

Ecole Doctorale

*Mathématiques et Sciences et Technologies de l'Information et de la Communication***HABILITATION A DIRIGER DES RECHERCHES****Avis de soutenance****Monsieur VERMET FRANCK**

présentera ses travaux en vue de l'habilitation à diriger des recherches, sur le sujet suivant :

**" Etude probabiliste de modèles neuronaux de mémoire associative et d'algorithmes utilisés en physique statistique et data science "****Le mercredi 29 mai 2019 à 14h**

à l'UFR Sciences et Techniques, amphithéâtre "F", 6 avenue Victor Le Gorgeu, Brest.

**Le jury sera ainsi composé :**

- **M. BOIVIN DANIEL, Professeur des universités**  
Univ. de Bretagne Occidentale - BREST
- **M. FRANKE BRICE, Professeur des universités**  
Univ. de Bretagne Occidentale - BREST
- **M. LE NY ARNAUD, Professeur des universités**  
Université Paris-Est Créteil - CRETEIL
- **MME LOUKIANOVA DASHA, Maître de conférences**  
Université d'Evry val d'Essone - EVRY
- **M. PETRITIS DIMITRI, Professeur des universités**  
Université Rennes 1 - RENNES
- **M. VAN ENTER AERNOU, Professeur**  
Université de Groningue - 9747 AG GRONINGEN - PAYS-BAS

A BREST, le 13 mai 2019

Le Président de l'Université de  
Bretagne Occidentale,**M. GALLOU**

# Etude probabiliste de modèles neuronaux de mémoire associative et d'algorithmes utilisés en physique statistique et data science

Franck VERMET

Laboratoire de Mathématiques de Bretagne Atlantique, Brest

## Résumé :

Je présente dans mon mémoire d'habilitation les résultats publiés depuis 2005. Mon intérêt pour le modèle de mémoire associative de Hopfield, qui a fait l'objet de ma thèse de doctorat, a été le point de départ de la majorité des articles écrits avec mes coauteurs. J'ai regroupé mes travaux en quatre parties, selon les thèmes de recherche abordés.

Dans la première partie, je présente les résultats sur des extensions du modèle de Hopfield. Les travaux portent sur la capacité de stockage, qui est influencée par les caractéristiques du modèle : loi et propriétés de dépendance des messages mémorisés, structure du graphe modélisant le réseau de neurones. Nous nous sommes également beaucoup intéressés avec mes coauteurs à la capacité de stockage d'autres modèles de mémoire associative, en particulier les modèles parcimonieux d'Amari, de Willshaw et les modèles biologiquement plus plausibles de Gripon et Berrou, intégrant la notion de cliques neurales. Nous avons proposé des variantes pour ces modèles et analysé leur dynamique.

La seconde partie est consacrée au système de télécommunication CDMA, pour lequel nous étudions la capacité (i.e. le nombre maximal d'utilisateurs possible) en fonction des algorithmes de décodage utilisés. Nous nous sommes en particulier intéressés aux algorithmes de décodage par filtre adapté et aux méthodes de suppression d'interférences HD-PIC et SD-PIC. Nous avons mis en évidence une analogie entre ce modèle et le modèle de Hopfield, ainsi qu'avec une variante du modèle de Hopfield proposée par Amari et Yanai. Grâce à un principe de déviations modérées pour des sommes de variables faiblement dépendantes que nous avons établi, nous avons démontré des résultats annoncés par Amari et Yanai pour leur modèle de mémoire associative et nous avons pu analyser le comportement du modèle CDMA dans le cas de l'application d'une étape de l'algorithme de suppression d'interférences SD-PIC.

Dans la troisième partie, nous nous intéressons aux algorithmes stochastiques de type Monte Carlo par chaîne de Markov. L'algorithme de Metropolis-Hastings est utilisé notamment en physique mathématique pour générer des échantillons de lois de probabilité complexes, mais peut avoir un temps de convergence exponentiel. Des variantes ont été proposées pour accélérer la convergence, dont les méthodes de *swapping* et *simulated tempering*, et nous nous sommes intéressés à ces algorithmes dans le cas du modèle de Hopfield. Nous avons pu traiter le cas où le modèle stocke deux images, généralisant ainsi le cas connu du modèle de Curie Weiss à un modèle désordonné en montrant un temps de convergence polynomial pour l'algorithme du *swapping*. Nous avons également étudié ces algorithmes dans le cas du modèle de Blume-Emery-Griffiths et mis en évidence sur cet exemple l'influence de la nature des transitions de phase du modèle sur la vitesse de convergence des algorithmes.

Enfin, la dernière partie de cette synthèse concerne des travaux récents en lien avec les réseaux de neurones à apprentissage profond. Nous avons proposé avec Matthias Löwe et Vincent Gripon une méthode de classification dans le cadre de technique de transfert d'apprentissage, qui consiste à réaliser l'apprentissage d'un réseau de neurones sur une base de

données conséquente, sur un ordinateur puissant et à utiliser le modèle obtenu pour traiter d'autres jeux de données, éventuellement sur des ordinateurs moins puissants. Les premières couches des réseaux de neurones de type *Deep Learning*, utilisés en traitement d'images notamment, sont le plus souvent des couches convolutives, qui permettent d'extraire des caractéristiques des données en entrée. Ces informations sont recodées en un vecteur de grande dimension qui sera injecté en entrée des dernières couches du réseau de neurones. S'il s'agit d'une tâche de classification, les dernières couches du modèle ont pour rôle de prédire la classe à partir de ce dernier vecteur. Nous pouvons utiliser des méthodes classiques de type plus proches voisins pour réaliser cette dernière tâche, mais nous avons montré sur des exemples réels qu'une méthode de classification segmentée peut améliorer légèrement les performances. Nous avons alors étudié mathématiquement la méthode proposée pour comprendre dans quel cas la segmentation peut avoir un effet positif.