = \frac{\sum_{n=1}^{\infty}(\alpha t)^n a_n}{\alpha t} - k_0$$

$$= \sum_{n=1}^{\infty}(\alpha t)^{n-1} a_n - k_0 = \sum_{n=0}^{\infty}(\alpha t)^n a_{n+1} - k_0$$

and since $a_1 = k_0$, we have

$$\sum_{n=0}^{\infty}(\alpha t)^n a_{n+1} - k_0 = \sum_{n=1}^{\infty}(\alpha t)^n a_{n+1} = \alpha t \sum_{n=0}^{\infty}(\alpha t)^n a_{n+2}.$$

Thus

$$\left\| \frac{U - e^{-\alpha t}\delta(x-1)}{\alpha t} - k_0 \right\|_{TV} \leq \alpha t \sum_{n=0}^{\infty}(\alpha t)^n \|a_{n+2}\|_{TV}.$$

Proof of the TV convergence result

The series converges (normal convergence) and thus it is bounded on any compact set, for instance for $t \in [0, T]$. Then the result holds for

$$K_{T,\alpha} = \alpha \max_{t \in [0,T]} \sum_{n=0}^{\infty} (\alpha t)^n \|a_{n+2}\|_{TV}.$$

By simple scaling:

Corollary 1 If U_λ is the solution with initial data $U_\lambda(0) = \delta(x - \lambda)$, for $t \in [0, T]$ and for some $K > 0$ depending on T, α, γ

$$\left\| \frac{U_\lambda(t) - e^{-\alpha t \lambda^\gamma} \delta(x - \lambda)}{\alpha t \lambda^\gamma} - \frac{1}{\lambda} k_0 \left(\frac{x}{\lambda}\right) \right\|_{TV} \leq K t \lambda^\gamma.$$

Corollary 2 If u is the solution with initial data u_0, for $t \in [0, T]$ and for some $K > 0$ depending on T, α, γ and $\|u_0\|_{L^1(\ell^{2\gamma} d\ell)}$

$$\left\| \frac{u(t) - e^{-\alpha t x^\gamma} \mu_0}{\alpha t} - \kappa * \mu_0 \right\|_{TV} \leq K t$$

Here $*$ denotes the multiplicative/Mellin convolution

Proof of the stability result in BL norm

Remember the hypothesis:

$$\|(u_{q,0} - \delta(x-1))\|_{BL} \leq q$$
$$\|u_{q,0,\varepsilon_0} - u_{q,0}\|_{BL} \leq \varepsilon_0, \quad \|u_{q,\varepsilon}(t) - u_q(t)\|_{BL} \leq \varepsilon$$

Then,

$$\left\| \frac{u_{q,\varepsilon}(t) - e^{-\alpha t} u_{q,0,\varepsilon_0}}{\alpha t} - k_0 \right\|_{BL} \leq \frac{\|u_{q,\varepsilon}(t) - u_q(t)\|_{BL}}{\alpha t} + \\ + \frac{\|u_q(t) - U(t)\|_{BL}}{\alpha t} + \left\| \frac{U(t) - e^{-\alpha t}\delta(x-1)}{\alpha t} - k_0 \right\|_{BL} + \\ + e^{-\alpha t} \frac{\|\delta(x-1) - u_{q,0}\|_{BL}}{\alpha t} + e^{-\alpha t} \frac{\|u_{q,0} - u_{q,0,\varepsilon_0}\|_{BL}}{\alpha t}$$

By the TV Theorem: $$\left\| \frac{U(t) - e^{-\alpha t}\delta(x-1)}{\alpha t} - k_0 \right\|_{BL} \leq Kt$$

Proof of the stability result in BL norm

For the last remaining term:

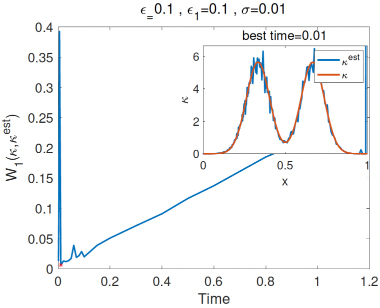
$$\|u_q(t) - u(t)\|_{BL} \leq C\|u_{q,0} - u_0\|_{BL},\ \forall \gamma \in (0,1],$$
$$\|u_q(t) - u(t)\|_{BL} \leq C\|u_{q,0} - u_0\|_{BL},\ \forall \gamma \geq 1$$

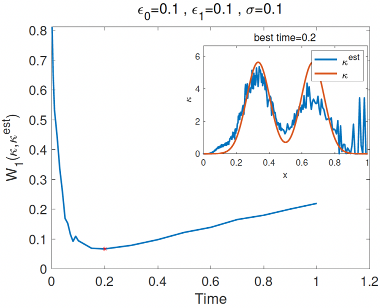
using the following.

Proposition There exists a constant $C > 0$ such that, for all bounded measure u_0 compactly supported in $[0, M]$, and either $Supp(u_0) \subset [m, M]$ with $m > 0$ or $\gamma \geq 1$, the weak solution u of the fragmentation equation satisfies, for all $t \in [0, T]$,

$$\|u(t)\|_{BL} \leq C(M, T)\|u_0\|_{BL}.$$

Numerical simulations





Extensions of the model

Variability:

$$\frac{\partial}{\partial t} n(t,x,v) + \frac{\partial}{\partial x}\big(vxn(t,x,v)\big) =$$

$$-B(x)n(t,x,v) + 2\int_x^\infty \int_0^\infty B(y)k(y,x)\rho(v',v)n(t,y,v')dy, dv'$$

with $\int_0^\infty \rho(v',v)dv = 1$

Extensions of the model

Variability:

$$\frac{\partial}{\partial t} n(t,x,v) + \frac{\partial}{\partial x}\big(vxn(t,x,v)\big) =$$

$$-B(x)n(t,x,v) + 2\int_x^\infty \int_0^\infty B(y)k(y,x)\rho(v',v)n(t,y,v')dy, dv'$$

with $\int_0^\infty \rho(v',v)dv = 1$

Age + variability:

$$\frac{\partial}{\partial t} n(t,a,x,v) + \frac{\partial}{\partial x}\big(vxn(t,a,x,v)\big) = -B(a,x)n(t,a,x,v),$$

$$n(t,a=0,x,v) = 2\int_x^\infty \int_0^\infty B(a,y)k(y,x)\rho(v',v)n(t,a,y,v')dy\,dv'\,da$$

(related (maturity) models: Lebowitz, Rubinow, 1977 - Rotenberg, 1983 - Mischler, Perthame, Ryzhik, 2002,...)

Incorporating variability

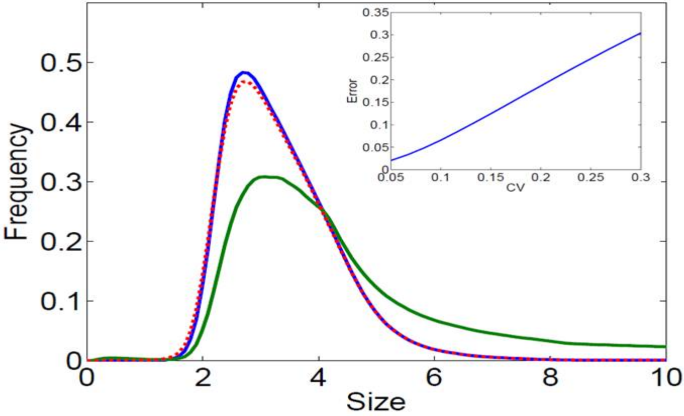


Figure: Effect on the distribution of growth rate variability

Incorporating variability

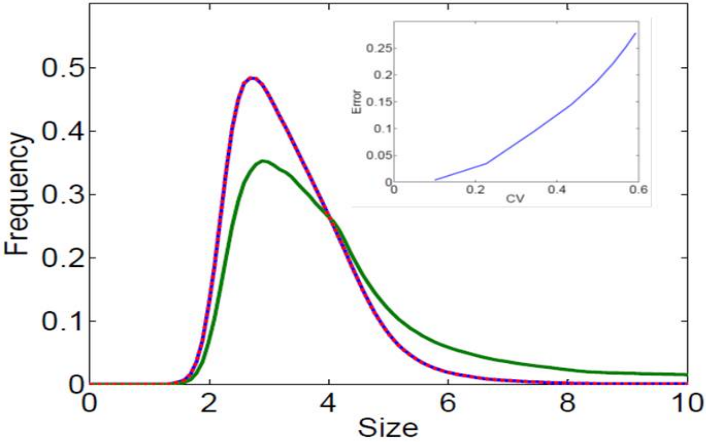


Figure: Effect on the distribution of variability in daughter sizes

Use the short time behaviour

First back to theory...

Hypothesis on k_0: contains no Dirac mass at $x = 0$ or $x = 1$, and

$$supp\,(k_0) \subset [0,1], \ \int_0^1 dk_0(z) < +\infty, \ \int_0^1 z dk_0(z) = 1.$$

Weak solution:
A family $(u(t))_{t\geq 0} \subset \mathcal{M}(\mathbb{R}^+)$ is called a measure solution with initial data $u_0 \in \mathcal{M}(\mathbb{R}^+)$ if for all $\varphi \in \mathcal{C}_c^0(\mathbb{R}^+)$ and all $t \geq 0$, $t \mapsto \int \varphi(x) u(t,dx)$ is continuous and

$$\int_{\mathbb{R}^+} \varphi(x) u(t,dx) = \int_{\mathbb{R}^+} \varphi(x) u_0(dx)$$
$$+ \alpha \int_0^t \int_{\mathbb{R}^+} \left(-x^\gamma \varphi(x) u(s,dx) + \int_0^1 \varphi(xz) x^\gamma k_0(dz) u(s,dx) \right) ds.$$

Existence and uniqueness ($\gamma > 0$) in $\mathcal{M}_+(\mathbb{R}^+)$ in Carrillo & al. 2012.