

# Table des matières

---

<b>Table des matières</b>	<b>1</b>
<b>1 Etat civil</b>	<b>2</b>
<b>2 Synthèse de carrière</b>	<b>2</b>
<b>3 Activités pédagogiques</b>	<b>2</b>
3.1 Présentation de l'activité d'enseignement . . . . .	2
<b>4 Activité scientifique</b>	<b>3</b>
4.1 Présentation synthétique des thématiques de recherche . . . . .	3
4.2 Publications . . . . .	6
4.3 Encadrement doctoral et scientifique . . . . .	6
4.4 Diffusion et rayonnement . . . . .	7
4.5 Responsabilités scientifiques . . . . .	8
4.6 Primes et rayonnement international . . . . .	8
4.7 Responsabilités collectives . . . . .	9
<b>A Liste classée des publications</b>	<b>10</b>

---

# Pierre LATOUCHE

Maître de conférences HDR en Mathématiques Appliquées  
Laboratoire SAMM, EA 4543, Université Paris 1 Panthéon-Sorbonne

## 1. Etat civil

---



Né le 07/09/1984 (33 ans) à Gien  
Marié, 2 enfants  
E-mail : pierre.latouche@univ-paris1.fr  
Web : <http://samm.univ-paris1.fr/Pierre-Latouche>

Laboratoire SAMM, EA 4543  
Université Paris 1 Panthéon-Sorbonne  
90 rue de Tolbiac, 75013 Paris, France  
Tél : +33 (0)1 44 07 88 26 – Fax : +33 (0)1 44 07 89 35

## 2. Synthèse de carrière

---

Depuis septembre 2011, je suis Maître de conférences à l'Université Paris 1 Panthéon-Sorbonne. J'ai obtenu l'habilitation à diriger les recherches en novembre dernier. Auparavant, j'ai effectué ma thèse à l'université d'Evry-val-d'Essonne de 2007 à 2010. Ma soutenance a eu lieu en décembre 2010 et j'ai eu un contrat d'ATER dans cette même université jusqu'à mon recrutement au mois de septembre qui a suivi.

2011-... **Maître de conférences en Mathématiques Appliquées, Université Paris 1 Panthéon-Sorbonne**

2007-2011 **Doctorant en Mathématiques appliquées puis ATER, Université d'Evry-val-d'Essonne**

## 3. Activités pédagogiques

---

### 3.1 Présentation de l'activité d'enseignement

Je décris ci-après mon activité d'enseignement depuis mon début de carrière, ainsi que les responsabilités pédagogiques que j'ai occupées.

#### 1. Enseignement :

2011-... **Maître de conférences UFR Mathématiques et Informatique Université Paris 1 Panthéon-Sorbonne** (~210h ETD/an)

- programmation en C, cours et TD, Licence 1 MASS,
- algèbre, TD, Licence 1 MASS,
- analyse numérique, TD, Licence 3 MASS,
- probabilités et statistiques, TD, Magistère éco,
- base de données, cours et TD, Master 2 Pro TIDE,
- programmation en R, cours et TD, Master 2 Pro TIDE,
- analyse de données, cours et TD, Master 2 Pro TIDE et Master 2 MIAGE,
- statistiques avancées, cours et TD, Master 1 MAEF,
- analyse statistique des réseaux, cours et TD, Master 2 Pro TIDE et Collège des écoles doctorales de Paris 1,
- statistiques computationnelles, cours et TD, Master 2 MO (Paris 1-Paris 7),
- + modèles de graphes aléatoires, cours, Rochester institute, USA (2017),
- + séminaire de modélisation statistique, cours, 2ème année ENSAE (2017).

2007-2011 **Moniteur puis ATER à l'Université d'Evry-val-d'Essonne** (290h ETD)

- probabilités et statistiques, cours et TD, Licence 1, 2, et 3 de biologie,
- algèbre, TD, Licence 1 mathématiques,
- programmation, TD, Licence 1 physique.

## 2. Responsabilités pédagogiques :

- 2014-... **Membre élu du conseil d'UFR Mathématiques et informatique, Université Paris 1**
- 2011-... **Membre du comité de sélection du Master 2 Pro TIDE**
- 2011-... **Membre du comité de perfectionnement du Master 2 Pro TIDE**
- 2011-... **Encadrement de 21 contrats d'apprentissage de Master 2 Pro**  
Apprentis au sein des entreprises Orange, AXA, Banque Postale, Total, BNP, BPCE, Caisse des dépôts.
- 2011-... **Encadrement de 5 stages de Master 2 Pro à l'université Paris 1**

## 3. Formations professionnelles :

- 2016-... **Formation « Analyse statistique des réseaux »**  
1 session/an, CNRS Formation entreprises.
- 2016 **Formation « étude des réseaux »**  
1 session, ateliers statistiques de la SFDS.

## 4. Organisation jury :

- 2015-... **Jury pour l'agrégation externe de mathématiques**  
Interrogation en modélisation A (probabilités et statistiques), B (calcul scientifique), et C (algèbre et calcul formel). J'ai également été secrétaire (responsable) de jury en 2016.

# 4. Activité scientifique

---

## 4.1 Présentation synthétique des thématiques de recherche

Mon activité de recherche s'articule autour de l'apprentissage statistique pour l'analyse des réseaux avec des modèles statiques et dynamiques, ainsi que sur l'étude des graphons et la grande dimension. Je détaille ci-dessous une partie de mes contributions pour ces thèmes.

### **Analyse statistique des réseaux : modèles statiques**

Les réseaux sont de plus en plus utilisés pour caractériser des liens entre des objets d'intérêt. Ainsi, ils apparaissent à l'heure actuelle dans de nombreux champs scientifiques. Ce domaine de recherche, historique, m'a, d'abord, poussé à développer des modèles de graphes aléatoires ainsi que des méthodes d'inférence associées dans un cadre statique, sans composante temporelle. L'objectif est d'extraire automatiquement des informations pertinentes à partir de ce type de données.

- *Inférence dans le modèle à blocs stochastiques* : j'ai, tout d'abord, proposé avec mes co-auteurs des méthodes d'inférence pour le modèle à blocs stochastiques, stochastic block model en anglais, publiées dans deux articles. Le premier article [16] propose notamment un nouveau critère de sélection de modèles. Le second [13] décrit une stratégie d'optimisation avec un faible coût algorithmique. J'ai adapté ce travail au modèle à blocs latents, latent block model, avec des co-auteurs de l'UCD à Dublin lors de ma venue en tant que professeur invité en septembre 2013. Un article [12] a été publié sur le sujet.
- *Recherche de clusters chevauchants* : je me suis également intéressé à la recherche de clusters chevauchants dans les réseaux, où un nœud peut appartenir à plusieurs clusters simultanément. Le premier article publié [17] propose un nouveau modèle ainsi qu'une stratégie d'inférence fréquentiste. L'identifiabilité du modèle est en outre démontrée. J'ai ensuite dérivé un algorithme variationnel dans un cadre Bayésien, pour la sélection de modèles et le clustering, dans un second article [14].
- *Caractérisation des sous-graphes* : l'objectif est ici d'intégrer une partition connue du praticien dans un dérivé du modèle à blocs stochastiques, en vue de chercher des clusters non-triviaux. Ce travail a été motivé par une collaboration avec des historiens médiévistes de l'université Paris 1. Le résultat est publié dans un article [15]. Ma collaboration avec Charles Bouveyron a commencé à cette occasion.
- *Analyse conjointe de réseaux et de textes* : de manière surprenante, les outils d'analyse de réseaux sociaux ne permettent normalement pas d'analyser les textes échangés. Certaines méthodes autorisent l'étude du réseau. D'autres se concentrent sur le corpus. Avec mes co-auteurs, nous avons proposé un nouveau modèle de graphes aléatoires, appelé STBM, ainsi qu'une stratégie d'inférence et un critère de sélection de modèles. La méthodologie permet l'analyse conjointe du réseau et des textes portés par les

connexions. Ce travail, publié dans un article [5], est à la base de la plate-forme web `linkage.fr` dont j'ai implémenté le source et qui s'inscrit dans un projet de valorisation soutenu par la SATT Idflnno. Un brevet associé au procédé a également été déposé aux États-Unis. L'article publié est joint à ce dossier. Un autre article publié [4] vulgarise ce travail.

### **Analyse statistique des réseaux : modèles dynamiques**

Les réseaux sont en réalité des objets dynamiques dans de nombreuses applications. J'ai donc proposé une série de contributions visant à adapter certaines stratégies statiques existantes, en intégrant une dimension temporelle. J'ai travaillé à temps discret et continu comme indiqué ci-dessous.

- *Modèles à temps discret* : dans un premier temps, j'ai proposé avec mes co-auteurs des modèles à temps discret où le réseau dynamique est vu comme une série d'images indexées par le temps. Deux travaux publiés [6, 7] sont basés sur des algorithmes de type glouton. Ces derniers recherchent à la fois des clusters de nœuds ainsi que des clusters temporels. Les deux autres [9, 11] sont basés sur des processus linéaires dynamiques et l'estimation est réalisée à l'aide des équations RTS et du filtre de Kalman. Rawya Zreik, dont j'ai co-encadré la thèse avec Charles Bouveyron, a principalement contribué sur ce thème. Enfin, ma collaboration avec l'UCD Dublin a donné lieu à un nouvel article [1] publié très récemment sur cette question, dont l'objectif premier est la sélection de modèles.
- *Modèles à temps continu* : la thèse de Marco Corneli que j'ai co-encadrée avec Fabrice Rossi a permis notamment le développement d'une méthodologie à temps continu. L'algorithme proposé autorise la segmentation du temps à partir de données d'interactions. Il rend donc possible la construction automatique d'images du réseau où le niveau de connectivité est homogène. Ce travail a été publié dans un article [2] et est joint au dossier.

### **Étude des graphons**

Cette thématique de recherche s'intéresse au modèle de  $W$ -graphe généralisant la plupart des modèles de graphes aléatoires existants. La recherche en statistique pour l'étude des réseaux s'est largement tournée vers ce modèle ces cinq dernières années. L'objectif premier est d'identifier des stratégies d'inférence pour le  $W$ -graphe sur données réelles.

- *Estimation dans le modèle de  $W$ -graphe* : ma collaboration avec Stéphane Robin a débuté avec ce projet. Comme d'autres chercheurs du domaine, nous sommes partis du constat que le modèle à blocs stochastiques était un cas particulier de  $W$ -graphe. L'estimation de la fonction graphon, caractérisant le modèle, étant particulièrement complexe à adresser directement, nous nous sommes donc appuyés sur une stratégie établie pour le modèle à blocs stochastiques. Le cadre Bayésien retenu permet alors d'intégrer le graphon bloc constant inféré sur la largeur et la hauteur des blocs afin d'obtenir une estimation plus pertinente du graphon sur données réelles. Ce travail est publié dans un article [10].
- *Qualité de l'ajustement (approche Bayésienne)* : Sarah Ouadah nous a ensuite rejoints afin de proposer une stratégie inférentielle permettant d'évaluer la qualité d'ajustement d'un modèle de régression logistique sur graphe où les covariables connues sont utilisées pour la prédiction des connexions. Ce type de modèle est couramment retenu par de nombreux praticiens. Une fois les paramètres estimés, il est alors crucial de déterminer si le modèle explique complètement la construction des connexions. Nous avons donc intégré au modèle de régression logistique un terme de type graphon caractérisant l'information de structure non expliquée par les covariables. L'inférence Bayésienne nécessite ici deux niveaux d'approximations variationnelles. Un test est alors mis en place à partir de facteurs de Bayes estimés pour évaluer la pertinence du modèle et des covariables associés. La méthodologie est publiée dans un article [3] et est également disponible au travers d'un package R appelé `gofNetwork` que j'ai implémenté. L'article est joint à ce dossier.
- *Qualité de l'ajustement (test)* : enfin, dans le cadre de cette collaboration, nous avons proposé une procédure de test afin d'évaluer la qualité d'ajustement de modèles de graphes aléatoires. Deux modèles particuliers sont ainsi considérés : le modèle d'Erdős-Rényi ainsi que le modèle de  $W$ -graphe. Des tests sont alors construits à partir d'une statistique basée sur les degrés observés. Dans l'article [21] qui est en phase de relecture nous prouvons la normalité asymptotique de cette statistique, pour ces deux modèles. Ces résultats rendent donc possible l'utilisation de cette statistique pour la mise en place de tests simples. La première preuve est essentiellement basée sur une décomposition de Hoeffding de la statistique et sur l'application du théorème de Linderberg-Lévy. La seconde s'appuie sur un résultat de convergence des probabilités d'apparition de motifs dans le cas du modèle de  $W$ -graphe. L'article est joint à ce dossier.

## La grande dimension

Ce dernier axe de recherche a pour objet la construction de modèles et de techniques d'inférence pour la grande dimension où le nombre de variables considérées est important au regard du nombre d'observations. Je me suis intéressé à ce problème au travers de ma collaboration avec Charles Bouveyron et Pierre-Alexandre Mattéi que Charles et moi avons co-encadré en thèse. La fin de ma propre thèse a coïncidé avec la démocratisation de l'utilisation de procédures de type Lasso et de travaux autour des pénalités  $L_1$  et  $L_2$ . Inspiré par les résultats de David McKay sur la régression ridge notamment, j'ai alors cherché un cadre d'optimisation Bayésien pour la régression permettant de réaliser la sélection des variables en même temps que l'apprentissage du modèle de régression. Le fruit de ce travail a été le développement de spinyreg pour la régression linéaire Bayésienne où une vraisemblance marginale est maximisée sur un chemin de modèles. Ce travail a motivé la thèse de Pierre-Alexandre Mattéi et nous avons depuis adapté ce cadre méthodologique à d'autres problèmes en grande dimension.

- *Régression linéaire parcimonieuse Bayésienne* : nous avons considéré un modèle de régression linéaire où le vecteur des poids de la régression est construit à partir d'un produit terme à terme entre un vecteur binaire, dont le rôle est de discriminer les variables actives des autres, et un vecteur caractérisé par une loi a priori Gaussienne. Une forme analytique pour la log vraisemblance marginale peut alors être obtenue. Nous avons montré que ce critère intégré pouvait être réécrit comme un terme des moindres carrés plus une pénalité de type rasoir d'Ockham, proche de l'elasticnet. L'inférence est réalisée à partir d'un algorithme EM et la sélection de variables à partir d'un chemin de modèles construits par ce dernier. Ce travail est publié dans un article [8] joint à ce dossier.
- *ACP globalement parcimonieuse* : nous nous sommes ensuite intéressés à l'analyse en composantes principales pour la représentation des données dans un espace propre (ACP). Le principe de base était d'étendre nos travaux en régression à ce cadre non supervisé. Une matrice binaire a donc été introduite pour la sélection des variables dans le modèle d'ACP probabiliste. Des lois a priori Gaussiennes ont alors été retenues pour la matrice des loadings. Afin d'obtenir une expression pour la log-vraisemblance marginale, nous avons eu recours à une décomposition du modèle sur la partie active du support des variables et la partie non active. L'expression pour la log vraisemblance marginale s'obtient lorsque la variance du bruit sur la partie active tend vers 0. Ce résultat est décrit dans un article [20], en phase de relecture, sous la forme d'un théorème. Un algorithme y est également présenté faisant intervenir un niveau d'approximation variationnelle. L'intérêt premier de ce travail est ainsi de permettre la sélection des variables d'un jeu de données, dans un cadre non supervisé, c'est à dire sans objectif prédictif défini.
- *Sélection de la dimension en ACP Bayésienne* : le dernier travail issu de cette collaboration à trois offre une méthode principale pour la sélection du nombre d'axes propres en ACP. Le point de départ est le modèle d'ACP probabiliste. Nous avons prouvé à nouveau, au travers d'un théorème, qu'une expression exacte pouvait être obtenue à partir de lois normale-gamma sur la matrice des loadings, dont les hyperparamètres sont interdépendants. Une contrainte de proportionnalité est en effet nécessaire. La preuve s'appuie essentiellement sur les propriétés théoriques de la loi multivariée asymétrique généralisée de Laplace. Ce travail est décrit dans un article [19] en phase de relecture.

## Applications

Mes travaux s'inscrivant en statistique computationnelle ont été essentiellement motivés par des questions soulevées dans le cadre d'applications sur données réelles. Je donne ici la liste des applications pour lesquelles mes développements méthodologiques ont fourni des résultats. J'ai été amené à travailler sur d'autres jeux de données réelles pour des publications. Sont indiqués ci-dessous seulement les jeux de données pour lesquels mes co-auteurs et moi avons réalisé une étude détaillée :

- réseau métabolique d'*Escherichia coli* ;
- réseau de transcription de *Saccharomyces cerevisiae* ;
- étude du cancer du sein / sélection de gènes ;
- réseau d'illustrations et de bandes dessinées ;
- blogosphere politique française ;
- réseau ecclésiastique en Gaule mérovingienne ;
- emails de l'entreprise Enron ;
- co-publications pour la conférence NIPS ;
- votes au congrès Américain ;
- base de données de films notés par des utilisateurs ;
- trois réseaux maritimes ;
- interactions entre les participants de la conférence ACM Hypertext à Turin ;

- flux de vélos entre les stations Santander de Londres ;
- prédiction de la fréquentation du musée d'Orsay en fonction des flux Vélib à Paris.
- analyse des tweets des français liés à la présidentielle 2017.

## 4.2 Publications

Le tableau ci-dessous présente une synthèse de mon activité en terme de publications. La liste classée de mes publications est donnée en Annexe A.

Type de publication	Quantité	Type de publication	Quantité
Preprint	4	Conférences internationales	7
Articles de journaux	17	Conférences nationales	8
Chapitres de livres	4	Logiciels	6
Articles de vulgarisation	5	Brevets	2
Conférences invitées	10	Séminaires	37

## 4.3 Encadrement doctoral et scientifique

### 1. Nombre de thèses soutenues et nombre de thèses en cours :

- 3 thèses soutenues : R. Zreik (2016), P-A Mattéi (2017), Marco Corneli (2017).
- 1 thèse en cours : N. Jouvin (2017-...).

### 2. Liste des thèses soutenues :

#### 2014-2017 **Encadrement de la thèse de doctorat de Marco Corneli**

Sujet : Dynamic stochastic block models, clustering and segmentation in dynamic graphs

Co-encadrant : Fabrice Rossi

Allocation de recherche, ED 386, Université Paris 1 Panthéon-Sorbonne

Publications associées : 1 preprint [18], 3 articles [2, 6, 7]

#### 2014-2017 **Encadrement de la thèse de doctorat de Pierre-Alexandre Mattei**

Sujet : Classification probabiliste parcimonieuse de données massives et de grande dimension

Co-encadrant : Charles Bouveyron

Allocation de recherche, ED 386, Université Paris Descartes

Publications associées : 2 preprints [19, 20], 1 article [8]

#### 2013-2016 **Encadrement de la thèse de doctorat de Rawya Zreik**

Sujet : Analyse statistique des réseaux et applications aux sciences humaines et sociales

Co-encadrant : Charles Bouveyron

Bourse Région IDF, thèse soutenue en novembre 2016, Universités Paris 1 & Paris Descartes

Publications associées : 3 articles [5, 9, 11], 2 chapitres de livre [22, 23]

### 3. Liste des thèses en cours :

#### 2017-... **Encadrement de la thèse de doctorat de Nicolas Jouvin**

Sujet : Apprentissage statistique pour la segmentation d'images guidée par une source textuelle : application en cancérologie

Bourse DIM FSMP, Université Paris 1 Panthéon-Sorbonne

### 4. Devenir des docteurs :

- Marco Corneli est actuellement en postdoctorat au laboratoire J.A. Dieudonné, Université Côte d'Azur.
- Pierre Alexandre Mattéi est actuellement en postdoctorat à l'université de Copenhague.
- Rawya Zreik est actuellement *data scientist* à Ellis-Car, The Data Driven Company

### 5. Encadrement de post-doctorants, ingénieurs et stagiaires :

- 2 post-doctorants : L. Bergé (2016, financement CNRS) ; M. Corneli (2017-..., financement SATT Idflnnov)
- 1 ingénieur d'étude : D. Marié (2017, financement SATT Idflnnov)
- 7 stages d'étudiants de Master 2 recherche : P.A. Mattei (ENS Cachan) ; R. El Hassan (Univ. Paris 6) ; Y. Jernite (Polytechnique) ; C-A. Dadi (ENS Cachan) ; N. Jouvin (ENS Cachan) ; Laetitia Nouedoui (Univ. Paris 1) ; A-C. Haury (ENSAE).

## 4.4 Diffusion et rayonnement

### 1. *Expertise* :

- 2018        **Evaluation IDEX, Université de Grenoble**
- 2011-...    **Expert pour l'agence nationale de la recherche (ANR)**

### 2. *Participation jurys de thèse* :

- 2011-...    **Rapporteur (et membre) du jury de thèse de doctorat**  
- Adrian Todeschini, Université de Bordeaux, novembre 2016, Probabilistic and Bayesian non parametric approaches for recommender systems and networks
- 2011-...    **Rapporteur et membre du jury d'évaluation à mi-parcours de thèse de doctorat**  
- Martin Royer, Université Paris-Sud, novembre 2017, Optimisation convexe et clustering
- 2011-...    **Membre du jury de 5 thèses de doctorat**

### 3. *Diffusion du savoir (vulgarisation), responsabilités au sein des sociétés savantes* :

#### (a) *Logiciels (détails en annexe)* :

- 5 paquets pour le logiciel R
- 1 source C++ pour la plate-forme web `linkage.fr`

#### (b) *Articles de vulgarisation* :

- P. Latouche & C. Bouveyron & D. Marié & G. Fouetillou, *Présidentielle 2017 : une réorganisation politique du web social?*, *Data analytics post*, 2017
- P. Latouche & C. Bouveyron & D. Marié & G. Fouetillou, *Présidentielle 2017 : une réorganisation politique du web social?*, *theconversation*, 2017
- P. Latouche & C. Bouveyron & D. Marié & G. Fouetillou, *Présidentielle 2017 : une réorganisation politique du web social?*, *Panthéon-Sorbonne magazine*, n° 21, 2017
- P. Latouche & C. Bouveyron, *Les échanges de données au peigne fin*, *Journal du CNRS*, n° 287, 2017
- P. Latouche & C. Bouveyron, *Des réseaux, des textes et de la statistique*, *La lettre de l'INSMI, CNRS*, édition de décembre, 2016.

#### (c) *Autres moyens de diffusion du savoir* :

- 2016-...    **Animation du compte Twitter @latouche\_pierre**  
J'assure par ce biais la diffusion des méthodologies que je développe. Ce compte est suivi par 179 abonnés à ce jour.
- 2014-...    **Organisateur séminaire Parisien de statistique (SemStats)**  
J'ai organisé deux sessions spéciales avec S. Gaïffas et M. Hoffmann pour SemStats en 2014 et 2015.
- 2011-...    **Organisateur workshops StatLearn**  
Le workshop Statlearn est aussi un excellent vecteur de diffusion du savoir car cette conférence est largement ouverte aux entreprises des secteurs du numérique. En effet, la part d'industriels parmi les participants à la conférence est de l'ordre de 25% en moyenne sur les dernières années. Cela représente en général une présence d'une trentaine d'entreprises à la conférence.

### 4. *Organisation colloques, conférences, journées d'étude* :

- 2010-...    **Organisateur de 7 éditions du workshop StatLearn**  
Workshop international sur « Challenging problems in Statistical Learning »  
4 sessions thématiques, 12 orateurs invités, ~200 participants / édition, budget ~15k€  
Statlearn est une conférence de la Société Française de Statistique depuis 2013
  - StatLearn'18, 4-6/04 2018, Nice
  - StatLearn'17, 6-7/04 2017, Lyon
  - StatLearn'16, 7-8/04 2016, Vannes
  - StatLearn'15, 2-3/04 2015, Grenoble
  - StatLearn'14, 7-8/04 2014, Paris
  - StatLearn'13, 8-9/04 2013, Bordeaux
  - StatLearn'12, 5-6/04 2012, Lille
- 2015        **Organisateur session « Graphs and machine learning »**  
21st European Symposium on Artificial Neural Networks (ESANN), Bruges, Belgique, 2015

- 2013 **Organisateur de la session « Statistical methodology for network data »**  
6th International Conference ERCIM, London, UK, décembre 2013
- 2008 **Co-organisateur de la session « Inférence et structure cachée dans les réseaux biologiques » aux journées MAS**  
Société des Mathématiques Appliquées et Industrielles (SMAI), Rennes, 2008
5. *Evaluation d'articles :*
- 2011-... **Evaluation d'articles pour 14 journaux scientifique internationaux**  
dont Journal of the American Statistical Association ; Proceedings of the National Academy of Sciences of America ; Journal of the Royal Statistical Society B ; Journal of the Royal Statistical Society C ; Statistics and Computing ; Annals of Applied Statistics ; Biometrics ; Bioinformatics ; Computational Statistics and Data Analysis ; Journal de la société française de statistique
6. *Participation à un réseau de recherche :*
- 2016-... **Membre de la coopération européenne (action) « COSTNET : Statistics of Network Data »**  
Groupe de travail avec financement européen sur l'analyse des réseaux organisés autour de Steffen L. Lauritzen (Univ. Oxford et Copenhague)
- 2013-... **Membre et porteur d'une action de l'ISC-PIF**  
Institut des Sciences Complexes de Paris - Ile de France
- 2011-... **Membre du « Working Group on Model-Based Clustering »**  
Groupe de travail fondé et animé par Adrian Raftery (University of Washington, USA), une réunion plénière chaque année

#### 4.5 Responsabilités scientifiques

1. *Contrats de recherche évalués suite à appel à projet :*
- 2011-... **Membre de 6 projets de recherche**  
ANR NeMo, PEPS (CNRS), 4 bilans qualité recherche (Paris 1) – Budget : 180 k€
- 2013-2017 **Porteur du projet ISC-PIF « Analyse statistique de réseaux en SHS »**  
Institut des Sciences Complexes Ile-de-France – Budget : 105 k€
- 2012-2013 **Porteur d'une action scientifique**  
Université Paris 1 Panthéon-Sorbonne - Budget : 20k€
2. *Contrats de recherche de gré à gré :*
- 2017-... **Porteur du projet de valorisation SATT IDFINNOV « TOPIX : textes et systèmes de recommandation »**  
Société accélératrice de transferts de technologie Idflnnov – Budget : 190 k€
- 2016-2018 **Porteur du projet de valorisation SATT IDFINNOV « STBM : clustering de réseaux et de textes »**  
Société accélératrice de transferts de technologie Idflnnov – Budget : 50 k€

#### 4.6 Primes et rayonnement international

1. *Primes :*
- 2014-... **Titulaire de la Prime d'Encadrement Doctoral et de Recherche (PEDR)**  
Université Paris 1 Panthéon-Sorbonne
2. *Rayonnement :*
- 2018 **Membre du conseil scientifique SFDS, organisation JDS**  
Nancy, 2019
- 2011-... **Membre du comité de programme de 4 conférences internationales**  
Conférences AISTATS, NIPS, compStat, ESANN
- 2011-... **Organisateur du workshop international StatLearn**  
Invitation de 4 orateurs étrangers à chaque édition
- 2011-... **Exposés invités dans 10 conférences internationales**



3. Participation jurys de thèse :

2018        **Membre du jury de thèse de James Ng**  
UCD Dublin, avril 2018, Statistical models for complex data

4. Encadrement :

2015        **Accueil pour séjour recherche**  
R. Rastelli en thèse avec N. Friel, UCD Dublin. Co-encadrement

5. Invitations :

2016        **Invitation pour séjour recherche au Newton Institute, Cambridge, UK**  
Session « Bayesian methods for networks »

2016        **Invitation pour séjour recherche au CIRM, Marseille**  
Session « Bayesian statistics and algorithms »

2013        **Professeur invité à l'UCD Dublin**  
J'ai été invité en tant que professeur en septembre 2013 à l'UCD Dublin, République d'Irlande

#### 4.7 Responsabilités collectives

2011-...    **Membre de 3 comités de sélection pour des postes MCF**  
- AgroParistech, Paris 1, CNAM

2011-...    **Membre élu du comité consultatif scientifique de l'Université Paris 1 Panthéon-Sorbonne**

## A. Liste classée des publications

Type de publication	Quantité	Type de publication	Quantité
Preprints	4	Conférences internationales	7
Articles de journaux	17	Conférences nationales	8
Chapitres de livres	4	Logiciels	6
Articles de vulgarisation	5	Brevets	2
Conférences invitées	10	Séminaires	37

### [revues avec comité de lecture]

- [1] R. Rastelli, P. Latouche, N. Friel, *Choosing the number of groups in a latent stochastic block model for dynamic networks*, 2018, in press, Network Science.
- [2] M. Corneli, P. Latouche, F. Rossi, *Multiple change points detection and clustering in dynamic networks*, 2017, in press, Statistics and Computing.
- [3] P. Latouche, S. Robin, S. Ouadah, *Goodness of fit of logistic regression models for random graphs*, 2017, in press, Journal of Computational and Graphical Statistics.
- [4] P. Latouche, C. Bouveyron, D. Marié, G. Fouetillou, *Présidentielle 2017 : l'analyse des tweets renseigne sur les recompositions politiques*, 2017, Vol. 5 (3), Statistique et Société.
- [5] C. Bouveyron, P. Latouche, R. Zreik, *The stochastic topic block model for the clustering of vertices in networks with textual edges*, 2016, Vol. 28 (1), 11-31, Statistics and Computing.
- [6] M. Corneli, P. Latouche, F. Rossi, *Block modelling in dynamic networks with non-homogeneous Poisson processes and exact ICL*, 2016, Vol. 6 (1), 1-14, Social Network Analysis and Mining.
- [7] M. Corneli, P. Latouche, F. Rossi, *Exact ICL maximization in a non-stationary temporal extension of the stochastic block model for dynamic networks*, 2016, Vol. 192, 81-91, Neurocomputing.
- [8] P. Latouche, P-A Mattei, C. Bouveyron, J. Chiquet, *Combining a relaxed EM algorithm with Occam's razor for Bayesian variable selection in high-dimension regression*, 2016, Vol. 146, 177-190, Journal of Multivariate Analysis.
- [9] R. Zreik, P. Latouche, C. Bouveyron, *The dynamic random subgraph model for the clustering of evolving networks*, 2016, Vol. 32 (2), 501-533, Computational Statistics.
- [10] P. Latouche, S. Robin, *Variational Bayes model averaging for graphon functions and motif frequencies inference in W-graph models*, 2016, Vol. 26 (6), 1173-1185, Statistical and Computing.
- [11] R. Zreik, P. Latouche, C. Bouveyron, *Classification automatique de réseaux dynamiques avec sous-graphes : étude du scandale Enron*, 2015, Vol. 156 (3), 166-191, Journal de la Société Française de Statistique.
- [12] J. Wyse, N. Friel, P. Latouche, *Inferring structure in bipartite networks using the latent block model and exact ICL*, 2015, Vol. 5 (1), 45-69, Network Science.
- [13] E. Côme, P. Latouche, *Model selection and clustering in stochastic block models based on the exact integrated complete data likelihood*, 2015, Vol. 15 (6), 654-589, Statistical Modelling.
- [14] P. Latouche, E. Birmelé, C. Ambroise, *Model selection in overlapping stochastic block models*, 2014, Vol. 8 (1), 762-794, Electronic Journal of Statistics.
- [15] Y. Jernite, P. Latouche, C. Bouveyron, P. Rivera, L. Jegou, S. Lamassé, *The random subgraph model for the analysis of an ecclesiastical network in Merovingian Gaul*, 2013, Vol. 8 (1), 377-405, Annals of Applied Statistics.
- [16] P. Latouche, E. Birmelé, C. Ambroise, *Variational Bayesian inference and complexity control for stochastic block models*, 2012, Vol. 12 (1), 93-115, Statistical Modelling.
- [17] P. Latouche, E. Birmelé, C. Ambroise, *Overlapping stochastic block models with application to the French political blogosphere*, 2011, Vol. 5 (1), 309-336, Annals of Applied Statistics.

### [Preprints]

- [18] M. Corneli, C. Bouveyron, P. Latouche, F. Rossi, *The dynamic stochastic topic block model for time evolving networks with textual edges*, 2017, révision, Journal of the American Statistical Association.
- [19] P-M Mattei, C. Bouveyron, P. Latouche, *Exact Dimensionality Selection for Bayesian PCA*, 2017, révision, Journal of Statistical Planning and Inference.
- [20] P-M Mattei, C. Bouveyron, P. Latouche, *Bayesian variable selection for globally sparse probabilistic PCA*, 2017, révision, Electronic Journal of Statistics.

- [21] S. Ouadah, S. Robin, P. Latouche, *Degree-based goodness-of-fit tests for heterogeneous random graph models : independent and exchangeable cases*, 2017, révision, Bernoulli.

## [Chapitres de livre]

- [22] R. Zreik, C. Ducruet, C. Bouveyron, P. Latouche, *Advances in Shipping Data Analysis and Modeling Tracking and Mapping Maritime Flows in the Age of Big Data -Cluster dynamics in the collapsing Soviet shipping network-*, 2017, Routledge.
- [23] C. Bouveyron, P. Latouche, R. Zreik, C. Ducruet, *Maritime Networks : Spatial Structures and Time Dynamics -Cluster identification in maritime flows with stochastic methods-*, 2015, 210-228, Routledge.
- [24] P. Latouche, E. Birmelé, C. Ambroise, *Handbook of Mixed Membership Models and Their Applications -Overlapping clustering methods for networks-*, 2014, 547-567, Chapman and Hall/CRC.
- [25] P. Latouche, E. Birmelé, C. Ambroise, *Advances in data handling and business intelligence -Bayesian methods for graph clustering-*, 2009, 229-239, Springer.

## [Vulgarisation]

- [26] P. Latouche, C. Bouveyron, D. Marié, G. Fouetillou, *Présidentielle 2017 : une réorganisation politique du web social ?*, 2017, Data analytics post.
- [27] P. Latouche, C. Bouveyron, D. Marié, G. Fouetillou, *Présidentielle 2017 : une réorganisation politique du web social ?*, 2017, theconversation.
- [28] P. Latouche, C. Bouveyron, D. Marié, G. Fouetillou, *Présidentielle 2017 : une réorganisation politique du web social ?*, 2017, Panthéon-Sorbonne magazine.
- [29] P. Latouche, C. Bouveyron, *Les échanges de données au peigne fin*, 2017, CNRS, le journal.
- [30] P. Latouche, C. Bouveyron, *Des réseaux, des textes, et de la statistique !*, 2016, Lettre de l'INSMI.

## [Brevets]

- [31] P. Latouche & C. Bouveyron, *Co-clustering and recommendation system with text data*, 2018, pending, EU patent.
- [32] P. Latouche & C. Bouveyron, *A system for clustering relational data with texts*, 2017, pending, US patent.

## [Logiciels]

- [33] Linkage.fr, *Plate-forme Web (source en C++, Javascript)*, 2017, Analyse conjointe de réseaux et de textes.
- [34] Gof network, *Package R (source en C++)*, 2016, Analyse du fit de modèles de régression logistique pour les réseaux.
- [35] SpinyReg, *Package R*, 2014, Régression sparse Bayésienne.
- [36] OSBM, *Package R (source en C)*, 2014, Méthodes d'inférence pour le modèle à blocs stochastiques.
- [37] Rambo, *Package R*, 2013-présent, Méthodes d'inférences pour le modèle à sous-graphes aléatoires.
- [38] Mixer, *Package R (source en C, Fortran)*, 2011-présent, Méthodes d'inférences pour le modèle à blocs stochastiques.

## [Invitations dans des conférences internationales]

- [39] P. Latouche, *The dynamic stochastic block model*, 2017, Working group on model based clustering, Pérouse, Italie.
- [40] P. Latouche, *Goodness of fit of logistic models for random graphs*, 2016, CIRM, Bayesian Statistics and Algorithms, Marseille.
- [41] P. Latouche, *Goodness of fit of logistic models for random graphs*, 2016, Newton Institute, Cambridge, UK.
- [42] P. Latouche, *Statistical learning on graphs : a short tutorial*, 2016, Conference on challenging issues in statistical learning, Vannes.
- [43] P. Latouche, *Graphs in machine learning : an introduction*, 2015, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, Bruges, Belgique.
- [44] P. Latouche, *Estimation of the posterior graphon function using stochastic block models*, 2013, International Conference of the ERCIM WG, London, UK.
- [45] P. Latouche, *Random graph models for social network analysis*, 2013, International biometric society channel network, Paris.

- [46] P. Latouche, *A model selection criterion for overlapping stochastic block models with L2 constraints*, 2012, Working group on model based clustering, Guelph, Canada.
- [47] P. Latouche, *Complexity control in overlapping stochastic block models*, 2012, Conference on challenging issues in statistical learning, Lille.
- [48] P. Latouche, *Structure analysis in networks*, 2008, Journées MAS, Rennes.

### [Conférences internationales avec comité de lecture]

- [49] P-A. Mattei, C.B. Bouveyron, P. Latouche, *Bayesian Variable Selection for Globally Sparse Probabilistic PCA*, 2016, AISTATS, Cadiz, Espagne.
- [50] Exact ICL maximization in a non-stationary time extension of the latent block model for dynamic networks, *M. Cornéli, P. Latouche, C. Bouveyron*, 2016, ESANN, Bruges, Belgique.
- [51] P. Latouche, S. Robin, *Estimation de la loi a posteriori de la fonction graphon d'un W-graphe*, 2014, AIGM.
- [52] Bayesian non parametric inference of discrete valued networks, *L. Nouedoui, P. Latouche*, 2013, ESANN, Bruges, Belgique.
- [53] Activity Date Estimation in Timestamped Interaction Networks, *F. Rossi, P. Latouche*, 2013, ESANN, Bruges, Belgique.
- [54] P. Latouche, E. Birmelé, C. Ambroise, *A latent logistic model to uncover overlapping clusters in networks*, 2009, Association Française pour l'Intelligence Artificielle, atelier AGS, Hammamet, Tunisie.
- [55] P. Latouche, E. Birmelé, C. Ambroise, *Bayesian methods for graph clustering*, 2008, Société Allemande de Classification (GFKL), Hambourg, Allemagne.

### [Conférences nationales avec comité de lecture]

- [56] P-A. Mattei, C. Bouveyron, P. Latouche, *Globally sparse probabilistic PCA*, 2016, Journées de Statistique, Montpellier.
- [57] P. Latouche, P-A. Mattei, C. Bouveyron, *Une relaxation continue du rasoir d'Occam pour la régression en grande dimension*, 2015, Journées de Statistique, Lille.
- [58] R. Zreik, P. Latouche, C. Bouveyron, *Un modèle de sous-graphes aléatoires pour l'analyse de la dynamique d'un réseau historique*, 2014, Journées de Statistique, Rennes.
- [59] P. Latouche, S. Robin, *Estimation de la loi a posteriori de la fonction graphon d'un W-graphe. Application au réseau de la blogosphere politique française*, 2013, Journées de Statistique, Toulouse.
- [60] C. Bouveyron, L. Jegou, Y. Jernite, S. Lamassé, P. Latouche, P. Rivera, *Un modèle de graphes aléatoires pour l'analyse d'un réseau ecclésiastique dans la Gaule mérovingienne*, 2013, Journées de Statistique, Toulouse.
- [61] P. Latouche, E. Birmelé, C. Ambroise, *Modèles de graphes aléatoires à classes chevauchantes pour l'analyse des réseaux*, 2010, Journées de Statistique, Marseille.
- [62] P. Latouche, E. Birmelé, C. Ambroise, *A latent logistic model to uncover overlapping clusters in networks*, 2009, Société Française de Classification, Grenoble.
- [63] P. Latouche, E. Birmelé, C. Ambroise, *Uncovering overlapping clusters in biological networks*, 2009, Journées Ouvertes en Biologie, Informatique et Mathématiques, Nantes.

### [Conférences sans acte]

- [64] P. Latouche, *Stochastic blockmodelling applied to global maritime flows*, 2012, ERC workshop.
- [65] P. Latouche, *Analyse de réseaux à l'aide de modèles de graphes aléatoires*, 2011, MASHS, Marseille.
- [66] P. Latouche, *Overlapping stochastic block models*, 2009, European Meeting of Statisticians, Toulouse.

### [Séminaires]

- [67] P. Latouche, *Statistical learning on graphs and networks*, 2018, Statlearn spring school, Université Côte d'Azur, Nice.
- [68] P. Latouche, *The dynamic stochastic topic block model*, 2018, Séminaire équipe SIERRA, INRIA, Paris.
- [69] P. Latouche, *The stochastic topic block model*, 2018, Statistical advances for real data problems, mini-conférence, Jussieu.

- [70] P. Latouche, *The stochastic topic block models*, 2017, Journée AppliBugs, AgroParisTech.
- [71] P. Latouche, *Introduction à l'algorithme variationnel EM*, 2017, Séminaire AgroCampus, Rennes.
- [72] P. Latouche, *Multiple change points detection and clustering in dynamic networks*, 2017, Séminaire unité MALAGE, INRA, Jouy-en-Josas.
- [73] P. Latouche, *The dynamic stochastic topic block models*, 2017, Séminaire de probabilité et statistique, Université Paris Est Marne.
- [74] P. Latouche, *The dynamic stochastic topic block models*, 2017, Séminaire de probabilité et statistique, Montpellier.
- [75] P. Latouche, *The dynamic stochastic topic block models*, 2017, Journée de rentrée du MAP5, Paris.
- [76] P. Latouche, *The dynamic stochastic topic block models*, 2017, Café des statistiques, Polytechnique, Paris.
- [77] P. Latouche, *Linkage analyse conjointe de réseaux et de textes*, 2017, INED, Les rencontres de statistique appliquée, Paris.
- [78] P. Latouche, *The dynamic stochastic topic block models*, 2017, Séminaire de probabilités et statistique, Grenoble.
- [79] P. Latouche, *The dynamic stochastic topic block models*, 2017, Séminaire, Institut de Mathématiques de Bourgogne, Dijon.
- [80] P. Latouche, *Goodness of fit of logistic models for random graphs*, 2016, Séminaire, AgroParisTech.
- [81] P. Latouche, *Goodness of fit of logistic models for random graphs*, 2016, Séminaire Parisien de Statistique, Institut Henri Poincaré.
- [82] P. Latouche, *The stochastic topic block model*, 2016, Groupe de travail sur les réseaux, AgroParisTech.
- [83] P. Latouche, *Goodness of fit of logistic models for random graphs*, 2015, Séminaire, département de mathématiques, Orsay.
- [84] P. Latouche, *Goodness of fit of logistic models for random graphs*, 2015, Bayes in Paris, ENSAE.
- [85] P. Latouche, *Estimation of the graphon function of a  $W$ -graph model*, 2015, Séminaire de modélisation statistique, laboratoire IRMA, Université de Strasbourg.
- [86] P. Latouche, *Estimation of the graphon function of a  $W$ -graph model*, 2015, Groupe de travail statistique, Jussieu.
- [87] P. Latouche, *Estimation of the graphon function of a  $W$ -graph model*, 2015, Groupe de travail statistique, Télécom ParisTech.
- [88] P. Latouche, *Estimation of the graphon function of a  $W$ -graph model*, 2014, Séminaire, Université de Rouen.
- [89] P. Latouche, *Random graph models for the clustering of nodes in networks and visualization*, 2014, Netbio, inférence de réseaux biologiques.
- [90] P. Latouche, *Estimation of the posterior distribution of the  $W$ -graph graphon function*, 2013, University college Dublin seminar, Dublin, République d'Irlande.
- [91] P. Latouche, *Estimation of the posterior distribution of the  $W$ -graph graphon function*, 2013, Bayes in Paris, ENSAE.
- [92] P. Latouche, *Estimation of the posterior distribution of the  $W$ -graph graphon function*, 2013, Séminaire MAP5, Université Paris Descartes.
- [93] P. Latouche, *Clustering de nœuds dans les réseaux à l'aide de mixer*, 2012, Semin-R, Muséum d'histoire naturelle, Paris.
- [94] P. Latouche, *Complexity control in overlapping stochastic block models*, 2012, Séminaire AgroCampus, Rennes.
- [95] P. Latouche, *Variational methods for overlapping and non-overlapping stochastic block models*, 2012, MSTGA.
- [96] P. Latouche, *Modèles de graphes aléatoires pour l'analyse des réseaux*, 2012, Orange Lab's, Lannion.
- [97] P. Latouche, *Modèles de graphes aléatoires pour l'analyse des réseaux*, 2012, Séminaire LIP6, Paris.
- [98] P. Latouche, *Model selection in overlapping stochastic block models*, 2011, Séminaire Statistique et Probabilités, Institut de Mathématiques, Toulouse.
- [99] P. Latouche, *Model selection in overlapping stochastic block models*, 2011, Statistics for Systems Biology, groupe réseau, AgroParisTech.
- [100] P. Latouche, *Model selection in overlapping stochastic block models*, 2011, Séminaire laboratoire SAMM, Université Paris 1.
- [101] P. Latouche, *Overlapping stochastic block models*, 2010, Séminaire Lyon 1.

- [102] P. Latouche, *Modèles de graphes aléatoires à classes chevauchantes pour l'analyse des réseaux*, 2010, Séminaire MAP5, Paris.
- [103] P. Latouche, *Modèles de graphes aléatoires pour l'analyse des réseaux*, 2009, Séminaire LIFO, Orléans.