2021
ACTIVITY REPORT

# Project-Team
# ASCII

## Analysis of Stochastic Cooperative Intelligent Interactions

**IN COLLABORATION WITH:** Centre de Mathématiques Appliquées (CMAP)

**DOMAIN**

**Applied Mathematics, Computation and Simulation**

**THEME**

**Stochastic approaches**

# Contents

# Project-Team ASCII

*Creation of the Project-Team: 2019 November 01*

# Keywords

## Computer sciences and digital sciences

A6.1.2. – Stochastic Modeling

A6.1.3. – Discrete Modeling (multi-agent, people centered)

A6.2.2. – Numerical probability

A6.2.3. – Probabilistic methods

A6.2.6. – Optimization

A6.2.7. – High performance computing

A6.3.4. – Model reduction

A6.3.5. – Uncertainty Quantification

## Other research topics and application domains

B1. – Life sciences

# 1   Team members, visitors, external collaborators

**Research Scientists**

- Denis Talay [Team leader, Inria, Senior Researcher, HDR]

- Carl Graham [CNRS, Researcher]

**Faculty Members**

- Mao-Fabrice Djete [Ecole Polytechnique]

- Josselin Garnier [École polytechnique, Professor]

- Nizar Touzi [École polytechnique, Professor]

**PhD Students**

- Naoufal Acharki [TOTAL-Pau]

- Leila Bassou [École polytechnique]

- Guillaume Chennetier [École polytechnique]

- Alexis Cousin [IFPEN, until Sep 2021]

- Assil Fadle [École polytechnique, from Sep 2021]

- Clement Gauchy [CEA]

- Corentin Houpert [CEA]

- Baptiste Kerleguer [CEA]

- Mehdi Talbi [École polytechnique]

- Songbo Wang [Ecole Polytechnique]

**Technical Staff**

- Maxime Colomb [Inria, Engineer, until Mar 2021]

- Nicolas Gilet [Inria, Engineer]

**Administrative Assistant**

- Marie Enee [Inria]

# 2   Overall objectives

The stochastic particles concerned by the ASCII research behave like agents whose interactions are **stochastic and intelligent**; these agents desire to optimally cooperate towards a common objective by solving strategic optimisation problems. ASCII's overall objective consists in developing their theoretical and numerical analyses with target applications in economics, neuroscience, physics, biology, stochastic numerics. To the best of our knowledge, this challenging objective is quite innovative.

In addition to modelling challenges raised by our innovating approaches to handle intelligent multi-agent interactions, we develop new mathematical theories and numerical methods to deal with interactions which, in most of the interesting cases we have in mind, are irregular, non Markovian, and evoluting. In particular, original and non-standard stochastic control and stochastic optimization methodologies are being developed, combined with original specific calibration methodologies.

To reach our objectives, we combine various mathematical techniques coming from stochastic analysis, partial differential equation analysis, numerical probability, optimization theory, and stochastic control theory.

# 3   Research program

Concerning particle systems with singular interactions, in addition to the convergence to mean–field limits and the analysis of convergence rates of relevant discretizations, one of our main challenges will concern the simulation of *complex, singular and large scale* McKean-Vlasov particle systems and stochastic partial differential equations, with a strong emphasis on the detection of numerical instabilities and potential big approximation errors.

The determination of blow-up times is also a major issue for spectrum approximation and criticality problems in neutron transport theory, Keller-Segel models for chemotaxis, financial bubble models, etc.

Reliability assessment for power generation systems or subsystems is another target application of our research. For such complex systems, standard Monte Carlo methods are inefficient because of the difficulties to appropriately simulate rare events. We thusdevelop algorithms based on particle filters methods combined with suitable variance reduction methods.

Exhibiting optimal regulation procedures in a stochastic environment is an important challenge in many fields. As emphasized above, in the situations we are interested in, the agents do not compete but concur to their regulation. We here give three examples: the control of cancer therapies, regulation and mechanism design by optimal contracting, and the distributed control for the planning problem.

Optimal contracting is widely used in economics in order to model agents interactions subject to the so-called moral hazard problem. This is best illustrated by the works of Tirole (2014 Nobel Prize in economics) in industrial economics. The standard situation is described by the interaction of two parties. The principal (e.g. land owner) hires the agent (e.g. farmer) in order to delegate the management of an output of interest (e.g. production of the land). The agent receives a salary as a compensation for the costly effort devoted to the management of the output. The principal only observes the output value, and has no access to the agent effort. Due to the cost of effort, the agent may divert his effort from the direction desired by the principal. The contract is proposed by the principal, and chosen according to a Stackelberg game: anticipating the agent's optimal response to any contract, she searches for the optimal contract by optimizing her utility criterion.

We are developing the continuous time formulation of this problem and allow for diffusion control and for possible competing multiple agents and principals. This is achieved by using crucially the recently developed second order backward stochastic differential equations, which act as HJB equations in the present non-Markovian framework.

The current environmental transition requires governments to incite firms to introduce green technologies as a substitute the the outdated polluting ones. This transition requires appropriate incentive schemes so as to reach the overall transition objective. This problem can be formulated in the framework of the last Principal-Agent problem as follows. Governments act as principals by setting the terms of an incentive regulation based on subsidies and tax reductions. Firms acting as agents optimize their production strategies given the regulation imposed by the governments. Such incentive schemes are also provided by the the refinancing channel through private investor, as best witnessed by the remarkable growth of green bonds markets.

Another motivation comes from Mechanism design. Modern decentralized facilities are present throughout our digitally connected economies. With the fragmentation of financial markets, exchanges are nowadays in competition. As the traditional international exchanges are now challenged by alternative trading venues, markets have to find innovative ways to attract liquidity. One solution is to use a make-taker fees system, that is, a rule enabling them to charge in an asymmetric way liquidity provision and liquidity consumption. The most classical setting, used by many exchanges (such as Nasdaq, Euronext, BATS Chi-X,...), consists in subsidizing the former while taxing the latter. In practice, this results in associating a fee rebate to executed limit orders and applying a transaction cost for market orders.

A platform aims at attracting two types of agents: market makers post bid and ask prices for some underlying asset or commodity, and brokers fulfill their trading needs if the posted prices are convenient. The platform takes its benefits from the fees gained on each transaction. As transactions only occur

when the market makers takes on more risky behavior by posting interesting bid and ask prices, the platform (acting as the principal) sets the terms of an incentive compensation to the market makers (acting as agents) for each realized transaction. Consequently, this optimal contracting problem serves as an optimization tool for the mechanism design of the platform.

Inspired by optimal transport theory, we formulate the above regulation problem as the interaction between a principal and a "crowd" of symmetric agents. Given the large number of agents, we model the limiting case of a continuum of agents whose state is then described by their distribution. The mean field game formulates the interacting agents optimal decision according to a Nash equilibrium competition. The optimal planning problem, introduced by Pierre-Louis Lions, seeks for an incentivizing scheme of the regulator, acting as a principal, aiming at pushing the crowd to some target distribution. Such a problem may be formulated for instance as a model for the design of smart cities. Then, one may use the same techniques as for the Principal-Agent problem in order to convert this problem to a more standard optimal transport problem.

In a situation where a Principal faces many interacting agents, distributed control may serve as an important tool to preserve the aggregate production of the agents, while distributing differently the contributions amongst agents.

The above approach needs now to be extended in order to accommodate more realistic situations. Let us list the following important extensions:

- The case of noisy observation of the output leads to control problems under partial observation for both types of agents, which are significantly more involved as they lead to infinite dimensional control problems after the filtering stage;

- Another important extension is the account for the so-called adverse selection: the principal has no access to the optimization criterion of the agent, instead she only has a prior on the distribution; in the economic literature, this is addressed in one period static models by allowing the prinicipal to offer a menu of incentivizing contracts which are chosen so that each agent chooses the one which is designed for him (incentive compatibility constraint).

Our research program on networks with interacting agents concerns various types of networks: electronic networks, biological networks, social networks, etc. The numerous mathematical tools necessary to analyse them depend on the network type and the analysis objectives. They include propagation of chaos theory, queing process theory, large deviation theory, ergodic theory, population dynamics, partial differential equation analysis, in order to respectively determine mean-field limits, congestion rates or spike train distributions, failure probabilities, equilibrium measures, evolution dynamics, macroscopic regimes, etc.

For example, recently proposed neuron models consist in considering different populations of neurons and setting up stochastic time evolutions of the membrane potentials depending on the population. When the number of populations is fixed, interaction intensities between individuals in different populations have similar orders of magnitude, and the total number of neurons tends to infinity, mean-field limits have been identified and fluctuation theorems have been proven.

However, to the best of our knowledge, no theoretical analysis is available on interconnected networks of networks with different populations of interacting individuals which naturally arise in biology and in economics.

We aim to study the effects of interconnections between sub-networks resulting from individual and local connections. Of course, the problem needs to be posed in terms of the geometry of the big network and of the scales between connectivity intensities and network sizes.

A related research concerns stochastic, continuous state and time opinion models where each agent's opinion locally interacts with other agents' opinions in the systemic. Due to some exogenous randomness, the interaction tends to create clusters of common opinion. By using linear stability analysis of the associated nonlinear Fokker-Planck equation that governs the empirical density of opinions in the limit of infinitely many agents, we can estimate the number of clusters, the time to cluster formation, the critical strength of randomness so as to have cluster formation, the cluster dynamics after their formation, the width and the effective diffusivity of the clusters.

Another type of network systems we are interested in derives from financial systemic risk modeling. We consider evolving systems with a large number of inter-connected components, each of which can

be in a normal state or in a failed state. These components also have mean field interactions and a cooperative behavior. We will also include diversity as well as other more complex interactions such as hierarchical ones. In such an inter-connected system, individual components can be operating closer to their margin of failure, as they can benefit from the stability of the rest of the system. This, however, reduces the overall margin of uncertainty, that is, increases the systemic risk: our research thus addresses QMU (Quantification of Margins of Uncertainty) problems.

We aim to study the probability of overall failure of the system, that is, its systemic risk. We therefore have to model the intrinsic stability of each component, the strength of external random perturbations to the system, and the degree of inter-connectedness or cooperation between the components.

Our target applications are the following ones:

- Engineering systems with a large number of interacting parts: Components can fail but the system fails only when a large number of components fail simultaneously.

- Power distribution systems: Individual components of the system are calibrated to withstand fluctuations in demand by sharing loads, but sharing also increases the probability of an overall failure.

- Banking systems: Banks cooperate and by spreading the risk of credit shocks between them can operate with less restrictive individual risk policies. However, this increases the risk that they may all fail, that is, the systemic risk.

One of our objectives is to explain that, in some circumstances, one simultaneously observes individual risk decreases and systemic risk increase.

## 4   Application domains

Our short and mid-term potential industrial impact concerns e.g. energy market regulations, financial market regulations, power distribution companies, nuclear plant maintenance. It also concerns all the industrial sectors where massive stochastic simulations at nano scales are becoming unavoidable and certified results are necessary.

We also plan to have impact in cell biology, macro-economy, and applied mathematics at the crossroads of stochastic integration theory, optimization and control, PDE analysis, and stochastic numerical analysis.

## 5   Highlights of the year

The team organised a two-day seminar in October where all the members presented their current work and recent results.

## 6   New software and platforms

### 6.1   New platforms

Our most topical activity concerns the PYCATSHOO toolbox developed by EDF which allows the modeling of dynamical hybrid systems such as nuclear power plants or dams. Hybrid systems mix two kinds of behaviour. First, the discrete and stochastic behaviour which is in general due to failures and repairs of the system's constituents. Second, the continuous and deterministic physical phenomena which evolve inside the system.

PyCATSHOO is based on the theoretical framework of Piecewise Deterministic Markov Processes (PDMPs). It implements this framework thanks to distributed hybrid stochastic automata and object-oriented modeling. It is written in C++. Both Python and C++ APIs are available. These APIs can be used either to model specific systems or for generic modelling i.e. for the creation of libraries of component models. Within PyCATSHOO special methods can be developed.

J. Garnier is contributing, and will continue to contribute, to this toolbox within joint Cifre programs with EdF. The PhD theses are aimed to add new functionalities to the platform. For instance, an importance sampling with cross entropy method

# 7 New results

## 7.1 Optimal stopping of dynamic stochastic interacting systems

This project is developed by Nizar Touzi, Mehdi Talbi (third year PhD student) and Jianfeng Zhang (University of Southerml California).

The problem of multiple optimal stopping In the context of a finite population has been considered by Kobylanski, Quenez and Rouy-Mironescu (Annals of Applied Probability, 2011). The value function is naturally expressed as a standard single stopping problem of one of the continuing particles, so that the value function can be characterized by a recursive cascade of single optimal stopping problems, and optimal stopping times are obtained as first hitting times of the corresponding value functions to their stopping values. Our objective is to study the mean field limit of these results under a symmetry assumption so as to guarantee the anonymity of the particules. The limiting multiple optimal stopping problem is an optimal stopping problem of a stochastic differential equation (SDE) with mean field interaction à la McKean-Vlasov, with optimization criterion depending on the law of the stopped state of the continuum of particules.

a) Our analysis of this problem builds on an appropriate version of the dynamic programming principle on the space of marginal laws of the state process of particules, and a general Itô's formula for flows of marginal distributions of càdlàg semimartingales. Our first paper derives the corresponding dynamic programming equation on the Wasserstein space as the infinitesimal version of the dynamic programming principle, and proves a verification argument in this context: any classical solution of this dynamic programming equation, with appropriate integrability, coincides with the value function of the mean field optimal stopping problem. We also obtain necessary and sufficient conditions for the optimality of some mean field stopping rule.

b) Unfortunately, the regularity conditions needed in the previous work are rarely satisfied, even in the standard finite dimensional optimal stopping problem. Our second work in this project contains a viscosity solution approach for the dynamic programming equation. As the Wasserstein space fails to be locally compact, the standard Crandall-Lions theory does not apply in the present context. Motivated by our previous experience for path)dependent PDEs, we introduce a notion of viscosity solutions which allow for existence and uniqueness results for the dynamic programming equation of the mean field stopping problem. Our notion of solutions shares the good properties of Crandall-Lions theory, namely the consistency with classical solutions and the stability by relaxed semilimits of semisolutions.

c) Our third work in this project proves a result of propagation of chaos in the context of mean field optimal stopping. Namely, we show that the $N-$particles multiple optimal stopping value function converges to the mean field optimal stopping value function. The proof is an adaptation to the present context of the Barles-Souganidis convergence of monotonic schemes, and requires the use of a notion of a variation of the finite dimensional viscosity solutions by considering the larger set of those test functions which are tangent in mean.

d) Finally, all previous results are adapted to the context of mean field optimal control by Mehdi Mrad. The notion of viscosity solution approach. The main difficulty is related to the control of the diffusion coefficient. Then, the collection of distributions induced by the controlled process is a non dominated set of singular measures. For this reason, the analysis of the mean field control problem uses nonlinear expectations, defined as supreme over the set of all possible distributions, in replacement for the standard expectation operator. The main results of this work consists in a complete wellposedness theory in the sense of an appropriate notion of viscosity solutions, and a result of propagation of chaos proved by a convenient adaptation of the Barles-Souganidis convergence of monotonic schemes.

## 7.2 Modeling systemic risk by stochastic differential games

This project is developed by Fabrice Djete, Nizar Touzi, and Leila Bassou (second year PhD student), Gaoyue Guo (CentraleSupelec).

The analysis of systemic risk of the financial sector is most important, especially since the last financial crisis which highlighted the complex effects induced the multiple connexions on the network of financial actors.

The existing literature on this topic is essentially empirical and suffers from a serious lack of theoretical foundations. Our contributions constitute the first step for understanding the interdependence structures constructed by the economic actors, and help to analyze their nature and to examine their consequences during the periods of stress. Our starting point is to formulate an interacting version of the standard portfolio optimization problem in finance so as to develop the analogue the the so called Capital Asset Pricing Model in the context of interaction by cross-holdings.

a) In the finite population cross holding problem, each agent is allowed to choose her investments in the others, while undergoing the investment strategies of the other, including on themselves. Our objective is to characterize the Nash equilibria of this game, if exists:

-Each actor chooses the optimal investment strategy in the others, given the strategies of all of the competitors, -A Nash equilibrium is a situation of interdependence such that no actor optimally chooses to deviate from it, given that the others' strategies are fixed to this equilibrium value.

We provide a general characterization of such Nash equilibria in the context of a special stochastic differential game. The actors preferences are defined by exponential utility functions so as to addressing the problem by means of backward stochastic differential equations, a most useful tool for the underlying path dependent dependent stochastic control problems.

b) By introducing an appropriate symmetry in the system, we obtain a scaling limit which represents the macroscopic description as a mean field stochastic differential equation à la McKean-Vlasov, of new type: the interaction term appears as an expectation of some stochastic integral with respect to a copy of the equilibrium system. When there is no common noise, we obtain a Nash equilibrium with a quasi-explicit characterization. At this equilibrium, the representative agent possesses a unit proportion of each of the competitors whose drift coefficient is above some constant characterized as the unique solution of an integral equation. We also show that this results induces an approximate equilibrium for the finite population optimal cross holding problem. Namely, taking the the optimal mean field policy as a feedback policy in the finite population, we prove that the induced criterion of each individual agent can not be improved up to some small tolerance threshold which vanishes to zero as the size of the population increases to infinity. As a by-product of this result, we obtain a result of propagation of chaos for the optimal cross holding problem.

c) In an ongoing work, we extend the last model so as to model the systemic risk of a system of actors interacting by mutual holding. To do this, we introduce the bankruptcy by forcing absorption at zero of the mean field stochastic differential equation. This leads to a yet new form of mean field game with the additional technical difficulties due to the absorption at zero. More importantly, we prove that the equilibrium distribution decomposes into a Dirac class at zero and a sub-probability measure which is absolutely continuous with respect to the Lebesgue measure. The evolution of the mass at zero represents the contagion effect of the systemic risk of the system. When the coefficients of the interacting SDE are deterministic, we obtain an autonomous equation which fully characterizes the mass at zero.

## 7.3 Mean field game of mutual holding

In order to study a financial model where each agent can a share of the assets of others, Fabride Djete and Nizar Touzi introduce a mean field model for optimal holding of a representative agent of her peers as a natural expected scaling limit from the corresponding $N-$agent model. The induced mean field dynamics appear naturally in a form which is not covered by standard McKean-Vlasov stochastic differential equations. We study the corresponding mean field game of mutual holding in the absence of common noise. Our main result provides existence of an explicit equilibrium of this mean field game, defined by a bang-bang control consisting in holding those competitors with positive drift coefficient of their dynamic value. Our analysis requires to prove an existence result for our new class of mean field SDE with the additional feature that the diffusion coefficient is irregular. The paper is available online at https://arxiv.org/pdf/2104.03884.pdf

## 7.4   Large population games with interactions through controls and common noise

In the presence of a common noise, Mao Fabrice Djete studies the convergence problems in mean field game (MFG) and mean field control (MFC) problem where the cost function and the state dynamics depend upon the joint conditional distribution of the controlled state and the control process. In the first part, he considers the MFG setting. He starts by recalling the notions of measure-valued MFG equilibria and of approximate closed-loop Nash equilibria associated to the corresponding $N$–player game. Then, he shows that all convergent sequences of approximate *closed–loop* Nash equilibria, when $N$ goes to $\infty$, converge to measure-valued MFG equilibria. And conversely, any measure-valued MFG equilibrium is the limit of a sequence of approximate closed-loop Nash equilibria. In other words, measure-valued MFG equilibria are the accumulation points of the approximate closed-loop Nash equilibria. Previous work has shown that measure-valued MFG equilibria are the accumulation points of the approximate open-loop Nash equilibria. Therefore, Mao Fabrice Djete obtains that the limits of approximate closed-loop Nash equilibria and approximate open-loop Nash equilibria are the same. In the second part, he deals with the MFC setting. After recalling the closed-loop and open-loop formulations of the MFC problem, he proves that they are equivalent. he also provides some convergence results related to approximate closed-loop Pareto equilibria. The paper is available online at https://arxiv.org/pdf/2108.02992.pdf.

## 7.5   Sensitivity of diffusion models to their noise Hurst parameter

In collaboration with Alexandre Richard (Ecole Centrale - Supelec, Saclay, France) Denis Talay developed a sensitivity analysis w.r.t. the long-range/memory noise parameter for probability distributions of functionals of solutions to stochastic differential equations.

This is an important stochastic modeling issue in many applications since Markov models may sometimes be seen as questionable idealizations of the reality. Empirical studies actually tend to show memory effects in biological, financial, physical data. Such empirical results justify to consider non-Markov models driven by noises with long-range memory such as fractional Brownian motions rather than by Lévy noises. But choosing a noise with long-range memory does not close the modeling problem since the parametric estimation of the model may be difficult and crude.

Therefore, one often needs to balance tractable Markov models against more realistic but complex non-Markov models. A natural question then arises: Is it really worth introducing complex models?

A. Richard and D. Talay have thus considered solutions $\{X_t^H\}$ to stochastic differential equations driven by fractional Brownian motions. They have developed sensitivity analyses when the Hurst parameter $H$ of the noise tends to the critical Brownian parameter $H = \frac{1}{2}$ from above or from below. Two classes of sensitivities have been studied. On the one hand, the probability distributions of smooth functionals in the Malliavin calculus sense. On the other hand, , the probability distributions of irregular functionals, namely, Laplace transforms of first passage times of $X^H$ at given thresholds.

Our results show that the Markov Brownian model is a good proxy model as long as the Hurst parameter remains close to $\frac{1}{2}$.

## 7.6   Entropic Fictitious Play

This is a project conducted by Songbo Wang (first-year PhD, CMAP, Ecole polytechnique) and one of his supervisors, Zhenjie Ren (CEREMADE, Université Paris-Dauphine). We study a general mean-field optimization problem with entropic regularizer. It is known that practical problems, including notably the convergence of deep neural networks, can be modelized by the mean-field approach. Residual neural networks have been studied in the previous works of Zhenjie Ren et al and the convergence of the gradient descent, arguably the most common training method in practice, is obtained. Motivated by the classical fictitious play from game theory, we construct a novel dynamics whose convergence to the optimal solution is shown. Additionally, the rate of convergence of this dynamics is obtained in the convex case, which is absent in the previous results of gradient flow.

## 7.7   Seismic probabilistic risk assessment

The key elements of seismic probabilistic risk assessment studies are the fragility curves which express the probabilities of failure of structures conditional to a seismic intensity measure. A multitude of procedures

is currently available to estimate these curves. For modeling-based approaches which may involve complex and expensive numerical models, the main challenge is to optimize the calls to the numerical codes to reduce the estimation costs. Adaptive techniques can be used for this purpose, but in doing so, taking into account the uncertainties of the estimates (via confidence intervals or ellipsoids related to the size of the samples used) is an arduous task because the samples are no longer independent and possibly not identically distributed. The main contribution of this work is to deal with this question in a mathematical and rigorous way. To this end, C. Gauchy, C. Feau, J. Garnier propose and implement an active learning methodology based on adaptive importance sampling for parametric estimations of fragility curves. They prove some theoretical properties (consistency and asymptotic normality) for the estimator of interest. Moreover, they give a convergence criterion in order to use asymptotic confidence ellipsoids. Finally, the performances of the methodology are evaluated on analytical and industrial test cases of increasing complexity.

## 7.8 Surrogate modeling of a complex numerical code

B. Kerleguer considers the surrogate modeling of a complex numerical code in a multifidelity framework when the code output is a time series. Using an experimental design of the low-and high-fidelity code levels, an original Gaussian process regression method is proposed. The code output is expanded on a basis built from the experimental design. The first coefficients of the expansion of the code output are processed by a co-kriging approach. The last coefficients are collectively processed by a kriging approach with covariance tensorization. The resulting surrogate model taking into account the uncertainty in the basis construction is shown to have better performance in terms of prediction errors and uncertainty quantification than standard dimension reduction techniques.

## 7.9 Communication networks and their algorithms

One of the main current topics of research of Carl Graham is the modeling, analysis, simulation, and evaluation of communication networks and their algorithms. Most of these algorithms must function in real time, in a distributed fashion, and using sparse information transfers and backlog. Of particular interest are load balancing algorithms, which aim to provide better utilization of the network resources and hence better quality of service for clients. Many such algorithms strive to avoid the starving of some servers and the build-up of queues at others by routing the clients so as to have them well spread out throughout the system. The main focus of the present work on these networks is prefect simulation in equilibrium. This in particular will allow to estimate by Monte Carlo methods a number of quality of service (QoS) indicators directly in the stationary regime.

## 7.10 A stochastic numerical method for the parabolic-parabolic Keller-Segel system

The parabolic-parabolic Keller-Segel model is a set of equations that model the process of cell movement. It takes into account the evolution of different chemical components that can aid, hinder or change the direction of movement, a process called chemotaxis.

In collaboration with Radu Maftei (who is a past Inria post-doc student) Milica Tomasevic (CMAP, Ecole Polytechnique) and Denis Talay analyse the numerical performances of a stochastic particle numerical method for the parabolic-parabolic Keller-Segel model. They also propose and test various algorithmic improvements to the method in order to substantially decrease its execution time without altering its global accuracy.

## 7.11 ICI: an epidemy propagation simulator

In 2020, ASCII team started the ICI (INRIA-Collaboration-IGN) project. This project is aimed to simulate a dynamic spreading of an epidemy at the individual scale, in a very precise geographic environment.

In addition to the permanent members of the team, Nicolas Gilet (INRIA engineer) and Maxime Colomb (INRIA-IGN engineer), this project initiated by Denis Talay in July 2020 gathered in 2021 the following contributors:

- Aline Carneiro Viana (Inria Project-team TriBE)

- Laura Grigori (Inria Project-team Alpines)

- Julien Perret (IGN)

- Razvan Stanica (Inria Project-team Agora)

- Milica Tomasevic (CMAP, Ecole Polytechnique)

Infection between inhabitants is calculated with the density of person that they stay with during the day, their epidemiologic status and probability laws. Statistical studies on the simulation results should allow one to better understand the propagation of an epidemy and to compare the performances of various public stop-and-go strategies to control the person-to-person contamination.

This year, Maxime Colomb (INRIA-IGN engineer) and Nicolas GILET (INRIA engineer) have developed a prototype of the modeling with the support of the ASCII permanent members and the ICI contributors. The modeling is based on the coupling of two different models: on the one hand, the modeling of the urban geographical area where the population live and move; on the other hand, the modeling of random choices of the daily travels and the contaminations due to the interactions between individuals. The simulation is for now applied on a sub-part of Paris's fifth arrondissement (Jussieu/St-Victor) and aimed to run on single arrondissement or small cities.

The geographic model is built from multiple geographic sources (IGN, INSEE, OpenStreetMap, Local authority open data portals, etc.). A three-layered synthetic population is generated in order to represent housing, populated by households, composed by individuals. The multiple characteristics added allows to represent conditions of living and inner household interactions of the population. Shops and activities are generated by matching multi-sourced data, allowing to enrich information about each amenity (opening hour, open during lockdown, etc.). We simulate the socio-professional structures and daily trips of the population by taking into account probability laws related to the urban space (probability of going out, going to work, shopping, etc.) and to social characteristics (age, job, etc.). Currently, the modeling is based on intuitive and simple laws of trips according to individuals groups (pupils, students, working people, retirees). The calibration of these probability laws is being improved by using data provided by precise surveys and mobile operators.

In addition, the person-to-person contamination has been modeled between individuals located in the same space at the same time using transmission probability laws specific to each individual, parameterized by the distance between a healthy and contamined individual, as well as by the contact duration. Since the model is stochastic, in order to obtain accurate and robust statistics on the evolution of the epidemic, we must be able to simulate a large number of independent socio-professional structures within a given urban area, and then, for each population, a large number of realizations of daily trips and contaminations.

Therefore, to carry out a very large number of simulations covering all parameters of the model, the model requires very high performance computing. The code is written in Julia language and is currently parallelized using MPI tool. At this time, the model is launched on the internal cluster of INRIA Saclay called Margaret (200 CPU cores corresponding to 10 CPU nodes) which allows to check the code for a few different epidemiological parameters. We have also obtained the support of AMD to launched our model on a cluster, equipped with AMD EPYC™ processors and AMD Instinct™ Accelerators, into the national GRID'5000/SILECS infrastructure (https://www.genci.fr/en/node/1122). Moreover, in September 2021, the ICI project has obtained 6 millions CPU hours from DARI/GENCI which can be used on the CEA cluster called Irene-Rome (up to 300 000 CPU cores) in order to launch simulations for a large panel of epidemiological parameters. These hours can be used until October 2022.

Finally, Maxime and Nicolas have developed a website that describes the ICI project and the characteristics of the ICI model (https://ici.gitlabpages.inria.fr/website/). They have also begin to develop an user interface from which it is possible to study the effect of health policies on the epidemic propagation by displaying the main epidemic indicators computed by the model (https://ici.gitlabpages.inria.fr/website/index_app_v2.html).

Our next step is to calibrate the model with epidemiologic data and compare the predictive capacities of ICI with simpler models (SIR/SEIR).

Concerning the data part, the first step will be to improve individual travels. Multiple Markov chains are constructed and calibrated for various geographical and socio-demographic profiles with the precise values of a global survey. Micro-spatialization of travel objective must be realized using mobile phone data and the list of available places, weighted by their capacity to receive public. The synthetic population generation have to be improved in order to give occupation to each individuals and to get more close to existing statistics. Those imporvements are made jointly with the redaction of a scientific article.

Finally, Maxime and Nicolas will continue to develop the user interface by including the back-end of the application. More precisely, they will put the interface on an INRIA web server and they will define an automatic pipeline between the interface and the server in order to display all of the simulations to the user. An undergraduate internship has been defined for a student in computer science in order to help Maxime and Nicolas to put the application on the web server.

# 8 Bilateral contracts and grants with industry

## 8.1 Bilateral contracts with industry

Several INRIA teams, including ASCII, are involved in the CIROQUO Research & Industry Consortium – Consortium Industrie Recherche pour l'Optimisation et la QUantification d'incertitude pour les données Onéreuses – (Industry Research Consortium for the Optimization and QUantification of Uncertainty for Onerous Data). Josselin Garnier is the INRIA Saclay representative on the steering committee.

The principle of the CIROQUO Research & Industry Consortium is to bring together academic and technological research partners to solve problems related to the exploitation of numerical simulators, such as code transposition (how to go from small to large scale when only small-scale simulations are possible), taking into account the uncertainties that affect the result of the simulation, validation and calibration (how to validate and calibrate a computer code from collected experimental data). This project is the result of a simple observation: industries using computer codes are often confronted with similar problems during the exploitation of these codes, even if their fields of application are very varied. Indeed, the increase in the availability of computing cores is counterbalanced by the growing complexity of the simulations, whose computational times are usually of the order of an hour or a day. In practice, this limits the number of simulations. This is why the development of mathematical methods to make the best use of simulators and the data they produce is a source of progress. The experience acquired over the last thirteen years in the DICE and ReDICE projects and the OQUAIDO Chair shows that the formalization of real industrial problems often gives rise to first-rate theoretical problems that can feed scientific and technical advances. The creation of the CIROQUO Research & Industry Consortium, led by the Ecole Centrale de Lyon and co-animated with the IFPEN, follows these observations and responds to a desire for collaboration between technological research partners and academics in order to meet the challenges of exploiting large computing codes.

**Scientific approach.** The limitation of the number of calls to simulators implies that some information – even the most basic information such as the mean value, the influence of a variable or the minimum value of a criterion – cannot be obtained directly by the usual methods. The international scientific community, structured around computer experiments and uncertainty quantification, took up this problem more than twenty years ago, but a large number of problems remain open. On the academic level, this is a dynamic field which is notably the subject of the French CNRS Research Group MascotNum since 2006 and renewed in 2020.

**Composition.** The CIROQUO Research & Industry Consortium aims to bring together a limited number of participants in order to make joint progress on test cases from the industrial world and on the upstream research that their treatment requires. The overall approach that the CIROQUO Research & Industry Consortium will focus on is metamodeling and related areas such as experiment planning, optimization, inversion and calibration. IRSN, STORENGY, CEA, IFPEN, BRGM are the Technological Research Partners. Mines Saint-Etienne, Centrale Lyon, CNRS, UCA, UPS, UT3 and Inria are the Academic Partners of the consortium.

**Scientific objectives.** On the practical level, the expected impacts of the project are a concretization of the progress of numerical simulation by a better use of computational time, which allows the determination of better solutions and associated uncertainties. On the theoretical level, this project

will allow to create an emulation around the major scientific locks of the discipline such as code transposition/calibration/validation, modeling for complex environments, or stochastic codes. In each of these scientific axes, a particular attention will be paid to the large dimension. Real problems sometimes involve several tens or hundreds of inputs. Methodological advances will be proposed to take into account this additional difficulty. The work expected from the consortium differs from the dominant research in machine learning by specificities linked to the exploration of expensive numerical simulations. However, it seems important to build bridges between the many recent developments in machine learning and the field of numerical simulation.

**Philosophy.** The CIROQUO Research & Industry Consortium is a scientific collaboration project aiming to mobilize means to achieve methodological advances. The project promotes cross-fertilization between partners coming from different backgrounds but confronted with problems related to a common methodology. It has three objectives: - the development of exchanges between technological research partners and academic partners on issues, practices and solutions through periodic scientific meetings and collaborative work, particularly through the co-supervision of students;
- the contribution of common scientific skills thanks to regular training in mathematics and computer science;
- the recognition of the Consortium at the highest level thanks to publications in international journals and the diffusion of free reference software.

## 8.2   Collaboration with EdF on industrial risks

This collaboration whose Josselin Garnier is the Ascii leader, has been underway for several years. It concerns the assessment of the reliability of hydraulic and nuclear power plants built and operated by EDF (Electricite De France). The failure of a power plant is associated with major consequences (flood, dam failure, or core meltdown), for regulatory and safety reasons EDF must ensure that the probability of failure of a power plant is sufficiently low.

The failure of such systems occurs when physical variables (temperature, pressure, water level) exceed a certain critical threshold. Typically, these variables enter this critical region only when several components of the system are deteriorated. Therefore, in order to estimate the probability of system failure, it is necessary to model jointly the behavior of the components and of the physical variables. For this purpose, a model based on a Deterministic Markovian Piecewise Process (DMPP) is used. The platform called PYCATSHOO has been developed by EDF to simulate this type of process. This platform allows to estimate the probability of failure of the system by Monte Carlo simulation as long as the probability of failure is not too low. When the probability becomes too low, the classical Monte Carlo estimation method, which requires a lot of simulations to estimate the probabilities of rare events, is much too slow to execute in our context. It is necessary to use methods using fewer simulations to estimate the probability of system failure: variance reduction methods. Among the variance reduction methods are "importance sampling" and "splitting" methods, but these methods present difficulties when used with PDMPs.

Work has been undertaken on the subject, leading to the defense of a first CIFRE thesis (Thomas Galtier, thesis defended in 2019) and the preparation of a new CIFRE thesis (Guillaume Chennetier, from 2021). Theoretical articles have been written and submitted to journals. New theoretical works on sensitivity analysis in rare event regimes are the subject of the new thesis. The integration of the methods in the PYCATSHOO platform is progressively done.

# 9   Partnerships and cooperations

## 9.1   International initiatives

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

**CIRCUS**

**Title:**  Columbia Inria Research on Collaborative Ultracritical Systems

**Duration:**  2020 ->

**Coordinator:** Philip Protter (pep2117@columbia.edu)

**Partners:**

- Columbia University

**Inria contact:** Denis Talay

**Summary:**

# 10  Dissemination

## 10.1  Member of the editorial boards

J. Garnier is a member of the editorial boards of the journals Asymptotic Analysis, Discrete and Continuous Dynamical Systems – Series S, ESAIM P&S, Forum Mathematicum, SIAM Journal on Applied Mathematics, and SIAM/ASA Journal on Uncertainty Quantification (JUQ).

D. Talay serves as an Area Editor of *Stochastic Processes and their Applications*, and as an Associate Editor of *Journal of the European Mathematical Society, Probability, Uncertainty and Quantitative Risk, ESAIM Probability and Statistics, Stochastics and Dynamics, Journal of Scientific Computing, Monte Carlo Methods and Applications, SIAM Journal on Scientific Computing, Communications in Applied Mathematics and Computational Science, Éditions de l'École Polytechnique*. He also served as Co-editor in chief of *MathematicS in Action*.

N. Touzi serves as an Area Editor of *Stochastic Processes and their Applications*, and as an Associate Editor of *Advances in Calculus of Variations, Stochastics: an International Journal of Probability and Stochastic Processes, Journal of Optimization Theory and Applications, Mathematical Control and Related Fields, Tunisian Journal of Mathematics, Springer Briefs in Mathematical Finance*. He also is Co-Editor of *Paris-Princeton Lectures in Mathematical Finance*.

## 10.2  Invited talks

C. Graham gave an online LAGA Probability seminar at Institut Galilée, Université Paris 13.

J. Garnier gave lectures at Oberwolfach (Germany) in March, at the RESIM21 Conference in May, at the SIAM Annual meeting in July, at the SMAI-MAS 21 Conference in August, at the Resonances, Inverse Problems and Seismic Waves Conference (Reims, France) in November. He gave an online seminar at Stanford university in October.

D. Talay gave an ICMS seminar in January, an invited lecture at the Conference in the honor of Gilles Pagès in May, a seminar at the Department of Mathematics and Institute of Mathematical Sciences at the Chinese University of Hong Kong in September.

N. Touzi gave lectures at the Conference Beyond the Boundaries, (Leeds) in May, at the Chebyshev 200 VI-th International Conference on Stochastic Methods (Moscow) in May, at the Summer School on Distributed Control: Decentralization and incentives" (CIRM Marseille) in June, at the Advances in Stochastic Analysis for Handling Risks in Finance and Insurance Conference (CIRM Marseille) in September, at the Journées Ateliers FiME in September, at the online Workshop on Mean-field reinforcement learning in October, at the First Florence-Paris Workshop on Mathematical Finance in October, and at the CIMPA School (Ben Guerir, Marocco) in November, at the 5th Conference on Mathematical Science and Applications (Kaust online) in November, at the Conference celebrating '30 ans de l'IUF' (Le Mans Université) in November, and at the IMSA Workshop (Chicago) in December.

## 10.3  Leadership within the scientific community

D. Talay is Vice-President of the Natixis Foundation which promotes academic research on financial risks.

N. Touzi is Scientific Director of the Risk Foundation, Louis Bachelier Institute.

D. Talay served in a recruitment committee for a Professor in Statistics position at the university of Lille.

## 10.4  Teaching - Supervision - Juries

### 10.4.1  Teaching

Josselin Garnier is a professor at the Ecole Polytechnique and he has a full teaching service. He also teaches the class "Inverse problems and imaging" at the Master Mathématiques, Vision, Apprentissage (M2 MVA).

D. Talay teaches the master course *Equations différentielles stochastiques de McKean–Vlasov et limites champ moyen de systèmes de particules stochastiques en interaction*, 24h, M2 Probabilité et Applications, LPSM, Sorbonne Université, France.

Nizar Touzi is a professor at the Ecole Polytechnique and he has a full teaching service.

### 10.4.2  Juries

J. Garnier was a member of the following juries: Charles-Edouard Brehier' HDR (Univ. Claude Bernard - Lyon 1) in June, Julien Reygner's HDR (Univ. Paris Est) in September, Ayao Ehara's PhD (Univ. College London) in October, Kilian Baudin's PhD (Univ. Dijon) in December.

C. Graham was a member of the following juries: Xavier Erny's PhD (Université d'Evry Val d'Essonne) in June, Josué Tchouanti Fotso's PhD (Institut Polytechnique de Paris) in September.

N. Touzi was a member of the following juries: Thibaut Mastrolia's HdR (Ecole Polytechnique) in January, Maxime Grangereau's PhD (Ecole Polytechnique) in March, Soufiane Mouchtabih's PhD (Université de Toulon) in March, Bastien Baldacci's PhD (Ecole Polytechnique) in May, Pierre Lavigne's PhD (Ecole Polytechnique) in December.

### 10.4.3  Supervision

J. Garnier is supervising the following Ph.D. students: Naoufal Acharki, Guillaume Chennetier, Alexis Cousin, Clement Gauchy, Corentin Houpert, Baptiste Kerleguer. Alexis Cousin got his Ph.D. degree (Ecole Polytechnique) in October.

C. Graham and S. Méléard's (Ecole Poytechnique) PhD student Josué Tchouanti Fotso obtained his PhD degree (Institut Polytechnique de Paris) in September.

N. Touzi is supervising the following Ph.D. students: Leila Bassou, Assil Fadle, Mehdi Talbi, Songbo Wang. Heythem Farhat obtained his PhD degree (Ecole Polytechnique) in February.

# 11  Scientific production

## 11.1  Publications of the year

**International journals**

[1]  L. Borcea, J. Garnier and K. Sølna. 'Onset of energy equipartition among surface and body waves'. In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 477.2246 (Feb. 2021), p. 20200775. DOI: 10.1098/rspa.2020.0775. URL: https://hal.archives-ouvertes.fr/hal-03147615.

[2]  J. Garnier. 'Passive Communication with Ambient Noise'. In: *SIAM Journal on Applied Mathematics* 81.3 (Jan. 2021), pp. 814–833. DOI: 10.1137/20M1366848. URL: https://hal.archives-ouvertes.fr/hal-03510634.

[3]  J. Garnier. 'Wave Propagation in Periodic and Random Time-Dependent Media'. In: *Multiscale Modeling and Simulation: A SIAM Interdisciplinary Journal* 19.3 (Jan. 2021), pp. 1190–1211. DOI: 10.1137/20M1377734. URL: https://hal.archives-ouvertes.fr/hal-03510670.

[4]  J. Garnier, P. Azam, A. Fusaro, Q. Fontaine, A. Bramati, A. Picozzi, R. Kaiser, Q. Glorieux and T. Bienaimé. 'Dissipation-enhanced collapse singularity of a nonlocal fluid of light in a hot atomic vapor'. In: *Physical Review A* 104.1 (July 2021). DOI: 10.1103/PhysRevA.104.013515. URL: https://hal.archives-ouvertes.fr/hal-03510666.

[5] J. Garnier, K. Baudin, A. Fusaro, N. Berti, K. Krupa, I. Carusotto, S. Rica, G. Millot and A. Picozzi. 'Energy and wave-action flows underlying Rayleigh-Jeans thermalization of optical waves propagating in a multimode fiber (a)'. In: *EPL - Europhysics Letters* 134.1 (17th May 2021), p. 14001. DOI: 10.1209/0295-5075/134/14001. URL: https://hal.archives-ouvertes.fr/hal-03510638.

[6] J. Garnier, K. Baudin, A. Fusaro and A. Picozzi. 'Coherent Soliton States Hidden in Phase Space and Stabilized by Gravitational Incoherent Structures'. In: *Physical Review Letters* 127.1 (June 2021). DOI: 10.1103/PhysRevLett.127.014101. URL: https://hal.archives-ouvertes.fr/hal-03510659.

[7] J. Garnier, K. Baudin, A. Fusaro and A. Picozzi. 'Incoherent localized structures and hidden coherent solitons from the gravitational instability of the Schrödinger-Poisson equation'. In: *Physical Review E* 104.5 (Nov. 2021). DOI: 10.1103/PhysRevE.104.054205. URL: https://hal.archives-ouvertes.fr/hal-03510675.

[8] J. Garnier and L. Borcea. 'Imaging in Random Media by Two-Point Coherent Interferometry'. In: *SIAM Journal on Imaging Sciences* 14.4 (Jan. 2021), pp. 1635–1668. DOI: 10.1137/21M142068X. URL: https://hal.archives-ouvertes.fr/hal-03510682.

[9] J. Garnier, H. Chraibi, A. Dutfoy and T. Galtier. 'Optimal potential functions for the interacting particle system method'. In: *Monte Carlo Methods and Applications* 27.2 (1st June 2021), pp. 137–152. DOI: 10.1515/mcma-2021-2086. URL: https://hal.archives-ouvertes.fr/hal-03510653.

[10] J. Garnier, E. Gay, L. Bonnet, C. Peyret and É. Savin. 'Coherent interferometric imaging in a random flow'. In: *Journal of Sound and Vibration* 494 (Mar. 2021), p. 115852. DOI: 10.1016/j.jsv.2020.115852. URL: https://hal.archives-ouvertes.fr/hal-03510577.

[11] J. Garnier and K. Sølna. 'Enhanced Backscattering of a partially coherent field from an anisotropic random lossy medium'. In: *Discrete and Continuous Dynamical Systems - Series B* 26.2 (2021), pp. 1171–1195. DOI: 10.3934/dcdsb.2020158. URL: https://hal.archives-ouvertes.fr/hal-03510623.

[12] C. Graham. 'Regenerative properties of the linear Hawkes process with unbounded memory'. In: *Annals of Applied Probability* 31.6 (2021), pp. 2844–2863. DOI: 10.1214/21-AAP1664. URL: https://hal-polytechnique.archives-ouvertes.fr/hal-02139998.

**Reports & preprints**

[13] C. Gauchy, C. Feau and J. Garnier. *Importance sampling based active learning for parametric seismic fragility curve estimation.* 26th Jan. 2022. URL: https://hal.inria.fr/hal-03544470.

[14] A. Richard and D. Talay. *Lipschitz continuity in the Hurst parameter of functionals of stochastic differential equations driven by a fractional Brownian motion.* 2021. URL: https://hal.inria.fr/hal-01323288.