

MASTER 2
spécialité :
”Probabilités et Modèles Aléatoires”

Sorbonne Université

Brochure 2020–2021

Nouveaux cours optionnels

01/04/2020

OBJECTIFS

L'objectif de la spécialité "PROBABILITÉS ET MODELES ALEATOIRES" de la seconde année du Master de Sorbonne Université, est de délivrer une formation de haut niveau en probabilités théoriques et appliquées.

En 2020-2021, nous proposons **deux orientations** aux étudiants en fonction des cours suivis et du sujet de mémoire ou de stage choisi : l'une plus centrée sur la

- *Théorie des Processus Stochastiques*,
- l'autre sur les
- *Probabilités Appliquées*.

La première orientation prépare les étudiants à une carrière de chercheur (ou enseignant-chercheur) en milieu académique, l'autre à une carrière en milieu industriel, en passant par des stages et des thèses CIFRE.

CO-HABILITATIONS

Le master est co-habilité avec l'Ecole Normale Supérieure de Paris et avec l'Ecole des Ponts ParisTech.

CRITÈRES DE SÉLECTION

Le principal critère de sélection est un prérequis solide en mathématiques générales, et en particulier en théorie de la mesure et de l'intégration, en calcul des probabilités.

Cette spécialité s'adresse aux étudiants provenant du Master 1 de mathématiques et aux titulaires d'un Master 1 ou d'une maîtrise de *Mathématiques*, de *Mathématiques Appliquées*, ou d'*Ingénierie Mathématique*. Elle est également ouverte aux titulaires d'une maîtrise M.A.S.S., ou d'un diplôme d'ingénieur (ou éventuellement de deux ans d'études dans une école d'ingénieurs), **sous réserve d'une formation suffisante en mathématiques, et notamment en probabilités.**

PREREQUIS POUR LES ETUDIANTS DE SORBONNE UNIVERSITE

Pour les étudiants de Sorbonne Université il est *indispensable* d'avoir validé avec de bonnes notes

- Le cours Intégration en L3

et

- Le cours "Probabilités de Base" (3M290) au second semestre de L3 et le cours "Probabilités Ap-profondies" (4M011) au premier semestre de M1.

Il est possible d'être admis à titre exceptionnel après avoir validé

- le cours "Probabilités de Base" (4M010) au premier semestre de M1 et le cours "Processus de sauts" (4M036) au second semestre de M1. Cette admission se fait d'après les recommandations personnelles des professeurs de ces cours et/ou des chargés de TD.

Il est *indispensable* d'avoir validé avec de bonnes notes d'autres modules de mathématiques pour avoir une culture large et solide.

Il est souhaitable d'avoir validé un cours de statistiques mathématiques.

Le cours "Programmation en C et C++" est très souhaitable pour ceux qui s'orienteront vers des stages, des thèses CIFRE, et une carrière en entreprise.

Par contre, d'autres cours en probabilités ne sont pas indispensables.

TÉLÉ-ENSEIGNEMENT

Les étudiants salariés peuvent préparer le Master 2 "Probabilités et Modèles Aléatoires" en télé-enseignement. Pour cela, vous devez vous connecter sur le site du télé-enseignement afin de pouvoir télécharger la fiche correspondante.

La liste de cours proposés en télé-enseignement est donnée dans la brochure de cette année, elle est restreinte par rapport à la liste de cours en enseignement en présentiel.

Les critères de sélection sont *les mêmes* qu'en enseignement en présentiel.

DÉBOUCHÉS

La vocation principale de la spécialité est de former de futurs chercheurs en calcul de probabilités. Pour cette raison, la majorité des étudiants qui réussissent cette spécialité s'orientent ensuite vers le **Doctorat**, troisième composante du LMD en préparant une thèse.

La thèse peut être préparée après l'obtention du Master en s'inscrivant en Doctorat pour une durée de 2 à 4 ans en principe. Cette inscription est subordonnée à l'acceptation de l'étudiant par un " directeur de thèse".

Des allocations de recherche sont attribuées aux étudiants en thèse, sur critères *pédagogiques*.

Plusieurs étudiants préparent une thèse au sein du *Laboratoire de Probabilités Statistiques et Modélisation*, voir le site: <http://www.proba.jussieu.fr> Ce laboratoire des Sorbonne Université et Université et Denis Diderot, associé au C.N.R.S., assure la maîtrise d'oeuvre de la spécialité. Il est l'un des centres de recherche les plus actifs dans le monde dans le domaine des probabilités et des processus stochastiques. Le laboratoire regroupe environ 60 enseignants et chercheurs, sans compter les nombreux étudiants en thèse.

Les étudiants peuvent aussi préparer une thèse au sein d'autres laboratoires de recherche (universitaires ou de grandes écoles) ainsi que dans des instituts de recherche (l'INRIA, l'INRA, et autres) en France et à l'étranger.

Les étudiants qui ont choisi l'orientation "Probabilités Appliquées" le font en entreprise sous la forme de contrats CIFRE (ORANGE, EDF, CEA, AREVA, SAFRAN, SCHLUMBERGER,...).

Cette formation fournit aussi des débouchés professionnels immédiats dans des entreprises, notamment dans des banques, des sociétés d'assurance, ou des organismes financiers.

Finalement, ce master constitue un atout de carrière pour les enseignants de mathématiques agrégés dans des lycées et des classes préparatoires.

ORGANISATION DE L'ANNÉE : Cours, stage, mémoire

Premier semestre (30ECTS) (Septembre-Janvier) L'année commence par *un cours intensif de prérentrée* en septembre : son objectif est de faire un point sur les connaissances en probabilités acquises en M1 qui seront constamment utilisées en M2. Ce cours dure deux semaines, 3h par jour, et ne donne pas lieu à un examen.

Au **premier semestre** les étudiants qui ont choisi l'orientation " *Processus stochastiques*" suivent les cours

- "Processus de Markov et Applications" (9ECTS)
- "Calcul Stochastique et Processus de Diffusions"(9ECTS)
- "Théorèmes Limites pour des Processus Stochastiques" (6ECTS)
- et un cours au choix parmi deux cours "Nuages Poissoniens, Processus de Levy et Excursions"(6 ECTS), ou "Statistique et Apprentissage"(6 ECTS).

Au **premier semestre** les étudiants qui ont choisi l'orientation " *Probabilités Appliquées*" suivent les cours

- "Modèles Markoviens sur des espaces discrets" (6ECTS),
- "Calcul Stochastique et Processus de Diffusions"(9ECTS),
- "Probabilités numériques et Méthodes de Monte Carlo" (9ECTS),
- "Statistique et Apprentissage"(6 ECTS).

Deuxième semestre (30 ECTS) (Février-Juin). Les étudiants ont le choix parmi les cours spécialisés :

- "Probabilités, Algèbre et Théorie Ergodique"(6ECTS)
- "Probabilités et Physique"(6ECTS)
- "Probabilités, Méthodes Numériques, Algorithmes et Internet"(6ECTS)
- "Méthodes Stochastiques et Statistiques II"(6ECTS)
- "Géométrie Aléatoire"(6ECTS)
- "Probabilités, Neurosciences et Sciences Médicales"(6ECTS)
- "Probabilités, Biologie et Analyse des Graphes"(6ECTS)

Ces cours présentent plusieurs domaines à la pointe de la recherche en Probabilités Théoriques et Appliquées. Le contenu de chacun des cours de cette année est décrit dans la brochure.

Les cours du second semestre conduisent les étudiants à une première confrontation avec la recherche sous la forme d'un mémoire ou d'un stage. **Le mémoire** consiste en général en la lecture approfondie d'un ou plusieurs articles de recherches récents, sous la direction d'un membre du Laboratoire de Probabilités, Statistiques et Modélisation ou d'un enseignant de la spécialité. Il doit être rédigé en Latex et soutenu devant un jury.

Le mémoire peut-être remplacé par un rapport de stage. **Le stage** s'effectue dans un organisme de recherche ou un bureau d'études, sous la direction conjointe d'un ingénieur de l'organisme d'accueil et d'un enseignant de la spécialité.

Comme le travail de mémoire ou de stage s'appuie sur les connaissances obtenus en cours, ce travail ne peut pas être entamé par un étudiant qui n'a pas validé les cours de base du premier semestre.

La travail de mémoire ou de stage de courte durée (moins de 3 mois) est accredité de **12ECTS**, les étudiants doivent le compléter par la validation de **TROIS** cours optionnels au choix pour valider 30 ECTS de second semestre.

Le travail de stage industriel de longue durée (à partir de 3 mois) est accredité de **18ECTS**, les étudiants le complètent par la validation de **DEUX** cours optionnels pour valider 30 ECTS au second semestre.

RESPONSABLES

Responsable pédagogique : Madame Irina KOURKOVA,
professeur du laboratoire de Probabilités, Statistiques et Modélisation
de Sorbonne Université

Responsable administrative : Madame Josette SAMAN
Couloir 16-26, 1er étage, bureau 08
Campus Jussieu, Laboratoire de Probabilités Statistiques et Modélisation
Sorbonne Université
B.C. 158
Campus Pierre et Marie Curie
4, place Jussieu, 75252 Paris Cedex 05.
Son bureau est ouvert le lundi, mardi (fermé le mercredi) jeudi, vendredi, de 10 h – 12 h, 14 h –17 h.
Tél : (33/0) 1 44 27 53 20
Fax : (33/0) 1 44 27 76 50
E-mail : Josette.Saman@upmc.fr

COMMENT POSTULER ?

Suivre les indications sur le lien universitaire ci-dessous du début Mai au 30 Juin 2020 :

<http://sciences.sorbonne-universite.fr/fr/formations/inscriptions.html>

Il est indispensable de joindre au dossier du candidat 2020–2021

LA FICHE DE SYNTHÈSE remplie soigneusement

qui se trouve le site de la spécialité :

<https://www.lpsm.paris/formation/masters/m2-probabilites-et-modeles-aleatoires/>

Sans cette fiche proprement remplie, le dossier sera incomplet.

CALENDRIER PREVISIONNEL 2020–2021.

- Inscription administrative sur le site de la scolarité de Sorbonne Université Mi-Avril – 30 Juin 2020.
- Le premier semestre débute le lundi 7 septembre et se termine en Janvier.
- Les examens des cours du premier semestre "Modèles Markoviens sur des espaces discrets", "Processus de Markov et Applications" auront lieu au mois de novembre, les examens pour les autres cours du premier semestre auront lieu lors de la deuxième moitié du mois de Janvier 2021.
- Une réunion d'information sur les cours du second semestre aura lieu à la fin du mois de Janvier 2021.
- Le deuxième semestre commencera en Février 2021 et se terminera en Mai 2021. Une pause de deux semaines est prévue pendant les vacances de Noël et les vacances de Pâques de la zone C.
- Les examens des cours du second semestre auront lieu lors de la première moitié du mois de Juin 2021.
- *Les examens de rattrapage pour les cours du premier semestre auront lieu au mois de mars* (pour permettre aux étudiants de partir en stage en avril), pour le deuxième semestre – lors de la deuxième moitié du mois de Juin.
- Pour les étudiants désirant valider le Master en Juillet, le mémoire (ou le stage) doit être soutenu au plus tard le 25 Juin 2021.
- Pour les étudiants désirant valider le Master en Octobre, le mémoire (ou le stage) doit être soutenu au plus tard le 30 Septembre 2021.

COURS

Cours préliminaire :

L. MAZLIAK : *Rappels de probabilités*

Cours fondamentaux du premier semestre pour l'orientation "Processus Stochastiques" :

N. FOURNIER : *Calcul Stochastique et Processus de Diffusion* (9 E.C.T.S)

I. KOURKOVA : *Processus de Markov et Applications* (9 E.C.T.S)

Th. LEVY : *Théorèmes Limites pour les Processus Stochastiques* (6 E.C.T.S)

Th. DUQUESNE : *Nuages Poissoniens, Processus de Levy, Excursions* (6 E.C.T.S)

Cours fondamentaux du premier semestre pour l'orientation "Probabilités Appliquées" :

N. FOURNIER : *Calcul Stochastique et Processus de Diffusion* (9 E.C.T.S)

I. KOURKOVA : *Modèles Markoviens sur des espaces discrets*(6 E.C.T.S)

G. PAGES, V. LEMAIRE : *Probabilités Numériques, Méthodes de Monte Carlo* (9 E.C.T.S)

L. ZAMBOTTI : *Statistique et Apprentissage* (6 E.C.T.S)

Cours spécialisés du deuxième semestre

Probabilités, Algèbre et Théorie Ergodique (6 E.C.T.S)

Ph. BIANE : Processus déterminantaux et matrices aléatoires

R. DUJARDIN : Introduction à la théorie ergodique

A. ERSCHLER : Marches aléatoires sur les groupes

Probabilités et Physique (6 E.C.T.S)

F. COMETS : Modèles d'entrelacs aléatoires

Q. BERGER : Polymères aléatoires

V. VARGAS : Chaos multiplicatif gaussien et gravité de Liouville

Probabilités, Méthodes Numériques, Algorithmes (6 E.C.T.S)

V. LEMAIRE : Arrêt optimal : théorie, méthodes numériques et applications

B. JOURDAIN : Algorithmes de Monte-Carlo pour chaînes de Markov et méthodes particulières

J. SALEZ : Temps de mélange des chaînes de Markov

Méthodes Stochastiques et Statistiques II (6 E.C.T.S)

L. DECREUSEFOND : Calcul de Malliavin

I. CASTILLO : Statistiques bayésiennes non-paramétriques : convergence de mesures a posteriori

D. TALAY : Equations différentielles stochastiques de McKean–Vlasov

Géométrie Aléatoire (6 E.C.T.S)

C. BOUTILLIER et **B. DE TILIÈRE** : Dimères et pavages aléatoires

B. BLASZCZYSZYN : Modèles Géométriques Aléatoires

N. BROUTIN : Limites d'échelles de graphes aléatoires

Probabilités, Neurosciences et Sciences Médicales(6 E.C.T.S)

M. THIEULLEN : Modèles probabilistes pour les neurosciences

G. NUEL : Propagation d'évidence dans les réseaux bayésiens

Probabilités, Biologie et Analyse des Graphes(6 E.C.T.S)

A. BEN-HAMOU : Inégalités de Concentration

Ph. ROBERT : Modèles Stochastiques de la Biologie Moléculaire

Cours exposés en télé-enseignement

Tous les cours du premier semestre pour suivre l'orientation "Processus stochastiques". (Attention, les cours de l'orientation "Probabilités Appliquées" ne sont pas disponibles en télé-enseignement.)

Les cours spécialisés "Probabilités, Algèbre et Théorie Ergodique" (la partie de R. Dujardin), "Probabilités, Neurosciences et Sciences médicales" (la partie de M. Thieullen), Géométrie aléatoire (la partie de N. Broutin) du second semestre.

Rappels de Probabilités

L. Mazliak

1er semestre, 2 semaines en septembre, 30h.

Ce cours préliminaire est destiné à faire le point sur des notions fondamentales de Probabilités abordées en M1 et utilisées systématiquement par la suite dans les cours du M2. Parmi ces notions on peut citer :

- éléments de la théorie de mesure et intégration
- le conditionnement
- les différentes notions de convergence
- les martingales à temps discret.

Le cours sera complété par des séances d'exercices et une bibliographie pouvant servir de référence tout au long de l'année de M2.

Calcul Stochastique et Processus de Diffusions

N. Fournier

1er semestre, (octobre – début janvier), 4h par semaine.

Dans ce cours nous allons introduire les techniques de bases du calcul stochastique :

- 1) le mouvement brownien, la continuité de ses trajectoires, la propriété de Markov (forte)
- 2) l'intégration stochastique par rapport à une martingale de carré intégrable, la formule d'Ito, le théorème de Girsanov
- 3) les équations différentielles stochastiques (EDS) et leurs solutions faibles ou fortes (dites diffusions), les liens avec les équations aux dérivées partielles
- 4) la formule d'Ito-Tanaka, le temps local du mouvement brownien, les EDS réfléchies
- 5) EDS à coefficients non-lipschitziens, processus de Bessel

Références bibliographiques

Ikeda, N. et Watanabe, S. : *Stochastic Differential Equations and Diffusion Processes ; 2e édition.* North Holland, 1988.

Le Gall, J.-F. : *Mouvement Brownien, Martingales et Calcul Stochastique.* Springer, Collection: Mathématiques et Applications, Vol. 71 2013, VIII.

Karatzas, I. et Shreve, S. : *Brownian Motion and Stochastic Calculus ; 2e édition corrigée.* Springer, 1994.

Mörters, P. et Peres, Y. : *Brownian Motion.* Cambridge University Press, 2010.

Revuz, D. et Yor, M. : *Continuous Martingales and Brownian Motion, 3e édition.* Springer, 1999.

Rogers, L.C.G. et Williams, D. : *Diffusions, Markov Processes and Martingales, Vol. II, Itô Calculus.* Wiley, 1987.

Processus de Markov et Applications

I. Kourkova

1er semestre, 1ère – 8ème semaines (septembre – début novembre), 6h par semaine.

Chaînes de Markov sur un espace d'états dénombrable : classification, mesures invariantes, comportement limite, théorèmes ergodiques. On considèrera des chaînes de Markov apparaissant en biologie (modèles de Wright-Fisher, de Moran), en dynamique des population (chaîne de naissance et de mort, modèle de Galton-Watson) et en mécanique statistique. Processus de Markov de saut pur (Chaînes de Markov sur un espace dénombrable en temps continu.), leur matrice d'intensité, les équations de Kolomogrov, le phénomène d'explosion, le comportement limite, la réversibilité. On considèrera des applications en biologie et des modèles d réseaux stochastiques.

Processus de Markov sur un espace mesurable général, familles de Markov Félieniennes, propriété de Markov forte. Le semi-groupe, le générateur et la résolvante, divers exemples de calculs dont le générateur du mouvement Brownien et autres. Le théorème affirmant que le générateur détermine le sémi-groupe à travers la résolvante. Calcul de la mesure invariante et de lois de fonctionnelles d'un processus de Markov à partir de son générateur. Formule de Feynman-Kac. Problème de martingales. Solutions des équations différentielles stochastiques en tant que processus de Markov. Leurs générateurs. Leurs fonctionnelles "arrêtées". Liens avec les équations aux dérivées partielles elliptiques avec de différentes conditions au bord.

Théorèmes Limites pour les Processus Stochastiques.

Th. Levy

1er semestre, 3ème – 14ème semaines (octobre – décembre), 2h par semaine pendant 12 semaines.

Théorème de Donsker. Introduction à la mécanique statistique. Introduction à la théorie de grandes déviations. Initiation aux probabilités non-commutatitves.

Nuages Poissoniens, Processus de Levy, Excursions.

Th. Duquesne

1er semestre, 10ème – 15ème semaines (fin novembre – début janvier), 6h par semaine sur 4 semaines, ou 4h par semaine sur 6 semaines.

Objectifs de l'UE : ce cours est un approfondissement du cours "Processus de Markov et Applications". On y étudie les processus de Lévy ainsi que certains processus de branchement. On introduira les mesures ponctuelles de Poisson et on exposera la théorie des excursions avec des applications aux processus de Lévy.

Prérequis : il est fortement recommandé d'avoir suivi le cours "Processus de Markov et Applications" et le cours sur les théorèmes limites.

Thèmes abordés :

- Retour sur les processus de Feller.
- Mesures ponctuelles de Poisson.
- Processus de Lévy.
- Processus de branchement à espace d'états continu.
- Théorie des excursions.

Probabilités Numériques, Méthodes de Monte-Carlo

G. Pages, V. Lemaire

1er semestre, 3ème – 15ème semaines (octobre – début janvier), 2h de cours, 3h d'illustrations numériques en C++11 par semaine.

Objectifs du cours :

Le but de ce cours est de présenter les méthodes de Monte-Carlo et de Quasi-Monte-Carlo d'usage courant . De nombreux exemples issus de problèmes de calcul de prix et de couverture d'options illustrent les développements. Une mise-en-oeuvre informatique des techniques abordés sera effectuée lors des séances de TD. Chaque étudiant devra réaliser, en binôme, un projet informatique (en langage C) implémentant soit des calculs de prix et de couvertures d'options soit des simulations d'autres modèles. Il remettra un rapport décrivant les méthodes utilisées et commentant les résultats obtenus. Ce cours aborde les thèmes suivants :

- 1.Introduction à la simulation : génération de variables aléatoires suivant les lois usuelles.
- 2.Méthode de Monte-Carlo : calcul d'espérance par simulation.
- 3.Méthodes de réduction de variance : variables de contrôle, échantillonnage préférentiel, variables antithétiques, stratification, conditionnement.
- 4.Quasi-Monte-Carlo : techniques de suites à discrédances faibles.
- 5.Méthodes de gradient stochastique et de Bootstrap.
- 6.Discrétisation en temps des équations différentielles stochastiques (schéma d'Euler, de Milshtein) : application au pricing d'options européennes.
- 7.Amélioration de la méthode dans le cas d'options path-dépendent : ponts browniens,...
- 8.Calcul des couvertures et sensibilités par méthode de Monte-Carlo
- 9.Introduction à la simulation des processus à sauts.

En parallèle de ce cours, 8 séances de 3 heures seront consacrées à la mise en oeuvre de ces algorithmes en C++11, en illustrant différents concepts de programmation: la programmation orientée objet, la programmation générique et la programmation fonctionnelle.

L'évaluation du cours se fait par un examen et le rendu d'un projet informatique (au second semestre): chaque étudiant devra réaliser, en binôme, un projet informatique (en langage C/C++) implémentant soit des calculs de prix et de couvertures d'options soit des simulations de modèles financiers. Il remettra un rapport décrivant les méthodes utilisées et commentant les résultats obtenus.

Modèles Markoviens sur des espaces discrets

I. Kourkova

1er semestre, 1ère – 4ème semaines (septembre – début novembre), 6h par semaine.

Chaînes de Markov sur un espace d'états dénombrable : classification, mesures invariantes, comportement limite, théorèmes ergodiques. On considèrera des chaînes de Markov apparaissant en biologie (modèles de Wright-Fisher, de Moran), en dynamique des population (chaîne de naissance et de mort, modèle de Galton-Watson) et en mécanique statistique. Processus de Markov de saut pur (Chaînes de Markov sur un espace dénombrable en temps continu), leur matrice d'intensité, les équations de Kolomogrov, le phénomène d'explosion, le comportement limite, la réversibilité. On considèrera des applications en biologie et des modèles d réseaux stochastiques.

Statistique et Apprentissage

L. Zambotti

*1er semestre, 10ème – 15ème semaines (fin novembre – début janvier), 6h par semaine sur 4 semaines,
ou 4h par semaine sur 6 semaines.*

Objectifs : Ce cours vise à donner aux étudiants les bases fondamentales du raisonnement et de la modélisation statistique, tout en présentant une ouverture vers des thématiques de recherche contemporaines. L'accent sera particulièrement mis sur l'utilisation pratique des nouveaux objets rencontrés.

Prérequis : Une bonne connaissance du calcul des probabilités et de l'algèbre linéaire.

Thèmes abordés :

- Rappels de probabilités, estimation ponctuelle, estimation par intervalles, tests.
- Modèle linéaire : estimation, intervalles de confiance et tests.
- Introduction à l'apprentissage statistique et à la classification supervisée.
- Minimisation du risque empirique, théorème de Vapnik-Chervonenkis.
- Règles de décision non paramétriques (méthode des k plus proches voisins et arbres de décision).
- Quantification et classification non supervisée.

Probabilités, Algèbre, Théorie Ergodique

2ème semestre, Février-Mai.

PARTIE I : **Ph. Biane**, Processus déterminantaux et matrices aléatoires

On étudiera les processus déterminantaux en lien avec la théorie des matrices aléatoires, notamment le GUE.

PARTIE II : **R. Dujardin**, Introduction à la théorie ergodique

Cette partie du cours porte sur la théorie ergodique des systèmes dynamiques déterministes et aléatoires, c'est à dire l'étude de ces systèmes dynamiques par la théorie de la mesure et des probabilités. Une question fondamentale dans ce contexte est la description du comportement asymptotique des orbites typiques au sens de la mesure.

On introduira dans un premier temps les concepts, exemples et résultats fondamentaux de la théorie déterministe:

- le vocabulaire des systèmes dynamiques, et les exemples de base (dynamique sur le cercle, sur les tores, systèmes symboliques) - les notions d'ergodicité, de mélange, les théorèmes ergodiques classiques.

La deuxième partie du cours portera sur les systèmes dynamiques aléatoires. Les notions suivantes devraient être abordées (liste non exhaustive):

- mesures stationnaires vs mesures invariantes, théorèmes ergodiques aléatoires - cocycles et théorème d'Oseledec - produits aléatoires de matrices - principe d'invariance

PARTIE III : **A. Erschler** : Marches aléatoires sur les groupes

Il y a une double motivation possible pour une étude des marches aléatoires sur les groupes. Du point de vue de la théorie des groupes, les marches aléatoires sur les groupes fournissent de nombreux invariants probabilistes du groupe. Dans le cas de groupes de type fini, ces invariants sont étroitement liés à la géométrie des graphes de Cayley correspondants. Un problème plus difficile qui reste un défi est de relier les invariants probabilistes aux propriétés algébriques du groupe en question.

Du point de vue des marches aléatoires, les espaces homogènes fournissent un contexte riche qui généralise des exemples classiques de marches aléatoires sur \mathbf{Z}^d . L'invariance du noyau de Markov par rapport à l'action de groupe impose la structure naturelle et des propriétés stochastiques intéressantes de marches aléatoires et de leurs trajectoires.

Les sujets du cours incluront la récurrence / la transience, les probabilités de transition, l'isopérimétrie, la vitesse de la fuite et l'entropie, comportement asymptotique des trajectoires, bord de Martin et de Poisson des marches aléatoires.

Contenu :

- les estimations gaussiennes de Carne Varopoulos
- les inégalités entre les probabilités de transition et le profil isopérimétrique
- l'absence de fonctions harmoniques positives pour les marches aléatoires symétriques sur les groupes nilpotents (Margulis)
- le critère d'entropie (Kaimanovich Vershik et Derriennic)
- caractérisation de groupes moyennables par l'existence de la mesure dont le bord de Poisson est trivial (Furstenberg, Rosenblatt, Kaimanovich Vershik)
- la construction des mesures dont le bord de Poisson est non triviale sur tous les groupes de type fini, l'exception de groupes virtuellement nilpotents (Hartmann, Frisch, Tamuz et Vahidi-Ferdowsi, 2018)

Bibliographie :

W. Woess. Random Walks on Infinite Graphs and Groups. Cambridge Tracts in Mathematics 138, Cambridge University Press, 2000

V.A.Kaimanovich, A.M.Vershik. Random walks on discrete groups: boundary and entropy.. Ann. Probab. 11 (1983), no. 3, 457 - 490.

R. Lyons, Yu. Peres. Probability on trees and networks.. Cambridge Series in Statistical and Probabilistic Mathematics, 42. Cambridge University Press, New York, 2016

J. Frisch, Y. Hartman, O. Tamuz, P. Vahidi-Ferdowsi. Choquet-Deny groups and the infinite conjugacy class property. preprint 2018 <https://arxiv.org/abs/1802.00751>

Probabilités et Physique

PARTIE I : F. Comets Modèles d'entrelacs alatoires

Introduit par A.-S. Sznitman en 2010, le modèle d'entrelacs alatoires est un processus de Poisson dont les points sont des chemins doublement infinis d'une promenade alatoire simple sur $\mathbf{Z}^d, d \geq 3$. La mesure d'intensité de Poisson est telle qu'un paramètre $u > 0$ détermine le nombre de trajectoires qui forment ensemble la structure d'entrelacement. A un tel niveau u , la probabilité qu'un ensemble fini $K \subset \mathbf{Z}^d$ appartienne à l'ensemble des points de réseau vacants non visités par aucune marche alatoire est $\exp\{-u \text{cap}(K)\}$, avec $\text{cap}(K)$ la capacité de l'ensemble K .

L'ensemble vacant présente des corrélations à décroissance lente, des propriétés percolatives intéressantes avec un phénomène de transition de phase pour sa connectivité. Les entrelacs alatoires donnent la limite locale de la trace laissée par des marches alatoires sur un tore de grande taille en d dimensions ($d \geq 3$) à des échelles de temps appropriées.

La construction se généralise à des graphes transients, et ainsi elle fournit encore des limites locales en dimension 1 ou 2 pour la marche conditionnée. Avec le mouvement Brownien elle se généralise aussi à \mathbf{R}^d , où elle jouit de propriétés analogues.

Mots-clé :

- Marche alatoire, mouvement Brownien, processus de Poisson
- théorie du potentiel, capacité, fonction de Green, transience, récurrence,
- couplage, inégalité FKG, percolation,
- champ libre Gaussien, théorème d'isomorphisme de Dynkin.

PARTIE II : Q. Berger, Polymères alatoires

Mots-clés: marches alatoires, processus de renouvellement, transition de phase, systèmes désordonnés.

Les polymères sont de très longues molécules, constituées d'unités de base (des monomères): en mathématiques, ils sont modélisés par la trajectoire d'une marche alatoire. Dans ce cours, on étudiera certains modèles de polymères, que l'on utilisera comme prétexte pour aborder certaines problématiques standard de mécanique statistique.

Un exemple sur lequel on se concentrera est le modèle d'accrochage: il considère une marche alatoire contrainte à être positive, mais "récompensée" à chaque retour en 0 (par un paramètre e^β). Cela modélise un polymère placé au-dessus d'une surface impénétrable, les monomères ayant une interaction attractive avec la surface. Un phénomène de transition de phase a lieu: il existe un β_c tel que les trajectoires typiques de la marche alatoire sont "repoussées" par la surface pour $\beta < \beta_c$, et sont "accrochées" à la surface pour $\beta > \beta_c$. Nous introduirons les outils qui permettent d'étudier cette transition de phase, que nous décrirons précisément.

On considérera aussi une version *désordonnée* de ce modèle, correspondant au cas où le polymère est inhomogène, et possède différents types de monomères. L'interaction avec la surface est différente pour chaque type de monomère, et la récompense associée au i^e monomère est $e^{\beta+\epsilon_i}$, où $(\epsilon_i)_{i \geq 0}$ est une suite de variables alatoires. Nous verrons ce que l'on peut dire sur ce système désordonné, et notamment sur le rôle du désordre dans la transition de phase.

Référence: G. Giacomin, *Random Polymer Models*, Imperial College Press.

PARTIE III : V. Vargas Chaos multiplicatif gaussien et gravité de Liouville

La théorie des champs de Liouville fut introduite en physique théorique par Polyakov en 1981 dans sa théorie de sommation des métriques sur une surface de Riemann (appelé "Quantum geometry of bosonic strings"). Par conséquent, la théorie de Liouville peut être considérée comme une version alatoire des surfaces de Riemann. En physique, cette théorie est définie formellement par une version 2d des intégrales de chemin à la Feynman. Le but de cette partie du cours est de donner une construction probabiliste rigoureuse de la théorie, en utilisant le champ libre gaussien (GFF) et l'exponentiel du champ libre appelé le chaos multiplicatif gaussien (d'après J.P. Kahane). On verra que la théorie possède des symétries qui en font une théorie conforme. On énoncera également des conjectures précises reliant la théorie de Liouville à la limite d'échelle des grandes cartes planaires plongées de façon conforme dans la sphère. Voici le résumé de cette partie du cours :

- Introduction au GFF et au chaos multiplicatif gaussien
- Construction des corrélations et de la forme volume associées à la théorie de Liouville. Propriétés de base de la théorie: invariance conforme, anomalie de Weyl, etc...
- Relation KPZ: relation (conjecturelle) entre les corrélations de la théorie de Liouville et certaines observables des grandes cartes planaires plongées dans la sphère (forme volume, champ de spin, etc...)

Bibliographie

Berestycki : An elementary approach to Gaussian multiplicative chaos arXiv:1506.09113

Gawedzki : Lectures on Conformal Field Theory. Disponible sur le site de l'IAS

Rhodes, Vargas : Lecture notes on Gaussian multiplicative chaos and Liouville Quantum Gravity arXiv:1602.07323

Sheffield : Gaussian free fields for mathematicians arXiv:math/0312099

Probabilités, Méthodes Numériques, Algorithmes

PARTIE I : V. Lemaire Arrêt optimal : théorie, méthodes numériques et applications

Voici le plan de cette partie du cours.

- Arrêt optimal en temps continu (cas régulier) : rappels sur les supremum essentiels, les surmartingales et martingales en temps continu (régularisation, décomposition de Doob-Meyer ...).
- Enveloppe de Snell, caractérisation des temps d'arrêt optimaux, plus petit et plus grand temps d'arrêt optimal, formulation duale de l'enveloppe Snell.
- Introduction à la théorie de l'AOA sur les marchés financiers (complets).
- Valorisation d'options américaines en marché complet (marché brownien avec actifs multidimensionnels sous forme de processus d'Itô) : lien avec l'arrêt optimal en temps continu, portefeuille de réplcation, stratégie de couverture.
- Formulations duales (Rogers 2002 ; Haugh-Kogan 2002 ; Jamshidian 2005).
- Etude analytique du prix de l'option américaine dans le cadre du modèle de Black-Scholes : propriété de continuité, de monotonie, de convexité, inéquations variationnelles, frontière libre, formule semi-fermée via la frontière libre, smooth-fit.
- Étude d'exemples.

Méthodes numériques (éléments)

Description et analyse succincte de quelques méthodes numériques de valorisation et de couverture pour les options américaines *via* des approximations bermudéennes.

- Itération sur les fonctions valeurs: régression non paramétrique (Carrière 1996), maillage aléatoire (Broadie-Glasserman 1997), quantification optimale (Bally-Pagès 2001), calcul de Malliavin (Lions-Régnier 2001).
- Itération sur les temps d'arrêt: approximation de la valeur de continuation par projection L^2 (Longstaff-Schwartz 2001).
- Calcul des couvertures: méthode de flot (Piterbarg 2002), méthodes de projection, de régression.

PARTIE II : B. Jourdain, Algorithmes de Monte-Carlo pour chaînes de Markov et méthodes particulières

Pour simuler suivant une loi cible π sur un espace E muni d'une tribu \mathcal{E} , l'algorithme de Metropolis-Hastings construit une chaîne de Markov $(X_n)_{n \geq 1}$ qui admet π comme mesure réversible. Cet algorithme est l'une des méthodes numériques probabilistes les plus largement utilisées en statistique, en physique, en chimie quantique comme en simulation moléculaire si bien qu'une recherche google sur son nom fournit plus de 350 000 résultats. Lorsque la chaîne de Markov est ergodique, la loi de X_n converge vers π lorsque n tend vers l'infini et pour $\varphi : (E, \mathcal{E}) \rightarrow \mathbf{R}$ mesurable et telle que $\int_E |\varphi(x)| \pi(dx) < \infty$, la loi forte des grands nombres ergodique assure la convergence presque sûre de $\frac{1}{n} \sum_{k=1}^n \varphi(X_k)$ vers $\int_E \varphi(x) \pi(dx)$ lorsque $n \rightarrow \infty$. Dans cette partie du cours, nous illustrerons sur l'exemple de cet algorithme la condition suffisante d'ergodicité pour une chaîne de Markov à valeurs dans un espace général donnée dans [?] et nous établirons sous cette condition la loi forte des grands nombres ergodique. Nous montrerons le théorème de la limite centrale associé sous une condition renforcée qui permet de résoudre l'équation de Poisson. Nous énoncerons le résultat de Peskun [?] qui explique comment minimiser la variance asymptotique. Nous verrons comment guider le choix de la variance des propositions de l'algorithme de Metropolis-Hastings par marche aléatoire à l'aide des résultats de scaling [?]: lorsque la loi cible est une probabilité produit sur un espace dont la dimension tend vers l'infini, la première composante de l'algorithme accélère en temps converge en loi vers une diffusion. Enfin, lorsque la convergence vers l'équilibre est très lente (phénomène de métastabilité), nous étudierons comment l'accélérer par des versions de l'algorithme avec échantillonnage préférentiel adaptatif comme Wang-Landau, SHUS, metadynamics ou well-tempered metadynamics.

Dans un second temps, nous étudierons les algorithmes particuliers génétiques [?] qui en alternant une étape de sélection où les particules interagissent avec une phase de mutation où elles évoluent de façon indépendante suivant un noyau markovien permettent de réduire la variance ou de générer des événements rares. Voici la BIBLIOGRAPHIE.

- Del Moral, P.: Feynman-Kac formulae. Genealogical and interacting particle systems with applications. Probability and its Applications (New York). Springer-Verlag, 2004.
- Gelman, A., Gilks, W. R. and Roberts, G. O., Weak convergence and optimal scaling of random walk Metropolis algorithms, Ann. Appl. Probab. 7(1):110-120, 1997
- Hairer, M., Mattingly, J. C., Yet another look at Harris' ergodic theorem for Markov chains. Seminar on Stochastic Analysis, Random Fields and Applications VI, 109-117, Progr. Probab., 63, Birkhauser/Springer Basel AG, Basel, 2011.
- Peskun, P. H., Optimal Monte-Carlo sampling using Markov chains, Biometrika 60(3):607-612, 1973.

PARTIE III : J. Salez, Temps de mélange des chaînes de Markov

Combien de fois faut-il battre un paquet de 52 cartes pour que la permutation aléatoire obtenue soit à peu près uniformément distribuée ? Cette partie du cours est une introduction sans pré-requis à la théorie moderne des temps de mélange des chaînes de Markov. Un intérêt particulier sera porté au célèbre phénomène de "cutoff", qui est une transition de phase remarquable dans la convergence de certaines chaînes vers leur distribution stationnaire. Parmi les outils abordés figureront les techniques de couplage, les martingales, l'analyse spectrale, et les inégalités fonctionnelles de Poincaré et log-Sobolev. En guise d'illustration, ces méthodes seront appliquées à divers exemples classiques issus de contextes variés: mélange de cartes, marches aléatoires sur les groupes, systèmes de particules en interaction, algorithmes de Metropolis-Hastings, etc. Une place importante sera accordée aux marches sur réseaux, qui sont aujourd'hui au coeur des algorithmes d'exploration d'Internet et sont massivement utilisées pour la collecte de données et la hiérarchisation des pages par les moteurs de recherche.

BIBLIOGRAPHIE: La partie du cours sera basée sur les deux ouvrages suivants, qui sont disponibles gratuitement en ligne:

- Markov Chains and Mixing Times (Levin, Peres and Wilmer)
- Mathematical Aspects of Mixing Times in Markov Chains (Montenegro and Tetali).

Méthodes Stochastiques et Statistiques II

Partie I : L. Decreusefond, Calcul de Malliavin

A la différence du calcul stochastique d'Itô qui utilise la structure temporelle des processus à travers la notion de martingale, le calcul de Malliavin exploite les propriétés des processus gaussiens, à l'instar du mouvement brownien, pour définir le cadre d'une analyse en dimension infinie. L'objet principal en est le gradient de Stroock-Malliavin et la divergence associée qui généralise l'intégrale d'Itô à des processus non adaptés. Ce cours est une introduction à ces concepts ainsi qu'à leurs applications: théorème de Girsanov généralisé, critère d'absolue continuité pour les solutions d'EDS, formule explicite d'Itô-Clark, calcul anticipatif, délit d'initié, calcul des grecques, transport optimal, etc.

Programme

Les éléments d'analyse fonctionnelle nécessaires (produits tensoriels d'espace de Hilbert, opérateurs Hilbert-Schmidt et opérateurs à trace, etc) seront introduits au fur et à mesure du cours.

La validation de cet enseignement se fait par analyse d'un article scientifique.

- Rappels sur les processus gaussiens
- Espace de Wiener abstrait, gradient.
- Divergence
- Processus d'Ornstein-Uhlenbeck
- Inégalités de Meyer
- Espaces de Sobolev
- Distributions et formule d'Itô-Clark
- Théorème de Girsanov-Ramer
- Calcul des grecques

PARTIE II : I. Castillo, Bayésien non paramétrique et modèles de grande dimension

L'approche bayésienne en statistiques consiste à munir l'espace des paramètres du modèle statistique d'une loi de probabilité, la loi a priori. Chaque loi du modèle est vue comme la loi des données conditionnellement au paramètre. L'estimateur bayésien est alors la loi conditionnelle du paramètre sachant les données. C'est une mesure de probabilité, aléatoire travers sa dépendance en les données. L'utilisation pratique d'algorithmes basés sur des lois a posteriori a connu un essor très important depuis la fin des années 1990, avec de nombreuses applications en statistique et apprentissage.

Dans ce cours, nous introduisons d'abord des outils généraux qui permettent d'étudier du point de vue mathématique la convergence de lois a posteriori, d'après des travaux de Ghosal et van der Vaart (2000-2010). Ensuite, nous considérons des applications dans deux cadres : les modèles non-paramétriques et les modèles de grande dimension avec contrainte de parcimonie.

Dans les modèles non-paramétriques, le paramètre est une fonction ou une autre quantité de dimension infinie. Pour construire une loi a priori sur de tels objets, on peut naturellement utiliser des processus stochastiques, tels que les processus gaussiens, le processus de Dirichlet, des lois de type cascades multiplicatives etc. Nous verrons plusieurs exemples en détail, et étudierons le comportement des lois a posteriori correspondantes.

Un deuxième champ applicatif est celui de modèles de grande dimension, où le paramètre est de dimension potentiellement très grande mais seul un petit nombre de ses coordonnées est réellement significatif. Ces modèles sont très utilisés notamment en génomique, où il peut s'agir par exemple d'identifier un petit nombre de gènes responsables d'une maladie. Nous introduirons des lois a posteriori 'parcimonieuses' et aborderons leurs propriétés de convergence.

Si le temps le permet, nous évoquerons brièvement des applications à d'autres aspects inférentiels: construction de régions de confiance, tests multiples, approche variationnelle.

Une bibliographie du cours et les informations pratiques se trouvent sur la page web d'Ismael Castillo.

PARTIE III : D. Talay Equations différentielles stochastiques de McKean–Vlasov

Ce cours a pour but d'introduire les équations différentielles stochastiques non linéaires au sens de McKean–Vlasov, la propagation du chaos de systèmes de particules correspondants, ainsi que quelques techniques permettant d'aborder des interactions singulières.

On insistera sur les problèmes de martingale non linéaires et les équations de Koker–Planck–McKean–Vlasov.

Le plan approximatif du cours est le suivant :

1. Prérequis divers : problèmes de martingales, estimations de densités de diffusions, distances de Wasserstein.
2. Équations différentielles stochastiques non linéaires au sens de McKean–Vlasov avec interactions régulières.
3. Systèmes de particules avec interactions McKean–Vlasov: caractère bien posé et propagation du chaos.
4. Ouvertures vers des modèles à interactions singulières.

Références :

Plusieurs sources sont partiellement utilisées, notamment le début du cours d'A-S. Sznitman à Saint-Flour, et divers travaux de M. Bossy, B. Jourdain, S. Méléard et moi-même.

Le cours valorisera les allers retours entre probabilités et EDP.

Géométrie Aléatoire

PARTIE I : C.Boutillier et B. De Tilière, Dimères et pavages aléatoires

Un *domino* est l'union de deux carrés unité partageant une arête. Étant donné un rectangle $m \times n$, est-il possible de le *paver* avec des dominos, c'est-à-dire de couvrir sa surface avec des dominos sans qu'il y ait de chevauchements ? Si oui, de combien de façons ? À quoi ressemble un pavage *typique* ? Qu'en est-il pour un autre domaine obtenu en découpant une portion du réseau \mathbf{Z}^2 le long d'arêtes ?

En remplaçant \mathbf{Z}^2 par le réseau triangulaire et les dominos par des losanges obtenus en accolant deux triangles adjacents, on obtient un modèle de pavages par losanges. Les pavages par dominos et par losanges sont des exemples de *modèles de dimères*. Ces modèles étudiés par les physiciens (Fisher, Kasteleyn, Temperley...) dans les années 1960 ont connu un regain d'intérêt dans la communauté mathématique à la fin des années 1990 qui a conduit à des développements impressionnants de la théorie (Cohn, Johansson, Kenyon, Okounkov, Propp, Sheffield, Wilson...)

Nous étudierons d'abord certains aspects combinatoires relatifs au dénombrement des configurations de ces modèles, ainsi que les relations avec d'autres modèles combinatoires (surfaces aléatoires, arbres couvrants, marches aléatoires à boucles effacées,...).

Ensuite, nous discuterons de la forme typique d'un pavage par dominos d'un grand domaine (phénomène du cercle arctique, forme limite déterministe). Les fluctuations autour du comportement limite macroscopique peuvent être reliées au spectre des grandes matrices aléatoires d'une part, et au champ libre gaussien sans masse d'autre part, impliquant des propriétés d'invariance conforme de ces modèles dans la limite d'échelle.

Puis, nous étudierons ces modèles sur des réseaux bipartis périodiques planaires infinis, en donnant une classification des mesures de Gibbs ergodiques, et en mettant en relief le lien entre quantités probabilistes et objets algébriques liés à la structure de ces réseaux.

Prérequis pour la partie 1. Programme de probabilités de M1. Des notions sur les processus gaussiens et les matrices aléatoires pourront être appréciées pour certaines parties du cours, mais seront rappelées le moment venu.

PARTIE II : B.Blaszczyszyn Modèles géométriques aléatoires

Le cours fournit un accès rapide à certains modèles populaires dans la théorie de graphes aléatoires, processus ponctuels et ensembles aléatoires. On rencontre ces modèles dans l'analyse mathématique de réseaux (sociaux, de communication, biologique, etc). Le cours est composé des quinze leçons suivantes:

- Percolation sur la grille carrée,
- Arbre de Galton-Watson,
- Graphe d'Erdős-Rényi — l'émergence de la composante géante,
- Modèle de configuration — graphe avec la distribution des degrés donnée,
- Graphes aléatoires unimodulaires — noeud typique du graphe,
- Graphe d'Erdős-Rényi — l'émergence de la connectivité,
- Processus ponctuel de Poisson,
- Probabilités de Palm — conditionnement par un point,
- Processus à noyau dur (hard core),
- Processus ponctuels stationnaires — principe de transport de masse,
- Mosaique stationnaire de Voronoi — formules inverse et d'échange de Neveu,
- Ergodicité et les point-shifts invariants,
- Ensembles fermés aléatoires,
- Modèle Booléen et les processus de couverture,
- Connexité des ensembles aléatoires et la percolation continue.

L'attention sera portée sur les relations entre les différents modèles et sur les concepts fondamentaux. Par exemple, le principe de transport de masse relie les graphes unimodulaires et les processus ponctuels stationnaire en permettant de définir l'élément typique dans des grandes structures aléatoires, régulières.

PARTIE III : N. Broutin Limites d'échelles de graphes aléatoires

A quoi ressemble un grand arbre choisi au hasard ? Et un graphe ? Voir par exemple les simulations données sur la page web de Nicolas Broutin.

Peut-on justifier formellement l'impression que donne les simulations, à savoir que ces structures on l'air d'être fractales ?

L'objet du cours est de tenter de répondre à ces questions et de présenter certaines limites de graphes aléatoires considérés en temps qu'espaces métriques. Il s'agira à la fois (a) d'introduire des objets centraux intimement liés au mouvement brownien, (b) de présenter un ensemble de techniques qui sont basées sur des représentations combinatoires explicites et (c) d'étudier les applications aux graphes aléatoires dans le régime dit critique. En particulier, les relations entre objets discrets et continus seront au centre de nos préoccupations.

Le cours comportera deux parties: dans un premier temps, nous considérerons des arbres aléatoires de type Galton–Watson et leurs limites. Nous parlerons en particulier des différents encodages des arbres, de leur convergences, ainsi que du point de vue 'objectif' qui consiste à les considérer comme des espaces métriques (mesurés); ce sera l'occasion de parler de la topologie de Gromov-Hausdorff sur les classes d'isométries d'espaces métriques compacts. Cela nous permettra d'introduire l'arbre continu brownien de plusieurs manières (en particulier comme métrique aléatoire sur $[0, 1]$, ou encore par découpage de \mathbf{R}_+ et reorganisation/recollage des morceaux), et d'étudier certaines de ses propriétés. Nous verrons en particulier qu'il s'agit d'un objet fractal qui est au coeur de la construction de nombreux objets limites de structures combinatoires (de la même manière que le mouvement brownien est central pour les convergences fonctionnelles).

Nous verrons ensuite comment, à partir des techniques d'explorations, il est possible d'obtenir la limites de certains graphes aléatoires dans le régime dit 'critique' qui précède l'émergence d'une composante connexe macroscopique. En particulier, nous construirons les objets limites à partir de l'arbre brownien continu. Là encore, nous nous efforcerons de développer plusieurs points de vue complémentaires. Nous parlerons aussi de processus de coalescence (en particulier le coalescent multiplicatif) que l'on peut définir comme le processus qui régit la dynamique des tailles des composantes connexes d'un graphe aléatoire lorsque l'on ajoute des arêtes à la bonne vitesse.

Nous aborderons ensuite certains de ces thèmes choisis parmi les suivants: universalité de l'arbre continu brownien, universalité des limites de graphes aléatoires, processus de fragmentation sur l'arbre continu brownien, généalogie de la fragmentation de l'arbre continu et arbre des coupes, constructions explicites des coalescents additifs et multiplicatifs standards.

Probabilités, Neurosciences et Sciences Médicales

PARTIE I : **M. Thieullen**, Modèles Probabilistes pour les Neurosciences

Les modèles stochastiques sont indispensables pour décrire la variabilité des phénomènes observés en Neurosciences. Les modèles probabilistes bâtis sur l'observation des phénomènes physiologiques ou sur des perturbations de modèles déterministes antérieurs conduisent à de nombreuses questions car ils ne satisfont pas toujours les hypothèses classiques. Le cours abordera certaines de ces questions: premier temps de passage, formule de Feynman-Kac, systèmes lents-rapides, applications des grandes déviations, comportement stationnaire, approximation diffusion, processus de Markov déterministes par morceaux, modèles champ moyen, propagation du chaos. Le lien avec les modèles basés sur des équations aux dérivées partielles sera souligné.

PARTIE II : **G. Nuel**, Propagation d'évidence dans les réseaux bayésiens

Notion de réseaux bayésien (vu comme une généralisation des modèles Markovien discrets) - notion d'évidence, marginalisation - notion de junction tree, heuristiques de construction - notion de messages, théorèmes fondamentaux - algorithmes de propagation, inward/outward, lois jointes - applications diverses (chaines de Markov conditionnées par ses deux extrémités, chaînes de Markov cachées sous contraintes, arbres Markoviens avec boucles, etc.) - calcul et maximisation de la vraisemblance en présence de données complètes - maximisation de la vraisemblance en présence de données incomplètes (par exemple par algorithme EM ou par optimisation multi-dimensionnelle directe)

L'ensemble du cours sera illustré par de nombreux exemples, notamment dans le contexte biomédical (diagnostique d'une maladie, prise en charge d'un patient aux urgences, génétique humaine, etc.), pour lesquels les calculs seront implémentés sous le logiciel R (pas de pré-requis car niveau technique de programmation assez faible).

NB: bien que le mot clef bayésien soit dans l'intitulé du cours, celui-ci ne traite absolument pas l'inférence bayésienne.

Probabilités, Biologie et Analyse des Graphes

PARTIE I : Ph. Robert, Modèles Stochastiques de la Biologie Moléculaire

Ce cours présente plusieurs modèles mathématiques fondamentaux de la biologie moléculaire où les phénomènes aléatoires jouent un rôle-clé. Aucune notion de biologie n'est prérequis.

On s'intéressera tout d'abord à l'expression du gène, i.e. la production de protéines dans les cellules prokaryotes (comme les bactéries). En raison du milieu désordonné du cytoplasme de ces cellules, les expériences montrent une grande variabilité du nombre de protéines d'un type donné dans les cellules d'une même culture. Les modèles dans ce contexte ont pour objectif d'identifier les paramètres de la cellule qui permettent de contrôler la variabilité de la production de protéines.

La deuxième partie s'intéressera aux phénomènes de polymérisation dans un cadre biologique. Certaines protéines à l'intérieur de la cellule ont la propriété de pouvoir s'assembler en longues fibres appelées polymères. De nombreux processus biologiques utilisent ces mécanismes qui contribuent au bon fonctionnement des cellules, pour l'élaboration du cytosquelette notamment. Dans certains cas cependant ces phénomènes peuvent être pathologiques, dans les cellules nerveuses notamment à des maladies comme celle d'Alzheimer semblent être liées à ce type de mécanismes. On observe dans les expériences in vitro que, au bout d'un temps très variable suivant les expériences, la concentration en polymères passe de la valeur 0 à une valeur élevée. Les modèles probabilistes utilisés ont pour objet de pouvoir expliquer la variabilité des phénomènes observés et d'étudier l'impact des différents paramètres sur la variance du temps de polymérisation.

Les méthodes probabilistes présentées utilisent plusieurs types de techniques

- Calcul stochastique pour les processus ponctuels de Poisson marqués
- Théorèmes limites pour les processus de sauts markoviens.
- Méthodes d'homogénéisation.

qui seront rappelées lors du cours.

1. Introduction

- Introduction au calcul stochastique pour les processus ponctuels de Poisson marqués. Rappels sur les martingales associées aux processus markoviens de sauts.
- Convergence en distribution des processus de sauts markoviens. Homogénéisation des processus de Markov.
- Modèles probabilistes des phénomènes chimiques. Loi d'action de masse, équations de Michaelis-Menten.

2. Expression du Gène.

- Modèles markoviens et non-markoviens de la production de protéines. Existence et caractérisation de la loi invariante de la concentration d'une protéine d'un type donné. Etude de la variance à l'équilibre.
- Compétition pour les ressources de la cellule dans la production de protéines: un modèle de champs moyen.
- Etude de l'impact de l'auto-régulation de la production de protéines sur la variabilité du nombre de protéines: méthodes d'homogénéisation.

3. Modèles de la Polymérisation.

- Un modèle simplifié de la polymérisation avec deux espaces de polymères. Théorèmes central-limite fonctionnels.
- Variations sur les renormalisations des taux de polymérisation.
- Impact des phénomènes de nucléation.

References

- [1] David F. Anderson and Thomas G. Kurtz, *Stochastic analysis of biochemical systems*, Mathematical Biosciences Institute Lecture Series. Stochastics in Biological Systems, vol. 1, Springer, Cham; MBI Mathematical Biosciences Institute, Ohio State University, Columbus, OH, 2015.
- [2] Marie Doumic, Sarah Eugène, and Philippe Robert, *Asymptotics of stochastic protein assembly models*, SIAM Journal on Applied Mathematics **76** (2016), no. 6, 2333–2352.
- [3] Vincent Fromion, Emanuele Leoncini, and Philippe Robert, *Stochastic gene expression in cells: A point process approach*, SIAM Journal on Applied Mathematics **73** (2013), no. 1, 195–211.
- [4] T.G. Kurtz, *Averaging for martingale problems and stochastic approximation*, Applied Stochastic Analysis, US-French Workshop, Lecture notes in Control and Information sciences, vol. 177, Springer Verlag, 1992, pp. 186–209.
- [5] Nicolaas Godfried Van Kampen, *Stochastic processes in physics and chemistry*, vol. 1, Elsevier, 1992.

PARTIE II : **A. Ben-Hamou** Inégalités de concentration

En probabilités comme en statistiques, on est souvent amené à étudier les déviations d'une variable aléatoire par rapport à son espérance. Alors que le théorème central limite nous renseigne sur les fluctuations asymptotiques, les inégalités de concentration fournissent des résultats non-asymptotiques (à n fixé). Les inégalités exponentielles classiques, comme l'inégalité de Hoeffding, concernent les sommes de variables indépendantes. Dans ce cours, nous verrons que le phénomène de concentration de la mesure apparaît aussi pour des fonctions bien plus complexes que la somme: "une variable qui dépend (de façon lisse) de beaucoup de variables indépendantes (mais pas trop de chacune d'entre elles) est essentiellement constante" (Michel Talagrand).

Parmi les méthodes et résultats abordés dans ce cours, citons par exemple les inégalités de Poincaré et de Sobolev, la méthode entropique, la méthode de transport, le lien entre concentration et isopérimétrie. Nous nous intéresserons aussi au cas de variables dépendantes et verrons notamment comment la méthode de Stein permet d'obtenir des résultats de concentration dans des cadres non-indépendants.

La théorie de la concentration trouve des applications dans de nombreux domaines, et le cours sera illustré par beaucoup d'exemples issus de la physique statistique, mais aussi d'autres contextes comme l'apprentissage statistique, les matrices et graphes aléatoires, le mélange de chaînes de Markov, la théorie de l'information.