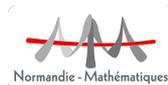


Journées MAS 2022

Rouen, 29-31 août



Modélisation dynamique et stochastique



L'image de couverture de ce livret, qui reprend celle de l'affiche des journées MAS 2022, est basée sur une huile de Charles Frechon « Rouen, île Lacroix, cours la Reine » (1890), Musée des Beaux Arts de Rouen.

La texture à travers laquelle on voit la peinture est obtenue par simulation d'un double processus de percolation de site orientée, dont le paramètre évolue avec la distance à l'axe des abscisses.

Bienvenue aux Journées MAS 2022

C'est avec un grand plaisir que nous vous accueillons à ces journées MAS 2022 qui ont lieu pour la première fois à Rouen. Voilà près d'un an et demi que nous préparons cet événement dans l'espoir de recréer ce qui a fait la réussite et le charme de ces journées lors des précédentes éditions, depuis 1996 : un programme scientifique qui témoigne de toute la richesse de la recherche récente dans le domaine de l'aléatoire, une place de choix faite aux jeunes avec notamment la remise des prix Neveu, la rencontre de plusieurs communautés en probabilités et statistique, issues des milieux académique ou industriel et surtout, l'occasion de se retrouver en présentiel dans une ambiance détendue et conviviale au retour de vacances.

Nous tenons à remercier chaleureusement ceux et celles qui nous ont permis de constituer un aussi beau programme scientifique : les membres du comité de liaison du groupe MAS de la SMAI présidé par Clément Marteau et ceux du comité scientifique présidé par Patricia Reynaud-Bouret. Un merci tout particulier à Clément et Patricia pour leur investissement depuis le début !

De nombreux acteurs nous ont apporté leur soutien financier, leur aide logistique ou leur parrainage, nous sommes extrêmement reconnaissants et fiers de cette marque de confiance : l'Agence pour les Mathématiques en Interaction avec l'Entreprise et la Société, l'Agence Nationale de la Recherche, le Centre National de la Recherche Scientifique, la Fédération Normandie Mathématiques, l'Institut Universitaire de France, le Labex Modèles Mathématiques et Économiques de la Dynamique, de l'Incertitude et des Interactions, le Laboratoire de Mathématiques Raphaël Salem, la Métropole de Rouen Normandie, la Région Normandie, la Société de Mathématiques Appliquées et Industrielles, l'Université de Rouen Normandie, l'U.F.R. de Droit, Sciences Économiques et Gestion et l'U.F.R. des Sciences et Techniques.

Enfin, un grand merci à toutes les personnes qui ont accepté d'organiser des sessions, tous les orateurs et oratrices, ceux et celles qui vont présenter des posters et tous les participants. Très bonnes journées MAS à tous !

Le comité d'organisation des journées MAS 2022

Comité scientifique

Présidente : Patricia Reynaud-Bouret (Université Côte d'Azur)

Membres :

- Thierry Lévy (Sorbonne Université)
- Sylvie Méléard (Polytechnique et IUF)
- Serge Pergamenchtkov (Université de Rouen Normandie)
- Anthony Réveillac (INSA Toulouse)
- Adeline Samson (Université Grenoble Alpes)
- Gwladys Toulemonde (Université de Montpellier)
- Romain Veltz (INRIA Sophia-Antipolis)

Comité d'organisation (LMRS, Université de Rouen Normandie)

Président : Pierre Calka

Membres :

- Vlad Barbu
- Jean-Baptiste Bardet
- Mohamed Ben-Alaya
- Jean-Yves Brua
- Gaëlle Chagny
- Antoine Channarond
- Mohamed El Machkouri
- Serge Pergamenchtkov
- Paul Raynaud de Fitte
- Thierry de la Rue
- Nicolas Vergne

Gestionnaires : Edwige Auvray, Sandrine Halé et Hamed Smail (LMRS), Noura Sahtout (SMAI)

Avec la participation du groupe MAS de la SMAI, dont le responsable est Clément Marteau.

Programme

Lundi 29 août

- 9.45-10.00** Ouverture des journées MAS 2022 - Amphi 600
- 10.00-11.00** Conférence plénière - Christophe Garban - Amphi 600
- 11.00-11.30** Prix Neveu - Barbara Dembin - Amphi 600
- 11.30-12.00** Prix Neveu - Jaouad Mourtada - Amphi 600
- 12.00-13.30** Déjeuner
- 13.30-14.00** Prix Neveu - Armand Riera - Amphi 600
- 14.05-16.00** Sessions parallèles
- Algorithmes stochastiques (Pierre Monmarché) - N103
 - Données fonctionnelles (Émilie Devijver) - N101
 - Maths - Entreprises (Nicolas Wicker) - C100
 - Matrices aléatoires (Marwa Banna) - N102
- 16.00-16.30** Pause - Café
- 16.30-18.25** Sessions parallèles
- Fiabilité et incertitudes (Mitra Fouladirad) - N101
 - Marches aléatoires (Kilian Raschel) - N103
 - Méthodes à noyaux, expérimentations informatiques (Luc Pronzato) - C100
 - Méthodes de pénalisation (Angelina Roche) - N102
- 19.00** Cocktail de bienvenue - Mairie de Rouen

Mardi 30 août

- 8.45-9.45** Conférence plénière - Gérard Biau - Amphi 600
- 9.45-10.15** Pause - Café
- 10.15-12.10** Sessions parallèles
- Statistique adaptative, bandits (Sébastien Gerchinovitz) - N101
 - Détection de ruptures (Farida Enikeeva et Nicolas Verzelen) - N102
 - Systèmes dynamiques aléatoires et graphes aléatoires (Jérôme Casse) - N103
 - Théorèmes limites (Olivier Durieu) - C100
- 12.10-13.30** Déjeuner
- 13.30-14.30** Conférence plénière - Vincent Bansaye - Amphi 600
- 14.35-16.30** Sessions parallèles
- Géométrie Stochastique (Raphaël Lachieze-Rey) - N101
 - Statistique des processus (Charlotte Dion et Céline Duval) - N103
 - Systèmes de particules (Oriane Blondel) - N102
 - Transport optimal en statistique (Claire Bréchet) - C100
- 16.30-17.00** Pause - Café - Session posters
- 17.00-18.00** Conférence plénière - Jean-François Coeurjolly - Amphi 600
- 18.00-19.00** Assemblée générale du groupe SMAI-MAS - Amphi 600
- 19.30** Dîner de la conférence - Hôtel de Bourgtheroulde

Mercredi 31 août

9.00-10.00 Conférence plénière - Irène Marcovici - Amphi 600

10.00-10.30 Pause - Café

10.30-12.25 Sessions parallèles

- Intelligence artificielle (Mohamed Heberi) - C100
- Optimisation en finance (Emmanuel Lépinette) - N101
- Percolation (Jean-Baptiste Gouéré) - N102
- Trajectoires rugueuses (Rémi Cattelier) - N103

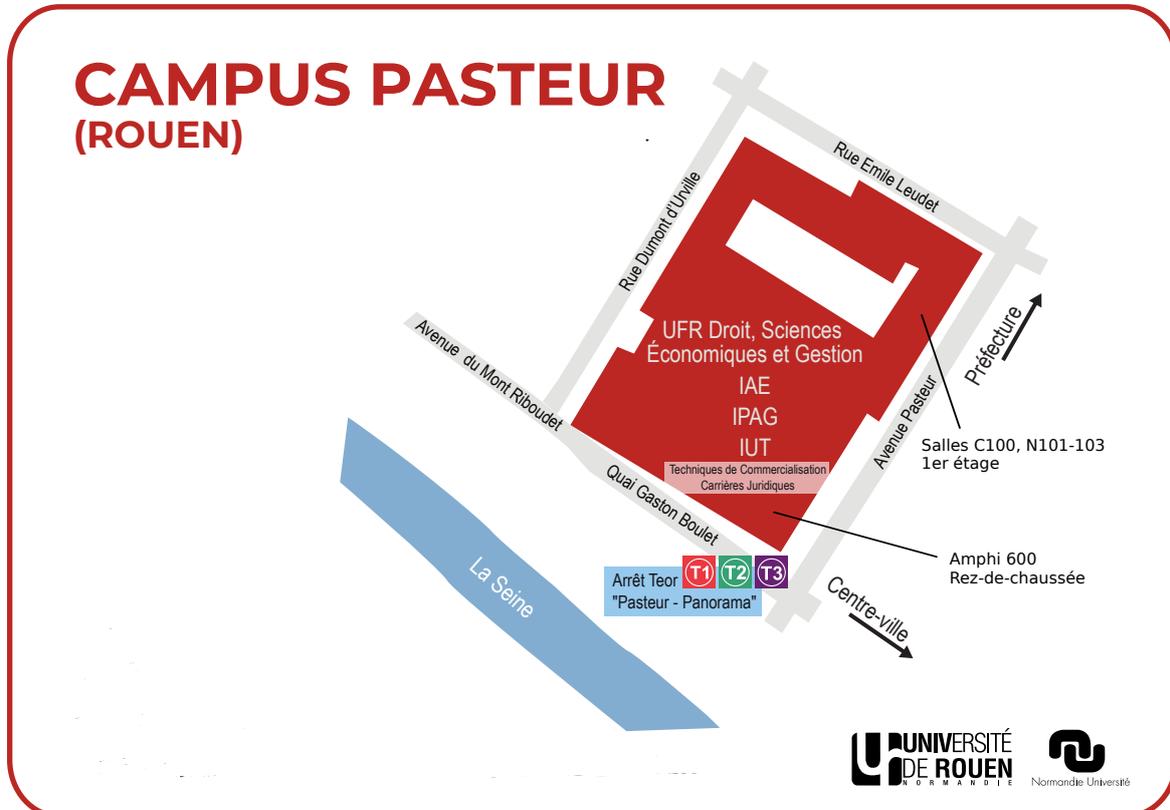
12.25-14.00 Déjeuner

14.00-15.00 Conférence plénière - Mathieu Rosenbaum - Amphi 600

15.00 Clôture

Plan du site des journées

Les journées se déroulent 3 avenue Pasteur à Rouen, sur le campus de l'UFR Droit, Sciences Économiques et Gestion. Les sessions plénières auront lieu en amphi 600, et les sessions parallèles dans les salles C100, N101, N102 et N103.



Résumés

Conférences plénières	1
Vincent Bansaye	1
Gérard Biau	1
Jean-François Coeurjolly	2
Christophe Garban	2
Irène Marcovici	2
Mathieu Rosenbaum	3
Prix Neveu	3
Barbara Dembin	3
Jaouad Mourtada	3
Armand Riera	3
Sessions parallèles – Lundi 29 août	4
Algorithmes stochastiques	4
Données fonctionnelles	5
Maths - Entreprises	7
Matrices aléatoires	8
Fiabilité et incertitudes	9
Marches aléatoires	11
Méthodes à noyaux, expérimentations informatiques	12
Méthodes de pénalisation	13
Sessions parallèles – Mardi 30 août	15
Statistique adaptative, bandits	15
Détection de ruptures	16
Systèmes dynamiques aléatoires et graphes aléatoires	18
Théorèmes limites	19
Géométrie stochastique	20
Statistique des processus	22
Systèmes de particules	23
Transport optimal en statistique	24
Sessions parallèles – Mercredi 31 août	25
Intelligence artificielle	25
Optimisation en finance	26
Percolation	28
Trajectoires rugueuses	29

Au sujet des sessions parallèles

Chaque session parallèle est composée d'un exposé introductif d'une durée de 40 minutes, puis de trois exposés d'une durée de 25 minutes chacun — questions comprises — hormis la session *Systèmes dynamiques aléatoires et graphes aléatoires*, composée de quatre exposés de durée égale, de 30 minutes chacun.

Dans la liste des résumés des sessions parallèles (pages 4 à 30), le premier résumé de chaque session est celui de l'exposé introductif. Sauf exception, l'ordre des résumés correspond ensuite à l'ordre des exposés de la session.

Conférences plénières

Processus de branchement et structure des contacts pour des modèles aléatoires en épidémiologie

Vincent Bansaye (CMAP, École Polytechnique)

Les processus de branchement apparaissent dans de nombreux modèles de dynamique des populations, notamment pour décrire des invasions. En particulier, ils interviennent dans la description des épidémies pour déterminer si le nombre d'infectés va exploser ou non et si oui à quelle vitesse. Dans ces modèles, la description des contacts joue un rôle important. Après une introduction sur ces problématiques, nous nous focaliserons sur un modèle incluant le traçage des contacts. Dans ce modèle, les individus infectent en population mélangée à taux fixe et l'information sur le contact infectieux est perdu à taux fixe, tandis que le test d'un individu infecté aboutit à l'isolement de la composante connexe associée aux contacts encore connus. Grâce à une propriété d'« éclatement » des arbres récursifs uniformes, nous pourrions réduire le modèle à un processus de croissance fragmentation isolation sur les tailles des composantes. Nous présenterons et exploiterons alors des techniques récentes d'analyse quantitative des semi groupes non conservatifs. Cela permettra d'obtenir des convergences fortes pour décrire la propagation de l'épidémie. Enfin, nous évoquerons des extensions de ce modèle et la prise en compte d'une structuration spatiale des contacts via de grands graphes aléatoires spatialisés, impliquant des techniques d'homogénéisation stochastique.

Ces travaux sont respectivement des collaborations avec Chenlin Gu et Linglong Yuan, Michele Salvi et Elisabeta Vergu.

Limite d'échelle pour les réseaux neuronaux résiduels profonds

Gérard Biau (LPSM, Sorbonne Université)

Les réseaux neuronaux résiduels (ResNets) profonds sont reconnus pour les résultats exceptionnels qu'ils obtiennent dans des tâches d'apprentissage automatique complexes. Les performances remarquables de ces architectures reposent cependant sur une procédure d'apprentissage délicate, qui doit être soigneusement mise en oeuvre afin d'éviter que les gradients ne s'annulent ou n'exploient, en particulier lorsque la profondeur L augmente. Il n'existe pas, à ce jour, de consensus clair sur la façon de régler cette difficulté, bien qu'une stratégie classique consiste à multiplier la sortie de chaque couche par un facteur d'échelle α_L . Nous montrerons dans cet exposé qu'avec des initialisations i.i.d. standards, la seule dynamique intéressante se produit pour $\alpha_L = 1/\sqrt{L}$, tous les autres choix conduisant à des comportements moins pertinents pour l'apprentissage. Ce facteur d'échelle correspond dans la limite temps continu à une équation différentielle stochastique, contrairement à une interprétation répandue selon laquelle les ResNets profonds sont des discrétisations d'équations différentielles ordinaires (dites neuronales). Pour ce dernier régime, la stabilité est obtenue avec des initialisations corrélées bien spécifiques et $\alpha_L = 1/L$. Au final, notre analyse suggère une interaction forte entre le facteur d'échelle α_L et la régularité des poids en fonction de l'indice des couches.

En collaboration avec Adeline Fermanian (Mines ParisTech), Pierre Marion (Sorbonne Université) et Jean-Philippe Vert (Google Research).

Régularisation et processus ponctuels spatiaux

Jean-François Coeurjolly (LJK, Université Grenoble Alpes)

L'analyse de processus ponctuels commence en général par l'estimation de la fonction d'intensité ou d'intensité conditionnelle (de Papangelou). Les méthodes paramétriques usuelles reposent sur des vraisemblances composites, quasi-vraisemblance, pseudo-vraisemblance ou régression logistique conditionnelle par exemple. La mise en place et l'étude de ces méthodes constituent des problèmes largement traités dans la littérature. Ceux-ci peuvent être considérés comme résolus tant d'un point de vue pratique que théorique pour des processus de points (inhomogènes) de "taille" modérée. Ces dernières années ont vu l'émergence de processus de points de grande dimension dans diverses applications. Par grande dimension, nous entendons dans cet exposé, observer une seule réalisation d'un motif de points dont on souhaite relier l'arrangement avec un très grand nombre de covariables spatiales. L'inférence devient alors problématique et il apparaît souhaitable et naturel de pénaliser les méthodes citées précédemment pour estimer les paramètres. Je discuterai comment ces idées sont appliquées dans le cadre de l'estimation d'intensité ou intensité conditionnelle, comment ces méthodes peuvent être implémentées et quels sont les types de résultats que nous pouvons obtenir.

(Travaux joints en partie avec I. Ba, A. Choiruddin, F. Letué, F. Cuevas Pacheco, M.-H. Descary and R. Waagepetersen)

Brisure de symétrie continue et problème de synchronisation

Christophe Garban (ICJ, Université Lyon 1)

Dans cet exposé, je commencerai par introduire des systèmes de spins sur Z^d dont la symétrie sous-jacente est continue (par opposition au célèbre modèle d'Ising dont la symétrie $\sigma \rightarrow -\sigma$ est discrète). Le but de l'exposé sera d'expliquer un lien surprenant qui est apparu récemment entre ces modèles de physique statistique et des questions de statistique bayésienne / reconstruction statistique. J'introduirai une nouvelle technique pour identifier un ordre à longue portée dans ces systèmes de spins à basse température ("brisure de symétrie") qui est basée sur le concept de "group synchronization" et repose en particulier sur un travail récent d'Abbe, Massoulié, Montanari, Sly and Srivastava (2018).

L'exposé ne nécessitera aucun pré-requis en physique statistique.

Travail en commun avec Thomas Spencer (IAS, Princeton).

Automates cellulaires et phénomènes d'auto-organisation

Irène Marcovici (IECL, Université de Lorraine)

Les automates cellulaires sont des systèmes dynamiques pour lesquels le temps et l'espace sont discrets. Ils permettent de modéliser l'évolution d'un ensemble de composantes interagissant entre elles de manière locale : au cours du temps, chacune actualise son état en fonction de ce qu'elle perçoit de son voisinage.

En étudiant certains automates cellulaires, on peut observer des phénomènes d'auto-organisation : à partir d'un état initial désordonné, les mises à jour successives des cellules par la règle locale conduisent à l'apparition d'une structure macroscopique.

À l'inverse, si l'on souhaite parvenir à un certain comportement global, on peut chercher à concevoir une règle locale permettant de l'atteindre de manière décentralisée. J'exposerai différents problèmes de ce type (obtention de consensus, synchronisation, correction d'erreurs, diagnostic de défaillances dans un réseau...), en étudiant l'influence que peut avoir l'introduction d'aléa dans les dynamiques.

La présentation reposera sur différents travaux effectués en collaboration avec Nazim Fatès, Régine Marchand, Mathieu Sablik et Siamak Taati.

Rough volatility : quand les outils de dynamique des populations permettent de comprendre les fluctuations financières

Mathieu Rosenbaum (CMAP, École Polytechnique)

Dans cet exposé, nous présenterons les éléments statistiques et probabilistes ayant conduit à l'introduction du paradigme "rough volatility" pour la modélisation des actifs financiers et la gestion de leurs risques. Nous nous intéresserons notamment à l'universalité de cette approche que nous pourrions expliquer grâce à l'utilisation d'outils probabilistes issus des modèles de dynamique des populations.

Prix Neveu

Percolation et percolation de premier passage

Barbara Dembin (ETH Zürich)

Le modèle de percolation sur le graphe $(\mathbb{Z}^d, \mathbb{E}^d)$ se définit comme suit. On conserve chaque arête indépendamment avec probabilité p . La question centrale est l'existence d'une composante connexe infinie selon la valeur de p . Il existe une généralisation de ce modèle appelée percolation de premier passage. Considérons G une distribution sur \mathbb{R}_+ . Pour chaque arête e , on tire indépendamment une variable aléatoire t_e de loi G . Cette variable t_e peut s'interpréter de deux façons différentes. On peut la voir comme le temps pour traverser l'arête e ou comme une capacité, i.e., le débit maximal d'eau pouvant traverser l'arête. Chaque interprétation amène à des questions différentes. Nous présenterons deux résultats obtenus pendant cette thèse qui illustrent comment des techniques de percolation de premier passage permettent d'obtenir de nouveaux résultats sur le modèle de percolation.

Quelques questions statistiques autour de la prédiction linéaire

Jaouad Mourtada (ENSAÉ, CREST)

En apprentissage statistique, l'objectif consiste à trouver une bonne fonction de prédiction d'une variable cible Y à partir de variables prédictives X , étant donné un jeu de données constitué de plusieurs variables aléatoires indépendantes de même loi que la paire (X, Y) . Dans cet exposé, nous discuterons de deux variantes classiques de ce problème faisant intervenir des prédicteurs linéaires, à savoir la régression linéaire (pour une cible Y numérique) et la régression logistique (pour une cible binaire). On s'intéressera aux garanties qu'il est possible d'obtenir sous des hypothèses aussi faibles que possible sur la loi des variables prédictives X et sur la loi conditionnelle de Y sachant X . Nous discuterons notamment de la nature de procédures atteignant des garanties optimales, et de leur caractère calculable (ou non). L'exposé présentera quelques résultats dans cette direction ainsi que certaines questions ouvertes.

Géométrie brownienne

Armand Riera (LMO, Université Paris Saclay)

La géométrie brownienne est une famille de surfaces aléatoires planaires. Ces objets peuvent être obtenus comme limite d'échelle de modèles de grands graphes aléatoires dessinés sur différentes surfaces, et peuvent être classés selon leur topologie : sphère, disque, plan, demi-plan... De manière informelle, la géométrie brownienne permet de définir une notion de surface uniforme sous certaines contraintes topologiques.

Le but de cet exposé est de présenter de manière imagée et sans pré-requis cette théorie, mais aussi d'expliquer comment explorer métriquement ces surfaces et les propriétés Markoviennes qui en découlent. Si le temps le permet, nous aborderons enfin la manière donc ces explorations permettent d'établir un certain nombre de propriétés de nature géométrique.

Sessions parallèles – Lundi 29 août

Algorithmes stochastiques

Organisation: Pierre Monmarché — Horaire: 14.05 — Salle: N103

Bounding the error of discretized Langevin algorithms for non-strongly log-concave targets

Arnak Dalalyan (CREST, ENSAÉ)

(joint with Lionel Riou-Durand and Avetik Karagulyan)

In this talk, we will provide non-asymptotic upper bounds on the error of sampling from a target density using three schemes of discretized Langevin diffusions. The first scheme is the Langevin Monte Carlo (LMC) algorithm, the Euler discretization of the Langevin diffusion. The second and the third schemes are, respectively, the kinetic Langevin Monte Carlo (KLMC) for differentiable potentials and the kinetic Langevin Monte Carlo for twice-differentiable potentials (KLMC2). The main focus is on the target densities that are smooth and log-concave on \mathbb{R}^p , but not necessarily strongly log-concave. Bounds on the computational complexity are obtained under two types of smoothness assumption : the potential has a Lipschitz-continuous gradient and the potential has a Lipschitz-continuous Hessian matrix. The error of sampling is measured by Wasserstein- q distances. We advocate for the use of a new dimension-adapted scaling in the definition of the computational complexity, when Wasserstein- q distances are considered. The obtained results show that the number of iterations to achieve a scaled-error smaller than a prescribed value depends only polynomially in the dimension.

Removing the mini-batching error in Bayesian inference using Adaptive Langevin dynamics

Inass Sekkat (CERMICS, École des Ponts)

Bayesian inference allows to obtain useful information on the parameters of models, either in computational statistics or more recently in the context of Bayesian Neural Networks. The computational cost of usual Monte Carlo methods for sampling a posteriori laws in Bayesian inference scales linearly with the number of data points. One option to reduce it to a fraction of this cost is to resort to mini-batching in conjunction with unadjusted discretizations of Langevin dynamics, in which case only a random fraction of the data is used to estimate the gradient. However, this leads to an additional noise in the dynamics and hence a bias on the invariant measure which is sampled by the Markov chain. We advocate using the so-called Adaptive Langevin dynamics, which is a modification of standard inertial Langevin dynamics with a dynamical friction which automatically corrects for the increased noise arising from mini-batching. We investigate the practical relevance of the assumptions underpinning Adaptive Langevin (constant covariance for the estimation of the

gradient), which are not satisfied in typical models of Bayesian inference, and quantify the bias induced by minibatching in this case. We also show how to extend AdL in order to systematically reduce the bias on the posterior distribution by considering a dynamical friction depending on the current value of the parameter to sample.

Multilevel-Langevin pathwise average for average for Gibbs approximation

Maxime Egéa (LAREMA, Université d'Angers)

In this talk, I will introduce a new multilevel-Langevin method dedicated to the Gibbs sampling. More precisely, this procedure both inspired by [Szpruch et al., 2016] and [Pagès-Panloup, 2018] is based on a multilevel combination of pathwise averages of discretized schemes of overdamped-Langevin diffusions. In a general setting, we first show that for any positive ε , an appropriate choice of the parameters (including steps/number of layers/length of the path, ...) leads to an ε -approximation with a cost proportional to ε^{-2} , i.e. proportional to a Monte-Carlo method without bias. In a second part, we investigate, in the uniformly convex setting, the dependence in the dimension and optimize the choice of the parameters of the multilevel strategy in terms of d . We show that it is possible to obtain a complexity proportional to $d\varepsilon^{-2}$. Finally, we will discuss some recent developments in the weakly convex case. The first part of the talk is based on a joint work with F. Panloup.

Convergence of the kinetic annealing for general potentials

Lucas Journal (LPSM, Sorbonne Université)

The goal of the simulated annealing algorithm is to find, via a stochastic process, the minimum of some function $U : \mathbb{R}^d \mapsto \mathbb{R}_+$. To this end we study the kinetic process :

$$\begin{cases} dX_t = Y_t dt \\ dY_t = -\nabla U(X_t) dt - \gamma_t Y_t dt + \sqrt{2\gamma_t \beta_t^{-1}} dB_t, \end{cases}$$

where $\beta_t = \frac{\ln(e^{c\beta_0} + t)}{c}$ and γ_t is some friction parameter. Let c^* be the largest energy barrier of U . We prove under mild assumptions on the potential U the convergence of the process towards the global minimum of U for $c > c^*$ (slow cooling schedules), as well as the non-convergence for $c < c^*$ (fast cooling schedules).

Données fonctionnelles

Organisation: Émilie Devijver — Horaire: 14.05 — Salle: N101

Introduction aux données fonctionnelles

Jairo Cugliari (Labo. ERIC, Université Lumière Lyon 2)

Dans cette introduction nous présenterons un état de l'art pédagogique et sélectif, l'objectif étant de permettre que les non-spécialistes puissent s'initier au domaine de l'analyse de données fonctionnelles. Nous évoquerons les fondements théoriques ainsi que d'outils informatiques pour la mise en œuvre des techniques sur de cas pratiques dans le contexte de données indépendantes et dépendantes. Pour finir, nous discuterons deux cas d'étude pour illustrer la contribution de l'approche fonctionnelle dans l'analyse de données.

Estimation adaptative dans le modèle linéaire fonctionnel à sortie fonctionnelle

Anouar Meynaoui (LIG, Université Grenoble Alpes)

Dans cet exposé, on considère un modèle de régression linéaire fonctionnel à sortie fonctionnelle : la variable explicative et la variable réponse sont des variables aléatoires "fonctionnelles", à valeurs dans un espace de Hilbert (typiquement un espace de fonctions). On s'intéresse à la question de l'estimation non-paramétrique adaptative de l'opérateur intégral reliant ces deux variables, à partir d'un échantillon. Une collection d'estimateurs par projection est d'abord construite : lorsque la base de projection choisie est la base de l'ACP empirique de la covariable, on obtient une décomposition biais-variance pour un risque quadratique moyen de prédiction. Une procédure de sélection de modèle (minimisation de contraste pénalisé) permet ensuite un choix automatique du meilleur estimateur dans la collection. Celui-ci satisfait une inégalité de type oracle, et atteint des vitesses de convergences minimax sur des espaces de régularité de type ellipsoïde : la borne supérieure du risque de prédiction correspond bien à la borne inférieure, que nous calculons également. Ces résultats théoriques sont illustrés par des simulations et des applications à des jeux de données réelles.

Il s'agit d'un travail en collaboration avec Gaëlle Chagny (LMRS, Université de Rouen) et Angelina Roche (Ceremade, Université Paris Dauphine).

Barycentres de données fonctionnelles : une nouvelle approche basée sur la méthode de la signature

Raphaël Mignot (IECL, Université de Lorraine)

La méthode de la signature a été largement utilisée pour l'analyse des séries temporelles multivariées. Cette approche a prouvé son efficacité pour de nombreuses applications en apprentissage statistique. La définition d'une notion de barycentre dans l'espace des signatures est un premier pas prometteur permettant de développer de nouvelles extensions de l'analyse en composantes principales ou de l'algorithme des k-moyennes aux données fonctionnelles.

High-dimensional nonparametric functional graphical models via the additive partial correlation operator

Eftychia Solea (CREST, ENSAI)

This article develops a novel approach for estimating a high-dimensional and nonparametric graphical model for functional data. Our approach is built on a new linear operator, the functional additive partial correlation operator, which extends the partial correlation matrix to both the nonparametric and functional setting. We show that its nonzero elements can be used to characterize the graph, and we employ sparse regression techniques for graph estimation. Moreover, the method does not rely on any distributional assumptions and does not require the computation of multi-dimensional kernels, thus avoiding the curse of dimensionality. We establish both estimation consistency and graph selection consistency of the proposed estimator, while allowing the number of nodes to grow with the increasing sample size. Through simulation studies, we demonstrate that our method performs better than existing methods in cases where the Gaussian or Gaussian copula assumption does not hold. We also demonstrate the performance of the proposed method by a study of an electroencephalography dataset to construct a brain network.

Maths - Entreprises

Organisation: Nicolas Wicker — Horaire: 14.05 — Salle: C100

Présentation de l'AMIES

Nicolas Wicker (LPP, Université de Lille)

L'AMIES (Agence pour les Mathématiques en Interaction avec les Entreprises et la Société) a été créée en 2011 et depuis cette date oeuvre pour favoriser les collaborations entre les laboratoires de mathématiques et les entreprises. Elle agit principalement en finançant des collaborations public/privé mais également en organisant des semaines maths-entreprise pour les doctorants ainsi que le Forum Emploi Math.

Construction of a surrogate model : multivariate time series prediction with random forests

Clara Carlier (CREST, ENSAÉ & Renault Group)

(joint with Arnaud Franju, Matthieu Lerasle and Mathias Obrebski)

In a context of evolution and expansion of advanced driver- assistance system, it is essential to validate and certify all these embedded technologies. Vehicles now have a large number of sensors, which gives us access to various information during testing. The number of tests to be carried out for the validation of the systems is constantly increasing and forces the automotive groups to develop methods of numerical simulations, to supplement and possibly replace certain tests currently carried out on track, for budget and time economy. Our main objective is to calibrate the simulator by comparing the simulations to the on-track experiments. In order to achieve this step, we simulate a large amount of data repeatedly in an iterative algorithm. This is why we are trying to develop a surrogate model that mimics the simulator as precisely as possible. This surrogate model, which is less expensive in terms of computation time, will allow us to predict the behavior of the car for given input parameters. We will see how to use random forests to build the most accurate model possible.

Keywords. time series prediction, surrogate model, generative model, random forest

Quelques méthodes d'apprentissage automatique pour la Maintenance de Systèmes Avion

Corentin Duprey & Anne-Françoise Yao (LMBP, Université Clermont Auvergne & Hop Air-France)

(collaboration avec Julie Galant et Nouridine Azzaoui)

Ce travail, s'inscrit dans le cadre des travaux de recherche, collaboration entre le laboratoire LMBP et l'entreprise HOP Air France filiale, dont l'objectif est de construire des modèles et algorithmes permettant l'amélioration des méthodes de maintenance et de traitement des pannes sur avion. Au cours de cet exposé, nous aborderons deux problématiques : celles liées au FCND d'une part et celles liées à un système avion en particulier d'autre part.

Le Flight Control No Dispatch (FCND) est une panne complexe récurrente sur les avions E190 et E170 de la marque Embraer. Une panne complexe est un aléa dont les causes peuvent être multiples.

Une mauvaise identification du système défaillant peut avoir des conséquences sur l'exploitation d'une compagnie aérienne : retards ou annulations. Notre objectif est de proposer aux équipes de maintenance, un outil d'aide à la décision ciblant le système mis en cause lors d'une panne complexe, accélérant ainsi les prises de décision et donc une meilleure anticipation dans la gestion des pannes. Nous utilisons des approches d'apprentissage supervisé et non-supervisé pour répondre à la problématique. Outre la procédure mise en place comme solution à la problématique du FCND, nous aborderons également celle concernant l'anticipation de panne d'un système avion en particulier.

Mots-clés. Détection de pannes, clustering, classification supervisée, Text mining, détection de rupture.

Recherche d'information dans la littérature médicale : l'apport des plongements de graphes

Marianne Clausel (IECL, Université de Lorraine)

Collaboration avec M. Constant (ATILF, Université de Lorraine) et P. Oudet (Cancéropôle Est).

Cet exposé présentera une collaboration entre le Cancéropôle Est et l'Université de Lorraine, qui a été initiée grâce à un PEPS2 AMIES. L'objectif était de développer des modèles permettant d'extraire un graphe de connaissances à partir des textes intégraux de publications scientifiques autour du cancer du sein. Nous avons pour cela utilisé des plongements de graphes et développé une interface permettant une exploration interactive de la littérature scientifique.

Ce travail préliminaire, a été prolongé par une thèse financée par l'Institut de Cancérologie Strasbourg Europe (ICANS) qui a démarré en Février 2022.

Bibliographie

Perozzi B, Al-Rfou R, Skiena S. Deepwalk : Online learning of social representations. In Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining 2014 Aug 24 (pp. 701-710).

Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Ł, Polosukhin I. Attention is all you need. Advances in neural information processing systems. 2017 ;30.

Matrices aléatoires

Organisation: Marwa Banna — Horaire: 14.05 — Salle: N102

Freeness over the diagonal and the spectrum of the sum of random matrices

Camille Male (IMB, Université de Bordeaux)

In this talk, I present a method to compute the spectrum of the sum of two large random matrices thanks to fixed points equations. I first review Pastur's equation and Voiculescu's subordination equations, that compute the free convolution of two probability measures and then can be applied for independent Wigner and unitarily invariant matrices. Then I introduce their generalization that has been discovered more recently to compute spectra in more general situations, for random matrices invariant in law by conjugation by permutation matrices. In collaboration with B. Au, G. Cébron, A. Dahlqvist and F. Gabriel.

La décomposition spectrale des matrices très creuses

Simon Coste (Centre Sciences des Données, ENS Paris)

Les valeurs propres de matrices aléatoires avec énormément d'entrées non nulles sont très peu comprises. Dans cet exposé, je montrerai quelques résultats partiels sur le spectre/polynôme caractéristique de ces matrices, dans deux modèles : modèles à entrées indépendantes, et sommes de matrices de permutations aléatoires (ce qui englobe le cas des graphes réguliers dirigés). Je mentionnerai aussi des questions (nées d'observations numériques) sur la localisation des entrées des vecteurs propres de ces matrices. L'exposé s'appuie sur une collaboration avec G. Lambert et Y. Zhu.

Landscape complexity of the finite-rank spiked tensor model

Vanessa Piccolo (UMPA, ENS Lyon)

In this talk, we focus on a class of disorder systems (called finite-rank spiked tensor model) characterized by the spherical pure p -spin Hamiltonian in the presence of a finite number of signals given by deterministic polynomials of degree p . It is known that these systems exhibit phase transitions in terms of the topology of their energy landscape. We study the average number of critical points and local minima of the landscape through the Kac-Rice formula, which allows one to obtain the complexity asymptotics of high-dimensional random landscapes via the determinants of large random matrices. We show that there is a threshold, below which the complexity is negative and the "good" critical points are exponentially rare, and above which the complexity vanishes and the critical points in this region have high correlations with the signals. This generalizes the work of Ben Arous-Mei-Montanari-Nica on the complexity of the rank-one spiked tensor model.

Conjecture de la variance dans les boules de Schatten

Benjamin Dadoun (Université NYU, Abu Dhabi)

Nous présentons des résultats en grande dimension sur la norme Hilbert-Schmidt d'une matrice uniformément distribuée dans la boule unité au sens de la p -norme des valeurs singulières. Nous nous appuyons sur la connexion avec l'analyse spectrale des β -ensembles en adaptant notamment certains résultats de fluctuation dus à Bekerman, Leblé et Serfaty. Lorsque $p > 3$, cela nous permet d'établir une version forte de la conjecture de la variance restreinte aux p -boules de Schatten auto-adjointes. Travail en commun avec Matthieu Fradelizi, Olivier Guédon et Pierre-André Zitt.

Fiabilité et incertitudes

Organisation: Mitra Fouladirad — Horaire: 16.30 — Salle: N101

Modélisation de dégradation par les processus stochastiques

Mitra Fouladirad (Labo. M2P2, École Centrale de Marseille)

Les grands systèmes tels que les ouvrages d'art, les systèmes de production d'énergie, se dégradent dans le temps dû à leur vieillissement, la sollicitation (fatigue), les conditions environnementales (corrosion), etc., La dégradation dépend des propriétés du systèmes ainsi que de nombreux facteurs externes. Afin de prédire la durée de vie et la période de bon fonctionnement du système, il est nécessaire de modéliser cette dégradation. La modélisation exacte de cette dégradation par

des modèles déterministes ne permet pas une prédiction qui prend en compte les différents facteurs environnementaux et les différents niveaux de sollicitations qui sont variables et souvent aléatoires. Les processus stochastiques sont des bons candidats pour cette modélisation. Dans cet exposé, à travers quelques exemples industriels, la modélisation stochastique de la dégradation est mise en exergue.

Stochastic Maintenance Optimization of Complex Industrial Systems under Uncertainty of Repair Time and Resources

Selma Khebbache (IRT System X, Paris)

The presentation focuses on maintenance optimization strategies in complex manufacturing systems under uncertainty of available resources and tasks duration constraints. We propose a deterministic formulation of the maintenance optimization problem and provide a chance constraints modeling to formulate uncertainty constraints. Our stochastic mathematical model is reformulated using approximations providing deterministic equivalent formulations that we solve in negligible times to reach near global solutions. Simulation results are provided to illustrate the efficiency of our mathematical models when highlighting the convergence of the chance constraints formulation to the deterministic form for given confidence levels.

Apprentissage et génération de séries temporelles pour la prédiction du productible éolien

Merlin Keller (EDF, Paris)

Au cours de la phase de projet d'une nouvelle ferme éolienne, un enjeu majeur réside dans l'évaluation de la ressource en vent à travers la zone où seront implantées les turbines, et sur la durée de vie contractuelle du projet, soit environ vingt ans. Cette évaluation permet de quantifier la rentabilité du projet, point clef pour obtenir son financement, et également d'en optimiser la conception (localisation des turbines et du câblage, choix technologiques...).

Cette évaluation, forcément entachée d'incertitudes, peut être vue comme un problème d'apprentissage statistique, pour laquelle on ne dispose que de données limitées. Celles-ci sont généralement issues d'un mat de mesures installé sur site et qui fournit des données précises mais sur une durée limitée (deux ou trois ans). Ces données peuvent être complétées par des prédictions de modèles météorologiques calés sur des données satellitaires, moins précises mais sur une plus grande période (de l'ordre de vingt ans).

Nous avons testé plusieurs méthodes pour estimer la loi jointe de ces séries temporelles multiples et partiellement observées : factorisation spectrale, bootstrap, modèles auto-régressifs. Nous comparons leurs avantages respectifs sur un cas d'usage précis issu d'une ferme éolienne expérimentale.

Prognostic expert system for railway fleet maintenance

Rémy Marion (SNCF Voyageurs, Saint-Pierre-Des-Corps)

To realize maintenance of a large rolling fleet, with operational constraints due to mass transit, a mixed maintenance solution based on real-time data analysis and condition-based maintenance has been integrated into the SNCF maintenance process in 2017. Based on a prognostic expert system, this solution relies on constant signalling thresholds defined using technical knowledge and physical models in order to assess the health state of a system. As the health of a system differs from one train to another and independently evolves in time, constant signalling thresholds does not always match maintenance load and infrastructure availabilities in the workshop. To overtake these limitations and enhance the current maintenance solution, an upgrade of the existing expert system is under development, using dynamic signalling threshold based on the distribution of health indicators across the whole fleet.

Marches aléatoires

Organisation: Kilian Raschel — Horaire: 16.30 — Salle: N103

Problèmes de persistance pour des marches aléatoires unidimensionnelles intégrées

Quentin Berger (LPSM, Sorbonne Université)

Je présenterai des problèmes de persistance (ou de survie) pour certains processus unidimensionnels non-Markoviens. Je m'intéresserai en particulier à des fonctionnelles additives de chaînes de Markov, comme l'intégrale (d'une fonction) de la chaîne de Markov. Le but est d'estimer la probabilité que le processus évite un demi-espace pendant un temps long. On verra comment, dans certains cas, on peut résoudre complètement le problème ; de nombreuses questions restent cependant encore ouvertes. Travail en collaboration avec Camille Tardif et Loïc Bettencourt.

Large fluctuations in multi-scale modeling for rest hematopoiesis

Céline Bonnet (UMPA, ENS Lyon)

Hematopoiesis is a biological phenomenon (process) of production of mature blood cells by cellular differentiation. It is based on amplification steps due to an interplay between renewal and differentiation in the successive cell types from stem cells to mature blood cells. We will study this mechanism with a stochastic point of view to explain unexpected fluctuations on the mature blood cell numbers, as surprisingly observed by biologists and medical doctors in rest hematopoiesis.

We consider three cell types : stem cells, progenitors, and mature blood cells. Each cell type is characterized by its own dynamics parameters : its division rate and the renewal and differentiation probabilities at each division event. We model the global population dynamics by a three-dimensional stochastic decomposable branching process. We show that the amplification mechanism is given by the inverse of the small difference between the differentiation and renewal probabilities. Introducing a parameter K which scales simultaneously the size of the first component, the differentiation and renewal probabilities and the mature blood cell death rate, we describe the asymptotic behavior of the process for large K .

We show that each cell type has its own size scale and its own time scale. Focusing on the third component, we prove that the mature blood cell population size, conveniently renormalized (in time and size), is expanded in a usual way inducing large fluctuations. The proofs are based on a fine study of the different scales involved in the model and on the use of different convergence and average techniques in the proofs.

La marche aléatoire de l'éléphant

Lucile Laulin (IMB, Université de Bordeaux)

La marche aléatoire d'éléphant (ERW) est une marche aléatoire discrète qui a été introduite au début des années 2000 par deux physiciens afin d'étudier l'influence d'un paramètre de mémoire sur le comportement de la marche aléatoire. Dans cet exposé, on présentera plusieurs possibilités pour étudier et obtenir des résultats sur l'ERW. En particulier, on s'intéressera à l'approche martingales, puis au lien avec les urnes de Polya ou encore avec les arbres aléatoires récurrents.

Marche aléatoire sur un arbre de Galton-Watson : des temps locaux à la dynamique

Loïc de Raphélis (IDP, Université d'Orléans)

Nous présenterons dans cet exposé un modèle de marche aléatoire récurrente sur un arbre de Galton-Watson. Notre objectif sera de présenter une méthode basée sur une observation statique, les temps locaux de la marche, pour en déduire des propriétés sur la dynamique de la marche. Nous présenterons des résultats obtenus par celle-ci :

- sur la fonction de hauteur de la marche dans l'arbre,
- sur l'ensemble des sites visités par la marche,
- sur les sites favoris de la marche, c'est-à-dire les sommets de l'arbre les plus visités.

En particulier, nous observerons que selon la loi de l'arbre deux régimes peuvent se mettre en place, le premier dans un cadre gaussien, et le second dans un cadre stable.

Méthodes à noyaux, expérimentations informatiques

Organisation: Luc Pronzato — Horaire: 16.30 — Salle: C100

Recent advances in kernel methods for computer experiments

Sébastien Da Veiga (Safran)

In the computer experiments community, kernel methods have now been used for several decades with great success on many industrial applications.

In particular, Gaussian process regression is commonly employed for building an approximation of complex numerical simulation models, as well as for designing efficient active learning strategies dedicated to optimization or uncertainty quantification problems.

Although kernel methods were thus mainly considered for regression tasks in this context, the recent introduction of the framework focusing on kernel embeddings of probability distributions offers a new perspective for computer experiments. More specifically, the Maximum Mean Discrepancy (MMD) for comparing probability distributions and the Hilbert-Schmidt Independence Criterion (HSIC) for testing independence are highly promising tools for our community.

In this talk, I will first briefly recap how kernel methods are used for regression, and then introduce the MMD and the HSIC. I will then discuss how they can be used both for proposing novel design of experiments and for giving a complementary view for global sensitivity analysis. Their potential, as well as even more recent advances based on the kernel Stein discrepancy, will be discussed and put on perspectives with respect to the current challenges in computer experiments.

Gaussian process regression on nested spaces

Thierry Gonon (ICJ, École Centrale Lyon)

As computer codes involve a very large number of variables, industrial experts build metamodels on restricted sets of variables, that are enlarged progressively to improve knowledge on the studied output. Several designs of experiment are generated, which belong to subspaces included in each other and of increasing dimension. The problematic is to create a metamodel, using all these designs, and adapted to their particular locations. An approach is introduced, called seqGPR (sequential Gaussian process regression), which uses Gaussian process regression. At each new step of the study (when new variables are released), the output is supposed to be the realization of the sum of two independent Gaussian processes. The first one models the output at the previous step.

The second one is a correction term which must be null on the subspace studied at the previous step, which constitutes a continuum of points. Firstly, some candidate Gaussian processes for the correction terms are suggested. Then, an EM (Expectation-Maximization) algorithm is implemented to estimate the parameters of the processes. Finally, the metamodel seqGPR is compared to a classic kriging metamodel on test cases and gives good results.

Model predictivity assessment : incremental test-set selection and accuracy evaluation

Joseph Muré (EDF R&D, Chatou)

Unbiased assessment of the predictivity of models learnt by supervised machine-learning methods requires knowledge of the learned function over a reserved test set (not used by the learning algorithm). The quality of the assessment depends, naturally, on the properties of the test set and on the error statistic used to estimate the prediction error. We tackle both issues, proposing a new predictivity criterion that carefully weights the individual observed errors to obtain a global error estimate, and using incremental experimental design methods to “optimally” select the test points on which the criterion is computed. Several incremental constructions are studied, including greedy-packing (coffee-house design), support points and kernel herding techniques. Our results show that the incremental and weighted versions of the latter two, based on Maximum Mean Discrepancy concepts, yield superior performance. An industrial test case provided by the historical French electricity supplier (EDF) illustrates the practical relevance of the methodology, indicating that it is an efficient alternative to expensive cross-validation techniques.

Determinantal point processes for coresets

Nicolas Tremblay (GIPSA-Lab, Université Grenoble Alpes)

When faced with a data set too large to be processed all at once, an obvious solution is to retain only part of it. In practice this takes a wide variety of different forms, and among them “coresets” are especially appealing. A coreset is a (small) weighted sample of the original data that comes with the following guarantee : a cost function can be evaluated on the smaller set instead of the larger one, with low relative error. For some classes of problems, and via a careful choice of sampling distribution (based on the so-called “sensitivity” metric), iid random sampling has turned to be one of the most successful methods for building coresets efficiently. However, independent samples are sometimes overly redundant, and one could hope that enforcing diversity would lead to better performance. The difficulty lies in proving coreset properties in non-iid samples. We show that the coreset property holds for samples formed with determinantal point processes (DPP). DPPs are interesting because they are a rare example of repulsive point processes with tractable theoretical properties, enabling us to prove general coreset theorems. We apply our results to both the k-means and the linear regression problems, and give extensive empirical evidence that the small additional computational cost of DPP sampling comes with superior performance over its iid counterpart.

Méthodes de pénalisation

Organisation: Angelina Roche — Horaire: 16.30 — Salle: N102

Analyse statistique des méthodes par régularisation en apprentissage statistiques

Guillaume Lecué (CREST, ENSAÉ)

Dans le contexte de l'apprentissage, je présenterai les étapes clefs utilisées pour l'analyse statistique des méthodes par régularisation. L'objectif de l'exposé est de montrer quelques techniques 'systématiques' sous-jacentes à l'étude de ce type d'estimateurs.

En particulier, pour la fonction de perte quadratique, les processus quadratiques et multiplicatifs ainsi que les arguments de localisation et d'homogénéité seront présentés pour l'analyse du minimiseur du risque empirique (MRE). Il en résultera une introduction de deux points fixes (le 'quadratique' et le 'multiplicatif') qui décrivent la vitesse de convergence du MRE.

On retrouvera ces mêmes outils pour l'analyse des méthodes par régularisation (que ce soit celles induisant de la régularité ou de la parcimonie) mais en plus, on montrera le rôle clef de la sous-différentielle de la régularisation pour l'analyse statistique de ce type d'estimateurs.

On mettra en œuvre ces techniques dans le modèle linéaire pour divers exemples de fonction de régularisation incluant le LASSO, SLOPE et la norme nucléaire. Si le temps le permet, on montrera comment adapter ces techniques et résultats aux cas de fonctions de perte Lipschitz (convexes ou non) et des données corrompues et faiblement concentrées grâce au median-of-means.

Références :

G. Chinot, G. Lecué and M. Lerasle. "Robust high dimensional learning for Lipschitz and convex losses". *Journal of machine Learning research*. (233) :147, 2020.

P. Alquier, V. Cottet and G. Lecué. "Estimation bounds and sharp oracle inequalities of regularized procedures with Lipschitz loss functions". *The Annals of Statistics*, 47(4) :2117-2144, 2019

G. Lecué and M. Lerasle. "Robust Machine Learning by median of means : theory and practice". *The Annals of Statistics***, Volume 48, Number 2 (2020), 906-931.

G. Lecué and S. Mendelson. *Learning Subgaussian classes : Upper and minimax bounds*.

Pénalisation pour les variables encodées one-hot en apprentissage supervisé

Mokhtar Alaya (LMAC, Université de Technologie de Compiègne)

Dans le cadre d'apprentissage supervisé en grande dimension, un pré-traitement classique (et souvent nécessaire) consiste à standardiser les colonnes de la matrice des variables explicatives. Une autre approche consiste à les discrétiser, par exemple par processus de binarisation appelé encodage one-hot. Nous nous intéressons à combiner l'astuce de l'encodage one-hot avec une nouvelle pénalisation appelée binarsity. Dans chaque groupe de variables binaires issues du l'encodage one-hot d'une seule variable brute, cette pénalisation utilise une régularisation de variation-totale accompagnée d'une contrainte linéaire supplémentaire pour éviter la collinéarité au sein des groupes. Nous proposons des inégalités d'oracle non asymptotiques pour les modèles linéaires généralisés, et nous illustrons les performances numériques de notre approche sur plusieurs datasets en régression logistique.

Sélection de modèle pour la construction de bandes de confiance de la fonction moyenne

Émilie Devijver (LIG, Université Grenoble Alpes)

Dans cet exposé, nous nous intéresserons à la construction de bandes de confiance pour la fonction moyenne. Si la question d'intervalle de confiance est bien traitée dans le cas multivarié, le cas fonctionnel est nettement moins étudié dans la littérature.

Formellement, sur un intervalle, on cherche à construire une bande de confiance à partir d'un estimateur de la fonction moyenne, dont on va contrôler théoriquement la couverture. Pour tirer profit de l'aspect fonctionnel des données, nous proposons de travailler sur une base orthonormée (la base de Fourier par exemple).

Le but de ce travail est de sélectionner la dimension de la base de façon à construire la bande la plus pertinente. Des résultats numériques illustreront la modélisation et l'analyse théorique. Ce travail est en collaboration avec Adeline Samson.

Comparaison de la stabilité et de la qualité des méthodes de sélection variable en analyse de survie

Mathilde Sautreuil (LBMC, ENS Lyon)

L'identification des gènes responsables d'une maladie en médecine de précision est un enjeu majeur. De nombreuses méthodes statistiques ont été développées pour faire de la sélection de variables. Les méthodes les plus utilisées sont les méthodes de régularisation et de screening. Cependant, le désavantage de ces méthodes concerne la stabilité et la qualité de la sélection en grande dimension. Il est donc essentiel d'étudier ces procédures afin de recommander la meilleure méthode aux cliniciens en fonction de leurs cas d'étude.

Dans cette présentation, on compare la performance des méthodes de sélection de variables dans le cadre de la survie en grande dimension à partir de plusieurs indices. On évaluera la stabilité et la qualité de ces méthodes à partir de données simulées et la stabilité de la sélection de gènes sur un jeu de données réelles.

L'étude de simulation a montré une mauvaise stabilité et qualité de la sélection des méthodes de régularisation. Les méthodes de screening ont de meilleurs résultats, mais la qualité de la sélection reste faible. L'ajout de la connaissance biologique dans les méthodes de screening a montré une amélioration de ces performances. Nous recommandons aux cliniciens d'utiliser les méthodes de screening couplées à la connaissance biologique en grande dimension afin de détecter les gènes les plus pertinents en analyse de survie.

Sessions parallèles – Mardi 30 août

Statistique adaptative, bandits

Organisation: Sébastien Gerchinovitz — Horaire: 10.15 — Salle: N101

Earning while Learning; a quick overview of online and bandit learning

Vianney Perchet (CREST, ENSAÉ)

In this introductory talk, I will present the main techniques, pitfalls, and results of online learning, a.k.a., bandits. The main difference between classical machine learning (and/or statistics) is that data are not obtained as a batch, but gathered and treated sequentially, on the fly. This setting can be quite intricate, as algorithms are building their own data-set; as a consequence, standard properties such as the law of large numbers might not hold, and estimation techniques must be carefully adapted. Moreover, the loss is evaluated on the whole dataset (and not on some test-set) which prevents the use of off-the-shelf naïve algorithms. The objective of this survey is to provide all the required concepts and knowledge to the audience for the three other talks of the session.

Novel bandit algorithms for early phase efficacy phase 2 trials

Sofia Villar (MRC Biostatistics Unit, Cambridge)

Although the scope of Multi-armed bandit problems is much more general, one of the most common applications chosen to motivate this methodology is that of a clinical trial which has the aim of balancing two separate goals : (1) correctly identifying the best treatment (exploration or learning) and (2) treating patients effectively during the trial (exploitation or earning). These two goals are naturally in conflict and appear all along the drug development phases. Correctly identifying the best treatment at the end of the trial (i.e. at a target treatment effect margin and with a high probability) requires a large number of patients to be assigned to all treatments to be able to compare them, and therefore goal (1) acts to limit goal (2) during the trial. In this talk I will briefly introduce a new response-adaptive randomised patient allocation procedure based on the Gittins index solution to the multi-armed bandit problem (Villar and Rosenberger, 2017). The rule explicitly aims to balance learning and ethical goals in a phase II setting. Then, using this rule as an example, I will illustrate some of the statistical and practical challenges that the use of bandit-based approaches to design clinical trials pose. If time allows, I will present a few solutions that have been proposed so far to make their implementation more appropriate in practice, in particular an approach first introduced in Barnett et al, 2021.

Bandit algorithms for early stage clinical trials

Émilie Kaufmann (CRIStAL, Université de Lille)

In this talk, we will investigate the use of bandit algorithms for dose-finding trials, which are often performed in phase I clinical trials. In oncology, under the classical assumption that the efficacy of a treatment is increasing with the dose, the goal is to identify the maximum tolerated dose (MTD), that is the highest dose whose toxicity is below some threshold. Doing so by adaptively allocating doses to patients in a trial poses a new bandit problem. We will propose several adaptations of bandit algorithms for MTD identification, with a focus on variants of Thompson Sampling. We will also discuss the challenges of other types of early stage trials.

Risk Minimization from Adaptively Collected Data : Guarantees for Policy

Antoine Chambaz (MAP5, Université de Paris)

Empirical risk minimization (ERM) is the workhorse of machine learning but its model-agnostic guarantees can fail when using data collected in an adaptive fashion, like in the setting of a contextual bandit algorithm for instance. In this setting, and focusing on policy learning, I will present a generic importance sampling weighted ERM algorithm and its regret guarantees, which close an open gap in the existing literature whenever exploration decays to zero. An empirical investigation validates the theory.

This is a joint work with Aurélien Bibaut (Netflix), Nathan Kallus (Cornell University and Netflix), Maria Dimakopoulou (Netflix) and Mark J. van der laan (UC Berkeley)

Détection de ruptures

Organisation: Farida Enikeeva et Nicolas Verzelen — Horaire: 10.15 — Salle: N102

Détection d'une rupture offline et online

Guillem Rigail (INRAE)

Au cours des dernières années, un grand nombre d'approches ont été proposées pour détecter des ruptures dans la moyenne d'un signal Gaussien. Je présenterai le problème plus simple de la détection d'une seule rupture offline qui me semble utile pour comprendre les idées sous-jacentes aux différentes méthodes de détection de ruptures (leurs similitudes et leurs différences) et les comparer. Puis, je décrirai une nouvelle méthode pour la détection d'une seule rupture online permettant de considérer simultanément toutes les tailles de fenêtre (ou toutes les valeurs possibles pour l'amplitude de la rupture) avec un coût de calcul attendu par itération logarithmique en le nombre d'observations.

Travaux joints en partie avec Gaetano Romano, Paul Fearnhead et Idris Eckley.

<https://onlinelibrary.wiley.com/doi/full/10.1002/sta4.291>

<https://arxiv.org/abs/2110.08205>

Détections de change-points dans un modèle en grande dimension avec des erreurs asymétriques ou avec des outliers

Gabriela Ciuperca (ICJ, Université Lyon 1)

On considère deux types de modèles avec change-points : a posteriori et en temps réel. Le nombre de variables explicatives peut être très grand et les erreurs peuvent être asymétriques ou comporter des outliers. Dans un modèle sans change-points, pour sélectionner automatiquement les variables pertinentes, la méthode d'estimation sera basée sur une fonction de perte expectile ou quantile avec une pénalité de type LASSO adaptative. Les estimateurs obtenus satisfont les propriétés oracle, c'est-à-dire la sparsité et la normalité asymptotique.

Les modèles avec change-points a posteriori sont analysés une fois que toutes les observations ont été réalisées. On se pose à ce moment-là la question si nous avons eu un seul modèle ou s'il y a eu plusieurs changements (phases), inconnus, dans le modèle. Nous estimons le nombre de changements, leur localisation, l'estimation et la sélection automatique des variables pertinentes dans chaque phase du modèle.

Pour la détection en temps réel d'un changement, on étudie par des tests d'hypothèses si le modèle construit sur les données historiques change à chaque nouvelle observation mesurée. La statistique de test est construite comme une somme cumulative (CUSUM) fonction des résidus LASSO adaptatifs. Nous montrons que la statistique de test converge sous l'hypothèse nulle vers la norme L^∞ du supremum d'un processus de Wiener et elle converge en probabilité sous l'hypothèse alternative vers ∞ .

Des simulations et des applications sur des données réelles seront présentées pour chaque type de modèle et chaque méthode d'inférence statistique.

Université Paris Nanterre

Émilie Lebarbier (MODAL'X)

Nous considérons un processus de Poisson dont l'intensité est supposée constante par morceaux. L'objectif est de détecter les instants de sauts, appelées ruptures, et d'en déterminer le nombre.

L'inférence des modèles de segmentation requiert de visiter l'espace de toutes les segmentations possibles. En raison de la nature continue des paramètres de ruptures, cet ensemble est infini et ne peut donc être exploré d'un point de vue computationnel. Nous montrons, que pour une classe générale de contrastes, satisfaisant une hypothèse de concavité en les instants de ruptures, le problème d'optimisation se réduit à un problème d'optimisation discret qui peut être résolu à

l'aide d'un algorithme efficace bien connu emprunté au cas de la détection de ruptures discrètes, l'algorithme de programmation dynamique.

Pour le choix du nombre de ruptures, nous proposons d'utiliser une méthode de validation croisée tirant avantage de l'algorithme rapide précédent et d'une propriété spécifique des processus de Poisson.

Des résultats de simulations seront présentés ainsi qu'une illustration sur des données d'éruptions volcaniques.

Détection et localisation minimax d'une rupture d'intensité dans un processus de Poisson

Ronan Le Guével (IRMAR, Université Rennes 2)

On considère un processus de Poisson observé sur un intervalle fixé borné, et l'on s'intéresse aux problèmes de détection et de localisation offline d'une rupture de son intensité constante. La réponse est donnée sous forme de procédures de tests minimax non-asymptotiques, en introduisant dans un premier temps une procédure de détection adaptée à la connaissance ou non des paramètres du modèle, puis en proposant une localisation de l'instant de rupture construite à l'aide d'une procédure de tests multiples minimax.

Systemes dynamiques aléatoires et graphes aléatoires

Organisation: 10.15 — Horaire: Jérôme Casse — Salle: N103

Estimations uniformes en temps pour la taille du chaos du système de McKean-Vlasov

Armand Bernou (LJLL, Sorbonne Université)

L'étude de la limite de champ moyen pour le système de McKean-Vlasov est un sujet dynamique, en particulier concernant la question de la propagation du chaos. Récemment, l'importation d'outils issus de la théorie des jeux à champ moyen a permis des avancées, et notamment l'obtention de résultats uniformes en temps pour des notions faibles de chaos dans le cas d'une interaction lisse (cf. Delarue-Tse 2021). Je présenterai un travail récent dans lequel ces techniques sont adaptées pour quantifier les corrélations entre trois et plus particules du système, ce qui permet d'obtenir des corrections à l'équation de champ moyen, de façon uniforme en temps. J'essaierai également de présenter les outils principaux de notre analyse. Travail en commun avec Mitia Duerinckx (ULB).

Chaînes de flips et cartes aléatoires

Alessandra Caraceni (INdAM, École Normale Supérieure de Pise)

Un problème de longue date proposé par David Aldous consiste à estimer le temps de mélange de la "triangulation walk", une chaîne de Markov définie sur l'ensemble de toutes les triangulations possibles du n -gone régulier. À chaque étape, on choisit une diagonale uniformément au hasard et, avec une probabilité de $1/2$, on la "flip", c'est-à-dire on la remplace par l'autre diagonale du quadrilatère formé par les deux triangles adjacents. Bien qu'il ait été démontré que le relaxation time pour la triangulation walk est polynomial en n et borné en dessous par $Cn^{3/2}$, une borne supérieure correspondante reste fermement hors de portée malgré l'apparente simplicité de la chaîne. Pour les chaînes des "flips" sur différents modèles – tels que les cartes planaires, les quadrangulations de la sphère, les lattice triangulations et autres graphes géométriques – on en sait encore moins. Dans cet exposé, je présenterai des techniques qui permettent d'étudier ce genre de problèmes et des résultats obtenus dans des travaux en commun avec Alexandre Stauffer.

Parking sur l'arbre binaire infini

Alice Contat (LMO, Université Paris Saclay)

On considère l'arbre binaire infini, chaque sommet étant interprété comme une place de parking pouvant accueillir une voiture. Sur cet arbre, arrive un nombre de voitures i.i.d. sur chaque sommet. Chaque voiture essaie de se garer sur son sommet d'arrivée et si la place est occupée, elle se dirige vers la racine de l'arbre jusqu'à trouver une place disponible. On observe alors une transition de phase : soit l'espérance du nombre de voitures qui ne trouvent pas de place disponible est finie (régime sous-critique), soit p.s toutes les places sont occupées et un nombre infini de voitures ne se gare pas (régime surcritique). Dans cet exposé, nous expliquerons comment caractériser ces régimes en fonction de la loi d'arrivée des voitures. Travail en commun avec David Aldous, Nicolas Curien et Olivier Hénard.

Graphes Aléatoires Dynamiques : point de vue de la Convergence Locale

Léo Dort (UMPA, ENS Lyon)

La limite locale de graphes, introduite par Itai Benjamini et Oded Schramm en 2001, décrit la notion qu'un graphe fini, vu d'un sommet spécifique, ressemble à un certains graphe limite. Plus précisément, ces objets limites sont des graphes aléatoires enracinés (infinis) qui décrivent la géométrie interne de grands graphes (finis) vus d'un sommet choisi uniformément au hasard. Dans cet exposé, je présenterai une notion de limite locale pour des graphes dynamiques, c'est-à-dire des graphes dans lesquels on autorise les arêtes apparaître et disparaître au cours temps. Puis nous déterminerons la limite locale pour un modèle simple : le graphe de Erdős-Rényi dynamique (ERD). Il s'agit d'une percolation dynamique sur le graphe complet à n sommets. Pour ce modèle, nous verrons que la limite locale est un arbre qui peut « croître » et « se segmenter » au cours du temps. Si le temps le permet, je présenterai une extension de ce résultat de convergence à une plus large classe de graphes dynamiques : les graphes aléatoires inhomogènes dynamiques dont le modèle ERD est un cas particulier. Cet exposé est basé sur un travail en collaboration avec Emmanuel Jacob (UMPA, ENS de Lyon).

Théorèmes limites

Organisation: Olivier Durieu — Horaire: 10.15 — Salle: C100

Inégalités de déviation pour des U-statistiques

Davide GIRAUDO (Université de Strasbourg)

Dans cet exposé, nous présenterons des inégalités de déviation pour des U-statistiques de données indépendantes à valeurs dans un espace de Banach. Elles consistent, dans le cas où la U-statistique vérifie de bonnes propriétés de martingales, à contrôler la queue d'une U-statistique à l'aide d'une fonctionnelle de la somme des carrés des accroissements. Nous formulerons ensuite des applications à la vitesse de convergence dans la loi des grands nombres ainsi qu'à des théorèmes limites fonctionnels.

Comportement local pour les champs aléatoires fractionnaires en dimension 2

Vytaute PILIPAUSKAITE (Université du Luxembourg)

Dans cet exposé on s'intéresse à une classe de champs aléatoires fractionnaires X en dimension 2, définis comme des intégrales stochastiques par rapport à une mesure aléatoire infiniment divisible. On considère deux types d'accroissements de X : accroissements ordinaires et accroissements rectangles. Dans les deux cas, on étudie les limites des accroissements de X quand l'échelle diminue dans les directions horizontale et verticale. Nos résultats révèlent un phénomène surprenant : selon la vitesse respective à laquelle l'échelle diminue dans les deux directions, différents champs aléatoires limites apparaissent. On discute des propriétés des champs aléatoires limites aussi. Travail en commun avec Donatas Surgailis (Université de Vilnius).

Théorèmes centraux limite quenched pour des champs de variables aléatoires

Lucas REDING (Université de Valenciennes)

Dans cet exposé nous présentons des théorèmes centraux limite quenched pour des champs stationnaires de variables aléatoires satisfaisant à une condition projective. Afin d'obtenir de tels résultats, les méthodes d'approximations fonctionnelles proposées par Gordin (1969) et Rosenblatt (1972) s'avèrent être un outil puissant. Depuis la publication de ces articles fondateurs, de nombreux résultats concernant à la fois les séries temporelles ou bien les champs de variables aléatoires ont été obtenus au travers de cette méthode. Cependant, le cadre quenched a été très peu étudié et c'est seulement récemment que Peligrad et Volný (2020) ont démontré la validité de théorèmes limite quenched pour les ortho-martingales. En faisant usage de ces résultats ainsi que de la méthode d'approximation par ortho-martingales, nous obtenons des théorèmes limite quenched pour des champs de variables aléatoires satisfaisant à différents critères projectifs similaires à la condition de Hannan.

Théorèmes asymptotiques pour processus cumulatifs et application aux processus de Hawkes

Laetitia COLOMBANI (Université de Toulouse)

Les processus cumulatifs, aussi nommés « renewal-reward process » ou « compound renewal process », sont une généralisation des processus de renouvellement. Il peut par exemple s'agir de sommer des quantités aléatoires associées à des intervalles de temps suivant un processus de renouvellement. Si la loi des grands nombres et le théorème central limite sont déjà bien connus, le principe de grande déviation est plus difficile à obtenir. Dans cet exposé, je présenterai ces processus ainsi que les inégalités que nous avons obtenu. J'indiquerai également notre motivation, à savoir étudier une certaine catégorie de processus de Hawkes, et les résultats obtenus dans ce cadre.

Géométrie stochastique

Organisation: Raphaël Lachieze-Rey — Horaire: 14.35 — Salle: N101

Menhirs aléatoires dans des déserts poissoniens

Nathanaël Enriquez (LMO, Université Paris-Saclay)

Après une présentation des techniques de base de l'étude de la mosaïque de Poisson-Voronoi, nous montrerons que le diagramme de Voronoi dans les zones désertiques d'un processus de Poisson donne lieu à des cellules très allongées dont nous étudierons la convergence après renormalisation convenable. (Travail en commun avec Pierre Calka et Yann Demichel).

Sur l'existence – ou non – d'une hypersurface nodale non bornée

Hugo Vanneuille (Institut Fourier, Université Grenoble Alpes)

Soit $f : \mathbb{R}^d \rightarrow \mathbb{R}$ une fonction analytique aléatoire telle que $(f(x))$ est un champ gaussien stationnaire et centré. Dans cet exposé, on étudiera la géométrie du lieu nodal, i.e. de $\{f = 0\}$. On se demandera en particulier si cet ensemble possède une composante connexe non bornée (la réponse dépendra de la dimension). Les deux difficultés principales sont : (a) la rigidité de tels objets ; (b) dans le cas où la covariance du champ n'est pas toujours à valeurs positives, le fait que cela entraîne la perte de nombreux outils qu'on aurait pu emprunter à la théorie de la percolation. Cet exposé se repose sur un travail en commun avec Hugo Duminil-Copin, Alejandro Rivera et Pierre-François Rodriguez (et, si le temps le permet, sur un travail en commun avec Stephen Muirhead et Alejandro Rivera).

Réseaux invariants d'échelle spatiaux

Emmanuel Jacob (UMPA, ENS Lyon)

Si l'on veut modéliser un réseau complexe (réseau social, réseau de communication ou de transport, internet, etc), on recherche généralement un graphe invariant d'échelle ou "scale-free", à savoir dont la loi empirique des degrés suit approximativement une loi puissance. Les graphes aléatoires invariants d'échelle les plus classiques - modèle de configuration ou modèle d'attachement préférentiel - décrivent cependant mal la structure locale de ces graphes, en ayant des coefficients de corrélation bien trop faibles. Pour y remédier, il est utile de travailler sur un espace métrique sous-jacent. Dans cet exposé je survolerai quelques modèles de ce type et leurs propriétés, notamment le modèle d'attachement préférentiel spatial et le modèle de graphe géométrique hyperbolique.

Full inference for the anisotropic fractional Brownian field

Frédéric Richard (I2M, Aix-Marseille Université)

The anisotropic fractional Brownian field is a non-stationary Gaussian field (Bonami and Estrade, 2003) which has been used for the modeling of image microtextures (Richard, 2016-18). Having stationary increments, its probability distribution is characterized by a semi-variogram whose spectral representation in polar coordinates

$$\forall x \in \mathbb{R}^d, v(x) = \frac{1}{2} \int_{-\pi/2}^{\pi/2} \tau(\theta) |\cos(\arg(x) - \theta)|^{2\beta(\theta)} |x|^{2\beta(\theta)} d\theta$$

is determined by two non-negative π -periodic functions τ and β called the topothesy and the Hurst functions, respectively. In this talk, we focus on the issue of estimating these two functions from a single realization of the field.

Solving that issue is a key point for the characterization and the classification of image textures. It would also pave the way to the simulation of realistic textures from the field model. In the literature, this issue has been partially tackled in two different studies. In (Biermé and Richard, 2008), a method was built upon the Radon transform of the field to estimate the Hurst function. But, due to discretization issues, its application is restricted to a few directions θ and leads to inaccurate results. In (Richard, 2018), an inverse problem was stated and solved to estimate the topothesy in directions where the Hurst function is minimal.

In this talk, we will present a method for the estimation of the whole topothesy and Hurst functions. This method is based on a turning-band field which was initially proposed for the simulation of anisotropic fractional Brownian field (Biermé, Moisan and Richard, 2015), and whose semi-variogram approximates the one of an anisotropic fractional Brownian field. We will set an

optimization problem to fit the semi-variogram of the turning-band field to the empirical semi-variogram. This problem is formulated as a non-linear separable least square problem (Golub, 2003). We then use a variable projection method to design an algorithm to solve numerically the problem. The design of this algorithm also includes a multigrid approach to improve the convergence. Eventually, we will present a numerical study of the performances of this algorithm on textures generated by the python package [PyAFBF](<https://fjprichard.github.io/PyAFBF/>).

Statistique des processus

Organisation: Charlotte Dion et Céline Duval — Horaire: 14.35 — Salle: N103

Quelques questions de statistiques des processus inspirées par des applications en biologie

Maud Delattre (Unité MaIAGE, INRAE)

Dans cet exposé introductif, je présenterai quelques modèles construits à partir de processus stochastiques, dont l'étude a été motivée par des problèmes concrets en biologie. Je parlerai ensuite plus spécifiquement des modèles de Markov cachés et des processus de diffusion. Je présenterai les difficultés rencontrées dans l'inférence de ces modèles et introduirai quelques outils classiques pour les traiter. Enfin, je présenterai quelques contributions dans le cas où ces processus stochastiques sont combinés à des effets aléatoires dans des modèles à effets mixtes paramétriques, ces modèles permettant de décrire une variabilité entre dynamiques individuelles lorsque plusieurs individus sont observés simultanément au cours du temps comme dans les essais cliniques de pharmacologie.

Inférence dans les modèles de Hawkes avec inhibition

Anna Bonnet (LPSM, Sorbonne Université)

Un processus de Hawkes est un processus ponctuel qui sert à modéliser l'apparition d'événements au cours du temps de sorte que la probabilité d'observer un nouvel événement dépende des observations précédentes. Il a historiquement été développé pour modéliser des phénomènes auto-excitants, c'est-à-dire qu'un événement augmente la probabilité d'en observer un nouveau. Au cours des dernières années, il y a eu un intérêt croissant pour modéliser l'effet opposé, qu'on appelle l'inhibition. Dans cet exposé, je présenterai une méthode d'estimation paramétrique adaptée aux cas d'excitation et d'inhibition, qui sera illustrée sur des données simulées et appliquée à un jeu de données issu des neurosciences.

Sur un estimateur de Nadaraya-Watson pour des copies indépendantes d'un processus de diffusion

Nicolas Marie (Modal'X, Université Paris Nanterre)

L'exposé portera sur un estimateur de Nadaraya-Watson de la fonction de drift d'un processus de diffusion, calculé à partir de copies indépendantes de ce dernier. A temps continu, puis à temps discret, nous présenterons des bornes de risque sur les estimateurs du numérateur et du dénominateur, puis sur l'estimateur de Nadaraya-Watson lui-même. Nous présenterons une extension de la méthode de sélection de fenêtre par cross validation leave-one-out et illustrerons sa pertinence via des expériences numériques. Enfin, nous donnerons le schéma de preuve d'une inégalité d'oracle pour l'estimateur adaptatif obtenu en sélectionnant la fenêtre du numérateur et la fenêtre du dénominateur séparément via une méthode type PCO. Il s'agit d'un travail en collaboration avec Amélie Rosier.

Modèles espace-état pour la prévision adaptative de consommation électrique

Joseph de Vilmarrest (LPSM, Sorbonne Université et EDF)

L'électricité étant difficile à stocker, prévoir la demande est un enjeu majeur pour maintenir l'équilibre entre la production et la consommation. L'évolution des usages de l'électricité, le déploiement des énergies renouvelables, et plus récemment la crise du coronavirus, motivent l'étude de modèles qui évoluent au cours du temps. Nous nous sommes intéressés au cadre des modèles espace-état pour tenir compte de ces changements de régime. Dans ce paradigme, l'environnement (ou le contexte) est représenté par un état caché. À chaque instant, la demande dépend de cet état que nous cherchons donc à estimer grâce aux observations dont nous disposons, et selon certaines hypothèses sur la dynamique du système. L'estimation de l'état nous permet ensuite de prévoir la demande. Nos prévisions s'appuient sur des modèles de prévision existants, par exemple le modèle additif généralisé, que nous cherchons à adapter. Ainsi, nous tirons parti de certaines dépendances complexes capturées par ces modèles existants, par exemple la sensibilité de la consommation d'électricité à la température, tout en profitant de la faculté d'adaptation des modèles espace-état.

Systèmes de particules

Organisation: Oriane Blondel — Horaire: 14.35 — Salle: N102

Quantitative homogenization of interacting particle systems

Jean-Christophe Mourrat (UMPA, ENS Lyon)

I will discuss a class of interacting particle systems in continuous space. Such models are known to "homogenize", in the sense that the behavior of the cloud of particles is approximately described by a partial differential equation over large scales. In the talk, I will describe a first step towards making this result quantitative. The approach is inspired by recent developments in the homogenization of elliptic equations with random coefficients. Joint work with Arianna Giunti and Chenlin Gu.

Large deviations for the SSEP in weak contact with reservoirs

Clément Éricignoux (INRIA Lille Nord-Europe)

(joint with A. Bouley and C. Landim)

The SSEP in contact with reservoirs is a simple and very rich model for nonequilibrium phenomena. In particular, the effect of slowing down the boundary interaction has been in recent years under considerable scrutiny. Twenty years ago, Derrida, Lebowitz and Speer, managed to derive the steady state large deviation principle for the SSEP in - strong - contact with reservoirs, and characterize the large deviations functional by means of the solution to a non-linear ODE. I will present a result recently obtained, in which we show that the non linear equation they obtain also characterizes, with properly changed boundary conditions, the large deviations principle for the steady state of the SSEP in weak contact with reservoirs.

Sharp spectral gap for regular and disordered chain of oscillators

Angeliki Menegaki (IHES, Université Paris-Saclay)

We consider one-dimensional chains and multi-dimensional networks of harmonic oscillators coupled to two Langevin heat reservoirs at different temperatures. Each particle interacts with its nearest neighbors by harmonic potentials and all individual particles are confined by harmonic potentials, too. In this talk, I will present the sharp N-particle dependence of the spectral gap of the associated generator in different physical scenarios and for different spatial dimensions. Our method of proof relies on a new approach to analyze a non self-adjoint eigenvalue problem involving low-rank non-hermitian perturbations of discrete Schroedinger operators. If time permits we will discuss some new results on the oscillator chains with longer-range interactions. This is a joint work with Simon Becker (Courant Institute).

Stability of interacting particle systems using Toom contours

Reka Szabo (Ceremade, Université Paris Dauphine)

We review and extend Toom’s classical result about stability of trajectories of cellular automata. This extension allows us to prove stability and derive bounds for certain interacting particle systems. Joint work with J. Swart and C. Toninelli.

Transport optimal en statistique

Organisation: Claire BréchetEAU — Horaire: 14.35 — Salle: C100

Introduction to Optimal Transport and its application in Machine Learning, with a focus on domain adaptation

Alain Rakotomamonjy (LITIS, Université de Rouen Normandie & Criteo AI Lab)

This talk presents a brief review on optimal transport and Wasserstein distance and will review some of its applications in machine learning as a loss for learning distributions or as a distance between embeddings of text. We will also consider the problem of unsupervised domain adaptation (UDA) between a source and a target domain under conditional and label shift a.k.a Generalized Target Shift (GeTarS) and how estimating a Monge mapping is key in this context.

Transport optimal pour l’apprentissage de modèles génératifs

Antoine Houdard (Ubisoft Bordeaux)

L’utilisation d’un coût de transport optimal pour l’apprentissage de modèles génératifs a été popularisé par les Wasserstein GANs (Generative Adversarial Networks). L’apprentissage d’un Wasserstein GAN nécessite de différentier le coût de transport optimal par rapport aux paramètres du modèle génératif. Dans cet exposé, nous fournissons des conditions suffisantes pour l’existence d’une formule de gradient dans deux cadres différents : le cas du transport optimal semi-discret (i.e. une distribution cible discrète) et le cas du transport optimal régularisé entropique. À partir de cette formule du gradient, nous proposons un algorithme alternatif pour l’apprentissage d’un modèle génératif.

JDCOT : an Algorithm for Transfer Learning in Incomparable Domains using Optimal Transport

Valérie Garès (IRMAR, INSA de Rennes)

(joint with Marion Jeamart, Renan Bernard, Nicolas Courty and Chloé Friguet)

Domain adaptation is a field of transfer learning where the training data (source) and the test data (target) come from different domains. The data in these two domains have therefore different underlying distributions, and the learned model should be adapted so that it can be used on the target data with good performance. We present here a domain adaptation algorithm to adapt heterogeneous domains, i.e. described by different features. The developed method uses optimal transport to map the distributions of the two domains and its implementation is illustrated on benchmark data.

Wasserstein Multivariate Autoregressive Models for distributional time series

Yiye Jiang (IMB, Université de Bordeaux)

In this work, we propose a new autoregressive model for multivariate distributional time series. We consider a collection of N series of probability measures supported over a bounded interval in R , which are indexed by distinct time instants. Especially, we wish to develop such a model which can identify the dependency structure in the temporal evolution of the measures. To this end, we adopt the Wasserstein metric. We establish the regression model in the Tangent space of the Lebesgue measure by first "centering" all the raw measures so that their Fréchet means turn Lebesgue. The uniqueness and stationarity results are provided. We also propose a consistent estimator for the model coefficient. In addition to the simulated data, the proposed model is illustrated on two real data sets : age distribution of countries and the bike sharing network in Paris.

Sessions parallèles – Mercredi 31 août

Intelligence artificielle

Organisation: Mohamed Hebiri — Horaire: 10.30 — Salle: C100

Vers un apprentissage profondément économique

Sébastien Loustau (LMAP, Université de Pau et des Pays de l'Adour)

Dans cet exposé, je présenterai les aspects théoriques et pratiques d'un apprentissage profond plus économe en électricité. Je débiterai par un survol de l'état de l'art en matière de machine learning (réseaux binaires, pruning notamment), de logiciels de mesures (rapl, nvidia-smi, et les comparaisons avec les mesures physiques), et de hardware. Puis j'introduirai les avancées récentes en optimisation et statistiques mathématiques pour obtenir de nouveaux compromis entre efficacité statistiques (risk/regret bounds), et consommation. Pour cela, je présenterai un nouvel optimiseur qui se passe de la back-propagation pour apprendre des réseaux binaires et parcimonieux.

DeepLearning léger et applications : Les Réseaux de Neurones Binaires

Matthieu François et Jordy Palafox (LMAP, Université de Pau et des Pays de l'Adour)

L'architecture, le format des poids, le choix des opérateurs, ont un impact significatif sur l'énergie nécessaire pour l'inférence et l'entraînement d'un réseau de neurones profond. Motivés par un contexte environnemental toujours plus sous pression, nous vous présenterons leur implémentation et les avantages que nous avons pu en tirer. Ces méthodes permettent d'embarquer des modèles sur des micro-processeurs en gardant une accuracy compétitive comparativement avec les modèles dis "classiques". Notre attention s'est portée sur les techniques de binarisation et d'early exit, qui feront l'objet de la présentation. La première permet de binariser les poids et les activations de réseaux de neurones et la seconde de limiter la durée d'une inférence.

Fujitsu Computer Vision Framework : industrialisation des projets de data science

Margarita Khokhlova (Fujitsu France, Paris)

L'équipe de data science Fujitsu est spécialisée dans les solutions permettant d'extraire de la valeur à partir de données visuelles - qu'il s'agisse d'obtenir des informations sur le comportement des clients pour optimiser le placement des produits dans les environnements de vente, de fabriquer des produits de meilleure qualité en détectant les défauts sur la chaîne de production, en détectant les fraudes aux caisses libre-service ou à l'identification des menaces dans les espaces publics.

Soutenue par une grande variété de technologies et d'outils Fujitsu, notre équipe FUJITSU Computer Vision (FCV) travaille avec les entreprises pour comprendre leur défi commercial et la meilleure façon de le surmonter, en utilisant une approche collaborative et indépendante de la technologie spécifique. Nous fournissons des solutions de reconnaissance d'images de bout en bout basées sur des algorithmes d'IA de pointe.

Dans ce discours, je vais présenter un cycle de vie type de projet CV adressé par Fujitsu, et les principaux défis auxquels nous sommes confrontés sur les exemples de quelques solutions sur lesquelles nous avons travaillé par le passé.

Classification multi-classes sous contrainte d'équité

Christophe Denis (LAMA, Université Gustave Eiffel)

L'équité algorithmique est un domaine de recherche en pleine expansion.

L'objectif est de corriger le biais inhérent aux données afin de rendre plus "équitable" les prédictions des algorithmes d'apprentissage. Dans cette présentation, nous présenterons des résultats récents dans ce domaine en apprentissage multi-classes.

Optimisation en finance

Organisation: Emmanuel Lépinette — Horaire: 10.30 — Salle: N101

Théorie des ensembles aléatoires : conditionnement et applications en optimisation et en finance

Emmanuel Lépinette (CEREMADE, Université Paris Dauphine)

Le contenu de l'exposé est motivé par un problème très classique en mathématiques financières. Il s'agit de l'évaluation de produits dérivés, en particulier les options Européennes. Une option Européenne est un titre financier qui s'achète à un certain prix et qui, en contrepartie, délivre en une certaine date (appelée maturité $T > 0$) une richesse (payoff) dont la valeur est une fonction

d'un actif sous-jacent, en général l'action d'une entreprise cotée en bourse ou un indice comme le Cac40.

On souhaite déterminer le prix minimal d'achat d'une telle option selon le marché. Par définition, un prix est la valeur initiale d'un portefeuille autofinancé dont la valeur terminale en T est au moins égale au payoff de l'option, avec probabilité 1. Depuis le modèle de Black et Scholes dans les années 70, des progrès considérables ont été réalisés dans ce domaine, tout au moins pour les modèles de marché financier sans coûts de transactions.

En effet, pour les modèles linéaires, la théorie du Pricing repose sur la dualité convexe qui permet de caractériser les prix des options via des éléments duaux que sont les probabilités de risque neutre sous lesquels les prix actualisés des actions du marché considéré sont des martingales. Ainsi, il a été démontré que, dans le cas d'une unique probabilité de risque neutre (marché complet), le prix minimal d'un option Européenne est l'espérance du payoff actualisé, calculée sous la probabilité de risque neutre.

En pratique, les transactions financières sont soumises à des coûts de transactions mais aussi, contrairement à ce qui est supposé dans la théorie classique, il peut y avoir plusieurs prix pour les actions, selon le volume d'achat ou de vente demandé. Ceci est visible dans les carnets d'ordre. Dans ce cas, on ne peut pas espérer modéliser de tels phénomènes par un marché financier linéaire si bien que toute la théorie jusqu'alors développée à partir des probabilités de risque neutre n'est plus applicable.

Depuis les modèles avec coûts de transactions proportionnelles de Kabanov, il est admis qu'un marché financier avec coûts de transaction est défini par un processus stochastique à valeurs des ensembles aléatoires modélisant, à un instant donné, l'ensemble des positions financières dont la valeur liquidative est supérieure ou égale à 0. Dès que le modèle n'est plus linéaire, en particulier avec des coûts fixes, les résultats de dualité convexe usuels ne s'appliquent plus d'où la nécessité de développer de nouvelles techniques de Pricing.

Nous présentons lors de cette conférence, des résultats nouveaux qui permettent de résoudre le problème d'évaluation des options pour de tels modèles. Pour cela, nous développons de nouvelles techniques de conditionnement des ensembles aléatoires qui apparaissent comme étant les outils naturels même dans le cas des modèles traditionnels car on peut se passer des probabilités de risque neutre qui ne sont pas si simple à estimer, et encore moins les caractérisations duales des prix, en particulier dans les marchés de risque neutre.

Le plan de l'exposé sera le suivant :

1. Rappel sur les options Européennes ; Pricing à partir de portefeuilles auto-financés. Théorème fondamental d'évaluation d'un actif contingent (FTAP).
2. Modèles de marché avec coûts de transactions.
3. Ensembles aléatoires graph-mesurable.
4. Conditionnement d'ensembles aléatoires.
5. Application à l'optimisation.

Méthodes alternatives de Pricing pour des modèles financiers avec coûts de transactions

Duc Thinh Vu (CEREMADE, Université Paris Dauphine)

Les modèles de marché financier avec coûts de transactions sont définis à partir d'un processus stochastique à valeurs dans des ensembles aléatoires modélisant, à un instant donné, l'ensemble des positions financières dont la valeur liquidative est supérieure ou égale à 0. Pour des modèles très généraux, on établit un principe de programmation dynamique qui permet de calculer les prix de sur-réplication d'un option Européenne de manière rétrograde, tout au moins en temps discret. Ce résultat est établi sans aucune condition d'absence d'opportunité d'arbitrage, contrairement à ce

qui est usuel de faire dans la littérature. Les notions de conditionnement des ensembles aléatoires jouent un rôle essentiel et nouveau, aussi l'essentiel supremum ou infimum conditionnel remplacent les espérances conditionnelles. Une condition faible naturelle de non-arbitrage est déduite et celle-ci est facilement vérifiable sur les marchés financiers. On illustre cette nouvelle approche par des exemples.

Optimisation de champ moyen régularisée par l'information de Fisher

Songbo Wang (CEREMADE, Université Paris Dauphine)

Récemment, il y a un intérêt croissant pour la recherche sur l'optimisation du champ moyen, en particulier en raison de son rôle dans l'analyse de la formation des réseaux de neurones. Dans cet exposé, en ajoutant l'information de Fisher (ou l'énergie cinétique de Schrodinger) comme régularisateur, nous relierons le problème d'optimisation de champ moyen à une diffusion McKean-Vlasov Birth-Death (MVBD). Nous développons une méthode d'énergie pour montrer que les distributions marginales du processus MVBD convergent vers l'unique minimiseur du problème d'optimisation régularisée. Il s'agit d'un travail commun en cours avec Julien Claisse, Giovanni Conforti et Zhenjie Ren.

Exposé de Yuri Kabanov (LMB, Université Franche-Comté)

Percolation

Organisation: Jean-Baptiste Gouéré — Horaire: 10.30 — Salle: N102

Sensibilité au bruit en percolation, sans outil spectral Hugo Vanneuville (Institut Fourier, Université Grenoble Alpes)

Dans cet exposé, je parlerai du phénomène de sensibilité au bruit en percolation critique planaire. Ce phénomène peut être résumé de la façon suivante : savoir qu'un rectangle macroscopique est traversé après avoir "un petit peu bruité" la configuration ne nous dit essentiellement rien sur la configuration initiale. Depuis que cette notion a été introduite (en 1999, par Benjamini, Kalai et Schramm), tous les travaux qui ont été réalisés sur ce sujet se reposent sur des outils spectraux. Ces outils sont très riches, mais ont des limitations que j'exposerai brièvement. Avec Vincent Taschon, nous avons proposé une nouvelle approche, non spectrale, qui consiste à étudier des inégalités différentielles qui décrivent comment les arêtes pivots sont affectées par le bruit. Le but de cet exposé est d'expliquer cette approche, après avoir énoncé certains résultats du domaine, tels que les théorèmes de Schramm–Steif et Garban–Pete–Schramm.

Modèle de percolation pour les réseaux D2D en milieu urbain David Corlin Marchand (Institut Mines Télécom Nord Europe)

L'arrivée de la 5G pourrait changer radicalement l'infrastructure des réseaux mobiles. S'il est nécessaire aujourd'hui pour un opérateur d'investir dans de nombreuses antennes relais afin que les usagers bénéficient partout d'une bonne couverture réseau, demain, les communications Device-to-Device (D2D) — ou communications directes et de courte portée entre deux terminaux mobiles — pourraient en théorie, au moins en milieu urbain, transformer les usagers eux-mêmes, via leurs appareils, en mini antennes relais. Mathématiquement, le bon fonctionnement d'un réseau ainsi construit peut s'exprimer comme un problème de percolation. On modélise la voirie d'une ville par une mosaïque aléatoire plane. Sur les arêtes de cette mosaïque, on jette un processus ponctuel de Cox, représentant les usagers. Enfin, un processus ponctuel de Bernoulli sur les sommets de la mosaïque décrit la présence d'antennes relais supplémentaires, réparties aléatoirement et plus ou moins densément, aux carrefours des rues de la ville. Le réseau mobile D2D est représenté par un graphe, dont les sommets sont les atomes des deux processus ponctuels précédents, et dont les arêtes relient les paires de sommets suffisamment proches et situés sur une même arête de la mosaïque. La percolation du graphe aléatoire, c'est-à-dire l'existence d'une composante connexe infinie, traduit le fonctionnement à grande échelle du réseau mobile D2D. Dans cet exposé, nous présenterons les différents régimes de connectivité du réseau selon la densité en utilisateurs et en antennes relais, quand la mosaïque sous-jacente est une triangulation de Delaunay. Travail en collaboration avec David Coupier et Benoît Henry.

Crossing probabilities for positively associated, planar percolation

Laurin Köhler-Schindler (ETH Zürich)

The study of crossing probabilities, known as Russo-Seymour-Welsh (RSW) theory, is a fundamental tool to understand the critical and near-critical behaviour of planar percolation processes. Initiated for Bernoulli percolation on the square lattice in 1978, it has later been extended to many other models such as FK-percolation, Voronoi percolation or level-set percolation of Gaussian fields. In this talk, we present a general RSW result valid for any invariant planar percolation process satisfying positive association. More precisely, we show that the probability of crossing a rectangle in the long direction is related by a universal homeomorphism to the probability of crossing it in the short direction. In contrast to previous results, our argument does not rely on any independence or mixing properties. This talk is based on joint work with Vincent Tassion (ETH).

Percolation arithmétique

Mike Liu (CREST, ENSAÉ)

Coloriez les points de \mathbb{Z}^d en blanc si leurs coordonnées sont premières entre elles et en noir sinon. Il n'est pas très difficile d'étudier l'existence ou non des composantes infinies du coloriage résultant. Qu'en est-il lorsque l'on regarde ce coloriage depuis un point "uniforme" de \mathbb{Z}^d ? Nous établissons, à l'aide de méthodes probabilistes et élémentaires, une nouvelle preuve du fait que le coloriage aléatoire résultant est surcritique.

Trajectoires rugueuses

Organisation: Rémi Cattelier — Horaire: 10.30 — Salle: N103

Rough paths : old and new

Laure Coutin (IMT, Université Toulouse 3)

During this talk we will provide a short presentation of rough paths theory using the framework of controlled rough paths introduced by M. Gubinelli and their applications.

Non-linear Young equations in the plane and pathwise regularization by noise

Fabian Harang (Norwegian Business School)

Regularization by noise for stochastic differential equations has been a long studied topic in the field of stochastic analysis. After the work of Gubinelli and Catellier in 2016 on a pathwise analogue to the probabilistic analysis of regularization by noise based on what they called averaged fields in combination with the concept of non-linear Young equations, this area of study has recently received much attention. In this talk we will discuss an extension of this approach to pathwise differential equations in the plane. These are essentially hyperbolic non-linear PDEs with additive noise. We extend the so-called local time formulation of the regularization by noise approach to these equations, and extend the concept of non-linear Young equations to rectangular domains in order to prove wellposedness of these equations, even when the nonlinear coefficient is a generalized function (e.g. distribution). We illustrate the application of this construction by proving wellposedness of a wave equation with a noisy boundary constructed from two independent fractional Brownian motions. In the end we will discuss further potential applications and extensions of this framework and present several open challenges.

This talk is based on a joint collaboration with Florian Bechtold (Bielefeld University) and Nimit Rana (Bielefeld University), see arXiv :2206.05360

Proper scoring rules and statistics for rough paths

Patric Bonnier (Mathematical Institute, University of Oxford)

Many forecasts consist not of point predictions but concern the evolution of quantities. For example, a central bank might predict the interest rates during the next quarter, an epidemiologist might predict trajectories of infection rates, a clinician might predict the behaviour of medical markers over the next day, etc. The situation is further complicated since these forecasts sometimes only concern the approximate "shape of the future evolution" or "order of events". Formally, such forecasts can be seen as probability measures on spaces of equivalence classes of paths modulo time-parametrization. We leverage the statistical framework of proper scoring rules with classical mathematical results to derive a principled approach to decision making with such forecasts. In particular, we introduce notions of gradients, entropy, and divergence that are tailor-made to respect the underlying non-Euclidean structure.

Hairer's multilevel Schauder estimates without regularity structures

Lucas Broux (LPSM, Sorbonne Université)

It is well-known that convolving a distribution against a singular kernel, such as the Heat kernel, results in a distribution with improved regularity : this is the content of the "classical" Schauder estimates. It was later understood by Hairer that such kernels still enjoy regularising properties even against "families of local approximations" of distributions : this has been formalised in the "multilevel" Schauder estimates of regularity structures, and are one of the crucial steps permitting deep results of renormalisation for singular stochastic PDEs. The purpose of this talk is to present and revisit those estimates as results in distribution theory, with the goal of making only minimal references to the formalism of regularity structures. (Work in collaboration with Francesco Caravenna and Lorenzo Zambotti)

Liste des participants

Agathe-Nerine	Zoé	MAP5, Univ. Paris Cité	zoe.agathe-nerine@u-paris.fr
Al Marjani	Aymen	UMPA, ENS Lyon	aymen.al_marjani@ens-lyon.fr
Alaya	Mokhtar Z.	LMAC, Univ. Technologie de Compiègne	elmokhtar.alaya@utc.fr
Angeliki	Menegaki	IHES, Univ. Paris Saclay	menegaki@ihes.fr
Auvray	Edwige	LMRS, Univ. Rouen	edwige.auvray@univ-rouen.fr
Balsollier	Lisa	LMJL, Univ. Nantes, INRIA Rennes	lisa.balsollier@univ-rennes1.fr
Banna	Marwa	NYU, Abu Dhabi	marwa.banna@nyu.edu
Bansaye	Vincent	CMAP, Ecole Polytechnique	vbansaye@gmail.com
Barbu	Vlad	LMRS, Univ. Rouen	vladstefan.barbu@univ-rouen.fr
Bardet	Jean-Baptiste	LMRS, Univ. Rouen	jean-baptiste.bardet@univ-rouen.fr
Barrier	Antoine	UMPA, ENS Lyon, LMO, Univ. Paris Saclay	antoine.barrier@ens-lyon.fr
Belhadj	Amel	Taher Moulay University, Algérie	amlblh50@gmail.com
Belmabrouk	Nadia	LAMA, Univ. Gustave Eiffel Créteil	nadia.belmabrouk@u-pec.fr
Ben Alaya	Mohamed	LMRS, Univ. Rouen	mohamed.ben-alaya@univ-rouen.fr
Benzoni	Séverin	LMRS, Univ. Rouen	severin.benzoni1@univ-rouen.fr
Berger	Quentin	LPSM, Sorbonne Univ.	quentin.berger@sorbonne-universite.fr
Bernou	Armand	LJLL, Sorbonne Univ.	armand.bernou@sorbonne-universite.fr
Biau	Gérard	LPSM, Sorbonne Univ.	gerard.biau@upmc.fr
Blanc	Guillaume	LMO, Univ. Paris Saclay	guillaume.blanc1@universite-paris-saclay.fr
Blondel	Oriane	ICJ, Univ. Lyon 1	blondel@math.univ-lyon1.fr
Bonnet	celine	UMPA, ENS Lyon	celine.bonnet@ens-lyon.fr
Bonnet	Anna	LPSM, Sorbonne Univ.	anna.bonnet@upmc.fr
Bonnier	Patric	Institut Math. Univ. d'Ox	patric.bonnier@maths.ox.ac.uk
Bouley	Angèle	LMRS, Univ. Rouen	angele.bouley@univ-rouen.fr
BréchetEAU	Claire	IRMAR, Univ. Rennes 2	claire.brecheteau@univ-rennes2.fr
Brouard	Vianney	UMPA, ENS Lyon	vianney.brouard@ens-lyon.fr
Broux	Lucas	LPSM, Sorbonne Univ.	lucas-broux@orange.fr
Brua	Jean-Yves	LMRS, Univ. Rouen	jean-yves.brua@univ-rouen.fr
Calka	Pierre	LMRS, Univ. Rouen	pierre.calka@univ-rouen.fr
Caraceni	Alessandra	INdAM, École Normale Supérieure Pise	alessandra.caraceni@gmail.com
Carlier	Clara	CREST ENSAÉ	clara.carlier@ensae.fr
Casse	Jérôme	LMO, Univ. Paris Saclay	jerome.casse.math@gmail.com
Catellier	Rémi	LJAD, Univ. Côte d'Azur	remi.catellier@univ-cotedazur.fr
Chagny	Gaëlle	LMRS, Univ. Rouen	gaelle.chagny@univ-rouen.fr
Chambaz	Antoine	MAP5, Univ. Paris Cité	antoine.chambaz@parisdescartes.fr
Channarond	Antoine	LMRS, Univ. Rouen	antoine.channarond@univ-rouen.fr
Chapron	AuréliE	LMRS, Univ. Rouen	aurelie.chapron@univ-rouen.fr
Ciuperca	Gabriela	ICJ, Univ. Lyon 1	Gabriela.Ciuperca@univ-lyon1.fr
Clausel	Marianne	IECL, Univ. Lorraine	marianne.clausel@univ-lorraine.fr
Coeurjolly	Jean-François	LJK, Univ. Grenoble Alpes	Jean-Francois.Coeurjolly@univ-grenoble-alpes.fr
Colombani	Laetitia	IMT, Univ. Paul Sabatier Toulouse	laetitia.colombani@math.univ-toulouse.fr
Contat	Alice	LMO, Univ. Paris Saclay	alice.contat@universite-paris-saclay.fr
Corlin Marchand	David	LMO, Univ. Paris Saclay	david.corlin.marchand@gmail.com
Costa	Manon	IMT, Univ. Paul Sabatier Toulouse	manon.costa@math.univ-toulouse.fr
Coste	Simon	Centre Science des Données, ENS Paris	simon.coste@inria.fr
Coupier	David	IMT Nord Europe	david.coupier@imt-nord-europe.fr
Coutin	Laure	IMT, Univ. Paul Sabatier Toulouse	laure.coutin@math.univ-toulouse.fr
Cugliari	Jairo	ERIC, Univ. Lyon 2	Jairo.Cugliari@univ-lyon2.fr
D'Achille	Matteo	LAMA, Univ. Paris-Est Créteil	matteo.dachille@u-pec.fr
Dadoun	Benjamin	NYU, Abu Dhabi	benjamin.dadoun@gmail.com
Dalalyan	Arnak	CREST, ENSAÉ	arnak.dalalyan@ensae.fr
Da Veiga	Sébastien	Safran	sebastien.da-veiga@safrangroup.com
Davydov	Michel	INRIA et ENS Paris	michel.davydov@inria.fr
de la Rue	Thierry	LMRS, Univ. Rouen	thierry.de-la-rue@univ-rouen.fr
de Raphélis	Loïc	IDP, Univ. Orléans	loic.de.raphelis@gmail.com
de Vilmarest	Joseph	LPSM, Sorbonne Univ.	josephdevilmarest@gmail.com
Delattre	Maud	Unité MaIAGE, INRAE	maud.delattre@inrae.fr
Dembin	Barbara	ETH Zürich	barbara_dembin@hotmail.fr
Denis	Christophe	LAMA, Univ. Gustave Eiffel	christophe.denis@u-pem.fr
Deshayes	Aurelia	LAMA, Univ. Paris-Est Créteil	aurelia.deshayes@u-pec.fr
Devijver	Emilie	LIG, Univ Grenoble Alpes	emilie.devijver@univ-grenoble-alpes.fr
Dort	Léo	UMPA, ENS Lyon	leo.dort@ens-lyon.fr

Duboux	Thibaut	IMB, Univ. Bourgogne	thibaut.duboux@u-bourgogne.fr
Duprey	Corentin	HOP! Air France	coduprey@hop.fr
Durieu	Olivier	IDP, Univ. Tours	olivier.durieu@lmpt.univ-tours.fr
Duval	Céline	LPP, Univ. Lille	celine.duval@univ-lille.fr
Egea	Maxime	LAREMA, Univ. Angers	maxime.egea@univ-angers.fr
El Machkouri	Mohamed	LMRS, Univ. Rouen	mohamed.elmachkouri@univ-rouen.fr
Enikeeva	Farida	LMA, Univ. Poitiers	farida.enikeeva@math.univ-poitiers.fr
Enriquez	Nathanael	LMO, Univ. Paris Saclay	nathanael.enriquez@u-psud.fr
Erignoux	Clément	INRIA Lille Nord Europe	clement.erignoux@inria.fr
Faipeur	Corentin	Institut Fourier, Univ. Grenoble Alpes	corentin.faipeur@univ-grenoble-alpes.fr
Fouladirad	Mitra	Labo. M2P2, École Centrale Marseille	mitra.fouladirad@centrale-marseille.fr
François	Matthieu	LMAP, Univ. Pau et Pays de l'Adour	matthieu.francois@univ-pau.fr
Freulon	Paul	IMB, Univ. Bordeaux	paul.freulon72@gmail.com
Garban	Christophe	ICJ, Univ. Lyon 1	garban@math.univ-lyon1.fr
Garès	Valérie	IRMAR, INSA de Rennes	valerie.gares@insa-rennes.fr
Gerchinovitz	Sébastien	IRT Saint Exupéry et IMT, Toulouse	sebastien.gerchinovitz@irt-saintexupery.com
Giraud	Davide	IRMA, Univ. Strasbourg	dgiraud@unistra.fr
Gkelsinis	Thomas	LMRS, Univ. Rouen	thomas.gkelsinis@univ-rouen.fr
Goncalves	Branda	LPTM, Cergy Paris Univ.	branda.goncalves@cyu.fr
Gonon	Thierry	ICJ, École Centrale Lyon	thierry.gonon@ec-lyon.fr
Gouéré	Jean-Baptiste	IDP, Univ. Tours	jean-baptiste.gouere@univ-tours.fr
Halé	Sandrine	LMRS, Univ. Rouen	sandrine.hale@univ-rouen.fr
Hammami	Mayssa	LAMA, Univ. Paris-Est Marne la Vallée	hammami.mayssa1@gmail.com
Harang	Fabian	Norwegian Business School	fabian.a.harang@bi.no
Hebiri	Mohamed	LAMA, Univ. Gustave Eiffel	Mohamed.Hebiri@u-pem.fr
Henry	Benoît	IMT Lille Douai	benoit.henry@imt-lille-douai.fr
Hoscheit	Patrick	Unité MaIAGE, INRAE	patrick.hoscheit@inra.fr
Houdard	Antoine	Ubisoft Bordeaux	antoine.houdard@ubisoft.com
Jacob	Emmanuel	UMPA, ENS Lyon	emmanuel.jacob@normalesup.org
Jacquet	Antonin	IDP, Univ. Tours	antonin.jacquet@univ-tours.fr
Janvresse	Elise	LAMFA, Univ. Picardie	Elise.Janvresse@u-picardie.fr
Jiang	Yiye	IMB, Univ. Bordeaux	yiyejiang93@gmail.com
Journel	Lucas	LPSM, Sorbonne Univ.	lucas.journel@upmc.fr
Kabanov	Yuri	LMB, Univ. Franche-Comté	ykabanov@univ-fcomte.fr
Kaufmann	Emilie	CRIStAL, Univ. Lille	emilie.kaufmann@gmail.com
Keller	Merlin	EDF, Paris	merlinkeller@gmail.com
Khebbache	Selma	IRT SystemX	selma.khebbache@irt-systemx.fr
Khezeli	Ali	Univ. Tarbiat Modares	alikhzezi@gmail.com
Khokhlova	Margarita	Fujitsu France	margokhokhlova@gmail.com
Köhler-Schindler	Laurin	ETH Zürich	laurin.koehler-schindler@math.ethz.ch
Krell	Nathalie	IRMAR, Univ. Rennes 1	nathalie.krell@univ-rennes1.fr
Lachieze-Rey	Raphael	MAP5, Univ. Paris Cité	raphael.lachieze-rey@parisdescartes.fr
Laruelle	Sophie	LAMA, Univ. Paris-Est Créteil	sophie.laruelle@u-pec.fr
Laulin	Lucile	IMB, Univ. Bordeaux	lucile.laulin@math.u-bordeaux.fr
Le Guével	Ronan	IRMAR, Univ. Rennes 2	ronan.leguevel@univ-rennes2.fr
Le Ny	Arnaud	LAMA, Univ. Paris-Est Créteil	arnaud.le-ny@u-pec.fr
Lebarbier	Emilie	Modal'X, Univ. Paris Nanterre	emilie.lebarbier@parisnanterre.fr
Lecué	Guillaume	CREST ENSAÉ	guillaume.lecue@ensae.fr
Ligonnière	Maxime	IDP, Univ. Tours	maxime.ligonniere@univ-tours.fr
LIU	Mike	CREST ENSAÉ	liu.mike1011@gmail.com
Loustau	Sébastien	LMAP, Univ. Pau et Pays de l'Adour	sebastien.loustau@univ-pau.fr
Lépinette	Emmanuel	CEREMADE, Univ. Paris Dauphine	emmanuel.lepinette@ceremade.dauphine.fr
Male	Camille	IMB, Univ. Bordeaux	camille.male@gmail.com
Manent	Emilien	IRMAR, Univ. Rennes 2	emilien.manent@gmail.com
Marcovici	Irène	IECL, Univ. Lorraine	irene.marcovici@univ-lorraine.fr
Marie	Nicolas	Modal'X, Univ. Paris Nanterre	nmarie@parisnanterre.fr
Marion	Rémy	SNCF Voyageurs, Saint Pierre des Corps	r.marion@sncf.fr
Marteau	Clément	ICJ, Univ. Lyon 1	marteau@math.univ-lyon1.fr
Mavrogiannis	Ioannis	LMRS, Univ. Rouen	ioannis.mavrogiannis@univ-rouen.fr
Meynaoui	Anouar	LIG, Univ. Grenoble Alpes	topanover@gmail.com
Mignot	Raphael	IECL, Univ. Lorraine	raphael.mignot@univ-lorraine.fr
Monmarché	Pierre	LJLL et LCT, Sorbonne Univ.	pierre.monmarche@upmc.fr
Mourragui	Mustapha	LMRS, Univ. Rouen	mustapha.mourragui@univ-rouen.fr
Mourrat	Jean-Christophe	UMPA, ENS Lyon	jean-christophe.mourrat@ens-lyon.fr
Mourtada	Jaouad	CREST, ENSAÉ	jaouad.mourtada@ensae.fr

Muré	Joseph	EDF R&D, Chatou	joseph.mure@edf.fr
Palafox	Jordy	LMAP, Univ. Pau et Pays de l'Adour	jordy.palafox@univ-pau.fr
Perchet	Vianney	CREST ENSAÉ	vianney.perchet@normalesup.org
Pergamenchtchikov	Serguei	LMRS, Univ. Rouen	Serge.Pergamenchtchikov@univ-rouen.fr
Piccolo	Vanessa	UMPA, ENS Lyon	vanessa.piccolo@ens-lyon.fr
Pilipauskaite	Vytaute	Univ. du Luxembourg	vytaute.pilipauskaite@gmail.com
Porté	Anne-Lise	??	porte.anne-lise@outlook.fr
Printems	Jacques	LAMA, Univ. Gustave Eiffel Créteil	printems@u-pec.fr
Pronzato	Luc	I3S, Univ. Côte d'Azur	pronzato@i3s.unice.fr
Quilan	Gauthier	LMRS, Univ. Rouen	gauthier.quilan@univ-rouen.fr
Rakotomamonjy	Alain	LITIS, Univ. Rouen et Criteo AI Lab	alain.rakotomamonjy@insa-rouen.fr
Raschel	Kilian	LAREMA, Univ. Angers	raschel@math.cnrs.fr
Raynaud de Fitte	Paul	LMRS, Univ. Rouen	prf@univ-rouen.fr
Reding	Lucas	Ecole des Mines St Etienne	lucas.reding@gmail.com
Ren	Zhenjie	CEREMADE, Univ. Paris Dauphine	zhenjie.ren@gmail.com
Richard	Frédéric	I2M, Aix-Marseille Univ.	frederic.richard@univ-amu.fr
Riera	Armand	LMO, Univ. Paris Saclay	armandlloberes@gmail.com
Rigail	Guillem	INRAE et LaMME, Univ. d'Evry	guillem.rigail@inra.fr
Roche	Angelina	CEREMADE, Univ. Paris Dauphine	roche@ceremade.dauphine.fr
Rosenbaum	Mathieu	CMAP, Ecole Polytechnique	mathieurosenbaum@hotmail.com
Rousselle	Arnaud	IMB, Univ. Bourgogne	arnaud.rousselle@u-bourgogne.fr
Rousselot	Mathilde	LMA, Univ. Poitiers	mathilde.rousselot@math.univ-poitiers.fr
Saada	Ellen	MAP5, Univ. Paris Cité	Ellen.Saada@mi.parisdescartes.fr
Sautreuil	Mathilde	LBMC, ENS Lyon	Mathilde.sautreuil@gmail.com
Schmisser	Emeline	LPP, Univ. Lille	emeline.schmisser@math.univ-lille1.fr
Sekkat	Inass	CERMICS, École des Ponts ParisTech	inass.sekkat@enpc.fr
Smail	Hamed	LMRS, Univ. Rouen	hamed.smail@univ-rouen.fr
Solea	Eftychia	CREST, ENSAI	eftychia.solea@ensai.fr
Szabo	Réka	CEREMADE, Univ. Paris Dauphine	reka.szabo@dauphine.psl.eu
Théret	Marie	Modal'X, Univ. Paris Nanterre	marie.theret@parisnanterre.fr
Thomassey	Loïc	MAP5, Univ. Paris Cité	loic.thomassey@etu.u-paris.fr
Tirel	Nicolas	LMAP, Univ. Pau	ntirel@univ-pau.fr
Tremblay	Nicolas	GIPSA-Lab, Univ. Grenoble Alpes	nicolas.tremblay@gipsa-lab.fr
Truong	Charles	Centre Borelli, ENS Paris Saclay	ctruong@ens-paris-saclay.fr
Vanneuville	Hugo	Institut Fourier, Univ. Grenoble Alpes	hugo.vanneuville@univ-grenoble-alpes.fr
Velasco	Sonia	MAP5, Univ. Paris Cité	sonia.velasco@u-paris.fr
Verges	Julien	IDP et MODAL'X, Univ. Tours et Paris Nanterre	julien.verges42@gmail.com
Vergne	Nicolas	LMRS, Univ. Rouen	nicolas.vergne@univ-rouen.fr
Vo	Tran Duy	IDP, Univ. Tours	duybieber195@gmail.com
Vu	Duc-Thanh	CEREMADE, Univ. Paris Dauphine	thinh.vu2016@qcf.jvn.edu.vn
Wicker	Nicolas	LPP, Univ. Lille	nicolas.wicker@univ-lille.fr
Yao	Anne-Françoise	LMBP, Univ. Clermont Auvergne	anne.yao@uca.fr
Zotsa Ngoufack	Arsène Brice	I2M, Aix-Marseille Univ.	arsene-brice.zotsa-ngoufack@univ-amu.fr

Plan du centre de Rouen

- A Site des journées MAS 2022, 3 avenue Pasteur, Rouen (voir page iv le plan du bâtiment)
- B Hôtel de Ville de Rouen (accueil des participants des journées par la Ville, lundi 29 août à 19h)
- C Hôtel de Bourgtheroulde (dîner de la conférence, mardi 30 août à 19h30), 15 place de la Pucelle

