

**Onzième Colloque
Jeunes Probabilistes et Statisticiens**

Forges-les-Eaux, du 6 au 11 Avril 2014

Résumés

Céline ABRAHAM

Quelques résultats de convergence pour les cartes planaires aléatoires

Les cartes aléatoires, qui peuvent être vues comme des discrétisations de surfaces aléatoires browniennes, jouent un rôle important dans de nombreux domaines des mathématiques et de la physique. Dans cet exposé, on donnera dans un premier temps la définition des cartes planaires aléatoires et quelques notations associées. Puis on exposera des résultats sur les limites d'échelle de telles cartes. Ces résultats nécessitent de faire le lien entre les cartes planaires et des arbres via des bijections combinatoires que l'on introduira. On énoncera aussi des théorèmes de convergence pour les arbres, qui sont les outils-clés dans la preuve du théorème principal.

Adil AHIDAR

Surfaces Quantile Multivarié et Théorèmes de Convergence

Dans cet exposé, nous introduisons une notion de quantile multivarié qui n'est pas basée sur une M-estimation globale mais une M-estimation directionnelle. Ceci est aussi une généralisation du quantile univarié. Nous nous intéresserons ensuite au quantile empirique associé qui présente de bonnes propriétés géométriques. Puis nous établirons une loi des grands nombres uniforme, un théorème central limite, ainsi qu'un principe d'invariance forte et enfin une généralisation du théorème de Bahadur-Kiefer au cas multivarié.

Anna AKSAMIT

Optional semimartingale decomposition and non-arbitrage condition in enlarged filtration

Our study addresses the question of how an arbitrage-free semimartingale model is affected when stopped at a random horizon or when a random variable satisfying Jacod's hypothesis is incorporated. Precisely, we focus on the No-Unbounded-Profit-with-Bounded-Risk condition, which is also known in the literature as the first kind of non-arbitrage. In the general semimartingale setting, we provide a necessary and sufficient condition on the random time for which the non-arbitrage is preserved for any process. Analogous result is formulated for initial enlargement with random variable satisfying Jacod's hypothesis. Moreover we give an answer to a stability of non-arbitrage question for fixed process. The crucial intermediate results in enlargement of filtration theory are obtained. For local martingales from the reference filtration we provide special optional semimartingale decomposition up to random time and in initially enlarged filtration under Jacod's hypothesis. An interesting link to absolutely continuous change of measure problem is observed. This is a joint work with Tahir Choulli, Jun Deng and Monique Jeanblanc.

Mélisande ALBERT

Tests d'indépendance entre deux processus ponctuels par bootstrap et permutation

Considérant un échantillon i.i.d de couples de processus ponctuels, observés sur une plage de temps donnée, on se pose la question de la détection des dépendances entre les processus sous-jacents. Cette question est motivée par l'étude des corrélations entre potentiels d'action en neurosciences. Suite au grand débat sur la modélisation de l'activité neuronale, notre but est de proposer un test d'indépendance ne nécessitant aucune hypothèse sur la répartition des potentiels d'action. On suit principalement Romano qui proposa des tests d'indépendance basés sur des approches de bootstrap et de permutation. Cependant, ici, en raison de la nature de nos variables (à savoir des processus ponctuels), nos statistiques de test sont plus complexes, et ne peuvent être vues comme des processus empiriques évalués sur des familles d'événements particulières. On doit donc pousser un peu plus loin les arguments de Romano afin d'obtenir, par exemple, les convergences des lois des statistiques bootstrappée et permutée vers la loi de la vraie statistique de tests sous l'hypothèse d'indépendance. Après avoir présenté les motivations d'un point de vue biologique, on montrera que nos tests sont de niveau asymptotique ou exact voulu en soulignant les principales variations avec les tests dans \mathbb{R}^d . On montrera également qu'ils sont consistants contre certaines alternatives.

Références :

Grün, S., Diesmann, M., and Aertsen, A.M. (2010). Unitary Events Analysis. *In Analysis of Parallel Spike Trains*, Grün, S., and Rotter, S., Springer Series in Computational Neuroscience.

Hoeffding, W. (1952). The large-sample power of tests based on permutations of observations. *The Annals of Mathematical Statistics* **23**, 2, 169–192.

Pipa, G. and Grün, S. (2003). Non-parametric significance estimation of joint-spike events by shuffling and resampling. *Neurocomputing*, 52-54 :31-37.

Romano, J.P. (1989). Bootstrap and Randomization Tests of some Nonparametric Hypotheses. *The Annals of Statistics* **17**, 1, 141-159.

Tuleau-Malot, C., Rouis, A., Grammont, F. and Reynaud-Bouret, P. (2013). Multiple Tests based on a Gaussian Approximation of the Unitary Events method with delayed coincidence count. Preprint, <http://hal.archives-ouvertes.fr/hal-00757323>. *To be published in Neural Computation in 2014.*

Julyan ARBEL

Bayesian nonparametric dependent model for the study of diversity in species data

We introduce a dependent Bayesian nonparametric model for the probabilistic modelling of species-by-site data, i.e. population data where observations at different sites are classified in distinct species. Our aim is to study the impact of additional factors (covariates), for instance environmental factors, on the data structure, and in particular on the diversity. To that purpose, we introduce dependence a priori across the covariates, and show that it improves posterior inference. The model is illustra-

ted on simulated data and on a real data set obtained in experiments conducted in Antarctica soil.

Benjamin ARRAS

Calcul stochastique par rapport au processus de Rosenblatt

Dans cet exposé, on présentera le processus de Rosenblatt, processus aléatoire auto-similaire et à accroissements stationnaires appartenant au deuxième chaos Gaussien de Wiener. On s'intéressera plus particulièrement aux différentes approches utilisables afin de définir un calcul stochastique par rapport à ce processus (étudiées par C. Tudor en 2008). On verra en quoi la théorie des distributions de Hida peut apporter un regard légèrement nouveau sur ce problème.

François BACHOC

Estimation paramétrique de la fonction de covariance dans le modèle de Krigeage par processus Gaussiens

Un modèle de Krigeage par processus Gaussiens consiste à modéliser une fonction déterministe par la réalisation d'un processus Gaussien. L'objectif final est, étant donné un ensemble de valeurs observées de la fonction, de construire un prédicteur de cette fonction. Ce modèle est couramment utilisé dans l'industrie (nucléaire, aéronautique...) lorsque la fonction représente un code de simulation numérique. Le modèle de Krigeage s'inscrit à ce titre dans la relativement récente discipline statistique de l'analyse des expériences numériques. Dans ce contexte, le problème de l'estimation de la fonction de covariance du processus Gaussien revêt une importance particulière. Dans cet exposé, l'estimation paramétrique de la fonction de covariance est étudiée, pour laquelle les estimateurs par Maximum de Vraisemblance et Validation Croisée sont considérés. Nous effectuons une analyse théorique de ces deux estimateurs dans un cadre asymptotique par expansion. Le plan d'expérience considéré est une grille régulière multidimensionnelle perturbée aléatoirement. Un résultat de consistance et de normalité asymptotique est montré pour les deux estimateurs, les matrices de covariance asymptotique ne dépendant que de l'amplitude de perturbation de la grille régulière. Il est alors d'abord montré que, bien que les deux estimateurs aient des vitesses de convergence identiques, la variance asymptotique est plus faible pour le Maximum de Vraisemblance. Il est ensuite montré numériquement que des amplitudes de perturbation importantes sont toujours préférables pour l'estimation par Maximum de Vraisemblance, ce qui confirme le constat général selon lequel les plans d'expériences irréguliers sont avantageux pour l'estimation de la fonction de covariance.

Charlotte BAEY

Modélisation de la variabilité interindividuelle dans les modèles de croissance de plantes

Il existe une forte variabilité génétique entre plantes, même au sein de la même variété, ce qui, combiné à la variation locale des conditions climatiques dans le champ,

peut conduire deux plantes voisines à se développer de façon très différentes. C'est l'une des raisons pour lesquelles les approches populationnelles dans les modèles de croissance de plantes suscitent un grand intérêt. Nous proposons dans cette étude une extension du modèle individu-centré Greenlab à l'échelle de la population, à l'aide d'un modèle non linéaire mixte. Deux variants stochastiques de l'algorithme EM (Espérance-Maximisation), le Monte-Carlo EM automatique (MCEM) et le SAEM seront comparés, en utilisant le fait que le modèle complet appartient à la famille exponentielle lorsque tous les effets sont considérés comme aléatoires.

Marwa BANNA

Matrices Aléatoires

Suite au développement technologique contemporain, l'étude des grandes matrices aléatoires est devenu un sujet important, qui trouve application dans de nombreux domaines comme le traitement de signal appliqué aux télécommunications, la finance et plusieurs autres. En particuliers, on s'intéresse à l'étude du spectre des grandes matrices aléatoires, ainsi qu'au comportement asymptotique de la distribution spectrale empirique. Des travaux de Wishart dans les années 30, étaient à l'initiative de l'étude des valeurs propres des matrices aléatoires, qui prit un nouvel envol dans les années 50, quand Wigner et Dyson commencèrent à s'y intéresser. Cet exposé commencera par une introduction aux matrices aléatoires et un rappel de certains outils utilisés dans ce domaine. Ensuite, on rappellera un résultat important sur la convergence des valeurs propres de matrices de covariance empirique avec des entrées i.i.d. En fin, nous nous intéresserons à des modèles de grandes matrices aléatoires associées à une classe de processus stationnaires.

Marco BANTERLE

Sufficient dimension reduction for ABC via RKHS

In an increasing number of research areas we are faced with complex statistical models where the likelihood function $p(y_{obs}|\theta)$ for the observed data y_{obs} (where $\theta \in \mathbb{R}^d$ is the parameter of interest) is unavailable/intractable either due to mathematical or computational reasons.

In these cases Bayesian inference, which relies on the posterior distribution

$$p(\theta|y_{obs}) \propto p(\theta)p(y_{obs}|\theta)$$

is impractical or even infeasible. We thus need to resort to some kind of practical approximation in order to conduct the inference.

Approximate Bayesian Computation (ABC) algorithms are a statistical tool for analysing this kind of models. ABC inference relies on samples from an approximate posterior where the only requirement is the feasibility of simulation from the model $y \sim p(\cdot|\theta)$. The ABC methodology induces however two types of approximation on the posterior. We focus on one of them, namely that the posterior on the full data $p(\theta|y_{obs})$ is replaced by $p(\theta|s_{obs})$ where $s_{obs} = S(y_{obs})$ is a lower-dimensional vector

(wrt the original y_{obs} dimension) of summary statistics on the real data. This is a good approximation if the set of statistics is highly informative on the parameter θ and it is exactly the posterior in the case of sufficiency.

While it's impossible to quantify the loss in taking a non-sufficient statistics, a low dimensional quasi-sufficient statistics helps in reducing the so-called **curse of dimensionality** common to many non-parametric techniques and hence is usual to resort to a collection of hopefully quasi-sufficient statistics, typically dependent on the problem, under the condition that this collection lies in a lower dimensional space to avoid the so-called curse of dimensionality. Following the idea underling dimension reduction techniques for ABC like Partial Least Squares, Wegman et al. (2009), we derive a procedure to select a sufficient dimension reduction subset u of the vector s by characterizing it as the minimal subset such that the distribution of θ is independent of s given u . PLS however operate only in the original space and the independence condition in their criterion is in fact replaced with an uncorrelation requirement. This simplification is known to fail outside Gaussian models, an assumption which is easily unsatisfied in complicate models where ABC is needed. We thus resort to kernel methods by examining the conditional independence requirement through mean embedding in a RKHS as in Fukumizu et al. (2008) and Zhang et al. (2012) where linear relations are known to represent non-linear dependencies in the original space. A proper testing procedure and a greedy procedure are compared in various simulation studies.

Oriane BLONDEL

Marcheurs sur verre

On présentera quelques résultats sur des marcheurs aléatoires en milieu aléatoire. Le milieu aléatoire sera ici donné par un modèle contraint cinétiquement (KCSM en anglais), introduit dans la littérature physique pour modéliser des systèmes vitreux. On donnera à la fois des résultats théoriques sur le comportement asymptotique des marcheurs aléatoires et des heuristiques apportant une compréhension plus profonde du comportement des KCSM à basse température.

Pierre BOSCH

Autour de la divisibilité infinie

La loi d'une variable aléatoire X est dite infiniment divisible (ID) si pour tout n , X se décompose en loi en une somme iid $X = X_1 + \dots + X_n$. Ces lois sont très répandues puisqu'il y a une correspondance entre les processus de Lévy et les lois ID. Nous verrons que des sous-classes apparaissent naturellement et qu'elles peuvent être très utile afin de prouver le caractère ID.

Anne BRIQUET

Hauteurs des arbres de Lyndon

Un mot de Lyndon est un mot plus petit (pour l'ordre lexicographique) que tous ses suffixes propres. On génère aléatoirement des mots de Lyndon de longueur n à partir d'un alphabet à m lettres, chaque lettre i ayant un poids p_i (i.e. on traite les cas uniformes et non-uniforme). Tout mot de Lyndon peut se décomposer en deux mots de Lyndon. Cette décomposition n'est pas unique. La décomposition qui a le facteur de gauche le plus court est appelée la décomposition standard. Les décompositions récursives créent une structure d'arbre. Les arbres de Lyndon sont utilisés pour décrire les algèbres de Lie libres et ont des applications pratiques en cryptographie et en physique statistique. On trouve que la hauteur de ces arbres converge en probabilité vers $\ln n$ fois une constante ne dépendant que de p_0 et de $\max(p_i, i \neq 0)$. Le niveau de saturation d'une lettre non nulle (i.e. la hauteur de la feuille étiquetée par cette lettre la plus basse) a la même forme, contrairement à celui de 0 (qui est le niveau de saturation de l'arbre, i.e. la hauteur de la feuille la plus basse) qui est en $O(1)$ et dont on a une expression explicite.

Rémi CATELLIER

Distributions Paracontrôlées et Équation de Quantisation Stochastique.

Un certain nombre d'équations aux dérivées partielles stochastiques très singulières présentent des problèmes dans leur définition même. C'est le cas, entre autres, de l'équation de KPZ, mais aussi de l'équation de quantisation stochastique en dimension 3

$$\partial_t u = \Delta u - u^3 + \xi \quad (1)$$

où $u \in C(\mathbb{R}_+; \mathcal{S}'(\mathbb{R}^3))$ et ξ est un bruit blanc espace temps. Dans cette équation, le produit u^3 ne peut pas être définie par des méthodes classiques d'analyse, et il s'avère que pour donner un sens à cette équation, il est nécessaire de soustraire une constante infinie et de considérer formellement le problème suivant

$$\partial_t u = \Delta u - (u^3 - \infty u) + \xi.$$

La théorie des distributions paracontrôlées, qui combine des idées de la théorie des Rough Paths avec la décomposition de Paley-Littlewood et le paraproduit, est un bon cadre pour donner un sens à cette renormalisation, et résoudre (localement) cette équation.

Dans une première partie nous introduirons donc la notion de distributions paracontrôlées, et dans une deuxième partie, nous montrerons comment cette théorie peut s'appliquer à l'équation (1).

Paul CHARTON*Gestion optimale d'une ferme éolienne couplée à un dispositif de stockage*

Du fait de la quantité limitée d'énergies fossiles et nucléaires disponibles et de leur coût écologique, la tendance actuelle est au développement des énergies renouvelables. En particulier l'énergie éolienne est en plein essor. Le principal inconvénient de l'énergie éolienne réside dans son caractère aléatoire. En effet, sur certains marchés, le producteur d'électricité doit déterminer à l'avance la puissance qu'il s'engage à délivrer. Le prix de vente de l'énergie est inconnu au moment du choix de la puissance à délivrer. De plus tout manquement à cet engagement entraîne des pénalités financières également inconnues lors de la livraison. Pour atténuer le caractère non constant de la production, nous proposons de coupler la ferme à un dispositif capable de stocker et de restituer une partie de la production. L'objectif est alors d'utiliser au mieux ce dispositif. Il faut pour cela prendre en compte la production future et les pénalités qui sont inconnues au moment de la livraison. Après avoir décrit plus en détails le marché de l'électricité, et les particularités de la production d'une ferme éolienne nous proposons une première modélisation probabiliste de notre problème. Dans ce modèle le problème de l'utilisation optimale de ce dispositif est un problème de contrôle stochastique optimal en temps continu. Nous montrons que ce problème est équivalent à la résolution d'une EDP au sens de viscosité. Enfin nous proposons une méthode numérique pour approcher la solution de cette EDP.

Claire CHRISTOPHE*Mathematical modeling of cytotoxic T lymphocyte efficacy during cancer immunoeediting*

The dynamics of Cytotoxic T Lymphocytes (CTL) interaction with tumor cells has been the subject of deep investigation using numerical simulations. However, stochastic mathematical models that take into account the phenomena of CTL displacement, lytic function and exhaustion together with tumor cell division and immunoeediting, are still missing. Here, we introduce a stochastic dynamical particle interaction model based on experimentally measured parameters. The model describes the competitive interaction between CTL and melanoma cell nodules and allows temporal and two-dimensional spatial progression. The model is designed to provide probabilistic estimates of tumor eradication in numerical simulations in which tunable parameters influencing CTL efficacy against a tumor nodule undergoing immunoeediting are tested. Our model shows that the rate of CTL/tumor nodule productive collisions during an initial time of interaction determines the success of CTL in tumor eradication. Two parameters emerge as crucial in allowing early productive collisions and tumor eradication : (i) a bias in CTL motility inducing a progressive attraction towards a few scout CTL, which have detected the nodule and (ii) the capacity of CTL to kill a large number of target cells before being exhausted. Taken together, our results show that CTL efficacy against a tumor nodule undergoing immunoeediting is strongly dependent on guidance of CTL trajectories by scout siblings and multiple

killing.

Quentin CLAIRON

Parameter estimation in Ordinary differential equations model using control theory

Ordinary differential equations (ODE's) are widespread tools to model biological process (SIR model, genetic regulation network). They are under the form $\dot{X} = f(t, X, \theta)$ where $X(t)$ represents the evolution of the biological variables of interest. They rely on parameters θ which are of critical importance in terms of dynamic and need to be estimated directly from noisy and partial observations of the system. The implicit definition of the regression function at the dynamic scale and not the observation scale as well as its great sensibility to parameter value often leads the classic statistic estimation methods (MLE, NLS) to fail and legitimate the search of alternative to classic approaches. Moreover the classic methods rarely propose misspecification criteria along the parametric estimation. We present here an estimation method based on a model relaxation at the dynamic scale by introducing the new model $\dot{X} = f(t, X, \theta) + u(t)$. Thanks to this relaxation we define an estimator minimizing both the discrepancy with the data and with the initial model. For doing so we use a result coming from control theory, Pontryagin maximum principle, in order to define the optimal function u and profile the cost in the space of possible functions u . We will see that on examples our method produces more precise estimation than classical estimators and the relaxation term allows us to explore the validity of the proposed model.

Antoine DAHLQVIST

Mouvements browniens sur les groupes compacts de matrices

Le but de cet exposé est de présenter le comportement du mouvement brownien multiplicatif sur des groupes compacts de matrices de grande taille. Philippe Biane, Thierry Lévy et Mylène Maïda ont décrit le comportement asymptotique de la mesure spectrale d'un mouvement brownien dans le groupe des matrices unitaires lorsque leur taille tend vers l'infini. On tentera de présenter ici des résultats de fluctuations de ces mesures spectrales et de montrer comment celles-ci dépendent du choix des groupes compacts.

Aurelia DESHAYES

Processus de contact avec vieillissement

Le Processus de Contact, introduit par Harris en 74, est un processus markovien décrivant l'évolution d'une infection sur \mathbb{Z}^d . Une généralisation à deux états a été introduite en 99 (Krone) : un site est d'abord infecté sans être contagieux puis devient contagieux dans un second temps. Dans cet exposé nous généraliserons ce modèle en introduisant une infinité d'états d'infection avec différents taux de contagion. Nous nous interrogerons d'abord sur la survie de l'infection. Nous évoquerons ensuite les problématiques des modèles non permanents (i.e. avec extinction possible).

Nous verrons comment utiliser les méthodes de comparaison à la percolation (type Bezuidenhout et Grimmett) pour obtenir des contrôles de l'évolution de l'infection et enfin conclure à un théorème asymptotique de la zone infectée (avec des méthodes presque sous additives type Garet et Marchand).

Xan DUHALDE

Processus de branchement avec immigration : temps d'entrée et critère de récurrence

Les CSBP (Continuous State Branching Processes) sont les processus de Markov en temps et espace continu qui s'obtiennent comme limite d'échelle de processus de Galton-Watson. Ils décrivent l'évolution au cours du temps d'une population avec reproduction asexuée. Leur loi est caractérisée par une fonction ψ de type Lévy-Khintchine. Considérons qu'au cours du temps, des individus rejoignent cette population selon une dynamique prescrite par un subordinateur d'exposant ϕ . La population totale constitue un processus de branchement avec immigration, dit $\text{CBI}(\psi, \phi)$. Le travail présenté ici a été réalisé avec Clément Foucart (Paris XIII) et Chunhua Ma (Nankai University). Pour $(Z_t, t \geq 0)$ un $\text{CBI}(\psi, \phi)$ partant de $x \in (0, \infty)$, on étudie pour $a \in [0, x]$, le temps d'entrée $\sigma_a = \inf\{t > 0 : Z_t \in [0, a]\}$. On déduit notamment de cette étude un critère de récurrence/transience pour le $\text{CBI}(\psi, \phi)$.

Frédéric EYI-MINKO

Processus max-stables stationnaires markoviens

Nous donnons une caractérisation des processus max-stables stationnaires markoviens : nous prouvons que cette classe de processus quitte à renverser le temps est égale à la classe des processus stationnaires max-autorégressifs d'ordre 1.

Coralie FRITSCH

Modèle individu-centré de chemostat

Le chemostat est un bioréacteur alimenté en continu dans lequel des bactéries dégradent un substrat. Les dynamiques de populations bactériennes sont généralement modélisées à l'aide d'équations différentielles. Or dans certaines situations, l'aléa ne peut être négligé, on doit alors avoir recours à d'autres modèles. Je présenterai un modèle individu-centré de chemostat dans lequel la population d'individus sera structurée en masse. Le modèle décrit des mécanismes discrets (division des individus et soutirage) et des mécanismes continus (croissance des individus et consommation du substrat). Ce modèle stochastique peut être comparé à des modèles déterministes, notamment à une équation intégro-différentielle qui décrit également la distribution en masse des individus, et à un système d'équations différentielles qui décrit uniquement les dynamiques des concentrations en biomasse et en substrat.

Pierre GAILLARD

Agrégation séquentielle d'experts avec application à la prévision de consommation électrique

Je m'intéresse dans cet exposé à prévoir séquentiellement une suite d'observations. À chaque instant, un nombre fini d'experts me propose des prévisions de la prochaine observation. Je forme alors ma prévision en mélangeant celles des experts. C'est le cadre de l'agrégation séquentielle d'experts. L'objectif classique est d'assurer un faible regret cumulé. En d'autres mots, je souhaite que ma perte cumulée ne dépasse pas trop celle du meilleur expert. Je donnerai une stratégie garantissant une borne déterministe du second ordre sur le regret cumulé. J'expliquerai l'intérêt de cette borne, par exemple dans le contexte d'experts spécialisés. Je finirai par une application à la prévision de consommation électrique.

Davide GIRAUDO

Approximation par ortho-martingales et principe d'invariance pour des champs aléatoires

On donne une condition suffisante sur une classe de champs aléatoires réels strictement stationnaire admettant un moment d'ordre deux pour garantir un principe d'invariance faible. Cette classe contient les champs aléatoires linéaires causaux avec des innovations indépendantes identiquement distribuées. Le résultat se base sur une approximation par ortho-martingales.

Renan GOBARD

Modèle de boules aléatoires avec dépendance

Nous proposons ici un modèle de boules aléatoires pondérées dans \mathbb{R}^d générées selon un processus ponctuel de Poisson, dont les centres et les rayons sont dépendants. On étudiera le comportement asymptotique de la masse totale générée par la configuration de boules à l'échelle macroscopique. Les modèles de boules aléatoires sont utilisés aussi bien en télécommunication pour modéliser des réseaux d'antennes, qu'en imagerie médicale.

David GODINHO PEREIRA

Propagation du chaos pour un modèle de Keller-Segel sous-critique

Nous allons étudier un modèle de Keller-Segel sous-critique. En partant d'un certain système de particules, dont nous montrons la tension, nous obtenons des résultats d'existence et d'unicité pour l'équation de Keller-Segel et son EDS associée, ce qui nous permet d'obtenir la propagation du chaos.

Matthias GORNY

Un modèle de Curie-Weiss de criticalité auto-organisée

Dans leur célèbre article de 1987, les physiciens Per Bak, Chao Tang et Kurt Wiesenfeld ont montré que certains systèmes complexes, composés d'un nombre important

d'éléments en interaction dynamique, évoluent vers un état critique, sans intervention extérieure. Ce phénomène, appelé criticalité auto-organisée (self-organized criticality, SOC, en anglais), peut être observé empiriquement ou simulé par ordinateur pour de nombreux modèles. Cependant leur analyse mathématique est très ardue. Même des modèles dont la définition est apparemment simple, comme les modèles décrivant la dynamique d'un tas de sable, ne sont pas bien compris mathématiquement. Après avoir introduit plus longuement cette notion, je m'inspirerai du modèle d'Ising Curie-Weiss pour construire un modèle probabiliste "simple" de particules en interaction présentant le phénomène de SOC. J'appuierai cette construction par des théorèmes limites.

Lorick HUANG

Density Bounds for some Degenerate Stable Driven SDEs

We consider a stable driven degenerate stochastic differential equation, whose coefficients satisfy a kind of weak Hörmander condition. Under mild smoothness assumptions we prove the uniqueness of the martingale problem for the associated generator. Also, in the scalar case we establish density bounds reflecting the multi-scale behavior of the process.

Émilien JOLY

Minimisation du risque empirique pour des fonctions de perte à queue lourde

Dans deux articles récents dus à [Minsker '13] et [Hsu & Sabado '13] des propositions de procédures d'estimateurs robustes dans des espaces de Banach sont avancées pour résoudre des problèmes de minimisation du risque empirique. Ils se basent tous deux sur une technique connue sous le nom de médiane des moyennes. En parallèle, [Catoni '12] propose une procédure légèrement différente et dont la simplicité de définition permet une adaptation à un cadre de minimisation empirique sur une classe de fonctions de perte très générale. L'erreur commise est à queue sous-gaussienne sous les hypothèses (relativement) faibles d'une variance uniformément bornée sur toute la classe de fonctions et d'une entropie pour la norme L^2 finie sur cette même classe. Nous verrons ensuite comment s'applique cette procédure pour deux types de pertes bien connues : la perte L^1 et la distorsion pour le problème des k -moyennes.

Daniel KIOUS

Stuck Walks : une conjecture d'Erschler, Tóth et Werner

En 2010, A. Erschler, B. Tóth et W. Werner ont défini, dans *Stuck walks*, une classe de marches aléatoires inter-agissantes sur \mathbb{Z} , dans lesquelles apparaît une compétition entre répulsion à petite distance et attraction à plus grande distance. Ils ont prouvé que, pour tout entier $L \geq 1$, si le paramètre α appartient à un certain intervalle (α_{L+1}, α_L) , alors ces marches se localisent sur $L + 2$ sites avec probabilité strictement positive. Ils ont également conjecturé que cela se produit presque sûrement.

Nous démontrons partiellement cette conjecture, en prouvant que la marche, sous les mêmes hypothèses, se localise sur $L + 2$ ou $L + 3$ sites presque sûrement. Nous prouvons également que, si $\alpha > 1$, alors la marche se localise p.s. sur 3 sommets.

Kevin KUOCH

Un processus de contact et ralentissements

On introduit une généralisation du processus de contact où les naissances des individus sont aléatoirement ralenties ou bloquées à taux r . On exhibe d'abord une transition de phase selon ce paramètre de ralentissement, au sens de la survie/extinction de la population. En particulier, le processus critique s'éteint. Puis, en ajoutant un phénomène de diffusion, on obtient une limite hydrodynamique dont la limite est un système de réaction-diffusion.

Jean-Maxime LE COUSIN

Limite d'échelle d'un processus de feu de forêt

On étudie une adaptation unidimensionnelle d'un processus introduit dans les années 80 : le processus de feu de forêt. Sur chaque site de \mathbb{Z} , des graines tombent selon un processus de Poisson de paramètre 1 et des allumettes tombent selon un processus de Poisson de paramètre λ . On suit les règles suivantes :

- si une graine tombe sur un site vide, un arbre pousse instantanément. On appelle *forêts* les composantes connexes. Si une graine tombe sur un site occupé, rien ne se passe ;
- si une allumette tombe sur un arbre, un feu démarre. Après un temps exponentiel de paramètre π , il se propage aux deux voisins et le site redevient vide. Si une allumette tombe sur un site vide ou si un feu se propage sur un site vide, rien ne se passe.

Chaque site possède trois états : vide (position 0), occupé (position 1) et en feu (position 2). On s'intéresse au comportement limite du processus (allure du processus, taille des "forêts") lorsque $\lambda \rightarrow 0$ et $\pi \rightarrow \infty$. En l'absence de feu (cas limite $\lambda = 0$), il s'agit simplement d'un processus de croissance. Après un changement d'échelle (accélération du temps/dilatation de l'espace), on distinguera trois comportements (ou régimes) limites.

Sarah LEMLER

High-dimensional estimation of counting process intensities

We consider the problem of obtaining a prognostic on the survival time adjusted on high-dimensional covariates. Towards this end, we consider two different approaches. First, we propose the construction of an estimator of the general conditional intensity function. We estimate it by the best Cox proportional hazards model given two dictionaries of functions. The first dictionary is used to construct an approximation of the logarithm of the baseline hazard function and the second to approximate

the relative risk. As we are in high-dimension, we consider the Lasso procedure to estimate the unknown parameters of the best Cox model approximating the conditional intensity. We provide non-asymptotic oracle inequalities for the Lasso estimator of the conditional intensity. Our results rely on an empirical Bernstein's inequality for martingales with jumps. Then, in a second part, we consider an intensity that relies on the Cox model. In this Cox model, two parameters are unknown : the baseline function, function of the time and the regression parameter connected to the covariates. We propose a two-step procedure to estimate the intensity : a Lasso procedure to estimate the regression parameter in high-dimension and then, a model selection procedure to estimate the baseline function. We establish a non-asymptotic oracle inequality on the baseline function.

Justine LEQUESNE

Tests d'adéquation basés sur l'entropie

Les tests d'adéquation basés sur l'entropie (E-tests) ont été introduits par Vasicek (1976). D'autre part, des tests basés sur l'entropie relative (ou divergence de Kullback-Leibler) ont été construits dans Song (2002).

Dans cet exposé, nous comparerons ces deux tests, dans le but d'ajuster la loi de Pareto à des données réelles. Nous justifierons mathématiquement des E-tests par une égalité de Pythagore appliquée à des lois de maximum d'entropie. Nous montrerons l'équivalence des deux tests en pratique dans le cas d'une hypothèse nulle composite.

Une extension possible à des entropies généralisées telles que les divergences de Rényi ou Tsallis sera considérée.

Références :

Vasicek, O. (1976). A test for normality based on sample entropy. *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol 38, No. 1 pp 54-59

Song, K.S. (2002). Goodness-of-fit tests based on Kullback-Leibler discrimination information. *IEEE Transactions on Information Theory*, Vol 48 pp 1103-1117

Girardin, V. & Lequesne, J. (2013). Relative entropy versus entropy difference in goodness-of-fit tests. Application to Pareto fitting. *preprint*

Lequesne, J. (2012). Entropy-based goodness-of-fit test : Application to the Pareto distribution. *AIP Conference Proceedings 1553* pp 155-162

Hélène LESCORNEL

Estimation of deformations between distributions by minimal Wasserstein distance

We consider that we observe a random variable ε and a deformation of this variable, denoted by X . We consider a semi-parametric model for the deformation since it is modelled through a function φ_θ . The shape of the deformation φ is known, but not its amount, represented by a deformation parameter θ . We will propose an estimator for this parameter, which enables us to estimate the law of ε by using the observations of X . For this, we will align the distribution of X onto the distribution of ε by using their

Wasserstein distance. We will propose an alternative procedure in the case where we observe several deformations in general spaces.

Julien LETEMPLIER

Unimodalité des temps d'atteinte pour des processus de Lévy stables

Nous nous intéressons dans cet exposé à l'étude de l'unimodalité de certaines variables aléatoires, c'est à dire au fait que celles-ci puissent être de fonction de répartition convexe puis concave. L'unimodalité des lois stables à été prouvée par Yamazato(78), et nous démontrerons que cette propriété est également partagée par les temps d'atteinte de processus de Lévy stables. Après quelques définitions et remarques sur l'unimodalité, nous établirons quelques égalités en loi vérifiées par notre temps d'atteinte, via le calcul de transformée de Mellin, à partir desquelles nous obtiendrons la propriété d'unimodalité recherchée.

Viviana LETIZIA

Fourier's law for the discrete Schrödinger equation

We consider the one-dimensional discrete linear Schrödinger (DLS) equation perturbed by a conservative stochastic dynamics, that changes the phase of each particle, conserving the total norm (or number of particles). The resulting total dynamics is a degenerate hypoelliptic diffusion with a smooth stationary state. We will show that the system has a hydrodynamical limit given by the solution of the heat equation. When it is coupled at the boundaries to two Langevin thermostats at two different chemical potentials, we prove that the stationary state, in the limit as $N \rightarrow \infty$, satisfies the Fourier's law.

Clément LEVRARD

Conditions de marge pour la quantification

Le point de vue adopté durant cette présentation est celui de la classification non supervisée : étant donné une distribution de probabilité sur un espace de dimension finie, le but est de déterminer une suite finie de points représentant au mieux cette distribution. N'ayant accès qu'à un échantillon tiré suivant la distribution source, la stratégie adoptée est celle de la minimisation d'un critère empirique de type moindre carré, dont le pendant implémenté est la procédure des k means introduite par Hartigan. En inscrivant le problème théorique des k means dans le contexte plus général des procédures de minimisation de fonctions de contraste, certaines conditions apparaissent comme favorables pour la vitesse de convergence de cette procédure en la taille de l'échantillon. Dans un premier temps, nous montrerons que ces conditions peuvent être vues comme l'analogie des conditions de marge introduites par Mammen et Tsybakov dans le domaine voisin de la classification supervisée. Nous discuterons ensuite de l'influence des autres paramètres introduits par ces conditions de marge sur la vitesse de convergence au travers d'inégalités oracles et de résultats de type minimax.

Shen LIN*Marche aléatoire indexée par un arbre et son nombre de points visités*

Considérons une marche aléatoire simple dans \mathbb{Z}^d indexée par un arbre aléatoire choisi uniformément au hasard dans l'ensemble des arbres planaires de n sommets, et soit R_n le nombre de points visités par cette marche. On montre que, si $d \geq 5$, $n^{-1}R_n$ converge vers une constante strictement positive, alors que lorsque $d = 4$, $n^{-1}R_n$ converge vers $\pi^2/2$. En petite dimension $d \leq 3$, on prouve que $n^{-d/4}R_n$ converge en loi vers la mesure de Lebesgue du support d'une mesure aléatoire \mathcal{S} sur \mathbb{R}^d , qui se définit comme la mesure d'occupation d'un serpent brownien d -dimensionnel dirigé par une excursion brownienne normalisée. Les résultats analogues sont obtenus pour des marches aléatoires plus générales et pour des structures de branchement plus générales. Ce travail est une collaboration avec Jean-François Le Gall.

Pierre-Yves MADEC*EDSR ergodiques et EDPs avec conditions de Neumann au bord dans un environnement faiblement dissipatif*

On étudie une classe d'EDSR ergodiques en lien avec les EDPs avec conditions de Neumann au bord. L'aléa du générateur de l'EDSR ergodique est donné par un processus faiblement dissipatif réfléchi et dont la matrice de diffusion est inversible. De plus ce processus est réfléchi dans un convexe de \mathbb{R}^d non nécessairement borné.

Nelo MAGALHAES*Vers une nouvelle validation croisée V-fold*

La validation croisée "V-fold" est une procédure bien connue de sélection de méthodes d'estimation dans laquelle les données sont partagées en V sous-échantillons pour être utilisées tantôt comme données d'apprentissage pour construire des estimateurs, tantôt comme données de test dans une étape de validation –fondée le plus souvent sur la minimisation d'un critère de coût empirique. Dans l'esprit de la procédure de hold-out proposée dans Birgé (2006), nous étudions l'effet produit par l'emploi d'une procédure de test robuste en lieu et place de la minimisation d'un critère de coût empirique dans l'étape de validation. Nous étudierons cette nouvelle procédure dans le cadre de l'estimation de densité sur la base d'observation indépendantes et de même loi.

Bastien MALLEIN*Marches aléatoires branchantes en environnement inhomogène*

Une marche aléatoire branchante est un processus aléatoire qui mime l'évolution au cours du temps d'une population, dans laquelle des individus se déplacent et se reproduisent de façon indépendante. Dans cet exposé, nous supposons que la façon dont les individus se reproduisent évolue lentement au cours du temps. Nous montrerons que dans ce cas, le plus grand déplacement est composé d'un premier ordre

linéaire, solution d'un problème d'optimisation convexe, plus une correction d'ordre $n^{1/3}$.

Sébastien MARTINEAU

Localité de la percolation sur les graphes abéliens

La percolation est l'étude des milieux poreux. Mathématiquement, on représente le milieu non-altéré par un graphe. Afin d'introduire la porosité, on choisit un paramètre $p \in [0, 1]$ et efface chaque arête avec probabilité $1 - p$, les arêtes étant traitées indépendamment les unes des autres. On peut alors définir un paramètre critique, le p où l'on passe du régime où toutes les composantes connexes du graphe sont finies presque sûrement à celui où ce n'est pas le cas. On démontre que, pour les "graphes abéliens de rang 2 ou plus", la valeur de ce paramètre critique ne dépend essentiellement que de la structure locale du graphe, pas de sa géométrie globale. Ce travail a été effectué en collaboration avec Vincent Tassion.

Pierre MONMARCHÉ

Convergence hypocoercive à l'équilibre

La loi d'un processus ergodique converge vers sa mesure invariante ; brique élémentaire de bien des algorithmes stochastiques. Pour quantifier la vitesse de cette convergence, les méthodes des inégalités fonctionnelles sont très efficaces (et plutôt bien comprises) pour des processus dit coercifs (typiquement, des processus réversibles). Pourtant de nombreux exemples non-coercifs présentent tout de même une convergence exponentielle vers l'équilibre, ce qui demande de nouvelles approches.

Peter NEJJAR

Fluctuations anormales du choc dans le processus d'exclusion

On considère le processus d'exclusion totalement asymétrique avec des conditions initiales (CI) et/ou des taux de saut tels que des chocs sont créés. Si la CI est déterministe, la largeur de la région du choc au temps t sera d'ordre $t^{1/3}$. On obtient la limite en temps long de la loi de la position d'une particule dans le choc dans quelques modèles. En particulier, on traite le cas où à gauche et à droite du choc le mouvement des particules est asymptotiquement décrit par un processus d'Airy₁. La distribution limite est donnée par un produit de deux fonctions de répartition, une conséquence du fait que deux lignes caractéristiques se croisent au choc, et que la décorrélation est ralentie le long de ces lignes. On montre que le résultat peut être étendu à un modèle général de percolation de dernier passage. Si le temps le permet, on abordera la relation avec les particules de deuxième classe. Ce travail est en collaboration avec Patrik Ferrari.

Geoffrey NICHIL

Provisionnement d'une compagnie d'assurance face au risque de défaut des emprunteurs

Un assureur doit indemniser une banque suite aux pertes liées aux défauts de paiement de ses emprunteurs. L'objectif pour cet assureur est d'évaluer le montant potentiel des pertes liées à ces défauts de paiement (montant de sinistres) pour une période donnée.

La quantité clé de notre modèle est le montant d'un sinistre. Pour un emprunteur donné et une date de défaut fixée, le montant de sinistres est égal à $\max(S_T^1 - S_T^2; 0)$, où S_T^1 est le montant dû par l'emprunteur et S_T^2 est le montant de la revente du bien immobilier financé par le prêt. Le montant S_T^1 est proportionnel au montant emprunté et le coefficient de proportionnalité dépend de la durée du prêt et de la date du défaut. La quantité S_T^2 est proportionnelle au montant emprunté; le coefficient de proportionnalité est modélisé par un mouvement Brownien géométrique et représente les fluctuations des prix de l'immobilier et l'apport personnel de l'emprunteur. On suppose que ce coefficient est identique pour tous les emprunteurs, ainsi il existe une dépendance entre les montants de sinistres de chaque emprunteur. La loi des couples (Date de fin du prêt, Durée du prêt) sera modélisée par un processus ponctuel de Poisson ce qui nous permettra de définir la provision comme une somme d'un nombre aléatoire de montants de sinistres individuel.

Contrairement aux modèles traditionnels (Chain Ladder, Cramer - Lundberg) notre modèle permet de prendre en compte trois types de dépendance : entre les dates des défauts de paiement et leurs montants, entre le montant de sinistres de chaque emprunteur et celle entre le nombre de sinistres et les montants de sinistres. A la fois théorique et appliquée cette modélisation permet de calculer l'espérance et la variance mais aussi de donner un algorithme de simulation de la provision. Le nombre important de données à notre disposition permet d'estimer les paramètres liés aux modèles et de fournir une valeur numérique à l'espérance, la variance et aux quantiles de la provision.

Adélaïde OLIVIER

Estimation non-paramétrique du taux de saut d'un processus de Markov déterministe par morceaux sur un arbre

Il s'agit d'étudier un système de particules. Chaque particule présente une caractéristique individuelle (sa taille, son âge, *etc.*) à laquelle on s'intéresse plus particulièrement. Deux phénomènes régissent l'évolution de ce système. D'une part, au cours de sa vie, une particule évolue de manière déterministe connue (la particule croît, vieillit, *etc.*). D'autre part, les particules se divisent aléatoirement : une particule de caractéristique x se divise en deux nouvelles particules (de taille initiale $x/2$, d'âge 0, *etc.*) selon un taux instantané $B(x)$ où B est la fonction inconnue à reconstituer. Dans ce cadre, différents schémas d'observation sont envisageables pour traiter le problème de l'estimation non-paramétrique de B . 1) Il est possible d'observer les

particules jusqu'à une génération donnée dans l'arbre généalogique des particules. Dans le modèle en taille par exemple, un estimateur adaptatif convergeant aux vitesses non-paramétriques classiques peut alors être construit. 2) Il est possible d'observer le système entre la date 0 et une date T. Des difficultés sont intrinsèquement liées à ce second schéma d'observation. Dans le cadre plus simple du modèle en âge, un estimateur non-adaptatif a été construit et sa vitesse de convergence étudiée.

Christophe POQUET

Synchronisation et comportement en temps long de rotateurs bruités en interaction

Nous nous intéresserons à un modèle classique de rotateurs bruités en interaction de type champ moyen. Sur des horizons de temps finis, ce modèle est décrit à la limite d'une infinité de rotateurs par une EDP de type Fokker-Planck. Ce modèle limite admet une transition de phase : lorsque l'interaction est assez forte l'EDP admet une famille de profils stationnaires synchronisés, qui constitue un attracteur du modèle. Nous verrons que l'approximation du système de taille finie par cette EDP n'est plus valable en temps longs (proportionnels au nombre de rotateurs), mais que sur cette échelle de temps la mesure empirique du système prend à la limite la forme d'un profil synchronisé dont le centre de synchronisation décrit une trajectoire aléatoire.

Hélène QUINTARD

Symétries de l'équation de la chaleur rétrograde avec potentiel et modèles de taux d'intérêt

Utilisation des isovecteurs pour l'équation de la chaleur éventuellement rétrograde avec potentiel V : $\gamma \frac{\partial \eta}{\partial t} = \frac{\gamma^2}{2} \frac{\partial^2 \eta}{\partial q^2} + V\eta$ dans le cadre de la résolution d'équations différentielles stochastiques et particulièrement en vue d'une application à certains modèles de taux d'intérêts.

Julien REYGNER

Comportement en temps long de diffusions interagissant à travers leur rang

On s'intéresse à des systèmes de particules browniennes évoluant sur la droite réelle, avec une dérive et une variance ne dépendant que de leur rang. De tels systèmes interviennent en particulier dans des modèles de marchés financiers, en mécanique statistique, et en théorie des files d'attente. On donnera un aperçu des résultats actuels sur le comportement en temps long de ces systèmes, ainsi que de leur limite de champ moyen.

Alexandre RICHARD

Wiener spaces and the sample paths properties of multiparameter fractional Brownian processes

In this talk, we will present briefly the notion of abstract Wiener space, in order to use it in the study of a certain class of fractional Brownian fields. In particular we will

present some results concerning the small ball probabilities of these fields as well as some regularity properties (e.g. Hölder regularity and modulus of continuity).

Angelina ROCHE

Estimation adaptative de la fonction de répartition conditionnellement à une covariable fonctionnelle

Ce travail est en collaboration avec Gaëlle Chagny.

Dans cet exposé, nous présentons une procédure d'estimation adaptative pour la fonction de répartition conditionnelle définie par

$$F^x(y) := \mathbf{P}(Y \leq y | X = x),$$

lorsque la covariable est fonctionnelle c'est-à-dire lorsque X est une variable aléatoire à valeur dans un espace de fonctions, que nous supposons être un espace de Hilbert séparable (c'est le cas par exemple de $\mathbf{L}^2(I)$ où I est un intervalle de \mathbf{R}).

Nous disposons pour l'estimation de F^x d'un échantillon $\{(X_i, Y_i), i = 1, \dots, n\}$ de copies du couple (X, Y) . Nous nous intéresserons ici plus particulièrement à l'estimateur de type Nadaraya-Watson proposé par Ferraty *et al.* (2006, 2010) :

$$\hat{F}_h^x(y) := \frac{\sum_{i=1}^n \mathbf{1}_{\{Y_i \leq y\}} K(\|X_i - x\|/h)}{\sum_{i=1}^n K(\|X_i - x\|/h)},$$

où K est un noyau (une application intégrable $K : \mathbf{R} \rightarrow \mathbf{R}_+$ vérifiant $\int_{\mathbf{R}} K(u) du = 0$). Nous étudierons dans un premier temps, d'un point de vue non-asymptotique, le risque de cet estimateur lorsque la fenêtre h est fixée. Nous verrons que le choix de la fenêtre h joue un rôle particulièrement important et que sa valeur optimale dépend de quantités inconnues en pratique. Nous proposerons une procédure de sélection de la fenêtre inspirée des récents travaux de Goldenshluger et Lepski (2011). L'estimateur ainsi obtenu atteint le meilleur compromis biais-variance et la vitesse de convergence minimax sous diverses hypothèses portant sur la loi de X .

Arnaud ROUSSELLE

Marches au hasard sur des graphes aléatoires engendrés par des processus ponctuels dans \mathbb{R}^d

Dans cet exposé, on considère des marches au hasard au plus proche voisin, sur des graphes construits géométriquement à partir de réalisations de processus ponctuels simples dans \mathbb{R}^d (par exemple la triangulation de Delaunay ou le graphe de Gabriel engendré par un processus ponctuel de Poisson). On présentera des résultats de récurrence/transience pour ces modèles, ainsi qu'un principe d'invariance.

Jean-Bernard SALOMOND

Adaptive Bayes test for monotonicity

We propose a Bayesian non parametric approach to test for monotonicity in a regression setting. In that context, the usual Bayes factor approach gives poor results in

practice. We thus study an alternative approach that is both efficient and straightforward to implement, which is a great improvement compared to the existing frequentists procedures. Furthermore we study its asymptotic properties and prove that our procedure attains the adaptive minimax separation rate for a wide variety Hölder smooth alternatives.

Marielle SIMON

Diffusion de l'énergie pour un système de particules en interaction

L'apparition de phénomènes diffusifs pour des systèmes macroscopiques est vue comme le résultat d'une dynamique microscopique suffisamment "chaotique". Pour cette raison, la dynamique Hamiltonienne classique du système de particules est perturbée par un bruit stochastique, et s'effectue dans un environnement aléatoire. Après changement d'échelle (si la distance entre deux particules est d'ordre N^{-1} , on accélère le temps par N^2) on observe un phénomène de diffusion. Plus précisément, les fluctuations de l'énergie évoluent suivant une EDS infini-dimensionnelle induite par l'équation de la chaleur linéaire.

Robin STEPHENSON

Scaling limits of k -ary growing trees

For each integer $k \geq 2$, we introduce a sequence of k -ary discrete trees constructed recursively by choosing at each step an edge uniformly among the present edges and grafting on "its middle" $k - 1$ new edges. When $k = 2$, this corresponds to a well-known algorithm which was first introduced by Rémy. Our main result concerns the asymptotic behavior of these trees as n becomes large : for all k , the sequence of k -ary trees grows at speed $n^{1/k}$ towards a k -ary random real tree that belongs to the family of self-similar fragmentation trees. This convergence is proved with respect to the Gromov-Hausdorff-Prokhorov topology. We also study embeddings of the limiting trees when k varies.

Vincent TASSION

Transition de phase du modèle de percolation FK planaire

La percolation FK est définie par un sous-graphe aléatoire $G_{p,q}$ du réseau carré. La loi de ce sous-graphe dépend de deux paramètres $p \in [0, 1]$ et $q \geq 1$. Le cas $q = 1$ correspond à la percolation indépendante : $G_{p,1}$ est obtenu à partir de \mathbb{Z}^2 en enlevant indépendamment chaque arête avec probabilité $1 - p$. Lorsque q augmente, le graphe $G_{p,q}$ cherchera à privilégier les configurations avec un maximum de composantes connexes. Pour q fixé, le modèle présente une transition de phase en p . Lorsque p est suffisamment petit, $G_{p,q}$ n'est formé (p.s.) que de composantes connexes finie, et dès que p devient strictement supérieur à une valeur critique $p_c(q)$, $G_{p,q}$ contient alors une composante connexe infinie. Dans cet exposé, nous étudierons le cas critique, à $(p_c(q), q)$: y a-t-il une composante connexe infinie ? Réponse : cela dépend de la valeur de q .

Maud THOMAS*Inégalités de concentration pour les statistiques d'ordre*

Le but de cet exposé est de montrer comment obtenir des inégalités de concentration non-asymptotiques pour les statistiques d'ordre d'un échantillon de variables aléatoires indépendantes et identiquement distribuées.

Les résultats de la théorie des valeurs extrêmes permettent de montrer que ces inégalités sont asymptotiquement tendues.

Lorsque la loi de l'échantillon possède un taux de hasard croissant, on montre que les statistiques d'ordre satisfont une inégalité exponentielle d'Efron-Stein, c'est-à-dire, une inégalité reliant le logarithme de la fonction génératrice des moments des statistiques d'ordre avec les moments exponentiels de l'estimée Efron-Stein de la variance. Nous commencerons par rappeler certaines inégalités de concentration et introduire quelques notions de la théorie des valeurs extrêmes. Nous montrerons ensuite comment les appliquer pour obtenir des inégalités de concentration pour les statistiques d'ordre et nous vérifierons qu'elle sont tendues grâce à la théorie des valeurs extrêmes. Enfin, nous les appliquerons aux échantillons gaussiens.

Sofia TSEPLETIDOU*Computational Bayesian Tools for Modeling the Aging Process in Escherichia coli*

This research studies the aging process at the bacterium {E. Coli} in a bayesian framework. Modeling appropriately this process, by reconstructing a hidden quantity that explains the physiological characteristics of each cell in the lineage tree, is the first step towards the estimation. The last one is possible through exploration of the posterior distribution of the constructed model. To this purpose, firstly, an Approximate Bayesian Computation methodology has been considered. Later, Monte Carlo Markov Chains methods, a Gibbs Sampler, have been also performed. Finally, the results of each approach are discussed, as well as the possible extensions.

Thomas VARESCHI*SVD par blocs et déconvolution semi-aveugle*

Cette présentation aura pour principal sujet le contrôle de l'imprécision sur un opérateur intervenant dans la résolution d'un problème inverse. La problématique habituelle des problèmes inverses en statistique est l'approximation d'un signal d'entrée à partir de son image par un opérateur régularisant. A l'incertitude habituelle contaminant l'observation du signal de sortie, on ajoute cette erreur commise sur l'opérateur que l'on modélise par un processus Gaussien d'une certaine amplitude, potentiellement différente de la précédente. Nous nous intéressons plus particulièrement au cas où l'opérateur en question est un opérateur à noyau, lorsque ce dernier est lui même bruité. Ce modèle recouvre par exemple les cas de la convolution de Fourier périodique, ou bien la convolution sphérique. Après un bref passage en revue des travaux antérieurs, je présenterai des procédures statistiques d'estimation dans chacun

de ces cas, en traitant de manière adéquate la nouvelle erreur commise sur le noyau selon la forme de la matrice associée à un schéma de Galerkin.

Elodie VERNET

Consistance des modèles non paramétriques bayésiens de Markov cachés à espace d'états fini

Les modèles de Markov cachés sont très utilisés en pratique, comme en génomique, reconnaissance de parole. . . Or la théorie est mal comprise. Lors de cet exposé, j'étudierai le comportement asymptotique de l'a posteriori dans ces modèles. Je commencerai par introduire les modèles bayésiens et je donnerai des comportements asymptotiques connus. J'introduirai ensuite les modèles de Markov cachés. Enfin, je présenterai un premier résultat asymptotique dans le cadre des chaînes de Markov cachées non paramétriques.

Raghid ZEINEDDINE

Une formule de type Ito pour le mouvement brownien fractionnaire en temps brownien

Dans un article écrit avec mon encadrant Ivan Nourdin, on a trouvé une formule de type Itô pour le mouvement brownien fractionnaire en temps brownien. Comme son nom l'indique, ce processus est un mouvement brownien fractionnaire indexé par un mouvement brownien indépendant au lieu d'être indexé par le temps. La formule obtenue admet deux formes possibles selon que $H :=$ l'indice d'Hurst du mouvement brownien fractionnaire est strictement plus grand que $1/6$ ou égale à $1/6$. Pour $H < 1/6$, il n'y a pas en générale une formule, on a donné un contre exemple.

Rania ZGHEIB

Test d'adéquation pour grandes matrices de covariance

Nous observons X_1, \dots, X_n , n variables aléatoires indépendantes, à valeurs dans \mathbb{R}^p pour $p \in \mathbb{N}^*$, suivant la même loi gaussienne $N_p(0, \Sigma)$. A partir de cet échantillon nous voulons tester $H_0 : \Sigma = Id$, contre une alternative non paramétrique $H_1 : \Sigma \in \mathcal{F}(\alpha, L)$ et $\|\Sigma - Id\|_F^2/p \geq \varphi^2$. L'écart entre une matrice Σ et l'identité est mesuré en norme de Frobenius et la classe $\mathcal{F}(\alpha, L)$ contient les matrices de covariance de diagonale 1, telle que $(1/p) \sum_{i=1}^p \sum_{\substack{j=1 \\ j>i}}^p \sigma_{ij}^2 |i - j|^{2\alpha} \leq L$. L'estimation des grandes matrices de

covariance a été souvent considérée sous l'aspect minimax. En revanche, le problème de test a été considéré jusqu'ici seulement par Cai et Ma 2013 pour l'alternative $H_1 : \|\Sigma - I\|_F^2/p \geq \varphi^2$. Nous proposons une procédure de test et calculons la borne asymptotique de séparation φ^2 pour lesquelles le risque maximal de test tend vers 0, quand n et p tendent vers l'infini. Dans le cas où Σ est Toeplitz et $n = 1$ ces vitesses sont connues minimax exactes (Ermakov 1994).