

INFÉRENCE GÉOMÉTRIQUE

MODALITÉS D'EXAMEN

Quand : Vendredi 5 avril

Où : Salle 15.16 - 201

Format : Vous présenterez individuellement et à l'oral un article de votre choix.

Durée : 20 minutes autonomes de présentation, 10 minutes de questions

Support : Libre (tableau, slides, les deux).

Si vous utilisez des slides, me les envoyer par email au plus tard la veille de la présentation.

INSTRUCTIONS POUR LE CHOIX DE L'ARTICLE

- Choisir le papier que vous souhaitez présenter pour l'examen.
- Indiquer vos nom et prénom et cocher la case « ☒ » correspondant à la colonne de l'article choisi sur le site suivant :

<https://date.chapril.org/iphI8BoIxvb6usZP>

- **Attention :**
 - Ne votez que pour un seul article. Un vote pour strictement plus qu'un article sera considéré comme invalide, puis supprimé.
 - L'attribution se fait selon la règle « premier arrivé, premier servi ».
- Si vous avez en tête un autre papier que vous souhaiteriez lire et présenter pour l'examen, c'est tout à fait possible, sous réserve de validation de ma part. Envoyez-moi un email pour en discuter.
- Me contacter pour toute question autour du papier que vous lisez.

LISTE ET DESCRIPTION DES ARTICLES

1. **Minimax Rates for Homology Inference**

S. Balakrishnan, A. Rinaldo, A. Singh, D. R. Sheehy, L. Wasserman
AISTATS : AI and Statistics 2012

Lien : <http://donsheehy.net/research/balakrishnan12minimax.pdf>

Thème : inférence de l'homologie, bornes minimax sous différents modèles de bruit.

Travail demandé : Quatre types de modèles sont donnés. "Noiseless" est le même modèle que celui considéré dans le cours pour l'inférence du type d'homotopie de compacts à reach positif. Il s'agit ici de choisir un (au moins) des autres modèles et de détailler et présenter les résultats du papier sous ce (ou ces) modèle(s).

2. **Rates of Convergence for Robust Geometric Inference**

F. Chazal, P. Massart, B. Michel
Electronic Journal of Statistics, vol. 10, Number 2 (2016)

Lien : <https://projecteuclid.org/euclid.ejs/1472125729>

Thème : étude statistique de la DTM pour des mesures empiriques.

Travail demandé : lecture et synthèse du papier. Choisir un point à détailler/développer.

3. **A Statistical Test of Isomorphism Between Metric-Measure Spaces Using the Distance-to-a-Measure Signature**

C. Brécheteau
Electronic Journal of Statistics, vol. 13, Number 1 (2019)

Lien : <https://projecteuclid.org/euclid.ejs/1553565705>

Thème : DTM dans les espaces métriques et test statistiques pour la comparaison d'espaces métriques mesurés.

Travail demandé : lecture et synthèse du papier. Choisir un point à détailler/développer.

4. **Stochastic Convergence of Persistence Landscapes and Silhouettes**

F. Chazal, B. T. Fasy, F. Lecci, A. Rinaldo, L. Wasserman
Journal of Computational Geometry, Vol 6, No 2 (2015)

Lien : article

Thème : *persistence landscapes* et processus empirique ; intervalles de confiance.

Travail demandé : lecture et synthèse du papier. Choisir un point à détailler/développer.

5. **Confidence Sets for Persistence Diagrams**

B. T. Fasy, F. Lecci, A. Rinaldo, L. Wasserman, S. Balakrishnan, A. Singh
Annals of Statistics, vol. 43, No. 6, 2014

Lien : article

Thème : construction d'intervalles de confiance pour l'homologie persistante.

Travail demandé : quatre approches différentes sont proposées pour calculer des intervalles de confiance autour de diagrammes de persistance. Il s'agira de donner une présentation synthétique du modèle statistique adopté et des 3 premières méthodes (la 4ème relevant d'un contexte totalement différent), puis de détailler les résultats obtenus pour une de ces méthodes.

6. Topological consistency via kernel estimation

O. Bobrowski, S. Mukherjee, J. E. Taylor

Bernoulli, 23(1), 2017

Lien : <https://arxiv.org/pdf/1407.5272v4.pdf>

Thème : inférence de l'homologie (et de l'homologie persistance) des niveaux de densité et de fonctions de régression. Approches par des méthodes à noyau.

Travail demandé : lecture du papier, brève synthèse et choix d'un point à détailler.

7. Convergence rates for persistence diagram estimation in Topological Data Analysis

F. Chazal, M. Glisse, C. Labruere, B. Michel

Journal of Machine Learning Research (JMLR), Vol. 16, p. 3603-3635, Dec. 2015

Lien : <http://www.jmlr.org/papers/volume16/chazal15a/chazal15a.pdf>

Thème : propriétés de convergence des diagrammes de persistance construits sur des nuages de points aléatoires.

Travail demandé : le résultat principal du papier aura été vu en cours. Il s'agit donc de se focaliser sur les parties non traitées en cours, i.e. la Section 4 et, surtout, le Theorem 5 et sa preuve qui permet d'obtenir une borne inférieure optimale en dimension 1.

8. Estimating the Reach of a Manifold

E. Aamari, J. Kim, F. Chazal, B. Michel, A. Rinaldo, L. Wasserman

Electronic Journal of Statistics, vol. 13, Number 1 (2019)

Lien : [Estimating the Reach of a Manifold.pdf](#)

Thème : géométrie et vitesse minimax d'estimation du reach d'une sous-variété fermée de classe $\mathcal{C}^3$.

Travail demandé : lecture et synthèse du papier. Choisir un point à détailler/développer. Article théorique relativement technique. Des bases en géométrie différentielle peuvent aider à la compréhension des preuves.

9. Minimax Rates for Estimating the Dimension of a Manifold

J Kim, A Rinaldo, L Wasserman

Journal of Computational Geometry, Volume 10, Number 1 (2019)

Lien : [article](#)

Thème : Vitesse minimax d'estimation de la dimension intrinsèque d'une sous-variété fermée de classe $\mathcal{C}^2$.

Travail demandé : lecture et synthèse du papier. Choisir un point à détailler/développer. La présentation de ce papier est très technique, bien que les idées sous-jacentes soient relativement simples. L'exposé devra présenter les idées générales de façon élémentaires.

10. Manifold Estimation and Singular Deconvolution under Hausdorff Loss

C. R. Genovese, M. Perone-Pacífico, I. Verdinelli, L. Wasserman

Annals of Statistics, Volume 40, Number 2 (2012), 941-963

Lien : projecteuclid.org

Thème : vitesses minimax d'estimation de support pour des sous-variétés de classe $\mathcal{C}^2$.

Travail demandé : lecture et synthèse du papier pour au moins deux des trois modèles considérés. Choisir un point à détailler/développer. Détailler la mesurabilité de l'estimateur par minimum de contraste proposé Section 4 peut faire partie de l'exposé et remplacer un modèle.

11. Stability and Minimax Optimality of Tangential Delaunay Complexes for Manifold Reconstruction

E. Aamari, C. Levrard

Discrete & Computational Geometry (2018) 59 : 923

Lien : Tangential Delaunay Complex.pdf

Thème : reconstruction effective de sous-variétés de classe $\mathcal{C}^2$ par complexes simpliciaux

Travail demandé : lecture et synthèse du papier. Choisir un point à détailler/développer.

12. Non-Asymptotic Rates for Manifold, Tangent Space and Curvature Estimation

E. Aamari, C. Levrard

The Annals of Statistics 2019, Vol. 47, No. 1, 177–204

Liens : Papier, Supplementary Material

Thème : Estimation de variété, d'espaces tangents et de courbure pour des sous-variétés fermées de classe $\mathcal{C}^k$, $k \geq 2$.

Travail demandé : lecture et synthèse du papier. Choisir un point à détailler/développer.

13. Adaptive Hausdorff estimation of density level sets

A. Singh, C. Scott, R. Nowak

Annals of Statistics, Volume 37, Number 5B (2009), 2760-2782

Lien : <https://www.jstor.org/stable/pdf/30243727.pdf>

Thème : estimation des sur-niveaux $\{f \geq \gamma\}$ d'une densité, qui peut être vu comme une généralisation de l'estimation de support $\{f > 0\}$.

Travail demandé : lecture et synthèse du papier. Choisir un point à détailler/développer.

14. Statistical Analysis and Parameter Selection for Mapper

M. Carrière, B. Michel, S. Oudot

Journal of Machine Learning Research 19 (2018) 1-39

Lien : <http://jmlr.org/papers/volume19/17-291/17-291.pdf>

Thème : étude statistique de la convergence de Mapper vers le graphe de Reeb, sélection de paramètres d'échelle, régions de confiance.

Travail demandé : présentation de la méthode Mapper, lecture et synthèse du papier. Choisir un point à détailler/développer.