



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

Présente

L'AVIS DE SOUTENANCE

De Monsieur Mahmoud BENTRIOU

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

"Statistical Inference and Verification of Chemical Reaction Networks"

Le 8 décembre 2021 à 10h00

À l'école CentraleSupélec, dans la **salle e.068** - Bâtiment Bouygues en présentiel et en
distanciel (pour le présentiel merci de bien vouloir contacter en avance l'assistante du
Laboratoire à l'adresse mail : Fabienne.brosse@centralesupelec.fr)

*Si vous souhaitez assister à la soutenance en distanciel, veuillez contacter en avance l'assistante du laboratoire pour obtenir
le lien.*

Membres du jury :

Adeline Leclercq-Samson, Professeure des universités (HDR), Laboratoire Jean Kuntzmann, Université
Grenoble-Alpes (Rapporteuse)

Blaise Genest, Directeur de recherche CNRS (HDR), IRISA, Université de Rouen (Rapporteur)

Benoit Barbot, Maître de conférences, LACL, Université Paris-Est Créteil (Examineur)

Pascale Le Gall, Professeure, MICS, CentraleSupélec (Examinatrice)

Paul-Henry Cournède, Professeur, MICS, CentraleSupélec (Directeur de thèse)

Paolo Ballarini, Maître de conférences, MICS, CentraleSupélec (Co-encadrant)

Résumé :

Les réseaux de réactions chimiques (CRN) constituent un formalisme utilisé pour modéliser des processus biologiques. Quand la population est de taille modérée et supposée bien mélangée, le processus stochastique sous-jacent pour décrire ses dynamiques est une chaîne de Markov en temps continu (CTMC). Ce processus est dit sans mémoire: l'état futur du système ne dépend que de l'état courant.

L'inférence statistique de ce type de CTMC est complexe: le calcul de la vraisemblance est en général difficile à résoudre. Les méthodes ABC (Approximate Bayesian Computation) forment une classe de méthodes bayésiennes sans calcul de vraisemblance qui permettent d'approcher la distribution postérieure avec des simulations de Monte-Carlo.

La vérification de modèles, qui fut à l'origine développée pour garantir la fiabilité de systèmes et logiciels informatiques, se penche de plus en plus sur la biologie des systèmes. En effet, il y a un réel besoin de comprendre les interactions complexes entre molécules dans les systèmes biologiques. Malheureusement, l'espace d'états d'un CTMC défini par un CRN explose généralement, voir est infini. Pour palier à cela, des méthodes de vérification statistiques ont été développées. Le principe est de simuler un certain nombre de fois le modèle et de calculer le ratio des simulations qui ont vérifié une propriété. Récemment, une logique temporelle appelée HASL a été introduite pour la vérification statistique de modèles: elle adopte intrinsèquement le point de vue statistique de la vérification.

Dans cette thèse, nous nous intéressons à l'inférence statistique et la vérification statistique de chaînes de Markov en temps continu définies par un modèle de réseau de réactions chimiques. Notre contribution tient principalement dans la formulation d'un algorithme ABC combiné avec le formalisme HASL appelé automate-ABC. Nous appliquons cette méthode haut niveau sur plusieurs tâches d'inférence statistique et de vérification pour des CTMCs issus de systèmes biologiques, impliquant notamment des modèles oscillatoires et des problèmes d'atteignabilité bornés en temps. L'implémentation des méthodes présentées est documentée et a conduit au développement d'une bibliothèque dans le langage de programmation Julia.

Abstract :

Chemical Reaction Networks (CRN) constitute a formalism used to model biological processes. When the population number is not significant and the system is well-stirred, a Continuous-Time Markov Chain describes its stochastic dynamics. This class of model is characterised by the memoryless property: the future state of the system only depends on the current state.

Statistical inference of such CTMCs is complex: likelihood computations are generally intractable. Approximate Bayesian Computation is a recent class of likelihood-free methods for Bayesian inference that allows approximating the posterior distribution with Monte Carlo simulations. It has proven its efficiency in the case of CTMCs.

Model-checking was initially developed for assessing hardware and software systems' reliability. There is a growing interest in the verification of models from Systems Biology to understand the complex molecular interactions within a biological system. Unfortunately, the state space of a CTMC modelled by a CRN quickly explodes or is infinite, which renders its complete exploration infeasible in practice. Statistical model checking methods have been developed to overcome this issue. They simulate the

model and compute the ratio of simulations that fulfils a property. Recently, Hybrid Automata Stochastic Logic (HASL) has been introduced for the statistical verification of stochastic models. This temporal logic inherently adopts the statistical point of view of model checking.

In this thesis, we focus on statistical inference and verification of CTMCs defined by CRNs. Our main contribution consists in the new formulation of an Approximate Bayesian Computation procedure combined with HASL called automaton-ABC. We apply this high-level method on several tasks of statistical inference and verification for biological CTMCs, including oscillatory models and time-bounded reachability problems. The implementation of our algorithms is documented and has led to a package in the Julia Programming language.