

RESEARCH CENTRE

**Inria Centre  
at Université de Lorraine**

IN PARTNERSHIP WITH:  
CNRS, Université de Lorraine

2023  
**ACTIVITY REPORT**

**Project-Team  
BIGS**

**Biology, genetics and statistics**

IN COLLABORATION WITH: Institut Elie Cartan de Lorraine (IECL)

**DOMAIN**

**Digital Health, Biology and Earth**

**THEME**

**Modeling and Control for Life Sciences**

*Inria*

# Contents

<b>Project-Team BIGS</b>	<b>1</b>
<b>1 Team members, visitors, external collaborators</b>	<b>2</b>
<b>2 Overall objectives</b>	<b>3</b>
<b>3 Research program</b>	<b>3</b>
3.1 Introduction	3
3.2 Stochastic modeling	3
3.3 Estimation and control for stochastic processes	3
3.4 Algorithms and estimation for graph data	4
3.5 Regression and machine learning	4
<b>4 Application domains</b>	<b>5</b>
4.1 Oncology: tumor growth and heterogeneity	5
4.2 Gene networks and single-cell data	5
4.3 Epidemiology and e-health	5
4.4 Dynamics of telomeres	5
<b>5 Highlights of the year</b>	<b>5</b>
5.1 Awards	5
<b>6 New software, platforms, open data</b>	<b>6</b>
6.1 New software	6
6.1.1 Harissa	6
6.1.2 MultiRNAflow	6
<b>7 New results</b>	<b>6</b>
7.1 Stochastic modeling	6
7.1.1 Reconstruction of epigenetic landscapes from single-cell data	7
7.1.2 Quasi-stationary distributions	7
7.1.3 Fluctuations of balanced urns with infinitely many colours	7
7.1.4 Adaptive dynamics in biological populations	8
7.1.5 Binary Branching Processes with Moran Type Interactions	8
7.1.6 Multi-type bisexual branching process	8
7.1.7 A branching model for intergenerational telomere length dynamics	8
7.1.8 Modeling of chronic obstructive pulmonary disease	9
7.1.9 Numerical simulation of diffusions	9
7.2 Regression and machine learning	9
7.2.1 Cramér-von Mises goodness-of-fit tests in regression models	9
7.2.2 Imprecise extension of the kernel density estimator	10
7.2.3 Online Big Data Analysis and Online Learning	10
7.3 Statistical learning and application in health	11
7.3.1 Invert emulsions alleviate biotic interactions in bacterial mixed culture	11
7.3.2 Prediction of silencing experiments on gene networks for chronic lymphocytic leukemia	11
7.3.3 Multidimensional statistical analysis of information for clinical use	11
7.3.4 Effects of adapted physical activity and education program on endometriosis symptoms	12
<b>8 Bilateral contracts and grants with industry</b>	<b>12</b>
8.1 Bilateral contracts with industry	12

<b>9 Partnerships and cooperations</b>	<b>12</b>
9.1 International initiatives	12
9.1.1 Inria associate team not involved in an IIL or an international program	12
9.2 International research visitors	13
9.2.1 Visits of international scientists	13
9.2.2 Visits to international teams	14
9.3 European initiatives	14
9.3.1 ERC projects	14
9.4 National initiatives	14
9.5 Regional initiatives	15
<b>10 Dissemination</b>	<b>15</b>
10.1 Promoting scientific activities	15
10.1.1 Scientific events: organisation	15
10.1.2 Scientific events: selection	16
10.1.3 Journal	16
10.1.4 Invited talks	16
10.1.5 Contributed talks, posters, workshops, seminars	17
10.1.6 Scientific expertise	17
10.1.7 Research administration	17
10.2 Teaching - Supervision - Juries	18
10.2.1 Teaching	18
10.2.2 Supervision	19
10.2.3 Juries	21
10.3 Popularization	21
10.3.1 Education	21
10.3.2 Interventions	21
<b>11 Scientific production</b>	<b>22</b>
11.1 Major publications	22
11.2 Publications of the year	23
11.3 Other	25
11.4 Cited publications	25

## Project-Team BIGS

*Creation of the Project-Team: 2011 January 01*

### Keywords

#### Computer sciences and digital sciences

- A3.1. – Data
- A3.2. – Knowledge
- A3.2.3. – Inference
- A3.3. – Data and knowledge analysis
- A3.3.1. – On-line analytical processing
- A3.3.2. – Data mining
- A3.3.3. – Big data analysis
- A3.4.1. – Supervised learning
- A3.4.2. – Unsupervised learning
- A3.4.4. – Optimization and learning
- A3.4.7. – Kernel methods
- A6. – Modeling, simulation and control
- A6.1. – Methods in mathematical modeling
- A6.1.2. – Stochastic Modeling
- A6.2. – Scientific computing, Numerical Analysis & Optimization
- A6.2.3. – Probabilistic methods
- A6.2.4. – Statistical methods
- A6.4. – Automatic control
- A6.4.2. – Stochastic control

#### Other research topics and application domains

- B1. – Life sciences
- B1.1. – Biology
- B1.1.2. – Molecular and cellular biology
- B1.1.10. – Systems and synthetic biology
- B1.1.11. – Plant Biology
- B2.2. – Physiology and diseases
- B2.2.1. – Cardiovascular and respiratory diseases
- B2.2.3. – Cancer
- B2.3. – Epidemiology
- B2.4. – Therapies

## 1 Team members, visitors, external collaborators

### Research Scientists

- Nicolas Champagnat [Team leader, INRIA, Senior Researcher, HDR]
- Coralie Fritsch [INRIA, Researcher]
- Ulysse Herbach [INRIA, Researcher]
- Bruno Scherrer [INRIA, Researcher, HDR]

### Faculty Members

- Thierry Bastogne [UL, Associate Professor, HDR]
- Sandie Ferrigno [UL, Associate Professor]
- Anne Gégout-Petit [UL, Professor, HDR]
- Jean-Marie Monnez [UL, Emeritus, HDR]
- Aurélie Muller-Gueudin [UL, Associate Professor]
- Pierre Vallois [UL, Emeritus, HDR]
- Denis Villemonais [UL, Associate Professor, HDR]
- Sophie Wantz-Mézières [UL, Associate Professor]

### PhD Students

- Sophie Baland [UL, from Oct 2023]
- Virgile Brodu [UL]
- Mathilde Gaillard [INRIA, from Oct 2023]
- Vincent Hass [UL, until Aug 2023]
- Vincent Kagan [INRIA, from Apr 2023 until Sep 2023]
- Rodolphe Loubaton [UL, ATER, until Aug 2023]
- Anouk Rago [UL]
- Nicolas Zalduendo Vidal [INRIA]

### Technical Staff

- Walid Laziri [INRIA, Engineer, until Nov 2023]
- Nathaniel Seyler [INRIA, Engineer, from Oct 2023]

### Interns and Apprentices

- Mathilde Gaillard [INRIA, from Apr 2023 until Sep 2023]

### Administrative Assistant

- Emmanuelle Deschamps [INRIA]

## 2 Overall objectives

BIGS is a joint team of Inria, CNRS and University of Lorraine, within the Institut Élie Cartan of Lorraine (IECL), UMR 7502 CNRS-UL laboratory in mathematics, of which Inria is a strong partner. One member of BIGS, T. Bastogne, comes from the Research Center of Automatic Control of Nancy (CRAN), with which BIGS has strong relations in the domain “Health-Biology-Signal”. Our research is mainly focused on stochastic modeling and statistics but also aims at a better understanding of biological systems. BIGS involves applied mathematicians whose research interests mainly concern probability and statistics. More precisely, our attention is directed on (1) stochastic modeling, (2) estimation and control for stochastic processes, (3) regression and machine learning, and (4) statistical learning and application in health. The main objective of BIGS is to exploit these skills in applied mathematics to provide a better understanding of issues arising in life sciences, with a special focus on (1) tumor growth and heterogeneity, (2) gene networks, (3) telomere length dynamics, (4) epidemiology and e-health.

## 3 Research program

### 3.1 Introduction

We give here the main lines of our research. For clarity, we made the choice to structure them in four items. Note that all of these items deal with stochastic modeling and inference, therefore they are all interconnected.

### 3.2 Stochastic modeling

Our aim is to propose relevant stochastic frameworks for the modeling and the understanding of biological systems. The stochastic processes are particularly suitable for this purpose. Among them, Markov processes provide a first framework for the modeling of population of cells [82, 64]. Piecewise deterministic processes are non-diffusion processes that are also frequently used in the biological context [51, 63, 52]. Among Markov models, we developed strong expertise about processes derived from Brownian motion and Stochastic Differential Equations [78, 62]. For instance, knowledge about Brownian or random walk excursions [83, 77] helps to analyse genetic sequences and to develop inference about them. We also have strong expertise in stochastic modeling of complex biological populations using individual-based models. These models can be used either from the point of view of asymptotic stochastic analysis [48], e.g. to study the long term Darwinian evolution of populations, or from the point of view of numerical analysis of biological phenomena [58, 39]. We also develop mathematical tools for the analysis of the long-time behavior of stochastic population processes accounting for possible extinction of (sub)populations [49].

### 3.3 Estimation and control for stochastic processes

We develop inference about the stochastic processes that we use for modeling. Control of stochastic processes is also a way to optimise administration (dose, frequency) of therapy, such as targeted therapies in cancer. Our team has a good expertise about inference of the jump rate and the kernel of piecewise-deterministic Markov processes (PDMP) [43, 42, 2], but there are many directions to go further into. For instance, previous work made the assumption of a complete observation of jumps and mode, which is unrealistic in practice. We also tackle the problem of inference of “hidden PDMP”. For example, in pharmacokinetics modeling inference, we want to account for the presence of timing noise and identification from longitudinal data. We have expertise on these subjects [44], and we also use mixed models to estimate tumor growth or heterogeneity [45].

We consider the control of stochastic processes within the framework of Markov Decision Processes [75] and their generalization known as multi-player stochastic games, with a particular focus on infinite-horizon problems. In this context, we are interested in the complexity analysis of standard algorithms, as well as the proposition and analysis of numerical approximate schemes for large problems in the spirit of [46]. Regarding complexity, a central topic of research is the analysis of the Policy Iteration algorithm, which has made significant progress in the last years [85, 74, 60, 80], but is still not fully understood. For large problems, we have an extensive experience of sensitivity analysis of approximate

dynamic programming algorithms for Markov Decision Processes [79, 67, 81], and we currently investigate whether/how similar ideas may be adapted to multi-player stochastic games.

### 3.4 Algorithms and estimation for graph data

Recently, our group has focused its attention on modeling and inference for graph data. A graph data structure consists of a set of nodes, together with a set of pairs of these nodes called edges. This type of data is frequently used in biology because they provide a mathematical representation of many concepts such as biological networks of relationships in a population or between genes in a cell.

Network inference is the process of making inference about the link between two variables, by taking into account the information about other variables. Reference [84] gives a very good introduction and many references about network inference and mining. Many methods are available to infer and test edges in Gaussian graphical models [84, 69, 57, 59]. However, the Gaussian assumption does not hold when dealing with typical “zero-inflated” abundance data, and we want to develop inference in this case.

Concerning gene networks, most studies have been based on population-averaged data: now that technologies enable us to observe mRNA levels in individual cells, a revolution in terms of precision, the network reconstruction problem paradoxically becomes more challenging than ever. Indeed, the traditional way of seeing a gene regulatory network as a deterministic system with some small external noise is being challenged by the probabilistic, bursty nature of gene expression revealed at single-cell level. Our objective is to propose dynamical models and inference methods that fully exploit the particular time structure of single-cell data. We described a promising strategy in which the network inference problem is seen as a calibration procedure for a new PDMP model that is able to acceptably reproduce real single-cell data [61, 76].

Among graphs, trees play a special role because they offer a powerful model for many biological concepts, from RNA to phylogenetic trees in heterogeneous tumors or through plant structures. Our research deals with several aspects of tree data. In particular, we work on statistical inference for this type of data under a given stochastic model. We also work on lossy compression of trees via directed acyclic graphs. These methods enable us to compute distances between tree data faster than from the original structures and with a high accuracy.

### 3.5 Regression and machine learning

Regression models and machine learning aim at inferring statistical links between a variable of interest and covariates. In biological studies, it is always important to develop adapted learning methods both in the context of *standard* data and also for data of high dimension (sometimes with few observations) and very massive or online data.

Many methods are available to estimate conditional quantiles and test dependencies [73, 65]. Among them we have developed nonparametric estimation by local analysis via kernel methods [55, 56] and we want to study properties of this estimator in order to derive a measure of risk based e.g. on confidence band and test. We study also other regression models like survival analysis, spatio-temporal models with covariates. Among the multiple regression models, we want to develop omnibus tests that examine several assumptions together.

Concerning the analysis of high dimensional data, our view on the topic relies on the *French data analysis school*, specifically on Factorial Analysis. In this context, stochastic approximation is an essential tool [66], which allows one to approximate eigenvectors in a stepwise manner [71, 70, 72]. We aim at performing accurate classification or clustering by taking advantage of the possibility of updating the information “online” using stochastic approximation algorithms [47]. We focus on several incremental procedures for regression and data analysis like linear and logistic regressions and PCA (Principal Component Analysis).

We also focus on the biological context of high-throughput bioassays in which several hundreds or thousands of biological signals are measured for a posterior analysis. We have to account for the inter-individual variability within the modeling procedure. We aim at developing a new solution based on an ARX (Auto Regressive model with eXternal inputs) model structure using the EM (Expectation-Maximisation) algorithm for the estimation of the model parameters.

## 4 Application domains

### 4.1 Oncology: tumor growth and heterogeneity

We want to propose stochastic processes to model the appearance of mutations and the evolution of their frequencies in tumor samples, through new collaborations with clinicians who measure a particular quantity called circulating tumor DNA (ctDNA). The final purpose is to use ctDNA as an early biomarker of the resistance to a targeted therapy: this is the aim of the project funded by ITMO Cancer that we coordinate. In the ongoing work on low-grade gliomas, a local database of 400 patients will be soon available to construct models. We plan to extend it through national and international collaborations (Montpellier CHU, Montreal CRHUM). Our aim is to build a decision-aid tool for personalised medicine.

### 4.2 Gene networks and single-cell data

We already mentioned in Section 3.4 our interest in the modeling and inference of transcriptomic bursting in gene regulatory networks from single-cell data. We are also currently working on the prediction and identification of therapeutic targets for chronic lymphocytic leukemia from gene expression data. Our goal is to propose new models allowing to make prediction of gene silencing experiments. Inference will be performed on gene expression data from patients' cells suffering from different forms of chronic lymphocytic leukemia. The goal is to identify therapeutic targets which could be silenced to reduce cell proliferation.

### 4.3 Epidemiology and e-health

In the context of personalized medicine, we have many ongoing projects with CHU Nancy. They deal with biomarkers research, prognostic value of quantitative variables and events, scoring, and adverse events. We also want to develop our expertise in rupture detection in a project with APHP (Assistance Publique Hôpitaux de Paris) for the detection of adverse events, earlier than the clinical signs and symptoms. The clinical relevance of predictive analytics is obvious for high-risk patients such as those with solid organ transplantation or severe chronic respiratory disease for instance. The main challenge is the rupture detection in multivariate and heterogeneous signals (for instance daily measures of electrocardiogram, body temperature, spirometry parameters, sleep duration, etc.). Other collaborations with clinicians concern foetopathology and we want to use our work on conditional distribution function to explain fetal and child growth. To that end, we use data from the "Service de foetopathologie et de placentologie" of the "Maternité Régionale Universitaire" (CHU Nancy).

### 4.4 Dynamics of telomeres

Telomeres are disposable buffers at the ends of chromosomes which are truncated during cell division; so that, over time, due to each cell division, the telomere ends become shorter. By this way, they are markers of aging. Through a collaboration with Pr A. Benetos, geriatrician at CHU Nancy, we recently obtained data on the distribution of the length of telomeres from blood cells [9]. We want to work in three connected directions: (1) refine methodology for the analysis of the available data; (2) propose a dynamical model for the lengths of telomeres and study its mathematical properties (long term behavior, quasi-stationarity, etc.); and (3) use these properties to develop new statistical methods.

## 5 Highlights of the year

### 5.1 Awards

D. Villemonais has been granted a delegation at Institut Universitaire de France from september 2023 to august 2028.



## 6 New software, platforms, open data

### 6.1 New software

#### 6.1.1 Harissa

**Name:** Hartree approximation for inference along with a stochastic simulation algorithm

**Keywords:** Gene regulatory networks, Reverse engineering, Molecular simulation

**Functional Description:** Harissa is a Python package for both inference and simulation of gene regulatory networks, based on stochastic gene expression with transcriptional bursting. It was implemented in the context of a mechanistic approach to gene regulatory network inference from single-cell data.

**News of the Year:** This software has a more user-friendly interface and several tutorial notebooks are now available.

**URL:** <https://github.com/ulysseherbach/harissa>

**Publications:** [hal-04208601](#), [hal-04071033](#), [hal-01646910](#)

**Contact:** Ulysse Herbach

#### 6.1.2 MultiRNAflow

**Name:** An R package for the analysis of RNAseq raw counts with multiple biological conditions and time points

**Keywords:** RNA-seq, Gene regulatory networks, Integrated data analysis, Complex experimental design, Multiple temporal and biological conditions, Differential expression

**Functional Description:** The R package MultiRNAflow provides an easy to use unified framework allowing to make both unsupervised and supervised analysis (differential expression analysis) for RNAseq datasets with an arbitrary number of biological conditions and time points. In particular, this package makes a deep downstream analysis of differential expression information, e.g. identifying temporal patterns across biological conditions and differentially expresses genes which are specific to a biological condition for each time.

**Release Contributions:** First version

**URL:** <https://bioconductor.org/packages/release/bioc/html/MultiRNAflow.html>

**Contact:** Nicolas Champagnat

**Participants:** Rodolphe Loubaton, Nicolas Champagnat, Pierre Vallois, Laurent Vallat

**Partner:** CHRU de Strasbourg

## 7 New results

### 7.1 Stochastic modeling

**Participants:** Sophie Baland, Virgile Brodu, Nicolas Champagnat, Coralie Fritsch, Mathilde Gaillard, Vincent Hass, Ulysse Herbach, Vincent Kagan, Nathaniel Seyler, Pierre Vallois, Denis Villemonais, Nicolás Zalduendo Vidal.

### 7.1.1 Reconstruction of epigenetic landscapes from single-cell data

*Joint work with E. Ventre (ENS Lyon), T. Espinasse (Univ. Lyon 1), G. Benoit (Univ. Rennes 1) and O. Gandrillon (ENS Lyon).*

The aim of this collaboration is to better understand how living cells make decisions (e.g., differentiation of a stem cell into a particular specialized type), seeing decision-making as an emergent property of an underlying complex molecular network. Indeed, it is now proven that cells react probabilistically to their environment: cell types do not correspond to fixed states, but rather to “potential wells” of a certain energy landscape (representing the energy of the possible states of the cell) that we are trying to reconstruct. The achievement of last year was to show that the same mathematical model driven by transcriptional bursting can be used simultaneously as an inference tool, to reconstruct biologically relevant networks, and as a simulation tool, to generate realistic transcriptional profiles emerging from gene interactions: the article presenting these results is now published [25]. In addition, the paper proposing a landscape reconstruction method with application to several datasets has also been published this year [22].

These results form the starting point of M. Gaillard’s thesis work, which will focus on making links with interpretable dimension reduction for single-cell RNA-seq data. Finally, we are working with software engineer N. Seyler on a refactoring of the “Harissa” Python package used in [10] for stochastic simulation and inference of gene regulatory networks, with the aim of making it modular and scalable. The latest stable version is available on [PyPI](#) and is presented in a dedicated tool paper [30].

### 7.1.2 Quasi-stationary distributions

We are continuing our research on quasi-stationary distributions (QSD), that is, distributions of Markov stochastic processes with absorption, which are stationary conditionally on non-absorption. For models of biological populations, absorption usually corresponds to extinction of a (sub-)population. QSDs are fundamental tools to describe the population state before extinction and to quantify the large-time behavior of the probability of extinction.

Thanks to the previous general result of the team in [50], together with B. Cloez (INRAE), we proved in [16] the exponential convergence of a chemostat model, whose dynamics are highly degenerate due to a deterministic part, towards a unique quasi-stationary distributions.

We also finalized an important work [15] that provides general criteria for the exponential convergence of conditional distributions of absorbed Markov processes when the convergence is not uniform with respect to the initial distribution. Our results allow to characterize a large subset of the domain of attraction of the minimal QSD and apply to a large range of stochastic processes, including diffusion processes and perturbed dynamical systems.

In collaboration with E. Strickler (Univ. Lorraine), we also studied in [34] the convergence of general penalized Markov processes with soft killing in  $L^1$  (Monge-Kantorovich) Wasserstein distance. We propose a simple criterion ensuring uniform convergence of conditional distributions to a unique quasi-stationary distribution. We give several examples of application where our criterion can be checked, including Bernoulli convolutions and piecewise deterministic Markov processes, for which convergence in total variation is not possible.

### 7.1.3 Fluctuations of balanced urns with infinitely many colours

*Joint work with Svante Janson (Uppsala Univ., Sweden) and Cécile Mailler (Univ. Bath, UK)*

In this collaborative study, we delve into the dynamics of measure-valued Pólya processes (MVPPs), commonly known as Pólya urns with infinitely-many colours. Our study introduces the first second-order results in the literature on MVPPs, extending classical fluctuation outcomes from finitely-many-colour Pólya urns to the infinite colour space scenario. The nature of fluctuations in MVPPs is intricately linked to the “spectral gap”, adding a layer of sophistication to our understanding of these processes.

By framing MVPPs as stochastic approximations operating within the set of measures on a measurable space  $E$  (the colour space), we employ martingale methods and standard operator theory to rigorously prove convergence and unravel the nuanced fluctuation patterns inherent in these stochastic approximations [23].

#### 7.1.4 Adaptive dynamics in biological populations

*Joint work with Sylvie Méléard (École Polytechnique), Sepideh Mirrahimi (Univ. Montpellier) and Viet Chi Tran (Univ. Paris Est Marne-la-Vallée).*

We continued our study of parameter scalings of individual-based models of biological populations under mutation and selection, taking into account the influence of negligible but non-extinct populations. In a work within the ERC SINGER [14], we were able to give an individual-based justification of the Hamilton-Jacobi equation of adaptive dynamics (see e.g. [68]), with a specific parameter scaling that is promising for the study of local (in space) extinction of sub-populations. The analysis of models allowing for such an extinction is the next step of this project. We also wrote an article [26] for the proceedings of the International Congress of Mathematicians (ICM 2022) where S. Méléard gave an invited talk on several large population scalings that can be used in evolutionary biology.

We also worked on general evolutionary models of adaptive dynamics under an assumption of large population and small mutations. We obtained in [13] existence, uniqueness and ergodicity results for a centered version of the Fleming-Viot process of population genetics, which are key steps to recover variants of the canonical equation of adaptive dynamics, which describes the long time evolution of the dominant phenotype in the population, under less stringent biological assumptions than in previous works such as [48]. We completed this second step in [33].

#### 7.1.5 Binary Branching Processes with Moran Type Interactions

*Joint work with Alexander Cox (Univ. Bath, UK) and Emma Horton (Univ. Warwick, UK).*

In this collaboration, our focus is on investigating the large population limit of a binary branching particle system with Moran type interactions. The novel model introduced in this paper features particles that evolve, reproduce, and die independently. It encompasses branching models and fixed size Moran type interacting particle systems. The death of a particle may trigger the reproduction of another, while a branching event may, in turn, lead to the demise of another particle. Our study [17] aims to elucidate the intricate dynamics of this model. We explore diverse applications of our model, including its relevance to the neutron transport equation and population size dynamics. We focus on the occupation measure of the new model, explicitly connecting it to the Feynman-Kac semigroup of the underlying Markov evolution. Additionally, we quantify the  $L^2$  distance between the normalizations of these measures, providing valuable insights into the convergence behavior of the system.

#### 7.1.6 Multi-type bisexual branching process

The asexual multi-type Galton-Watson branching processes as well as the single-type bisexual processes have been studied in the literature. In particular, survival condition of the processes are well known in both cases. However, until now, the multi-type bisexual branching processes have only been studied in very specific situations and no general mathematical description has been established yet.

In [21], we studied general multi-type bisexual branching processes with superadditive mating function. We exhibited a necessary and sufficient condition for almost sure extinction, we proved a law of large numbers for our model and we studied the long-time convergence of the rescaled process.

#### 7.1.7 A branching model for intergenerational telomere length dynamics

*Joint Work with Athanasios Benetos (Univ. Lorraine), Lionel Lenôtre (Univ. Haute Alsace) and Simon Toupance (Univ. Lorraine).*

In this study, we construct and analyze an individual-based model capturing the evolution of telomere length in a population across multiple generations [32]. The model, a continuous-time typed branching process, incorporates individual characteristics such as gamete mean telomere length and age. Our investigation delves into the Malthusian behavior of the model, and we complement our findings with numerical simulations to elucidate the impact of biologically relevant parameters on telomere length dynamics on an evolutionary time scale.

### 7.1.8 Modeling of chronic obstructive pulmonary disease

*Joint work with Isabelle Dupin (Univ. Bordeaux), Élise Maurat (Univ. Bordeaux) and Jean-Marc Sac-Epée (IECL).*

Lung exposure to various types of particules, such as those present in cigarette smoke, can lead to chronic obstructive pulmonary disease (COPD). COPD bronchi are an area of intense immunological activity and tissue remodeling, as evidenced by the extensive immune cell infiltration and changes in tissue structures. This allows the persistent contact between resident cells and stimulated immune cells. Our hypothesis is that the contact between cells is a major cause of chronic destructive or fibrotic manifestations. We aim to analyze the potential cell-cell interactions in situ in human tissues, to characterize in vitro the dynamics of the interplay, and to define a computational model with intercellular interactions which fits to experimental measurements and explains the macroscopic properties of cell populations. The effects of potential therapeutic drugs modulating local intercellular interactions will be tested by simulations. A paper has been submitted this year [19] (see also [54]).

### 7.1.9 Numerical simulation of diffusions

In a collaboration with A. Lejay (Inria PASTA team) and their PhD student A. Anagnostakis, D. Villemonais proposed a method for approximating general, singular diffusions by discrete time and state space processes [11]. One of the main interests compared to existing methods is to propose a numerical method whose main computational cost is done upstream and thus represents a fixed cost, independently of the number of simulations performed afterwards.

## 7.2 Regression and machine learning

**Participants:** Sandie Ferrigno, Jean-Marie Monnez.

### 7.2.1 Cramér-von Mises goodness-of-fit tests in regression models

*Join work with R. Azaïs (Inria, ENS Lyon) and M.-J. Martinez (Univ. Grenoble Alpes).*

Many goodness-of-fit tests have been developed to assess the different assumptions of a (possibly heteroscedastic) regression model. Most of them are ‘directional’ in that they detect departures from a given assumption of the model. Other tests are ‘global’ (or ‘omnibus’) in that they assess whether a model fits a dataset on all its assumptions. We focus on the task of choosing the structural part of the regression and the variance functions because they contain easily interpretable informations about the studied relationship. We consider two nonparametric ‘directional’ tests and one nonparametric ‘global’ test, all based on generalizations of the Cramér-von Mises statistic.

To perform these goodness-of-fit tests, we have developed the R package `cvmgof` [41], an easy-to-use tool for practitioners, available from the Comprehensive R Archive Network (CRAN). The package was updated in 2022 (this is its third version) [40]. This latest version currently allows testing the “regression function” part of the model. In 2023, we worked to enrich the package by allowing the user to test the

homoskedasticity/heteroskedasticity of the model. This new version will be submitted to CRAN in 2024 and an associated article is currently being written.

To complete this work, we plan to assess the other assumptions of a regression model such as the additivity of the random error term. The implementation of these directional tests would enrich the `cvmgof` package and offer a complete easy-to-use tool for validating regression models. Another perspective of this work would be to develop a similar tool for other statistical models widely used in practice such as generalized linear models.

### 7.2.2 Imprecise extension of the kernel density estimator

*Join work with Bilal Nehme (IECL, Nancy).*

The estimation of the probability density function underlying a finite set of observations is a fundamental problem that covers a broad range of applications including machine learning. We propose a new nonparametric method to estimate this function that combines both the Schwartz distribution theory and the possibility theory. It is an extension of the kernel density estimator that leads to imprecise estimation, based on a new type of kernel called maxitive kernel. The form of the obtained estimation is an interval. In collaboration with B. Nehme, S. Ferrigno demonstrated several theoretical properties of the imprecise estimator. We implement this method using very low complexity algorithms and illustrate some theoretical properties of the proposed imprecise density estimation as well as a comparative analysis with other estimation intervals. An associated article is currently being written.

### 7.2.3 Online Big Data Analysis and Online Learning

A tool for analyzing streaming data is stochastic approximation introduced by Robbins and Monro in 1951, that can be used for example to estimate online parameters of a regression function [53] or centers of clusters in unsupervised classification [47]. Another type of stochastic approximation processes was introduced by Benzécri in 1969 for estimating eigenvectors and eigenvalues of the unknown  $Q$ -symmetric expectation of a random matrix  $A$  using independent observations of  $A$ . In all these processes, it is assumed that independent observations of the random matrix are observed and that one or a mini-batch of observations per step are taken into account. We are interested in the study of cases where we cannot have independent observations and we define processes where at each step all the observations up to this step are taken into account without storing them. Experiments we have conducted show that this second type of process generally converges faster than the first type.

**Stochastic approximation of eigenvectors and eigenvalues of the  $Q$ -symmetric expectation of a random matrix** In the article [24], we establish an almost sure convergence theorem of an extension of the stochastic approximation process of Oja for estimating eigenvectors of the unknown  $Q$ -symmetric expectation  $B$  of a random matrix, under a correlation model between the incoming random matrices. This theorem generalizes previous theorems and extends them to the case where the metric  $Q$  is unknown and estimated online in parallel. We suggest constructing processes using past and current observations at each step without storing them. We prove the almost sure convergence of specific processes to corresponding eigenvalues. We apply these results to streaming principal component analysis (PCA) of a random vector  $Z$ , when a mini-batch of observations of  $Z$  is used at each step or all the observations up to the current step. We deal with the case of streaming generalized canonical correlation analysis, interpreted as a PCA with a metric estimated online in parallel.

**An extended Oja process for streaming canonical analysis** In the article [36], after recalling an almost sure convergence theorem of an extended Oja process [24], we present the canonical correlation analysis (CCA) of two random vectors  $Z^1$  and  $Z^2$  such that there is no affine relation between their components. Couples of canonical components are interpreted as couples of principal components of the respective PCA of the linear regression function of  $Z^1$  with respect to  $Z^2$  and  $Z^2$  with respect to  $Z^1$ , or as canonical components of the generalized canonical correlation analysis (gCCA) of  $Z = (Z^1, Z^2)$ . In the case of streaming data, we estimate online in parallel a regression function and canonical components, using at

each step a mini-batch of current data or all the data up to the current step to have a faster convergence. We define two algorithms, the second being extended to gCCA. Using the same methodology for streaming factorial correspondence analysis (FCA) when the components of  $Z^1$  and  $Z^2$  are respectively the indicators of the exclusive modalities of two categorical variables, we define two algorithms to estimate online the canonical components, the second being extended to multiple correspondence analysis (MCA). Finally, we apply this methodology to streaming factorial discriminant analysis (FDA), when there is no affine relation between the components of  $Z^1$  and the components of  $Z^2$  are the indicators of the exclusive modalities of a categorical variable.

### 7.3 Statistical learning and application in health

**Participants:** Nicolas Champagnat, Sandie Ferrigno, Anne Gégout-Petit, Ulysse Herbach, Walid Laziri, Rodolphe Loubaton, Sophie Wantz-Mézières, Anouk Rago, Pierre Vallois.

#### 7.3.1 Invert emulsions alleviate biotic interactions in bacterial mixed culture

*Joint work with A. Dijamentiuk, C. Mangavel and E. Borges from LIBio, Univ. Lorraine.*

The large application potential of microbiomes has led to a great need for mixed culture methods. However, microbial interactions can compromise the maintenance of biodiversity during cultivation in a reactor. In particular, competition among species can lead to a strong disequilibrium in favor of the fittest microorganism. The aim of this study was to evaluate the potential of single invert emulsions to alleviate competition during the culture of antagonistic microorganisms and therefore to maintain diversity in a more complex mixed culture. Experimental data obtained in this study were analyzed using a two-way analysis of variance using a fixed effects model, followed by Tukey's HSD test. In the droplet size distributions of the invert emulsions, factors involved were the presence or absence of bacteria, and the incubation of invert emulsions. In bacterial enumerations, factors were the cultivation system used and the incubation. In community cultivation experiments, differences in Shannon diversity index between groups of samples were tested using one-way analysis of variance, followed by a Tukey's HSD test. An article [18] has been published on this work in 2023.

#### 7.3.2 Prediction of silencing experiments on gene networks for chronic lymphocytic leukemia

*Joint work with Laurent Vallat (CHRU Strasbourg).*

In this collaboration, we work on the inference of dynamical gene networks from RNAseq and proteome data. The goal is to infer a model of gene expression allowing to predict gene expression in cells where the expression of specific genes is silenced (e.g. using siRNA), in order to select the silencing experiments which are more likely to reduce the cell proliferation. We expect the selected genes to provide new therapeutic targets for the treatment of chronic lymphocytic leukemia. This year, we have developed a new method of prediction of the effect of gene silencing, based on the re-exploitation of expression data of genes not influenced by the silenced gene [27]. We also have developed the package MultiRNAflow (see Section 6.1.2) for the statistical analysis of temporal gene expression datasets with several biological conditions (in particular for exploratory analysis and the detection of differentially expressed genes). The package is described in the application note [35].

#### 7.3.3 Multidimensional statistical analysis of information for clinical use

The start-up EMOSIS develops blood tests relying on flow cytometry in order to improve in vitro diagnosis of vascular thrombosis. This technology leads to multiparametric measurements on tens of thousands cells collected from each blood sample. Manual methods of analysis classically used in flow cytometry are based on data visualization by means of histograms or scatter plots. Computational algorithmic approach

that would automate and deepen the search of differences or similarities between cell subpopulations could thus increase the quality of diagnosis.

Recent progresses in the active area of computational methods for dimension reduction suggest many directions of improvement of the classical approaches for the analysis of flow cytometry data. The approach that we considered is information geometry, whose principle is to lower the dimensionality of multiparametric observations by considering the subspace of the parameters of the statistical model describing the observation, whose points are probability density functions, and which is equipped with a special geometrical structure. The objective of the reported study is to use an algorithm belonging to the field of information geometry known as Fisher Information Non-parametric Embedding (FINE) to analyze flow cytometry data in the context of the specific severe disorder called heparin-induced thrombocytopenia. This work led to two communications in conferences [28, 29].

Unfortunately the start-up EMOSIS no longer exists, which put an end to our collaboration.

### 7.3.4 Effects of adapted physical activity and education program on endometriosis symptoms

*Joint work with Géraldine Escriva-Boulley and Lionel Lenôtre (Univ. Haute-Alsace).*

Endometriosis is a chronic disease characterized by growth of endometrial tissue outside the uterine cavity which could affect 200 million women worldwide. One of the most common symptoms of endometriosis is pelvic chronic pain associated with fatigue. This pain can cause psychological distress and interpersonal difficulties. As for several chronic diseases, adapted physical activity could help to manage the physical and psychological symptoms.

We are participating in both design and statistical analysis of a randomized-controlled trial, led by G. Escriva-Boulley, to investigate the potential effects of a videoconference-based adapted physical activity combined with endometriosis-based education program [20]. This study is one of the first trials to test the effects of a combined adapted physical activity and education program for improving endometriosis symptoms and physical activity.

## 8 Bilateral contracts and grants with industry

### 8.1 Bilateral contracts with industry

**Participants:** Anne Gégout-Petit, Walid Laziri, Sophie Wantz-Mézières.

As part of the French “Plan de relance”, we obtained funds for a 2-year engineering contract with the start-up EMOSIS based in Strasbourg (from October 1, 2022). Project MOSAiC : MultidimensiOnal Statistical Analysis of Information for Clinical use. Unfortunately EMOSIS ordered to file for bankruptcy in 2024 and the project was stopped.

## 9 Partnerships and cooperations

**Participants:** Sophie Baland, Nicolas Champagnat, Coralie Fritsch, Mathilde Gailard, Anne Gégout-Petit, Vincent Hass, Ulysse Herbach, Rodolphe Loubaton, Anouk Rago, Pierre Vallois, Denis Villemonais, Sophie Wantz-Mézières, Nicolas Zalduendo Vidal.

### 9.1 International initiatives

#### 9.1.1 Inria associate team not involved in an IIL or an international program

MAGO

**Title:** Modelling and analysis for growth-fragmentation processes

**Duration:** 2022-2024

**Coordinator:** Denis Villemonais

**Partners:**

- Inria Nancy (C. Fritsch, D. Villemonais)
- University College London London (E.X. Briol, O. Key, A. Watson)

**Summary:** Growth-fragmentation(GF) refers to a collection of mathematical models in which objects – classically, biological cells – slowly gather mass over time, and fragment suddenly into multiple, smaller offspring. These models may be used to represent a range of biological processes, in which an individual reproduces by fission into two or more new individuals, such as the evolution of plasmids in bacteria populations and protein polymerisation. It is crucial to understand the long-term behaviour of GF processes so that they can be used to build algorithms to simulate real-world processes and estimate quantities such as the growth rate of the system, the steady state behaviour, and the fragmentation rate and kernel, allowing scientists to gain a better understanding of the behaviour of these complex systems. In this project, we aim to combine probabilistic and statistical tools to study these processes. In particular, we will employ methods from branching processes, quasi-stationary distributions and interacting particle systems to study their long-term behaviour and develop numerical simulations. Further, we will develop likelihood-free methods to estimate the model parameters, followed by goodness-of-fit tests to analyse the strength of these methods when working with real data.

## 9.2 International research visitors

### 9.2.1 Visits of international scientists

#### Other international visits to the team

##### Emma Horton

**Status** Researcher

**Institution of origin:** University of Melbourne

**Country:** Australia

**Dates:** May 30 - June 2

**Context of the visit:** collaboration on the growth-coagulation-fragmentation processes.

**Mobility program/type of mobility:** research stay

##### Alex Watson

**Status** Researcher

**Institution of origin:** University College London

**Country:** UK

**Dates:** May 30 - June 2

**Context of the visit:** collaboration on the growth-coagulation-fragmentation processes in the framework of the MAGO Inria associate team.

**Mobility program/type of mobility:** research stay



### 9.2.2 Visits to international teams

#### Research stays abroad

##### **Coralie Fritsch & Denis Villemonais**

**Visited institution:** University College London

**Country:** UK

**Dates:** October 30 - November 3

**Context of the visit:** collaboration with Alex Watson and Emma Horton on the growth-coagulation-fragmentation processes in the framework of the MAGO Inria associate team.

**Mobility program/type of mobility:** research stay

##### **Nicolas Champagnat**

**Visited institution:** Pontificia Universidad Católica de Chile, Universidad de Valparaiso

**Country:** Chile

**Dates:** March 18 - March 27

**Context of the visit:** Collaboration with Pablo Marquet and Rolando Rebolledo on niche construction. After this visit, we applied for the Inria associate team aStoNiche (a Stochastic framework for modeling Niche construction), that is funded for the period 2024-2027.

**Mobility program/type of mobility:** research stay

##### **Pierre Vallois**

**Visited institution:** Turku University

**Country:** Finland

**Dates:** November 21 - November 25

**Context of the visit:** Collaboration with Paavo Salminen.

**Mobility program/type of mobility:** research stay

## 9.3 European initiatives

### 9.3.1 ERC projects

N. Champagnat is scientific collaborator of the ERC SINGER (AdG 101054787) on Stochastic dynamics of sINgle cells, coordinated by S. Méléard (Ecole Polytechnique). He is involved in the research axes “From stochastic processes to singular Hamilton-Jacobi equations” and “Lineages and time reversed trajectories” of this project.

## 9.4 National initiatives

- A. Gégout-Petit was in the committee interviewed by ANR for the IHU Infiny on the subject of chronic inflammatory bowel diseases. PI: L. Peyrin-Biroulet (Univ. Lorraine and CHRU Nancy).
- ITMO Physics, Mathematics applied to Cancer (from October 2023): “Quantifying and predicting the evolution of clonal heterogeneity in chronic lymphocytic leukemia”. **Funding organisms:** ITMO Cancer, ITMO Technologies pour la santé de l’alliance nationale pour les sciences de la vie et de la santé (AVIESAN), INCa. **Partners:** Inria and IECL (Institut Élie Cartan de Lorraine) and CHRU Strasbourg. **Leader:** N. Champagnat. **Participants:** C. Fritsch, U. Herbach, P. Vallois, D. Villemonais.

- PEPR Exploratoire Maths-VivES, (starting in spring 2024), target project DyLT (Dynamics of Telomere Length) on “Influence of telomere length dynamics and environmental conditions on biological and clinical aspects of aging”. **Funding organisms:** ANR. **Partners:** Inria Nancy and Saclay, Institut Élie Cartan de Lorraine (Nancy), CHRU Nancy, Centre de Recherche en Cancérologie de Marseille and Institut de recherche sur le cancer et le vieillissement (Nice). **Coordinators:** N. Champagnat and A. Benetos (CHRU Nancy). **Participants:** C. Fritsch, A. Gégout-Petit, D. Villemonais, S. Baland.
- PEPR Santé Numérique (started in July 2023), project AI4scMed (Multiscale AI for single-cell-based precision medicine) including WP3: “Regulatory network inference: from dynamical models to logical models”. **Funding organisms:** ANR. **Partners:** Inria, Inserm, CNRS. **Coordinator:** F. Picard (CNRS, ENS Lyon). **Participants:** M. Gaillard, U. Herbach.
- FHU CARTAGE (Fédération Hospitalo Universitaire Cardial and ARterial AGEing). **Leader:** Pr. A. Benetos. **Participants:** J.-M. Monnez, A. Gégout-Petit.
- ANR JCJC project CRESCENDO (inCrease physical Exercise and Sport to Combat ENDOmetriosis, AAPG 2022). **Coordinator:** G. Escriva-Boulley (LISEC, Université de Haute-Alsace). **Participant:** U. Herbach.
- GDR 720 IASIS (funded by CNRS). **Leader:** C. Richard. **Participant:** S. Wantz-Mézières.
- Réseau Thématique MathSAV (funded by CNRS). **Leader:** F. Crauste. **Participants:** N. Champagnat, C. Fritsch, V. Hass, U. Herbach, R. Loubaton, A. Rago, N. Zalduendo Vidal.
- Chair “Modélisation Mathématique et Biodiversité” between VEOLIA, Ecole Polytechnique, Museum National d’Histoire Naturelle and Fondation X (funded by VEOLIA). **Leader:** S. Méléard. **Participants:** V. Brodu, N. Champagnat, C. Fritsch, V. Hass, D. Villemonais, N. Zalduendo Vidal.

## 9.5 Regional initiatives

A. Gégout-Petit is one the two PIs of the interdisciplinary program “Life Travel” of the I-Site “Lorraine Université d’Excellence” on life trajectories and longevity (under construction).

## 10 Dissemination

**Participants:** Sophie Baland, Virgile Brodu, Nicolas Champagnat, Sandie Ferrigno, Coralie Fritsch, Mathilde Gaillard, Anne Gégout-Petit, Vincent Hass, Ulysse Herbach, Vincent Kagan, Rodolphe Loubaton, Jean-Marie Monnez, Anouk Rago, Pierre Vallois, Denis Villemonais, Sophie Wantz-Mézières, Nicolas Zalduendo Vidal.

### 10.1 Promoting scientific activities

#### 10.1.1 Scientific events: organisation

##### Member of the organizing committees

- N. Champagnat co-organized the conference [A Random Walk in the Land of Stochastic Analysis and Numerical Probability](#) (September 4-8, CIRM, Luminy) in the honor of Denis Talay.
- C. Fritsch co-organized the [21st INFORMS Applied Probability Society Conference](#) (Centre Prouvé, Nancy, June 28-30). S. Baland, V. Brodu, M. Gaillard, V. Hass, R. Loubaton, A. Rago and N. Zalduendo Vidal were members of the logistic crew during the conference.

- N. Champagnat and D. Villemonais organized the invited session “Quasi-stationary distributions in numerical stochastic methods and statistics” in the **21st INFORMS Applied Probability Society Conference** (Centre Prouvé, Nancy, June 28-30).
- D. Villemonais co-organized the **GdR Branchement first conference** in November 2023, Toulouse, France.
- U. Herbach has been co-organizing the Probability and Statistics weekly seminar at IECL in Nancy until September.

### 10.1.2 Scientific events: selection

#### Chair of conference program committees

- A. Gégout-Petit is chair program committee of the **ENBIS meeting 2024** that will be held in Leuven, Belgium in September 2024.

#### Member of the conference program committees

- A. Gégout-Petit was member of program committee of the **ENBIS meeting 2023** that was held in Valencia, Spain in September.

### 10.1.3 Journal

#### Member of the editorial boards

- N. Champagnat is associate editor for *ESAIM: Probability & Statistics* and *Stochastic Models*.
- A. Gégout-Petit was guest editor with L. Marco-Almagro (Univ. Politècnica de Catalunya, Barcelona, Spain) for the Quality and Reliability Engineering International **special issue related to the 22nd Annual Conference** of the European Network for Business and Industrial Statistics (ENBIS).

### 10.1.4 Invited talks

- N. Champagnat gave a plenary talk at the 11ème Biennale Française des Mathématiques Appliquées et Industrielles (**Congrès SMAI 2023**) in Le Gosier, Guadeloupe in May. He has been also invited to give talks at the **43rd Conference on Stochastic Processes and their Applications** in Lisbonne, Portugal in July, the conference **Celebrating the mathematics of Michel Benaïm** in Bernoulli Center, Lausanne in August, the conference **A random walk in the land of stochastic analysis and numerical probability** at CIRM, Luminy in September and the **International Conference on Recent Developments of Theory and Methods in Mathematical biology**, conference of the IRN ReaDiNet network at NCTS Taipei, Taiwan in October.
- C. Fritsch has been invited to give talks at the **21st INFORMS Applied Probability Society Conference** in Nancy in June and at the **Première conférence du GDR Branchement** in Toulouse in November.
- U. Herbach has been invited to give talks at **Statistics seminar of LMA** in Avignon in April, at **GT Bioss workshop** in Marseille in July, at **Inria MUSCA team seminar** in Saclay in September, at **Statistics seminar of IRMA** in Strasbourg in October and at **LCSB Systems Control group seminar** in Luxembourg in November.
- P. Vallois has been invited to give talks at the **Journées de Probabilités 2023** in Angers in June and at the **21st INFORMS Applied Probability Society Conference** in Nancy in June.
- D. Villemonais has been invited to give a talk at the **21st INFORMS Applied Probability Society Conference** in Nancy in June and at the conference **Celebrating the mathematics of Michel Benaïm** in Bernoulli Center, Lausanne in August.
- N. Zalduendo Vidal has been invited to give talks at the **21st INFORMS Applied Probability Society Conference** in Nancy in June and at the conference **Discrete Randomness** in Créteil in December.

### 10.1.5 Contributed talks, posters, workshops, seminars

- V. Brodu has presented a poster at the [Conférence internationale Mathematical Population Dynamics, Ecology and Evolution, MPDEE 2023](#) in Marseille in April, at the [21st INFORMS Applied Probability Society Conference](#) in Nancy in June (where he has been awarded a Best Poster prize), at the [43rd Conference on Stochastic Processes and their Applications](#) in Lisbon in July, and at the [GdR Branchement first conference](#) in Toulouse in November.
- N. Champagnat has been invited to give a talk at the [Ecodep-Biostochastic Workshop: Modeling Time Series and Stochastic Processes](#) at Las Cruces Marine Station, Chile in March. He has been also invited to give a (remote) seminar talk at the [Seminar of differential equations, Instytut Matematyczny Wrocław](#), Poland in November.
- C. Fritsch gave talks at the [Journées INRAE - Inria 2023](#) in Nancy in July and at the [Journée de la donnée en Meurthe et Moselle](#) in Nancy in October.
- U. Herbach has given talks at [Statistical Methods for Post Genomic Data workshop \(SMPGD 2023\)](#) in Ghent (Belgium) in February, at [21st International Conference on Computational Methods in Systems Biology \(CMSB 2023\)](#) in Luxembourg in September and at [CENTURI Conference on Information networks in biological systems](#) in Cargèse in October.
- A. Rago has presented a poster at [Journée scientifique autour de l'IA](#) in Nancy in February and at [Statlearn23](#) in Montpellier in April, and has given a talk [27] at [Journées de Statistique 2023](#) in Bruxelles in July.
- P. Vallois has given a seminar talk at Univ. Sorbonne Paris Nord in November.
- D. Villemonais gave a talk at the [Journée Santé Numérique](#) in Nancy in November.
- N. Zaluendo Vidal gave talks at the [Seminario de Probabilidades de Chile](#) (online) in August, at the [Workshop  \$L^2\$  in Probability and Statistics](#) in Metz in September and at the [Séminaire Image Optimisation et Probabilités](#) in Bordeaux in October.

### 10.1.6 Scientific expertise

- N. Champagnat evaluated a research project submitted to [Shape-Med@Lyon](#).
- C. Fritsch has been a member of the Committee for junior permanent research positions of Inria Nancy - Grand Est.
- A. Gégout-Petit was expert for the [Messidore call 2023 \(Méthodologie des ESSais cliniques Innovants, Dispositifs, Outils et Recherches Exploitant les données de santé et biobanques\)](#) of INSERM.
- A. Gégout-Petit was member of two different selection committees for Professor (Univ. de Pau et de l'Adour (UPPA) and Univ. Rennes 1) and one selection committee for a Maître de conférence for Aix-Marseille Université.
- U. Herbach evaluated a research project submitted to ANR AAPG 2023 as a scientific expert for the “Interfaces: mathematics, digital sciences - biology, health” panel.

### 10.1.7 Research administration

- V. Brodu is an elected representative of doctoral students at the doctoral school committee (local scale), and also at the doctoral college committee (regional scale).
- N. Champagnat is elected member of the *Commission d'Evaluation* of Inria since September, member of the COMIPERS (hiring committee for non-permanent positions) of Inria Nancy – Grand Est, substitute member of the *Comité de Centre* of Inria Nancy – Grand Est, local researcher (*correspondant local*) representing the COERLE (Inria's Ethic Committee) at Inria Nancy – Grand Est, and he was *responsable scientifique* for the library of Mathematics of IECL until October.

- C. Fritsch was a member, until August, of the *Commission du Développement Technologique* of the Inria Research Center of Nancy - Grand Est and of the *Commission du personnel* of IECL. She is an elected member of the *Commission d'Évaluation* of Inria since September.
- A. Gégout-Petit is director of the research unit IECL (Institut Elie Cartan de Lorraine), Mathematics laboratory of Univ. Lorraine (200 members).

## 10.2 Teaching - Supervision - Juries

### 10.2.1 Teaching

BIGS faculty members have teaching obligations at Univ. Lorraine and are teaching at least 192 hours each year. They teach probability and statistics at different levels (Licence, Master, Engineering school). Many of them have pedagogical responsibilities.

- T. Bastogne is in charge of the research master program “Santé Numérique et Imagerie Médicale” with the Faculty of Medicine, Univ. Lorraine.
- S. Ferrigno is in charge (since september 2023) of the “DU Big Data and Data Science” in ENSMN, Univ. Lorraine.
- D. Villemonais is the head of the Mathematical Engineering Major of ENSMN, Univ. Lorraine.
- Licence: V. Brodu, Probability theory tutorial, 40h, L3, first year of ENSMN, Univ. Lorraine.
- Licence: V. Brodu, Numerical Analysis tutorial, 20h, L3, first year of ENSMN, Univ. Lorraine.
- Master: N. Champagnat, Introduction to Quantitative Finance, 12h, M1, second year of ENSMN, Univ. Lorraine.
- Master: N. Champagnat, Introduction to Quantitative Finance, 9h, M2, third year of ENSMN, Univ. Lorraine.
- Master: S. Ferrigno, Experimental designs, 6h, M1, fourth year of EEIGM, Univ. Lorraine.
- Master: S. Ferrigno, Data analyzing and mining, 36h, M1, second year of ENSMN, Univ. Lorraine.
- Master: S. Ferrigno, Modeling and forecasting, 32h, M1, second year of ENSMN, Univ. Lorraine.
- Master: S. Ferrigno, Training projects, 18h, M1/M2, second and third year of ENSMN, Univ. Lorraine.
- Licence: S. Ferrigno, Descriptive and inferential statistics, 60h, L2, second year of EEIGM, Univ. Lorraine.
- Licence: S. Ferrigno, Statistical modeling, 60h, L2, second year of EEIGM, Univ. Lorraine.
- Licence: S. Ferrigno, Mathematical and computational tools, 20h, L3, third year of EEIGM, Univ. Lorraine.
- Licence: S. Ferrigno, Training projects, 40h, L1/L3, first, second and third year of EEIGM, Univ. Lorraine.
- Master: C. Fritsch, Inverse problem, 18h, M1, second year of ENSMN, Univ. Lorraine.
- Licence: C. Fritsch, Probability Theory tutorial, 27h, L3, first year of ENSMN, Univ. Lorraine.
- License: M. Gaillard, Numerical analysis and Optimization tutorial, 23h, L3, first year of ENSMN, Univ. Lorraine.
- Master: A. Gégout-Petit, Statistics, modeling, data analysis, 80h, master in applied mathematics, Univ. Lorraine.
- Licence: V. Hass, Mathématiques FIGIM 1A, 38h, L1/L2, first year of ENSMN, Univ. Lorraine.

- Licence: V. Hass, Mathématiques FIGIM 2A, 19h, L2, second year of ENSMN, Univ. Lorraine.
- Licence: V. Hass, Probabilités, 40h, L3, first year of ENSMN, Univ. Lorraine.
- Licence: V. Hass, Analyse numérique et optimisation, 45h, L3, first year of ENSMN, Univ. Lorraine.
- Licence: V. Hass, Recherche opérationnelle, 18h, L3, first year of ENSMN, Univ. Lorraine.
- Master: V. Hass, Méthodes stochastiques pour le calcul, 14h, M1, second year of ENSMN, Univ. Lorraine.
- Licence: R. Loubaton, Modélisation Statistique, 21.5h, L2, prépa intégrée de l'EEIGM, Univ. Lorraine.
- Licence: R. Loubaton, EDP, 20h, L2, prépa intégrée de l'EEIGM, Univ. Lorraine.
- Licence: R. Loubaton, Algèbre des matrices, 21.5h, L1, prépa intégrée de l'EEIGM, Univ. Lorraine.
- Licence: R. Loubaton, alcul différentiel, 21.5h, L1, prépa intégrée de l'EEIGM, Univ. Lorraine.
- Master : A. Rago, Analyse de données, 18h, M1, second year of ENSMN, Univ. Lorraine.
- Master : A. Rago, Statistiques pour la grande dimension, 18h, M2 IMSD/third year of ENSMN, Univ. Lorraine.
- Master: D. Villemonais, Probability Theory II, 63h, M1, second year of ENSMN, Univ. Lorraine.
- Master: D. Villemonais, Stochastic processes, 32h, Master 2 MFA, Univ. Lorraine.
- Master: D. Villemonais, Modeling and forecasting, 14h, M1, second year of ENSMN, Univ. Lorraine.
- License: D. Villemonais Probability Theory I, 57h, L3, first year of ENSMN, Univ. Lorraine.
- Master: S. Wantz-Mézières, Learning and analysis of medical data, 36h, with J.M. Moureaux, M2 SNIM, Univ. Lorraine.
- Licence: S. Wantz-Mézières, Applied mathematics for management, financial mathematics, Probability and Statistics, 160h, IUT Nancy-Charlemagne (L1/L2/L3), Univ. Lorraine.
- Licence: S. Wantz-Mézières, Probability, 100h, first year in TELECOM Nancy (initial and apprenticeship cursus), Univ. Lorraine.
- Licence: N. Zalduendo Vidal, Probability Theory tutorial, 20h, L3, first year of ENSMN, Univ. Lorraine.

### 10.2.2 Supervision

#### PhD

- PhD in progress: Sophie Baland, “Telomere length dynamics : modelisation, estimation and application to diagnostic support systems” since September 2023, funding LUE. Advisors: S. Toupance (Univ. Lorraine) and D. Villemonais.
- PhD in progress: Virgile Brodu, “Émergence des allométries dans les systèmes écologiques : comportement stationnaire de modèles déterministes et stochastiques de flux d'énergie et de biomasse”, grant ENS Lyon. Advisors: S. Billiard (Univ. Lille), N. Champagnat, C. Fritsch.
- PhD in progress: Mathilde Gaillard, “Processus de Markov déterministes par morceaux et inférence bayésienne de réseaux de gènes”, grant PEPR Santé Numérique, since October 2023. Advisors: A. Gégout-Petit, U. Herbach.

- PhD: Vincent Hass, “Individual-based models in adaptive dynamics and long time evolution under assumptions of rare advantageous mutations” [31], ATER in ENSMN. Defense on the 21st September. Advisor: N. Champagnat.
- PhD in progress: Anouar Jeddi, “Convergence of individual-based population models to Hamilton-Jacobi equations” since September 2023, grant ERC SINGER (Ecole Polytechnique). Advisors: S. Méléard (Ecole Polytechnique) and N. Champagnat.
- PhD in progress: Vincent Kagan, “Asymptotic behavior of epidemiological epidemiological models with individual viral load” since September 2023, funding Université de Lorraine. Advisors: E. Strickler (Univ. Lorraine) and D. Villemonais.
- PhD: Rodolphe Loubaton, “Caractérisation des cibles thérapeutiques dans un programme génique tumoral”, ATER in EEIGM. Defense on the 21st December. Advisors: N. Champagnat and L. Vallat (CHRU Strasbourg).
- PhD in progress: Anouk Rago, “Inférence de réseaux de gènes dynamiques et prédiction d’expériences d’interventions biologiques dans des cellules cancéreuses”, grant Région Grand-Est and Inria. Advisors: N. Champagnat, A. Gégout-Petit.
- PhD: Nicolás Zalduendo Vidal, “Processus de branchement bi-sexués multi-types”, grant Inria-Cordi. Defense on the 18th December. Advisors: C. Fritsch, D. Villemonais.

#### Other

- Engineer: Walid Laziri, “Flow cytometry data analysis” (Plan de relance, contract with the start-up EMOSIS), until November. Advisors: A. Gégout-Petit, S. Mézières.
- M2 internship: Mathilde Gaillard, “Couplage et vitesse de convergence de processus de Markov déterministes par morceaux” (Univ. Lyon 1). Advisor: U. Herbach.
- M2 internship: Anouar Jeddi, “Convergence of individual-based population models to Hamilton-Jacobi equations” (M2 MSV, Univ. Paris-Saclay) Advisors: S. Méléard (Ecole Polytechnique), S. Mirrahimi (Univ. Montpellier), V.C. Tran (Univ. G. Eiffel) and N. Champagnat.
- M2 internship: Vincent Kagan, “Asymptotic behavior of epidemiological epidemiological models with individual viral load”, funding Inria. Advisors: E. Strickler (IECL) and D. Villemonais.
- M2 SNIM Research Project: Marie Camonin, “Etude prospective de l’impact de certaines mutations génétiques dans le traitement des gliomes de bas grade”, Advisors: S. Wantz-Mézières et J.-M. Moureaux (CRAN).
- M2 IMSD Research Project: Nicolas Dinant, “Apport des statistiques spatiales dans l’étude de la localisation des tumeurs cérébrales”, Advisors: S. Wantz-Mézières and J.-M. Moureaux (CRAN).
- PIDR 2A Telecom-Nancy: A. Chevallier, A. Crivelli and T. Kieffer, “Machine Learning pour la mesure automatique du volume des tumeurs cérébrales”, Advisors: S. Wantz-Mézières et J.-M. Moureaux (CRAN).
- M1 ENSMN Research project: Antonin Clerc, “Émergence des allométries dans les écosystèmes” (full-year research project). Advisors: V. Brodu and C. Fritsch.
- M1 ENSMN Research project: two 2nd year students, “Continuous-time Markov chains” (from September). Advisor: M. Gaillard.

### 10.2.3 Juries

- N. Champagnat was referee for the PhD theses of Léo Meyer (Univ. Orléans, 09/10/2023), Van Hai Thai (Univ. Nantes, 28/09/2023), Imane Akjouj (Univ. Lille, 29/06/2023) and Anaïs Rat (Aix-Marseille Univ., 31/05/2023). He was also president of the PhD committee of Aleksian Ashot (Univ. Saint-Etienne, 20/11/2023). He was also examiner for the PhD thesis of Vincent Hass (Univ. Lorraine, 26/09/2023).
- N. Champagnat and P. Vallois were examiner for the PhD thesis of Rodolphe Loubaton (Univ. Lorraine, 21/12/2023).
- C. Fritsch and D. Villemonais were examiner for the PhD thesis of Nicolás Zalduendo-Vidal (Univ. Lorraine, 18/12/2023).
- A. Gégout-Petit was examiner for the HdR jury of Nathalie Krell (Univ. Rennes 1, 01/06/2023).
- A. Gégout-Petit was referee for the PhD theses of Maéva Kyheng (Université de Lille, 01/03/2023) ; Fatima Ezzahra MANA (Univ. Troyes, 03/09/2023). She was president of the PhD committee of Jérémie Frigério (Univ. Dijon, 18/12/2023) ; Jeremy Borderieu (AgroParistech, 14/12/2023) ; Rodolphe Loubaton (Univ. Lorraine, 21/12/2023) and examiner for the PhD of Nicolas Zalduendo (Univ. Lorraine, 18/12/2023).
- V. Brodu has been part of the jury for the thesis prize awarded by *Métropole du Grand Nancy* (rewarding PhD theses that are creative, innovative, and/or implanted in the local territory).

## 10.3 Popularization

### 10.3.1 Education

- V. Brodu supervised a maths club for highschool students in Lycée Jeanne d’Arc, Nancy. This club hosted a dozen students for two hours sessions, on a weekly basis.
- J.-M. Monnez wrote an unupdated version of the lecture notes [37] presented in the 2022 Activity Report.
- S. Wantz-Mézières organised a Research Training Week on Neuro-oncology and Numerics, for medical and engineering students in Nancy in March.

### 10.3.2 Interventions

- S. Ferrigno: Advisor of a group of EEIGM students, “Ateliers expérimentaux : Mécanique et Statistique” Project, various high schools, Nancy.
- S. Ferrigno: Advisor of a group of EEIGM students, “La main à la Pâte” Project, Institut médico-éducatif (IME), Commercy.
- S. Ferrigno: Advisor of a group of EEIGM students, “La main à la Pâte”, “CGénial” Projects, Collèges Paul Verlaine in Malzéville and La Craffe in Nancy.
- S. Ferrigno: Advisor of a group of EEIGM students, “La main à la Pâte” Project, elementary schools, Nancy.
- C. Fritsch gave three talks as part of the “Chiche!” program at Lycée De La Salle, Metz, in December.
- M. Gaillard gave a talk as part of the “Chiche!” program at Lycée René Cassin, Mâcon, in October. She presented her educational background and possible directions for a doctoral thesis in mathematics; explained what mathematical modeling is, and illustrated it with a stochastic gene expression model.



- U. Herbach gave a talk “Les maths peuvent-elles servir à vaincre le cancer ?” as part of a training course organized by [Maison pour la science en Lorraine](#) for secondary and high school teachers, at Faculté de Pharmacie de Nancy (campus Brabois) in October.
- A. Rago participated to the regional finals (Université de Lorraine) of “[MT180](#)”.

## 11 Scientific production

### 11.1 Major publications

- [1] R. Azaïs, S. Ferrigno and M.-J. Martinez. ‘cvmgof: an R package for Cramér-von Mises goodness-of-fit tests in regression models’. In: *Journal of Statistical Computation and Simulation* 92.6 (2022), pp. 1246–1266. DOI: [10.1080/00949655.2021.1991346](https://doi.org/10.1080/00949655.2021.1991346). URL: <https://hal.science/hal-03101612>.
- [2] R. Azaïs and A. Muller-Gueudin. ‘Optimal choice among a class of nonparametric estimators of the jump rate for piecewise-deterministic Markov processes’. In: *Electronic Journal of Statistics* 10.2 (2016), pp. 3648–3692. DOI: [10.1214/16-EJS1207](https://doi.org/10.1214/16-EJS1207). URL: <https://hal.science/hal-01168651>.
- [3] N. Champagnat, S. Méléard and V. C. Tran. ‘Multiscale eco-evolutionary models: from individuals to populations’. In: International Congress of Mathematicians, ICM 2022. Vol. 7. 1. fully virtually, Russia: EMS Press, 15th Dec. 2023, pp. 5656–5678. DOI: [10.4171/ICM2022/24](https://doi.org/10.4171/ICM2022/24). URL: <https://inria.hal.science/hal-04407958>.
- [4] N. Champagnat and D. Villemonais. ‘General criteria for the study of quasi-stationarity’. In: *Electronic Journal of Probability* (2023). DOI: [10.1214/22-EJP880](https://doi.org/10.1214/22-EJP880). URL: <https://inria.hal.science/hal-01672979>.
- [5] Y. Efroni, G. Dalal, B. Scherrer and S. Mannor. ‘How to Combine Tree-Search Methods in Reinforcement Learning’. In: AAAI 19 - Thirty-Third AAAI Conference on Artificial Intelligence. Honolulu, Hawaii, United States, 27th Jan. 2019. URL: <https://inria.hal.science/hal-02273713>.
- [6] E. Eyraud, E. Maurat, J.-M. Sac-Epee, P. Henrot, M. Zysman, P. Esteves, T. Trian, H. Bégueret, P.-O. Girodet, M. Thumerel, R. Hustache-Castaing, R. Marthan, F. Levet, P. Vallois, C. Contin-Bordes, P. Berger and I. Dupin. *Short-range interactions between fibrocytes and CD8+ T cells in COPD bronchial inflammatory response*. 21st Oct. 2022. DOI: [10.1101/2022.10.21.513138](https://doi.org/10.1101/2022.10.21.513138). URL: <https://hal.science/hal-03958836>.
- [7] C. Fritsch, D. Villemonais and N. Zaldueño. ‘The Multi-type Bisexual Galton-Watson Branching Process’. In: *Annales de l’I.H.P. Probabilités et statistiques* (2024). URL: <https://hal.science/hal-03696115>.
- [8] J.-M. Monnez. ‘Stochastic approximation of eigenvectors and eigenvalues of the Q-symmetric expectation of a random matrix’. In: *Communications in Statistics - Theory and Methods* 53.5 (2024), pp. 1669–1683. DOI: [10.1080/03610926.2022.2107225](https://doi.org/10.1080/03610926.2022.2107225). URL: <https://hal.science/hal-03956687>.
- [9] S. Toupance, D. Villemonais, D. Germain, A. Gégout-Petit, E. Albuissou and A. Benetos. ‘The individual’s signature of telomere length distribution’. In: *Scientific Reports* 9.1 (24th Jan. 2019), pp. 1–8. DOI: [10.1038/s41598-018-36756-8](https://doi.org/10.1038/s41598-018-36756-8). URL: <https://inria.hal.science/hal-01925000>.
- [10] E. Ventre, U. Herbach, T. Espinasse, G. Benoit and O. Gandrillon. ‘One model fits all: Combining inference and simulation of gene regulatory networks’. In: *PLoS Computational Biology* 19.3 (27th Mar. 2023), e1010962. DOI: [10.1371/journal.pcbi.1010962](https://doi.org/10.1371/journal.pcbi.1010962). URL: <https://inria.hal.science/hal-04071033>.

## 11.2 Publications of the year

### International journals

- [11] A. Anagnostakis, A. Lejay and D. Villemonais. ‘General diffusion processes as the limit of time-space Markov chains’. In: *The Annals of Applied Probability* 33.5 (2023), pp. 3620–3651. DOI: [10.48550/arXiv.2206.03713](https://doi.org/10.48550/arXiv.2206.03713). URL: <https://inria.hal.science/hal-02897819>.
- [12] T. Bastogne, L. Hassler, J. Bruniaux, M. Thomassin, X. Gidrol, E. Sulpice and F. Navarro. ‘A Bayesian implementation of quality-by-design for the development of cationic nano-lipid for siRNA transfection’. In: *IEEE Transactions on NanoBioscience* 22.3 (July 2023), pp. 455–466. DOI: [10.1109/TNB.2022.3213412](https://doi.org/10.1109/TNB.2022.3213412). URL: <https://hal.science/hal-04212321>.
- [13] N. Champagnat and V. Hass. ‘Existence, uniqueness and ergodicity for the centered Fleming-Viot process’. In: *Stochastic Processes and their Applications* 166 (Dec. 2023), p. 104219. DOI: [10.1016/j.spa.2023.09.006](https://doi.org/10.1016/j.spa.2023.09.006). URL: <https://hal.science/hal-03596777>.
- [14] N. Champagnat, S. Méléard, S. Mirrahimi and V. Chi Tran. ‘Filling the gap between individual-based evolutionary models and Hamilton-Jacobi equations’. In: *Journal de l’École polytechnique — Mathématiques* 10 (2023), pp. 1247–1275. DOI: [10.5802/jep.244](https://doi.org/10.5802/jep.244). URL: <https://hal.science/hal-03656608>.
- [15] N. Champagnat and D. Villemonais. ‘General criteria for the study of quasi-stationarity’. In: *Electronic Journal of Probability* (2023). DOI: [10.1214/22-EJP880](https://doi.org/10.1214/22-EJP880). URL: <https://inria.hal.science/hal-01672979>.
- [16] B. Cloez and C. Fritsch. ‘Quasi-stationary behavior for a piecewise deterministic Markov model of chemostat: the Crump-Young model’. In: *Annales Henri Lebesgue* (2024). URL: <https://inria.hal.science/hal-03660678>.
- [17] A. M. Cox, E. Horton and D. Villemonais. ‘Binary branching processes with Moran type interactions’. In: *Annales de l’I.H.P. Probabilités et statistiques* (2024). URL: <https://inria.hal.science/hal-03715678>.
- [18] A. Dijamentiuk, C. Mangavel, A. Elfassy, F. Michaux, J. Burgain, E. Rondags, S. Delaunay, S. Ferrigno, A.-M. Revol-Junelles and F. Borges. ‘Invert emulsions alleviate biotic interactions in bacterial mixed culture’. In: *Microbial Cell Factories* 22.1 (Dec. 2023), p. 16. DOI: [10.1186/s12934-022-02014-w](https://doi.org/10.1186/s12934-022-02014-w). URL: <https://hal.univ-lorraine.fr/hal-03953726>.
- [19] I. Dupin, E. Eyraud, É. Maurat, J.-M. Sac-Épée and P. Vallois. ‘Probabilistic Cellular Automata modeling of intercellular interactions in airways : complex pattern formation in patients with Chronic Obstructive Pulmonary Disease’. In: *Journal of Theoretical Biology* (5th Mar. 2023), p. 111448. DOI: [10.1016/j.jtbi.2023.111448](https://doi.org/10.1016/j.jtbi.2023.111448). URL: <https://hal.science/hal-03572045>.
- [20] G. Escriva-Boulley, C.-A. Philip, S. Warembourg, L. Lenôtre, P. Flore, P. Faure, T. Michy, V. Letouzey, C. Arnold, C. Piluso, L. Chalmel, R. Kacem, G. F. Blum, R. Detayrac, C. Trocmé, I. Brigaud, U. Herbach, P. Branche, E. Faller and A. Chalabaev. ‘Effects of a physical activity and endometriosis-based education program delivered by videoconference on endometriosis symptoms: the CRESCENDO program (inCREase physical Exercise and Sport to Combat ENDometriosis) protocol study’. In: *Trials* 24.1 (27th Nov. 2023), p. 759. DOI: [10.1186/s13063-023-07792-1](https://doi.org/10.1186/s13063-023-07792-1). URL: <https://hal.science/hal-04365506>.
- [21] C. Fritsch, D. Villemonais and N. Zaldueño. ‘The Multi-type Bisexual Galton-Watson Branching Process’. In: *Annales de l’I.H.P. Probabilités et statistiques* (2024). URL: <https://hal.science/hal-03696115>.
- [22] N. P. Gao, O. Gandrillon, A. Páldi, U. Herbach and R. Gunawan. ‘Single-cell transcriptional uncertainty landscape of cell differentiation’. In: *F1000Research* 12 (20th Apr. 2023), p. 426. DOI: [10.12688/f1000research.131861.2](https://doi.org/10.12688/f1000research.131861.2). URL: <https://hal.science/hal-04082461>.
- [23] S. Janson, C. Mailler and D. Villemonais. ‘Fluctuations of balanced urns with infinitely many colours’. In: *Electronic Journal of Probability* 28 (2023), pp. 1–72. DOI: [10.1214/23-EJP951](https://doi.org/10.1214/23-EJP951). URL: <https://hal.science/hal-03508437>.

- [24] J.-M. Monnez. ‘Stochastic approximation of eigenvectors and eigenvalues of the Q-symmetric expectation of a random matrix’. In: *Communications in Statistics - Theory and Methods* 53.5 (2024), pp. 1669–1683. DOI: [10.1080/03610926.2022.2107225](https://doi.org/10.1080/03610926.2022.2107225). URL: <https://hal.science/hal-03956687>.
- [25] E. Ventre, U. Herbach, T. Espinasse, G. Benoit and O. Gandrillon. ‘One model fits all: Combining inference and simulation of gene regulatory networks’. In: *PLoS Computational Biology* 19.3 (27th Mar. 2023), e1010962. DOI: [10.1371/journal.pcbi.1010962](https://doi.org/10.1371/journal.pcbi.1010962). URL: <https://inria.hal.science/hal-04071033>.

#### International peer-reviewed conferences

- [26] N. Champagnat, S. Méléard and V. C. Tran. ‘Multiscale eco-evolutionary models: from individuals to populations’. In: International Congress of Mathematicians, ICM 2022. Vol. 7. 1. fully virtually, Russia: EMS Press, 15th Dec. 2023, pp. 5656–5678. DOI: [10.4171/ICM2022/24](https://doi.org/10.4171/ICM2022/24). URL: <https://inria.hal.science/hal-04407958>.
- [27] A. Rago, N. Champagnat, A. Gégout-Petit and L. Vallat. ‘Simulation d’expériences d’intervention biologique dans des cellules cancéreuses à partir de données temporelles d’expression de gènes’. In: 54es Journées de Statistique de la SFdS (JdS 2023). Bruxelles, Belgium, 2023. URL: <https://inria.hal.science/hal-04408929>.

#### Conferences without proceedings

- [28] W. Laziri, F. Allemand, A. Gégout-Petit and S. Wantz-Mézières. ‘Méthodes de réduction de dimension basées sur l’algorithme FINE pour le clustering de patients à partir de données de cytométrie en flux’. In: 54es Journées de Statistiques de la SFdS. Bruxelles (BEL), Belgium, 3rd July 2023. URL: <https://hal.science/hal-04406372>.
- [29] W. Laziri, A. Gégout-Petit, F. Allemand and S. Wantz-Mézières. ‘Dimension reduction methods based on FINE algorithm for clustering patients from flow cytometry data’. In: ENBIS 2023 Conference. Valencia, Spain, 10th Sept. 2023. URL: <https://hal.science/hal-04406413>.

#### Scientific book chapters

- [30] U. Herbach. ‘Harissa: Stochastic Simulation and Inference of Gene Regulatory Networks Based on Transcriptional Bursting’. In: *Computational Methods in Systems Biology*. Vol. 14137. Lecture Notes in Computer Science. Springer, 9th Sept. 2023, pp. 97–105. DOI: [10.1007/978-3-031-42697-1\\_7](https://doi.org/10.1007/978-3-031-42697-1_7). URL: <https://hal.science/hal-04208601>.

#### Doctoral dissertations and habilitation theses

- [31] V. Hass. ‘Individual-based models in adaptive dynamics, asymptotic behaviour and canonical equation : the case of small and frequent mutations’. Université de Lorraine, 26th Sept. 2023. URL: <https://hal.univ-lorraine.fr/tel-04295339>.

#### Reports & preprints

- [32] A. Benetos, O. Coudray, A. Gégout-Petit, L. Lenôtre, S. Toupance and D. Villemonais. *A branching model for intergenerational telomere length dynamics*. 15th Oct. 2023. URL: <https://hal.science/hal-04242595>.
- [33] N. Champagnat and V. Hass. *Convergence of individual-based models with small and frequent mutations to the canonical equation of adaptive dynamics*. 17th Mar. 2023. URL: <https://hal.science/hal-04034027>.
- [34] N. Champagnat, E. Strickler and D. Villemonais. *Uniform Wasserstein convergence of penalized Markov processes*. 28th June 2023. URL: <https://hal.science/hal-04145204>.

- [35] R. Loubaton, N. Champagnat, P. Vallois and L. Vallat. *MultiRNAflow: integrated analysis of temporal RNA-seq data with multiple biological conditions*. 2023. URL: <https://inria.hal.science/hal-04407596>.
- [36] J.-M. Monnez. *An extended Oja process for streaming canonical analysis*. 2023. URL: <https://hal.science/hal-04351700>.

### 11.3 Other

#### Educational activities

- [37] J.-M. Monnez. ‘Analyse des données en flux. Analyse en composantes principales et méthodes dérivées’. Doctoral. France, 19th June 2023. URL: <https://hal.science/hal-03690459>.

#### Softwares

- [38] [SW] U. Herbach, *Harissa: tools for mechanistic gene network inference from single-cell data* version 3.0.7, 2023. LIC: BSD 3-Clause. HAL: [hal-03370296](https://hal.science/hal-03370296), URL: <https://hal.science/hal-03370296>, VCS: <https://github.com/ulysseherbach/harissa>, SWHID: [sw:1:dir:315fa8cd-b8abd2801f545dca59dd136584572a3d](https://sw.hal.science/sw/1/dir/315fa8cd-b8abd2801f545dca59dd136584572a3d).

### 11.4 Cited publications

- [39] M. Andrade-Restrepo, N. Champagnat and R. Ferrière. ‘Spatial eco-evolutionary dynamics along environmental gradients: multi-stability and cluster dynamics’. In: *Ecology Letters* 22.5 (May 2019), pp. 767–777. URL: <https://hal.inria.fr/hal-01732325>.
- [40] R. Azaïs, S. Ferrigno and M.-J. Martinez. ‘cvmgof: an R package for Cramér-von Mises goodness-of-fit tests in regression models’. In: *Journal of Statistical Computation and Simulation* 92.6 (2022), pp. 1246–1266. DOI: [10.1080/00949655.2021.1991346](https://doi.org/10.1080/00949655.2021.1991346). URL: <https://hal.science/hal-03101612>.
- [41] [SW] R. Azaïs, S. Ferrigno and M.-J. Martinez, *cvmgof: Cramer-von Mises goodness-of-fit tests* version 1.0.0, Nov. 2018. LIC: CeCILL. HAL: [hal-02014516](https://hal.archives-ouvertes.fr/hal-02014516), URL: <https://hal.archives-ouvertes.fr/hal-02014516>, SWHID: [sw:1:dir:6b9e75ac30a78bdf2d928b2177e8edf538c16748](https://sw.hal.science/sw/1/dir/6b9e75ac30a78bdf2d928b2177e8edf538c16748).
- [42] R. Azaïs. ‘A recursive nonparametric estimator for the transition kernel of a piecewise-deterministic Markov process’. In: *ESAIM: Probability and Statistics* 18 (2014), pp. 726–749.
- [43] R. Azaïs, F. Dufour and A. Gégout-Petit. ‘Nonparametric estimation of the jump rate for non-homogeneous marked renewal processes’. In: *Annales de l’Institut Henri Poincaré, Probabilités et Statistiques*. Vol. 49. 4. Institut Henri Poincaré. 2013, pp. 1204–1231.
- [44] T. Bastogne, S. Mézières-Wantz, N. Ramdani, P. Vallois and M. Barberi-Heyob. ‘Identification of pharmacokinetics models in the presence of timing noise’. In: *Eur. J. Control* 14.2 (2008), pp. 149–157. DOI: [10.3166/ejc.14.149-157](https://doi.org/10.3166/ejc.14.149-157). URL: <http://dx.doi.org/10.3166/ejc.14.149-157>.
- [45] T. Bastogne, A. Samson, P. Vallois, S. Wantz-Mézières, S. Pinel, D. Bechet and M. Barberi-Heyob. ‘Phenomenological modeling of tumor diameter growth based on a mixed effects model’. In: *Journal of theoretical biology* 262.3 (2010), pp. 544–552.
- [46] D. Bertsekas and J. Tsitsiklis. *Neurodynamic Programming*. Athena Scientific, 1996.
- [47] H. Cardot, P. Cénac and J.-M. Monnez. ‘A fast and recursive algorithm for clustering large datasets with k-medians’. In: *Computational Statistics and Data Analysis* 56 (2012), pp. 1434–1449. DOI: [10.1016/j.csda.2011.11.019](https://doi.org/10.1016/j.csda.2011.11.019). URL: <https://hal.science/hal-00644683>.
- [48] N. Champagnat, P.-E. Jabin and S. Méléard. ‘Adaptation in a stochastic multi-resources chemostat model’. In: *Journal de Mathématiques Pures et Appliquées* 101.6 (June 2014), pp. 755–788. DOI: [10.1016/j.matpur.2013.10.003](https://doi.org/10.1016/j.matpur.2013.10.003). URL: <https://hal.inria.fr/hal-00784166>.

- [49] N. Champagnat and D. Villemonais. ‘Exponential convergence to quasi-stationary distribution and Q-process’. In: *Probability Theory and Related Fields* 164.1 (2016), 46 pages, pp. 243–283. DOI: [10.1007/s00440-014-0611-7](https://doi.org/10.1007/s00440-014-0611-7). URL: <https://hal.archives-ouvertes.fr/hal-00973509>.
- [50] N. Champagnat and D. Villemonais. ‘Practical criteria for R-positive recurrence of unbounded semigroups’. In: *Electronic Communications in Probability* 25.6 (2020), pp. 1–11. DOI: [10.1214/20-ECP288](https://doi.org/10.1214/20-ECP288). URL: <https://hal.science/hal-02101680>.
- [51] M. H. Davis. ‘Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models’. In: *Journal of the Royal Statistical Society. Series B (Methodological)* (1984), pp. 353–388.
- [52] M. Doumic, M. Hoffmann, N. Krell and L. Robert. ‘Statistical estimation of a growth-fragmentation model observed on a genealogical tree’. In: *Bernoulli* 21.3 (2015), pp. 1760–1799.
- [53] K. Duarte, J.-M. Monnez and E. Albuissou. ‘Sequential linear regression with online standardized data’. In: *PLoS ONE* (2018), pp. 1–27. DOI: [10.1371/journal.pone.0191186](https://doi.org/10.1371/journal.pone.0191186). URL: <https://hal.science/hal-01538125>.
- [54] E. Eyraud, E. Maurat, J.-M. Sac-Epee, P. Henrot, M. Zysman, P. Esteves, T. Trian, H. Bégueret, P.-O. Girodet, M. Thumerel, R. Hustache-Castaing, R. Marthan, F. Levet, P. Vallois, C. Contin-Bordes, P. Berger and I. Dupin. ‘Short-range interactions between fibrocytes and CD8+ T cells in COPD bronchial inflammatory response’. In: *eLife* (Oct. 2022). Microscopy was performed at BIC, a service unit of the CNRS-INSERM and Bordeaux University, a member of the national BioImaging infrastructure of France supported by the French National Research Agency (ANR-10-INBS-0004). DOI: [10.1101/2022.10.21.513138](https://doi.org/10.1101/2022.10.21.513138). URL: <https://hal.science/hal-03958836>.
- [55] S. Ferrigno and G. Ducharme. ‘Un test d’adéquation global pour la fonction de répartition conditionnelle’. In: *C. R. Math. Acad. Sci. Paris* 341.5 (2005), pp. 313–316. DOI: [10.1016/j.crma.2005.07.003](https://doi.org/10.1016/j.crma.2005.07.003). URL: <http://dx.doi.org/10.1016/j.crma.2005.07.003>.
- [56] S. Ferrigno, M. Maumy-Bertrand and A. Muller-Gueudin. ‘Uniform law of the logarithm for the local linear estimator of the conditional distribution function’. In: *C. R. Math. Acad. Sci. Paris* 348.17-18 (2010), pp. 1015–1019. DOI: [10.1016/j.crma.2010.08.003](https://doi.org/10.1016/j.crma.2010.08.003). URL: <http://dx.doi.org/10.1016/j.crma.2010.08.003>.
- [57] J. Friedman, T. Hastie and R. Tibshirani. ‘Sparse inverse covariance estimation with the graphical lasso’. In: *Biostatistics* 9.3 (2008), pp. 432–441.
- [58] C. Fritsch, F. Campillo and O. Ovaskainen. ‘A numerical approach to determine mutant invasion fitness and evolutionary singular strategies’. In: *Theoretical Population Biology* 115 (2017), pp. 89–99. DOI: [10.1016/j.tpb.2017.05.001](https://doi.org/10.1016/j.tpb.2017.05.001). URL: <https://hal.science/hal-01413638>.
- [59] C. Giraud, S. Huet and N. Verzelen. ‘Graph selection with GGMselect’. In: *Statistical applications in genetics and molecular biology* 11.3 (2012).
- [60] T. Hansen and U. Zwick. ‘Lower Bounds for Howard’s Algorithm for Finding Minimum Mean-Cost Cycles’. In: *ISAAC (I)*. 2010, pp. 415–426.
- [61] U. Herbach, A. Bonnaïffoux, T. Espinasse and O. Gandrillon. ‘Inferring gene regulatory networks from single-cell data: a mechanistic approach’. In: *BMC Systems Biology* 11.1 (Nov. 2017), p. 105. DOI: [10.1186/s12918-017-0487-0](https://doi.org/10.1186/s12918-017-0487-0). URL: <https://hal.science/hal-01646910>.
- [62] S. Herrmann and P. Vallois. ‘From persistent random walk to the telegraph noise’. In: *Stoch. Dyn.* 10.2 (2010), pp. 161–196. DOI: [10.1142/S0219493710002905](https://doi.org/10.1142/S0219493710002905). URL: <http://dx.doi.org/10.1142/S0219493710002905>.
- [63] J. Hu, W.-C. Wu and S. Sastry. ‘Modeling subtilin production in bacillus subtilis using stochastic hybrid systems’. In: *Hybrid Systems: Computation and Control*. Springer, 2004, pp. 417–431.
- [64] R. Keinj, T. Bastogne and P. Vallois. ‘Multinomial model-based formulations of TCP and N TCP for radiotherapy treatment planning’. In: *Journal of Theoretical Biology* 279.1 (June 2011), pp. 55–62. DOI: [10.1016/j.jtbi.2011.03.025](https://doi.org/10.1016/j.jtbi.2011.03.025). URL: <http://hal.inria.fr/hal-00588935/en>.
- [65] R. Koenker. *Quantile regression*. 38. Cambridge university press, 2005.

- [66] L. Lebart. 'On the Benzecri's method for computing eigenvectors by stochastic approximation (the case of binary data)'. In: *Compstat 1974 (Proc. Sympos. Computational Statist., Univ. Vienna, Vienna, 1974)*. Vienna: Physica Verlag, 1974, pp. 202–211.
- [67] B. Lesner and B. Scherrer. 'Non-Stationary Approximate Modified Policy Iteration'. In: *ICML 2015*. Lille, France, July 2015. URL: <https://hal.inria.fr/hal-01186664>.
- [68] A. Lorz, S. Mirrahimi and B. Perthame. 'Dirac mass dynamics in multidimensional nonlocal parabolic equations'. In: *Communications in Partial Differential Equations* 36.6 (2011), pp. 1071–1098.
- [69] N. Meinshausen and P. Bühlmann. 'High-dimensional graphs and variable selection with the lasso'. In: *The Annals of Statistics* (2006), pp. 1436–1462.
- [70] J.-M. Monnez. 'Approximation stochastique en analyse factorielle multiple'. In: *Ann. I.S.U.P.* 50.3 (2006), pp. 27–45.
- [71] J.-M. Monnez. 'Convergence d'un processus d'approximation stochastique en analyse factorielle'. In: *Publ. Inst. Statist. Univ. Paris* 38.1 (1994), pp. 37–55.
- [72] J.-M. Monnez. 'Stochastic approximation of the factors of a generalized canonical correlation analysis'. In: *Statist. Probab. Lett.* 78.14 (2008), pp. 2210–2216. DOI: [10.1016/j.spl.2008.01.088](https://doi.org/10.1016/j.spl.2008.01.088). URL: <http://dx.doi.org/10.1016/j.spl.2008.01.088>.
- [73] E. Nadaraya. 'On non-parametric estimates of density functions and regression curves'. In: *Theory of Probability & Its Applications* 10.1 (1965), pp. 186–190.
- [74] I. Post and Y. Ye. *The simplex method is strongly polynomial for deterministic Markov decision processes*. Tech. rep. arXiv:1208.5083v2, 2012.
- [75] M. Puterman. *Markov Decision Processes*. Wiley, New York, 1994.
- [76] A. Richard, L. Boullu, U. Herbach, A. Arnaud, V. Morin, E. Vallin, A. Guillemin, N. Papili Gao, R. Gunawan, J. Cosette, O. Arnaud, J.-J. Kupiec, T. Espinasse, S. Gonin-Giraud, O. Gandrillon and S. Teichmann. 'Single-Cell-Based Analysis Highlights a Surge in Cell-to-Cell Molecular Variability Preceding Irreversible Commitment in a Differentiation Process'. In: *PLoS Biology* 14.12 (Dec. 2016). DOI: [10.1371/journal.pbio.1002585](https://doi.org/10.1371/journal.pbio.1002585). URL: <https://hal.science/hal-01934489>.
- [77] B. Roynette, P. Vallois and M. Yor. 'Brownian penalisations related to excursion lengths, VII'. In: *Annales de l'IHP Probabilités et statistiques*. Vol. 45. 2. 2009, pp. 421–452.
- [78] F. Russo and P. Vallois. 'Stochastic calculus with respect to continuous finite quadratic variation processes'. In: *Stochastics: An International Journal of Probability and Stochastic Processes* 70.1-2 (2000), pp. 1–40.
- [79] B. Scherrer. 'Approximate Policy Iteration Schemes: A Comparison'. In: *ICML - 31st International Conference on Machine Learning - 2014*. Pékin, China, June 2014. URL: <https://hal.inria.fr/hal-00989982>.
- [80] B. Scherrer. 'Improved and Generalized Upper Bounds on the Complexity of Policy Iteration'. In: *Mathematics of Operations Research* (Feb. 2016). Markov decision processes ; Dynamic Programming ; Analysis of Algorithms. DOI: [10.1287/moor.2015.0753](https://doi.org/10.1287/moor.2015.0753). URL: <https://hal.inria.fr/hal-00829532>.
- [81] B. Scherrer, M. Ghavamzadeh, V. Gabillon, B. Lesner and M. Geist. 'Approximate Modified Policy Iteration and its Application to the Game of Tetris'. In: *Journal of Machine Learning Research* 16 (2015). A paraître, pp. 1629–1676. URL: <https://hal.inria.fr/hal-01091341>.
- [82] P. Vallois. 'The range of a simple random walk on  $Z$ '. In: *Advances in applied probability* (1996), pp. 1014–1033.
- [83] P. Vallois and C. S. Tapiero. 'Memory-based persistence in a counting random walk process'. In: *Phys. A*. 386.1 (2007), pp. 303–307. DOI: [10.1016/j.physa.2007.08.027](https://doi.org/10.1016/j.physa.2007.08.027). URL: <http://dx.doi.org/10.1016/j.physa.2007.08.027>.
- [84] N. Villa-Vialaneix. *An introduction to network inference and mining*. <http://wikistat.fr/>. (consulté le 22/07/2015). 2015. URL: [http://www.nathalievilla.org/doc/pdf/wikistat-network\\_compiled.pdf](http://www.nathalievilla.org/doc/pdf/wikistat-network_compiled.pdf).

- [85] Y. Ye. ‘The Simplex and Policy-Iteration Methods Are Strongly Polynomial for the Markov Decision Problem with a Fixed Discount Rate’. In: *Math. Oper. Res.* 36.4 (2011), pp. 593–603.