Recovering the History of a Fibril Population Undergoing Fragmentation from Its Asymptotic Profile

Magali Tournus, Institut Mathématique de Marseille
Collaborators : Marie Doumic, Miguel Escobedo and Wei Feng Xue

CIMPA School Mathematical models in biology and medicine
Mauritius - 2016

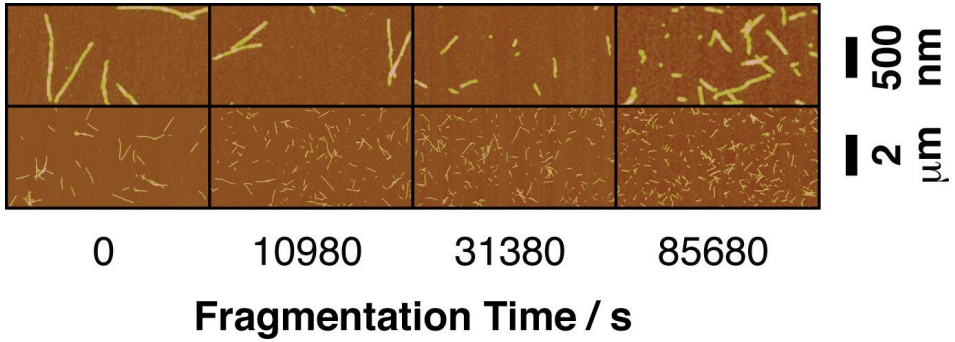
Goal

Understand the dynamical behaviour of a system of fibrils undergoing a mechanichal fragmentation.

Context : Diseases involving Amyloid fibrils (Prion, Alzeihmer).

Tools : Deterministic models of the fragmentation mechanism.

Data : Experiment performed at the University of Kent, UK, by the team of Prof Xue.



A fragmentation equation

Deterministic equation describing the evolution of a population of particles with characteristic size $x \rightsquigarrow$ Size-structured PDE.

Unknown: f(t,x) = Density of particles of size x at time t.

$$\underbrace{\frac{\partial f}{\partial t}(t,x)}_{\text{Evolution of the number of particles}} = \underbrace{-B(x)f(t,x)}_{\text{Death}} + \underbrace{\int_{y=x}^{y=\infty} k(x,y)B(y)f(t,y)dy}_{\text{Creation}}$$

Parameters:

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

$$supp\,(k_0) \subset [0,1], \quad \int_0^1 dk_0(z) = 2, \quad \int_0^1 z dk_0(z) = 1.$$

Related size-structured PDE

- **Pure Fragmentation.** Particles can only break up.
  Example: Amyloid fibrils

$$\frac{\partial f}{\partial t}(t,x) = -B(x)f(t,x) + \int_{y=x}^{y=\infty} k(x,y)B(y)f(t,y)dy$$

- **Transport-Fragmentation.** Particles can break up and grow.
  Example: Cells, microtubules

$$\frac{\partial f}{\partial t}(t,x) + \frac{\partial}{\partial x}\left(g(x)f(t,x)\right) = -B(x)f(t,x) + \int_{y=x}^{y=\infty} k(x,y)B(y)f(t,y)dy$$

- **Coagulation-Fragmentation.** Particles can break up or merge.
  Example: Dispersion of dust, smoke, or pollutants.

$$\frac{\partial f}{\partial t}(t,x) = -B(x)f(t,x) + \int_{y=x}^{y=\infty} k(x,y)B(y)f(t,y)dy$$
$$-f(t,x)\int_0^{+\infty} k_c(x,y)f(y)dy + \int_{y=0}^{y=x} k_c(y,x-y)f(t,y)f(t,x-y)dy$$

Existence and basic properties

The fragmentation equation :

$$\begin{cases} \dfrac{\partial f}{\partial t}(t,x) = -B(x)f(t,x) + \displaystyle\int_{y=x}^{y=\infty} \frac{B(y)}{y} k_0\left(\frac{x}{y}\right) f(t,y)dy, \\ f(x,0) = f_0(x). \end{cases} \tag{1}$$

Under the assumptions on the fragmentation rate ($B(x) = x^\gamma$) and on the scaling of the kernel, **the fragmentation equation has a global solution** (Smith, Thieme, Escobedo, Michel, Perthame, Mishler, ...) which lies in $\mathcal{C}([0,T); L^1(\mathbb{R}^+, xdx)) \cap L^1(0,T; L^1(\mathbb{R}^+, x^{\gamma+1}dx)$.

Conservation properties :

$$\frac{d}{dt}\int_0^{+\infty} f(t,x)dx = \int_0^{+\infty} B(x)f(t,x)dx \qquad \text{Number of clusters increases}$$

$$\frac{d}{dt}\int_0^{+\infty} xf(t,x)dx = 0 \qquad \text{Mass conservation}$$

A Non-preserving mass solution (Stewart)

Uniqueness is true only in $\mathcal{C}([0,T);L^1(\mathbb{R}^+,xdx))\cap L^1(0,T;L^1(\mathbb{R}^+,x^{\gamma+1}dx)$. The fact that the solution is in $L^1(0,T;L^1(\mathbb{R}^+,xB(x)dx)$ is crucial.

For

$$B(x)=\frac{x}{2}(1+x)^{-r},\ r\in(0,1),\ k_0=2\mathbb{1}_{[0,1]},$$

and for the initial condition $f_0\in L^1(\mathbb{R}^+,xdx)\setminus L^1(\mathbb{R}^+,xB(x)dx)$ defined as

$$f_0(x)=\exp\left(-\int_0^x\frac{3-ru(1+u)^{-1}}{2\lambda(1+u)^r+u}du\right),$$

there is a solution

$$f(t,x)=\exp(\lambda t)f_0(x)$$

belonging to $\mathcal{C}([0,T);L^1(\mathbb{R}^+,(1+x)dx))$ and for which the total mass of the system increases exponentially fast.

Theorem (Michel-Perthame-Escobedo-Mishler-Ricard – 2005)

Under reasonable technical assumptions, for large time, the profile f tends to distribute according to the self-similar profile g In other words, $f(t,x)$ satisfies

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

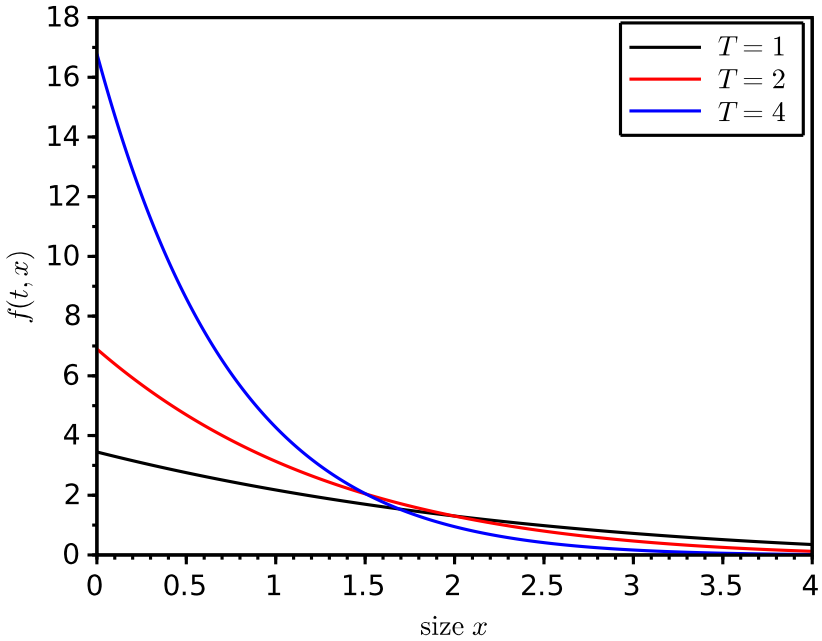
where the self-similar profile g is the unique solution of

$$z\frac{dg}{dz}(z) + (2+\alpha\gamma z^\gamma)g(z) = \alpha\gamma \int_{u=z}^{u=\infty} \frac{1}{u} k_0\left(\frac{z}{u}\right) u^\gamma g(u)\,du, \qquad \int_0^\infty zg(z)dz = \rho.$$

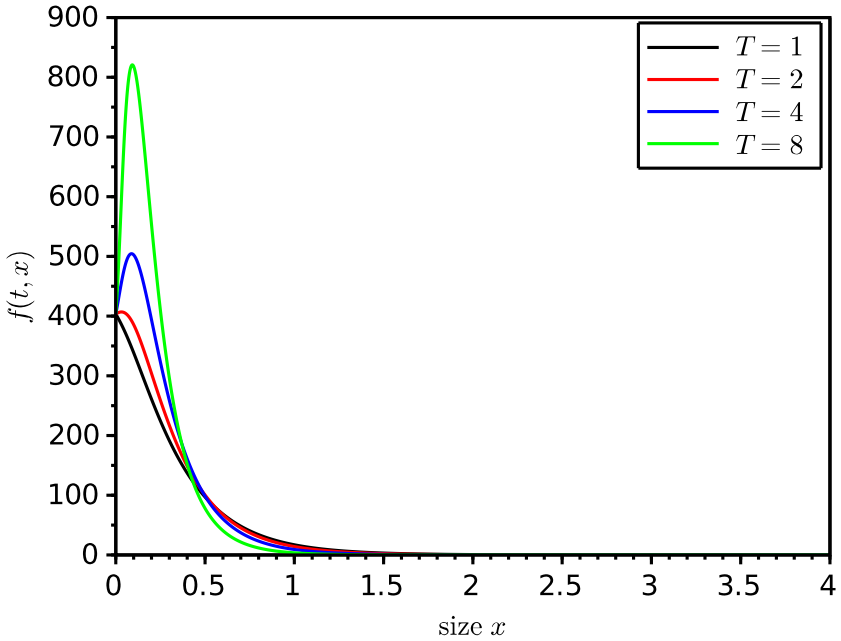
- No admissible stationary state. Change of coordinates.
- In the new coordinates, the fragmentation equation rewrites as a transport/fragmentation equation.
- The moments of f for large t are then related to the moments of g via the following formula

$$\int_0^\infty y^s g(y)dy \quad \sim \quad t^{(s-1)/\gamma} \int_0^\infty y^s f(t,y)dy, \qquad s \in [0,\infty),\ t \text{ large}.$$

Two examples.



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



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

Remark

Mitosis is not included in the model. We exclude $k_0(x) = 2\delta_{x=1/2}(x)$.

Formulation of the inverse problem

Inverse Problem (IP): Given a (noisy) measurement of $g(x)$ solution in $L^1((1+x^{\gamma+1})dx)$ of the stationary equation, is it possible to estimate the functional parameters of the evolution equation, namely the triplet $(\alpha, \gamma, k_0) \in (0, +\infty) \times (0, +\infty) \times \mathcal{M}^+(0,1)$?

A new approach: To understand the fragmentation of Amyloid fibrils, MEASURE the time-evolution of the SIZE DISTRIBUTION

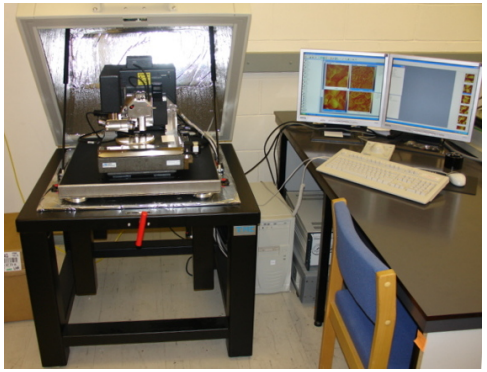
- **Proteins under study**: β2-Amyloid, Lysozyme, a-Synuclein, $\beta-$Lactoglobuline.
- Step 1. Prepare a pure sample of proteins.
- Step 2. **Recreating the experimental conditions of mechanical fragmentation.** The suspension is continuously agitated during two weeks by a magnetic agitator.

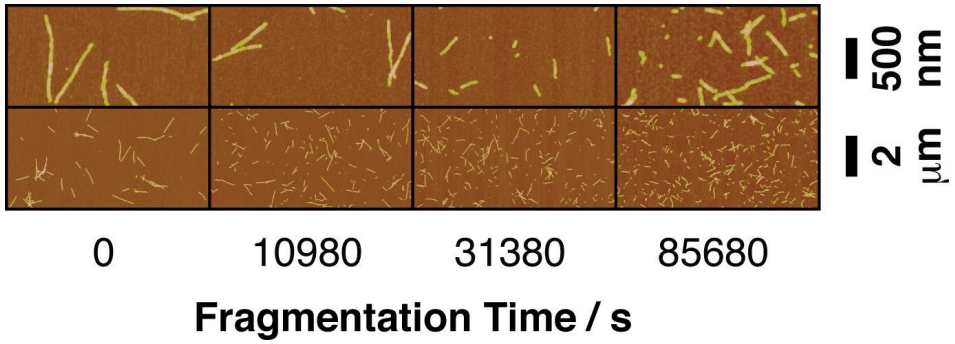
  From time to time, a small sample is taken out from the aliquot and saved.

- Step 3. A picture of each set taken out from the sample is done.

  Tool: **Atomic Force Microscopy (AFM)**.

  $\rightsquigarrow$ Fibrils can be identified visually.

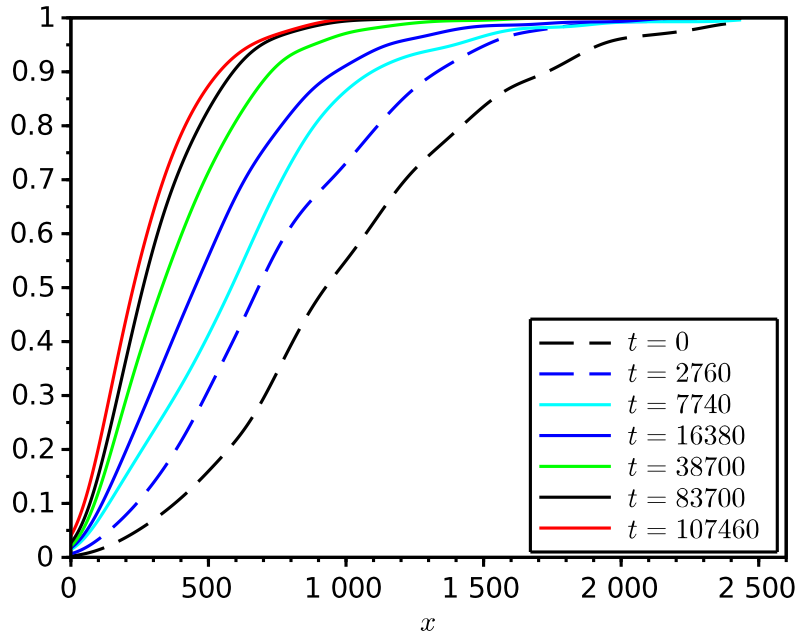


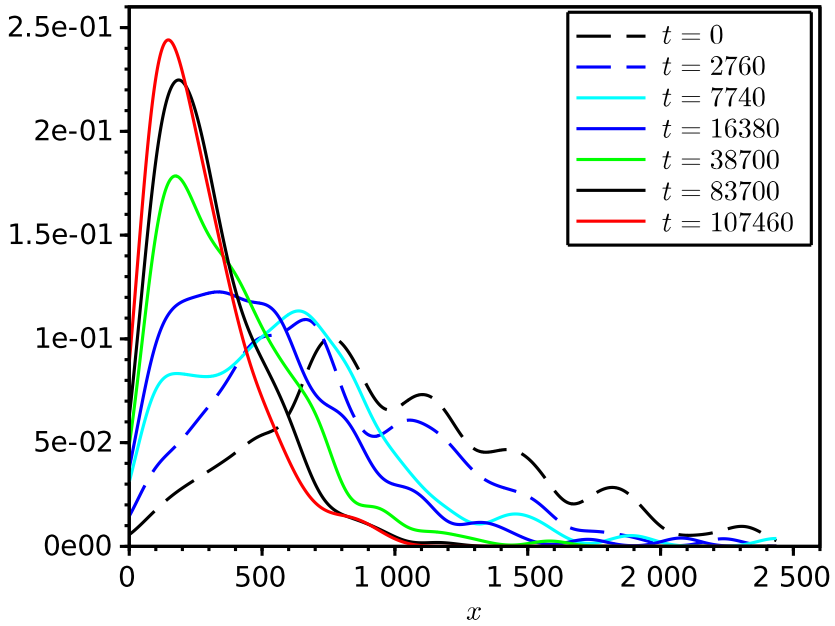


- Step 4. **Determine** the localization and a **measure of the size of each fibril of the picture.**
  Large fibrils are harder to detect than small fibrils:
  $\boxed{a}$ Fibrils touching the border.
  $\boxed{b}$ Fibrils which are crossing each other in a complicated way so that it is tricky to determine which is which.
  $\rightsquigarrow$ **It seems then that the number of large fibrils is underestimated.**

- Step 5. Estimate of much we lose and **modify the frequencies** by allocating a higher weight to the large sized fibrils. Very unusual technique since most experimentalists use directly the frequencies they measure as final data.

- We access **normalized cumulative distribution** function h of the length distribution of each fibril sample.





What is done by the Experimentalists (Pr Xue team)

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

- **Regularization of the data**. Approximate the data by polynomial functions.
- Make an **hypothesis on the fragmentation kernel** k_0: Parametrization.
  $\rightsquigarrow$ The problem becomes : Determine $\gamma, \alpha, k1, k2, k3, k4 \in \mathbb{R}^6$
- Solve the direct problem for the comprehensive set of admissible parameters $\gamma, \alpha, k1, k2, k3, k4 \in \mathbb{R}^6$.
- Determine which set of parameters gives rise to the dynamical behaviour observed experimentally. **Total linear least square analysis** (Minimization Problem).

Definition

Let μ be a measure over $\mathbb{R}$. We denote by $M[\mu]$ the Mellin transform of μ, defined by the integral

$$M[\mu](s) = \int_{x=0}^{x=+\infty} x^{s-1}\, d\mu(x),$$

for those values of s for which the integral exists.

Properties.

1. **Riemann-Lebesgue theorem.**
If $f \in L^1_{\text{loc}}(\mathbb{R})$, then $F(u+iv) \underset{v\to\pm\infty}{\to} 0$.

2. If $M[\mu](a)$ exists for some real number a,b then $M[\mu]$ is **holomorphic** in the open bande $\{s \in \mathbb{C} | a < Res < b\}$.

3. The **inverse-Mellin** transform is defined by

$$f(x) := \frac{1}{2i\pi} \int_{q^*-i\infty}^{q^*+i\infty} F(q^*+it) x^{-q^*} x^{-it}\, dt.$$

Assumptions

- $\exists \varepsilon > 0, \quad 0 < \eta_1 < \eta_2 < 1$ such that $\quad k_0(z) \geq \varepsilon, \quad z \in [\eta_1, \eta_2]$,
- There exists $\varepsilon > 0$ such that k_0 is a bounded continuous function on $[1-\varepsilon, 1]$ and on $[0, \varepsilon]$.
- There exists $s_- < 2$ such that $K_0(s) \neq 1$ for all $s \in \mathbb{C}$ such that $\Re e(s) \in (s_-, 2)$.
- There exists $h_0 \in L^1(0,1) \setminus \{0\}$, $n \geq 0$, $a_j \in (0,1)$ and $C_j \geq 0$, for $j = 1 \dots n$, such that

$$k_0(x) = h_0(x) + \sum_{j=1}^{n} C_j \delta_{a_j}(x),$$

where the a_j are such that there exists $m_j \in \mathbb{Q}^n$ and θ such that $a_j = \theta^{m_j}$, $j = 1 \dots n$.

The Mellin transform of g

Proposition

Suppose that a function g, satisfying

$$x^k g \in W^{1,1}(0,\infty)\ \forall k \geq 0,\ \ x^k g \in L^\infty(0,\infty),\ \forall k \geq 1;\ g \in W^{1,\infty}_{loc}(0,\infty)$$

is a solution of the stationary fragmentation equation in the sense of distributions on $(0,\infty)$, for some kernel $k_0 \in \mathcal{M}^+(0,1)$. Then, its Mellin transform:

$$G(s) = \int_0^\infty x^{s-1} g(x) dx$$

is well defined for all $s \in \mathbb{C}$ such that $\Re e(s) \geq 1$, it is analytic in the domain

$$D_1 = \{s \in \mathbb{C}; \Re e(s) > 2\ \}$$

and it satisfies:

$$(2-s)G(s) = \alpha\gamma(K_0(s)-1)G(s+\gamma) \quad in\ \ D_1.$$

The stationary fragmentation equation in Mellin-coordinates: used to recover parameters

Define $G(s) := M[g](s)$, $K_0(s) := M[k_0](s) \quad (K_0(2) = 1)$.

Turn an integro-differential equation into a functional equation

$$\boxed{(2-s)G(s) + \alpha\gamma G(s+\gamma) = \alpha\gamma K_0(s) G(s+\gamma), \qquad s \in D_1.} \tag{2}$$

- **Determine γ.**

- **Determine α.** Plug $s = 2$.

$$\alpha = \frac{G(1)}{\gamma G(1+\gamma)}.$$

- **Determine K_0. ($\rightsquigarrow k_0$)**

$$K_0(s) = 1 + \frac{G(s)(2-s)}{\alpha\gamma G(s+\gamma)}, \qquad s \in D_1.$$

Theorem (Doumic,Escobedo, T.)

For all $g \in L^1(\mathbb{R}^+)$ such that for all $k \geq 0$, $\int_0^\infty y^k g(y)dy < \infty$, there exists at most one triplet $(\gamma, \alpha, k_0) \in \mathbb{R}^+ \times \mathbb{R}^+ \times \mathcal{M}(0,1)$ such that g is the solution of the stationary equation in the sense of distributions.

Proof of the uniqueness for (γ, α) :

Proof 1. Relies on the estimates obtained by Balague, Canizo, Gabriel: There exists a constant $C > 0$ and a power $p \geq 0$ such that

$$g(x) \underset{x\to\infty}{\sim} Cx^p e^{-\frac{\alpha}{\gamma}x^\gamma} \to \quad \log\left(\frac{1}{g}\right) \underset{x\to\infty}{\sim} \frac{\alpha}{\gamma}x^\gamma,$$

Proof 2. Relies on estimates we obtain for $G(s)$ when $\mathrm{Re}(s)$ is large.

$$G(s+\gamma) = \Phi(s)G(s), \qquad \Phi(s) = \frac{2-s}{\alpha\gamma(K_0(s)-1)}.$$

Both proofs relie strongly on computations on the moments of g, and their behaviour when the power of the moment tends to infinity.

Uniquenes of γ: Proof based on the Mellin transform

Proposition (Necessary condition for γ)

Suppose that $g \in L^1_{loc}(\mathbb{R}^+)$ such that $\int_{\mathbb{R}^+} x^k g(x)dx < +\infty$, $k \geq 1$, and g solves the stationary fragmentation equation for some parameters $\gamma > 0$, $\alpha > 0$ and some non negative measure k_0, compactly supported in $[0,1]$. Let G be the Mellin transform of g. Then, given any constant $R > 0$:

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

Idea of the proof: Equation satisfied by G is close, for s large, to equation

$$AG(s+\gamma) = sG(s), \qquad s \in \mathbb{R}, \qquad G(2) = \rho.$$

It has a unique analytical solution $\Gamma_{A,\gamma}(s) = K\left(\frac{\gamma}{A}\right)^{\frac{s}{\gamma}} \Gamma\left(\frac{s}{\gamma}\right)$,

$$\Gamma_{A,\gamma}(s) \sim K\sqrt{\frac{2\pi\gamma}{s}}\left(\frac{\gamma}{A}\right)^{\frac{s}{\gamma}}\left(\frac{s}{e\gamma}\right)^{\frac{s}{\gamma}}\left(1 + \underset{s\to\infty}{\mathcal{O}}\left(\frac{1}{s}\right)\right) = K\sqrt{2\pi\gamma}\, s^{-\frac{1}{2}} e^{\frac{s}{\gamma}(\log s - 1 - \log A)}$$

Proof of the uniqueness for k_0

Since $G(s+\gamma)$ is strictly positive for $s \in (-\gamma, +\infty)$ we deduce that for $s \in [1, \infty)$ we can define $K_0(s)$ by

$$K_0(s) := 1 + \frac{G(s)(2-s)}{\alpha\gamma G(s+\gamma)}, \qquad 1 \leq s < +\infty.$$

$$K_0(s) := \int_0^1 x^{s-1} k_0(x) dx = \int_0^\infty e^{-sy} \mathcal{K}(y) dy = \hat{\mathcal{K}}(s)$$

The Mellin transform K_0 can be interpreted as a Laplace transform $\hat{K}$ of the function $\mathcal{K}(y) := k_0(e^{-y})$ with $y \in (0, +\infty)$, (if k_0 is diffuse) or of the pushforward $T\#k_0$ of k_0 by a function T .

Using the uniqueness property of the Laplace transform for measures , we conclude that this defines uniquely $\mathcal{K}$ and thus k_0.

We have a constructive way to determine γ and α, but not k_0.

Idea : Define its Mellin transform using the fomula

$$K_0^{\text{est}} := 1 + \frac{(2-s)G(s)}{\alpha\gamma G(s+\gamma)}, \qquad \Re(s) \geq 2 \tag{3}$$

To define k_0 from this formula, 2 points need to be solved:

- be allowed to divide by $G(s+\gamma)$, *i.e.* prove that G does not cancel at least on a vertical strip of the complex plane,
- be allowed to define the inverse Mellin transform of K_0^{est}, and prove that this corresponds to the original k_0.

Theorem

Suppose that $g \in \cap_{k=1}^{\infty} L^1(x^k dx) \cap L^1_{loc}(0,\infty)$ is the unique solution of the stationary equation for some given parameters α and γ and where k_0 is a non negative measure, compactly supported in $[0,1]$. Let $G(s)$ be the Mellin transform of the function g . Then, there exists $s_0 > 0$ such that

$$(i) \quad |G(s)| \neq 0, \ \ \forall s \in \mathbb{C};\ \Re e(s) \in [s_0, s_0+\gamma],$$

$$(ii) \quad K_0(s) = 1 + \frac{(2-s)G(s)}{\alpha\gamma G(s+\gamma)}, \ \ \text{for } \Re e(s) = s_0$$

$$(iii) \quad k_0(x) = \frac{1}{2i\pi} \int\limits_{\Re e(s)=s_0} x^{-s}\left(1 + \frac{(2-s)G(s)}{\alpha\gamma G(s+\gamma)}\right) ds.$$

Direct problem. Goal : check that G does not cancel.

Based on the Wiener-Hopf method, we can exhibit an explicit solution to the functional equation

$$\boxed{G(s+\gamma) = \Phi(s)G(s), \qquad \Phi(s) = \frac{2-s}{\alpha\gamma(K_0(s)-1)}, \; G(2) = \rho.} \tag{4}$$

Useful to

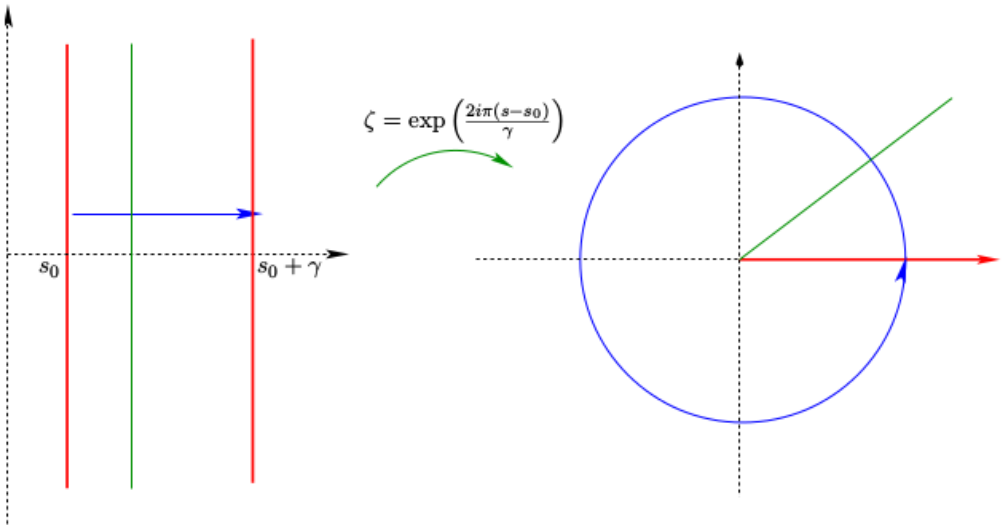
- check that the Mellin transform G of the function g never cancels,
- provide estimates on the inverse problem and derive a stability result.

Not useful to

- invert the problem, and to get the parameters from the solution.

$$\boxed{\text{Solve} \qquad G(s+\gamma) = \Phi(s)G(s), \qquad \Phi(s) = \frac{2-s}{\alpha\gamma(K_0(s)-1)}.} \tag{5}$$

$\boxed{1.}$ Transform the strip $\{z \in \mathbb{C} \mid s_0 < Im(z) < s_0 + \gamma\}$ into the exterior of $\mathbb{R}^+$



We denote by $\quad F(\zeta) = G(s), \qquad \tilde{\varphi}(\zeta) = \Phi(s), \qquad \zeta \in \mathbb{C} \setminus \mathbb{R}^+,$

The jump of the function F through the half line $\mathbb{R}^+$ satisfies

$$F(x - i0) = \varphi(x)F(x + i0), \qquad x \in \mathbb{R}^+,$$

where $\quad F(x+i0) := \lim_{\epsilon \to 0} F(xe^{i\epsilon}), \qquad F(x-i0) := \lim_{\epsilon \to 0} F(xe^{i(2\pi-\epsilon)}),$

$\boxed{2.}$ **Change of variable** to obtain a "typical" Carleman equation (additive),

$$P(x - i0) = \log(\varphi(x)) + P(x + i0), \qquad x \in \mathbb{R}^+.$$

we look for a solution of the shape

$$F(\zeta) = \exp(P(\zeta)), \qquad \zeta \in \mathbb{C} \setminus \mathbb{R}^+.$$

$\boxed{3.}$ **A candidate.** The function P is analytic in $\mathbb{C} \setminus \mathbb{R}^+$ and its jump through the half-line $\mathbb{R}^+$ is determined by the Carleman equation. A candidate for the function P is the function

$$P(\zeta) = -\frac{1}{2i\pi} \int_0^{+\infty} \log(\varphi(w)) \Big\{ \frac{1}{w - \zeta} - \frac{1}{w + 2} \Big\} \, dw, \qquad \zeta \in \mathbb{C} \setminus \mathbb{R}^+.$$

$\boxed{4.}$ **Back to F and G**

$$\overline{F}(\zeta) = \exp\Big(-\frac{1}{2i\pi} \int_0^{+\infty} \log(\varphi(w)) \Big\{ \frac{1}{w - \zeta} - \frac{1}{w + 2} \Big\} \, dw \Big) \rightsquigarrow \overline{G}. \qquad (6)$$

$\boxed{5.}$ **Check** that $\overline{G} \equiv G$.

Both functions $\overline{G}$ and $G =$ the Mellin transform of g are analytic and satisfy the equation (5), but nothing guarantees that $G = \overline{G}$.
Indeed, for $\Phi(s) = s$ and $\gamma = 1$, the functions

$$s \mapsto \frac{1}{2}\Gamma(s) \quad \text{and} \quad s \mapsto \frac{1}{2}\Gamma(s)\Big[1 + \sin(2\pi s)\Big]$$

are two distinct solutions.

Theorem

Let g be the solution to the stationary equation, and $\overline{G}$ the function defined in (6). *Then*

$$g(x) = \frac{1}{2i\pi}\int_{\Re e(s)=u} \overline{G}(s)x^{-s}ds, \ ,\forall u > s_0. \tag{7}$$

Determination of γ

Recall the Mellin transform of f

$$M[f(t,.)](s+1) = \int_0^\infty x^s f(t,x)dx.$$

Assuming that the distribution f has reached equilibrium for a given time t,

$$M[f(t,.)](s+1) = \int_0^\infty x^s t^{2/\gamma} g(t^{1/\gamma}x)dx = t^{(1-s)/\gamma} \int_0^\infty y^s g(y)dy, \qquad s \in \mathbb{R}.$$

Then, there exists a constant C_s such that

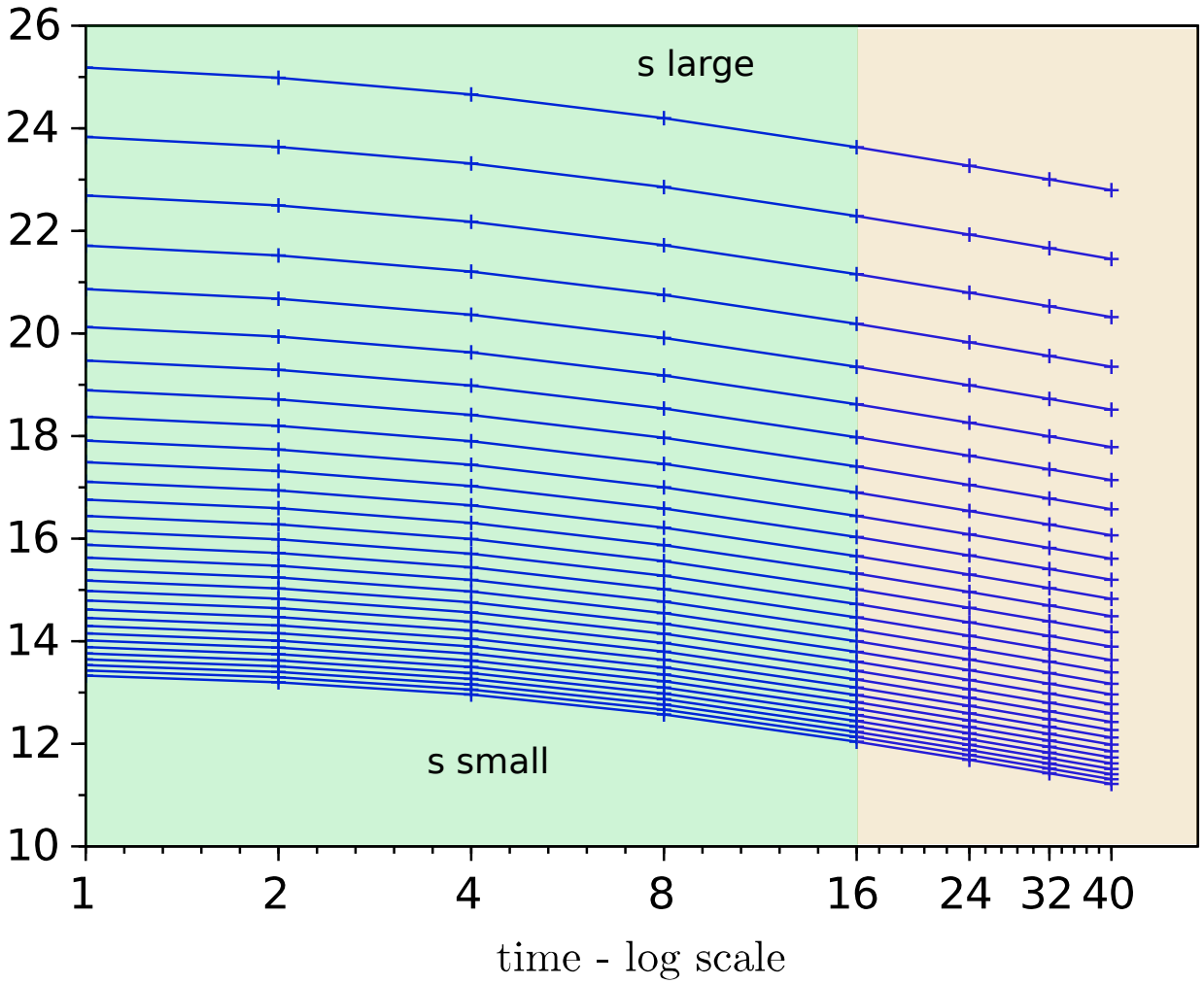
$$\log\left(M[f(t,.)](s+1)\right) = \frac{1-s}{\gamma}\log(t) + \log(C_s).$$

Then, we expect that for any given $s \in [0,+\infty]$, and for t large enough,

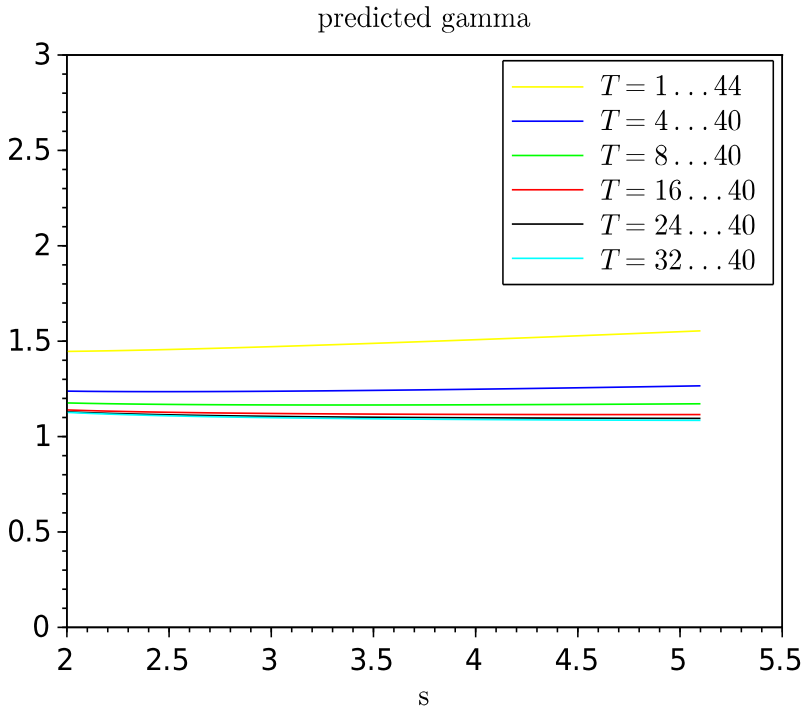
$$\text{the function} \quad I_s : log(t) \mapsto \frac{\log\left(M[f(t,.)](s+1)\right)}{1-s} \text{ has a slope of } 1/\gamma.$$

Simulated data. Determination of γ

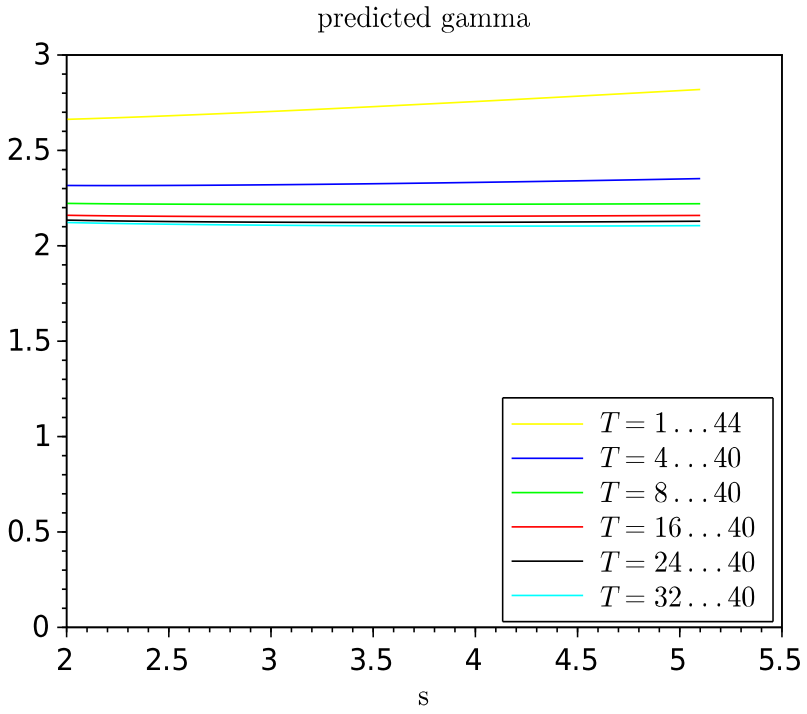
Plots of the function I_s at different times and for $s \in [1, 6.2]$.



Simulated data. Determination of γ



$\gamma = 1$



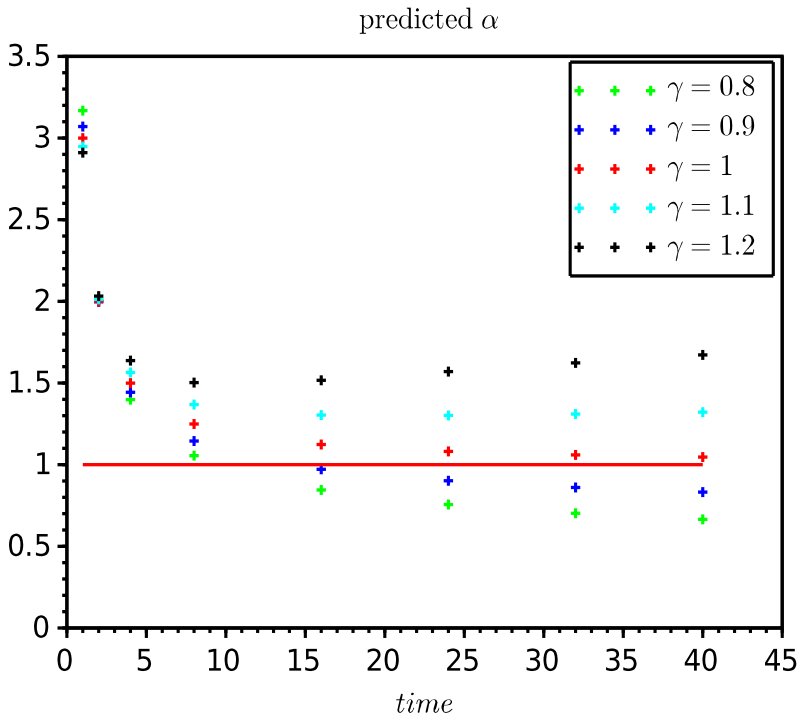
$\gamma = 2$

Simulated data. Determination of α

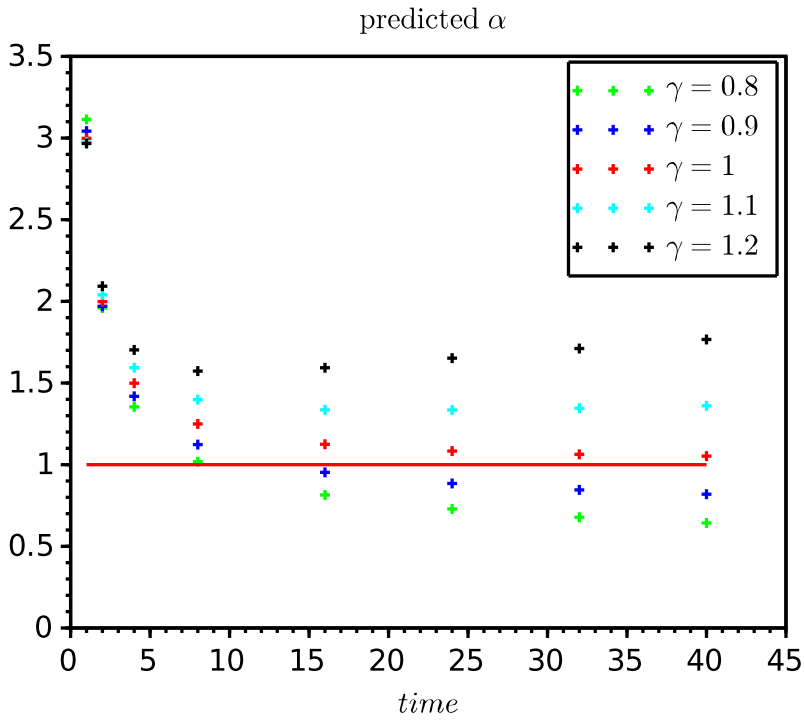
The value for α is directly obtained as a function of γ and G (which depends on γ) by using $K_0(1) = 2$

$$\alpha_{est} = \frac{G(1)}{\gamma_{est} G(1 + \gamma_{est})}.$$

.



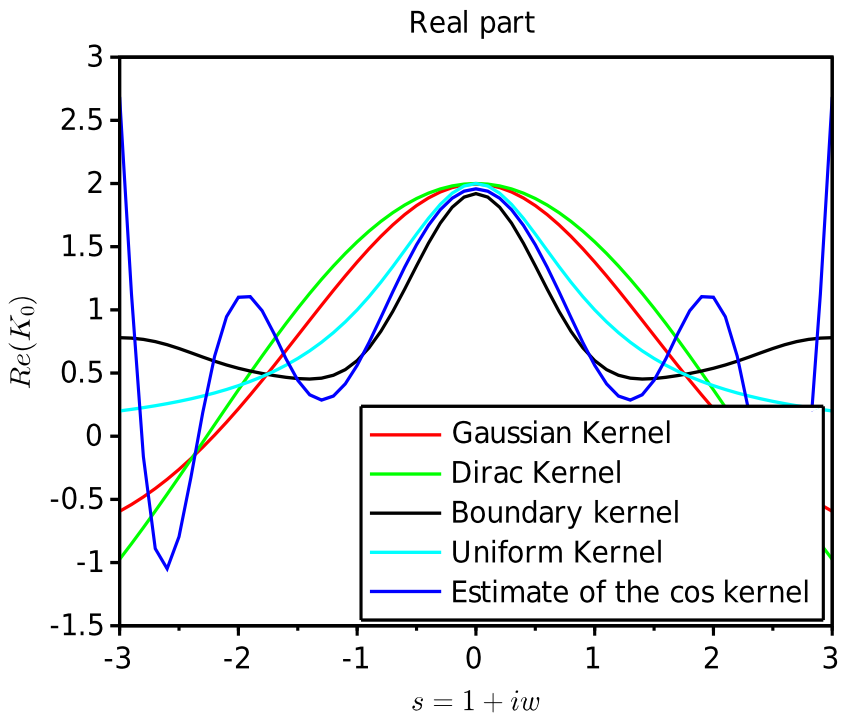
Uniform kernel



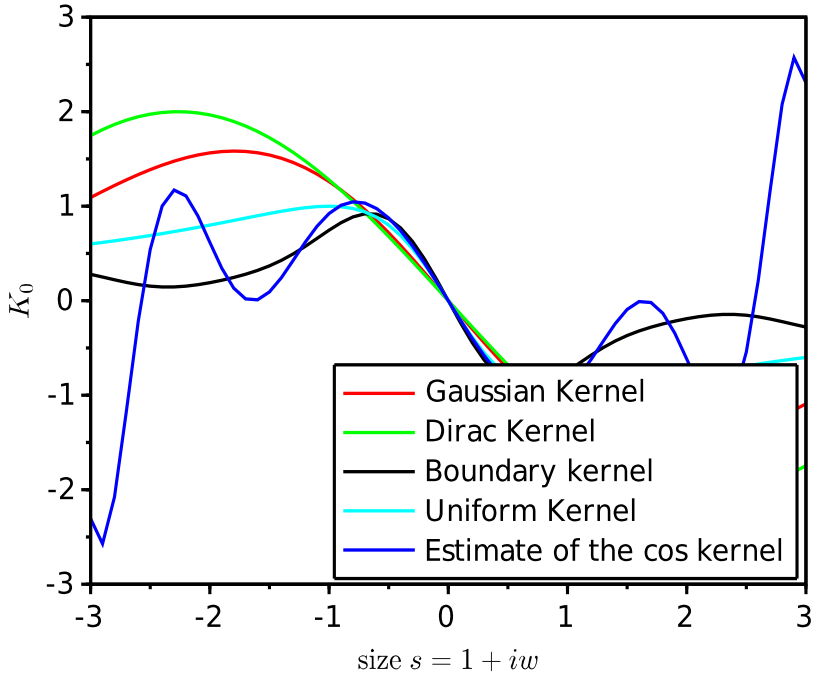
Gaussian kernel

Simulated data. Determination of k_0

$$K_0^{est}(s) = 1 + \frac{(2-s)G(s)}{\alpha_{est}\gamma_{est}G(s+\gamma_{est})}, \quad Re(s) > 2, \quad k_0^{est}(x) = \frac{1}{2\pi}\int_{u-i\infty}^{u+i\infty} K_0^{est}(s)x^{-s}ds.$$



Real part



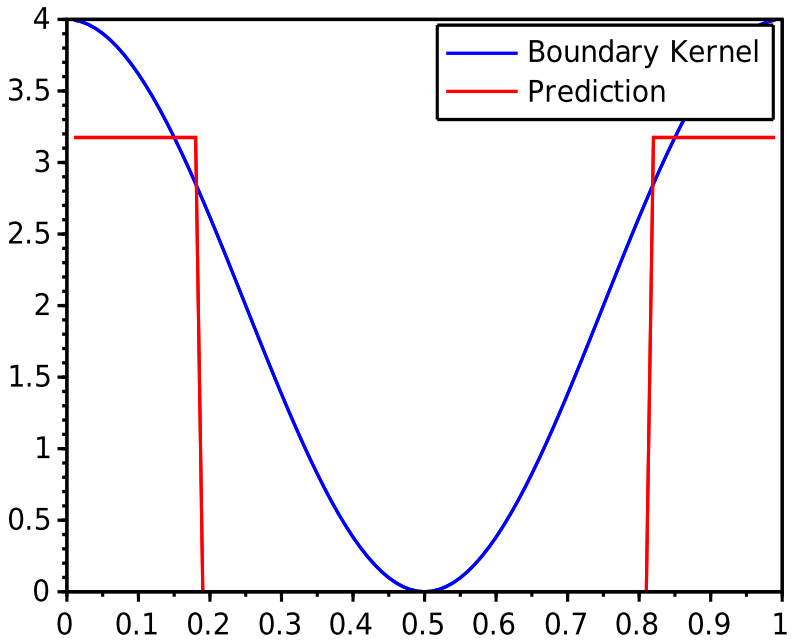
Imaginary part

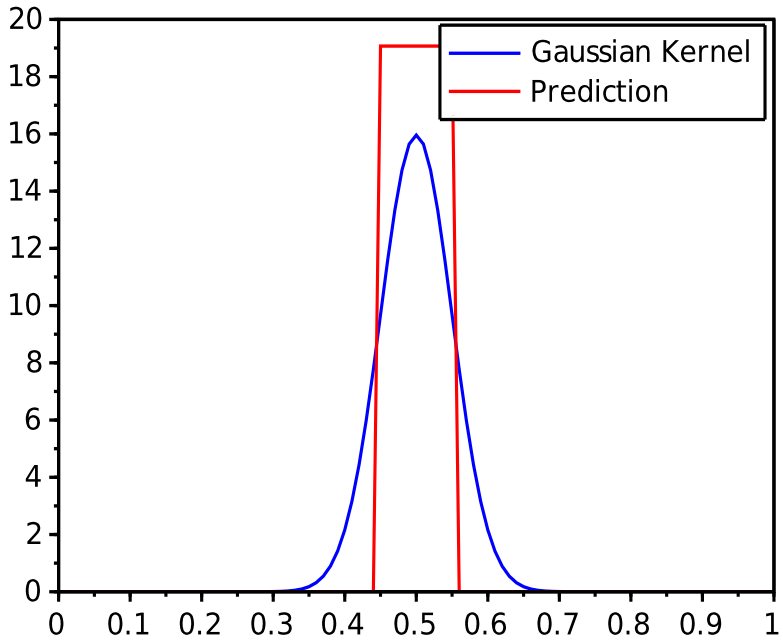
Simulated data. Determination of k_0

$\rightsquigarrow$ Complicated to estimate k_0 using the Mellin inverse formula.

Questions of interest Does k_0 charge the boundaries, or the center of the fibrils? If the center is more likely to break apart, what is the spread of the gaussian?

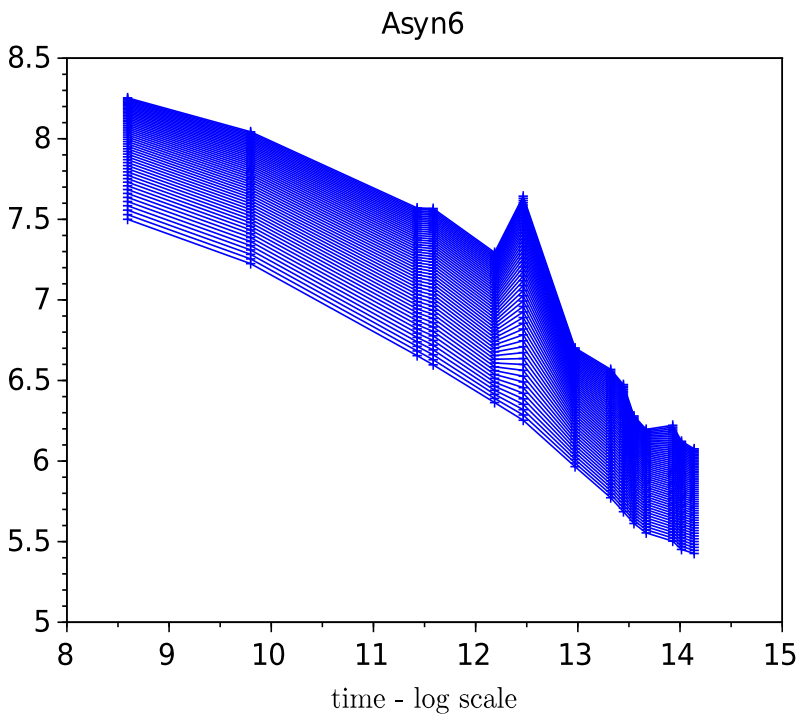
What we do Assume $k_0^{a,b,\eta}(x) = a\mathbf{1}_{[\eta,1-\eta]} + b\mathbf{1}_{([0,\eta]\cup[1-\eta,\eta])}$, and find a, b and η such that $\|K_0^{est} - K_0^{a,b,\eta}\|_{L^2(1+i[-1,1])}$ is minimal.

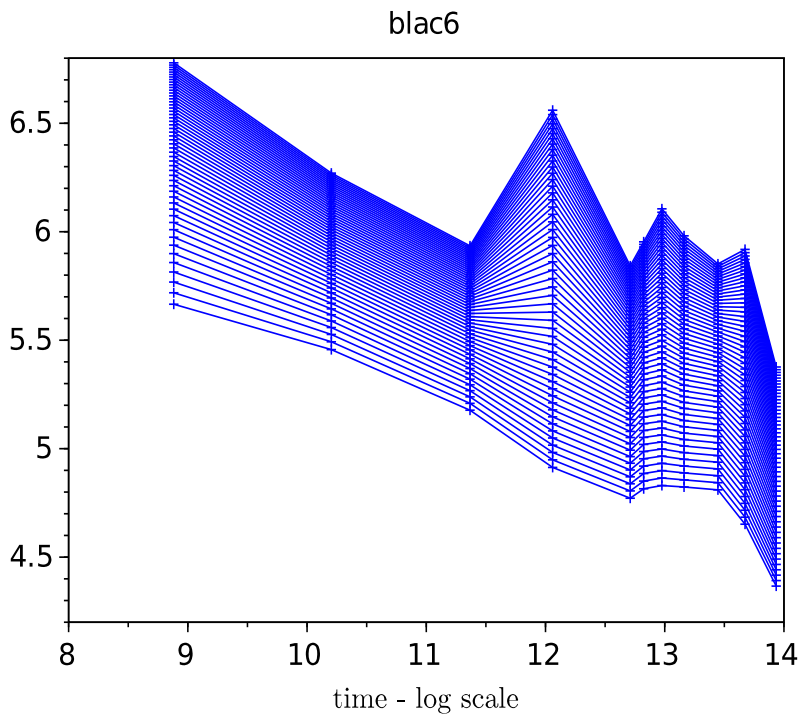




Real data: Determination of γ

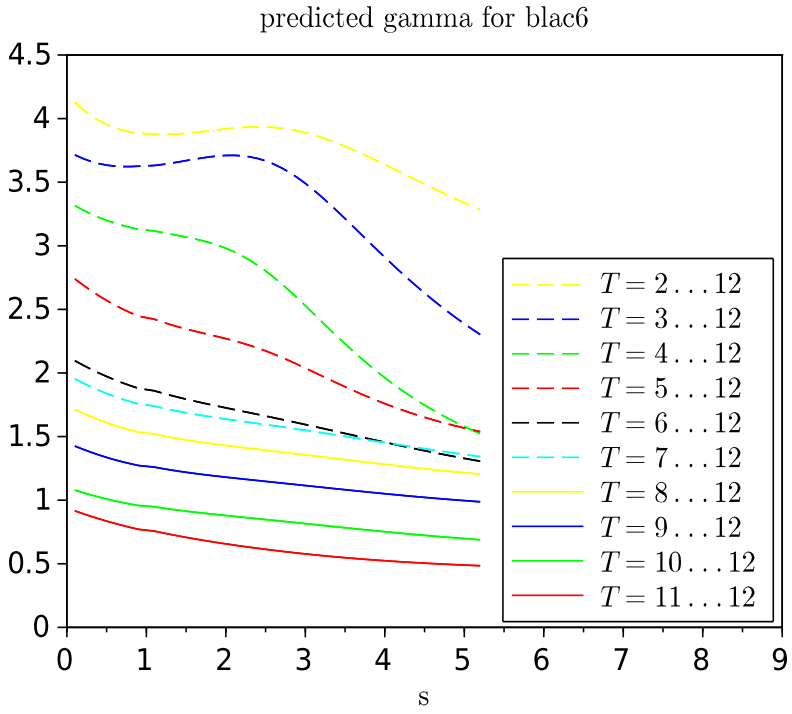
Plots of the function I_s at different times and for $s \in [1, 6.2]$.



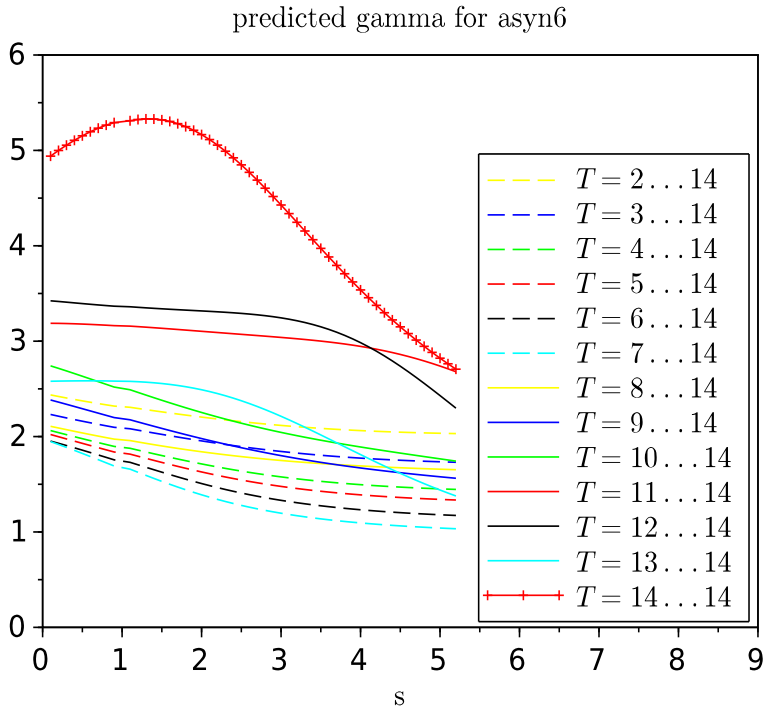


Data are noisy: Should we take into account all the data points to predicts γ?

Real data: Determination of γ

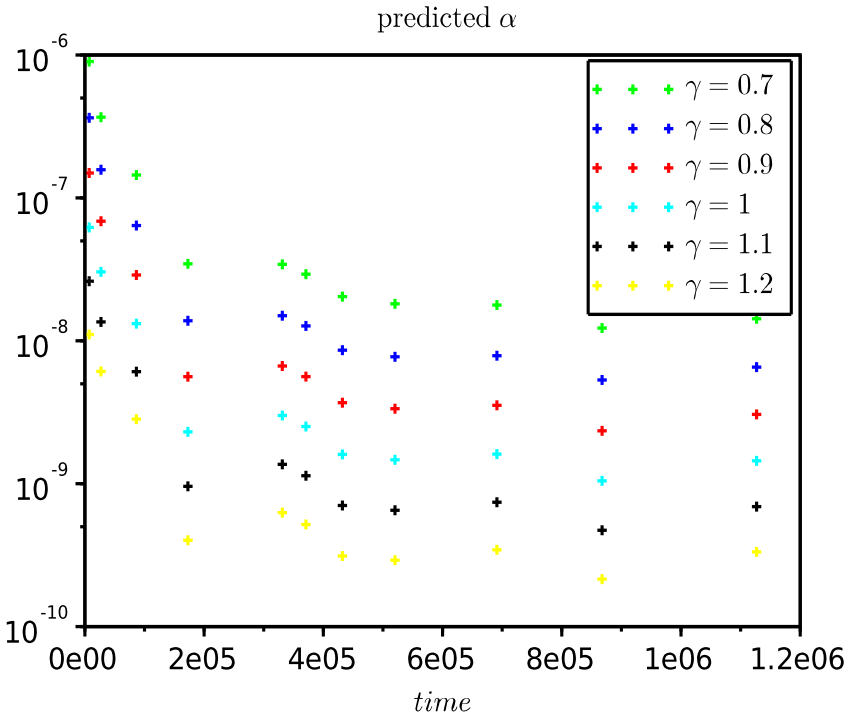


β-lactamase

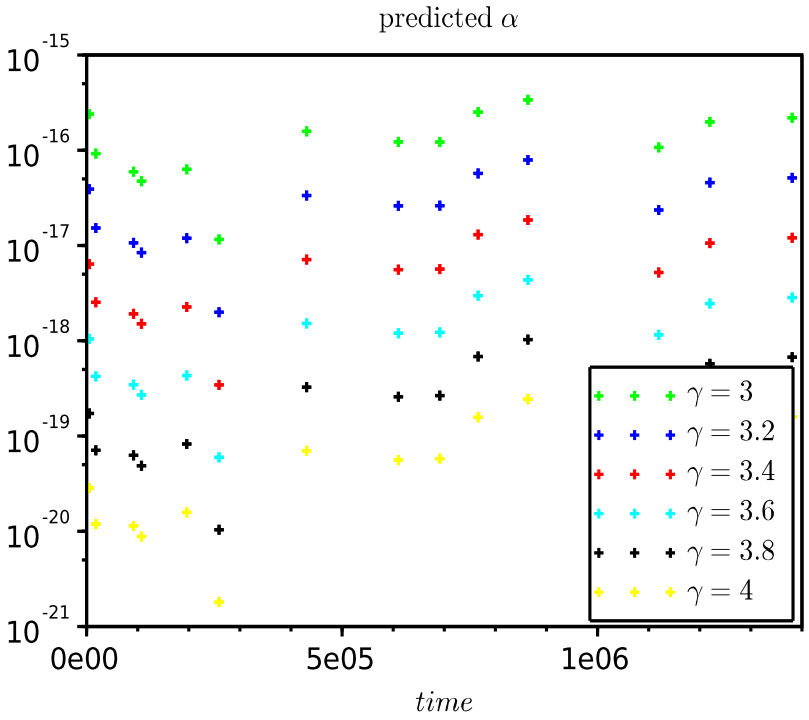


α-synuclein

Real data: Determination of α

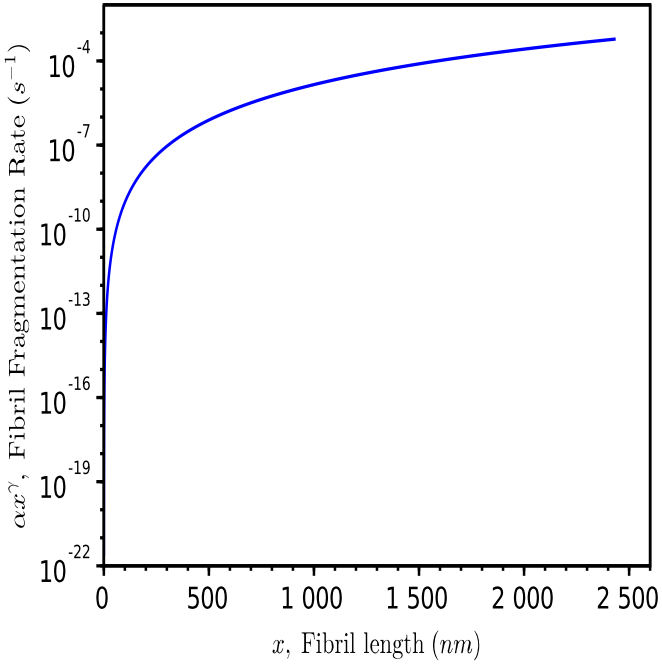


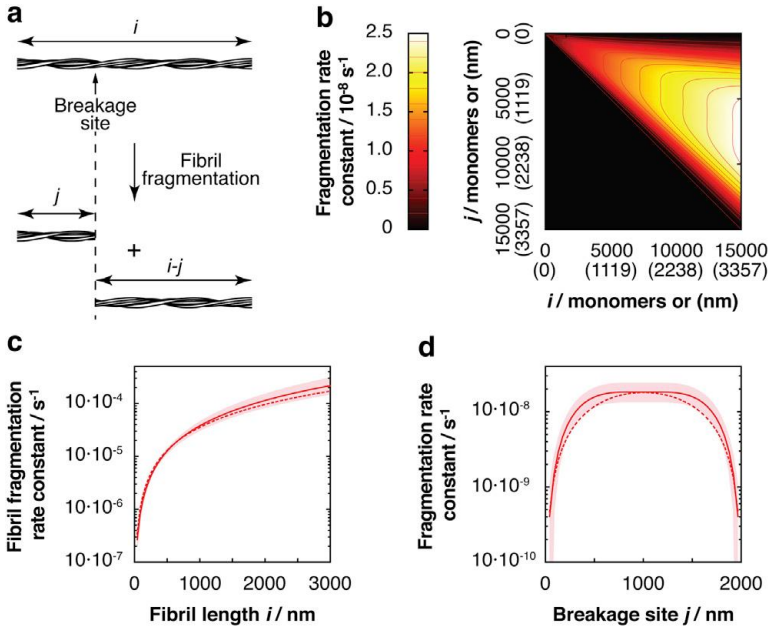
β-lactamase



α-synuclein

Test case: β2-amyloid





Comparison between the value we predict for $B(x)$ (left) and what the biologists found (right).

Conclusion and Perspectives

- Well posedness of the inverse problem
- Determination of a constructive and non-parametric method to recover the parameters
- Test of our method on simulated data
- Need to determine a protocol to deal with real data.

Thank you!