Bayesian inverse problem in computational Biology: application to protein transcription.

J.-C. Croix
Mines Saint-Etienne

Supervisor(s): Prof. M. Batton-Hubert (Mines Saint-Etienne), Dr. X. Bay (Mines Saint-Etienne) and Dr. E. Touboul (Mines Saint-Etienne). This is a joint work with Dr. N. Durrande (Mines Saint-Etienne) and Dr. M. Alvarez (Sheffield University)

Ph.D. expected duration: 2015-2018

Adress: Institut Henri Fayol, Espace Fauriel, 29 rue Ponchardier, F-42023 Saint-Etienne

Email: jean-charles.croix@emse.fr

Abstract:

In the sake of advanced understanding of living organisms, biologists continuously formulate theories and compare them against real experimental data. The phenomenon of protein transcription, that is the production of proteins from mRNA, is crucial in the early development of embryos and is yet not so well detailed. In particular, even for simple bodies such as the Melanogaster drosophilia, a quantitative estimation of the diffusion process and the rate of decay remain a challenge. This is mainly due to difficult, noisy and partial measurements of involved mRNA and protein concentrations.

The embryo body is modelled as 1D space, the anterior/posterior axis of total length 1. In this work, we consider the transcription process as a linear parabolic partial differential equation (PDE) with dirichlet boundary conditions $\forall (t, x) \in [0, 1] \times [0, 1]$:

$$\frac{\partial y}{\partial t}(t, x) + \lambda y(t, x) - D \Delta y(t, x) = s(x, t), \tag{1}$$

$$y(0, x) = 0, \tag{2}$$

$$y(t, 0) = g(t), \tag{3}$$

$$y(t, 1) = 0. \tag{4}$$

where y is the protein concentration, s the mRNA concentration, λ the factor of decay and D the diffusion coefficient which are all positive quantities. Our objective is to give an estimation, including uncertainty quantification of the parameters (λ, D, s), from noisy point-wise observations of the solution y.

The Bayesian Inverse problems methodology has been recently developped to tackle such challenge. It consists in choosing a prior distribution μ_0 for the parameters (λ, D, s) and characterize the posterior distribution μ_y given the dataset. This methodology is particularly well tailored for infinite dimensional problems, such as PDE operators [3].

In this work, the PDE solution operator:

$$T : (\lambda, D, s) \rightarrow y \tag{5}$$

is non-linear. This and the positivity contraint prevent the use of gaussian distributions and Green's function, and to the best of our knowledge, no closed formulae are available for inversion. We thus use a finite element approximation, coupled with state-of-the-art infinite dimensional MCMC algorithms [1] [2], which are robust to mesh refinement and use geometric information to speed-up computations.

References

[1] Alexandros Beskos. Geometric MCMC for Infinite-Dimensional Inverse Problems. *arXiv*, 2016.

[2] Daniel Rudolf and Bjrn Sprunk. On a generalization of the preconditioned Crank-Nicolson Metropolis algorithm. *arXiv*, 2016.

[3] Andrew M. Stuart. Inverse Problems: A Bayesian perspective. *Acta Numerica*, pages 451–559, 2010.

Short biography – After a Master Degree in Engineering at Mines Saint-Etienne and a Master Degree in Actuarial science from Lyon University, and a few years as an Actuary at AXA, Jean-Charles Croix is currently in second year of a PhD in applied mathematics at Mines Saint-Etienne on a public funding. His research interest are about inverse problems, uncertainty quantification and model order reduction in physical models.