Contributions à l'étude de modèles biologiques, d'inégalités fonctionnelles, et de matrices aléatoires

Djalil Chafai

Université Paul Sabatier – Toulouse III

Institut de Mathématiques de Toulouse

École Doctorale de Mathématiques

Contributions à l'étude de modèles biologiques, d'inégalités fonctionnelles, et de matrices aléatoires

Document de synthèse présenté pour l'obtention d'une

Habilitation à Diriger des Recherches

par Djalil Chafaï

14 octobre 2008

Neil O'Connell	Université de Warwick	(rapporteur)
Persi Diaconis	Université Stanford	(rapporteur)
Gábor Lugosi	Université Pompeu Fabra, Barcelone	(rapporteur)
Jean-Marc Azaïs	Université de Toulouse	(examinateur)
Dominique Bakry	Université de Toulouse	(coordinateur)
Michel Ledoux	Université de Toulouse	(examinateur)
Sylvie Méléard	École Polytechnique, Palaiseau	(présidente)
Pierre Del Moral	Université de Bordeaux & INRIA	(examinateur)

À la fée curiosité et à sa sœur jumelle dispersion. . .

Presque rien sur presque rien. . .

Avant propos

Ce mémoire fait brièvement le point sur ma modeste activité scientifique depuis ma thèse. Au delà des mathématiques elles-mêmes, la thèse avait été l'occasion de découvrir un métier, d'apprendre l'autonomie, d'apprendre également à vivre avec ses doutes et ses limites, en côtoyant une communauté et quelques maîtres comme Michel Ledoux et Dominique Bakry. Cette époque des grands apprentissages me semble à présent assez lointaine, bien que je sois perpétuellement à la recherche de nouvelles remises en question et relativisations, pour mieux progresser. Comme le dit et le pratique si bien Fabrice Baudoin, rien ne vaut l'action, ce qui n'exclue pas la réflexion. Il faut dire que je suis un (ancien) timide, et que la confiance en soi m'a souvent manqué dans mes plus jeunes années. De ce point de vue, le temps m'a apporté un peu d'apaisement et de sérénité, sans toutefois émousser mon enthousiasme.

Je tiens à remercier vivement Gabor Lugosi, Neil O'Connell, et Persi Diaconis d'avoir consacré un peu de leur temps pour rédiger les rapports sur ce mémoire d'habilitation. Leur qualités scientifiques et humaines sont exemplaires. J'aimerais également remercier les trois toulousains Jean-Marc Azaïs, Dominique Bakry, et Michel Ledoux, qui, par leur présence dans ce jury, me font prendre conscience du poids des ans… J'aimerais enfin remercier Pierre Del Moral et Sylvie Méléard, dont je ne cesse d'apprécier l'enthousiasme communicatif et le dynamisme renouvelé.

J'ai eu la chance de passer une année postdoctorale dans le laboratoire de Terry Lyons à Oxford. Terry est un personnage passionné, au tempérament bouillonnant, et au français savoureusement anglais. J'ai apprécié sa puissance intellectuelle, toujours prête à tenter de digérer de nouvelles questions quel qu'en soit la nature. Voilà un mathématicien exigeant et de bon niveau qui n'hésite pas à s'intéresser à la fois à des questions théoriques ardues et à des problèmes concrets, à expérimenter ses idées en écrivant des programmes et à discuter avec conviction des tares de la dernière norme du C++. Je me souviens avec émotion de nos batailles sur les mérites et inconvénients des logiciels libres, sur la meilleure façon de calculer l'exponentielle de matrices, les bases de Hall, les coefficients de la formule de Campbell-Baker-Hausdorff, ou la signature en intégrales itérées de chemins multivariés. J'ai progressé pendant cette année, à son contact, mais aussi en fréquentant les séminaires, et en côtoyant avec enthousiasme d'autres chercheurs de divers horizons comme les probabilistes Andreas Eberle, Tom Fawcett, Ben Hambly, et Nicolas Victoir, le logicien Xavier Vidaux, et les géomètres Stelios Koundouros et Xiang-Dong Li, dont les destins sont si différents. Je me souviens encore avec sourire des *tea time* avec l'équipe de logique du Professeur Boris Zilber, ainsi que de la *high table* de Wooster College. Terry m'a dit un jour que « *dans le monde actuel, la curiosité et l'éclectisme sont rarement rentables à court terme et un peu moins rarement à long terme* ». J'ai souvent médité cette cruelle réalité dans mes moments de doute. J'aimerais également saluer Pierre Fougères et Bogusław Zegarliński, qui m'ont gentiment accueilli à cette époque à Londres lors de visites pluvieuses ou ensoleillées, à Imperial College, pour parler d'énergie et d'entropie, deux vieilles amies communes.

J'ai parfois repoussé les choix trajectoriels canoniques que les systèmes imposent aux individus, en partie par besoin de liberté. À mon arrivée, en tant que chargé de recherches, dans l'unité INRA dirigée par Pierre-Louis Toutain, ce dernier m'a tenu en substance ce discours : « *Monsieur Chafaaaille, vous êtes maintenant un mathématicien dans un laboratoire de biologie, aussi il vous faut maintenir une activité et une reconnaissance dans votre discipline. Ne coupez pas les ponts avec votre communauté et l'Université. Mais de temps à autre, intéressez vous à nos problèmes et à nos modèles. C'est pour cela qu'on vous a recruté. N'oubliez pas aussi de ne pas vous endormir sur vos lauriers* ». Cette tirade, toute toutinienne et pleine de bon sens, correspond bien à ma vision des choses. Aussi, j'ai tenté ces dernières années de maintenir, autant qu'il m'était possible, un équilibre entre activités théoriques et appliquées. J'apprécie toujours autant le contact avec la nouveauté, apprendre, comprendre, explorer, synthétiser, tenter de mettre à jour l'harmonie tapie sous le désordre apparent. J'ai cru un temps que la curiosité me condamnait inexorablement à creuser une certaine forme d'échec dans la dispersion. Faut-il vraiment choisir entre presque tout sur presque rien et presque rien sur presque tout ? L'éclectisme est-il réservé aux génies universels comme Andrei Kolmogorov, John von Neumann, ou Terence Tao ? Nombreux sont mes camarades qui préfèrent le penser, par commodité. Il m'a semblé malgré tout possible, à ma toute modeste échelle, de faire fructifier la curiosité. J'ai donc suivi mon imagination, en prenant des risques, et en faisant preuve de ténacité. Peu importe la nature des questions intellectuelles que l'on considère, c'est le rapport à ces questions qui compte avant tout. Il faut aussi tenter de terminer ce qui est commencé, ou tout au moins de fabriquer de la récurrence et de faire confiance au temps. J'ai été frappé par exemple de parvenir à résoudre certains petits problèmes mathématiques laissés en sommeil de nombreuses années, par manque d'idée féconde ou de bon point de vue. Chaque mathématicien transporte dans ses pensées une petite collection personnelle de problèmes familiers qu'il réexamine régulièrement, au gré des lectures et des rencontres, au bout des chemins de montagne, ou entre deux avions... J'aimerais souligner que dans mon mode de pensée, il n'y a pas de différence entre appliqué et fondamental. Pour moi, il y a des questions, et des réponses. Si la question est concrète, la réponse doit l'être, mais les questions concrètes peuvent aussi féconder d'autres questions abstraites, et la réciproque est vraie. Peu importe, du reste, je ne retiens de ce jeu intellectuel que ce qui me semble subjectivement pertinent, surprenant, ou esthétique. Hélas, il m'arrive régulièrement de constater, piteusement, qu'aucun de ces trois critères n'est vérifié ! L'entropie et l'énergie sont toujours mes compagnes... Malgré tout, ces petits échecs qui jalonnent tout travail exigeant ne doivent pas décourager, mais plutôt inciter à redoubler d'efforts, à faire preuve d'audace, d'imagination, et à rechercher ses limites.

Ces quelques années passées dans un laboratoire de biologie de l'INRA m'ont beaucoup apporté sur bien des plans. Aussi, j'aimerais remercier sincèrement Pierre-Louis Toutain, Alain Bousquet-Mélou, et Didier Concordet pour la confiance qu'ils m'ont toujours accordé. J'aimerais également saluer plus généralement toute l'équipe pour la bonne ambiance qui y règne et les échanges enrichissants tissés au fil du temps. Je ne dirais jamais assez combien j'apprécie les discussions passionnées avec Didier Concordet, à la recherche d'un bout de vérité, où les egos sont mis entre parenthèses, et où l'humour est toujours au rendez-vous. Les années d'après thèse sont également l'occasion de se lancer dans la formation doctorale de plus jeunes que soi. J'ai ainsi entrepris le co-encadrement de la thèse de Julie Antic, à qui je souhaite de poursuivre son épanouissement dans ces métiers si passionnants de la science appliquée en prise avec l'industrie et la société.

Mon activité scientifique a également bénéficié de la proximité de l'Institut de Mathématiques de Toulouse. J'ai ainsi pu interagir au gré de mes visites hebdomadaires avec des chercheurs fort différents. Je pense par exemple à Dominique Bakry, Franck Barthe, Fabrice Baudoin, Bernard Bercu, Charles Bordenave, Mireille Capitaine, Muriel Casalis, Patrick Cattiaux, Komla Domelevo, Béatrice Laurent, Michel Ledoux, et Jean-Michel Loubes, et à ceux juste de passage comme Pietro Caputo et Neil O'Connell. J'ai une pensée spéciale pour Laurent Miclo, dont l'âme profondément probabiliste est tout à fait remarquable. Je ne résiste pas également à mentionner Gérard Letac l'ancien et son fameux *« côm-ment â-llez-vous, mon gâr-çon ? »* ainsi que les multiples déménagements de Xiang-Dong Li, qui se terminaient parfois à Ramonville...

J'aimerais aussi profiter de ces lignes pour encourager vivement les « jeunes » chercheuses et chercheurs que j'ai côtoyés plus récemment ici et là. Je pense par exemple à Julie Antic, Anne-Laure Basdevant, Michel Bonnefont, Aude Ferran, Nolwen Huet, Aldéric Joulin, Agnès Lagnoux, Sébastien Motsch, et Clément Rau. Je leur souhaite un bel épanouissement dans leur métier, sans oublier toutefois que la vie est malgré tout ailleurs !

J'ai pris beaucoup de plaisir à enseigner pendant plusieurs années la modélisation stochastique aux agrégatifs, d'abord seul, puis aux côtés de Bernard Bercu. Fabrice Gamboa nous a encouragé à transformer en livre le manuel polycopié que j'avais rédigé à l'époque. C'est ainsi que j'ai entrepris avec Bernard l'écriture d'un livre sur la modélisation et la simulation. Notre projet n'en finissait pas de mûrir, et nous avons dû chasser le doute à plusieurs reprises. J'éprouvais malgré tout un certain plaisir à remettre sur l'établi chaque chapitre, mois après mois. Notre aboutissement au printemps 2007 me restera comme une victoire de l'énergie sur l'entropie, malgré les inévitables imperfections de toute œuvre humaine. Bravo Bernard d'avoir si bien su faire confiance au temps !

Mon expérience à Météo-France et plus récemment à l'INRA m'incite à penser qu'il ne faut pas opposer science fondamentale et appliquée. Nombre de tensions proviennent du petit nombre de ceux qui tentent de mener ces deux activités de front. Il m'avait fallu, en thèse, admettre que les mathématiciens sont bien plus fragmentés et segmentés que les mathématiques elles-mêmes. Ces quelques années supplémentaires dans le monde de la recherche m'ont conduit à poursuivre cette réflexion sur ces humains qui font la science. Il y a par exemple chez certains mathématiciens, la tentation d'apposer un verni appliqué sans se soucier vraiment des applications réelles, et chez certains scientifiques une recherche de respectabilité par la mathématisation sans se soucier de sa pertinence. Ces deux attitudes sont de plus parfaitement compatibles et peuvent constituer le socle de collaborations assez productives surfant sur les modes du moment. Cette sorte de *mythe du quantitatif* est peut-être le prix à payer pour faire émerger parfois des interactions interdisciplinaires fécondes. Parallèlement, le mathématicien fondamental ordinaire méprise les applications et réduit les mathématiques appliquées à de la technologie ou à de l'ingénierie. Pourtant, on ne se grandit pas en méprisant autrui, mais plutôt en tentant de faire mieux. Au delà de l'utilitarisme primaire, se trouve une attitude qui consiste à toujours s'intéresser à des problèmes concrets, et à ne développer de théorie abstraite qu'au travers de cette motivation. C'est par exemple ce que prône Persi Diaconis. Cela nécessite une certaine ouverture d'esprit, et tente d'éviter le développement d'*abstract non-sense* dans les mathématiques appliquées.

Il faut cependant tenir compte d'aspects sociologiques supplémentaires. Les domaines scientifiques sont si vastes aujourd'hui que les chercheurs sont fortement poussés à l'hyper-spécialisation. Rares sont ceux qui sont à la fois généralistes et spécialistes dans un domaine. Il est moins risqué et plus efficace de mettre en œuvre un nombre limité de techniques bien maîtrisées dans un cadre précis pendant de nombreuses années. Un chercheur agissant de la sorte ne sortira que rarement de son moule doctoral ou post-doctoral. Cela constitue presque la règle aujourd'hui. Cependant, le système actuel est malgré tout assez souple pour tolérer toutes sortes d'écarts au modèle dominant. Cela m'est bien agréable car j'aurais sans doute renoncé à ce métier s'il me condamnait, à cause de mes limites, à une forme de fermeture, à une optimisation coût-bénéfice trop éloignée de ma manière de penser. Je dois dire que rencontrer mes semblables m'a beaucoup apporté. Au bout du compte, sur le plan humain, c'est le plaisir à exercer son métier qui compte le plus, tandis que sur le plan scientifique, c'est la (relative) qualité de la production qui importe. Mais tout cela n'est-il pas que l'expression de la comédie humaine ? Au delà des activités dérisoires de chacun, la production scientifique globale subit une lente digestion collective, et l'Histoire fait son tri implacable. Combien avez-vous lu d'articles originaux de Kolmogorov ?

Les années d'après thèse sont également l'occasion de découvrir plusieurs facettes de l'envers du décor, dans les conseils d'administration, de gestion, et surtout les commissions de spécialistes, bientôt remplacées par les comités de sélection. La comédie humaine est pleine d'enseignements, bien que celle qui se joue au moment des recrutements ait un rapport assez curieux à l'enseignement. Idéalement, un recrutement devrait étoffer le groupe sur le plan scientifique, le dynamiser, l'ouvrir, tout en évitant les catastrophes pédagogiques. Cet idéal semble assez peu compatible avec le micro-étiquetage des profils de postes pratiqué ici et là, à l'Université comme à l'INRA. Il faut dire que l'absence d'étiquetage n'est envisageable que si le groupe est raisonnable et ne donne pas dans l'ostracisme. Dans cette comédie où les destins se jouent, recruter meilleur ou différent fait parfois peur. La hiérarchie intra et inter-disciplinaire s'exprime aussi, et les plus puissants ne sont pas toujours les plus ouverts. Malgré tout, certains groupes parviennent à faire preuve de raison et d'audace, et en sont récompensés sur le long terme. J'aime à penser que les choses ne peuvent changer, lentement, que si certains font tout pour qu'elles changent rapidement. Ce paradoxe de l'engagement, qui fait fondre les illusions, encourage pourtant à agir inlassablement.

J'aimerais terminer ces lignes en saluant celles et ceux avec qui j'ai partagé des moments qui ont compté, au delà des mathématiques. Je pense par exemple à ces voyages à Biskra en compagnie d'Amine, à travers les Aurès, à Brahim, et à tous les mathématiciens en Algérie, qui tentent d'exercer leur métier dans un pays riche au peuple pauvre et meurtri. Je pense également à Christophe, Pierre, et Sylvain, qui ont rendu mes escapades niçoises si agréables. Je pense à ces fraternelles visites parisiennes chez Anne-Laure et Ivan. Je pense à Liming, et à son si sympathique accueil à Wuhan, en plein été chinois. Je pense à Nicolas le new-yorkais, qui faisait des « rough paths » le dimanche, après le pain et avant le tennis. Je pense encore à mes visites à Saint-Malon-sur-Mel, chez Florent et Lise, à nos discussions toujours enrichissantes, au bonheur de vivre des enfants rayonnants. Je pense à Fabrice et à Alice, à Bucarest, qui m'a tant rappelé mon enfance, dans un « pays de l'Est » maghrébin, cette enfance entre deux cultures qui m'a tôt fait comprendre que ma patrie n'était pas faite de terres et de religions mais de pensée et de sentiments. Je pense aussi à toutes celles et tous ceux chez qui j'apprécie l'humanité et la bienveillance, sans trouver les mots pour le leur dire. Il y a enfin tout ce que je ne dis pas ici...

Djalil Chafaï — Quartier des Minimes, 17 mai 2008.

Table des matières

Productions scientifiques

Ce bref chapitre recense mes principales productions scientifiques liées à la recherche : articles, livres, etc, issues d'un travail solitaire ou en collaboration. Sont ignorées ici un certain nombre de productions déjà finalisées mais jamais soumises pour publication.

Thèse de Doctorat

[T] D. Chafaï, Sur les inégalités de Sobolev logarithmiques en théorie de l'information et pour des systèmes de spins conservatifs en mécanique statistique *Thèse de Doctorat, Université Paul Sabatier*, 2002.

Articles dans des revues internationales à comité de lecture

[A1] D. Chafaï, M. Ledoux
Méthodes fonctionnelles pour des grandes déviations quasi-gaussiennes
Comptes Rendus de l'Académie des Sciences de Paris, Série I, Mathématiques 329, 523–526 (1999)

[A2] F. Rabier, N. Fourrié, D. Chafaï, P. Prunet
Channel selection methods for infrared atmospheric sounding interferometer radiances
Quarterly Journal of The Royal Meteorological Society 128, 1011–1027 (2002)

[A3] D. Chafaï
Gaussian maximum of entropy and reversed log-Sobolev inequality
Séminaire de Probabilités, XXXVI, Lecture Notes in Mathematics 1801, Springer, 194–200 (2003)

[A4] D. Chafaï
Glauber vs Kawasaki for spectral gap and log-Sobolev ineq. of some unbounded conservative spin systems
Markov Process. Related Fields 9, 341–362 (2003)

[A5] D. Chafaï
Entropies, convexity, and functional inequalities : on Φ-entropies and Φ-Sobolev inequalities
Journal of Mathematics of Kyoto University 44, 325–363 (2004)

[A6] D. Chafaï
Binomial-Poisson entropic inequalities and the $M/M/\infty$ queue
ESAIM Probability and Statistics 10, 317–339 (2006)

[A7] D. Chafaï, J.-M. Loubes
On nonparametric maximum likelihood for a class of stochastic inverse problems
Statistics and Probabability Letters 76, 1225–1237 (2006)

[A8] D. Chafaï, D. Concordet
On the strong consistency of asymptotic M-estimators
Journal of Statististical Planning and Inference 137, 2774–2783 (2007)

[A9] D. Chafaï, D. Concordet
Explicit formulas for a continuous stochastic maturation model. Application to anticancer drug PK/PD
Stochastic Models 24, 376-400 (2008)

[A10] D. Chafaï, D. Concordet
A new method for the estimation of variance matrix with prescribed zeros in nonlinear mixed effects models
accepté dans *Statistics and Computing* (2008)

[A11] J. Antic, C. Laffont, D. Chafaï, D. Concordet
Comparison of nonparametric methods in nonlinear mixed effects models
accepté pour publication dans *Computational Statistics and Data Analysis* (2008)

[A12] D. Bakry, F. Baudoin, M. Bonnefont, D. Chafaï
On gradient bounds for the heat kernel on the Heisenberg group
accepté pour publication dans *Journal of Functional Analysis* (2008)

Livre

[**L1**] C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto, G. Scheffer – Sur les inégalités de Sobolev logarithmiques, *Panoramas et Synthèses* 10, Société Mathématique de France (SMF) Paris (2000), 217 p. avec une préface de Dominique Bakry et Michel Ledoux

Prépublications

[**A13**] D. Chafaï
Circular law for non-central random matrices
soumis, 2007

[**A14**] D. Chafaï
The Dirichlet Markov Ensemble
soumis, 2007

[**A15**] D. Chafaï, D. Concordet
Confidence regions for the multinomial parameter with small sample size
soumis, 2008

[**A16**] D. Chafaï, F. Malrieu
On fine properties of mixtures with respect to concentration of measure and Sobolev type inequalities
en cours de révision mineure pour les Annales de l'IHP, 2008

[**A17**] P. Caputo, D. Chafaï, Ch. Bordenave
Spectrum of large random reversible Markov chains
en préparation, 2008

Livre pédagogique

[**L2**] B. Bercu, D. Chafaï – Modélisation stochastique et simulation – Cours et applications, *Collection Sciences Sup - Mathématiques appliquées pour le Master*, Société de Mathématiques Appliquées et Industrielles (SMAI), éditions Dunod (2007), 352 p.

Code informatique

[**C**] D. Chafaï, sous l'impulsion de T. Lyons – *Computational Rough Paths* : libraries of C++ template classes using the STL and compatible with GNU MP, disponible sur `http://CoRoPa.SourceForge.net/`

Nota Bene. Le livre pédagogique [**L2**] et le code informatique [**C**] sont à la frontière de la recherche. Ils constituent des productions en accord avec mes missions et en jonction avec mes activités scientifiques. Il m'a donc semblé indispensable de les mentionner dans cette liste.

Parcours scientifique

Ce court chapitre constitue une forme d'autobiographie scientifique. Il donne une vue d'ensemble de mon parcours et en présente l'articulation et la construction, sans formalisme mathématique. Une présentation synthétique de mes travaux, bien plus conforme aux standards académiques, est donnée dans le chapitre suivant.

J'ai soutenu le 17 mai 2002 une thèse de doctorat [**T**] intitulée *Sur les inégalités de Sobolev logarithmiques en théorie de l'information et pour des systèmes de spins conservatifs en mécanique statistique*, sous la direction de Michel Ledoux. Certains aspects de la théorie de l'information et de la mécanique statistique font désormais partie de mon bagage culturel, bien que je n'ai plus travaillé sur ces deux thèmes depuis ma thèse. Deux concepts importants sous-tendaient mon travail de thèse : celui d'entropie et celui de courbure pour les dynamiques markoviennes, l'exemple emblématique étant constitué par les processus d'Ornstein-Uhlenbeck, dont la courbure est constante, et dont la mesure invariante est gaussienne. Mon travail de thèse [**A1, A3, A4, L1**] m'a conduit plus tard à développer dans [**A5**] une unification d'une famille de cadres pour lesquels une inégalité de Sobolev logarithmique a lieu, en utilisant le concept de Φ-entropie et ses liens avec la convexité et l'interpolation par semi-groupe. Plus tard encore, j'ai obtenu une caractérisation complète de la propriété de tensorisation des Φ-entropies en terme de convexité, à l'occasion d'un travail [**A6**] concernant la file d'attente $M/M/\infty$, qui constitue un analogue discret remarquable du processus d'Ornstein-Uhlenbeck. Cette file d'attente m'est apparue lors de mon travail de modélisation en cancérologie [**A9**].

Mon travail de thèse a été interrompu par une année de service national, passée au Centre National de Recherches en Météorologie de Météo-France, sous la direction de Florence Rabier. J'ai alors travaillé sur des méthodes de sélection de canaux pour l'interféromètre infrarouge IASI, qui fait à présent partie de l'équipement d'un satellite météorologique à orbite polaire de la famille Metop. Le problème inverse stochastique associé est lié au transfert radiatif le long d'une colonne atmosphérique. L'une des méthodes de sélection envisagées était basée sur la différence d'entropie. Cette étude a nécessité l'écriture d'un code informatique en Fortran 90 pour le super-calculateur de Météo-France. L'ensemble de ce travail a fait l'objet d'une publication en météorologie [**A2**]. J'ai retrouvé les problèmes inverses stochastiques quelques années plus tard, en pharmacologie.

Après ma thèse, j'ai travaillé pendant une année en tant que post-doctorant sur le projet *digital description of serial data streams* mené par Terry Lyons à l'Université d'Oxford. Je me suis intéressé au problème de la reconstitution d'un signal multivarié à partir de la suite des intégrales itérées du signal entre deux instants, appelée signature. Cette approche est tout naturellement issue de la théorie des *rough paths*. Sous l'impulsion de Terry Lyons, j'ai en particulier développé des bibliothèques de classes templates C++ implémentant des conteneurs d'arbres et d'algèbres de Lie libres, incluant la formule de Campbell-Baker-Hausdorff et les bases de Hall. Deux autres post-doctorants ont successivement poursuivi ce travail. Le code est disponible sur le site du projet CoRoPa [**C**]. J'ai retrouvé beaucoup plus récemment certains aspects des rough paths à l'occasion d'un travail sur le noyau de la chaleur sur le groupe d'Heisenberg. Le couple formé par le mouvement brownien plan et son aire de Lévy constitue un exemple simple de *rough path*. Ce couple est markovien, et correspond exactement au noyau de la chaleur sur le groupe d'Heisenberg. Le semi-groupe associé peut être vu comme une famille de convolution gaussienne sur un groupe non-commutatif à dilatation de type Carnot. Le générateur infinitésimal associé est hypoelliptique mais non-elliptique, ce qui rend inefficaces les techniques riemanniennes usuelles basées sur la notion de courbure. Dans ce contexte, j'ai étudié (en collaboration avec D. Bakry, F. Baudoin, et M. Bonnefont [**A12**]) des bornes gradients pour ce noyau et leurs conséquences en termes d'inégalités fonctionnelles. Ce travail simplifie les approches de Driver & Melcher, et de H.-Q. Li et permet d'établir des inégalités isopérimétriques locales de type Cheeger et Bobkov. Ces résultats constituent de modestes premiers pas dans le vaste de champ inexploré de la géométrie des dynamiques markoviennes hypoelliptiques non-elliptiques.

Après mon année post-doctorale, j'ai rejoint le laboratoire de pharmacologie dirigé par Pierre-Louis Toutain (à présent par Alain Bousquet-Mélou), en tant que chargé de recherches à l'INRA (département de Mathématiques et Informatique Appliquées). La pharmacologie est la branche de la biologie qui étudie les médicaments et leurs interactions avec les organismes vivants. Cette nouvelle mission m'a tout naturellement conduit à aborder les problèmes de modélisation stochastique et de statistique qui apparaissent en pharmacologie. Ces problèmes mettent en œuvre des modèles compartimentaux. Les problèmes statistiques associés mettent en jeu à leur tour des modèles non-linéaires à effets mixtes, qui peuvent être vus à la fois comme des problèmes inverses stochastique et comme des modèles de mélange semi-paramétriques. Dans ce cadre, je me suis intéressé en particulier à la consistance du maximum de vraisemblance non-paramétrique pour des problèmes inverses stochastiques (en collaboration avec J.-M. Loubes [**A7**]), à la consistance de M-estimateurs généralisant un résultat de Pfanzagl pour les modèles concaves (en collaboration avec D. Concordet [**A8**]), à la modélisation de la pharmacodynamie des anticancéreux en utilisant des files d'attente et des formules de Feynman-Kac (en collaboration avec D. Concordet [**A9**]), à un algorithme de calcul du maximum de vraisemblance de modèles non-linéaires à effets mixtes couplant EM et ICF (en collaboration avec D. Concordet [**A10**]), ainsi qu'à la comparaison numérique par simulation de méthodes non-paramétriques en pharmacocinétique de population (en collaboration avec J. Antic, C. Laffont, et D. Concordet [**A11**]). Les modèles de la pharmacologie sont anciens et ont beaucoup été étudiés. Cependant, certaines questions restent ouvertes ou assez peu traitées, probablement en raison de leur difficulté, comme par exemple la classification supervisée de cinétiques d'une population, la sélection de modèles compartimentaux, ou la détection de sous-populations à partir de l'observation de cinétiques. La prise en compte des covariables rend ces questions encore plus ardues.

Les modèles compartimentaux à échanges linéaires donnent lieu à des systèmes d'équations différentielles. Mon travail sur les anticancéreux [**A9**] m'a fait remarquer que ces équations peuvent être vues comme la dynamique moyenne de systèmes de particules en interaction de type « zero-range » en mécanique statistique, ainsi que comme des files d'attente $M/M/\infty$ en interaction sous forme de réseau de Jackson. Cette observation m'a conduit à étudier plus spécifiquement dans [**A6**] la file d'attente $M/M/\infty$, qui s'est avéré être un excellent équivalent discret du processus d'Ornstein-Uhlenbeck, pour lequel j'ai en particulier exploré des inégalités fonctionnelles à base de Φ-entropie. Ce travail a été l'occasion de caractériser la propriété de tensorisation des Φ-entropies en terme de propriétés de convexité de la fonction Φ sous-jacente. Les liens entre mécanique statistique et réseaux de files d'attentes sont sous-développés, et mériteraient d'être explorés, notamment en ce qui concerne la convergence à l'équilibre. Le cas non-réversible est probablement à la fois le plus difficile et le plus intéressant.

La statistique des modèles à compartiments suggère de considérer le comportement asymptotique de processus de Markov à dynamique aléatoire. Cette observation m'a amené à l'étude du spectre de grandes matrices markoviennes aléatoires dans [**A14**], qui met eu jeu une loi sur le polytope de Markov. Cette étude est liée au fait que le trou spectral d'une grande chaîne de Markov aléatoire à n états est de l'ordre de $1-n^{-1/2}$. Ces matrices aléatoires sont non-centrées, non-symétriques, et à entrées corrélées. Cela m'a conduit à établir un théorème de la loi du cercle de type Girko-Bai pour des matrices aléatoires à entrées i.i.d. non-centrées dans [**A13**]. J'ai enfin étudié (en collaboration avec P. Caputo et Ch. Bordenave [**A17**]) le comportement asymptotique du spectre de matrices aléatoires réversibles en utilisant des graphes à poids aléatoires. Ces modèles de matrices markoviennes aléatoires peuvent être vus comme des modèles de chaînes de Markov finies en environnement aléatoire. Ce travail mène à tout un champ de questions ouvertes, entre graphes aléatoires et matrices aléatoires.

Les modèles non-paramétriques à effets mixtes de la pharmacocinétique de population peuvent être vus comme des problèmes de mélanges de lois. Ceci m'a conduit à étudier (en collaboration avec F. Malrieu [**A16**]) les propriétés fines des mélanges de lois vis à vis de la concentration de la mesure et des inégalités fonctionnelles de Poincaré et de Sobolev logarithmique. Nous montrons en particulier que même pour un mélange à deux composantes, la constante optimale de l'inégalité de Sobolev logarithmique peut exploser lorsque l'un des composants du mélange disparaît, tandis que la concentration de la mesure est beaucoup plus stable. Ce travail montre également le rôle joué par le transport optimal à contrainte de support dans la géométrie des mélanges.

J'ai souvent été sollicité, en tant que mathématicien à l'INRA, par des collègues à propos de problèmes concrets issus de la biologie. Bien souvent, la solution se résume à l'utilisation d'une technique connue et balisée comme la régression, le modèle linéaire, la classification supervisée, etc. Cependant, un problème récurrent a particulièrement excité mon esprit. Il s'agit de construire une zone de confiance pour le paramètre d'une loi multinomiale à partir d'un échantillon de petite taille. Ce problème, beaucoup plus subtil qu'il n'y parait, m'a conduit (en collaboration avec D. Concordet [**A15**]) à proposer une nouvelle méthode basée sur le concept de familles couvrantes et sur les ensembles de niveau, qui semble plus performante que les approches existantes.

Présentation synthétique

Ce chapitre présente de manière synthétique et structurée l'essentiel de mes travaux scientifiques. J'ai choisi de ne pas aborder le contenu de mon travail de thèse [**T, A1, A3, A4, L1**], des articles "marginaux" [**A2, A15**], du livre pédagogique [**L2**], et du code informatique [**C**]. Ce chapitre est organisé en trois parties **autonomes**, dont voici les thèmes :

1. **Modèles biologiques et statistique.**
   – modèles compartimentaux, pharmacocinétique et pharmacodynamie de population
   – estimateurs pour problèmes inverses stochastiques, modèles non-linéaires à effets mixtes
   – modèles de mélanges, algorithmes de type EM et ICF, modèles graphiques de covariance
   – modélisation en cancérologie, processus ponctuels, systèmes de particules, files d'attentes
   – renormalisation de processus markoviens inhomogènes et formules de Feynman-Kac
2. **Inégalités fonctionnelles.**
   – inégalités fonctionnelles, concentration de la mesure, isopérimétrie
   – rôle de la convexité dans les inégalités entropiques, tensorisation
   – noyau de la chaleur, groupe d'Heisenberg et dynamiques hypoelliptiques
   – files d'attentes, mélanges de lois
3. **Matrices aléatoires.**
   – spectre des matrices markoviennes aléatoires, graphes à poids aléatoires
   – théorèmes de type Wigner, Marchenko-Pastur, et Girko-Bai
   – convergence des valeurs propres extrémales, déformations de rang un

Le concept le plus récurrent ici est celui de dynamique markovienne. Dans la première partie, ce sont les modèles à compartiments de la pharmacologie qui sont liés à des dynamiques markoviennes. La seconde partie traite d'inégalités fonctionnelles associées à la vitesse et à la géométrie de dynamiques markoviennes. Enfin, la troisième partie traite de dynamiques markoviennes aléatoires. Ces trois parties ne se réduisent pas à l'étude de facettes de problèmes markoviens. Leur contenu balaye un spectre à la fois théorique et appliqué, et met en œuvre des techniques et des concepts issus de différents pans de l'analyse, des probabilités, et de la statistique.

1.1 Modèles biologiques et statistique

Cette partie présente brièvement le travail que j'ai mené à l'INRA concernant différents aspects des problèmes inverses stochastiques de la pharmacologie, liés aux modèles compartimentaux. Les modèles compartimentaux sont largement utilisés en science appliquée, et en particulier en pharmacologie, comme en témoignent les livres [GP82, JS93, DG95, MK00, MI06, RT89].

La pharmacologie est la branche de la médecine qui étudie les médicaments et leur interaction avec les organismes. Nous abordons ici certains aspects des modèles compartimentaux à taux linéaires mais pouvant toutefois dépendre du temps. Un modèle compartimental est constitué par des boites appelées compartiments, contenant chacune une certaine quantité de matière.

Ces quantités évoluent au cours du temps selon trois mécanismes : apport extérieur au système, transfert entre compartiments, perte de matière vers l'extérieur du système. En pharmacologie, un compartiment correspond par exemple à un tissu, ou encore à un stade de maturation, tandis que la matière correspond à une quantité de médicament, ou de cellules. Considérons pour commencer que la quantité de matière est représentée par un nombre réel positif déterministe. On note $Q_i(t)$ la quantité de matière dans le compartiment $i \in I$ au temps $t \geq t_0$. On note $Q(t)$ le vecteur $i \in I \mapsto Q_i(t)$, et on identifie $\mathbb{R}^I$ à $\mathbb{R}^n$ où $n = \mathrm{card}(I) < \infty$. Considérons la dynamique de $(Q(t))_{t \geq t_0}$ décrite par le système différentiel linéaire

$$\forall i \in I,\ \forall t \geq t_0,\ \partial_t Q_i(t) = \lambda_i(t) + \sum_{j \neq i} \rho_{j,i}(t) Q_j(t) - Q_i(t) \left\{ \kappa_i(t) + \sum_{j \neq i} \rho_{i,j}(t) \right\} \tag{1.1}$$

avec condition initiale $Q(t_0)$. Cela correspond à une équation bilan instantanée (cf. figure 1.1). Les taux λ, κ et ρ sont positifs. Les composantes nulles de λ, κ, et ρ déterminent une topologie des échanges entre compartiments et avec l'extérieur. Cela peut être vu comme un graphe pouvant fluctuer au cours du temps. Par ailleurs, il est toujours possible d'inclure κ dans ρ en ajoutant un compartiment supplémentaire. Les taux κ et ρ agissent proportionnellement au contenu du compartiment qui donne de la matière. Matriciellement, la dynamique s'écrit

$$\partial_t Q(t) = \mathbf{M}(t) Q(t) + \lambda(t) \tag{1.2}$$

où $\lambda(t)$ désigne le vecteur $(\lambda_i(t); i \in I)$, et où $\mathbf{M}(t)$ est la matrice de « transfert » définie par

$$\mathbf{M}_{i,j}(t) = \begin{cases} \rho_{j,i}(t) & \text{if } i \neq j \\ -\sum_{k \neq i} \rho_{i,k}(t) - \kappa_i(t) & \text{if } i = j \end{cases}$$

pour tout $i, j \in I$. La théorie traditionnelle des équations différentielles ordinaires donne

$$Q(t) = \mathbf{R}(t, t_0) Q(t_0) + \int_{t_0}^{t} \mathbf{R}(t, u) \lambda(u)\, du$$

pour tout $t \geq t_0$, où la résolvante $\mathbf{R}$ est solution de l'équation différentielle linéaires matricielle $\mathbf{R}(t_0, t_0) = \mathbf{I}$ et $\partial_t \mathbf{R}(t, t_0) = \mathbf{M}(t) \mathbf{R}(t, t_0)$ pour tout $t \geq t_0$. Si $\mathbf{M}$ *ne dépend pas du temps*, alors $\mathbf{R}(v, u) = \exp((v-u)\mathbf{M})$, et la diagonalisation de $\mathbf{M}$ permet d'exprimer $Q(t)$ comme combinaison linéaire de fonctions exponentielles liées au spectre de $\mathbf{M}$. En revanche, si $\mathbf{M}$ *dépend du temps*, alors il n'existe pas de formule fermée pour $Q(t)$.

Kelly donne dans [Kel79, sec. 4.5 p. 113–117] une interprétation stochastique des modèles compartimentaux à taux constants. Cette interprétation subsiste lorsque les taux dépendent du temps. À présent, chaque compartiment contient des particules, et on note N_t^i le nombre de particules dans le compartiment $i \in I$ au temps $t \geq t_0$. Les particules sont indistingables, et peuvent être créées, détruites, ou se mouvoir d'un compartiment à un autre, selon une dynamique markovienne. Plus précisément, on suppose que $(N_t)_{t \geq t_0}$ est un processus de Markov inhomogène en temps, d'espace d'état $E = \mathbb{N}^I$, et de générateurs $(L_t)_{t \geq t_0}$. Pour tout $x \in E$ et tout $t \geq s \geq t_0$, on pose $Q(s, t, x) = \mathbb{E}(N_t \mid N_s = x)$. Les équations de Chapman-Kolmogorov permettent d'établir que si pour tout $t \geq t_0$, la fonction $x \in E \mapsto A_t(x) = \sum_{y \in E} L_{t,x,y} y$ est affine, alors Q est la solution de l'équation différentielle linéaire

$$\partial_t Q(s, t, x) = A_t(Q(s, t, x)),$$

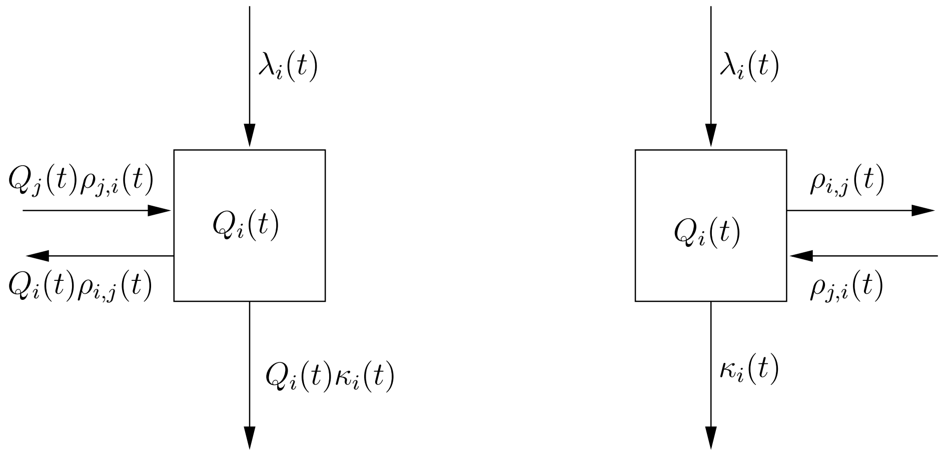

FIG. 1.1 – Le diagramme de gauche montre les véritables taux, qui dépendent linéairement du compartiment qui donne de la matière. On a coutume de ne pas figurer cette dépendance, comme dans le diagramme de droite.

pour tout $t \geq s \geq t_0$, avec condition initiale $Q(s,s,x) = x$. Considérons à présent le cas particulier où L_t est donné pour $x \neq y \in E$ par

$$L_{t,x,y} = \begin{cases} \lambda_i(t) & \text{si } y = x + e_i \text{ pour } i \in I \text{ (naissance)} \\ x_i\rho_{i,j}(t) & \text{si } y = x - e_i + e_j \text{ pour } i \neq j \text{ dans } I \text{ (transfert)} \\ x_i\kappa_i(t) & \text{si } y = x - e_i \text{ pour } i \in I \text{ (mort)} \\ 0 & \text{sinon} \end{cases} \tag{1.3}$$

où $e_1, \ldots, e_n$ désigne la base canonique de $\mathbb{R}^n \equiv \mathbb{N}^I = E$. Cela donne pour tout $x \in E$ et $i \in I$,

$$A_t(x)_i = \sum_{y \in E} L_{t,x,y} y_i = \lambda_i(t) + \sum_{j \in I, j \neq i} x_j \rho_{j,i}(t) - x_i \left\{ \kappa_i(t) + \sum_{j \in I, j \neq i} \rho_{i,j}(t) \right\}. \tag{1.4}$$

On reconnait immédiatement l'équation (1.2). Ainsi, le nombre moyen de particules par compartiments se comporte comme un système compartimental déterministe. Cela suggère deux interprétations stochastiques des modèles compartimentaux linéaires.

- **Files d'attentes M/M/∞ en interaction.** Le générateur (1.3) suggère d'interpréter le processus $(N_t)_{t \geq t_0}$ comme un système de n files d'attentes M/M/∞ en interaction. Ici, chaque particule est un client, et chaque compartiment est une file M/M/∞. La quantité N_t^i est donc le nombre de clients dans la file i au temps t. Les clients peuvent changer de file, selon un taux linéaire en fonction de la file qu'ils quittent. Ces files en interaction peuvent être vues comme des systèmes de particules en interaction markovien spéciaux [Dur95, Lig05]. En mécanique statistique, un tel système est une dynamique "zero-range" sur le graphe des compartiments [CP07]. Ces files en interaction constituent également des réseaux de Jackson particuliers [Jac57, Jac63, Rob03] inhomogènes en temps. Ces files sont en tandem dans le cas d'une chaîne de compartiments avec élimination en queue de chaîne.
- **Particules indépendantes sur le graphe des compartiments.** Comme l'a remarqué Kelly [Kel79, sec. 4.5 p. 113–117] dans le cas homogène, N_t peut être vu comme le vecteur d'occupation de la superposition de marches aléatoires indépendantes sur le graphe des

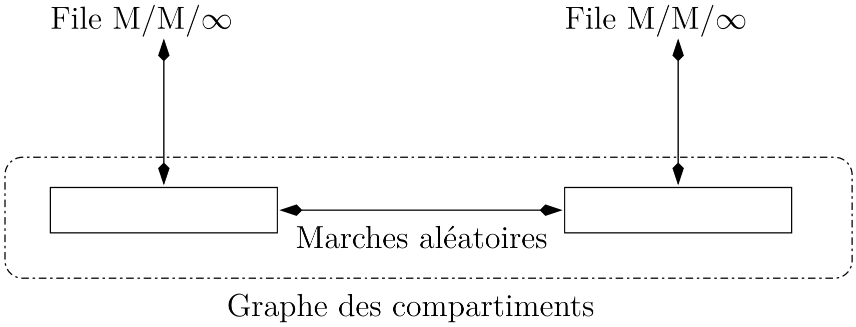

Fig. 1.2 – Deux interprétations de la dynamique markovienne (1.3) de (N_t). La première est verticale et correspond à des files d'attentes M/M/∞ en interaction, tandis que la seconde est horizontale et correspond à des marches aléatoires indépendantes sur le graphe des compartiments, sujettes à création et destruction. Tous les taux peuvent dépendre du temps.

compartiments. Ces marches sont créées par des processus ponctuels de Poisson indépendants d'intensité λ, et peuvent mourir avec taux κ (cela peut s'écrire avec une formule de Feynman-Kac [DM04]). La nature affine de A_t se traduit par le fait que la dynamique L_t est une superposition de dynamiques aléatoires indépendantes. D'autre part, lorsque le système est réduit à un unique compartiment, on retrouve la construction de la file M/M/∞ à partir de clients indépendants avec intensité d'arrivée λ et taux de service κ.

1.1.1 Pharmacocinétique de population

Pour un individu (un organisme donné), l'évolution au cours du temps de la quantité de principe actif (un antibiotique par exemple) dans un tissu s'écrit $t \mapsto f(t, \varphi)$ où $f(\cdot, \varphi)$ est une combinaison linéaire de fonctions exponentielles provenant d'un modèle compartimental déterministe. Cette fonction n'est pas linéaire en φ. Le paramètre $\varphi \in \mathbb{R}^p$ est spécifique à l'individu. Ses composantes, qui possèdent une interprétation biologique précise, correspondent aux taux de transfert entre les compartiments (les tissus). Le principal objectif de la pharmacocinétique de population est d'estimer la loi $\mathcal{P}$ du paramètre φ dans la population, en observant pour une sous-population d'individus de taille n la cinétique individuelle, de manière approchée et éparse. Un modèle statistique naturellement associé s'écrit par exemple sous la forme

$$Y_{i,j} = f(t_{i,j}, \varphi_i)(1 + \sigma\varepsilon_{i,j})$$

où $1 \leq i \leq n$ et $1 \leq t_{i,j} \leq n_i$ et où les $\varepsilon_{i,j}$ sont typiquement gaussiens standard et indépendants des φ_i et des $t_{i,j}$. Ce modèle correspond à l'observation pour l'individu i de n_i valeurs bruitées de la cinétique aux temps $t_{i,1}, \ldots, t_{i,n_i}$. Ces temps sont typiquement aléatoires mais connus. Ici, f est connue, les φ_i ne sont pas observés, et le but est d'estimer la loi $\mathcal{P}$ des φ_i à partir de l'observation des $Y_{i,j}$. La loi des $t_{i,j}$ n'est pas connue et peut être vue comme un paramètre de nuisance. Le paramètre fixe σ peut-être supposé connu dans certains cas. L'asymptotique porte sur n tandis que $\max_{1 \leq i \leq n} n_i$ reste borné et petit. La loi des observations Y peut-être vue comme un mélange où la loi mélangeante est précisément $\mathcal{P}$. Le modèle statistique ci-dessus apparaît également comme un modèle non-linéaire, à effets mixtes (fixes et aléatoires). En pratique, il est courant de tenir compte de co-variables en utilisant un modèle linéaire de second niveau sur le paramètre φ. L'analyse mathématique de ces modèles a été développée surtout dans un cadre paramétrique où $\mathcal{P}$ est par exemple un mélange fini à nombre de composants prescrit de lois de type prescrit. Ces modèles, dans leurs versions paramétriques, sont réellement utilisés par l'industrie pharmaceutique et le monde médical pour fixer les doses des médicaments.

Le problème non-paramétrique de l'estimation de $\mathcal{P}$ à partir de l'observation de Y peut-être vu comme un problème inverse stochastique particulier, pour lequel les propriétés de l'estimateur de maximum de vraisemblance sont mal connues. Dans [**A7**], nous montrons que si le coefficient de variation σ est connu, alors le maximum de vraisemblance non-paramétrique de $\mathcal{P}$ sur une classe de distributions convexe est consistant. La méthode consiste à se ramener à un résultat de Pfanzagl sur les modèles de mélange concaves [Pfa88]. Cette approche est poursuivie dans [**A8**], où nous établissons la consistance de M-estimateurs en généralisant le résultat de Pfanzagl pour les modèles concaves. Dans [**A11**], nous menons une comparaison par simulation des performances de plusieurs estimateurs non-paramétriques, dont les propriétés mathématiques sont mal connues, dans la situation où la loi de population $\mathcal{P}$ est bimodale (présence de sous-populations).

1.1.2 Modèles de maturation-survie en cancérologie

Certains médicaments anti-cancéreux ont pour effet d'éliminer les cellules en cours de division. En particulier, ils ont pour effet secondaire d'éliminer les cellules blanches du sang (leucocytes) qui sont essentielles au système immunitaire. Le contrôle de la cinétique est donc crucial pour maintenir en vie le patient. Karlsson et ses collaborateurs [FFSK00, FK03] ont développé depuis de nombreuses années des modèles de maturation-survie pour modéliser la pharmacodynamie des anti-cancéreux à partir de leur pharmacocinétique. Ces modèles sont constitués par une longue chaîne de compartiments à taux d'élimination variable au cours du temps, qui représente la maturation-survie des leucocytes. Le problème du choix du nombre de compartiments n'est pas élucidé, et l'approche nécessite de lourds calculs de résolution de systèmes d'équations différentielles. Dans [**A9**], nous montrons comment un passage à la limite en espace permet de simplifier ces modèles tout en conservant leurs proprités qualitatives. Ce passage à la limite fait apparaître un modèle à deux compartiments (cf. figure 1.3), avec un taux d'élimination qui dépend de la cinétique de l'anti-cancéreux et un retard qui dépend de l'individu. L'étude mathématique fait apparaître naturellement des processus ponctuels de Poisson amincis, des formules de Feynman-Kac, et des files d'attente de type $M/M/\infty$ inhomogènes en temps, en liaison avec des travaux de Schuhmacher & Thieme [ST88] et Keilson & Servi [KS94]. Les formules explicites obtenues permettent également l'étude de certains problèmes d'identifiabilité, qui sont innacessibles pour les modèles constitués d'une longue chaîne de compartiments.

1.1.3 Couplage des algorithmes EM et ICF

Le calcul effectif de l'estimateur paramétrique de maximum de vraisemblance dans les modèles non-linéaires à effet mixtes avec un effet aléatoire gaussien peut être mené en utilisant un algorithme de type Expectation Maximization (EM). Cette approche pose problème lorsque la matrice de covariance de l'effet aléatoire a une structure de zéros prescrite, car cette dernière n'est pas facile à maintenir dans l'étape M de l'algorithme. Dans ce cadre, nous avons montré dans [**A10**] que EM peut être naturellement couplé avec un algorithme de type Iterated Proportional Fitting (ICF). Cela a pour effet de produire des approximations qui appartiennent toujours au cône des matrices de covariances le long de l'algorithme. L'algorithme ICF a été développé récemment par Chaudury & al [CDR07] pour estimer la covariance d'un vecteur gaussien à structure de zéro prescrite à partir de son observation directe. Il est basé sur l'exploitation de la décomposition de la vraisemblance obtenue en utilisant le complément de Schur. L'optimization est menée colonne par colonne. La méthodologie proposée par [**A10**] résoud de manière satisfaisante un vieux problème algorithmique concret posé par l'usage des modèles non-linéaires à effets mixtes.

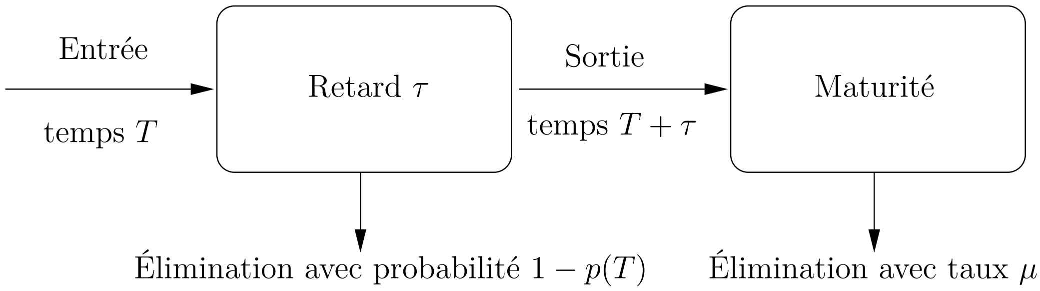

FIG. 1.3 – Schéma du système compartimental inhomogène simplifié obtenu par passage à la limite dans [**A9**]. Ici, τ représente un temps de maturation déterministe. La quantité $p(T)$ représente la probabilité de survie d'une particule qui a commencé sa maturation au temps T. Le compartiment de gauche correspond à un amincissement poissonnien avec retard, tandis que le compartiment de droite correspond à un processus de comptage de type M/M/∞ inhomogène en temps.

1.1.4 Quelques perspectives

Les modèles compartimentaux et les modèles non-linéaires à effets mixtes posent de nombreux problèmes mathématiques et algorithmiques ouverts. Voici quelques pistes parmi d'autres :
- classification supervisée et non-supervisée de cinétiques, méthodes à noyau
- détection de sous-population en pharmacocinétique. Problèmes de tests associés
- sélection de modèles pour loi de population dans les modèles non-linéaires à effets mixtes et estimation adaptative. Lien avec les mélanges
- algorithmes EM stochastiques en compétition pour optimiser la vraisemblance
- liens entre mécanique statistique et réseaux de files d'attentes

1.2 Inégalités fonctionnelles

Les inégalités fonctionnelles de type Sobolev et leur liens avec les dynamiques markoviennes constituent un champ de recherche actif depuis de nombreuses années, comme en témoignent par exemple les ouvrages synthétiques [Hel02, GZ03, Roy99, Gro06, Bak94] et [**L1**].

1.2.1 Entropies, convexité, et inégalités Phi-Sobolev

Je me suis attaché à développer dans mon travail [**A5**] et [**A6**] un formalisme et des outils basés sur la notion de Φ-entropie, qui permettent d'englober une collection de résultats connus et d'en proposer des généralisations. L'idée directrice ici est de mettre à jour le rôle de la convexité dans les inégalités de Poincaré et de Sobolev logarithmique et leur preuves par semi-groupe. Les méthodes développées dans [**A5**] sont par exemple utilisées dans [GI07] pour l'étude d'équations de Fokker-Planck dirigées par certains processus de Lévy.

Soit Q une loi de probabilité sur $(\Omega,\mathcal{A},\mathbb{P})$ et $\Phi:\mathcal{I}\to\mathbb{R}$ une fonction convexe définie sur un intervalle $\mathcal{I}\subset\mathbb{R}$. La Φ-entropie d'une fonction mesurable $f:\Omega\to\mathbb{R}$ est définie par

$$\mathbf{Ent}_Q^{\Phi}(f)=\mathbf{E}_Q(\Phi(f))-\Phi\left(\mathbf{E}_Q(f)\right).$$

On retrouve la variance de f pour Q lorsque $\Phi(u)=u^2$ avec $\mathcal{I}=\mathbb{R}$ et l'entropie de f pour Q lorsque $\Phi(u)=u\log(u)$ avec $\mathcal{I}=\mathbb{R}_+$. Formellement, si $(P_t)_{t\geq 0}$ est un semi-groupe markovien de générateur L et de loi invariante Q, alors pour toute fonction f raisonnable et tout temps $t \geq 0$

$$\partial_t \mathbf{Ent}_Q^\Phi(P_t f) = \mathbf{E}_Q(\Phi'(P_t f) L P_t f) \leq 0.$$

Cette décroissance est une conséquence immédiate de la convexité de Φ et de la nature markovienne de L. Il en découle que pour toute constante $\rho > 0$, il y a équivalence entre

$$\forall f,\ \forall t, \quad \mathbf{Ent}_Q^\Phi(P_t f) \leq e^{-\rho t} \mathbf{Ent}_Q^\Phi(f)$$

et

$$\forall f, \quad \rho\, \mathbf{Ent}_Q^\Phi(f) \leq -\mathbf{E}_Q(\Phi'(f) L f).$$

Cette dernière inégalité est qualifiée d'*inégalite de* Φ*-Sobolev* pour Q et L. Elle devient, selon la fonction Φ choisie, une inégalité de Poincaré, de Sobolev logarithmique, ou de Beckner-Latała-Oleszkiewicz (les détails se trouvent dans [**A5**]). Lorsque Q est réversible et L est un générateur de diffusion, le second membre s'écrit alternativement

$$-\mathbf{E}_Q(\Phi'(f) L f) = -\mathbf{E}_Q(\Phi''(f) \Gamma(f,f))$$

où $\Gamma(f,f) = \frac{1}{2}(L(f^2) - 2fLf)$. Lorsque par exemple L est une diffusion sur $\mathbb{R}^d$ sans terme constant et de matrice du second ordre égale à l'identité alors $\Gamma(f,f) = |\nabla f|^2$. Dans ce cadre, il est montré dans [**A5**] que les preuves par semi-groupe des inégalités de Poincaré et de Sobolev logarithmique sous hypothèse de courbure minorée peuvent être généralisées aux inégalités de Φ-Sobolev lorsque la fonction $(u,v) \mapsto \Phi''(u) v^2$ est convexe. Cette observation permet également d'obtenir des inégalités similaires pour la mesure de Wiener sur l'espace des chemins de variétés riemanniennes, en généralisant certain résultats de [CHL97]. Cette avancée basée sur la convexité permet également de traiter le cas de l'espace des mesures ponctuelles de Poisson, en utilisant cette fois-ci la convexité de la « divergence de Bregman [Bre67] » $(u,v) \mapsto \Phi(u+v) - \Phi(u) - \Phi'(u)v$. Les inégalités de Φ-Sobolev à énergies modifiées obtenues de la sorte généralisent certains résultats de [Wu00]. En particulier, cette approche permet de donner une preuve élémentaire par interpolation semi-groupe des inégalités de Sobolev logarithmiques modifiées pour la loi de Poisson, avec la meilleure énergie connue. Tous ces aspects sont présentés dans [**A5**]. Cette approche est approfondie dans [**A6**], en introduisant les transformées A, B, C de Φ définies par

$$\begin{aligned} A^\Phi(u,v) &= \Phi(u+v) - \Phi(u) - \Phi'(u)v \\ B^\Phi(u,v) &= (\Phi'(u+v) - \Phi'(u))v \\ C^\Phi(u,v) &= \Phi''(u)v^2. \end{aligned}$$

La A-transformée de Φ apparaît naturellement sur l'espace à deux points $\Omega = \{0,1\}$ équipé de la loi de Bernoulli $Q = p\delta_1 + q\delta_0$, puisque pour toute fonction $f : \{0,1\} \to \mathcal{I}$,

$$\mathbf{Ent}_Q^\Phi(f) = q\Phi(a) + p\Phi(b) - \Phi(qa+pb) = pA^\Phi(u,v) - A^\Phi(u,pv) \tag{1.5}$$

où $(a,b) = (f(0), f(1))$ et $(u,v) = (a, b-a)$. Les transformées $A - B - C$ sont les germes de formes de Dirichet modifiées discrètes via les identités

$$\begin{aligned} A^\Phi(f, Df) &= D\Phi(f) - \Phi'(f) Df \\ B^\Phi(f, Df) &= D(\Phi'(f)) Df \\ C^\Phi(f, Df) &= \Phi''(f) |Df|^2 \end{aligned}$$

où $(Df)(x) = f(x+1) - f(x)$. La table 1.1 donne quelques exemples. Les transformations $A - B - C$ permettent d'obtenir de nombreuses inégalités entropiques. La convexité de Φ est équivalente à la positivité de ses transformées $A-B-C$. Qu'en est-il de la convexité de la Φ-entropie elle-même ? Le résultat le plus important de [**A6**] consiste en l'équivalence des propriétés suivantes, pour une fonction lisse $\Phi : \mathcal{I} \to \mathbb{R}$ (les détails techniques sont omis ici pour simplifier la présentation) :

1. A^Φ est convexe
2. B^Φ est convexe
3. C^Φ est convexe
4. Φ est affine ou bien $\Phi'' > 0$ et $-1/\Phi''$ est convexe
5. la fonction $(a,b) \mapsto t\Phi(a) + (1-t)\Phi(b) - \Phi(ta + (1-t)b)$ est convexe pour tout $t \in [0,1]$
6. **convexité fonctionnelle.** $f \mapsto \mathbf{Ent}_Q^\Phi(f)$ est convexe pour tout espace probabilisé $(\Omega, \mathcal{A}, Q)$
7. **formule variationnelle.** pour tout espace probabilisé $(\Omega, \mathcal{A}, Q)$ et toute fonction f

$$\mathbf{Ent}_Q^\Phi(f) = \sup_g \left\{ \mathbf{E}_Q((\Phi'(g) - \Phi'(\mathbf{E}_Q g))(f-g)) + \mathbf{Ent}_Q^\Phi(g) \right\} \tag{1.6}$$

8. **inégalité de tensorisation.** pour tout espace de probabilité produit

$$(\Omega, \mathcal{A}, Q) = (\Omega_1 \times \cdots \times \Omega_n, \mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_n, Q_1 \otimes \cdots \otimes Q_n)$$

et toute fonction f

$$\mathbf{Ent}_Q^\Phi(f) \leq \mathbf{E}_Q(\mathbf{Ent}_{Q_1}^\Phi(f) + \cdots + \mathbf{Ent}_{Q_n}^\Phi(f)). \tag{1.7}$$

En particulier, la convexité des transformées $A-B-C$ est équivalente à la convexité de la Φ-entropie ainsi qu'à la propriété de tensorisation de la Φ-entropie. Cette équivalence unifie certains résultats de [BBLM05, LO00, Mas07] et [**A5**]. La formule variationnelle (1.6) n'a rien de mystérieux et correspond à l'expression de la fonctionnelle convexe Φ-entropie comme enveloppe de ses tangentes directionnelles. Par ailleurs, lorsque $\Phi(u) = u^2$ et $\mathcal{I} = \mathbb{R}$ alors $2A^\Phi = B^\Phi = C^\Phi$ tandis que lorsque $\Phi(u) = u\log(u)$ et $\mathcal{I} = \mathbb{R}_+$ alors $A^\Phi \leq \max(B^\Phi, C^\Phi)$. Dans les deux cas, A^Φ, B^Φ, et C^Φ sont convexes. D'autres exemples sont donnés dans [**A6**]. Cette observation apporte un nouvel éclairage sur les liens entre les énergies modifiées dans les inégalités de Poincaré et de Sobolev logarithmiques par exemple.

Fonction Φ	$\boldsymbol{\mathcal{I}}$	$\boldsymbol{A^\Phi}$	$\boldsymbol{A^\Phi(f, Df)}$
$u\log(u)$	$\mathbb{R}_+^*$	$(u+v)(\log(u+v) - \log(u)) - v$	$(f+Df)D(\log f) - Df$
u^2	$\mathbb{R}$	v^2	$\|Df\|^2$
		$\boldsymbol{B^\Phi}$	$\boldsymbol{B^\Phi(f, Df)}$
$u\log(u)$	$\mathbb{R}_+^*$	$v(\log(u+v) - \log(u))$	$D(f)D(\log f)$
u^2	$\mathbb{R}$	$2v^2$	$2\|Df\|^2$
		$\boldsymbol{C^\Phi}$	$\boldsymbol{C^\Phi(f, Df)}$
$u\log(u)$	$\mathbb{R}_+^*$	v^2u^{-1}	$\|Df\|^2 f^{-1}$
u^2	$\mathbb{R}$	$2v^2$	$2\|Df\|^2$

TAB. 1.1 – Exemples de transormées $A-B-C$ et énergies modifiées associées.

L'article [**A6**] propose également une étude détaillée de la file d'attente M/M/∞, qui joue pour le processus de Poisson le rôle joué par le processus d'Ornstein-Unlenbeck pour le mouvement brownien. La loi instantannée du processus M/M/∞ noté $(X_t)_{t\geq 0}$ est donnée par la formule Poisson-binomiale

$$\mathcal{L}(X_t|X_0 = n) = \mathcal{B}(n, e^{-\mu t}) * \mathcal{P}(\rho(1 - e^{-\mu t}))$$

qui constitue un analogue de la formule gaussienne de Mehler pour le processus d'Ornstein-Uhlenbeck. La loi de Poisson $\mathcal{P}(\rho)$ de moyenne $\rho = \lambda/\mu$ est invariante et réversible pour le processus. Le semi-groupe $(P_t)_{t\geq 0}$ de $(X_t)_{t\geq 0}$ possède une courbure constante en ce sens que

$$DP_t f = e^{-\mu t} P_t Df.$$

Cette commutation exacte est l'analogue de celle du semi-groupe d'Orstrein-Uhlenbeck, et s'écrit infinitésimalement $[D, L] = DL - LD = -\mu D$ où $(Lf)(n) = \lambda(Df)(n) + n\mu(D^* f)(n)$ est le générateur infinitésimal de $(P_t)_{t\geq 0}$. Il est alors possible d'obtenir des inégalités fonctionnelles à base de Φ-entropies et d'énergies modifiées liées aux transformations $A - B - C$ de Φ, par interpolation semi-groupe, à la fois pour P_t et pour la loi invariante $\mathcal{P}(\rho)$. Une autre approche consiste à établir des inégalités sur l'espace à deux points puis à utiliser la tensorisation des Φ-entropies et le théorème central limite poissonien. À titre d'exemple, cela permet d'établir que pour toute fonction f raisonnable

$$\mathbf{Ent}^{\Phi}_{\mathcal{P}(\rho)}(f) \leq \rho\, \mathbf{E}_{\mathcal{P}(\rho)}(A^{\Phi}(f, Df)).$$

En utilisant le théorème central limite, cette inégalité à base de A-transformée donne sa contre-partie optimale gaussienne à base de C-transformée. Par ailleurs, la décroissante de la Φ-entropie le long de $(P_t)_{t\geq 0}$ est équivalente à une inégaliré de Φ-Sobolev modifiée à base de B-transformée, qui n'est toutefois pas suffisante pour produire l'inégalité gaussienne optimale par passage à la limite. Plus généralement, il est possible d'étudier le comportement des inégalités pour P_t le long du procédé de type « limite fluide » qui fait apparaître le processus d'Ornstein-Uhlenbeck. Tous ces aspects sont développés dans [**A6**]. La file d'attente M/M/∞ constitue un processus remarquable, qui apparaît comme un exemple utile dans les travaux récents [Oll07] et [CDPP07].

	Loi de Gauss	**Loi de Poisson**
Convolution	Mouvement Brownien	Processus de Poisson
Ergodicité	Ornstein-Uhlenbeck	File d'attente M/M/∞

TAB. 1.2 – Des inégalités de Φ-Sobolev pour la gaussienne peuvent être obtenues par interpolation en utilisant le semi-groupe de convolution du mouvement brownien sur l'intervale de temps $[0, t]$, ou bien en utilisant le semi-groupe ergodique d'Ornstein-Uhlenbeck sur l'intervalle de temps $[0, \infty]$. De même, pour la loi de Poisson, des inégalités à énergie modifiée peuvent être obtenues en utilisant le semi-groupe de convolution du processus de Poisson sur l'intervalle de temps $[0, t]$ ou bien en utilisant le semi-groupe ergodique de la file d'attente M/M/∞ sur l'intervalle de temps $[0, \infty]$. Malgré tout, le cas poissonien est moins souple, notamment pour les inégalités locales.

1.2.2 Noyau de la chaleur sur le groupe d'Heisenberg

Il est bien connu que le couple formé par le mouvement brownien plan et son aire de Lévy conduit au noyau de la chaleur sur le groupe d'Heisenberg $\mathbb{H}$ [Bau04, Neu96, Mon02]. Ce groupe constitue un exemple élémentaire d'espace sous-riemannien. L'opérateur de diffusion associé est hypo-elliptique mais non-elliptique, ce qui rend les techniques de courbure développées dans le cadre riemannien inopérantes. Malgré tout, certaines sous-commutations subsistent. Dans [**A12**], nous proposons des preuves simples de ces sous-commutations, et nous en explorons les conséquences. Il est commode d'identifier $\mathbb{H}$ à $\mathbb{R}^3$, muni du générateur infinitésimal $L = X^2 + Y^2$ où les champs de vecteurs X et Y sont donnés par $X = \partial_x - \frac{y}{2}\partial_z$ et $Y = \partial_y + \frac{x}{2}\partial_z$. L'opérateur L est auto-adjoint pour la mesure de Lebesgue. La matrice de diffusion de L est dégénérée (non-ellipticité) mais $Z = [X, Y] = \partial_z$ et $[X, Z] = [Y, Z] = 0$ (hypo-ellipticité au sens de Hörmander). Le semi-groupe $(P_t)_{t\geq 0} = (e^{tL})_{t\geq 0}$ associé à L admet une densité lisse, et possède une structure de convolution liée aux dilatations et à la structure de groupe. L'opérateur carré du champ est donné par $\Gamma(f, f) = |\nabla f|^2$ où $\nabla f = (Xf, Yf)$. Un calcul simple montre que la courbure de L n'est pas minorée. Malgré tout, Driver & Melcher [DM05] ont montré l'existence d'une constante C_2 telle que pour tout $t \geq 0$ et f assez lisse,

$$|\nabla P_t f|^2 \leq C_2 P_t(|\nabla f|^2).$$

Cette sous-commutation fournit des inégalités de Poincaré et de Poincaré inverse locales. Dans [**A12**], nous déterminons la constante optimale de l'inégalité de Poincaré inverse locale, et nous proposons une preuve élémentaire de la sous-commutation de Driver & Melcher, basée sur l'exploitation des dilatations. Plus récemment, H.-Q. Li [Li06] a renforcé le résultat de Driver & Melcher en montrant l'existence d'une constante C_1 telle que pour tout $t \geq 0$ et f assez lisse

$$|\nabla P_t f| \leq C_1 P_t(|\nabla f|).$$

Dans [**A12**], nous proposons une preuve conceptuellement plus simple de cette sous-commutation, basée sur une partition de l'unité et des inégalités de type Cheeger sur les boules et leur complémentaire. Nous proposons également une autre preuve par prolongement complexe du noyau de la chaleur, qui met en lumière la subtilité de ce noyau par rapport au noyau de la chaleur usuel sur $\mathbb{R}^3$. Nous montrons également dans [**A12**] que l'inégalité de H.-Q. Li entraîne des inégalités de Φ-Sobolev locales, ainsi que des inégalités isopérimétriques locales du type Cheeger et Bobkov. L'idée principale ici est de voir $P_t(\cdot)(0)$ comme une sorte de loi gaussienne sur $\mathbb{H}$, et d'adapter les preuves riemanniennes des inégalités en les réduisant à l'usage de la sous-commutation de H.-Q. Li. Cela fournit des inégalités isopérimétriques pour $P_t(\cdot)(0)$ pour lesquelles la notion de bord est liée via ∇ à la distance de Carnot-Carathéodory naturellement associée à $\mathbb{H}$.

1.2.3 Concentration et inégalités pour les mélanges

Un mélange est une combinaison convexe de lois. Malgré cette définition simple, un mélange peut être beaucoup plus subtil que ses composants. Mélanger des lois gaussiennes par exemple peut faire apparaitre un potentiel avec de multiples puits. Nous étudions dans [**A16**] certaines propriétés fines des mélanges relatives à la concentration de la mesure et aux inégalités de type Poincaré et Sobolev logarithmique. Nous fournissons pour des mélanges généraux des bornes sur la concentration pour les fonctions Lipschitz, qui font intervenir le diamètre des lois mélangées pour la distance de transport W_1. Notre analyse des inégalités de type Sobolev logarithmique pour des mélanges à deux composants de la forme $p\mu_1 + q\mu_0$ révèle des liens naturels avec une forme d'isopérimétrie pour les bandes et avec une interpolation par transport à support prescrit. Nous montrons que la constante de Poincaré d'un mélange à deux composants peut rester bornée lorsque $\min(p, q) \to 0$, tandis que la constante de Sobolev logarithmique peut exploser. Ce résultat contre intuitif ne se réduit pas à un phénomène de déconnexion de support, mais provient du comportement de la loi de Bernoulli asymétrique sous-jacente. Ce travail montre que pour les mélanges, la concentration de la mesure est beaucoup plus stable que les inégalités de Sobolev logarithmique. Nous illustrons notre étude par une collection d'exemples concrets laissant entrevoir la grande variété de comportement possibles pour les mélanges.

1.2.4 Quelques perspectives

Le travail mené dans [**A12**] constitue un premier pas vers une étude plus générale des diffusions hypo-elliptiques non-elliptiques. Les questions ouvertes sont nombreuses, à commencer par l'étude du noyau de la chaleur sur les groupes d'Heisenberg d'ordre plus élevé, ou encore l'étude de la mesure de Wiener du groupe d'Heisenberg. Idéalement, il serait remarquable de pouvoir procéder par comparaison à un modèle de référence à partir des informations contenues dans l'algèbre de Lie des champs de vecteurs. Dans une toute autre direction, j'ai commencé l'étude probabiliste d'équations de Fokker-Planck cinétiques hypo-elliptiques associées à un modèle de déplacement en biologie. Le travail sur les mélanges mené dans [**A16**] suggère de poursuivre par exemple la piste de l'interpolation par transport à contrainte de support. Il me semble également intéressant d'explorer les conséquences de ce travail pour l'étude statistique de certains modèles comprenant des mélanges.

1.3 Matrices aléatoires

L'étude des propriétés spectrales de grandes matrices aléatoires constitue un champ de recherche très actif ces dernières années, comme en témoignent les livres [Meh04, HP00, BS06].

1.3.1 Ensemble Dirichlet Markov

Je me suis intéressé dans [**A14**] à la loi uniforme sur le polytope de Markov constitué par les matrices markoviennes $n \times n$. Cette loi revient à considérer des matrices markoviennes aléatoires dont les lignes sont i.i.d. de loi de Dirichlet $\mathcal{D}_n(1, \dots, 1)$. Les entrées de ces matrices aléatoires ne sont pas centrées et ne sont pas indépendantes. Malgré tout, si M est une telle matrice aléatoire, alors la loi

$$\frac{1}{n} \sum_{k=1}^{n} \delta_{s_k(\sqrt{n}M)}$$

formée avec les valeurs singulières de $\sqrt{n}M$ converge presque sûrement lorsque $n \to \infty$ vers une loi du quart de cercle. Ce théorème de type Marchenko-Pastur peut être obtenu très simplement en utilisant la méthode développée dans [Aub06], et complète les résultats qualitatifs déjà connus [GONS00, GN03]. L'ensemble des matrices markoviennes $n \times n$ forme également un semi-groupe compact pour le produit matriciel. Il est alors naturel de s'interroger sur l'existence d'une loi invariante par translation. Je montre dans [**A14**] que seule les translations associées aux matrices de permutation laissent stable la loi Dirichlet Markov. De plus, la théorie des semi-groupes compacts [Ros71, ch. 5] permet d'établir qu'il n'existe pas de loi non triviale sur le polytope de Markov qui soit stable par toutes les translations. Le travail mené dans [**A14**] fait apparaître de nombreuses questions ouvertes, en liaison avec l'actualité de la théorie des matrices aléatoires.

1.3.2 Loi du cercle et théorèmes de Girko-Bai

Le spectre d'une matrice markovienne est inclus dans le disque unité du plan complexe et contient toujours la valeur 1. Lorsque M est de type Dirichlet Markov, des simulations mentionnées dans [**A14**] suggèrent que la loi

$$\frac{1}{n}\sum_{k=1}^{n}\delta_{\lambda_k(\sqrt{n}M)}$$

formée avec le spectre complexe de $\sqrt{n}M$ converge presque sûrement quand $n \to \infty$ vers la loi uniforme sur le disque unité. Cette conjecture résiste encore à mes efforts. En revanche, je suis parvenu dans [**A13**] a établir que si X est une matrice aléatoire $n \times n$ à entrées i.i.d. de moyenne $m \in \mathbb{C}$, de variance 1, et de moment d'ordre 4 fini, alors la loi empirique formée avec le spectre complexe de $n^{-1/2}X$ converge presque sûrement lorsque $n \to \infty$ vers la loi uniforme sur le disque unité. Ce résultat de type Girko-Bai généralise au cas $m \neq 0$ le résultat obtenu en 2007 par Pan & Zhou [PZ07]. La preuve qui se trouve dans [**A14**] consiste à mêler la méthode à base de potentiel logarithmique de Pan & Zhou [PZ07] avec la borne polynomiale sur la plus petite valeur singulière de Tao & Vu [TV07].

1.3.3 Matrices markoviennes réversibles aléatoires

Dans [**A17**], nous construisons des matrices markoviennes réversibles aléatoires à partir de graphes à poids aléatoires. Nous ramenons l'étude du spectre réel de ces matrices aléatoires à celui de certaines matrices symétriques aléatoires à entrées dépendantes. Ces modèles peuvent également être vus comme des chaînes de Markov réversibles finies en environnement aléatoire. Nous étudions deux cas particuliers. Le premier provient du graphe complet, pour lequel nous montrons que la loi spectrale limite est une loi du demi-cercle de type Wigner. Le second provient du graphe chaîne (vie et mort), pour lequel nous déterminons la suite des moments de la loi spectrale limite, qui n'est pas une loi de type Wigner mais qui est reliée à la loi de l'arc-sinus. Dans les deux cas, nous étudions le comportement du rayon spectral utile et du trou spectral, révélant des comportements tout à fait différents. La méthode est basée dans le premier cas sur l'usage d'une loi des grands nombres uniforme similaire à celle apparaissant dans [BY93] et dans le second cas sur la méthode des moments comme dans [Pop08]. Le graphe complet muni de poids exponentiels fournit un ensemble de matrices markoviennes réversibles aléatoires dont les lignes sont dépendantes et de même loi de Dirichlet. Il peut être vu en quelque sorte comme un Ensemble Dirichlet Markov Reversible. Cette observation suggère l'utilisation d'une construction graphique à base de graphes orientés à poids pour définir des lois sur le polytope de Markov considéré dans [**A14**].

1.3.4 Quelques perspectives

Un théorème du type Girko-Bai pourrait être obtenu pour l'Ensemble Dirichlet Markov de [**A14**] en utilisant une borne polynomiale sur la plus petite valeur singulière, similaire à celle de Tao & Vu. Ici, la nature continue de la loi de Dirichlet permet d'éviter la complexité de l'argument générique de Tao & Vu. Plus généralement, il est probablement possible d'étendre le résultat de Tao & Vu à des matrices dont les lignes sont i.i.d. de loi échangeable et continue, dans l'esprit d'un résultat de Chatterjee [Cha06]. Le théorème de type Girko-Bai obtenu dans [**A13**] constitue un cas particulier du problème plus général de la mesure spectrale limite de matrices aléatoires de la forme $X + A$ où A est déterministe mais pas forcément de rang 1. Cette question difficile, que la théorie des probabilités libres n'a pas élucidé, est explorée actuellement par Krishnapur & Vu (communication personnelle). Les modèles et l'approche de [**A17**] suggèrent de nombreux développements, entre graphes aléatoires et matrices aléatoires. Le cas des graphes à poids aléatoires à queue lourde fait sans doute apparaître des lois de Poisson-Dirichlet déformée via un processus ponctuel. D'autre part, le problème de la fluctuation des valeurs propres extrémales des matrices markoviennes aléatoires reste ouvert. Enfin, j'aimerais mentionner le problème ouvert des propriétés spectrales du polytope de Birkhoff formé par les matrices bistochastiques.

Bibliographie exogène

[Aub06] G. Aubrun – « Random points in the unit ball of l_p^n », *Positivity* **10** (2006), no. 4, p. 755–759. 1.3.1

[Bak94] D. Bakry – « L'hypercontractivité et son utilisation en théorie des semigroupes », Lectures on probability theory. École d'été de probabilités de St-Flour 1992, Lecture Notes in Math., vol. 1581, Springer, Berlin, 1994, p. 1–114. 1.2

[Bau04] F. Baudoin – *An introduction to the geometry of stochastic flows*, Imperial College Press, London, 2004. 1.2.2

[BBLM05] S. Boucheron, O. Bousquet, G. Lugosi et P. Massart – « Moment inequalities for functions of independent random variables », *Ann. Probab.* **33** (2005), no. 2, p. 514–560. 1.2.1

[Bre67] L. Bregman – « The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming. », *Zh. Vychisl. Mat. Mat. Fiz.* **7** (1967), p. 620–631 (Russian). 1.2.1

[BS06] Z. D. Bai et J. W. Silverstein – *Spectral Analysis of Large Dimensional Random Matrices*, Mathematics Monograph Series 2, Science Press, Beijing, 2006. 1.3

[BY93] Z. D. Bai et Y. Q. Yin – « Limit of the smallest eigenvalue of a large-dimensional sample covariance matrix », *Ann. Probab.* **21** (1993), no. 3, p. 1275–1294. 1.3.3

[CDPP07] P. Caputo, P. Dai Pra et G. Posta – « Convex entropy decay via the Bochner-Bakry-émery approach », prépublication, `arXiv :0712.2578 [math.PR]`, 2007. 1.2.1

[CDR07] S. Chaudhuri, M. Drton et T. S. Richardson – « Estimation of a covariance matrix with zeros », *Biometrika* **94** (2007), no. 1, p. 199–216. 1.1.3

[Cha06] S. Chatterjee – « A generalization of the Lindeberg principle », *Ann. Probab.* **34** (2006), no. 6, p. 2061–2076. 1.3.4

[CHL97] M. Capitaine, E. P. Hsu et M. Ledoux – « Martingale representation and a simple proof of logarithmic Sobolev inequalities on path spaces », *Electron. Comm. Probab.* **2** (1997), p. 71–81 (electronic). 1.2.1

[CP07] P. Caputo et G. Posta – « Entropy dissipation estimates in a zero-range dynamics », *Probab. Theory Related Fields* **139** (2007), no. 1-2, p. 65–87. 1.1

[DG95] M. Davidian et D. M. Giltinan – *Nonlinear Models for Repeated Measurement Data*, Chapman & Hall, 1995. 1.1

[DM04] P. Del Moral – *Feynman-Kac formulae*, Probability and its Applications (New York), Springer-Verlag, New York, 2004, Genealogical and interacting particle systems with applications. 1.1

[DM05] B. K. Driver et T. Melcher – « Hypoelliptic heat kernel inequalities on the Heisenberg group », *J. Funct. Anal.* **221** (2005), no. 2, p. 340–365. 1.2.2

[Dur95] R. Durrett – « Ten lectures on particle systems », Lectures on probability theory (Saint-Flour, 1993), Lecture Notes in Math., vol. 1608, Springer, Berlin, 1995, p. 97–201. 1.1

[FFSK00] L. Friberg, A. Freijs, M. Sandström et M. Karlsson – « Semiphysiological model for the time course of leukocytes after varying schedules of 5-fluorouracil in rats », *J. Pharmacol. Exp. Therap.* **2** (2000), no. 295, p. 734–740. 1.1.2

[FK03] L. Friberg et M. Karlsson – « Mechanistic models for myelosuppression », *Invest. New Drug* (2003), no. 21, p. 183–194. 1.1.2

[GI07] I. Gentil et C. Imbert – « The Lévy-Fokker-Planck equation : Φ-entropies and convergence to equilibrium », à paraître dans Asymptotic Analysis, 2007. 1.2.1

[GN03] G. Goldberg et M. Neumann – « Distribution of subdominant eigenvalues of matrices with random rows », *SIAM J. Matrix Anal. Appl.* **24** (2003), no. 3, p. 747–761 (electronic). 1.3.1

[GONS00] G. Goldberg, P. Okunev, M. Neumann et H. Schneider – « Distribution of subdominant eigenvalues of random matrices », *Methodol. Comput. Appl. Probab.* **2** (2000), no. 2, p. 137–151. 1.3.1

[GP82] M. Gibaldi et D. Perrier – *Pharmacokinetics, 2nd edition*, Dekker, 1982. 1.1

[Gro06] L. Gross – « Hypercontractivity, logarithmic Sobolev inequalities, and applications : a survey of surveys », Diffusion, quantum theory, and radically elementary mathematics, Math. Notes, vol. 47, Princeton Univ. Press, Princeton, NJ, 2006, p. 45–73. 1.2

[GZ03] A. Guionnet et B. Zegarlinski – « Lectures on logarithmic Sobolev inequalities », Séminaire de Probabilités, XXXVI, Lecture Notes in Math., vol. 1801, Springer, Berlin, 2003, p. 1–134. 1.2

[Hel02] B. Helffer – *Semiclassical analysis, Witten Laplacians, and statistical mechanics*, Series on Partial Differential Equations and Applications, vol. 1, World Scientific Publishing Co. Inc., River Edge, NJ, 2002. 1.2

[HP00] F. Hiai et D. Petz – *The semicircle law, free random variables and entropy*, Mathematical Surveys and Monographs, vol. 77, American Mathematical Society, Providence, RI, 2000. 1.3

[Jac57] J. R. Jackson – « Networks of waiting lines », *Operations Res.* **5** (1957), p. 518–521. 1.1

[Jac63] — , « Jobshop-like Queuing Systems », *Management Sciences* **10** (1963), no. 1, p. 131–142. 1.1

[JS93] J. Jacquez et C. Simon – « Qualitative theory of compartmental systems », *SIAM Rev.* **35** (1993), no. 1, p. 43–79. 1.1

[Kel79] F. P. Kelly – *Reversibility and stochastic networks*, John Wiley & Sons Ltd., Chichester, 1979, Wiley Series in Probability and Mathematical Statistics. 1.1, 1.1

[KS94] J. Keilson et L. D. Servi – « Networks of nonhomogeneous $M/G/\infty$ systems », *J. Appl. Probab.* **31A** (1994), p. 157–168, Studies in applied probability. 1.1.2

[Li06] H.-Q. Li – « Estimation optimale du gradient du semi-groupe de la chaleur sur le groupe de Heisenberg », *J. Funct. Anal.* **236** (2006), no. 2, p. 369–394. 1.2.2

[Lig05] T. M. Liggett – *Interacting particle systems*, Classics in Mathematics, Springer-Verlag, Berlin, 2005, Reprint of the 1985 original. 1.1

[LO00] R. Latała et K. Oleszkiewicz – « Between Sobolev and Poincaré », Geometric aspects of functional analysis, Springer, Berlin, 2000, p. 147–168. 1.2.1

[Mas07] P. Massart – *Concentration inequalities and model selection*, Lecture Notes in Mathematics, vol. 1896, Springer, Berlin, 2007, Lectures from the 33rd Summer School on Probability Theory held in Saint-Flour, July 6–23, 2003, With a foreword by Jean Picard. 1.2.1

[Meh04] M. L. Mehta – *Random matrices*, third éd., Pure and Applied Mathematics (Amsterdam), vol. 142, Elsevier/Academic Press, Amsterdam, 2004. 1.3

[MI06] P. Macheras et A. Iliadis – *Modeling in biopharmaceutics, pharmacokinetics, and pharmacodynamics*, Interdisciplinary Applied Mathematics, vol. 30, Springer, New York, 2006, Homogeneous and heterogeneous approaches. 1.1

[MK00] J. H. Matis et T. R. Kiffe – *Stochastic population models*, Lecture Notes in Statistics, vol. 145, Springer-Verlag, New York, 2000, A compartmental perspective. 1.1

[Mon02] R. Montgomery – *A tour of subriemannian geometries, their geodesics and applications*, Mathematical Surveys and Monographs, vol. 91, American Mathematical Society, Providence, RI, 2002. 1.2.2

[Neu96] D. Neuenschwander – *Probabilities on the Heisenberg group*, Lecture Notes in Mathematics, vol. 1630, Springer-Verlag, Berlin, 1996, Limit theorems and Brownian motion. 1.2.2

[Oll07] Y. Ollivier – « Ricci curvature of Markov chains on metric spaces », prépublication, `arXiv :math/0701886 [math.PR]`, 2007. 1.2.1

[Pfa88] J. Pfanzagl – « Consistency of maximum likelihood estimators for certain nonparametric families, in particular : mixtures », *J. Statist. Plann. Inference* **19** (1988), no. 2, p. 137–158. 1.1.1

[Pop08] I. Popescu – « General tridiagonal random matrix models, limiting distributions and fluctuations », `arXiv :math/0610827 [math.PR]` à parraître dans Probab. Theory Related Fields, 2008. 1.3.3

[PZ07] G. Pan et W. Zhou – « Circular law, extreme singular values and potential theory », prépublication `arXiv :0705.3773 [math.PR]`, 2007. 1.3.2

[Rob03] P. Robert – *Stochastic networks and queues*, french éd., Applications of Mathematics (New York), vol. 52, Springer-Verlag, Berlin, 2003, Stochastic Modelling and Applied Probability. 1.1

[Ros71] M. Rosenblatt – *Markov processes. Structure and asymptotic behavior*, Springer-Verlag, New York, 1971, Die Grundlehren der mathematischen Wissenschaften, Band 184. 1.3.1

[Roy99] G. Royer – *Une initiation aux inégalités de Sobolev logarithmiques*, Société Mathématique de France, Paris, 1999. 1.2

[RT89] M. Rowland et T. N. Tozer – *Clinical Pharmacokinetics : Concepts and Applications*, Lea & Febige, 1989. 1.1

[ST88] K. Schuhmacher et H. Thieme – « Some theoretical and numerical aspects of modelling dispersion in the development of ectotherms », *Comput. Math. Appl.* **15** (1988), no. 6-8, p. 565–594, Hyperbolic partial differential equations. V. 1.1.2

[TV07] T. Tao et V. Vu – « Random Matrices : The circular law », prépublication `arXiv :0708.2895 [math.PR]`, 2007. 1.3.2

[Wu00] L. Wu – « A new modified logarithmic Sobolev inequality for Poisson point processes and several applications », *Probab. Theor. Relat. Fields* **118** (2000), no. 3, p. 427–438. 1.2.1