

Centre National de la Recherche Scientifique (CNRS)
Programme inter-départements Math-STIC

Chaînes de Markov Cachées
et Filtrage Particulaire

Équipes participantes :

LSP, Toulouse (UMR C5583)
IRISA, Rennes (UMR 6074)
CEREMADE, Paris (UMR 7534)
Laboratoire de Mathématique, Orsay (UMR C8628)
LMC, Grenoble (UMR 5523)
LTCI, Paris (URA 820)

Financement demandé :

150 kF

Correspondants :

François LeGland
IRISA / INRIA
Campus de Beaulieu
35042 RENNES Cédex
02 99 84 73 62
legland@irisa.fr

Éric Moulines
LTCI, ENST
46, rue Barrault
75634 PARIS Cédex 13
01 45 81 77 03
moulines@tsi.enst.fr

1 Motivations

Au cours des cinquante dernières années, la modélisation statistique s'est peu à peu imposée comme un paradigme central dans tous les domaines relevant du traitement et de la communication de l'information (traitement du signal, théorie de l'information, reconnaissance de formes, fouille de données, etc.) Dans ce contexte, les modèles Markov cachés (plus souvent désigné par l'acronyme HMM pour *Hidden Markov Models*), constituent une classe de modèles à l'origine de quelques unes des plus spectaculaires avancées de ces dernières années. Le principe central de cette famille de modèles est de supposer l'existence d'un état caché (latent) évoluant suivant une dynamique markovienne, les observations étant des fonctions (déterministes ou aléatoires) de cet état caché. Les modèles HMM sont très flexibles du fait de l'introduction de variables latentes (non observées), qui permettent de modéliser des structures de dépendances temporelles complexes. La structure markovienne des variables latentes permet l'utilisation de procédures numériques de traitement dont la complexité de calcul reste compatible avec les moyens informatiques actuels. L'exemple le plus élémentaire de chaînes de Markov cachées est incontestablement le modèle d'état linéaire gaussien, à la base de nombreux développements en traitement du signal et des images. Au cours de ces dernières années, de nombreuses variantes de ce paradigme de base ont été étudiées dans des applications aussi variées que la reconnaissance de parole, la poursuite en environnement complexe, la détection et le diagnostic de pannes, la bio-informatique, la modélisation et le contrôle de réseaux, les communications numériques, etc. (voir [2] pour une bibliographie plus complète).

L'objectif du présent projet est de capitaliser sur ces avancées récentes pour développer de nouveaux éléments méthodologiques concernant les points qui, à notre sens, constituent les limites des connaissances actuelles :

- (i) le développement d'une théorie complète de l'inférence des chaînes de Markov cachées, incluant l'estimation paramétrique pour des espaces d'états généraux, la validation et la sélection de modèles, les ouvertures vers la modélisation semi-paramétrique et non-paramétrique,
- (ii) l'approximation numérique des procédures de filtrage optimal via les méthodes particulières, et leur application à la mise en œuvre d'algorithmes statistiques pour les chaînes de Markov cachées.

Cette proposition regroupe des chercheurs rattachés à des unités de recherche relevant des départements SPM et STIC du CNRS, qui ont tous contribué au domaine dans les années récentes (voir bibliographie en fin du document). L'objectif de cette proposition est de renforcer les synergies entre ces deux communautés, et d'établir des liens forts entre les recherches méthodologiques dans ces domaines et les applications. Le financement demandé (150 kF) est destiné à couvrir les frais de visites entre les partenaires du projet, d'invitation de chercheurs étrangers sur des courtes durées, et de participation à des workshops ou des conférences spécialisés sur les thèmes du projet. Sous une forme qui reste à préciser (session spéciale dans une conférence, organisation d'un workshop spécialisé, etc.), le projet est également susceptible de donner lieu à une manifestation spécifique, ouverte à des participants extérieurs. Un site web sera créé pour mettre en valeur les ressources (références bibliographiques, preprints, codes et données de simulation) mises à disposition par les participants dans le cadre du projet.

2 Etat de l'art

2.1 Chaînes de Markov cachées

L'étude de ces modèles a réellement débuté dans les années soixante, par l'analyse des modèles d'états linéaires gaussiens, qui a suscité un engouement fort dans la communauté automatique et traitement du signal. L'essentiel des résultats pour ces modèles ont été obtenus dans les années soixante-dix, mais certains problèmes essentiels (comme la stabilité du filtre de Kalman) n'ont trouvé de solutions que très récemment [15]. L'extension aux modèles d'états non-linéaires a été entreprise très tôt, mais peu de résultats réellement satisfaisants ont été obtenus jusqu'à l'apparition des méthodes de filtrage particulière. Les techniques de linéarisation sont en effet inadaptées aux modèles fortement non-linéaires, et les méthodes numériques d'intégration ne permettent pas de traiter les problèmes de grandes dimensions.

En parallèle se sont développées dès la fin des années soixante des études sur les HMM à états discrets (le processus latent étant une chaîne de Markov prenant ses valeurs dans un ensemble fini d'états). Les liens entre ces deux modèles n'ont d'ailleurs pas été mis en évidence avant le début des années quatre-vingt. Les modèles discrets ont connu un vif succès tant en traitement de parole (les HMM forment l'élément de base des systèmes de reconnaissance de parole) qu'en bio-informatique (où les HMM sont utilisés pour la segmentation et le séquençage de génomes).

Si l'on excepte les modèles d'états linéaires gaussiens et les modèles discrets à observations discrètes (pour lesquels des réponses, d'ailleurs tout à fait non triviales, peuvent être apportés), les problèmes même les plus élémentaires liés à l'inférence des paramètres des HMM par des méthodes de vraisemblance n'ont trouvé de solutions que très récemment. Les résultats existant à ce jour ne sont que partiellement satisfaisants car ils ont été obtenus sous des hypothèses restrictives de mélange uniforme de la chaîne de sous-jacente, hypothèse mécaniquement vérifiée pour des chaînes à états finis mais qui s'avère très restrictive pour des chaînes à états continus, tout particulièrement lorsque l'espace d'états est non-compact. La mise en œuvre pratique des algorithmes statistiques obtenus passe nécessairement par les méthodes de simulation, et tout particulièrement, les techniques particulières (voir ci-dessous).

Pour la modélisation de processus, le choix de l'espace d'états de la chaîne cachée est un problème important. En particulier, lorsque l'espace d'état de la chaîne est fini, le choix du nombre d'états (ordre du modèle HMM) est un paramètre important de la modélisation (problème de choix de modèle). Ce problème est très difficile, mais certaines avancées ont eu lieu récemment. En particulier, la consistance de l'estimateur du maximum de vraisemblance pénalisée a été établie (pour des chaînes finies à émission continue) [13], et des résultats d'optimalité de la vitesse ont été obtenus [14], sous certaines hypothèses restrictives.

2.2 Filtrage particulière

L'étude des HMM à état général soulève immédiatement la question du calcul, fût-il seulement approché, du filtre optimal et de quantités connexes, comme par exemple la dérivée du filtre optimal par rapport à un paramètre inconnu du modèle (plus généralement, le filtre optimal linéarisé tangent), ou bien d'autres filtres associés au calcul récursif d'espérances conditionnelles

(par rapport aux observations) de fonctionnelles additives de l'état caché (un exemple important est la fonction auxiliaire de l'algorithme EM).

Une réponse extrêmement attractive et prometteuse a été proposée récemment (il y a moins de dix ans) sous le nom générique de filtrage particulaire, et fait l'objet de recherches très actives, tant dans le domaine de la mise en œuvre pratique, que dans celui de l'extension à des modèles et des problèmes plus généraux. Les principaux résultats mathématiques sont présentés dans l'article de synthèse [7], et de nombreux aspects théoriques et pratiques sont abordés dans l'ouvrage collectif [11].

Dans sa version la plus simple, la méthode consiste à approcher le filtre optimal à l'aide de la distribution de probabilité empirique d'un système de particules. Entre deux instants d'observation, les particules se déplacent de façon indépendante selon la dynamique de l'état caché, et sitôt qu'une nouvelle observation est disponible, un rééchantillonnage a lieu, où les particules sont choisies en fonction de leur adéquation à la nouvelle observation (quantifiée par la fonction de vraisemblance). Sous l'effet du rééchantillonnage, qui constitue l'étape essentielle de la méthode, les particules se concentrent automatiquement dans les régions d'intérêt de l'espace d'état. La méthode est très facile à mettre en œuvre, puisqu'il suffit de simuler de façon indépendante des trajectoires de l'état caché, l'interaction ayant lieu uniquement lors du rééchantillonnage.

Si les méthodes de filtrage particulaire ont permis de réaliser des avancées très significatives, beaucoup de problèmes méthodologiques et pratiques restent ouverts. L'expérience accumulée à ce jour a mis en particulier en relief l'importance des méthodes de proposition de nouvelles particules (filtres auxiliaires) ainsi que la nécessité d'utiliser des techniques de réduction de dimension par intégration conditionnelle explicite (inspirées des méthodes de Rao-Blackwell en statistique bayésienne). Ces méthodes alternatives sont encore aujourd'hui très mal connues et maîtrisées.

3 Objectifs

Comme nous l'avons souligné dans l'introduction, l'objectif principal de cette proposition est de **fédérer** les efforts des chercheurs et de faire émerger une véritable **communauté** nationale autour de ces thèmes trans-disciplinaires.

Les objectifs scientifiques que nous nous assignons dans ce projet sont multiples et ambitieux. Nous avons séparé un peu artificiellement les objectifs pour les filtres particuliers et les chaînes de Markov cachées, bien que les progrès dans ces deux domaines soient très étroitement liés.

3.1 Estimation pour les chaînes de Markov cachées

Comme nous l'avons souligné dans l'état de l'art, la compréhension que nous avons des problèmes d'inférence dans les chaînes de Markov cachées est encore très incomplète. Un des objectifs principaux de cette étude sera de parvenir à une théorie de l'inférence paramétrique complète et permettant de prendre en compte les espaces d'états non-compacts. Il est vraisemblable que les extensions au cas non-compact tireront partie des études concernant la stabilité asymptotique des filtres non-linéaires, pour lesquels des résultats récents ont été obtenus [6],

[16]. Nous chercherons aussi à résoudre les problèmes de sélection de modèles, en étendant les études existantes sur le choix d'ordre. De nombreux résultats sont encore attendus sur le sujet, en particulier

- l'analyse des techniques de vraisemblance pénalisées dans le cas général — les résultats de [12] montrent que le choix de la pénalisation n'est pas standard,
- des résultats à horizon fini (et pas seulement asymptotiques), en utilisant des techniques de concentration tels que ceux de Katalin Marton,
- la calibration des pénalisations.

La problématique liée au choix de modèle ouvre naturellement vers les méthodes statistiques d'inférence semi- et non-paramétriques, qui n'ont jamais été considérées pour de tels modèles, en dépit de leur importance pratique.

3.2 Applications statistiques du filtrage particulière

Un autre objectif de cette proposition concerne l'application des filtres particulières à divers problèmes statistiques pour les HMM à état général

- identification récursive des paramètres du modèle,
- surveillance et diagnostic précoce de petits changements dans les paramètres du modèle.

Dans ces deux problèmes, il est nécessaire de calculer aussi le filtre optimal linéarisé tangent, en plus du filtre optimal lui-même. Sous une hypothèse naturelle et peu restrictive, on a montré récemment [4, 5] que le filtre optimal linéarisé tangent est une mesure signée de masse nulle, absolument continue par rapport au filtre optimal. De la même façon, on a montré récemment [1] que le filtre associé au calcul récursif d'espérances conditionnelles de fonctionnelles additives de l'état caché est une mesure éventuellement signée, absolument continue par rapport au filtre optimal.

Il est alors naturel d'approcher ces filtres dérivés à l'aide du même système de particules déjà utilisé pour approcher le filtre optimal, et d'affecter chaque particule d'un poids éventuellement signé, représentatif de la densité du filtre dérivé par rapport au filtre optimal. On se propose d'étudier les algorithmes ainsi obtenus, en particulier d'obtenir des estimations de l'erreur d'approximation quand le nombre de particules tend vers l'infini. La variante séquentielle proposée dans [16] pour le filtre optimal, où le nombre de particules est choisi à chaque instant de façon à garantir un niveau d'erreur donné, sera adaptée aux filtres dérivés. Comme dans l'approximation particulière du filtre optimal [16], il faudra d'abord établir que l'erreur globale résulte de la propagation d'erreurs locales par les équations du filtre dérivé, et pourvu que ces équations vérifient une propriété de contraction, on pourra obtenir des estimations d'erreur uniformes en temps. L'étude des propriétés de contraction pour le filtre optimal et les filtres dérivés fait partie du programme de recherche de cette proposition sur les HMM à état général.

Une partie importante du travail consistera à mettre en œuvre sur différents exemples, les algorithmes proposés et leurs nombreuses variantes.

3.3 Nouvelles directions en filtrage particulière

Un des objectifs de cette proposition concerne les applications des systèmes de particules en interaction et avec branchements à la résolution de problèmes d'estimation de trajectoires

entières de l'état caché, et pas seulement de l'état courant, et plus généralement les formules de Feynman-Kac sur l'espace des trajectoires.

Si l'on interprète l'évolution de ces systèmes particuliers comme des processus de naissance et mort, la notion de ligne ancestrale d'un individu apparaît de façon naturelle. L'ensemble des ces généalogies permet de construire de façon markovienne l'arbre généalogique complet de la population du système à chaque étape.

De récents travaux [6] ont montré que, dans le cadre du filtrage, la mesure empirique associée à ce processus historique converge vers la loi conditionnelle de la trajectoire du signal par rapport à une suite d'observations. D'autre part ces distributions peuvent se représenter par des formules de Feynman-Kac dans l'espace des chemins et divers outils d'analyse permettent de dégager des résultats fins de convergence. En particulier des résultats de type propagation croissante du chaos permettent de quantifier l'indépendance entre les trajectoires ainsi que l'adéquation de leur loi avec la distribution conditionnelle limite.

Ces résultats nouveaux donnent lieu à des applications dans trois directions

- simulation de lois conditionnelles de processus, par exemple processus restreints à des domaines, simulation d'évolution de processus dans des environnements avec obstacles,
- étude des liens entre les processus de coalescence avec ces arbres généalogiques rétrogrades et les formules de type Feynman-Kac trajectorielles,
- régulation et contrôle optimal par sélection de trajectoires particulières.

Les études numériques se déroulent essentiellement autour de deux points

- étude fine des vitesses de convergence de propagation du chaos pour des blocs de particules et des horizons temporels — les résultats de ce type présentés dans [6] utilisent une hypothèse de régularité et de mélange qui n'est pas satisfaite dans ce cadre,
- étude des théorèmes limites classiques tels les théorèmes de la limite centrale et de Donsker, et recherche de vitesses de convergence exponentielle.

Un dernier objectif de cette proposition concerne la conception et l'analyse de filtres particuliers pour des HMM où les hypothèses classiques sont en partie relâchées. On s'intéressera par exemple aux modèles d'états où les bruits d'état et d'observation sont corrélés. Le cas des bruits gaussiens corrélés sera abordé en premier lieu.

Références

- [1] Olivier Cappé. Recursive computation of smoothed functionals of hidden Markovian processes using a particle approximation. *Monte Carlo Methods and Applications*, 7(1-2),81-92, 2001.
- [2] Olivier Cappé. Ten years of HMMs. <http://www.tsi.enst.fr/~cappe/docs/hmbib.html>.
- [3] Olivier Cappé, Christian Robert and Tobias Rydén. Reversible jump MCMC converging to birth-and-death MCMC and more general continuous time samplers. <http://www.ceremade.dauphine.fr/~xian/bdrjmcmc.ps.gz>.
- [4] Frédéric Cérou and François Le Gland. Efficient particle filters for residual generation in partially observed SDE's. In *Proceedings of the 39th IEEE Conference on Decision and Control, Sydney 2000*, pp. 1200-1205. December 2000.

- [5] Frédéric Cérou, François Le Gland, and Nigel J. Newton. Stochastic particle methods for linear tangent filtering equations. In José-Luis Menaldi, Edmundo Rofman, and Agnès Sulem, editors, *Optimal Control and Partial Differential Equations. In honour of professor Alain Bensoussan's 60th birthday*, pp. 231-240. IOS Press, Amsterdam, 2001.
- [6] Pierre Del Moral and Laurent Miclo. Genealogies and increasing propagations of chaos for Feynman-Kac and genetic models. *The Annals of Applied Probability*. (accepted for publication).
- [7] Pierre Del Moral and Laurent Miclo. Branching and interacting particle systems approximations of Feynman-Kac formulae with applications to nonlinear filtering. In Jacques Azéma, Michel Émery, Michel Ledoux, and Marc Yor, editors, *Séminaire de Probabilités XXXIV*, volume 1729 of *Lecture Notes in Mathematics*, pp. 1-145. Springer-Verlag, Berlin, 2000.
- [8] Randal Douc and Catherine Matias. Propriétés asymptotiques de l'estimateur de maximum de vraisemblance pour des modèles de Markov cachés généraux. *Comptes Rendus de l'Académie des Sciences, Série I*, 330(2),135-138, 2000.
- [9] Randal Douc and Catherine Matias. Asymptotic properties of the maximum likelihood estimator for general hidden Markov models. *Bernoulli*, 7(3),381-420, 2001.
- [10] Randal Douc, Éric Moulines and Tobias Rydén. Asymptotic properties of the maximum likelihood estimator in autoregressive models with Markov regime. *The Annals of Statistics*. (submitted, 2001).
- [11] Arnaud Doucet, Nando de Freitas and Neil Gordon, editors. *Sequential Monte Carlo Methods in Practice*. Statistics for Engineering and Information Science. Springer-Verlag, New York, 2001.
- [12] Elisabeth Gassiat and Christine Kéribin. The likelihood ratio test for the number of components in a mixture with Markov regime. *ESAIM: Probability and Statistics*, 4,25-52, 2000.
- [13] Elisabeth Gassiat. Likelihood ratio inequalities with applications to various mixtures. Prépublication d'Orsay 2001-20, Université Paris-Sud, April 2001. http://www.math.u-psud.fr/~biblio/pub/2001/fic/ppo_2001_20.ps.
- [14] Stéphane Boucheron and Elisabeth Gassiat. Optimal error exponents in hidden Markov models order estimation. Prépublication d'Orsay 2001-41, Université Paris-Sud, September 2001. http://www.math.u-psud.fr/~biblio/pub/2001/fic/ppo_2001_41.ps.
- [15] Lei Guo. Stability of recursive stochastic tracking algorithms. In *Proceedings of the 32nd IEEE Conference on Decision and Control, San Antonio 1993*, pp. 2062-2067. December 1993.
- [16] François Le Gland and Nadia Oudjane. Stability and uniform approximation of nonlinear filters using the Hilbert metric, and application to particle filters. Publication Interne 1404, IRISA, June 2001. <ftp://ftp.irisa.fr/techreports/2001/PI-1404.ps.gz>.

4 Description des équipes

4.1 Laboratoire de Statistique et Probabilités (LSP, UMR C5583), Université Paul Sabatier

Les recherches menées au LSP couvrent tous les domaines de l'aléatoire, à la fois théoriques et appliqués. Parmi les nombreuses orientations de recherche, les plus pertinentes pour la proposition sont les suivantes :

- systèmes de particules en interaction et applications (résolution d'équations non-linéaires et filtrage non-linéaire),
- analyse des processus et des semi-groupes de Markov,
- grandes déviations et applications,
- méthodes statistiques asymptotiques,
- traitement du signal et méthodes MCMC.

Le laboratoire a joué un rôle très actif dans l'animation de la communauté autour des systèmes de particules en interaction et du filtrage particulaire, grâce au soutien du CNRS dans le cadre du programme "Modélisation et Simulation Numérique" (projet 97N23 / 0019, "Méthodes Particulaire et Filtrage Non-linéaire"), et à travers l'organisation (ou la co-organisation) de workshops à Toulouse (avril et décembre 1997), à Rennes (juin 1998), à Cambridge (décembre 1999) et à Paris (IHP, juin 2001). Au niveau international, outre de nombreuses collaborations individuelles, on peut signaler la participation au réseau européen TMR sur l'analyse stochastique et ses applications (<http://sag.maths.ox.ac.uk/>) et au réseau MITACS sur la classification, la localisation et la poursuite d'objets en interaction, avec l'université d'Alberta (PINTS, <http://www.math.ualberta.ca/pints/>).

Chercheurs impliqués Pierre Del Moral (CR CNRS), Pascal Lezaud (chercheur au CENA), Laurent Miclo (CR CNRS).

4.2 Équipe Signaux, Modèles et Algorithmes (Sigma2), Institut de Recherche en Informatique et Systèmes Aléatoires (IRISA, UMR 6074)

Les objectifs de l'équipe **Sigma2** (projet commun CNRS / INRIA) sont la conception, l'analyse et la mise en œuvre d'algorithmes statistiques basés sur l'utilisation de modèles, pour l'identification, la surveillance et le diagnostic de systèmes industriels complexes. Les modèles considérés sont d'une part les modèles d'état de l'automatique stochastique, avec une importance croissante des modèles non-linéaires, et d'autre part des modèles partiellement stochastiques (HMM, réseaux de Petri, réseaux d'automates, etc.) sur des structures discrètes (arbres, graphes, etc.), par exemple pour modéliser les systèmes distribués à événements discrets.

Les contributions méthodologiques les plus importantes de l'équipe, et qui constituent les bases scientifiques de ses activités actuelles, sont

- l'utilisation de l'approche asymptotique locale pour la surveillance et le diagnostic des systèmes continus,
- le développement de filtres particuliers pour la statistique des modèles de Markov cachés à état général,
- la conception d'algorithmes approchés d'estimation d'état dans les modèles graphiques et les réseaux bayésiens, et leur application aux algorithmes turbo,
- et la conception d'algorithmes répartis de reconstruction d'état, de type Viterbi, pour la surveillance et le diagnostic des systèmes distribués à événements discrets.

Les principales applications considérées sont la surveillance et le diagnostic des structures mécaniques en vibration (automobile, aéronautique, génie civil), la surveillance et le diagnostic d'organes de véhicules automobiles, la modélisation de risque dans les applications critiques, et le diagnostic de pannes dans les réseaux de télécommunications.

Des partenariats académiques internationaux ont été établis au cours des dernières années, à travers la participation au réseau européen TMR sur l'identification des systèmes (<http://www.cwi.nl/~schuppen/ernsi/tmrsi.html>), et au réseau européen IHP sur les méthodes statistiques pour les modèles dynamiques stochastiques (DYNSTOCH, <http://www.math.ku.dk/~michael/dynstoch/>).

Chercheurs impliqués Frédéric Cérou (CR INRIA), Jean-Pierre LeCadre (DR CNRS), François LeGland (DR INRIA), Laurent Mevel (CR INRIA).

4.3 Centre de Recherche en Mathématiques de la Décision (CEREMADE, UMR 7534), Université Paris-Dauphine

L'équipe de Statistique, du fait de son appartenance au CEREMADE, laboratoire de Mathématiques Appliquées, a comme domaine principal d'application la modélisation financière et économique. A ce titre, elle s'intéresse à divers aspects de la modélisation des processus en temps discret comme en temps continu, à la fois sur le plan théorique de la convergence d'estimateurs, comme sur celui du calcul de ces estimateurs. A côté de ce thème commun aux divers membres du laboratoire, les directions de recherche propres à Ch. Robert sont les suivantes :

- choix bayésien de modèles et sélection de variables,
- modélisation bayésiennes des modèles à données latentes (mélanges, chaînes de Markov cachées, volatilité stochastique),
- simulation de lois non-standards par les méthodes de Monte Carlo par chaînes de Markov,
- calibration et évaluation d'algorithmes de Monte Carlo par chaînes de Markov.

En particulier, Ch. Robert a animé la partie française du réseau européen TMR sur les statistiques spatiales et computationnelles (<http://www.maths.lancs.ac.uk/~breyer/essn/>) de 1996 a 2001, dont le thème principal était le développement d'outils de simulation dans des systèmes complexes comme ceux rencontrés en statistique bayésienne et en modélisation spatiale. Il est également impliqué à ce titre dans la nouvelle proposition de réseau TMR soumise au printemps dernier.

Chercheurs impliqués Christian P. Robert (professeur).

4.4 Équipe Probabilités, Statistique et Modélisation, Laboratoire de Mathématique (UMR C8628), Université d'Orsay

L'équipe possède un double aspect théorique et appliqué, qui ne recouvre pas toujours la distinction entre calcul des probabilités et statistique, un équilibre très rare en France. Les thèmes actuels de travail de l'équipe pouvant interagir avec le travail envisagé sont les suivants :

- choix adaptatifs de modèles et théorèmes limites,
- estimation efficace dans des modèles statistiques, et en particulier dans les modèles à données latentes (modèles de mélanges, chaînes de Markov cachées),
- modélisation et statistique, problèmes inverses et applications au traitement statistique du signal.

Le thème “Chaînes de Markov Cachées” est un thème actif de recherches depuis plus de cinq ans, et des résultats novateurs ont été obtenus sur l'inférence paramétrique (thèse de Catherine Matias, sous la direction de Elisabeth Gassiat) et la sélection de modèles (Elisabeth Gassiat, en collaboration avec Stéphane Boucheron).

Chercheurs impliqués Stéphane Boucheron (CR CNRS, LRI), Elisabeth Gassiat (professeur), Marc Lavielle (professeur).

4.5 Équipe Statistique et Modélisation Stochastique (SMS), Laboratoire de Modélisation et Calcul (LMC, UMR 5523), Institut de Mathématiques Appliquées de Grenoble (IMAG)

Les recherches menées dans l'équipe Statistique et Modélisation Stochastique (SMS) du LMC portent sur des aspects théoriques et applicatifs variés de la statistique et des probabilités. Les principales orientations de recherche sont les suivantes :

- analyse des séries chronologiques et traitement du signal (séparation de sources, assimilation de données),
- inférence fonctionnelle (méthodes d'ondelettes, analyse de ruptures, tests fonctionnels),
- statistique des processus (estimation et filtrage de processus fractionnaires, fonctionnelles quadratiques),
- fiabilité et sûreté de fonctionnement (fiabilité des systèmes matériels et logiciels),
- algorithmes stochastiques, probabilités appliquées (filtrage particulière, méthodes de simulation de Monte-Carlo par chaîne de Markov).

La période 1998-2001 a été marquée notamment par l'émergence de projets en partenariat visant au renforcement de l'orientation "Approches Statistiques en Traitement du Signal et Géophysique". Des relations étroites existent depuis de nombreuses années avec le Centre d'Étude des Phénomènes Aléatoires et Géophysiques (CEPHAG) et le laboratoire de Traitement des Images et Reconnaissance des Formes (TIRF), désormais regroupés dans le Laboratoire des Images et Signaux (LIS) de Grenoble. Récemment, à travers l'équipe SMS, le LMC et le LIS se sont associés dans le projet ELESA / IMAG Statistique Avancée pour le Signal et les Images (SASI), centré sur la séparation de sources et les algorithmes stochastiques. Une proposition conjointe vient aussi d'être retenue dans le cadre du projet européen IST sur la séparation de sources et ses applications (BLISS, <http://www.bliss-project.org/>).

Chercheurs impliqués Anestis Antoniadis (professeur), Alexis Bienvenüe (MdeC, ENSI-MAG), Gersende Fort, (CR CNRS), Marc Joannides (MdeC, Université Pierre Mendès-France), Dinh-Tuan Pham (DR CNRS).

4.6 Département de Traitement du Signal et des Images (TSI), Laboratoire de Traitement et de Communication de l'information (LTCI, URA820), Ecole Nationale Supérieure des Télécommunications

Les recherches menées au sein de l'équipe Traitement Statistique Appliqués aux Communications (TSAC) du département TSI se situent à l'interface entre les statistiques mathématiques et computationnelles, les probabilités appliquées, le traitement du signal, le traitement de l'information et les communications. Les principales orientations de recherche sont les suivantes :

- méthodes de simulation MCMC (applications, analyse de convergence), filtrage particulier, couplage des méthodes de simulation et approximation stochastique,
- modèles de Markov cachés et modèles à régimes markoviens : estimation des paramètres, filtrage, etc.
- statistique des processus (estimation spectrale des processus fractionnaires),
- applications à l'analyse de trafic, aux communications numériques, à la segmentation, et à la restauration de signaux, etc.

La période 1998-2001 a été marquée par le développement de partenariats plus étroits avec la communauté "Statistique Computationnelle" à travers la participation active au réseau européen TMR sur les statistiques spatiales et computationnelles (<http://www.maths.lancs.ac.uk/~breyer/essn/>). L'équipe a assumé un rôle important dans l'animation communautaire autour de ces domaines, à la fois à travers l'animation du groupe de travail "Information" du GdR Information, Signal, Images et Vision (ISIS, <http://www-isis.enst.fr/>), l'organisation d'une Ecole d'été "Méthodes de Monte Carlo pour l'Inférence Statistique" à Luminy, et la coordination d'un numéro spécial de la revue "Signal Processing" sur les applications des méthodes de simulation en traitement du signal. L'activité de l'équipe est toujours importante autour des thèmes liés aux communications numériques, et en particulier, égalisation, séparation de sources. Récemment, l'équipe a démarré une action autour de l'analyse du trafic informatique, en liens avec l'équipe Réseaux à Haut Débit du LTCI.

Chercheurs impliqués Olivier Cappé (CR CNRS), Randal Douc (MdeC, École Polytechnique), Éric Moulines (professeur), Philippe Soulier (MdeC, Université d'Evry Val d'Essonne, en délégation au LTCI).