Inverse problems in Biology

Thematic trimester program at IHP, March 7-9, 2022

Tentative Program

Venue: Amphithéâtre Hermite Institut Henri Poincaré, 11 Rue Pierre et Marie Curie, 75005 Paris.

Monday, March 7

- **14h-14h45:** Mélanie Prague (INRIA Bordeaux)
  Model building strategies in nonlinear mixed effects models.
- **14h45-15h30:** Chloé Audebert (Sorbonne Université)
  Parameters estimation and medical measurements: two problems to illustrate some difficulties and how to overcome them.
- 15h30-16h: Break.
- **16h-16h45:** Vincent Rivoirard (Université Paris-Dauphine)
  Nonparametric estimation for size-structured population of cells.
- **16h45-17h30:** Magali Tournus (Université d'Aix-Marseille)
  An inverse problem: recovering the fragmentation kernel from the short-time behaviour of the fragmentation equation

Tuesday, March 8

- **10h15-11h :** Pierre-Alexandre Bliman (Sorbonne Université)
  Modelling, analysis, observability and identifiability of epidemic dynamics with reinfections.

- 11h-11h30: Break.
- **11h30-12h15:** Christian Klingenberg (University of Würzburg)
  How to use data for PDE modeling in biology.
- **12h15-13h:** Kathrin Hellmuth (University of Würzburg)
  An inverse problem for bacterial movement: unique reconstruction of the chemotactic tumbling rate from macroscopic measurements.
- 13h-14h30: Break.
- **14h30-15h15:** Franck Picard (ENS Lyon)
  A probabilistic Graph Coupling View of Dimension Reduction.
- 15h15-15h30: Break.
- **15h30-16h15:** Kolyan Ray (Imperial College, London)
  Bayesian inference for multidimensional diffusions.
- **16h15-17h00:** Piotr Gwiazda (Polish Academy of Sciences)
  [TBA]

Wednesday, March 9

- **10h15-11h:** Carola Schoenlieb (University of Cambridge)
  Inverse problems in biomedical imaging.
- 11h-11h30: Break.
- **11h30-12h15:** Karim Ramdani (Inria Nancy and Université de Lorraine)
  Geometric inverse problems for the 2D-Laplace operator.
- **12h15-13h:** Philippe Moireau (Inria Saclay)
  Asymptotic analysis in inverse problems for depolymerization estimation.
- 13h-14h30: Break.
- **14h30-15h15:** Hermine Biermé (Université de Tours)
  Spike detection for calcium activity.
- **15h15-16h00:** Randolf Altmeyer (University of Cambridge)
  Statistical inference for SPDEs applied to an example in cell biology.

Abstracts

Randolf Altmeyer (University of Cambridge)

Statistical inference for SPDEs applied to an example in cell biology.

In this talk we propose an extension of a classical reaction-diffusion model for cell repolarisation by adding dynamic noise. This stochastic partial differential equation (SPDE) generates data, which differ qualitatively from the deterministic PDE model corrupted by measurement errors. The dynamic noise has interesting effects on the repolarisation behaviour such as larger noise levels speeding up the time to repolarisation instead of destroying the pattern formation. Apart from a qualitative description of the SPDE model, results on parameter estimation are presented to calibrate the model to data.

Chloé Audebert (Sorbonne Université)

Parameters estimation and medical measurements: two problems to illustrate some difficulties and how to overcome them.

Mathematical models are able to reproduce and explain some biological systems. To be reliable these models need to be calibrated with the available measurements. This calibration step raises several questions: are all parameters identifiable with the available data? are the model outputs sensitive to the parameters of interest? which method of estimation is the most relevant to my problem? are the measurements sufficient? We will address these questions for two different medical problems. First we will study the effects of liver partial ablation on pressure and flow waveforms. The model is a 1D-0D closed-loop model. In order to correctly represent the hepatic artery tree, which is the arterial tree of interest here, we will first work with a RCR model. We will present the calibration of this model with the available data, the parameters identifiability will be studied and parameters estimation will be performed with unscented Kalman filter (UKF) algorithm. The second medical problem we will present is the study of the effects of the treatment of chronic lymphocytic leukemia (CLL) by Btk Inhibitor Ibrutinib. The overall objective is to elaborate a computational model that accounts for the physical and biological evolution of the CLL leukemic cell population during ibrutinib therapy, to predict, at the start of the treatment, the clinical outcome and adapt the treatment. From the analyses of the data our medical collaborators are able to establish two groups of patients with either poor or good response to the treatment. We will describe with ordinary differential equations leukemic and normal lymphocytes dynamics in CLL patients treated with ibrutinib. We want to account for inter-patient variability in cell count dynamics. To

do so parameters estimation will be performed with non-linear mixed effects models. The idea of this method is to characterize the average behavior and to extract the individuals dynamics from this behavior. We aim at discriminating the different responses to treatment based on parameters of the mathematical model.

Hermine Biermé (Université de Tours)

Spike detection for calcium activity.

We present a global methodology for the spike detection in a biological context of fluorescence recording of GnRH-neurons calcium activity. For this purpose we first propose a simple stochastic model that could mimic experimental time series by considering an autoregressive AR(1) process with a linear trend and specific innovations involving spiking times. Estimators of parameters with asymptotic normality are established and used to set up a statistical test on estimated innovations in order to detect spikes. We compare several procedures and illustrate on biological data the performance of our procedure.

Pierre-Alexandre Bliman (Sorbonne Université)

Modelling, analysis, observability and identifiability of epidemic dynamics with reinfections.

In order to understand if counting the number of reinfections may provide supplementary information on the evolution of an epidemic, we consider in this paper a general SEIRS model describing the dynamics of an infectious disease including latency, waning immunity and infection-induced mortality. We derive an infinite system of differential equations that provides an image of the same infection process, but counting also the reinfections. Well-posedness is established in a suitable space of sequence valued functions, and the asymptotic behavior of the solutions is characterized, according to the value of the basic reproduction number. This allows to determine several mean numbers of reinfections related to the population at endemic equilibrium. We then show how using jointly measurement of the number of infected individuals and of the number of primo-infected provides observability and identifiability to a simple SIS model for which none of these two measures is sufficient to ensure on its own the same properties. This is a joint work with Marcel Fang. More details may be found in the report https://arxiv.org/abs/2011.12202

Piotr Gwiazda (Polish Academy of Sciences, Poland)

[TBA]
[TBA]

Kathrin Hellmuth (University of Würzburg)

An inverse problem for bacterial movement: unique reconstruction of the chemotactic tumbling rate from macroscopic measurements.

E. coli bacteria move by running along a straight line, until they stop and choose a new direction. When the movement is directed, e.g. induced by a gradient in the concentration of an attracting chemical substance, this phenomenon is called chemotaxis. On the population level, this 'run-and-tumble' process if often summarized by a kinetic chemotaxis partial differential equation (PDE) in the mesoscopic phase space. In this model, the tumbling process is typically assumed to be a Poisson process, whose rate is given by the tumbling rate. We study the inverse problem, namely we want to reconstruct this tumbling rate using given data. The data for the reconstruction consists of measurements of the bacteria density after running for a given time. Since velocity dependent measurements are difficult to obtain, we need to use macroscopic measurements, i.e. the velocity averaged bacteria density is measured. Inferring the tumbling rate from these data constitutes the inverse problem we are considering. Assuming we have tight control over the initial condition and measurement functions, we show the reconstruction of the tumbling rate is unique. For the proof we use a singular decomposition technique. This is joint work with Christian Klingenberg (Würzburg, Germany), Qin Li (Madison, Wisc., USA) and Min Tang (Shanghai, China).

Christian Klingenberg (University of Würzburg)

How to use data for PDE modeling in biology.

Say you observe a biological phenomena and are able to take measurements (data). Say this phenomenon is described by a partial differential equation (PDE) model, where some of the coefficients in this model still need to be determined. Now you would like to use your data to determine these coefficients of the PDE. In this lecture we will show examples for which solving such an inverse problem is possible. We study inverse problems in the framework of chemotaxis. On the population level, say a 'run-and-tumble' process of moving bacteria is modeled by a kinetic chemotaxis type equation. In the scaling limit, these models are typically linked to macroscopic PDEs, e.g. the Keller-Segel equation, describing the bacteria density in space and time. Biologists and practitioners are interested in modeling the collision kernel in order to study and control the bacterial movement e.g. in bioreactors. The collision kernel shall be reconstructed from data. Notice we have two related PDE models (mesoscopic chemotaxis type and macroscopic Keller-Segel type). The collision kernel in the chemotaxis type model corresponds to the diffusion/advection coeffi-

cients of the Keller-Segel type equation. Typically the inverse problem for the macroscopic model is ill posed, whereas this is usually not the case for the kinetic model. Hence we shall use our data to solve the kinetic inverse problem. We will then infer the solution to the macroscopic inverse problem in a Bayesian sense. At this point the underlying PDE model needs to be supplied, albeit in a way where coefficients of the PDE can to be determined by data. One could dream of a future where the data contains enough information to determine even the PDE model.

Juliette Leblond (INRIA Sophia Antipolis)

Inverse current source estimation problems and brain imaging issues related to EEG (ElectroEncephaloGraphy).

We will discuss inverse source estimation and Cauchy data transmission problems for EEG, for spherical or realistic layered head geometries, the layers corresponding to the brain, the skull, the scalp, assumed to be homogeneous. From given data consisting in measurements of the electric potential on a part of the outer surface of scalp, the aim is to recover the corresponding current source term supported within the brain. These issues are ill-posed and require regularization for their resolution. We propose Tykhonov-like schemes for data approximation in appropriate function spaces, depending on smoothness assumptions on the domain and the data, and on the source term model (deep pointwise dipolar or distributed on the gray-white matter interface). We may also use other measurements from other modalities, like sEEG (stereoElectroEncephaloGraphy) or MEG (MagnetoEncephaloGraphy), simultaneously available or not. The situation of a non-homogeneous skull will be discussed as well, for which the data transmission step requires a volumic scheme, in order to account for the variable conductivity. Joint work with J.-M. Badier, L. Baratchart, M. Darbas, J.-P. Marmorat, M. Nemaire, P.-H. Tournier.

Philippe Moireau (Inria Saclay)

Asymptotic analysis in inverse problems for depolymerization estimation.

We consider a model of depolymerization based on a Becker-Doring model. As only depolymerization is considered, the model reduces to a linear system where we want to estimate the initial distribution when measuring the time evolution of the second moment of the distribution during the depolymerization. In this respect, we propose to mathematically evaluate the impact of using asymptotic models of different orders as a surrogate of the initial Becker-Doring model. At order 0, the asymptotic model is a simple transport model which was proven to be mildly ill posed of order 3 when used to inverse the second moment

measurements. At order 1, the asymptotic model becomes an advection diffusion with an original transparent boundary condition in 0, a much more accurate approximation of the Becker-Doring model. However, the resulting inversion is still injective but severely ill posed. This leads to a classic dilemma of accuracy versus stability in the inversion strategy that we propose to fully quantify.

Franck Picard (ENS Lyon)

A probabilistic Graph Coupling View of Dimension Reduction.

Dimension reduction is a standard task in machine learning, to reduce the complexity and represent the data at hand. Many (and more than many !) methods have been proposed for this purpose, among which the seminal principal component analysis (PCA), that approximates the data linearly with a reduced number of axes. In recent years, the field has witness the emergence of new non linear methods, like the Stochastic Neighbor Embedding method (SNE) and the Uniform Manifold Approximation and Projection method (UMAP), that proposes very efficient low-dimensional representations of the observations. Though widely used, these approaches lack clear probabilistic foundations to enable a full understanding of their properties and limitations. A common feature of these techniques is to be based on a minimization of a cost between input and latent pairwise similarities, but the generative model is still missing. In this work we introduce a unifying statistical framework based on the coupling of hidden graphs using cross entropy. These graphs induce a Markov random field dependency structure among the observations in both input and latent spaces. We show that existing pairwise similarity dimension reduction methods can be retrieved from our framework with particular choices of priors for the graphs. Moreover this reveals that these methods suffer from a statistical deficiency that explains poor performances in conserving coarse-grain dependencies. Our model is leveraged and extended to address this issue while new links are drawn with Laplacian eigenmaps and PCA.

Mélanie Prague (INRIA Bordeaux)

Model building strategies in nonlinear mixed effects models.

[TBA]

Karim Ramdani (Inria Nancy and Université de Lorraine)

Geometric inverse problems for the 2D-Laplace operator

In this talk, we will present some results from an ongoing work with Anthony Gerber-Roth and Alexandre Munnier on the solution of two-dimensional geometric inverse problems for the Laplace operator. Such inverse problems arise naturally in many different frameworks, such as electrical impedance tomography in medicine (reconstruction of a cavity from boundary measurements) or inverse gravimetry in geophysics (reconstruction of a domain from its exterior gravitational (Newtonian) potential). The key ingredient we use to tackle these inverse problems is to rephrase them as shape-from-moments problems. In this context, the notion of quadrature domains used in potential theory plays a crucial role. Numerical examples will be given to illustrate the efficiency of the proposed reconstruction method.

Kolyan Ray (Imperial College, London)

Bayesian inference for multidimensional diffusions

A diffusion process, where a particle moves for instance in a fluid, is a common model in applications such as molecular dynamics. We study the problem of estimating the drift function driving such dynamics using a Bayesian approach, which provides popular uncertainty quantification methodology among other things. We place a Gaussian prior distribution on the drift function and study its theoretical properties, including converge rates for the posterior distribution and whether Bayesian uncertainty quantification provides reliable frequentist uncertainty quantification ('confidence sets'). We provide theoretical guarantees for such Bayesian methods, showing that they provide reliable inference for certain tasks in this setting.

Vincent Rivoirard (Université Paris-Dauphine)

Nonparametric estimation for size-structured population of cells.

We consider a stochastic individual-based model in continuous time to describe a size-structured population for cell divisions. This model is motivated by the detection of cellular aging in biology. The size of the system evolves according to a transport-fragmentation equation: each individual grows with a given transport rate and splits into two offsprings. In the first part of this talk, we assume that offsprings have the same size, following a binary fragmentation process with unknown division rate that depends on its size. Our nonparametric procedure to deal with the problem of rate estimation relies on kernel methods with automatic bandwidth selection performed by the Goldenshluger-Lepski methodology. In the second part, we do no longer assume that offsprings have the same size and we then

address the problem of nonparametric estimation of the kernel ruling the divisions based on the eigenvalue problem related to the asymptotic behavior in large population. This inverse problem involves a multiplicative deconvolution operator. Using Fourier technics we derive a nonparametric estimator whose consistency is studied.

Carola Schoenlieb (University of Cambridge)

Inverse problems in biomedical imaging.

In this talk I will review some prototypical inverse problems in biomedical imaging, including tomographic image reconstruction, image super resolution, motion estimation from dynamic imaging data, image segmentation and classification. Mathematical solution strategies discussed are variational regularisation approaches, partial differential equations, graphical models and deep neural networks.

Magali Tournus (Université d'Aix-Marseille)

An inverse problem: recovering the fragmentation kernel from the short-time behaviour of the fragmentation equation.

We consider a size-structured population of particles that undergo fragmentation. We address the question of estimating the fragmentation kernel from measurements of the size distribution at various times. The application that drives our work is the study of mechanical properties of amyloid fibrils that undergo fragmentation. In this talk, I will first present the biological context, and then detail how short time measurements can be used to recover the fragmentation kernel when initial data is close to a delta function. To this puropose, we provide a representation of the solution to the fragmentation equation as a power series in the Banach space of Radon measures endowed with the total variation norm.

Denis Villemonais (Université de Lorraine)

Stochastic branching models for the telomeres dynamics including telomerase.

Telomeres are short sequences of nucleotides at the end of chromosomes, whose evolution over time is intrinsically related to biological ageing. In most somatic cells, with each cell division, telomeres shorten due to the so-called end replication problem, which can lead to replicative senescence and a variety of age-related diseases. On the other hand, in

certain cells, the presence of the enzyme telomerase can lead to the lengthening of telomeres, which may delay or prevent the onset of such diseases but can also increase the risk of cancer. In this article, we propose a type branching stochastic representation for the evolution of telomeres in order to study their long-term behaviour in the presence of telomerase. This work is part of a collaboration with Athanasios Benetos, Coralie Fritsch, Anne Gégout-Petit, Emma Horton, Lionel Lenotre and Simon Toupance.