



**Laboratoire de Mathématiques et Informatique pour la Complexité et les Systèmes
MICS**

Présente

L'AVIS DE SOUTENANCE

De Monsieur Antonin Della Noce

Laboratoire MICS, CentraleSupélec, Université Paris Saclay, soutiendra publiquement
ses travaux de thèse de doctorat intitulés :

**"Symmetric and heterogeneous population models: approximations for simulation
and statistical inference "**

Le lundi 29 mars 2021 à 14h00

À l'école CentraleSupélec, dans la **salle sc.046 (Peugeot)** - Bâtiment Bouygues en
présentiel et en distanciel

*Si vous souhaitez assister à la soutenance en distanciel, veuillez contacter en avance l'assistante du laboratoire pour obtenir
le lien (fabienne.brosse@gmail.com)*

Membres du jury :

Amandine Véber, Directrice de recherche, CNRS, Université de Paris, UMR MAP5 (Rapporteuse et
examinatrice)

Nicolas Vauchelet, Professeur des universités, Université Sorbonne Paris Nord, UMR LAGA (Rapporteur
et examinateur)

Pauline Lafitte, Professeure des universités, Université Paris-Saclay, CentraleSupélec, laboratoire MICS,
(Examinatrice)

Samis Trevezas, Maître de conférences, National and Kapodistrian University of Athens, Department of
Mathematics (Examinateur)

Paul-Henry Cournède, Professeur des universités, Université Paris-Saclay, CentraleSupélec, laboratoire
MICS (Directeur de thèse)

Amélie Mathieu, Maîtresse de conférences, INRAE, AgroParisTech, UMR ÉcoSys (Co-encadrante de thèse)

Résumé :

Les mouvements collectifs décrivent des populations dans lesquelles les interactions entre individus sont le moteur de leurs déplacements dans l'espace et de leurs transformations dans le temps. La compréhension et le contrôle des mouvements collectifs constituent des enjeux majeurs dans de nombreux domaines, notamment pour l'étude des écosystèmes (dynamique des essaims d'animaux), la sécurité dans les grands rassemblements et les bâtiments (mouvement de foule), ou encore l'agriculture (étude de la croissance des plantes). Les modèles de population que nous considérons sont des systèmes d'équations différentielles ayant la propriété d'être hétérogènes, i.e., constituées d'individus avec des caractéristiques différentes, et ces caractéristiques ont une influence sur la dynamique. Cette hypothèse est motivée par l'application agricole, où il est question d'étudier les interactions entre plantes de différentes variétés, voire de différentes espèces. Ces systèmes sont également supposés symétriques, i.e., ayant une dynamique invariante par permutation des individus, ce qui est une caractéristique largement répandue au sein des modèles de mouvement collectif, et qui permet de nombreuses simplifications. Un certain nombre de défis restent toutefois à relever pour que ces modèles soient utilisés dans des cas d'application concrets, et nous nous concentrons en particulier sur les problèmes liés à l'inférence statistique, i.e., la confrontation du modèle à des données expérimentales, des observations réalisées sur le système réel étudié.

Un premier niveau de difficulté est d'ordre computationnel : la simulation de grandes populations en interaction peut s'avérer trop coûteuse en temps de calcul, et elle constitue ainsi un premier obstacle à l'étude de la population à une échelle macroscopique. Un second niveau de difficulté a trait à la qualité des données: du fait de la complexité du système modélisé, les observations expérimentales ne peuvent permettre de caractériser exactement la dynamique du système (en particulier car elles ne portent généralement que sur une sous-partie de la population), et il est nécessaire de quantifier les incertitudes liées aux imperfections dans l'acquisition de ces données.

Dans cette thèse, nous caractérisons l'ensemble des sources d'incertitude liées aux observations partielles des systèmes symétriques et hétérogènes dans un cadre bayésien. Certaines sources d'incertitude, notamment celle venant d'une connaissance inexacte de la taille de la population et des propriétés des individus non observés, donnent lieu à des problèmes d'inférence particulièrement difficiles, que nous nous proposons d'approcher en utilisant des représentations macroscopiques de la population. Ces approximations statistiques du mouvement global de la population sont basées sur des simulations numériques des distributions limites de champ moyen associées au mouvement collectif, distribution s'exprimant comme la solution d'une équation de transport non-locale.

Abstract :

Collective motions describe populations in which individuals' interactions are the driving force behind their displacements and their transformation over time. Understanding and controlling collective motions are significant issues in many fields, especially for the study of ecosystems (swarm dynamics), safety in large gatherings and buildings (crowd movement), or agriculture (the study of plant growth). The population models we consider are systems of differential equations with the property of being heterogeneous, i.e., made up of individuals with different characteristics influencing the dynamics.

This assumption is motivated by the agricultural application, to study interactions between plants of different varieties or even different species. These systems are also assumed to be symmetric, i.e., having dynamics invariant by permutation of individuals' labels, which is a widespread property within collective motion models, enabling numerous simplifications. However, several challenges remain to be addressed before these models can be used in real-life applications. We focus on the problems related to statistical inference, i.e., matching the model with experimental data and observations made on the system under study.

The first level of difficulty is computational: the simulation of a sizeable interacting population can be too costly in terms of computing time, and, therefore, it is a first impediment to the study of the population at a macroscopic scale. The second level of difficulty relates to the quality of the data: because of the complexity of the modeled system, experimental observations cannot characterize the system's dynamics exactly, in particular because they generally only concern a subset of the population. It is necessary to quantify the uncertainties related to the imperfections in the acquisition of such data.

In this thesis, we characterize all the uncertainty sources related to partial observations of symmetric and heterogeneous systems in a Bayesian framework. Some sources of uncertainty, notably the ones arising from inaccurate knowledge of the population size, result in particularly complex inference problems, which we propose to approach using a macroscopic representations of the population. This statistical approximation of the global population motion is based on numerical simulations of the mean-field limit distribution, i.e., a probability distribution expressed as the solution of a non-local transport equation associated with the symmetric system.}