

RAPPORT D'ACTIVITÉ

**Centre de calculs
AZZURRA**



**UNIVERSITÉ
CÔTE D'AZUR**

2022

Maeva ANTOINE

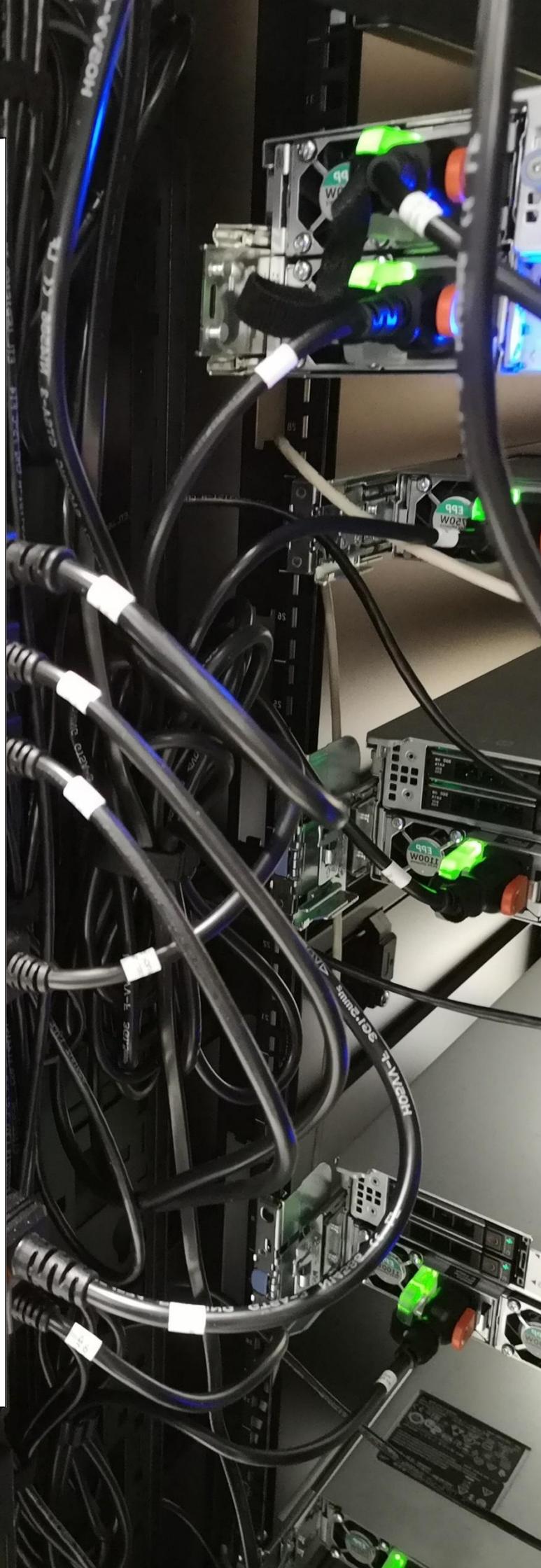


Table des matières

AZZURRA EN CHIFFRES.....	4
FAITS MARQUANTS.....	5
MATERIEL	6
FINANCEMENT	8
UTILISATEURS.....	10
SOUTIEN NUMERIQUE A LA KHARKIV NATIONAL UNIVERSITY	11
STATISTIQUES D'UTILISATION	12
FORMATIONS.....	14
FORMATIONS ET COURS AYANT EU ACCES À AZZURRA.....	15
PROFIL DES PARTICIPANTS.....	16
TEMOIGNAGES DE PARTICIPANTS.....	17
RÉSULTATS DE L'ENQUÊTE DE SATISFACTION.....	18
COMMUNICATION AUTOUR D'AZZURRA.....	21
BILAN SCIENTIFIQUE.....	23
RÉCAPITULATIF DES PROJETS DE RECHERCHE.....	25
RECAPITULATIF DES PUBLICATIONS.....	26
DESCRIPTIF DES PROJETS.....	31
Projets MSI.....	32
Chimie	48
Physique non linéaire, fluides complexes.....	73
Optique	93
Dynamique des fluides	97
Géophysique interne.....	99
Mathématiques, informatique.....	117
Magnétohydrodynamique.....	131
Astrophysique.....	142

Biologie, santé.....	156
Droit, économie.....	167
Projets industriels.....	172
OPAL.....	174
Projets accrédités OPAL par l'université ne calculant pas sur Azzurra.....	175

AZZURRA EN CHIFFRES

49 nœuds dédiés au calcul CPU, soit **1 864** cœurs.

3 nœuds dédiés au calcul GPU, soit **8** cartes GPU Tesla V100 et **4** cartes A100.

320 To pour le stockage des données.

701 201,00 € HT de matériel, prestations d'installation / support et licence logicielle depuis 2019.

En 2022 :

7 157 037 heures de calcul ont été utilisées. (**+8,4%** par rapport à 2021)

42 projets de recherche ont bénéficié d'une allocation de ressources. (**+44,83 %**)

107 personnes ont bénéficié d'un compte utilisateur sur Azzurra. (**+24,42 %**)

137 étudiants et personnels UCA ont été accrédités OPAL par l'université. (**+8,73 %**)

Depuis son lancement en mai 2020, **15 232 124** heures de calcul ont été utilisées.

FAITS MARQUANTS

Avril 2022 : mise en service du 3^{ème} nœud GPU, équipé de 4 cartes Nvidia A100 80 Go

Juillet 2022 : ouverture d'accès pour 4 chercheurs ukrainiens en exil, issus de la School of Chemistry, V.N. Karzain Kharkiv National University

Septembre 2022 : formations MPI et Introduction pratique au deep learning (42 participants)

MATERIEL



Azzurra fin 2022

Azzurra est physiquement hébergée par Inria Sophia Antipolis Méditerranée. Il est prévu qu'Azzurra soit transférée vers 2024 dans une future salle machine, antenne locale du Data Center régional. A ce moment-là, davantage de matériel pourra être installé dans le cluster.

Ajout en 2022 :

1 nœud de type R750XA financé par un laboratoire :

- 4 cartes GPU Nvidia A100 80 Go ;
- 2 CPU Intel Xeon Gold 5320 @ 2.2GHz x 26 coeurs
- 512 Go de mémoire RAM



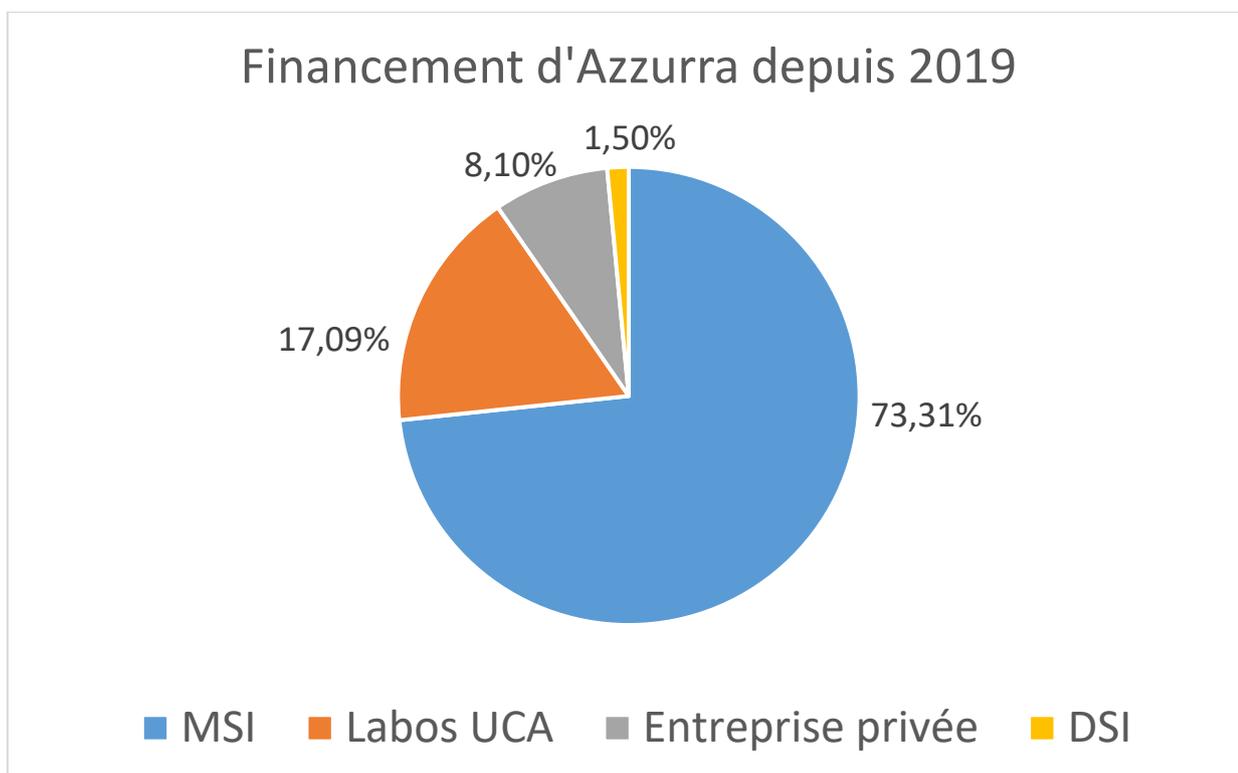
PowerEdge R750XA Server

Mise en service : avril 2022

La liste intégrale du matériel équipant Azzurra peut être consultée dans le [Rapport d'Activité Azzurra 2020-21](#).

FINANCEMENT

Depuis 2019, les **701 201,00 € HT** de dépenses en matériel, prestations d'installation, de support et licence logicielle se répartissent ainsi :



Dans le détail :

MSI (Idex) : 514046 €

Différents laboratoires de recherche UCA : 119820 €

Entreprise privée : 56835€

DSI (UCA) : 10500€

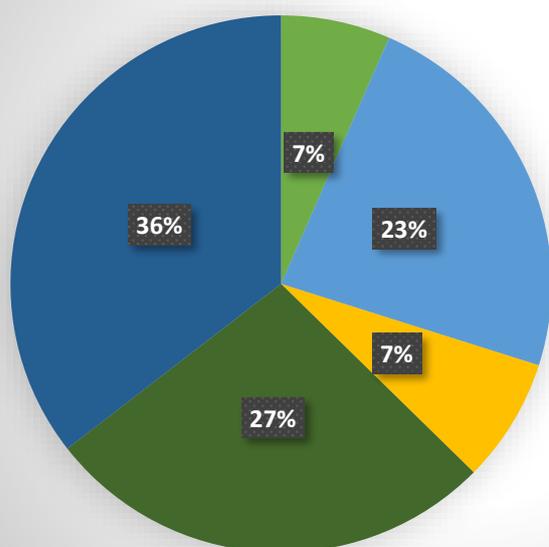
La mutualisation des ressources de calcul est encouragée auprès des laboratoires d'Université Côte d'Azur. Les équipes de recherche disposant d'un budget suffisant sont invitées à greffer leurs achats matériels à Azzurra. L'équipement acquis peut ainsi être mis à disposition de l'ensemble de la communauté de chercheurs UCA lorsqu'il n'est pas utilisé par l'équipe l'ayant financé. La maintenance et le support à l'utilisation sont assurés par la DSI / MSI.

UTILISATEURS

Durant l'année 2022, **107** personnes ont bénéficié d'un compte utilisateur sur Azzurra. **(+24,42 % par rapport à 2021)**

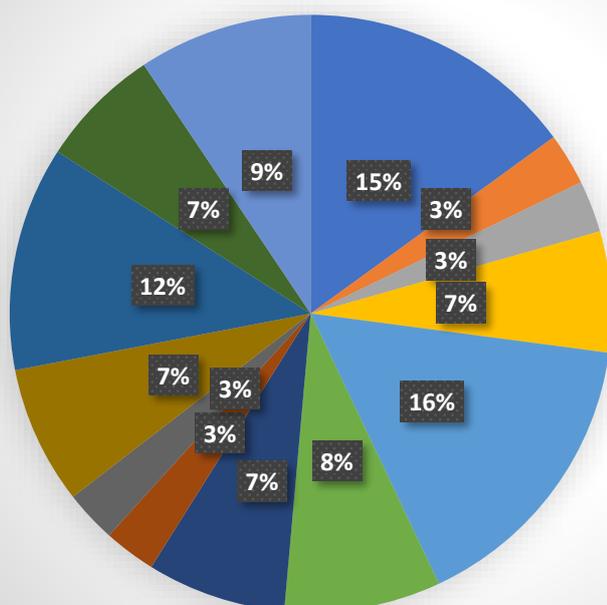
Le profil des utilisateurs est le suivant :

Utilisateurs d'Azzurra 2022



- Etudiants niveau Master
- Doctorants
- Post-doc
- Ingénieurs
- Enseignants et/ou chercheurs

Appartenance des utilisateurs



- Géoazur
- I3S
- IBV
- ICN
- INPHYNI
- INRIA
- IPMC
- IRCAN
- ISA
- Lagrange
- LJAD
- MSI
- Autres

SOUTIEN NUMERIQUE A LA KHARKIV NATIONAL UNIVERSITY

La MSI et la DSI d'Université Côte d'Azur sont mobilisées auprès des enseignants chercheurs de la V.N. Karazin Kharkiv National University pour leur permettre l'accès aux moyens de calcul et leur offrir un accompagnement nécessaire (installation de logiciels, gestion des authentifications, etc.).

Quatre chercheurs (2 professeurs et 2 doctorants) ont ainsi pu bénéficier d'un accompagnement complet, leur permettant de continuer leurs travaux de recherche malgré la situation de guerre à Kharkiv et les dégâts fait par les bombardements dans leur université et parfois dans leur logement.

Les deux projets de recherche de nos collègues ukrainiens sont décrits à partir de la [page 67](#).

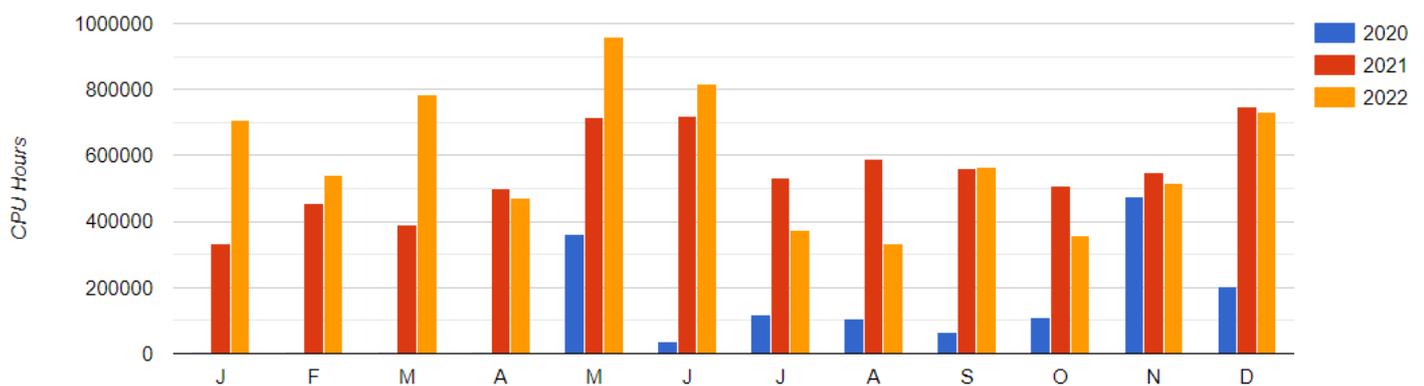
STATISTIQUES D'UTILISATION

Heures consommées

Depuis son lancement officiel en mai 2020, **15 232 124** heures de calcul ont été consommées.

En 2022, ce sont **7 157 037** heures qui ont été utilisées (+8,4% par rapport à 2021).

	2022
Janvier	706010
Février	541898
Mars	785389
Avril	473181
Mai	957964
Juin	816208
Juillet	374781
Août	333910
Septembre	563075
Octobre	358657
Novembre	514243
Décembre	731720



OCCUPATION DU CLUSTER

CPU

De janvier à mars 2022 : **1 425 103** heures disponibles par mois, en moyenne.

Depuis avril 2022 (ajout d'un nœud GPU) : environ **1 511 808** heures disponibles par mois soit 48768 heures de calcul disponibles quotidiennement.

Taux d'occupation du cluster (CPU) sur l'année : **40.47%**

2022

<i>janvier</i>	47.93%
<i>février</i>	40.73%
<i>mars</i>	53.37%
<i>avril</i>	32.34%
<i>mai</i>	63.37%
<i>juin</i>	55.79%
<i>juillet</i>	24.79%
<i>août</i>	22.09%
<i>septembre</i>	38.49%
<i>octobre</i>	23.72%
<i>novembre</i>	35.15%
<i>décembre</i>	48.40%

GPU

De janvier à mars 2022 : 5 760 heures GPU disponibles par mois.

Depuis avril 2022 : 8640 heures GPU disponibles par mois.

Taux d'occupation des ressources GPU sur l'année : **38.61%**

2022

<i>janvier</i>	22.70%
<i>février</i>	45.47%
<i>mars</i>	25.83%
<i>avril</i>	48.65%
<i>mai</i>	50.05%
<i>juin</i>	42.62%
<i>juillet</i>	39.90%
<i>août</i>	67.18%
<i>septembre</i>	42.78%
<i>octobre</i>	26.49%
<i>novembre</i>	18.38%
<i>décembre</i>	26.46%

FORMATIONS



Depuis 2021, Azzurra est utilisée comme plateforme servant de support à différentes formations. Des comptes temporaires sur Azzurra sont fournis aux inscrits le temps du cours.

Université Côte d'Azur a accueilli en septembre 2022, dans l'amphithéâtre du LJAD, six formateurs de l'[IDRIS](#) (Orsay) venus dispenser deux formations (MPI et Introduction Pratique au Deep Learning). Ces formations ont pu avoir lieu grâce au soutien de la MSI.

FORMATIONS ET COURS AYANT EU ACCES À AZZURRA

- **MPI**

- Formation
- Durée : 21 heures
- Nombre de participants : 18
- Public : ouvert à tout UCA
- Informations : sur le [site de l'IDRIS](#)

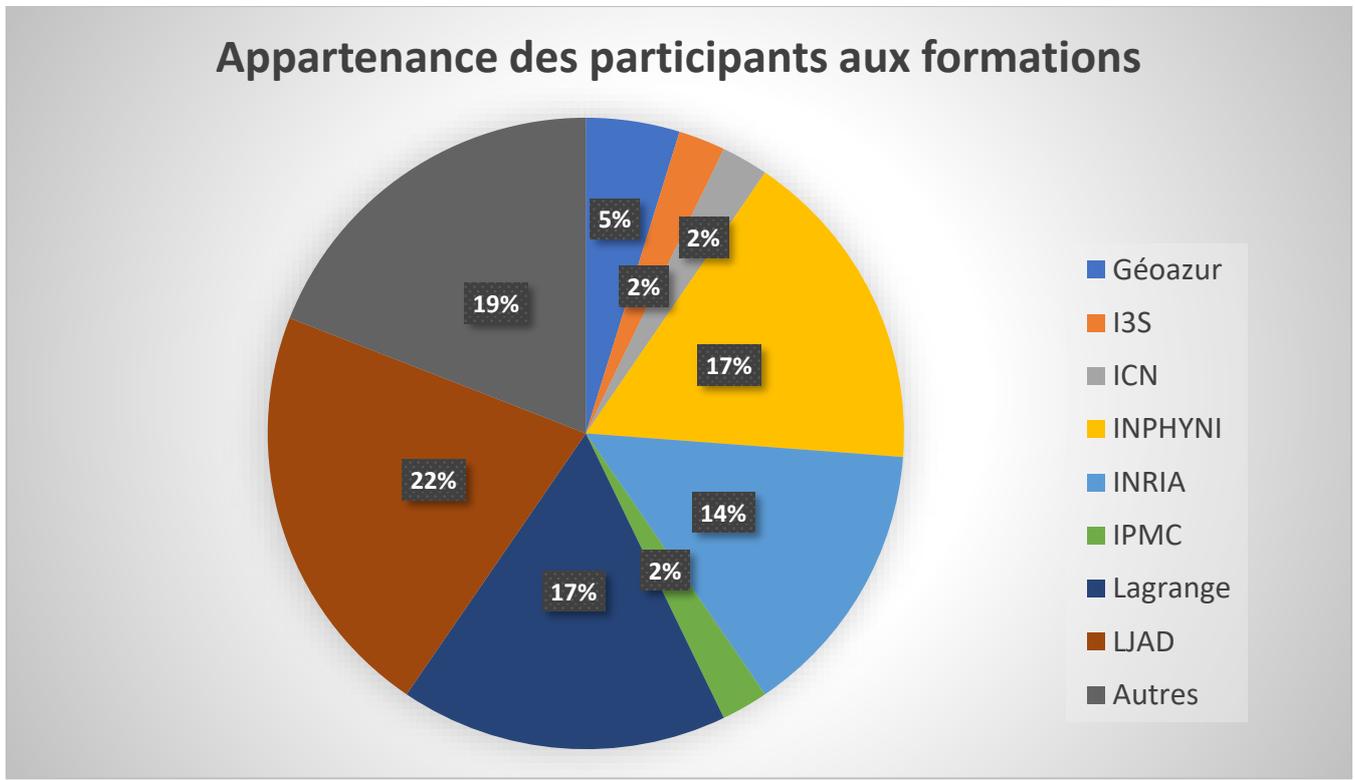
- **Introduction Pratique au Deep Learning**

- Formation
- Durée : 17 heures
- Nombre de participants : 24
- Public : ouvert à tout UCA
- Informations : sur le [site de l'IDRIS](#)

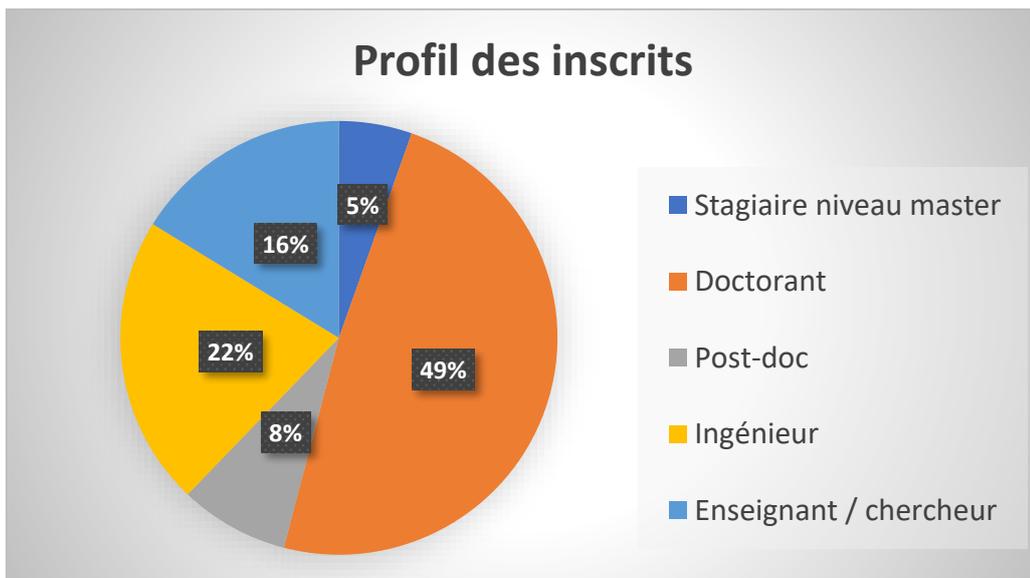
- **ECUE Données massives**

- Cours
- Durée : 4 heures
- Nombre de participants : 11
- Public : M1 Biologie, informatique et mathématiques (EUR LIFE)

PROFIL DES PARTICIPANTS



Appartenance des participants inclus dans « Autres » : NeuroMod, Artemis (OCA), Ermes, ISA, CEPAM, IRCAN, Polytech'Lab.



TEMOIGNAGES DE PARTICIPANTS

Mireille Bossy, Directrice de Recherche à Inria :

Merci aux membres de la MSI pour l'excellente organisation de cette formation. Les formateurs de l'IDRIS ont très bien su cibler les besoins et les niveaux, tout en s'adaptant à la durée de la formation. L'accès à Azzurra pour les TPs a permis une très bonne mise en situation. Un format sur trois jours très efficace. Je recommande !

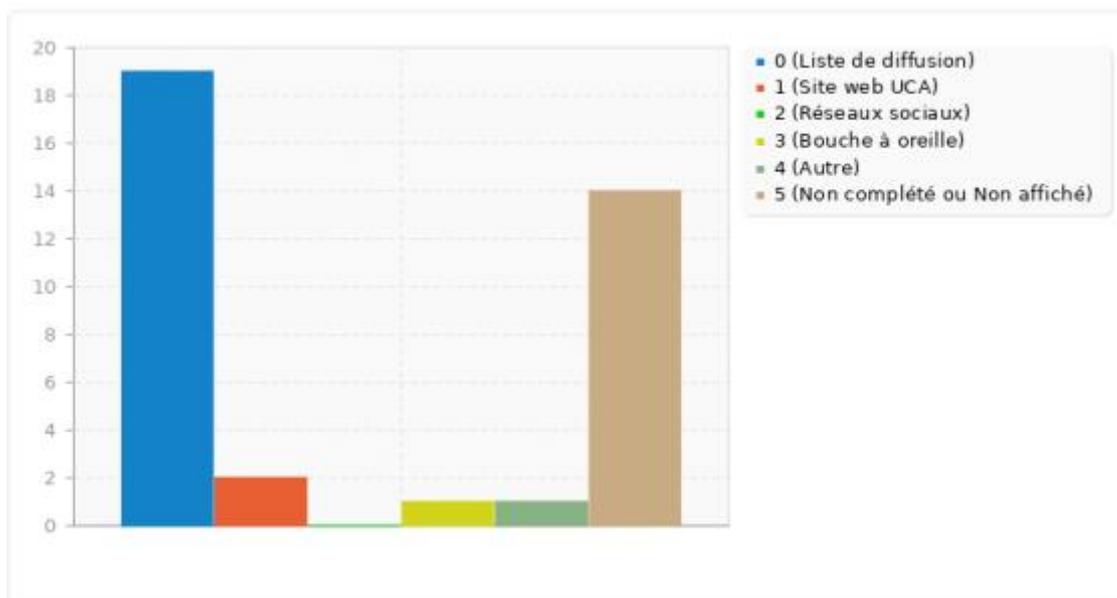
Michel ORSI, doctorant à l'INPHYNI :

J'avais participé l'année dernière à la formation OpenMP et, étant donné qu'elle avait été excellente, j'ai décidé de m'inscrire aussi à cette-ci sur MPI. Je fais une thèse numérique sur la rhéologie des suspensions et j'ai écrit mon propre code, il est donc fondamental pour moi savoir le paralléliser pour diminuer les temps de calculs. Cette formation a été simplement géniale : la durée, le contenu, les intervenants, la qualité de la formation, tout a été parfait. Ça m'a sûrement donné envie de me replonger dans mon code et essayer de l'optimiser encore plus !

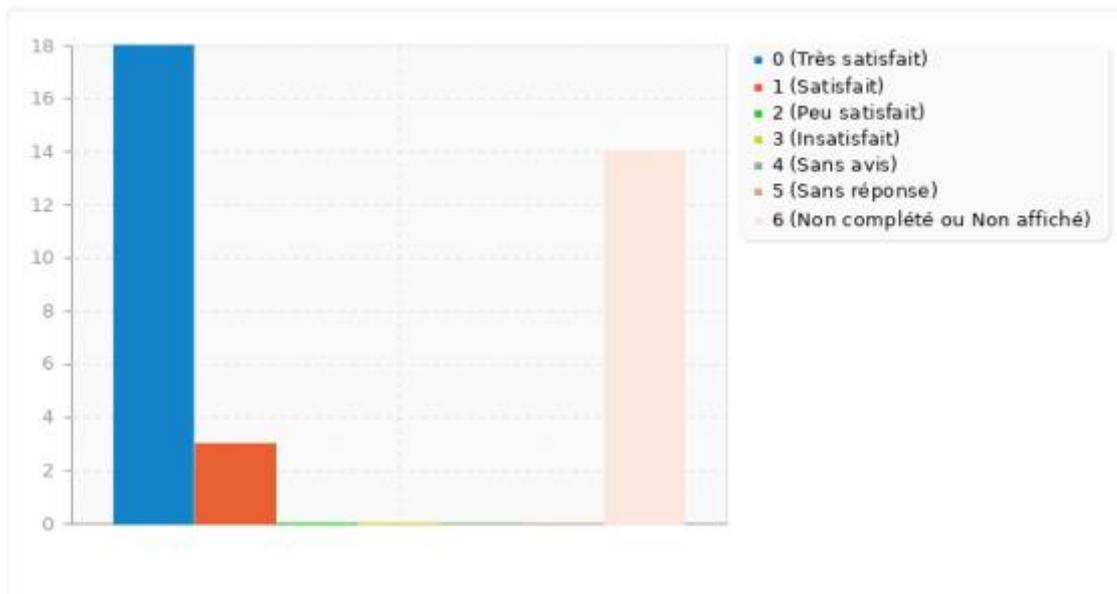
RÉSULTATS DE L'ENQUÊTE DE SATISFACTION

Suite aux formations, une enquête de satisfaction anonymisée a été envoyée aux participants. 21 personnes y ont répondu :

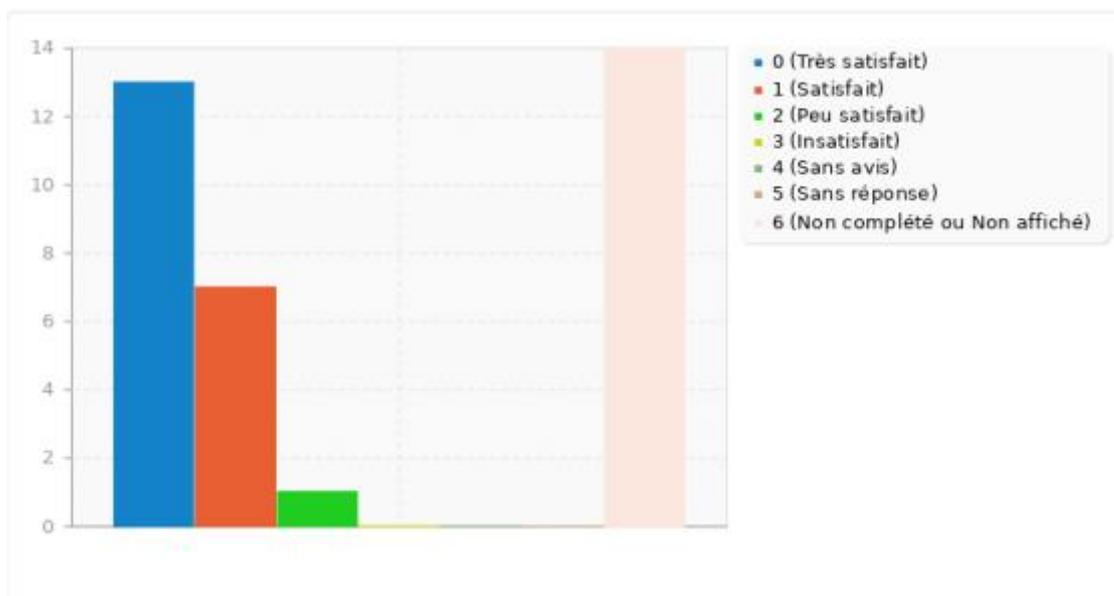
Comment avez-vous pris connaissance de l'organisation de ces formations ?



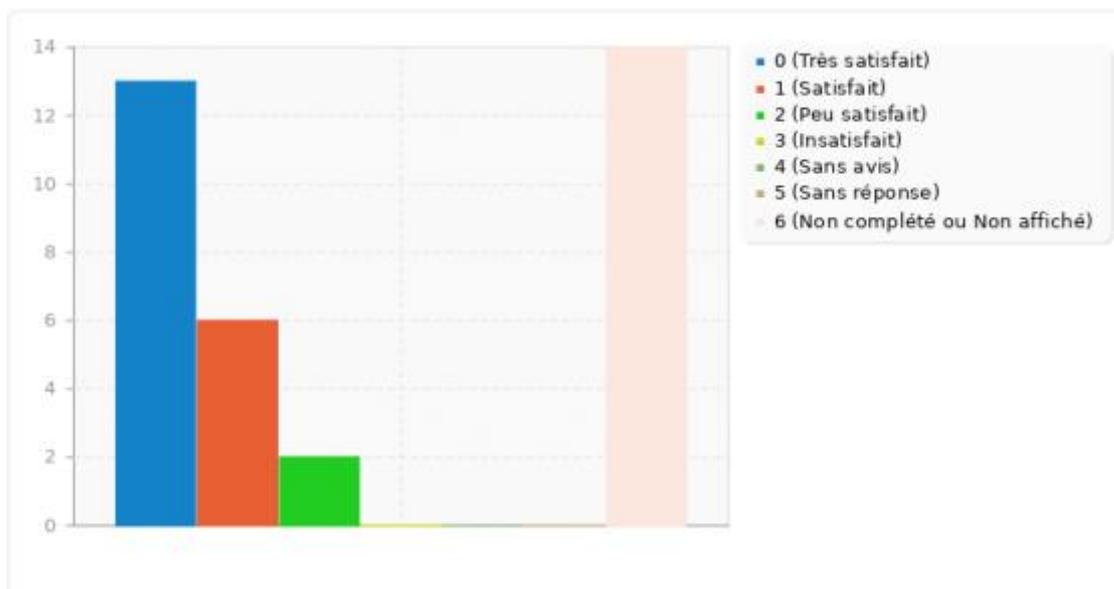
Organisation générale des formations (accueil, salle, communication, etc.)



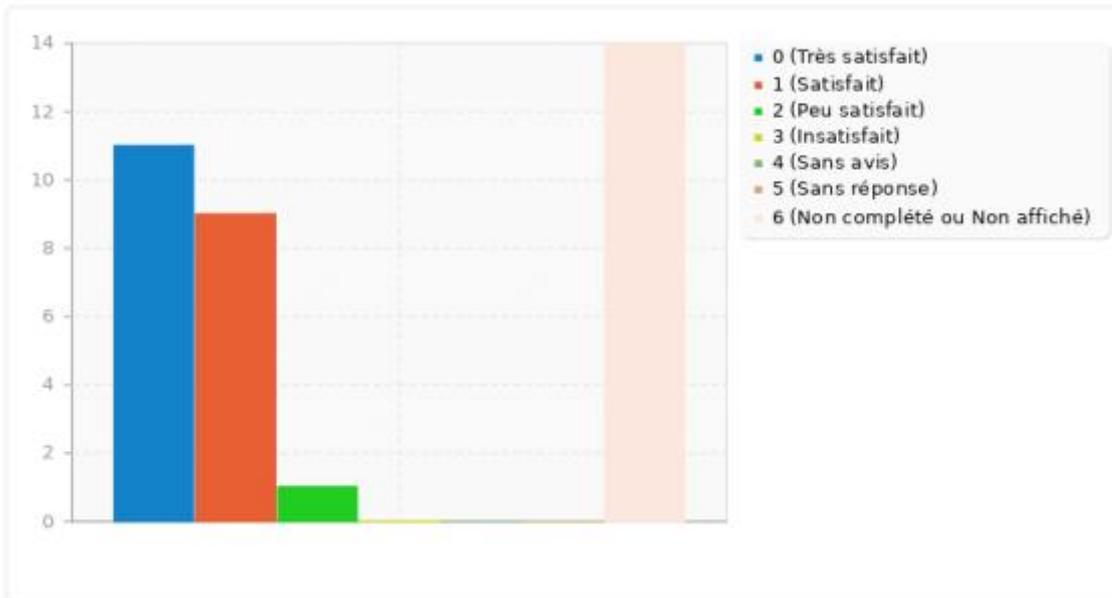
Thème et contenu des cours



Connaissances acquises grâce au cours



Durée de la formation



Niveau de difficulté de la formation



COMMUNICATION AUTOUR D'AZZURRA

L'actualité relative à Azzurra est relayée via différents canaux de communication au sein d'UCA. Quelques exemples cette année :

Soutien numérique à la Kharkiv National University :

<https://univ-cotedazur.fr/universite/ukraine>

Vidéo tournée lors des formations sur la chaîne YouTube UCA :

https://www.youtube.com/watch?v=XW_HM2gU888

Compte Twitter de la MSI :

https://twitter.com/MSI_UCA/status/1575780013647224833

https://twitter.com/MSI_UCA/status/1566759904823721991

https://twitter.com/MSI_UCA/status/1567810236748021762

https://twitter.com/MSI_UCA/status/1496505439877570564

Mention d'articles dans les Newsletters IDEX de :

mars 2022

décembre 2022

Mention sur Azzurra et les formations dans les actualités diffusées par e-mail « Quoi de neuf cette semaine ? » #16 (17/10/2022)

Articles régulièrement postés sur le site UCA MSI :

<https://msi.univ-cotedazur.fr/>

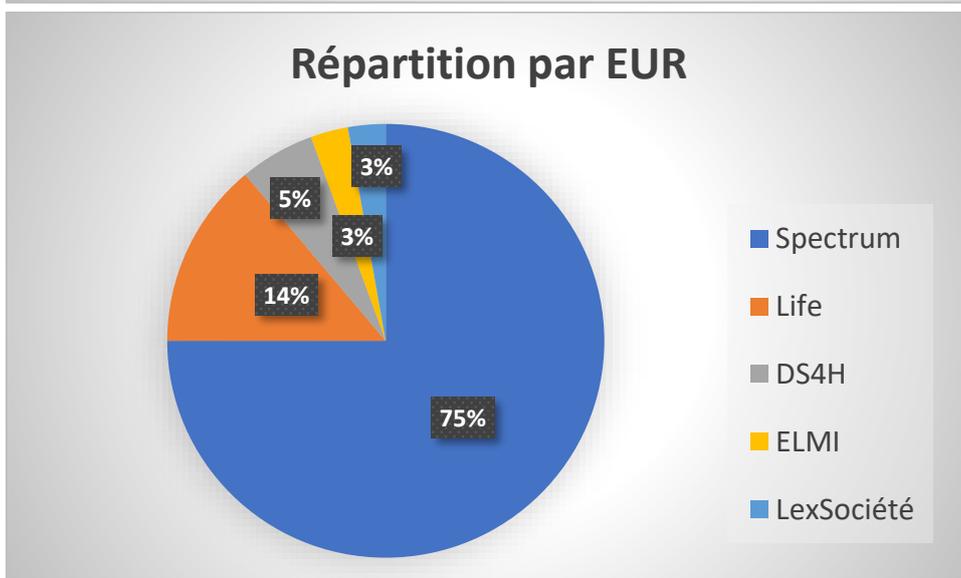
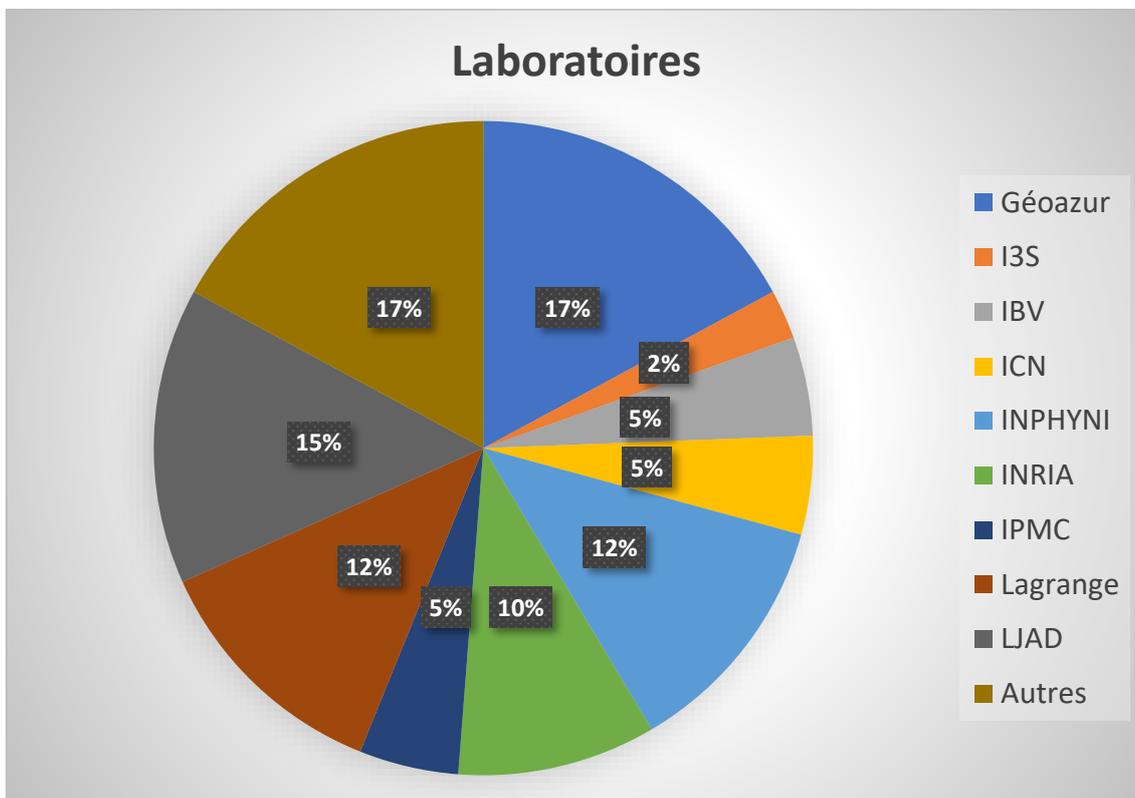
Octobre : <https://msi.univ-cotedazur.fr/actualites/calcul-scientifique-formations-passees-et-futures-par-la-msi>

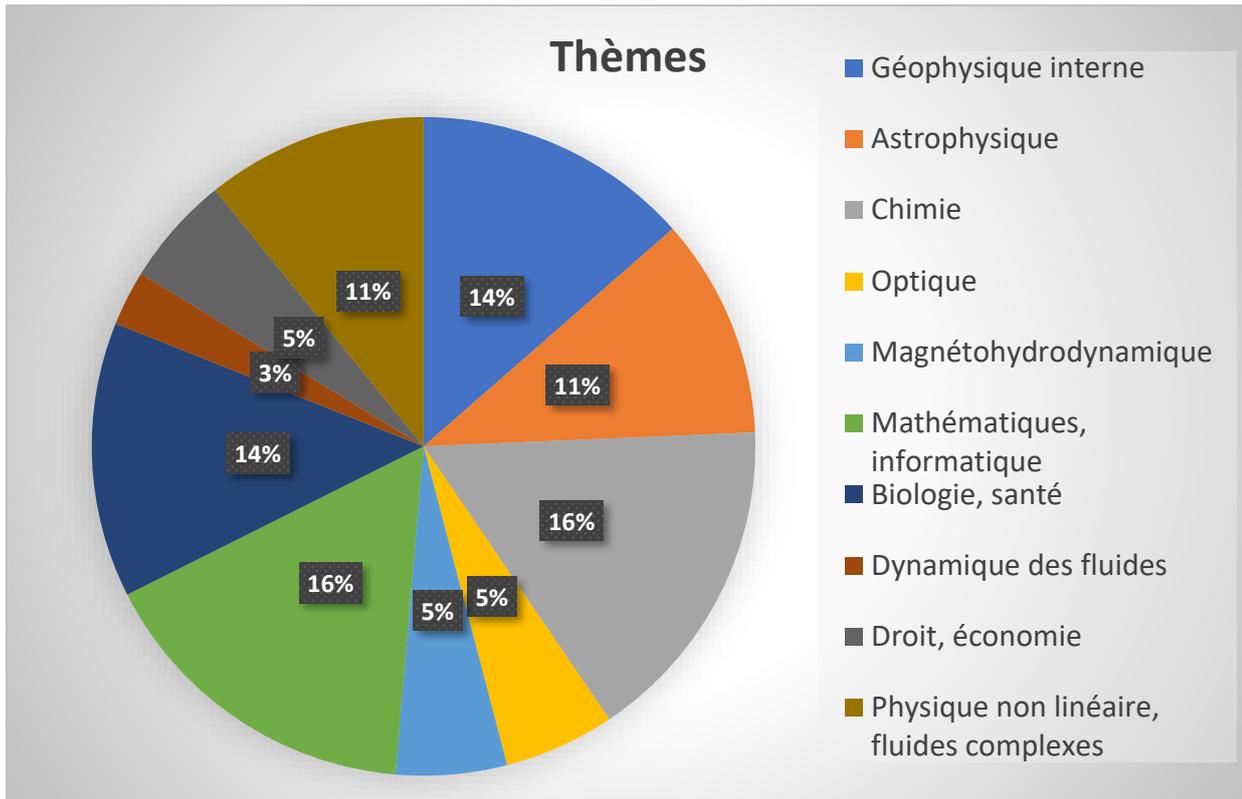
Septembre : <https://msi.univ-cotedazur.fr/actualites/graines-de-dynamo>

Juin : <https://msi.univ-cotedazur.fr/offre-de-formation-en-calcul-intensif-et-ia-par-la-msi>

BILAN SCIENTIFIQUE

En 2022, **42** projets ont bénéficié d'une allocation de ressources sur Azzurra. Ces projets se répartissent ainsi :





Dans le détail :

Géophysique interne (Géoazur; EUR Spectrum)	5
Astrophysique (Géoazur, Lagrange; EUR Spectrum)	4
Chimie (ICN, LJAD; EUR Spectrum)	6
Optique (Inphyni; EUR Spectrum)	2
Magnétohydrodynamique (Lagrange, LJAD; EUR Spectrum)	2
Mathématiques, informatique (I3S, Inria, LJAD; EUR DS4H, Spectrum)	6
Biologie, santé (IPMC, ISA; EUR Life)	5
Dynamique des fluides (CEMEF, LJAD; EUR Spectrum)	1
Droit, économie (ERMES; EUR LexSociété – INSTAR ; EUR ELMi)	2
Physique non linéaire, fluides complexes (Inphyni; EUR Spectrum)	4

RÉCAPITULATIF DES PROJETS DE RECHERCHE

Porteur	Laboratoire	Discipline	EUR	CLASSIFICATION
Becavin	IPMC	Pharmacologie	Life	Biologie, santé
Masson	LJAD	Mathématiques	Spectrum	Mathématiques, informatique
Operto	Geoazur	Géophysique	Spectrum	Géophysique interne
Lanteri	Inria	Mathématiques	Spectrum	Mathématiques, informatique
Dolean	LJAD	Mathématiques	Spectrum	Mathématiques, informatique
Fienga	Geoazur	Astrophysique	Spectrum	Astrophysique
Sigloch	Geoazur	Géophysique	Spectrum	Géophysique interne
Aubry	InPhyNi	Physique	Spectrum	Optique
Doya	InPhyNi	Physique	Spectrum	Optique
Vincenzi	LJAD	Mécanique des fluides	Spectrum	Dynamique des fluides
Danchin	Agrobiotech	Génomique	Life	Biologie, santé
Glaichenhaus	IPMC	Pharmacologie	Life	Biologie, santé
Bossy	Inria	Mathématiques	Spectrum	Mathématiques, informatique
Fiorucci	ICN	Chimie	Spectrum	Chimie
Sladen	Geoazur	Géophysique	Spectrum	Géophysique interne
Méheut	Lagrange	Astrophysique	Spectrum	Astrophysique
Cassam-Chenaï	LJAD	Chimie	Spectrum	Chimie
Nazarenko	InPhyNi	Physique/Mécanique des fluides	Spectrum	Physique non linéaire, fluides complexes
Marcotte	LJAD	Mathématiques	Spectrum	Magnétohydrodynamique
Ponty	Lagrange	Mécanique des fluides	Spectrum	Magnétohydrodynamique
Pasquier	I3S	Informatique	DS4H	Mathématiques, informatique
Fontaine-Vive	ICN	Chimie	Spectrum	Chimie
Duclaux	Geoazur	Géophysique	Spectrum	Géophysique interne
Peters	InPhyNi	Physique/Mécanique des fluides	Spectrum	Physique non linéaire, fluides complexes
Seminara	InPhyNi	Physique	Spectrum	Physique non linéaire, fluides complexes
Ampuero	Geoazur	Géophysique	Spectrum	Géophysique interne
Krstulovic	Lagrange	Physique/Mécanique des fluides	Spectrum	Physique non linéaire, fluides complexes
Henry	Inria	Chimie/Mathématiques	Spectrum	Chimie
Roux	IRCAN	Médecine	Life	Biologie, santé
Villata	Inria	Informatique	DS4H	Mathématiques, informatique
Rousseau	Lagrange	Astrophysique	Spectrum	Astrophysique
Czellar	INSTAR	Economie	ELMI	Droit, économie
Constant	Ermes	Droit	LexSociété	Droit, économie
Besse	IBV	Biologie	Life	Biologie, santé
Chiavassa	Lagrange	Astrophysique	Spectrum	Astrophysique
Kalugin	-	Chimie	-	Chimie
Kyrychenko	-	Chimie	-	Chimie

RECAPITULATIF DES PUBLICATIONS

Journaux :

- Badia, A., D'Angelo, Y., Peters, F., & Lobry, L. (2022). Frame-invariant modeling for non-Brownian suspension flows. *Journal of Non-Newtonian Fluid Mechanics*, 309, 104904.
- Rigolli, N., Magnoli, N., Rosasco, L., & Seminara, A. (2022). Learning to predict target location with turbulent odor plumes. *Elife*, 11, e72196. <https://doi.org/10.7554/eLife.72196>
- Rigolli, N., Reddy, G., Seminara, A., & Vergassola, M. (2022). Alternation emerges as a multi-modal strategy for turbulent odor navigation. *Elife*, 11, e76989. <https://doi.org/10.7554/eLife.76989>
- Zhu, Y., Semisalov, B., Krstulovic, G., & Nazarenko, S. (2022). Testing wave turbulence theory for the Gross-Pitaevskii system. *Physical Review E*, 106(1), 014205. <https://doi.org/10.1103/PhysRevE.106.014205>
- Chlortetracycline, a Novel Arf Inhibitor That Decreases the Arf6-Dependent Invasive Properties of Breast Cancer Cells, E. Macia, M. Vazquez-Rojas, A. Robiolo, R. Fayad, S. Abélanet, I. Mus-Veteau, F. Fontaine-Vive, M. Mehiri, F. Luton, M. Franco, *Molecules*, 2021, 26(4), 969
- Yazhgur, P., Aubry, G. J., Froufe-Pérez, L. S., & Scheffold, F. (2022). Scattering phase delay and momentum transfer of light in disordered media. *Physical Review Research*, 4(2), 023235. <https://doi.org/10.1103/PhysRevResearch.4.023235>
- Nonoverlapping gene networks regulated by mating. Claude Pasquier, Alain Robichon. *Life Science Alliance* Nov 2021, 5 (2) e202101119; DOI: 10.26508/lsa.202101119
- Paul M. Mannix, Yannick Ponty, and Florence Marcotte "Systematic Route to Subcritical Dynamo Branches" *Phys. Rev. Lett.* 129, 024502 – Published 5 July 2022 <https://doi.org/10.1103/PhysRevLett.129.024502>
- Martinuzzi, E., S. Barbosa, P. Courtet, E. Olie, S. Guillaume, E. C. Ibrahim, D. Daoudlarian, L. Davidovic, N. Glaichenhaus and R. Belzeaux (2021). "Towards a blood test to differentiate bipolar disorder and major depressive disorder during a major depressive episode." *Brain Behav Immun Health*, Mar10;13:100232; <https://doi.org/10.1016/j.bbih.2021.100232>
- Barbosa, S., O. Khalfallah, A. Forhan, C. Galera, B. Heude, N. Glaichenhaus and L. Davidovic (2020). "Serum cytokines associated with behavior: A cross-sectional study in 5-year-old children." *Brain Behav Immun*, Jul;87:377-387. <https://doi.org/10.1016/j.bbi.2020.01.005>
- Barbosa, S., O. Khalfallah, A. Forhan, C. Galera, B. Heude, N. Glaichenhaus and L. Davidovic (2020). "Immune activity at birth and later psychopathology in childhood." *Brain Behav Immun Health* 8: 100141. <https://doi.org/10.1016/j.bbih.2020.100141>

- Corentin Lothode, Jules Poncin, Didier Lemosse, David Gross, Eduardo Souza de Cursi, Investigation of blade-mast fluid-structure interaction of a tidal turbine, *Ocean Engineering*, Volume 261, 2022, 112046, ISSN 0029-8018, <https://doi.org/10.1016/j.oceaneng.2022.112046>
- Michel Orsi, Laurent Lobry, François Peters, Frame-invariant sub-grid corrections to the Fictitious Domain Method for the simulation of particulate suspensions in nonlinear flows using OpenFOAM, *Journal of Computational Physics*, Volume 474, 2023, 111823, ISSN 0021-9991, <https://doi.org/10.1016/j.jcp.2022.111823>

Conférences :

- Cédric Vincent-Cuaz, Rémi Flamary, Marco Corneli, Titouan Vayer, and Nicolas Courty. Template based graph neural network with optimal transport distances. In *Advances in Neural Information Processing Systems*, 2022b. URL: <https://openreview.net/forum?id=seYcx6CqPe>
- Djampa Kozłowski. Nerd : a tool to predict proteins involved in nematode plant parasitism. SophIA Summit 2022.
- Djampa Kozłowski. Nerd : a tool to predict proteins involved in nematode plant parasitism. Journée Annotation de génomes, Intelligence Artificielle et Text-Mining' organisée par le PEPI IBIS (communauté d'ingénieurs INRAe) le 28/11/2022.
- Gremmel, J., Duclaux, G., Corsini, M., Maillet, A., Jourdon, A., & Bascou, J. (2022). Structure, strain partitioning and exhumation mechanism during the late stage oblique tectonic evolution of the Variscan Tanneron massif (SE France) (No. EGU22-967). *Copernicus Meetings*.
- Cerpa, N., Sigloch, K., Garel, F., Davies, R., and Heuret, A.: Subduction dynamics through the mantle transition zone in the presence of a weak asthenospheric layer, *EGU General Assembly 2022*, Vienna, Austria, 23–27 May 2022, EGU22-3822, <https://doi.org/10.5194/egusphere-egu22-3822> , 2022.
- Sahar Borzooei, Victorita Dolean, Claire Migliaccio and Pierre-Henri Tournier, "Solution of time-harmonic Maxwell's equations by a domain decomposition method based on PML transmission conditions" - Poster at SCEE 2022, The 14th International Conference on Scientific Computing in Electrical Engineering
- Sahar Borzooei. A domain decomposition method with PML transmission conditions for Maxwell's equation. The 27th International Domain Decomposition Conference, DD27, in Prague, Czech Republic
- Tests of GR with planetary ephemerides with BepiColombo simulations. Agnès Fienga (Géoazur, OCA), 56th Rencontres de Moriond "Gravitation", January 30th - February 6th 2022.

https://moriond.in2p3.fr/2022/Gravitation/transparencies/02_tuesday/01_morning/04_fie nga.pdf

- Faye, Antoine & Carrere, N & Sacher, Matthieu & Hauville, F & Nême, A. (2022). Determination of the equivalent structural properties of a composite hydrofoil. 25ème Congrès Français de Mécanique. Nantes, 20 août au 2 septembre 2022
https://www.researchgate.net/publication/363406203_Determination_of_the_equivalent_structural_properties_of_a_composite_hydrofoil
- Identification systématique de branches dynamo sous-critiques. P.M. Mannix, Y. Ponty, F. Marcotte, 25ème Congrès Français de Mécanique, Nantes, 29 août-2 septembre 2022 - présentation orale (Paul Mannix)
- Identification of subcritical dynamos and minimal seeds. Festival de Théorie d'Aix-en-Provence, Chaos Control, Feedback & Model Reduction, 4 juillet 2022 - séminaire invité (Florence Marcotte)
- A systematic approach to identify subcritical dynamo instabilities. Courant Institute of Mathematical Sciences, New York University, Magneto-Fluid Dynamics Seminar, 7 décembre 2021 - séminaire invité (Florence Marcotte)

Workshops :

- Yanqi Wu, Hossein S. Aghamiry, Stephane Operto, Jianwei Ma. Wave simulation in non-smooth media by PINN with quadratic neural network and PML condition. Workshop Artificial Intelligence for Waves - AI4W 2022. https://wavecomplexity.univ-cotedazur.fr/events/ai4w2022/contribs/yanqi_wu/
- 4th BepiColombo GGWG Virtual Meeting (21st January 2022) : Agnes Fienga, "INPOP Planetary Ephemerides and BepiColombo : from Orbit Improvements to GR and Alternative Theories tests"
<https://www.cosmos.esa.int/web/bepicolombo-ggwg>
- Easy paths to subcritical dynamo branches. Isaac Newton Institute for Mathematical Sciences, University of Leeds, Fluid flow and magnetic field generation in fluids and plasmas - theory and laboratory experiments, 17 octobre 2022 - présentation orale (Yannick Ponty)
- Minimal seeds for subcritical dynamos. Isaac Newton Institute for Mathematical Sciences, University of Cambridge, Dynamos in planets and stars: similarities and differences, 14 septembre 2022 - séminaire invité (Florence Marcotte)

Thèses :

- Simulation de suspensions non-Browniennes concentrées frictionnelles et adhésives en écoulement linéaire et non-linéaire par Michel Orsi. Soutenue le 2 décembre 2022. <https://theses.fr/s317644>
- Navigation olfactive : comment prendre des décisions avec un signal brisé par Nicola Rigolli. Soutenue le 19 mai 2022. <https://theses.fr/2022COAZ4023>
- Modélisation moléculaire des récepteurs chimiosensoriels par Jody Pacalon. Soutenue en novembre 2022 <https://theses.fr/s317702>

Posters :

- Bayesian planetary numerical ephemerides B-INPOP with MCMC. Vincenzo Mariani. Poster pour Bayesian deep learning for cosmology and time-domain astrophysics 2022 (20/06 - 24/06 2022).
- Imaging High-Frequency Radiation of Large Earthquakes with a Bayesian Inversion Method. Carolina Paz Gutiérrez Morales. Poster pour the IV Assembly of the Latin American and Caribbean Seismological Commission, Quito, Ecuador, Oct. 2022.
- Predicting the combinatorial code via graph neural networks and representation learning M. Hladiš, S. Fiorucci, J. Topin, AChemS XLIV, Bonita Springs (USA), 20-23 avril 2022.
- Integrative structural modeling reveals functional molecular switches of human G protein-coupled bitter-taste receptors J. Topin, C. Bouysset, Y. Kim, M. Rhyu, J. Golebiowski, S. Fiorucci. AChemS XLIV, Bonita Springs (USA), 20-23 avril 2022.
- Structure-based virtual screening of bitter taste receptors M. Lalis, C. Belloir, J. Topin, L. Briand, S. Fiorucci. ERNEST 6, webinaire, 28-31 mars 2022.
- Graines minimales de transition dynamo. 25e Rencontre du Non-Linéaire, Université Paris Cité, 29 – 31 mars 2022, P. M. Mannix, Y. Ponty, Florence Marcotte, présenté par Florence Marcotte.
- Advances in geophysical and astrophysical turbulence TURW05, Isaac Newton Institute, University of Cambridge, 16-20 mai 2022, présenté par Paul Mannix.

Pré-publications ou en cours d'écriture :

- Fanny Simões, Charles Bouveyron, Frédéric Precioso. DeepWILD: Wildlife Identification, Localisation and estimation on camera trap videos using Deep learning. <https://hal.archives-ouvertes.fr/hal-03797530>
- Polymer stretching in laminar and random flows: entropic characterization. Stefano Musacchio, Victor Steinberg and Dario Vincenzi. <https://arxiv.org/pdf/2112.01344.pdf>

-
- The effect of a weak asthenospheric layer on surface kinematics, subduction dynamics and slab morphology in the lower mantle. Nestor Cerpa, Karin Sigloch, Fanny Garel, Arnaud Heuret, D. Rhodri Davies, Mitchell G. Mihalynuk <https://eartharxiv.org/repository/view/3230/>
 - Zhu Y, Semisalov B, Krstulovic G, et al. Direct and inverse cascades in turbulent Bose-Einstein condensate. arXiv preprint arXiv:2208.09279, 2022. <https://doi.org/10.48550/arXiv.2208.09279>
 - Bastien Buchwalter, Veronika Czellar, Engin Iyidogan. Multifractal Cryptocurrencies. (soumis au Journal of Econometrics fin 2022)
 - Discrete adjoint-based control: A robust gradient descent procedure for optimisation with PDE and norm constraints (<https://arxiv.org/pdf/2210.17194.pdf>). Mannix, P.M., Skene, C.S., Auroux, D. & Marcotte, F.

Stages et projets étudiants niveau master :

- Minh Nhat Do, Marco Milanesio. Image Classification on Corals using ensemble deep Transfer-learning approach. Rapport de stage Master 2 Data science & Artificial intelligence (2022).
- Towards the identification of sequences involved in alternative splicing caused by the loss of LncRNA SNHG21 par Anna Diamant, 2022, stage de Master 2 SV (Sciences du Vivant) Biology, Informatics and Mathematics (UCA)
- Association entre cytokines à la naissance et troubles émotionnels chez l'enfant à 5 ans : réplication et validation. Stage de Benjamin Lapijover à l'IPMC en 2022 (étudiant Master 2 SV BIM à UCA).

DESCRIPTIF DES PROJETS

Projets MSI

La Maison de la Modélisation, de la Simulation et des Interactions (MSI) doit sa création au projet UCA^{JEDI} pour lequel en 2016 la ComUE Université Côte d'Azur a remporté le prix Initiatives d'Excellence (IDEX) du gouvernement français.

Son but : favoriser et structurer les échanges transdisciplinaires dans la recherche académique et les partenariats de recherche public privé dans la région SUD sur tous les domaines qui sollicitent des expertises fortes en matière de modélisation, de simulation et de traitement de données massives. La MSI intervient en appui du réseau d'acteurs qui constituent et animent Université Côte d'Azur.

La MSI est pilotée par un comité composé de huit chercheurs issus de différents laboratoires d'Université Côte d'Azur et présidé par un directeur nommé par le Président du projet IDEX UCA^{JEDI} ; ce comité s'appuie sur une équipe de chercheurs et d'ingénieurs experts et sur une chargée de mission.

Les ingénieurs de la MSI s'investissent sur des projets scientifiques développés au sein de centres d'expertise de la MSI ou au niveau d'Université Côte d'Azur. De plus, dans le cadre d'appels à projets, la MSI soutient des projets de recherche interdisciplinaires et/ou tournés vers des partenariats industriels. Elle propose également une offre de formation pour chercheurs et ingénieurs via des séminaires réguliers et des écoles thématiques.

La MSI ayant financé Azzurra, ses ingénieurs ont accès au cluster pour leurs différents projets, décrits ci-dessous.

ARCH-AI-STORY

Structure d'accueil

- Lab : CEPAM, Université Côte d'Azur
- Author(s) : Marco Corneli (Marco.CORNELI@unice.fr), Elysandre Puech (PhD student, CEPAM), Isabelle Théry-Parisot (CNRS-CEPAM)
- Object : Supervised classification of archaeological wood charcoals (SEM images) - *performed on the Azzurra computing center*

Description

In the context of the [ARCH-AI-STORY project](#), aiming at exploiting the recent IA advances to boost the research in history and archaeology, we trained a classifier (a deep convolutional neural net) onto scanning electron microscopic (SEM) images (see Figure 1) in order to identify archaeological remains of wood charcoals. In more details, we have a reference collection of wood charcoal MES images, from the site of Busham Rock Shelter (South Africa). Each image corresponds to a specimen belonging to a family that we seek to predict.

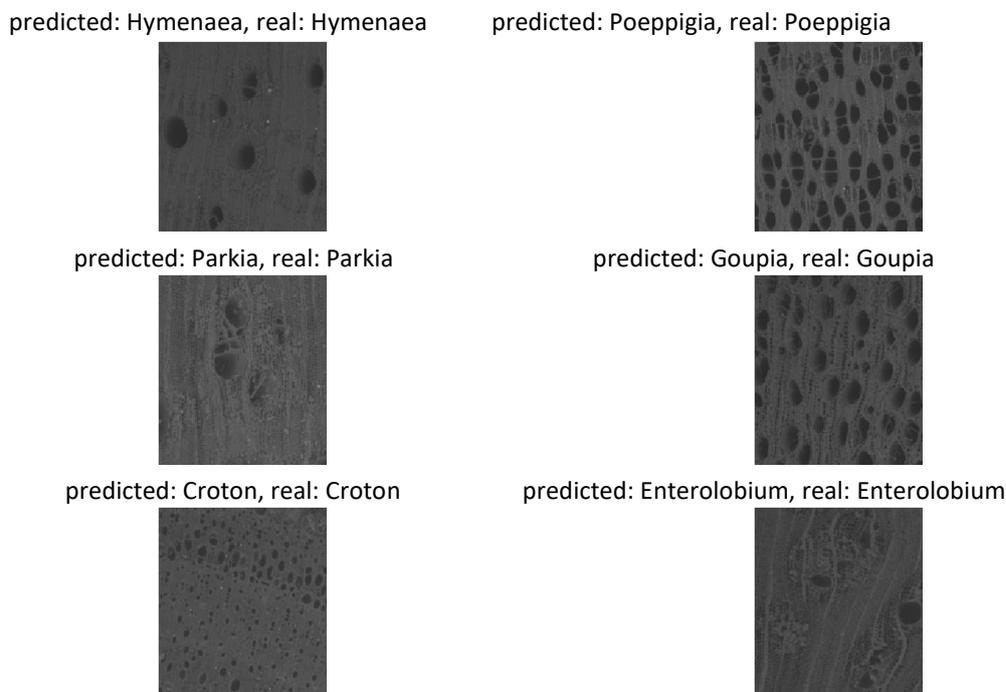


Figure 1 – Images to classify : real vs predicted family

Then, the idea is to adopt the trained CNN (ResNet 18 architecture) to predict the class/family or archaeological wood fragments, whose SEM images are available at the CEPAM. The neural net was trained on the Azzurra computing center, exploiting GPUs acceleration (Figure 2) In order to entrust the researchers in history/archaeology toward IA, we are currently adopting gradient based approaches in order to capture the regions of the images mostly adopted by the classifier for the identification (see Figure 3). This might suggest to the anthracologist some useful insights and allowing them to interpret AI.

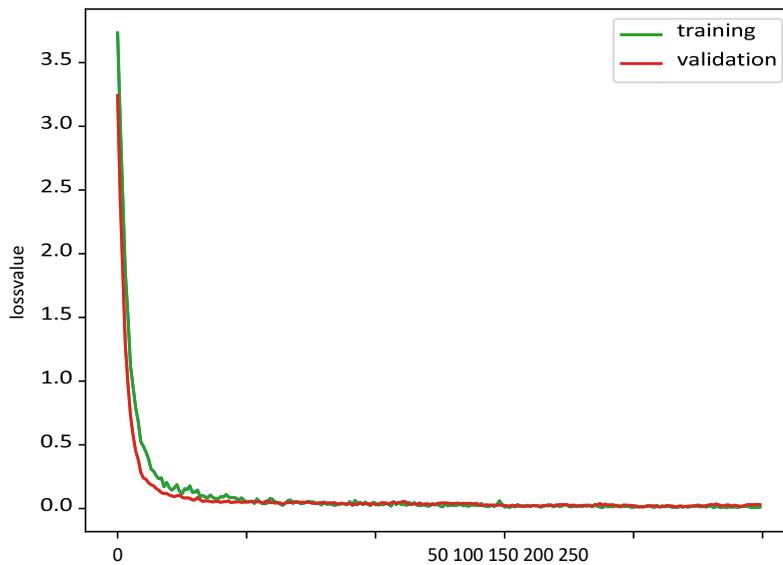


Figure 2 – Training and validation losses per number of epochs

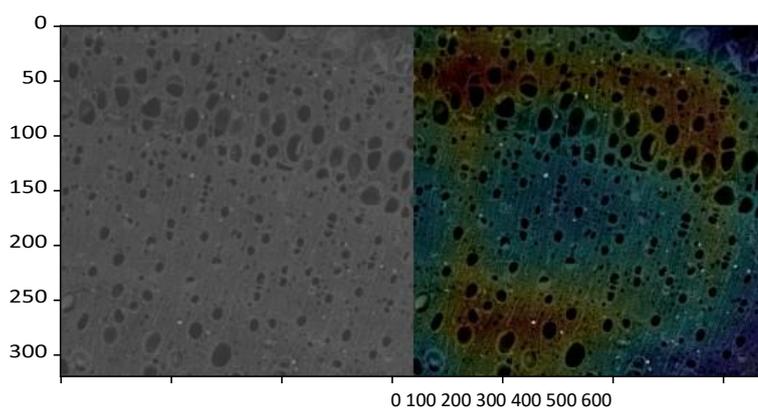


Figure 3 – GradCam heat-map highlighting the areas of the SEM images being used by the classifier for the class attribution.

Optimal Transport for structured data

Sujet de thèse de Cédric Vincent Cuaz, co-encadré par Marco Corneli, ingénieur MSI en partenariat avec l'équipe INRIA – MASAAI. <http://www.theses.fr/s319700>

Résultats obtenus grâce aux moyens informatiques attribués :

Cédric Vincent-Cuaz, Rémi Flamary, Marco Corneli, Titouan Vayer, and Nicolas Courty. Template based graph neural network with optimal transport distances. In Advances in Neural Information Processing Systems, 2022b. URL: <https://openreview.net/forum?id=seYcx6CqPe>

Description :

Les récents progrès en matière de transport optimal (OT) régularisé ont apporté de nouveaux solveurs efficaces et ouvert la porte à de nombreuses applications dans l'apprentissage automatique et l'intelligence artificielle, tels que l'apprentissage des modèles génératifs (GAN), l'apprentissage par transfert, l'adaptation de domaine et la réduction de dimensionnalité. Ces dernières années, l'OT et la distance de Wasserstein correspondante se sont révélés être un outil fondamental pour l'apprentissage automatique, en raison de sa capacité à fournir des gradients significatifs sur les distributions empiriques (nécessaires pour la formation des modèles ML) et la possibilité de coder des informations géométriques dans la distance (qui peut aider à l'interprétabilité).

L'objectif principal de ce poste de doctorant est d'étudier l'utilisation de la distance de transport relativement nouvelle de Gromov-Wasserstein (GW) [Peyré et al, 2016] pour mesurer la similitude entre des graphes labellisés. Les résultats préliminaires dans cette direction ont proposé une variante de la distance GW, la distance Fused Gromov-Wasserstein (FGW), qui peut également encoder les caractéristiques des nœuds de ces graphes et il a été démontré que c'est une très bonne mesure de similarité pour la classification des graphes [Vayer et al, 2019]. Certaines premières études suggèrent

également son utilisation pour l'exploration de graphes (à la fois pour le clustering d'échantillons de graphes et pour le clustering de nœuds), mais doivent encore être comparées et positionnées par rapport aux meilleures approches de clustering de graphes.

Le doctorant étudiera l'utilisation de FGW pour le clustering de graphes lorsque des connaissances préalables sont à disposition sur le type de graphe. Par exemple, la structure du graphe est encodée dans FGW comme une matrice de distance, et on pourrait imposer que la distance soit basée sur un nombre entier, ce qui fournira un graphe bien défini. La matrice de distance pourrait également être modélisée comme une matrice de faible rang qui peut plutôt bien coder une structure arborescente. La question de savoir comment mieux encoder la structure du graphe dans la matrice de «distance» de FGW est également encore ouverte et devra être étudiée. Si les avancées liées à l'usage de FGW dans un cadre d'apprentissage non-supervisé sont concluantes, le doctorant explorera l'intégration de FGW dans des architectures type réseaux de neurones, qui à aujourd'hui mènent aux meilleures performances dans un contexte d'apprentissage supervisé.

La méthodologie qui en résultera sera enfin appliquée à l'analyse de graphes du monde réel, avec des données éventuellement hétérogènes sur les arêtes, provenant de réseaux sociaux et de communication, ou de données biologiques telles que des molécules, des enzymes et des protéines.

Développement d'une approche de médecine personnalisée pour améliorer le diagnostic des maladies mitochondriales

Sujet de thèse de Justine Labory, encadrée par Silvia Bottini (MSI).

Contexte

Les maladies mitochondriales (MM) sont des pathologies rares provoquées par une déficience de la chaîne respiratoire (CR) mitochondriale, qui fournit l'énergie aux cellules par phosphorylation oxydative (1). Ce sont les plus fréquentes des maladies métaboliques avec une incidence estimée à 1/5000 naissances. Les MM sont extrêmement hétérogènes, tant sur le plan clinique que génétique, avec un large éventail d'âges et des symptômes très différents, ce qui rend leur diagnostic très difficile (2).

Les MM sont dues à des variants dans des gènes impliqués dans la biogénèse mitochondriale (3). Bien que les mitochondries possèdent leur propre génome (4), plus de 1000 gènes nucléaires sont nécessaires au fonctionnement de la CR (5). Par conséquent, les MM peuvent être causées par des variants pathogènes affectant soit l'ADNmt, soit les gènes nucléaires. Leur diagnostic nécessite l'identification du gène responsable.

Aujourd'hui, plus de 300 gènes nucléaires sont impliqués dans les MM et la liste ne cesse de s'allonger (6). Malgré le développement du séquençage à haut débit (NGS), plus d'un patient sur deux est en impasse diagnostique, le gène responsable restant inconnu. L'arrivée du séquençage de l'exome (WES) et du séquençage du génome (WGS) a considérablement accéléré l'identification de variants sur des gènes de maladies jusqu'alors inconnus (7). Néanmoins, leur taux de réussite pour la détection des variants responsables se situent seulement entre 25 et 50 %. Plusieurs variants détectés par le WES restent des variants de signification inconnue (VUS) ou sont omis en raison de l'impossibilité de les classer par ordre de priorité. En outre, les variants présents dans des régions non codantes ne peuvent pas être identifiés par WES. Le WGS détecte les variants dans les régions codantes et non codantes, mais il présente un défi important (et actuellement non résolu) d'interprétation des données pour les plus de 3 millions de variants par échantillon (8). Des études récentes ont montré l'intérêt du séquençage de l'ARN (RNAseq), pour sélectionner des gènes candidats présentant une expression altérée ou monoallélique et pour identifier des variants introniques profonds affectant l'épissage, lorsqu'il est utilisé en combinaison avec le WES (9). Pour les MM, l'utilisation du RNAseq chez 48 patients non diagnostiqués après un WES a permis d'augmenter le taux de réussite de 10% (10). Dans ce contexte, le développement de nouvelles approches bio-informatiques est essentiel pour résoudre l'impasse diagnostique et améliorer notre connaissance des maladies mitochondriales (11).

Objectifs :

1. Mise en place d'un protocole standardisé pour analyser les données de séquençage (WES et RNAseq) des patients avec suspicion de MM
2. Développement d'une nouvelle stratégie bio-informatique pour stratifier les variants identifiés par l'analyse des données de séquençage
3. Étude fonctionnelle des variants prioritaires pour identifier le gène responsable candidat des MM
4. Validation de l'outil nouvellement développé sur des données publiques et sur une cohorte de patients réalisée en interne.

Résultats attendus

L'identification des gènes responsables de MM permettra de proposer un conseil génétique, un diagnostic prénatal et d'envisager de nouvelles approches thérapeutiques. Avec ce projet, nous espérons développer de nouveaux outils en élargissant la gamme des variants qui peuvent être identifiés par RNAseq, en particulier par une analyse approfondie des transcrits non codants et des événements d'épissage qui peuvent affecter la structure et la fonctionnalité des gènes. Grâce au Centre de Référence sur les Maladies Mitochondriales, nous disposons d'une cohorte de 20 patients en impasse diagnostique (80 patients supplémentaires sont en cours d'inclusion). Le RNAseq a été réalisé après l'échec du WES pour identifier le gène responsable. Des échantillons provenant d'individus sains de la base de données GTEx (Genotype Tissue Expression) seront sélectionnés pour identifier les transcrits spécifiques des tissus sains et ceux spécifiques de la pathologie. Nous utiliserons une approche de machine learning pour développer une nouvelle méthode de stratification des variants afin d'identifier des gènes candidats potentiels à tester ensuite expérimentalement. Le défi consistera à identifier la meilleure stratégie pour traiter un grand nombre de variables et peu d'observations (patients), car généralement ce type d'algorithme fonctionne mieux avec la configuration de données inverse (peu de variables et beaucoup d'observations). Comme l'expression des gènes dépend des tissus, les données de séquençage d'individus sains provenant de multiples tissus (fibroblastes, muscles, sang) seront incluses dans l'analyse afin de distinguer l'hétérogénéité des tissus. L'analyse fonctionnelle des variants identifiés permettra de hiérarchiser les gènes candidats. Pour caractériser les variants de manière fonctionnelle, nous devons intégrer des données provenant de différentes sources : des données de séquençage aux informations cliniques en passant par les voies métaboliques. Nous utiliserons les réseaux de co-expression afin d'identifier les modules de gènes fonctionnels co-exprimés associés aux MM. Cette approche permettra de relier les variants génétiques trouvés en WES aux

voies métaboliques identifiées par les données RNAseq, extraites par exemple des bases de données de la Gene Ontology, RegNetwork et KEGG. Nous exploiterons les données des bases de données publiques pour étalonner et valider nos approches, en particulier pour entraîner nos algorithmes. Nous utiliserons notamment des données du consortium Geuvadis (12), fournissant des données RNAseq pour 460 échantillons pour lesquels le WES est disponible dans le cadre du projet « 1000 Genome ». Enfin, la validation fonctionnelle des variants identifiés se fera sur des modèles cellulaires et murins en lien avec l'équipe « Génétique des MM » à l'IRCAN.

Nous espérons que ce projet contribuera à la mise en place de nouveaux outils de diagnostic dans le cadre de la médecine personnalisée afin de réduire le nombre de patients en impasse diagnostique. Cette étude sera une étape importante dans le transfert de l'utilisation conjointe du WES et du RNAseq dans le domaine de la recherche au diagnostic. Enfin, le développement d'une nouvelle stratégie pour prioriser les variants permettra de relever l'un des plus grands défis dans le domaine des maladies mendéliennes pour améliorer la prise en charge des patients.

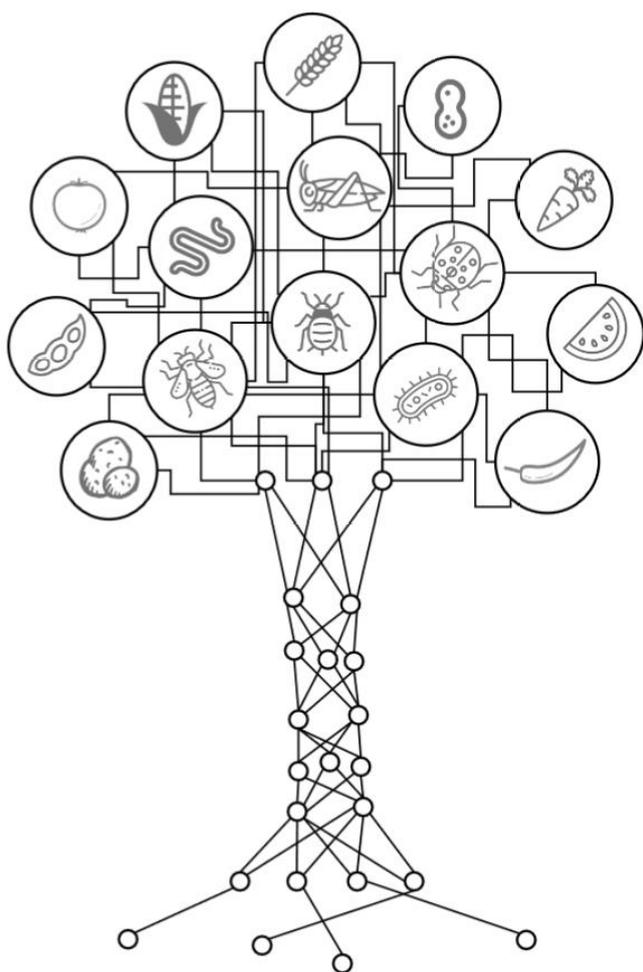
Valeurs ajoutées

Ce projet transdisciplinaire sera développé avec le "Centre de Référence sur les Maladies Mitochondriales" dont le Professeur Paquis, le superviseur de ce projet de doctorat, est le coordinateur, en collaboration avec le « Medical Data Laboratory (MDLab) ». Le MDLab, dont le Dr. Bottini, co-superviseur de ce projet, est le directeur, est un centre d'excellence en « data sciences » appartenant à la Maison de la Modélisation, de la Simulation et des Interactions (Université Côte d'Azur, UCA). Le MDLab a été créée pour renforcer le dialogue entre les chercheurs en « data sciences » et ceux du domaine médical. Enfin l'équipe « Génétique des Maladies Mitochondriales » à l'IRCAN assurera la validation fonctionnelle *in vivo* grâce à ses compétences sur l'étude des fonctions mitochondriales et aux modèles cellulaires et murins disponibles.

Références

1. Elliott H, Samuels D, Eden J, Relton C, Chinnery P. Pathogenic Mitochondrial DNA Mutations Are Common in the General Population. 2008;
2. Munnich A, Rustin P. Clinical spectrum and diagnosis of mitochondrial disorders. 2001;
3. Saneto R. Genetics of Mitochondrial Disease. In 2017.
4. Robin DE, Wong R. Mitochondrial DNA molecules and virtual number of mitochondria per cell in mammalian cells. 1988;
5. Rath S, Sharma R, Gupta R, Ast T, Chan C, Durham TJ, et al. MitoCarta3.0: an updated mitochondrial proteome now with sub-organelle localization and pathway annotations. *Nucleic Acids Research*. 8 janv 2021;49(D1):D1541-7.
6. Craven L, Alston C, Taylor R, Turnbull D. Recent Advances in Mitochondrial Disease. 2017;
7. Wortmann S, Koolen D, Smeitink J, van den Heuvel L, Rodenburg R. Whole exome sequencing of suspected mitochondrial patients in clinical practice. 2015;
8. Lohmann K, Klein C. Next Generation Sequencing and the Future of Genetic Diagnosis. *Neurotherapeutics*. 1 oct 2014;11(4):699-707.

-
9. Rahman J, Rahman S. Mitochondrial Medicine in the Omics Era. 2018;
 10. Kremer L, Bader D, Mertes C, Kopajtich R, Pichler G, Iuso A, et al. Genetic diagnosis of Mendelian disorders via RNA sequencing. 2017;
 11. Labory J, Fierville M, Ait-El-Mkadem S, Bannwarth S, Paquis-Flucklinger V, Bottini S. Multi-omics approaches to improve mitochondrial disease diagnosis: challenges, advances and perspectives. *Frontiers*. 2020;
 12. Lappalainen T, Sammeth M, Friedländer MR, 't Hoen PAC, Monlong J, Rivas MA, et al. Transcriptome and genome sequencing uncovers functional variation in humans. *Nature*. 1 sept 2013;501(7468):506-11.



IA et multi-omiques en santé des plantes

Résultats obtenus grâce aux moyens informatiques attribués :

Djampa Kozlowski. Nerd : a tool to predict proteins involved in nematode plant parasitism – Présenté aux conférences :

- SophIA Summit 2022, le 23/11/2022.
- Journée Annotation de génomes, Intelligence Artificielle et Tex-Mining' organisée par le PEPI IBIS (communauté d'ingénieurs INRAe) le 28/11/2022.

Description :

Produire des ressources agricoles suffisantes pour alimenter une population mondiale estimée à plus de 9 milliards d'individus d'ici 2050, tout en réduisant

l'impact sur notre planète est l'un des principaux défis auxquels l'humanité est confrontée.

Les plantes sont engagées dans divers types d'interactions avec les organismes présents dans leur environnement. Ces interactions peuvent être soit bénéfiques pour leur développement (par exemple, la symbiose avec les champignons et les bactéries), soit préjudiciables pour leur santé et leur survie (par exemple, les ravageurs, les agents pathogènes et les parasites). Les ravageurs agricoles sont eux-mêmes impliqués dans des interactions avec des prédateurs, des pathogènes et des parasites qui peuvent être utilisés comme agents de bio-contrôle (par exemple, les guêpes parasitoïdes).

L'approfondissement de nos connaissances sur le fonctionnement et le succès de ces différentes interactions biotiques est ainsi d'une importance capitale pour améliorer la santé des plantes et éviter une crise agricole à venir, tout en limitant notre impact environnemental.

Dans ce contexte Université Cote D'Azur (UCA) via la MSI et l'Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement (INRAE) ont développé un projet de recherche commun visant à étudier des interactions entre organismes en lien avec la santé des plantes grâce à l'intégration de données multi-omiques et à l'utilisation d'algorithmes d'Intelligence Artificielle (IA).

Les principales thématiques abordées concernent l'étude :

- Des interactions entre les plantes et divers bio-agresseurs (nématodes, oomycètes, insectes).
- Des relations symbiotiques bénéfiques à la plante telles que celles entretenues avec différentes bactéries et champignons.
- Des interactions entre insectes parasitoïdes et bio-agresseurs des plantes dans le cadre de leur bio-contrôle.

Pour y parvenir, divers pipelines d'analyse de données biologiques devront être mis en place, ce qui nécessitera l'accès à des ressources informatiques conséquentes telles que celles proposées par le cluster Azzurra. La réalisation de certaines des études précédemment citées nécessitera en effet d'avoir accès :

i) à une capacité de stockage importante compte tenu de la quantité de données à traiter

ii) à une grande puissance de calcul via l'utilisation de GPU afin d'entraîner et utiliser des modèles d'intelligence artificielle.

Projet DeepWILD



© Parc national du Mercantour

Projet MSI et 3IA Côte d'Azur en partenariat avec le Parc national du Mercantour.

Résultats obtenus grâce aux moyens informatiques attribués :

Fanny Simões, Charles Bouveyron, Frédéric Precioso. DeepWILD: Wildlife Identification, Localisation and estimation on camera trap videos using Deep learning. 2022. [hal-03797530](https://hal.archives-ouvertes.fr/hal-03797530) (pré-publication)

L'objectif du projet est de localiser, classifier et compter les espèces du Parc national du Mercantour à l'aide d'un algorithme de deep learning. Les images utilisées sont issues de vidéos extraites de pièges photographiques installés dans la Vallée de la Roya. L'ensemble des vidéos et des images utilisées pour ce projet sont stockées sur Azzurra. Azzurra a permis le traitement de ce gros volume de données et l'application d'un modèle de détection d'objet (Faster R-CNN avec Inception-ResNet-v2) capable de reconnaître les différentes espèces présentes dans le Parc national du Mercantour.

L'équipe travaillant sur ce projet est composée de Charles Bouveyron, Frédéric Precioso et Fanny Simões pour UCA, et Nathalie Siefert et Stéphane Combeaud pour le Parc national du Mercantour.

Projet Tara



© Megan Clampitt

Projet de l'IRCAN dont le but est d'évaluer si une approche IA permet d'estimer la santé des colonies coralliennes à l'aide d'un dispositif automatisé et facilement applicable.

Résultats obtenus grâce aux moyens informatiques attribués :

Minh Nhat Do, Marco Milanesio. Image Classification on Corals using ensemble deep Transfer-learning approach. Rapport de stage Master 2 Data science & Artificial intelligence (2022).

On utilise en premier lieu des images sous-marines collectées au cours de l'expédition Tara Pacific. Cette expédition a produit ~14000 photos (correspondant à 2216 colonies individuelles) qui étaient annotées manuellement. Ces annotations comprennent plusieurs facteurs : l'identification taxonomique jusqu'au rang du genre, la prédation, le contact avec d'éventuelles algues, la présence d'organismes perforants, le contact avec les sédiments, ainsi que d'autres facteurs directement liés à la santé des coraux (i.e. coloration,

blanchissement, maladies, pigmentation). Grâce aux études de l'expédition Tara, plusieurs autres bases de données sont pareillement disponibles, elles contiennent notamment des facteurs biologiques (biomarqueurs et télomères) et environnementaux.

Nous avons débuté une approche IA pour extraire les catégories d'annotations qui avaient été faites manuellement. Plus spécifiquement, nous utilisons la méthode dite du « Convolutional neural network (CNN) ». La première étape de cette approche est d'identifier le genre auquel appartient les colonies photographiées. Avec notre réseau qui comprenait les photographies et les annotations manuelles (qui ne visaient pas à identifier le genre), nous avons réussi à identifier les genres avec une exactitude de 66.28%.

L'étape en cours est d'augmenter l'exactitude de cette prédiction avec des approches IA comme le « transfer learning » ou les modèles génératifs. De plus, nous prévoyons de compléter ce réseau avec davantage de variables qui étaient annotées manuellement (blanchissement, présence de maladies, prédation, etc.) et nous espérons aussi pouvoir utiliser l'algorithme pour prédire les facteurs biologiques (télomères et biomarqueurs) qui sont considérablement plus difficile à obtenir. Mais il nous faudra plus d'échantillons pour construire l'algorithme.

Chimie

EUR Spectrum

Simulations de complexes biomoléculaires impliqués dans les sens chimiques (olfaction, gustation).

Responsable scientifique du projet : Sébastien Fiorucci

Laboratoire : ICN

Résultats obtenus grâce aux moyens informatiques attribués :

Journaux :

- Meinert, C., Garcia, A.D., Topin, J. et al. Amino acid gas phase circular dichroism and implications for the origin of biomolecular asymmetry. Nat Commun 13, 502 (2022).

<https://doi.org/10.1038/s41467-022-28184-0>

- Shim, T., Pacalon, J., Kim, W. C., Cong, X., Topin, J., Golebiowski, J., & Moon, C. (2022). The Third Extracellular Loop of Mammalian Odorant Receptors Is Involved in Ligand Binding. International Journal of Molecular Sciences, 23(20), 12501.

- Garcia, A. D., Topin, J., Bocková, J., Jones, N. C., Hoffmann, S. V., & Meinert, C. (2022). Chiroptical activity of gas-phase propylene oxide predicting the handedness of interstellar circular polarization in the presolar nebula. Science Advances, 8(46), eadd4614.

Conférences :

- Odorant binding and receptor activation deciphered at the molecular level. M. Hladiš, J. Pacalon, M. Lalis, S. Fiorucci, J. Topin. ECRO 2022, Berlin (Allemagne), 31-août – 4 septembre 2022.

- Predicting odorant-receptor activation with protein language and graph neural networks. M. Hladiš, M. Lalis, S. Fiorucci, J. Topin. SophIA summit 2022, Sophia Antipolis (France), 23-25 novembre 2022.

- J. Topin, N. C. Jones, S. V. Hoffmann, L. Briand, C. Meinert, Tracking ligand-binding effect on protein stability by CD spectroscopy. 18th International Conference on Chiroptical Spectroscopy (CD 2022), New-York (USA), July 2022.

- J. Topin, A. Garcia, J. Bocková, N. C. Jones, S. V. Hoffmann, U. J. Meierhenrich, C. Meinert, CD spectroscopy and ab initio calculations revealed hidden conformation of amino-acids. Chirality, Chicago (USA), July 2022.

Thèse :

Modélisation moléculaire des récepteurs chimiosensoriels par Jody Pacalon. Soutenue en novembre 2022 <https://theses.fr/s317702>

Posters :

- Predicting the combinatorial code via graph neural networks and representation learning M. Hladiš, S. Fiorucci, J. Topin, AChemS XLIV, Bonita Springs (USA), 20-23 avril 2022.
- Integrative structural modeling reveals functional molecular switches of human G protein-coupled bitter-taste receptors J. Topin, C. Bouysset, Y. Kim, M. Rhyu, J. Golebiowski, S. Fiorucci. AChemS XLIV, Bonita Springs (USA), 20-23 avril 2022.
- Structure-based virtual screening of bitter taste receptors M. Lalis, C. Belloir, J. Topin, L. Briand, S. Fiorucci. ERNEST 6, webinaire, 28-31 mars 2022.
- Predicting the combinatorial code for any odorant molecule: a graph neural networks approach M. Hladiš, M. Lalis S. Fiorucci, J. Topin. ECRO 2022, Berlin (Allemagne), 31 août-3 septembre 2022.
- TAARs response to amines are largely affected by sequence variants J. Pacalon, C. Belloir, S. Fiorucci, L. Briand, J. Topin. ECRO 2022, Berlin (Allemagne), 31 août-3 septembre 2022.
- AI-guided reverse chemical ecology applied to pest control J. Pacalon, H. Hocine, V. Monetti, G. Caballero-Vidal, C. Bouysset, N. Montagné, E. Jacquin-Joly, S. Fiorucci. ECRO 2022, Berlin (Allemagne), 31 août-3 septembre 2022.
- L. Moitrier, C. Belloir, M. Lalis, Y. Hou, J. Topin, L. Briand. Deciphering the ligand binding properties of the mouse odorant-binding protein OBP5 from *Mus musculus*. European Chemoreception Research Organisation (ECRO), Berlin (Germany), September 2022.
- M. Lalis, M. Hladiš, Y. Hou, N. Baldovini, J. Topin. ChemOddity: Machine Learning to predict odorant detection threshold. European Chemoreception Research Organisation (ECRO), Berlin (Germany), September 2022
- J. Topin, J. Pacalon, G. Audic, J. Golebiowski, C. Moreau. Identification of entry pathway and binding site of insect odorant co-receptor. Association of Chemoreception Sciences, Bonita spring (USA), April 2022.
- M. Hladiš, S. Fiorucci, J. Topin. Predicting the combinatorial code of olfaction via graph neural networks and representation learning. Association of Chemoreception Sciences, Bonita spring (USA), April 2022.
- S. Fiorucci, C. Bouysset, Y. Kim, M. Rhyu, J. Golebiowski, J. Topin. Integrative structural modelling reveals functional molecular switches of human G protein-coupled bitter-taste receptors. Association of Chemoreception Sciences, Bonita spring (USA), April 2022.
- M. Hladiš, J. Pacalon, M. Lalys, C. Bouysset, S. Fiorucci, J. Topin. Deciphering Chemical Senses at the Molecular Scale. Gordon Research Conference on Ligand Recognition and Molecular Gating, Lucca (Italy), March 2022.

Description du projet :

Nous cherchons à décrypter les bases moléculaires de la perception chimiosensorielle (i.e. les sens chimiques : olfaction et gustation) qui est par nature un processus neuronal

complexe. Pour tenter de déchiffrer le code combinatoire de la perception chimiosensorielle nous mettons au point des modèles numériques basés sur des approches de modélisation moléculaire, de machine learning, ou encore de bioinformatique structurale. Pour décrire les mécanismes moléculaires sous-jacents, nous tirons profit de l'expertise de nos collaborateurs, notamment en biologie moléculaire et neurobiologie (Figure 1).

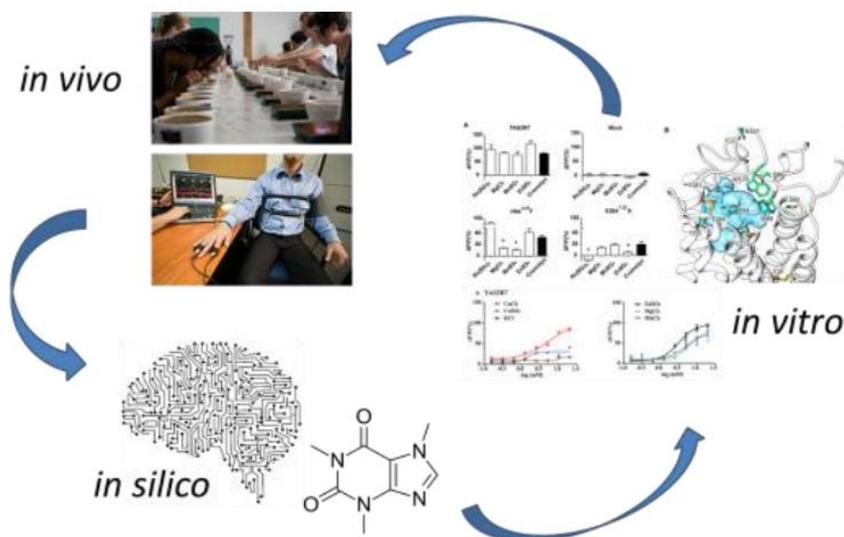


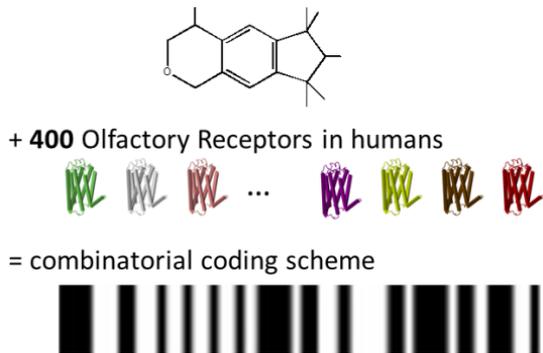
Figure 1 Synergie de l'approche pluridisciplinaire

1. Décrypter le codage de l'information chimique par notre système sensoriel.

Notre cerveau est fait de milliards de neurones et est considéré comme l'organe le plus complexe de notre corps. Parmi nos cinq sens, l'odorat et le goût sont les deux sens chimiques permettant aux êtres vivants d'identifier et de réagir aux substances présentes dans leur environnement. L'extraordinaire capacité de discrimination de notre cerveau provient de l'organisation de notre système de détection. Nous possédons environ 400 récepteurs olfactifs et 30 récepteurs gustatifs. Sachant qu'une molécule (odorante ou sapide) peut activer plusieurs récepteurs sensoriels, qu'un récepteur peut répondre à plusieurs molécules différentes et que la perception d'une odeur ou d'une saveur peut provenir d'une seule molécule ou d'un mélange de molécules, la combinatoire est virtuellement infinie¹ (Figure 2). L'homme peut par exemple distinguer plus de mille

¹ Topin et al. Submitted. <https://doi.org/10.1101/2020.10.23.348706>

Figure 2 Codage combinatoire de la perception des odeurs



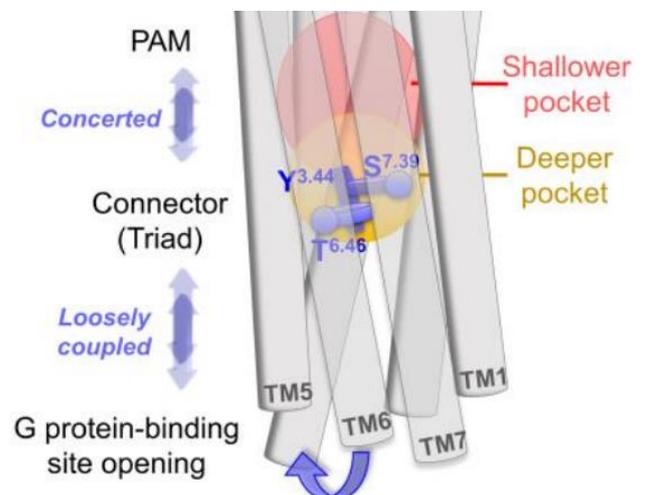
milliards de molécules odorantes. L'espace chimique des molécules pouvant activer notre odorat ou notre goût est immense et l'un des objectifs de nos travaux est d'identifier les liens chémo-génomiques de ces espaces de grande dimensionalité et de mettre au point des modèles moléculaires et numériques capables de prédire de façon rationnelle des composés olfactifs et sapides nouveaux.

Faits marquants : Les travaux de l'équipe ont notamment permis de mieux comprendre les relations entre la structure d'un composé et son activité biologique² et permis d'élargir le spectre de reconnaissance de récepteurs olfactifs humains³, de poissons⁴ ou encore d'insectes⁵ mais également de récepteurs gustatifs humains^{6,7}.

2. Développer de nouveaux composés pour moduler l'activation des récepteurs chimiosensoriels.

La détection des odeurs et des saveurs sucré, amère et umami nécessitent l'activation des récepteurs de la famille des Récepteurs Couplés aux Protéines G (RCPG) dont le mécanisme d'activation est similaire à celui de la Rhodopsine, récepteur prototypique des RCPG de classe A. Ils agissent comme des « pinces » moléculaires qui peuvent « s'ouvrir » lorsqu'un agoniste (odorant ou molécule sapide) s'associe

Figure 3 Mécanisme de modulation allostérique d'un RCPG (ici de classe C)



² Bushdid et al. Science, 2014, 343, 1370-1372

³ Licon et al. Plos Comput. Biol. 2019, 15(4), e1006945

⁴ Cong et al. J. Biol. Chem., 2019, 294, 6762-6771

⁵ Caballero-Vidal et al. Sci. Rep. 2020,10, 1655

⁶ Chéron et al. Food. Chem. 2017, 221, 1421-1425

⁷ Bouysset et al. Food. Chem. 2020, 324, 126864

avec eux, ou « se refermer » s'il s'agit d'un agoniste inverse (Figure 3). La nature versatile de ces récepteurs (subtil équilibre entre différents états conformationnels dont la forme active même en l'absence de ligand, i.e. état basal) rend leur pharmacologie complexe. En l'absence de structure cristallographique de récepteur chimiosensoriel, le développement rationnel de nouveaux composés olfactifs et gustatifs reste un défi. L'un des objectifs de nos travaux est de tirer parti des outils de simulation moléculaire (validés par des études de mutagénèse dirigée) pour étudier la dynamique des interactions ligands-récepteurs et développer de nouveaux modulateurs du goût et de l'odorat. Faits marquants : Les travaux de l'équipe ont notamment montré qu'il était possible de prédire le mécanisme d'activation et le spectre de reconnaissance de récepteurs chimiosensoriels à l'aide de modèles moléculaires et numériques.

On peut mettre en avant les résultats sur les récepteurs olfactifs de mammifères^{8 9 10 11} mais également sur la famille des récepteurs gustatifs^{12 13} et plus largement sur les RCPG^{14 15 16}, famille de récepteurs impliqués dans la signalisation cellulaire et naturellement une des premières cibles pharmacologiques pour de nombreuses pathologies.

3. Collaborations et importance pour le tissu socio-économique

Le projet s'appuie sur une approche pluridisciplinaire chimie-informatique-neurobiologie. Il permettra d'envisager la mise au point d'un « nez ou d'une langue computationnelle » physiologiquement inspiré(e) capable de prédire l'odeur, la saveur et/ou la réponse émotionnelle d'un composé sur la base de sa structure moléculaire. Pour l'industrie des arômes et parfums, la recherche de nouvelles molécules est un enjeu majeur. Par analogie avec d'autres molécules connues, il serait possible de prédire les structures de nouvelles molécules ayant des effets de modulation de l'odorat ou du goût (inhibition notamment).

⁸ Yu et al. PNAS, 2015, 112(48), 14966-1497

⁹ de March et al. JACS, 2015, 137(26), 8611-8616

¹⁰ de March et al. Angew. Chem. Int. Ed. 2018, 57, 4554-4558

¹¹ Bushdid et al. CMLS, 2019, 76(5), 995-1004

¹² Chéron et al. Chem. Senses, 2019, 44, 303-310

¹³ Wang et al. Chem. Senses, 2019, 44, 339-347

¹⁴ Cong et al. J. Chem. Theory Comput., 2018, 14(8), 4467-4473

¹⁵ Cong et al. J. Chem. Info. Model. 2019 59(6), 2871-2878

¹⁶ Cong et al. PCCP, 2018, 20, 24915-24920

De plus, les récepteurs chimiosensoriels sont également exprimés dans d'autres organes que ceux de nos sens du goût et de l'odorat (expression ectopique) et leur rôle sur notre santé est encore mal connu. On peut donc envisager des applications de notre projet dans de grands secteurs économiques aussi divers que la santé, l'industrie des arômes et parfums ou l'industrie agroalimentaire, particulièrement bien représentés dans la région PACA. Notre travail de recherche mobilise de nombreux partenaires au niveau local avec des experts du domaine de la santé (CHU Nice, IPMC) et de l'intelligence artificielle (thèse 3IA qui démarrera en octobre 2020 en collaboration avec l'IS), au niveau national avec des experts des sens chimiques chez l'homme (CRNL-Lyon, CSGA-Dijon) ou chez l'insecte (INRAE Versailles) ainsi qu'au niveau international avec des experts en neurosciences à Duke (USA), au Monell (USA) ou encore au DGIST (Corée). A noter également que le projet entre dans les thématiques prioritaires de l'Université Côte d'Azur liées à la science des odorants à travers l'Institut d'Innovation et de Partenariat en Arômes, Parfums & Cosmétiques et aux neurosciences computationnelles et aux sciences de la cognition notamment soutenues à travers l'institut Neuromod ou encore la structure CoCoLab.

4. Méthodologies employées

D'un point de vue computationnel, nos travaux s'appuient sur des outils de modélisation moléculaire et d'apprentissage statistique (machine / deep learning). La modélisation d'un récepteur chimiosensoriel dans un environnement physiologiquement inspiré (prise en compte explicite de la membrane, du solvant ou encore des ions) nécessite des simulations moléculaires d'environ 60 000 à 100 000 atomes et pouvant aller jusqu'à 400 000 atomes pour des systèmes multimériques (Figure 4). Nous étudions les changements de conformations qui ont lieu lors des processus de liaison d'un odorant ou d'une molécule sapide avec un récepteur. Cela nécessite des simulations de dynamique moléculaire de plusieurs dizaines de microsecondes. Pour cela nous utilisons les logiciels Amber et Gromacs, par ailleurs déjà installées sur l'ancien supercalculateur de l'Université (CICADA), que ce soit pour des calculs sur CPU ou GPU. Les analyses des trajectoires de simulations se font souvent sur les machines de l'équipe et il peut arriver que ponctuellement nous ayons besoin de réaliser une analyse (calculs de mode normaux par ex.) qui nécessite une ressource mémoire plus importante. Dans ce cas, Azzurra nous permettra de réaliser ces calculs. Par ailleurs, les structures de ces récepteurs et les

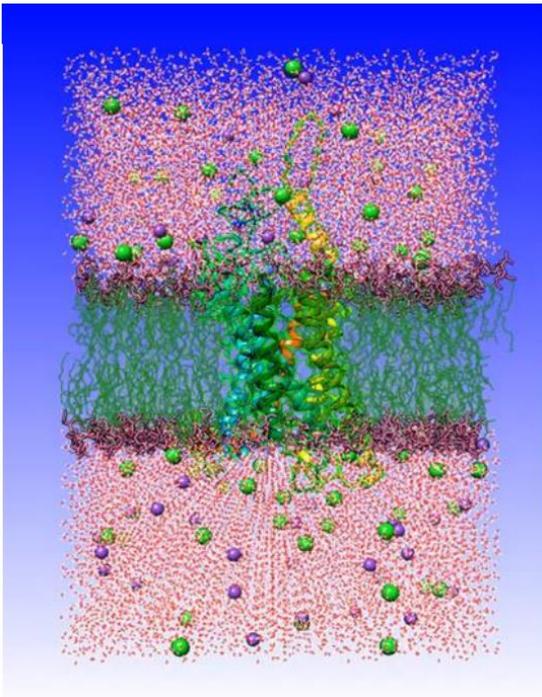


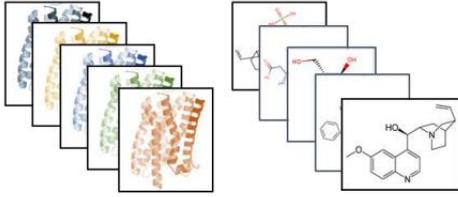
Figure 4 Boîte de simulation d'un RCPG dans environnement physiologique.

interactions ligand-récepteurs sont déduites de contraintes expérimentales et théoriques par des approches de reconstruction par homologie et de docking.

Pour cela nous utilisons les logiciels Modeller et Autodock, également installés sur CICADA et parfaitement adaptés pour du calcul parallèle et distribué sur CPU. Ponctuellement nous pouvons utiliser des logiciels de chimie quantique (Gaussian) pour le calcul de propriétés électroniques des systèmes étudiés. L'identification et la prédiction des liens chemo-génomiques et des relations structure-activité nécessite d'intégrer de grandes bases de données sur la structure des molécules odorantes ou sapides et sur la structure des récepteurs

chimiosensoriels. Les descripteurs moléculaires permettront d'encoder les structures chimiques et serviront de millions d'entrées aux modèles de machine learning (SVM, RF, ...) ou de deep learning qui seront ensuite validés par des expériences in vitro chez nos collaborateurs (Figure 5). Ces nouveaux résultats alimenteront in fine les bases de données initiales et permettront selon un principe de cercle vertueux d'améliorer les modèles in silico. L'ensemble de ces méthodes sont notamment implémentées dans différentes bibliothèques Python (Scikit-learn, Pytorch) optimisées pour le calcul CPU et GPU ou inclus dans les extensions du logiciel R.

1. Database (DB) of chemosensory receptors and odorants/tastants



2. Encoding molecular structures into numerical descriptors

0.42	0.18	0.37	0.25	0.35	0.58	0.35	0.48	
0.17	0.29	0.63	0.30	0.70	0.26	0.39	0.08	
0.80	0.08	0.18	0.23	0.25	0.68	0.03	0.71	
0.29	0.53	0.51	0.49	0.61	0.64	0.00	0.13	
0.13	0.15	0.50	0.13	0.61	0.28	0.33	0.08	
0.15	0.34	0.75	0.15	0.99	0.00	0.33	0.08	
0.07	0.17	0.17	0.48	0.11	0.86	0.03	0.21	
0.26	0.22	0.50	0.34	0.49	0.43	0.36	0.17	
0.57	0.12	0.30	0.23	0.36	0.55	0.14	0.29	
0.20	0.15	0.16	0.44	0.12	0.85	0.00	0.33	

5. Feeding DB

Computational modeling of olfaction & taste

4. Prediction & *in vitro* validation

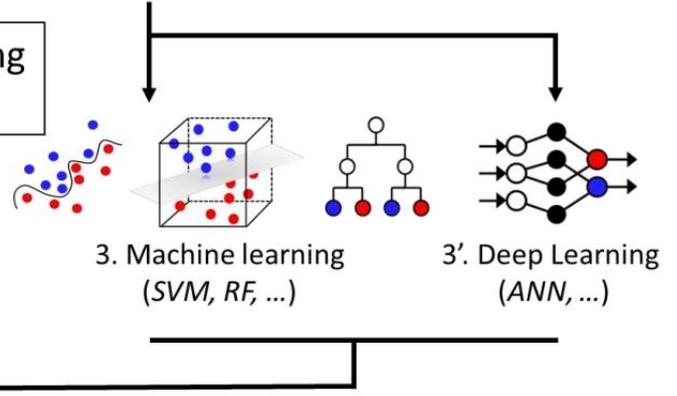
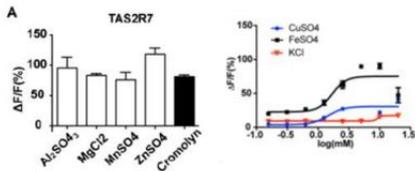


Figure 5 Différentes étapes de la construction du modèle numérique prédictif basé sur des approches d'apprentissage automatique.

Modélisation de molécules d'intérêt biologique, catalytique et éco-compatible

Responsable scientifique : Fabien Fontaine-Vive

Laboratoire : ICN

Résultats obtenus grâce aux moyens informatiques attribués :

Durant l'année écoulée, les calculs effectués sur le cluster AZZURRA ont permis de développer de nouveaux inhibiteurs à la protéine Arf6 et ainsi de publier un article dans la revue *Molecules* (IF 4.9) :

"Chlortetracycline, a Novel Arf Inhibitor That Decreases the Arf6-Dependent Invasive Properties of Breast Cancer Cells", E. Macia, M. Vazquez-Rojas, A. Robiolo, R. Fayad, S. Abélanet, I. Mus-Veteau, F. Fontaine-Vive, M. Mehiri, F. Luton, M. Franco, *Molecules*, 2021, 26(4), 969

Description des activités :

Le Service de Modélisation Moléculaire est un service de la Plateforme Technologique de Chimie dont les missions premières sont d'appuyer les travaux de recherche de l'ICN en créant des modèles numériques corroborant les expériences et permettant ainsi de les interpréter et les prédire. J'ai la responsabilité (maintenance, mise en œuvre, conception de méthodes et de modèles numériques, disponibilité et formation) pour les équipes de recherche de l'ICN ainsi que pour le tissu industriel local et national. Les problématiques que l'on est à même d'étudier avec les matériels et logiciels du service peuvent se classer comme suit :

Déterminer les conformations d'énergie minimale, dynamique moléculaire (AMBER)

Calculs de propriétés moléculaires et spectroscopiques (UV, ECD, IR, RMN, neutrons)

Etude de la réactivité chimique par les méthodes de la chimie quantique.

Applications aux sciences des matériaux.

Depuis 10 ans, le service de modélisation moléculaire est spécialisé dans l'étude des propriétés catalytiques de complexes métalliques et de nanoparticules et la compréhension des mécanismes d'inhibition de molécules antivirales.

Les 3 types de calculs de modélisation moléculaire peuvent être présentés ci-après :

Calculs ab-initio	Calculs semi-empiriques	Calculs empiriques
Optimisation de structures, calculs d'énergie électroniques, spectres électroniques de dichroïsme circulaire et UV, spectres vibrationnels IR	Détermination des conformations d'énergie minimale, dynamique moléculaire	Détermination des conformations d'énergie minimale, dynamique moléculaire

Présentation scientifique du projet :

Pour le groupe « Produits Naturels Marins » de l'équipe « Molécules Bioactives », les calculs de TD-DFT avec GAUSSIAN permettent d'élucider la configuration absolue de molécules bioactives chirales grâce aux calculs de spectres électroniques de dichroïsme circulaire électronique. Au préalable, une recherche conformationnelle au moyen de méthodes de dynamique moléculaire empirique et semi-empirique avec AMBER est nécessaire afin d'identifier les conformations les plus probables des molécules.

La figure 1 illustre la recherche de la configuration absolue de molécules extraites d'éponges marines : les anchinopeptolides qui diffèrent par les groupements R et R'.

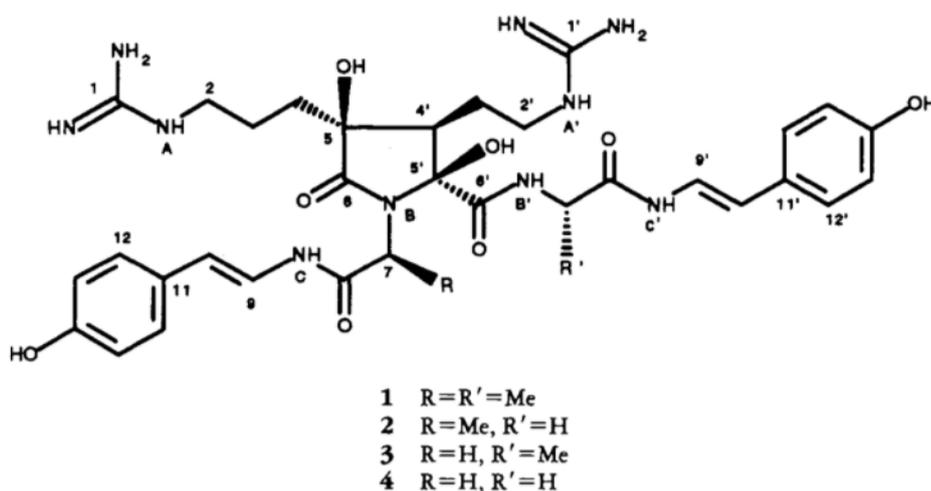


Fig. 1. Stéréochimie relative des molécules anchinopeptolides extraites d'éponges marines

Dix conformations pour chaque molécule conduisent à un calcul d'optimisation DFT puis à un calcul d'états excités TD-DFT, et la moyenne des spectres pondérée par le facteur de

Boltzmann a été affichée dans la figure 2 pour une molécule. La publication sur ces travaux est en cours de rédaction et la configuration absolue des trois autres molécules restent à identifier.

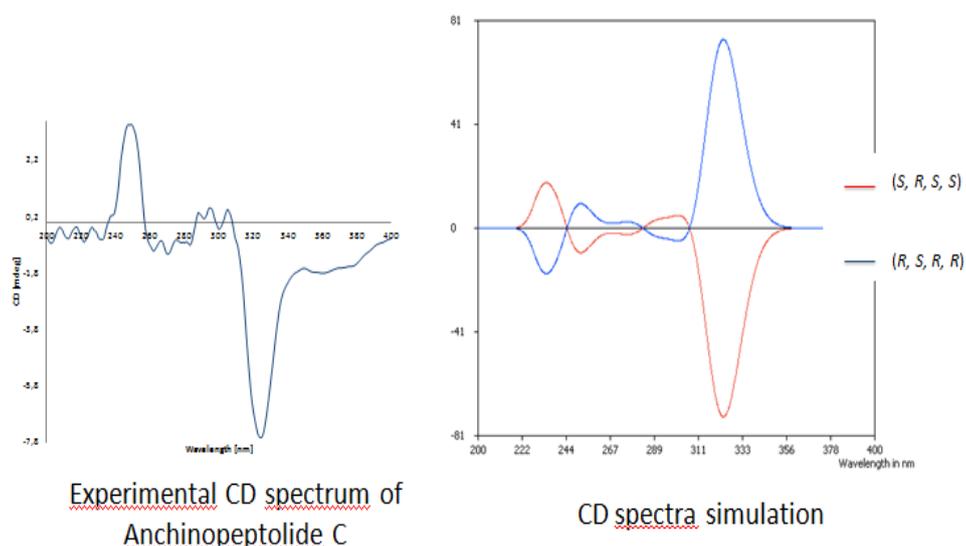


Fig. 2. Spectres ECD expérimental et calculé par TD-DFT de la molécule chirale anchinopeptolide C.

Des calculs similaires, conduits au sein du cluster CICADA ont d'ores et déjà permis de publier plusieurs articles de rang A référencés ci-dessous ^{17,18,19}.

Pour le groupe « Catalyse, Méthodologies de Synthèse et Chimie Fine » de l'équipe « Arômes Parfums Synthèse et Modélisation », les calculs DFT et TD-DFT permettront d'élucider la configuration absolue de molécules chirales et d'effectuer une étude mécanistique afin d'élucider le mécanisme réactionnel. La figure 3 illustre le projet en cours concernant la cyclisation catalysée par un complexe d'or. Le but du calcul sera de proposer une étude mécanistique permettant d'expliquer la catalyse asymétrique et déterminer si la réaction est cinétiquement ou thermodynamiquement limitante.

¹⁷ New bioactive chlorinated cyclopentene derivatives from the marine-derived Fungus *Phoma* sp., M. Elsebai, H.A. Ghabbour, N. Legrave, F. Fontaine-Vive, M. Mehiri, *Med Chem Res* 27, 1885–1892 (2018).

¹⁸ Pan-genotypic Hepatitis C Virus Inhibition by Natural Products Derived from the Wild Egyptian Artichoke, M. Elsebai, G. Koutsoudakise, V. Saludese, G. Pérez-Vilaró, A. Turpeinen, S. Mattila, A. M. Pirttilä, F. Fontaine-Vive, M. Mehiri, A. Meyerhans and J. Dieze, *J. of Virology*, 2016, vol. 90 no. 4 1918-1930.

¹⁹ New isocyanide antibiotics with broad spectrum against Gram-negative pathogens produced by cultures of an unidentified fungus, N El Aouad, V Gonzalez, M De la Cruz, L Lorenzo, JR Tormo, J Martin, I Perez-Victoria, C Diaz, F Vicente, F Fontaine-Vive, OP Thomas, O Genilloud, F Reyes, *Planta Med* 201480-PIL72.

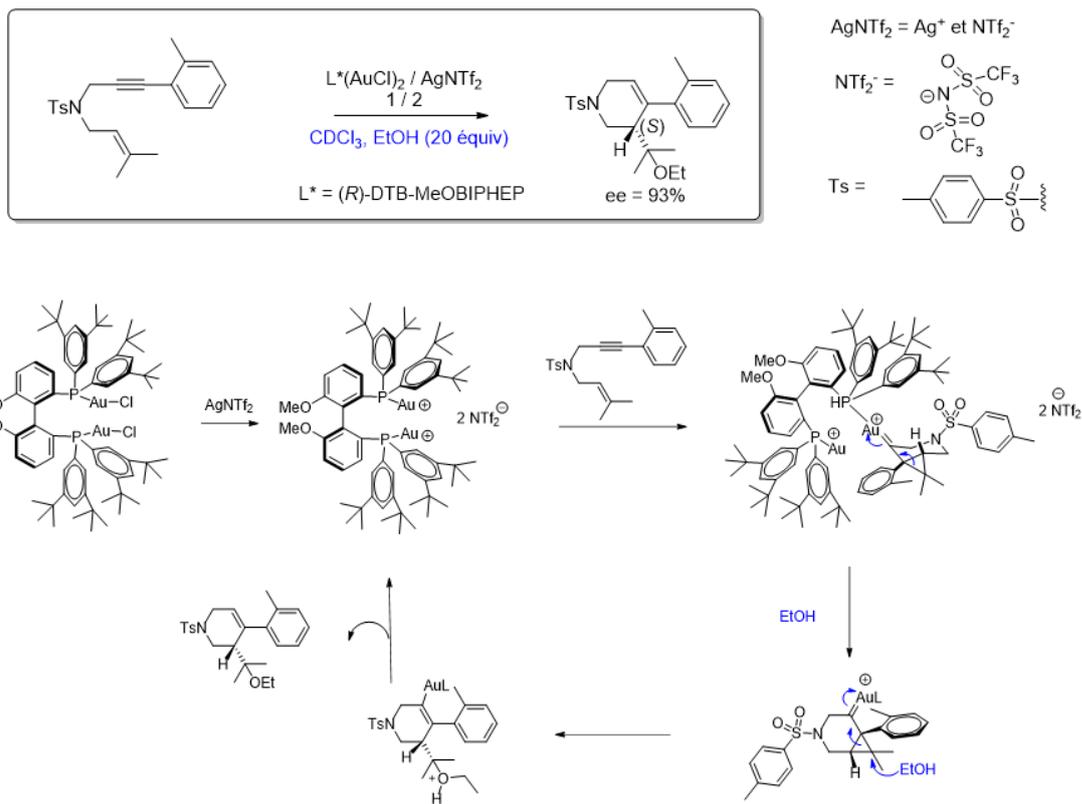


Fig. 3. Mécanisme de cyclisation asymétrique catalysée par des complexes organiques d'or

Des calculs similaires, conduits au sein du cluster CICADA ont d'ores et déjà permis de publier plusieurs articles de rang A référencés ci-dessous ^{20,21,22,23,24,25}.

Pour l'équipe « Matériaux et Polymères Eco-Compatibles », les calculs DFT dans l'état solide avec le logiciel VASP nous permettront d'interpréter les spectres vibrationnels et de prédire la structure de polymères biosourcés destinés à remplacer les dérivés de produits pétroliers. Les changements conformationnels induits par des étirements pourront ainsi

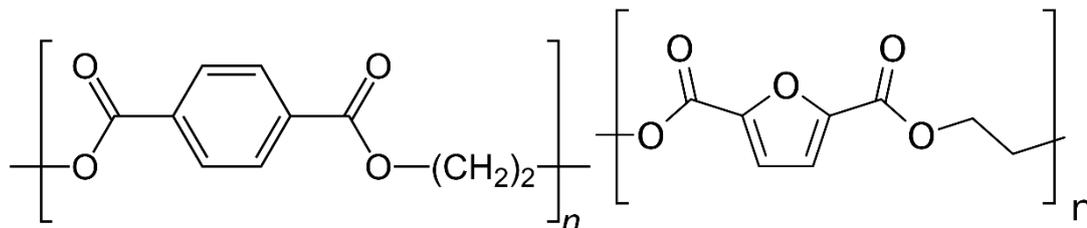


Fig. 4. Unités de répétition de PET (gauche) et PEF (droite)

être élucidés. Par exemple, le PET est un polymère thermoplastique pétrosourcé dont la structure chimique est donnée à la figure 4. Nous travaillons actuellement sur un projet qui vise à l'industrialisation de son « analogue » 100% biosourcé, le PEF. Les structures chimiques de ces 2 polymères sont proches mais diffèrent par la présence d'un noyau furanique dans le PEF à la place du noyau aromatique du PET. Il a été montré expérimentalement que certaines propriétés du PEF sont très différentes de celles du PET, ce qui s'explique par les différences de structures. Il a notamment été proposé que ces différences pourraient provenir du fait que le cycle benzénique peut tourner alors que le

²⁰ Gold-catalyzed cycloisomerization of 1,6-cyclohexenylalkyne: an efficient entry to bicyclo[3.2.1]oct-2-ene and bicyclo[3.3.1]nonadiene, Davenel V., Nisole C., Fontaine-Vive F., Fourquez J.-M., Chollet A.-M., Michelet V., *The Journal of Organic Chemistry*, 2020.

²¹ Silver-catalyzed intramolecular [4+2] cycloaddition reaction of amide-1,6-enynes, X. Chen, F. Fontaine-Vive, S. Martini, V. Michelet, *Catalysis Communications*, 2020, 106117.

²² In(OTf)₃-Catalysed Easy Access to Dihydropyranocoumarin and Dihydropyranochromone Derivatives, Boufroua N., Dunach E., Fontaine-Vive F., Achouche-Bouzroua S., Poulain-Martini S., *New Journal of Chemistry*, 2020.

²³ Synergies in the catalytic activity of bimetallic nanoparticles and new synthetic methods for the preparation of fine chemicals, I. Francesco, F. Fontaine-Vive, S. Antoniotti, *Chem. Cat. Chem.*, 2014, 6, 2784–2791.

²⁴ Novel radical tandem 1,6-enynes thioacylation / cyclisation : Au-Pd nanoparticles catalysis versus thermal activation as a function of the substrate specificity, I. Francesco, J. Giauffret, F. Fontaine-Vive, J. Edwards, G. Hutchings, S. Antoniotti, *Tetrahedron*, 2014, Volume 70, Issue 51, 9635-9643.

²⁵ Access to polycyclic derivatives by triflate-catalysed intramolecular hydroarylation, B. Cacciuttolo, S. Poulain-Martini, F. Fontaine-Vive, M. A. H. Abdo, H. El Kashef, E. Dunach, *EuroJOC* 2014, Issue 33, 7458-7468.

cycle furanique ne le peut pas. La présence de l'oxygène sur le cycle furanique est également un élément important (interactions dipolaires) pour expliquer les différences de propriétés. Nous souhaitons donc réaliser un travail plus fondamental, en modélisant les structures de ces composés et en faisant les prédictions des spectres IR, afin d'essayer d'expliquer de façon plus académique les différences de comportement observées expérimentalement.

Calcul ab initio de spectres de molécules réservoirs pour le carbone interstellaire et circumstellaire

Responsable scientifique : Patrick Cassam-Chenaï

Laboratoire : LJAD

Résultats obtenus grâce aux moyens informatiques attribués :

Article en cours de rédaction avec les nouveaux résultats obtenus qui doit mettre à jour le preprint hal-00785488v1 : <https://hal.univ-cotedazur.fr/hal-00785488>

Description :

Le projet se concentrera sur l'étude de 2 familles de molécules considérées comme piégeant une grande partie du carbone interstellaire et circumstellaire, et qui pour cela jouent un rôle clé dans l'astrochimie organique, et possiblement prébiotique. Il s'agit des hydrocarbures aromatiques polycycliques (HAPs) et des fullerènes.

Les HAPs forment une famille de composés organiques constitués de cycles aromatiques d'atomes de carbones dont les valences périphériques sont saturées par des atomes d'hydrogènes. Les HAPs (ou des espèces dérivées proches) sont considérés comme les porteurs des "Bandes Infra-rouges Aromatiques" (BIAs) à ~3.3, 6.7, 7.7, 8.6, 11.3 μm . Ces bandes caractéristiques sont parmi les plus fortes observées en infra-rouge (IR) dans le milieu interstellaire et présentent une certaine variabilité en fonction des sources observées. Ainsi, la spectroscopie IR des HAPs est d'importance capitale pour l'astrochimie et l'astrophysique: une meilleure compréhension de celle-ci est essentielle pour déduire des observations IR, des informations sur la composition des HAPs et les paramètres physiques de leur environnement [1].

Le fullerène le plus emblématique est le buckminsterfullerène, C₆₀, parfois également appelé footballène. Ses propriétés physiques uniques sont l'objet de recherches très actives. Elle a été détectée dans une grande variété d'environnements circumstellaires et interstellaires, y compris dans des nébuleuses protoplanétaires [2], grâce à son spectre

dans l'infra-rouge (IR) résultant de transitions entre ses niveaux d'énergie vibrationnels. De plus, une expérience récente de spectroscopie haute résolution sur molécules isolées vient de livrer des données d'une qualité sans précédent [3], dont l'interprétation reste très incomplète. C'est pourquoi, une étude théorique du spectre de rotation-vibration du C60 est du plus vif intérêt.

Nous avons mis au point au laboratoire un code, CONVIV, qui résout l'équation de Schrödinger avec une précision qui correspond à l'état de l'art, pour les degrés de liberté de rotation-vibration moléculaires [4]. Il peut être utilisé pour calculer ab initio le spectre IR de molécules de tailles arbitraires, y compris le C60 qui comporte 174 degrés de liberté internes. Il a permis de tester les champs de force actuellement disponibles pour ce système dans la littérature [5]. Les résultats ab initio obtenus avec CONVIV seront confrontés aux données expérimentales et observationnelles, afin d'extraire le plus d'informations possible de ces dernières, en particuliers les données de l'interféromètre Matisse et surtout du James Webb space telescope qui devrait être mis en orbite cette année.

Références :

- [1] G. Mulas, C. Falvo, P. Cassam-Chenai, C. Joblin, *J. Chem. Phys.* 149, 144102 (2018).
- [2] Y. Zhang, S. Kwok, *Astrophys. J.* 730, 126 (2011).
- [3] P. B. Changala, M. L. Weichman, K. F. Lee, M. E. Fermann, J. Ye, *Science* 363, 49 (2019).
- [4] P. Cassam-Chenai, J. Liévin, *Journal of Computational Chemistry* 27, 627-640 (2006).
- [5] Daniel A. Jelski, Laszlo Nemes, and S Broughton *Journal of Cluster Science* 16, 1 (2005).

Uncertainty in nano-particule agglomeration

Responsable scientifique : Christophe Henry

Laboratoire : Inria (équipe Calisto)

Description :

Aim and Objectives :

The objective of this project is to refine existing models for particle agglomeration using Molecular Dynamics simulations. In particular, we aim at analyzing these results to extract statistical information on the morphology of the aggregates formed and the sensitivity of these results to some of the input parameters. The analysis of these results will be used to refine existing macroscopic models based on Population Balance Equation (PBE).

Methodology :

The methodology used here consists in simulating the agglomeration of nanoscopic carbon quantum dots undergoing purely diffusive motion (Brownian motion).

Carbon quantum dots (CQD) are particles formed by carbon sheets (see Paloncova et al, 2018). As displayed in Figure 1, we consider in this study CQDs composed of 3 benzene rings on the edge size and of a total of 5 layers. This means that each CQD has a diameter of the central layer equal to 1.6 nm and a height of 1.74 nm. Each CQD is composed of 288 atoms.

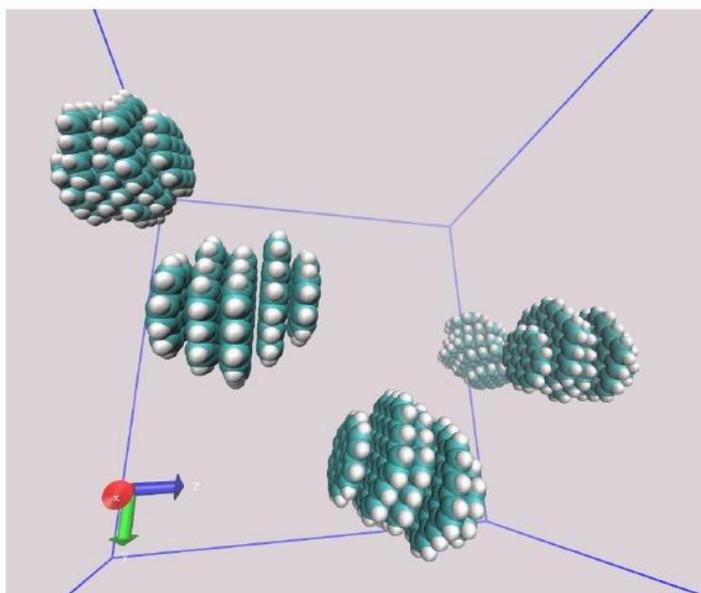


Figure 1 : snapshot of 5 carbon quantum dots (CQD) in a simulation box.

Each simulation is performed by randomly introducing 5 CQDs within a cube of size $L = 10$ nm filled with water molecules (not shown in Figure 1 for the sake of clarity). Periodic boundary conditions are applied in all directions. The temperature is fixed at 300 K. Then, Molecular Dynamics (MD) simulations are performed using the GROMACS simulation package with a simulation time step of 2 fs over an equilibration period of 20 ns. This total simulation time has been chosen long enough to ensure that a single aggregate is formed by the end of the simulation (i.e. all 5 CQDs form one large cluster).

With this project, we aim at running several simulations starting from various random initial configurations of the CQDs and to extract statistical information on the agglomeration dynamics. Among other outputs that will be measured, we plan to measure the aggregation time (i.e. the time needed to have a single cluster formed of all 5 CQDs), the spatial and temporal correlations between consecutive aggregation events, the aggregate morphology (overall shape, porosity, orientation, fractal dimension).

Outcome and impact :

This work is part of a project funded from the European Union's Horizon 2020 research and innovation programme through grant agreement no. 760907, Virtual Materials Marketplace. Within the framework of this EU project, we aim at running here a new scientific workflows that coupled both microscopic approaches (MD here) and macroscopic approaches (PBE here).

The results and analysis of these simulations will be used to write a scientific paper.

Bibliography :

- [1] Palonc'ová, M., Langer, M. and Otyepka, M., 2018. Structural dynamics of carbon in water and *n*, *n*-dimethylformamide probed by all-atom molecular dynamics simulations. *Journal of chemical theory and computation*, 14(4), pp.2076-2083. <https://pubs.acs.org/doi/10.1021/acs.jctc.7b01149#citeThis>
- [2] Jauja-Ccana, V. R., Cordova-Huaman, A. V., Feliciano, G. T., & Gómez, A. L. R. T. (2021). Experimental and molecular dynamics study of graphene oxide quantum dots interaction with solvents and its aggregation mechanism. *Journal of Molecular Liquids*, 335, 116136. <https://doi.org/10.1016/j.molliq.2021.116136>
- [3] Izvekov, S., & Violi, A. (2006). A coarse-grained molecular dynamics study of carbon nanoparticle aggregation. *Journal of chemical theory & computation*, 2(3), 504-512. <https://doi.org/10.1021/ct060030d>

Adsorption Modes of Main Protease of Coronavirus SARS-COV-2 onto Gold Nanoparticles of Different Shapes and Sizes

Responsible scientifique : Oleg N. Kalugin, Alexander Kyrychenko

Laboratoire : V.N. Karazin Kharkiv National University, Department of Inorganic Chemistry

Description :

Introduction

The outbreak of the novel coronavirus disease, COVID-19, caused by the new coronavirus 2019-nCoV, assigned as Severe Acute Respiratory Syndrome-related coronavirus (SARS-CoV-2), represents a pandemic threat to worldwide public health. Extraordinary efforts have been undertaken to suggest the treatment of SARS-CoV-2 infection, spanning various approaches ranging from traditional medicinal chemistry, and drug repurposing, up to numerous computational methods.

The main peptidase (M^{pro}) from the coronavirus SARS-CoV-2 is one of the most promising molecular targets for the development of anti-SARS agents. Recently, gold nanoparticles (AuNPs) have been reported as a promising alternative to available strategies based on traditional medicinal chemistry and drug repurposing.¹ It has recently been reported that inorganic nanoparticle (NPs) can interact strongly with viruses due to morphological similarities and size domain of SARS-CoV-2.¹⁻³ M^{pro} binding to gold nanoparticles leads to the formation of the protein corona, which, however, may alter the protein structure and enzymatic activity. Despite biological importance of this phenomena, our understanding of the molecular aspects M^{pro} -gold interactions remains incomplete.

Project description

The project is devoted to studying the binding interactions of the main protease (M^{pro}) of virus SARS-CoV-2 with bare gold nanoparticles of different sizes and shapes by all-atom molecular dynamics (MD) simulations (Figures 1-2). Crystal structures of M^{pro} (PDB code: 6LU7)⁴ was obtained from the Protein Data Bank. Protonation states for all titratable residues were set at their default state at pH = 7. The ribbon representation of M^{pro} and its catalytic dyad residues His41 and Cys145 are shown as stick models in Figure 1. The

CHARMM27 force field⁵ and TIP3P model⁶ were used to treat the protein and explicit water molecules, respectively.

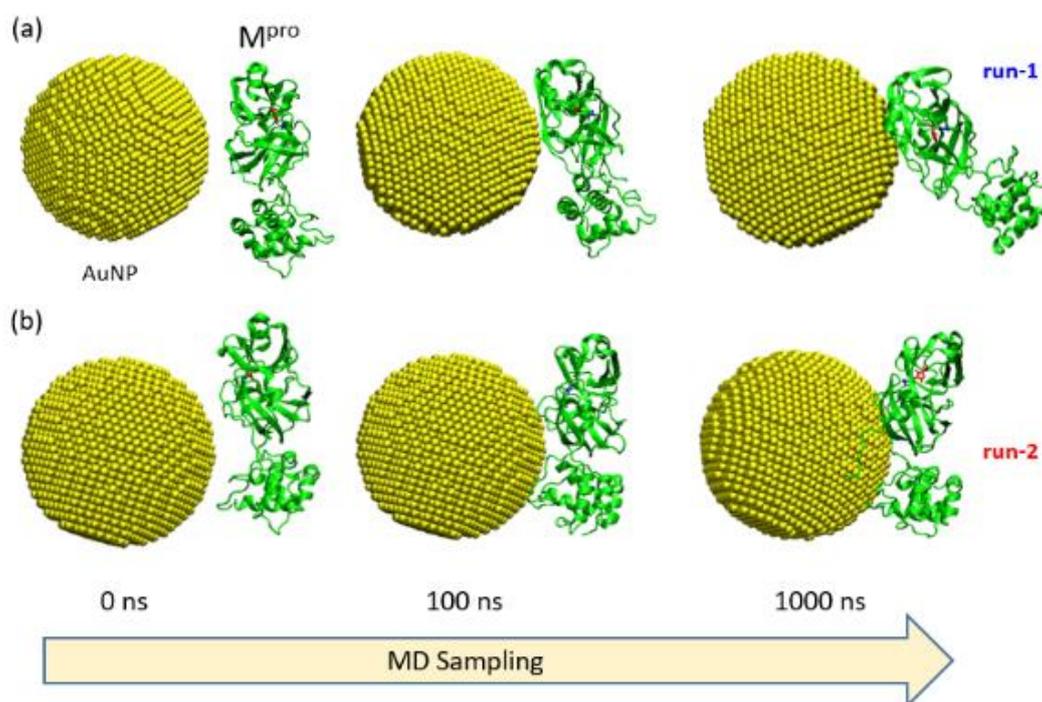


Figure 1. The simulation results of the interaction of M^{pro} with the gold nanoparticle (AuNP₆₆₃₂). Typical snapshots of the initial and adsorbed structures of M^{pro} onto AuNP for two independent MD runs, referred to as run-1/run-2, respectively. The protein is shown in green. Water molecules and ions are omitted for clarity.

To model AuNPs, the repulsion and dispersion terms of nonbonded interactions between Au atoms were computed by using the Lennard-Jones (LJ) 12-6 pairwise potential energy function with $\sigma=0.2629$ nm and $\epsilon=22.13$ kJ/mol, respectively. The performance of these 12-6 LJ pair potentials has recently been validated by modeling the adsorption behavior of small ligand and peptides onto gold nanoparticles.⁷ The Lorentz-Berthelot combination rules were used to calculate mixed nonbonded interactions between the protein, gold, and water atoms. All MD simulations and analysis were carried out by using the GROMACS set of programs, version 2021.4.

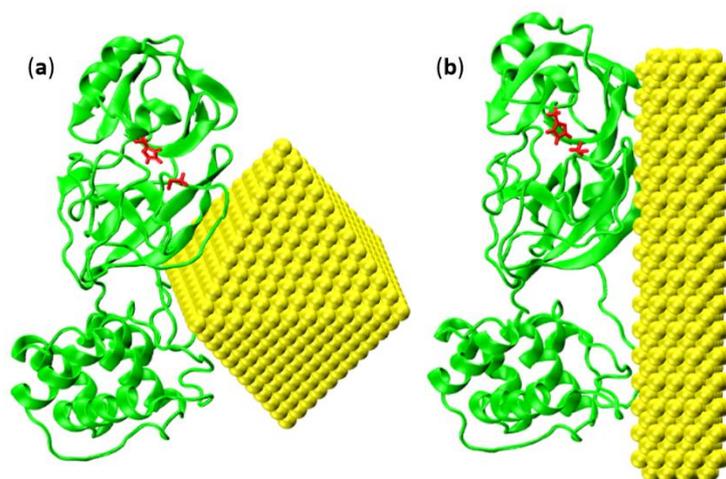


Figure 2. Structure of M^{pro} (PDB 6LU7) adsorbed onto AuNPs of different shapes: (a) bipyramidal (AuNB₁₁₅₆) and (b) rod (AuNR₆₀₀)

Preliminary results of this project suggest that M^{pro} reveals a high binding affinity towards all studied AuNPs (Figures 1-2). Importantly, inhibitory effects of AuNPs were observed as local protein refolding and the loss of the native spatial structure of its active site residues His41 and Cys145 accompanied the M^{pro} binding onto small sharp-edged AuNPs (Figure 2). Our findings point out that the local misfolding of M^{pro} on an AuNP surface may lead to partial or complete loss of its catalytic function, inhibiting the replication of the virus SARS-CoV-2

References:

1. Medhi, R.; Srinoi, P.; Ngo, N.; Tran, H.-V.; Lee, T. R. Nanoparticle-Based Strategies to Combat Covid-19. *ACS Applied Nano Materials* 2020, 3(9), 8557-8580.
2. Chen, L.; Liang, J. An Overview of Functional Nanoparticles as Novel Emerging Antiviral Therapeutic Agents. *Materials Science and Engineering: C* 2020, 112110924.
3. Mallakpour, S.; Azadi, E.; Hussain, C. M. The Latest Strategies in the Fight against the Covid-19 Pandemic: The Role of Metal and Metal Oxide Nanoparticles. *New Journal of Chemistry* 2021, 45(14), 6167-6179.
4. Jin, Z.; Du, X.; Xu, Y.; Deng, Y.; Liu, M.; Zhao, Y.; Zhang, B.; Li, X.; Zhang, L.; Peng, C.; Duan, Y.; Yu, J.; Wang, L.; Yang, K.; Liu, F.; Jiang, R.; Yang, X.; You, T.; Liu, X.; Yang, X.; Bai, F.; Liu, H.; Liu, X.; Guddat, L. W.; Xu, W.; Xiao, G.; Qin, C.; Shi, Z.; Jiang, H.; Rao, Z.; Yang, H. Structure of M^{pro} from Sars-Cov-2 and Discovery of Its Inhibitors. *Nature* 2020, 582(7811), 289-293.

-
5. Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell Jr, A. D. Charmm General Force Field: A Force Field for Drug-Like Molecules Compatible with the Charmm All-Atom Additive Biological Force Fields. *Journal of Computational Chemistry* 2010, *31* (4), 671-690.
 6. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *Journal of Chemical Physics* 1983, *79* (2), 926-935.
 7. Slavgorodska, M. V.; Gurova, Y. O.; Kyrychenko, A. Γ -Cyclodextrin as a Capping Agent for Gold Nanoparticles. *Computational and Theoretical Chemistry* 2021, *1194* 113060.

Molecular Dynamics Simulation of the Ionic-molecular Systems of a Different Nature: Influence of Ion Aggregation on Transport Properties

Responsible scientifique : Oleg N. Kalugin

Laboratoire : V.N. Karazin Kharkiv National University, Department of Inorganic Chemistry

Description :

Ion-molecular systems are the main component of batteries, supercapacitors, lithium- and sodium-ion batteries, as well as solar cells. The operational properties of such electrochemical devices mainly depend on the viscosity and electrical conductivity of their components and the nature of interparticle interactions, in particular ion-ion interactions. Currently, several hypotheses have been proposed to explain the decrease in electrical conductivity of ion-molecular systems when certain electrolyte concentrations are reached. This may be due to the increase in the viscosity of solutions with increasing concentration, as well as the formation of associates.

The phenomenon of ionic association has been widely studied in the literature, but quantitative characteristics describing this process in highly concentrated systems used in electrochemical devices are scarcely available for experimental research. The use of the molecular dynamic modeling method provides an opportunity to solve this problem, allowing to obtain structural and dynamic characteristics of the studied objects, which are extremely difficult to obtain from "direct" methods of studying condensed matter.

In this work we are studying several types of systems that has maximum on conductance dependence on concentration:

- LiPF_6 in the mixture of ethylene carbonate and dimethyl carbonate
- SBPBF_4 in acetonitrile (AN).
- Bmim-based ionic liquids with a set of anions and solvents (BF_4^- , PF_6^- , TFO^- , TFSI^- in acetonitrile, propylene carbonate and γ -butyrolactone)

As an example Fig. 1 and Fig. 2 demonstrate the changes in ion aggregations and transport properties in the solutions of SBPBF_4 in AN obtain by MD simulation.

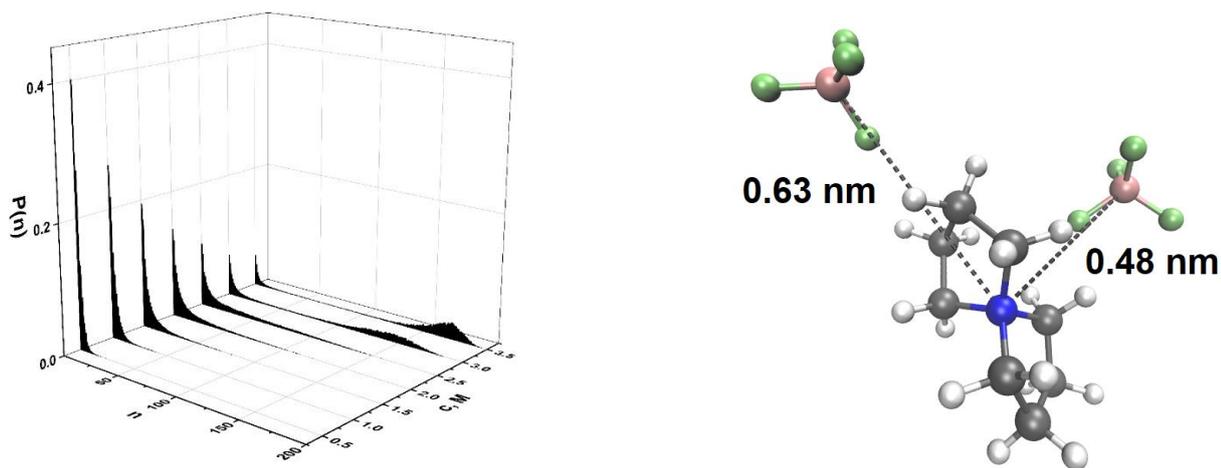


Figure 1. Ionic association composition analysis SBPBF₄ in AN solutions of various concentrations (left) and the snapshot of the aggregate [SBP(BF₄)₂]⁻ with axial and equatorial relative positions of anions.

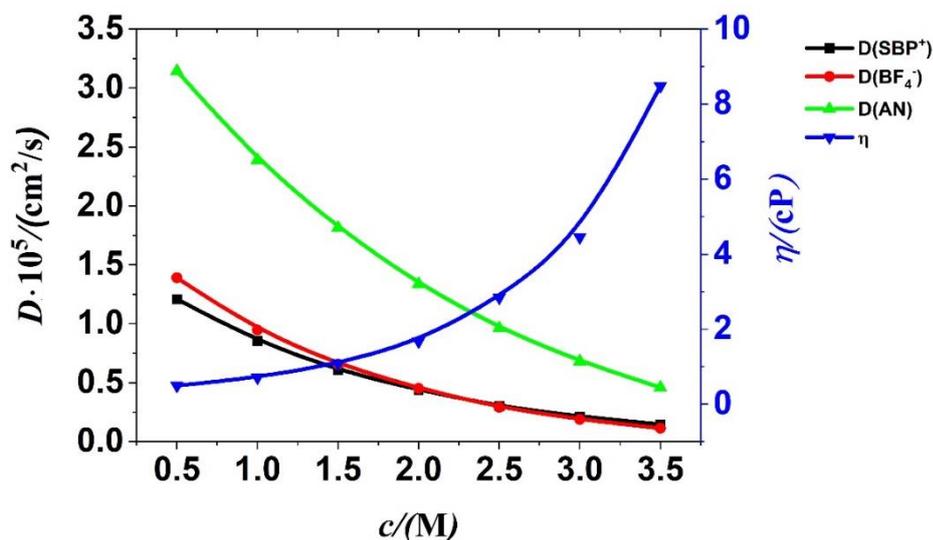


Figure 2. Diffusion coefficients for cation, anion and solvent molecules and viscosity values of the mixtures at various molar concentration of SBPBF₄ in AN mixture.

As can be seen from the Fig. 2, the diffusion coefficients of all components as well as the viscosity concentration dependence are in good agreement with each other. With the increasing of electrolyte concentration, the diffusion of all ions and molecules becomes lower while the viscosity is increasing drastically.

Preliminary results shows that in the investigated ion-molecular systems the ion aggregation play a major role in the concentration dependence of all transport properties.

Physique non linéaire, fluides complexes

EUR Spectrum

Rhéologie des suspensions concentrées

Responsable scientifique : François Peters

Laboratoire : INPHYNI

Résultats obtenus grâce aux moyens informatiques attribués :

2 articles acceptés :

Badia, A., D'Angelo, Y., Peters, F., & Lobry, L. (2022). Frame-invariant modeling for non-Brownian suspension flows. *Journal of Non-Newtonian Fluid Mechanics*, 309, 104904.

Michel Orsi, Laurent Lobry, François Peters, Frame-invariant sub-grid corrections to the Fictitious Domain Method for the simulation of particulate suspensions in nonlinear flows using OpenFOAM, *Journal of Computational Physics*, Volume 474, 2023, 111823, ISSN 0021-9991, <https://doi.org/10.1016/j.jcp.2022.111823>

Présentations

IUTAM Symposium: From Stokesian suspension dynamics to particulate flows in turbulence. 29 août – 2 septembre 2022. Toulouse, France).

M. Orsi, L. Lobry et F. Peters. Simulation of particulate suspensions in non-linear flows: formulation of frame-invariant sub-grid corrections to the Fictitious Domain Method

Annual European Rheology Conference (26-28 avril 2022, Sevilla, Spain)

M. Orsi, L. Lobry et F. Peters. INFLUENCE OF COUPLED FRICTION AND ADHESION ON THE RHEOLOGY OF NON-BROWNIAN SUSPENSIONS

Thèse :

Simulation de suspensions non-Browniennes concentrées frictionnelles et adhésives en écoulement linéaire et non-linéaire par Michel Orsi. Soutenue le 2 décembre 2022.

<https://theses.fr/s317644>

Description du projet :

L'équipe Rhéologie des Suspensions Concentrées de l'Institut de Physique de Nice s'intéresse aux propriétés d'écoulement des suspensions de particules solides en écoulement lent. Du point de vue macroscopiques, ces propriétés sont prises en compte

par des lois de comportement de fluides continus décrivant des relations mathématiques entre grandeurs macroscopiques (contraintes, taux de déformation, flux de particules). Les fonctions matérielles (viscosité généralisées, coefficient de diffusion de particules etc.) mises en jeu reflètent, de façon encore partiellement incomprise, les mécanismes physiques actifs à l'échelle des particules (interactions hydrodynamiques, interactions directes entre particules, en particulier contact). Ces mécanismes microscopiques mettent en jeu des paramètres microscopiques assez divers (propriétés rhéologiques du fluide suspendant, taille, forme des particules, ainsi que leur élasticité et leurs propriétés de surface). Depuis une dizaine d'année, l'équipe RSC tente d'approfondir la compréhension des liens entre ces mécanismes à l'échelles microscopiques et le comportement mésoscopique à une échelle supérieure. Pour ce faire, nous nous appuyons sur des approches combinées d'expériences de rhéologie macroscopique classique, de mesures locales (à l'échelle de la particule), de simulations résolues à l'échelle des particules, de modélisation et de calcul d'écoulement en milieux continus. Plusieurs de ces approches nécessitent la mise en œuvre de calculs assez lourds, c'est pourquoi le projet présenté ici se décline en plusieurs volets qui seront détaillés ci-dessous. On notera qu'une partie de ces travaux fait l'objet d'un projet ANR (AMARhéo 2019- 2022) qui a notamment financé la machine SMP du cluster Azzurra.

Simulations discrètes de suspensions

Les simulations discrètes ont pour but de décrire la dynamique des particules et du fluide en tenant compte des interactions pertinentes au sein de la suspension, et d'en déduire, au moyen de moyenne, les lois de comportements macroscopiques permettant en retour de décrire le matériau comme un milieu continu. La difficulté de cette démarche réside dans le haut niveau de couplage existant entre la dynamique du fluide (partie « fluide » du code) et celle des particules (partie DEM: Discrete Element Method). Nous utilisons pour ce faire une méthode de domaines fictifs, qui permet de calculer un écoulement dans tous le volume représentatif de la suspension. La présence des particules, notamment leur caractère indéformable, est prise en compte via une densité de force calculée à chaque pas de temps, et qui traduit les relations de conservations mécaniques fondamentales. L'écoulement ainsi déterminé permet de calculer les interactions hydrodynamiques entre particules qui en résultent. Les écoulements à une échelle plus petite que la maille ne sont évidemment pas accessibles et donnent lieu à des corrections

de force et de moment hydrodynamiques « sousmaille » sur les particules. Du point de vue des particules, outre ces interactions hydrodynamiques, les interactions directes sont de première importance. Dans notre cas, il s'agit d'interactions de contact, mettant en jeu à la fois des forces élastiques, de frottement et d'adhésion. Cette approche, qui nécessite de décrire des échelles inférieures à la taille des particules dans des systèmes comptant un nombre suffisant de particules pour obtenir des statistiques suffisantes, demande une puissance de calcul importante. Elle a contribué ces dernières années à clarifier l'influence du frottement entre particules sur la rhéologie des suspensions.^{26 27 28 29}

Deux codes différents fondés sur deux méthodes légèrement différentes, seront mis en oeuvre. Le premier utilise la méthode FCM (Force Coupling Method). L'écoulement est calculé en utilisant des méthodes spectrales (typiquement F.F.T.) implémenté dans le langage Fortran. Ce code, faiblement parallélisé (procédure OpenMP), permet de gérer des systèmes de petite taille (typique 1000 particules) pour des durées de calcul d'une quinzaine de jours sur 4 coeurs. Ce code a déjà été validé et utilisé sur le cluster Cicada, et a donné lieu à plusieurs publications, . Un deuxième code (Domaine fictif sous OpenFOAM) a donné lieu sous une forme différente à plusieurs publication, . Son adaptation à l'environnement OpenFOAM est en cours de mise au point. Il s'agit ici de paralléliser massivement les calculs en vue de simuler des systèmes de très grande taille. Ceci permettra de décrire la dynamique des particules jusqu'à des échelles intermédiaires (mésoscopiques) entre la taille des particules et les distances typiques caractéristiques des écoulements pertinents pour les applications. C'est notamment crucial pour les écoulements où la fraction volumique de solide et le taux de cisaillement présentent des gradients spatiaux et imposent donc au système une telle échelle intermédiaire. Ces

²⁶ Stany Gallier, Elisabeth Lemaire, Laurent Lobry, and François Peters. A fictitious domain approach for the simulation of dense suspensions. *Journal of Computational Physics*, 256:367–387, 2014.

²⁷ Stany Gallier, Elisabeth Lemaire, François Peters, and Laurent Lobry. Rheology of sheared suspensions of rough frictional particles. *Journal of Fluid Mechanics*, 757:514–549, 2014.

²⁸ François Peters, Giovanni Ghigliotti, Stany Gallier, Frédéric Blanc, Elisabeth Lemaire, and Laurent Lobry. Rheology of non-brownian suspensions of rough frictional particles under shear reversal: A numerical study. *Journal of rheology*, 60(4):715–732, 2016.

²⁹ Laurent Lobry, Elisabeth Lemaire, Frédéric Blanc, Stany Gallier, and François Peters. Shear thinning in non-brownian suspensions explained by variable friction between particles. *Journal of Fluid Mechanics*, 860:682–710, 2019.

gradients sont à l'origine de flux de particules qui, s'ils ont été étudié depuis longtemps du point de vue expérimental, sont actuellement mal compris du point de vue microscopique, et imparfaitement décrits par les modèles de milieux continus actuels. La majorité des simulations s'est jusqu'ici limitée à des systèmes homogènes. La simulation discrète de tels systèmes hétérogènes est donc d'un très grand intérêt dans ce domaine. Le calcul des écoulements est réalisé par la méthode des volumes finis dans l'environnement libre OpenFOAM qui offre des outils standards de résolution des équations de Navier-Stokes / Stoke de façon parallèle (procédure MPI). L'adaptation de la méthode des domaines fictifs à cet environnement, avec en particulier la parallélisation de la partie DEM de la méthode, constitue une partie du travail de thèse de Michel Orsi, dont le contrat est également financé par l'ANR AMARhéo. Il est prévu que des jobs puissent être lancés sur typiquement 100-200 coeurs.

Simulations en milieux continus

L'objet de cette partie est de développer et utiliser des modèles en milieux continus capables de décrire l'écoulement des suspensions, y compris la diffusion des particules, dans des géométries d'écoulement complexes rencontrées soit dans des expériences de recherche menées dans l'équipe, soit dans le traitement de problématiques industrielles. Le verrou principal de cette activité réside dans la nécessité de décrire les équations de transport de la quantité de mouvement et de la masse de la suspension vue comme un milieu continu, ainsi que le transport des particules (diffusion) par rapport à ce milieu continu. L'approche suivie, classique, consiste en la définition, en plus de la phase moyenne (suspension), d'une phase solide, certes transportée par la suspension, mais également capable de mouvement relativement à elle. La problématique principale dans ce domaine consiste à définir les lois de comportement décrivant le transport de cette phase solide à partir de contraintes développées dans la phase particulaire. Les modèles actuels peinent en effet à reproduire les résultats expérimentaux de l'équipe RSC et plus largement de la communauté scientifique qui s'intéresse à ces problèmes. Les équations aux dérivées partielles issues des modèles adoptés sont résolus par la méthode des volumes finis dans l'environnement OpenFOAM, qui comme cela a déjà été rappelé, permet une parallélisation par la procédure MPI. Ce travail, déjà en cours depuis quelques

années³⁰, a été mis en oeuvre au sein du cluster Cicada, et constitue le travail de thèse d'Ayoub Badia, dirigé conjointement par Yves D'Angelo au laboratoire J.A.D. et Laurent Lobry au laboratoire InPhyNi.

Traitement d'images expérimentales par des méthodes d'intelligence artificielle

L'équipe RSC a développé des méthodes de rhéologie locale capable de mesurer, en plus des grandeurs rhéologiques intégrées classiques (couple, taux de cisaillement), des grandeurs locales au sein de la suspension (fraction volumique locale, taux de déformation local). Ces mesures locales constituent une très grande valeur ajoutée pour l'étude quantitative des lois responsables du transport de la suspension et de la phase solide au sein d'un écoulement. Elles permettent en effet de mesurer directement les gradients de vitesse, taux de cisaillement, fraction volumique qui déterminent les différents flux (matière, quantité de mouvement etc). Elles ont notamment permis de mettre en évidence le rôle primordial des contacts entre particules dans les propriétés rhéologiques des suspensions non-browniennes.^{31 32}

Ceci se fait au prix d'expériences complexes et d'un traitement de données assez lourd. En effet les suspensions utilisées doivent être rendues transparentes en adaptant l'indice des particules à celui du fluide, qui est lui-même rendu visible par l'adjonction d'un colorant fluorescent activé par une nappe laser, dont l'image est enregistré par une caméra. L'épaisseur de la nappe laser étant plus fine que le taille des particules, celles-ci apparaissent comme des disques noirs sur les images enregistrées par la caméra. L'essentiel du traitement de ces images consiste à les segmenter pour séparer l'intérieur des particules du liquide avant de déterminer la position des centres des particules. Ce traitement est classiquement réalisé par des méthodes morphologiques, mais cela

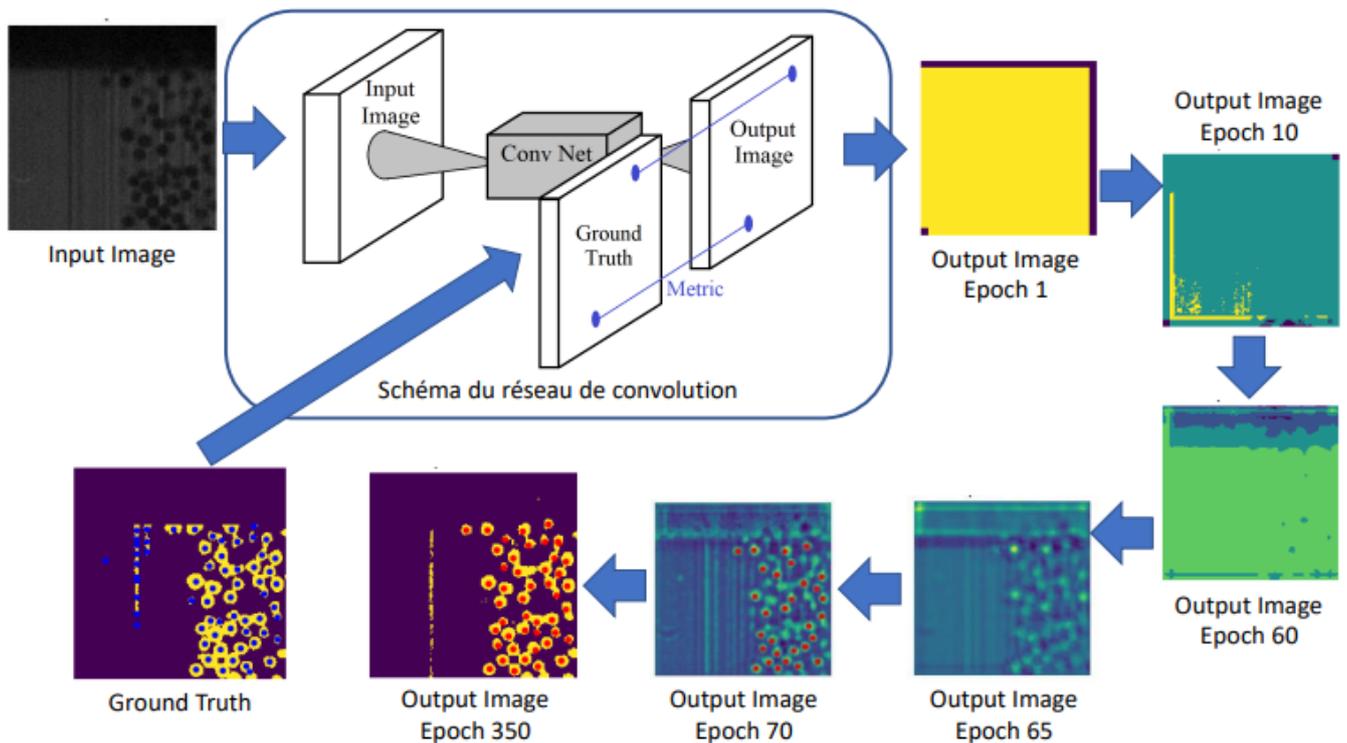
³⁰ Talib Dbouk, Elisabeth Lemaire, Laurent Lobry, and F Moukalled. Shear-induced particle migration: Predictions from experimental evaluation of the particle stress tensor. *Journal of Non-Newtonian Fluid Mechanics*, 198:78–95, 2013.

³¹ Frédéric Blanc, François Peters, and Elisabeth Lemaire. Experimental signature of the pair trajectories of rough spheres in the shear-induced microstructure in noncolloidal suspensions. *Physical review letters*, 107(20):208302, 2011.

³² Frédéric Blanc, Elisabeth Lemaire, Alain Meunier, and François Peters. Microstructure in sheared non-brownian concentrated suspensions. *Journal of rheology*, 57(1):273–292, 2013.

nécessite le réglage de nombreux paramètres qui sont fortement dépendant des conditions d'éclairages liées à de subtiles variations de l'indice de réfraction de la suspension. Pour accélérer le traitement et ne plus dépendre du réglage de ces paramètres, nous nous sommes tournés vers des techniques d'intelligence artificielle. Ce travail, effectué en collaboration avec Bruno Figliuzzi du Centre de Morphologie Mathématique de Mines ParisTech, est basé sur un algorithme de réseau neuronal convolutif³³ (CNN). « L'apprentissage » du réseau se fait en utilisant des images de nos expériences de rhéologies locales, segmentées par une méthode classique. Une image alimente l'entrée du réseau et la sortie calculée par le réseau est comparée à la « vérité terrain » (i.e l'image segmentée). Une fonction de « coût » calcule alors l'écart entre la sortie et cette vérité terrain. Pour minimiser ce « coût », les nombreux paramètres du réseau sont subtilement modifiés par un processus de « descente de gradient ». Ce processus est itéré avec une série d'images différentes jusqu'à minimiser le « coût ». Le schéma ci-dessous illustre cet apprentissage. Les images de sorties sont calculées à partir d'une même image d'entrée sur la base des paramètres d'apprentissages obtenus au bout d'un nombre d'itération (epoch). Les points rouges sur l'image de l'epoch 350 correspondent aux particules détectées par le réseau sur l'image segmentée. Ce résultat contient moins d'erreurs de détection que sur la vérité terrain (points bleus).

³³ Chen, Qifeng, Jia Xu, and Vladlen Koltun. "Fast image processing with fully-convolutional networks." Proceedings of the IEEE International Conference on Computer Vision. 2017.



Fort de ces premiers résultats prometteurs, nous orientons nos travaux vers l'apprentissage sur une banque d'images de synthèse dans le but d'adapter la détection à tout type de forme de particules (cubes, fibres, etc ...) ainsi qu'à une extension à des images en trois dimensions. La programmation et l'apprentissage supervisé du réseau sont réalisés en python 3.7 à l'aide des bibliothèques Pytorch-Cuda, Numpy, scikit-learn et scikit-image. Les calculs sont effectués sur GPU (Pytorch gère CUDA de manière transparente). La base de données d'images nécessite une volumétrie d'environ une dizaine de Go (voir plus dans le cas de la 3D) stockée transitoirement sur Azzurra.

Olfactory navigation

Responsable scientifique : Agnese Seminara

Laboratoire : INPHYNI

Résultats obtenus grâce aux moyens informatiques attribués :

Navigation olfactive : comment prendre des décisions avec un signal brisé par Nicola Rigolli.

Thèse soutenue le 19 mai 2022. <https://theses.fr/2022COAZ4023>

Rigolli, N., Magnoli, N., Rosasco, L., & Seminara, A. (2022). Learning to predict target location with turbulent odor plumes. *Elife*, 11, e72196. <https://doi.org/10.7554/eLife.72196>

Rigolli, N., Reddy, G., Seminara, A., & Vergassola, M. (2022). Alternation emerges as a multi-modal strategy for turbulent odor navigation. *Elife*, 11, e76989. <https://doi.org/10.7554/eLife.76989>

Description du projet :

Living systems continuously operate decision-making, based on a large array of sensory inputs that inform their understanding of the world. Chemical cues bear a fundamental source of information, and all domains of life have evolved sophisticated mechanisms to extract, store and process this information. Olfaction is one of the major distant senses and animal species use it to target their computational needs under different physical constraints. Remarkably, the neural architecture underlying the sense of smell is widely conserved. Olfactory navigation often targets the source of an odor, e.g. to survive predators and forage for food. At microscopic scales, where diffusion dominates over advection, chemical signals vary smoothly in space and their gradient typically points to the source. It is well known that individual bacteria take advantage of this property to move towards attractants or away from repellents. However, gradient climbing requires that signals vary smoothly in space, and this is rarely the case at macroscopic scales, where advection dominates over diffusion. Animals living within the atmospheric surface layer, often experience highly fluctuating conditions due to turbulence. The gradients of a chemical substance transported passively by a turbulent flow does not necessarily point to its source. Deriving useful information from turbulent signals is notoriously hard, and

this makes olfactory navigation both fascinating and obscure. In this project, we target a three dimensional numerical simulation of odor transport a turbulent channel flow. The work is divided in three parts:

(1) develop a massive dataset of turbulent odor fields, varying the Reynolds numbers based on mean flow as well as on fluctuations. To this end we will vary geometry and we will introduce an obstacle of various sizes which will tune the level of fluctuations (which is otherwise quite universal in channel flows). We will employ a spectral element algorithm to solve the Navier Stokes equations as well as equation of transport of a passive scalar emitted from a concentrated source. We will use the open source code Nek5000, which is optimized for parallel computing.

(2) develop supervised learning algorithms based on classical regularized empirical risk minimization, using linear; linear-non linear and fully non linear models (using dictionary of features, kernel methods, neural networks – collaboration Lorenzo Rosasco and Nicodemo Magnoli University of Genova). This will elucidate what features of these spatio temporal signals hold the most useful information about the location of the odor source. We will study how inference varies in different flow conditions, especially depending on intermittency, which is a crucial challenge for olfactory navigation (Gire et al, 2016). Apply our understanding to analyze neural representation of odors in mice olfactory bulb (collaboration D Gire).

(3) Use these data to develop reinforcement learning algorithms for navigation in an abstract setting (collaboration with M. Vergassola ENS Paris, G. Reddy and C Rycroft at Harvard University); and to model animal behavior (experiments on mice, octopuses and fishes conducted by D Gire U Washington and N Bellono Harvard University)

Wave turbulence simulations and theory

Responsable scientifique : Sergey Nazarenko

Laboratoire : INPHYNI

Résultats obtenus grâce aux moyens informatiques attribués :

Zhu, Y., Semisalov, B., Krstulovic, G., & Nazarenko, S. (2022). Testing wave turbulence theory for the Gross-Pitaevskii system. *Physical Review E*, 106(1), 014205. (Editor Suggestion)

<https://doi.org/10.1103/PhysRevE.106.014205>

Zhu Y, Semisalov B, Krstulovic G, et al. Direct and inverse cascades in turbulent Bose-Einstein condensate. arXiv preprint arXiv:2208.09279, 2022. <https://doi.org/10.48550/arXiv.2208.09279>

Description :

The proposed research project is part of the Simons Foundation international collaboration project "Wave Turbulence". Our research aims at developing solid theoretical foundations for wave turbulence systems and modelling experiments in wave turbulence performed in laboratory by partners of this collaboration. In this work, large-scale numerical simulations of different wave systems will be performed. We focus on those systems (e.g. Bose-Einstein Condensate (BEC) and nonlinear optics) which can be described by the Gross-Pitaevskii equation (GPE). Then we propose to confront the numerical and theoretical results coming from Weak Wave Turbulence theory (WWT) for BEC, and also numerical and experimental results for nonlinear optics. The Gross-Pitaevskii equation

$$i\hbar \frac{\partial \psi(\mathbf{x}, t)}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi(\mathbf{x}, t) + g |\psi(\mathbf{x}, t)|^2 \psi(\mathbf{x}, t) + F + D$$

is also known as the nonlinear Schrödinger equation, where $\psi(\mathbf{x}, t)$ is the order parameter indicating the condensate wave function in BEC ($\psi(\mathbf{x}, t)$ may represent different physical quantities in other physical contexts), F and D represent possible external forcings and dissipation mechanisms, $\hbar = h / 2\pi$ and $g = (4\pi \hbar^2 a_s) / m$ is the coupling constant fixed by the s-wave scattering length as that models a local interaction between bosons. Here we set $g=1$ corresponding to the defocusing GP.

Etude de la turbulence d'ondes dans les fluides stratifiés et en rotation

Responsable scientifique : Giorgio Krstulovic

Laboratoire : Lagrange / OCA

Description en pages suivantes.

Résumé

Comme de nombreux écoulements géophysiques, les écoulements atmosphériques et océaniques sont dans un état *turbulent*, et sont fortement impactés par la *stratification* et la *rotation*. La compréhension de ces écoulements est encore aujourd'hui un sujet de recherche fondamentale en turbulence et physique statistique. Il s'agit également d'un enjeu majeur dans l'amélioration des prédictions faites par les modèles météorologiques et climatiques.

Une particularité des écoulements stratifiés et/ou en rotation est leur capacité à supporter la propagation d'*ondes internes* : les *ondes de gravité* dans les fluides stratifiés et les *ondes d'inertie* dans les fluides en rotation. L'énergie de l'écoulement est alors répartie entre trois composantes : l'énergie cinétique contenue par les ondes, l'énergie cinétique contenue par les tourbillons, et l'énergie potentielle. En plus de la dynamique des tourbillons, qui constituent les écoulements incompressibles, il est donc nécessaire de décrire le comportement des ondes, et de leurs interactions avec les tourbillons et les structures aux grandes échelles.

La *Théorie de la Turbulence d'Ondes* est utilisée pour décrire les systèmes d'ondes dispersives, interagissant de façon faiblement non-linéaire, avec une dissipation négligeable. Elle est utilisée pour étudier de nombreux systèmes d'ondes, comprenant les écoulements stratifiés et/ou en rotation dans la limite des faibles nombres de *Froude* (forte stratification) et de *Rossby* (forte rotation) et d'un grand nombre de *Reynolds* (régime turbulent). Cette théorie permet, entre autres, d'écrire une *équation cinétique* qui décrit l'évolution de l'amplitude des modes, et fournit des prédictions concernant les quantités statistiques d'intérêt.

Les campagnes de mesures atmosphériques et océaniques permettent d'obtenir les spectres d'énergie cinétique et de température, mais ne permettent de quantifier ni les proportions des différents types d'énergie, ni les échanges entre celles-ci. Bien que des avancées expérimentales importantes ont permis d'améliorer notre compréhension des écoulements stratifiés et/ou en rotation, les expériences ne permettent pas encore d'observer les régimes d'intérêt géophysique, notamment à cause des effets de tailles finis et de la présence de parois. Ces limitations motivent l'utilisation de simulations numériques pour tenter d'obtenir plus d'informations sur ces écoulements.

Le but de ce projet est de confronter les prédictions fournies par la Théorie de la Turbulence d'Ondes à des simulations numériques directes idéalisées, et de quantifier l'importance des ondes dans les différents régimes turbulents, c.-à-d. lorsque les nombres de Froude, de Rossby et de Reynolds varient. Nous utiliserons les équations de Navier-Stokes dans l'*approximation de Boussinesq* en imposant des *conditions aux limites périodiques*. Elles seront résolues à partir d'un *code pseudo-spectral* parallélisé avec *MPI*.

2 Présentation générale

La stratification et la rotation rendent les écoulements turbulents anisotropes. La force de gravité tend à ramener les particules fluides à leur niveau de flottabilité moyen. Dans le cas d'une stratification stable, la gravité tend donc à inhiber les mouvements verticaux. L'écoulement est alors marqué par la présence de structures allongées horizontalement en forme de "pancakes", et une partie importante de l'énergie est transférée vers les modes de cisaillement verticaux. Dans le cas purement tournant, quand le nombre de Rossby est faible, l'écoulement est caractérisé par un équilibre selon l'horizontale appelé équilibre géostrophique (le gradient de pression horizontal équilibre la force de Coriolis). Si cet équilibre est réalisé exactement, on montre que l'écoulement est invariant selon la verticale (théorème de Taylor-Proudman), ce qui explique la présence de structures en forme de colonnes, dites de Taylor. Une partie importante de l'énergie de l'écoulement est alors contenu dans ces colonnes.

Par définition des écoulements turbulents, la dissipation est négligeable sur une grande gamme d'échelle, appelée gamme inertielle. L'équilibre entre stratification d'une part, et l'inertie d'autre part, est atteints à une échelle intermédiaire, appelée échelle d'Ozmidov. En dessous de cette échelle, l'écoulement présente les caractéristiques d'une turbulence homogène et isotrope. En revanche, aux grandes échelles, la turbulence est fortement anisotrope.

Les écoulements stratifiés et/ou en rotation ont la capacité à supporter la propagation d'ondes internes : les ondes de gravité dans les fluides stratifiés et les ondes d'inertie dans les fluides en rotation. L'énergie de l'écoulement est alors répartie entre l'énergie cinétique contenue par les ondes, l'énergie cinétique contenue par les tourbillons, et l'énergie potentielle. On peut donc tenter de décomposer l'écoulement en ondes et en tourbillons et d'étudier les interactions entre ces structures [Lelong and Riley, 1991].

Les écoulements océaniques sont caractérisés pas une forte stratification (faible Froude), une forte rotation (faible Rossby), et une turbulence développée (grand Reynolds). Il est maintenant reconnu que les ondes internes jouent un rôle important dans ce régime, et donc dans les processus océaniques aux grandes échelles [MacKinnon et al., 2017], notamment dans le transport de masse, d'impulsion, et d'énergie. La compréhension de la dynamique de ces ondes est donc importante pour l'amélioration des modèles atmosphériques et océaniques [Gregg et al., 2018]. Bien qu'une équation d'état réaliste soit nécessaire pour modéliser l'atmosphère, les équations de Navier-Stokes dans l'approximation de Boussinesq sont souvent utilisée pour étudier les fluides stratifiés [Brethouwer et al., 2007]. Elles sont même utilisées dans certains modèles océaniques.

D'un point de vue théorique, la théorie de la turbulence d'ondes [Zakharov et al., 2012, Nazarenko, 2011] fournit un cadre d'étude des systèmes d'ondes dispersives, interagissant de façon faiblement non linéaire, faiblement impactées par la dissipation. Elle s'applique aux écoulements stratifiés [Caillol and Zeitlin, 2000] et/ou en rotation [Galtier, 2003, Medvedev and Zeitlin, 2007] dans la limite des faibles nombres de Froude et de Rossby et d'un grand nombre de Reynolds. Cette théorie permet, entre autres, d'écrire une équation cinétique qui décrit l'évolution de l'amplitude des modes (qui est l'analogie de l'équation de Boltzmann pour un système d'ondes, dans le domaine spectral). Elle fournit des prédictions concernant les spectres d'énergie, la direction de la cascade d'énergie, les fonctions de structures et des autres objets statistiques d'intérêt.

Bien que des avancées expérimentales importantes ont permis de vérifier certaines prédictions de la turbulence d'onde dans les écoulements stratifiés [Davis et al., 2020] et/ou en rotation [Monsalve et al., 2020], les expériences ne permettent pas d'observer les régimes d'intérêt géophysique, notamment à cause des effets de tailles finis et de la présence de parois. Les informations expérimentales sur les ondes internes viennent principalement des capteurs de vitesse et de températures immergées, utilisées lors des campagnes de mesures maritimes [Garrett and Munk, 1979, Polzin and Lvov, 2011]. Certaines mesures de spectres océaniques sont cohérentes avec la théorie de la turbulence d'ondes [Lvov et al., 2004]. Il existe également de nombreuses simulations de turbulence fortement stratifiée [Waite and Bartello, 2006, Pouquet et al., 2017, Maffioli, 2017]. La principale difficulté à observer le régime de turbulence d'ondes vient du fait qu'il faut un faible nombre de Froude, mais aussi un grand nombre de Reynolds de flottabilité [Brethouwer et al., 2007]. Si l'on souhaite étudier les propriétés statistiques des ondes, il faut également que les simulations soient suffisamment longues devant le temps typique d'interaction des ondes.

Le but de ce projet est d'effectuer des simulations numériques des équations de Navier-Stokes dans l'approximation de Boussinesq pour les fluides stratifiés et en rotation afin de vérifier les prédictions de la théorie de la turbulence d'ondes. Le projet se décompose en 3 parties :

1. Fluides stratifiés :

- Nous souhaitons réaliser des écoulements statistiquement stationnaires en utilisant la projection poloidale. Le but est de se rapprocher au mieux d'un système d'ondes de gravité en éliminant les modes vorticaux et les modes de cisaillement de la dynamique.
- Nous voulons également réaliser des simulations des équations sans projection poloidale, à faible Froude. Ces simulations sont importantes pour observer le comportement des ondes dans un écoulement plus réaliste. En particulier, ces simulations nous aideront à comprendre l'effet des modes vorticaux et le rôle des ondes de gravité dans le mélange d'énergie.

2. Fluides en rotation :

- Nous souhaitons réaliser des écoulements statistiquement stationnaires en éliminant les colonnes de Taylor de la dynamique. Le but est de se rapprocher au mieux d'un

système d'ondes d'inertie en éliminant les colonnes de Taylor de la dynamique.

- Nous voulons également réaliser des simulations des équations complètes, à faible Rossby. Ces simulations sont nécessaires pour observer le comportement des ondes d'inertie dans un écoulement plus réaliste. En particulier, ces simulations nous aideront à comprendre l'effet des colonnes de Taylor et le rôle des ondes d'inertie dans le mélange d'énergie.

3. Fluides stratifiés et en rotation :

Dans un dernier temps, nous souhaiterions faire une simulation avec une forte rotation et une forte stratification pour comprendre le comportement des ondes gravito-inertielles. Ce cas est le plus pertinent pour les applications géophysiques.

Ce travail théorique et numérique vise à compléter les expériences menées actuellement aux laboratoires FAST (Orsay) par l'équipe de Pierre Philippe Cortet, et au LEGI (Grenoble) par l'équipe de Nicolas Mordant. Il s'inscrit dans le cadre de notre participation au projet de la Simons Foundation sur la turbulence d'ondes.

Nous utiliserons le code pseudo-spectral fluidsim (<https://fluidsim.readthedocs.io/en/latest/index.html>) développé par Pierre Augier (LEGI) avec qui nous collaborons directement. Ce code a déjà utilisé sur différents clusters pour effectuer des calculs intensifs.

3 Méthode

3.1 Modèle théorique

Le modèle le plus simple pour décrire les fluides stratifiés en rotation consiste à considérer un fluide Newtonien dans l'approximation de Boussinesq, dans lequel le gradient de densité moyen est constant. Plus précisément, on considère que le champ de densité ρ peut se décomposer sous la forme suivante :

$$\rho = \rho_0 + \frac{d\bar{\rho}}{dz}(z - z_0) + \rho' \quad (1)$$

où $\rho_0 + \frac{d\bar{\rho}}{dz}(z - z_0)$ représente la densité moyenne (profil linéaire), et ρ' est la perturbation de densité. Dans ce cas, les équations du mouvement du fluide s'écrivent

$$\nabla \cdot \mathbf{v} = 0, \quad (2)$$

$$\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} = -\nabla p - b \mathbf{e}_z - 2\boldsymbol{\Omega} \times \mathbf{v} + \nu \Delta \mathbf{v}, \quad (3)$$

$$\frac{\partial b}{\partial t} + \mathbf{v} \cdot \nabla b = N^2 v_z + \kappa \Delta b, \quad (4)$$

où $\mathbf{v} = (v_x, v_y, v_z)$ est le champ de vitesse, p est la pression totale (incluant la pression hydrostatique et la force centrifuge),

$$N \equiv \sqrt{-\frac{g}{\rho_0} \frac{d\bar{\rho}}{dz}}$$

est la fréquence de Brunt-Väisälä, g est l'accélération de pesanteur,

$$b \equiv -\frac{g\rho'}{\rho_0}$$

est la flottabilité, $\boldsymbol{\Omega} = \Omega \mathbf{e}_z$ est la vitesse angulaire du référentiel, ν est la viscosité cinématique, et κ est la diffusivité. Ici, on considère le cas où l'axe de rotation est parallèle à l'axe de stratification. Nous utiliserons des conditions aux limites périodiques en espace.

On considère un écoulement de vitesse horizontale caractéristique U_h (par exemple la fluctuation typique de la vitesse horizontale) et de taille L_h . Lorsque la fréquence de Brunt-Väisälä devient grande devant la fréquence de retournement à grande échelle, i.e. $N \gg U_h/L_h$, la stratification influence fortement l'écoulement. L'effet de la stratification est donc contrôlé par le *nombre de Froude*

$$Fr = \frac{U_h}{NL_h}. \quad (5)$$

De façon similaire, l'effet de la rotation est quantifié par le *nombre de Rossby*

$$Ro = \frac{U_h}{2\Omega L_h}. \quad (6)$$

L'effet de la viscosité et de la diffusion sont quantifiés par les nombres de Reynolds et de Peclet :

$$Re = \frac{U_h L_h}{\nu} \quad \text{et} \quad Pe = \frac{U_h L_h}{\kappa}. \quad (7)$$

Nous considérerons un fluide pour lequel $\nu = \kappa$. Par conséquent, les paramètres de contrôle de l'écoulement sont Fr , Ro , et Re uniquement. Il a été montré [Brethouwer et al., 2007] que le régime de turbulence fortement stratifiée est atteint quand le nombre de Reynolds de flottabilité

$$R_b = Re Fr^2 \quad (8)$$

est grand, et que le nombre de Froude est faible (typiquement < 0.02). Ce régime est donc difficile à atteindre dans les expériences et dans les simulations numériques [Pouquet et al., 2017].

La relation de dispersion des ondes gravito-inertielles, dans la limite inviscide, est

$$\omega_{\mathbf{k}} = \pm \sqrt{N^2 \sin^2 \theta_{\mathbf{k}} + 4\Omega^2 \cos^2 \theta_{\mathbf{k}}}, \quad (9)$$

où $\omega_{\mathbf{k}}$ représente la pulsation de l'onde ayant pour vecteur d'onde $\mathbf{k} = (k_x, k_y, k_z)$ et $\theta_{\mathbf{k}}$ est l'angle entre le vecteur d'onde et l'axe \mathbf{e}_z . La vitesse de phase de ces ondes est

$$\mathbf{c}_{p\mathbf{k}} = \frac{\omega_{\mathbf{k}}}{k} \mathbf{e}_{\mathbf{k}} = \pm \frac{\sqrt{N^2 \sin^2 \theta_{\mathbf{k}} + 4\Omega^2 \cos^2 \theta_{\mathbf{k}}}}{k} \mathbf{e}_{\mathbf{k}}, \quad (10)$$

tandis que leur vitesse de groupe est

$$\begin{aligned} \mathbf{c}_{g\mathbf{k}} = \nabla_{\mathbf{k}} \omega_{\mathbf{k}} &= \frac{\partial \omega_{\mathbf{k}}}{\partial k} \mathbf{e}_{\mathbf{k}} + \frac{1}{k} \frac{\partial \omega_{\mathbf{k}}}{\partial \theta_{\mathbf{k}}} \mathbf{e}_{\theta_{\mathbf{k}}} + \frac{1}{k \sin \theta_{\mathbf{k}}} \frac{\partial \omega_{\mathbf{k}}}{\partial \varphi_{\mathbf{k}}} \mathbf{e}_{\varphi_{\mathbf{k}}} \\ &= \pm \frac{(N^2 - 4\Omega^2) \sin \theta_{\mathbf{k}} \cos \theta}{k \sqrt{N^2 \sin^2 \theta_{\mathbf{k}} + 4\Omega^2 \cos^2 \theta_{\mathbf{k}}}} \mathbf{e}_{\theta_{\mathbf{k}}}. \end{aligned} \quad (11)$$

où l'on a introduit le repère sphérique $(\mathbf{e}_{\mathbf{k}}, \mathbf{e}_{\theta_{\mathbf{k}}}, \mathbf{e}_{\varphi_{\mathbf{k}}})$ associé au vecteur d'onde \mathbf{k} .

La transformée de Fourier du champ de vitesse peut alors se décomposer comme

$$\hat{\mathbf{v}}_{\mathbf{k}} = \hat{v}_{\mathbf{k}\theta} \mathbf{e}_{\theta_{\mathbf{k}}} + \hat{v}_{\mathbf{k}\varphi} \mathbf{e}_{\varphi_{\mathbf{k}}} \quad (12)$$

où $\hat{v}_{\mathbf{k}\theta}$ est la composante poloïdale et $\hat{v}_{\mathbf{k}\varphi}$ est la composante toroïdale. La composante selon $\mathbf{e}_{\mathbf{k}}$ est nulle par incompressibilité. Pour un fluide purement stratifié, on montre que les ondes ne sont pas portées par la composante toroïdale. La projection du champ de vitesse sur sa composante poloïdale permet d'inhiber la formation de mode vorticaux, et ainsi favoriser l'observation d'un régime de turbulence d'ondes. Pour la même raison, il est possible de ne forcer que la partie poloïdale du champ de vitesse pour favoriser l'excitation des ondes.

Les modes de cisaillement correspondent aux modes de la vitesse horizontale pour lesquels $k_x = k_y = 0$ (invariance de la vitesse horizontale selon l'horizontale), tandis que les colonnes de Taylor correspondent aux modes de la vitesse horizontale pour lesquels $k_z = 0$ (invariance selon la verticale). Comme dans des expériences récentes [Monsalve et al., 2020], nous souhaitons dans un premier temps retirer ces modes de la dynamique pour simplifier le problème.

Deux possibilités sont envisagées pour favoriser l'observation de la gamme inertielle dans ces régimes. La première est d'utiliser un rapport d'aspect plus faible pour limiter la taille de la grille de résolution. La seconde est d'utiliser une hyperviscosité et une hyperdiffusivité

$$-\nu_{\alpha}(-\Delta)^{\alpha/2} \mathbf{u} \quad \text{et} \quad -\kappa_4(-\Delta)^{\alpha/2} b. \quad (13)$$

dans les équations (2-4) avec, typiquement, $\alpha = 4$. Cela a pour effet de repousser la dissipation à des plus petites échelles, pour une dissipation totale égale. L'hyperviscosité sera ajustée de façon à stabiliser les simulations quelle que soit la résolution choisie.

Nous effectuerons d'abord des simulations à des résolutions modérées (typiquement $160 \times 640 \times 640$ ou $240 \times 960 \times 960$) depuis l'état de repos pour atteindre un régime statistiquement stationnaire. Nous raffinerons ensuite progressivement la résolution de ces simulations. Cela permettra de limiter le temps de calcul durant le régime transitoire.

3.2 Modèle numérique

Les équations (2-4) seront résolues avec le code pseudo-spectral `Fluidsim` [Mohan et al., 2019b] du projet `Fluiddyn` [Augier et al., 2019]. Il est principalement écrit en Python. La pression est traitée de façon implicite par projection du champ de vitesse sur l'espace des champs à divergence nulle. Nous utilisons la méthode de Runge-Kutta d'ordre 4 pour le schéma temporel, mais le code permet aussi d'utiliser la méthode de Runge-Kutta d'ordre 2. Le code a été testé sur la machine `Jean-Zay` lors d'un accès préparatoire (projet AP012A13209).

Les tâches coûteuses en calculs sont effectuées par des fonctions optimisées et compilées, ce qui permet au code d'avoir des performances comparables à celles d'un code Fortran ou C++ [Augier et al., 2021]. En particulier, il utilise le module `Fluidfft` [Mohan et al., 2019a] qui gère l'utilisation des bibliothèques standards de transformées de Fourier rapides (`fftw1d`, `fftw3d`, `pfft`, et `p3dfft`) déjà installées dans le cluster, ou dans le répertoire de travail de l'utilisateur. `Fluidfft` est écrit en C++, Cython et Python. Le code utilise `Pythran`, un compilateur Python qui produit des exécutables via C++11. Les fichiers de données sont sauvegardés au format `hdf5`. La bibliothèque `hdf5/1.12.0-mpi` est utilisée pour gérer l'enregistrement parallélisé des données de sortie.

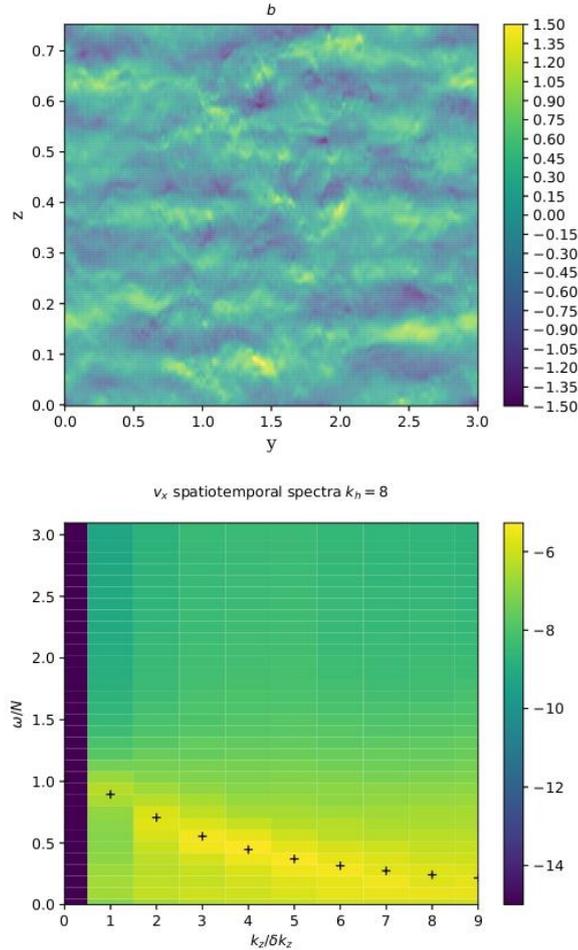


FIGURE 1 – Coupe verticale du champ de flottabilité (haut) et spectre spatio-temporel $k_n - \omega$ (bas) pour une simulation avec $N = 40$, $R_b = 40$ et la projection poloidale réalisée sur `Jean-Zay`. Les croix représentent la relation de dispersion théorique pour les ondes de gravité.

Beaucoup d'outils sont déjà implémentés dans le code : les projections sur la composante poloidale, les forçages des différents champs (différentes composantes de la vitesse et flottabilité), la suppression des modes de cisaillement, l'utilisation d'une hyperviscosité, et le traitement de données. En particulier, le code permet d'enregistrer les spectres spatiaux, temporels et spatio-

temporels (figure 1). Fluidsim a déjà utilisé sur le cluster Occigen (CINES) pour effectuer des calculs intensifs avec une bonne parallélisation jusqu'à $\sim 10^4$ processeurs pour une grille de 1152^3 . Il a été utilisé pour effectuer des simulations de turbulence stratifiée à 2D sur ce cluster [Calpe Linares, 2020]. Il est aussi utilisé sur le mésocentre CRIMSON de l'Observatoire de la Côte d'Azur et sur le cluster du LEGI.

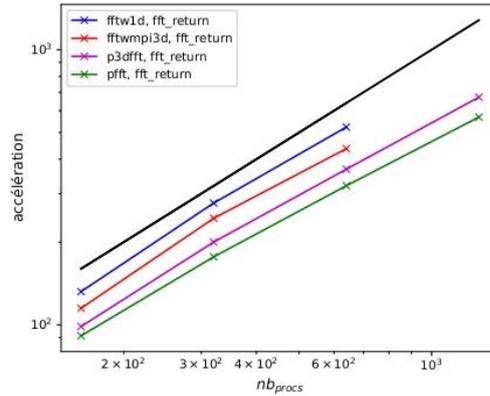


FIGURE 2 – Facteur d'accélération en fonction du nombre de processeurs pour réaliser les transformées de Fourier rapides par Fluidfft . Le test a été réalisé sur la machine Jean-Zay pour une grille de $1280 \times 1280 \times 640$ points et 4 bibliothèques : `fftw1d`, `fftw3d`, `pfft`, et `p3dfft`.

La figure 2 montre que les opérations de transformées de Fourier rapides sont bien parallélisées, avec une loi d'échelle quasi-optimale jusqu'à 640 – 1280 processeurs pour une grille de taille $1280 \times 1280 \times 640$. La bibliothèque `fftw1d` est la plus efficace, mais est limitée en termes de nombre de processeurs utilisables, car elle utilise une décomposition en tranche. Les bibliothèques avec une décomposition en crayon (`pfft` et `p3dfft`) sont plus lentes à nombre de processeurs égal, mais permettent d'utiliser plus de processeurs. La bibliothèque `p3dfft` est plus rapide que `pfft`, et est environ 1.5 fois plus lente que `fftw1d`. La figure 3 montre que le solveur utilisé pour la résolution des équations (`ns3d.strat`) présente des propriétés de parallélisation très similaires à celles de Fluidfft.

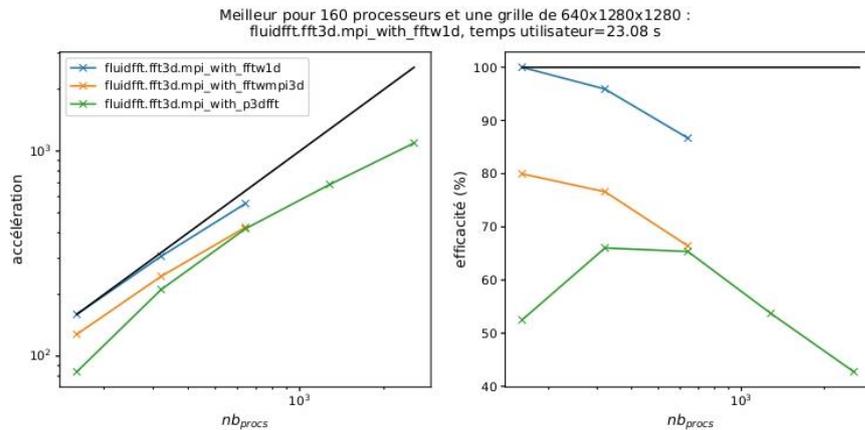


FIGURE 3 – Propriétés de parallélisation du solveur `ns3d.strat` de Fluidsim. Le test a été réalisé sur la machine Jean-Zay pour une grille de $1280 \times 1280 \times 640$ points et 100 pas de temps. À gauche : accélération relative du temps d'exécution t par rapport au cas de référence (1 noeud = 40 processeurs), t/t_0 . À droite : efficacité de la parallélisation $(tN_p)/(tN_p)_0$, où N_p est le nombre de processeurs utilisés.

3.3 Justification de l'usage des ressources sur les machine demandées

Ce projet fait l'objet d'une demande d'attribution d'heures sur la machine Jean-Zay (projet DARI AP012A13209). Afin de préparer au mieux l'utilisation du cluster national, nous souhaiterions disposer de ressources supplémentaires pour effectuer des simulations en faisant varier les paramètres de contrôle (les nombres sans dimensions Fr , Ro , et Re). Ces simulations auront pour but de calibrer les détails du forçage afin de favoriser l'observation du régime de turbulence d'ondes, de faire une première étude paramétrique du problème, et de générer des champs dans les états statistiquement stationnaires. Nous pourrions ensuite augmenter la résolution de ces champs pour relancer les simulations sur Jean-Zay. Nous avons déjà réalisé plusieurs simulations à une résolution de $640 \times 640 \times 160$ sur notre mésocentre.

Pour adresser les différents volets de ce projet, nous comptons réaliser les simulations suivantes :

1. Fluides stratifiés :

- Nous souhaitons réaliser des écoulements statistiquement stationnaires en utilisant la projection poloïdale. Nous voulons obtenir une vingtaine de simulations dans le cas purement stratifié avec une résolution de $640 \times 640 \times 160$. Le coup de chaque simulation est estimé à 10 k heures CPU.
- Nous voulons également réaliser des simulations des équations sans projection poloïdale, à faible Froude. Ces simulations sont importantes pour observer le comportement des ondes dans un écoulement plus réaliste. Le coup de chaque simulation est estimé à 20 k heures CPU.

Nous estimons le coût de cette partie à 600 k heures CPU.

2. Fluides en rotation :

- Nous souhaitons réaliser des écoulements statistiquement stationnaires en éliminant les colonnes de Taylor de la dynamique. Nous voulons réaliser une vingtaine de simulations dans le cas purement rotatif avec une résolution de $640 \times 640 \times 160$. Le coup de chaque simulation est estimé à 8 k heures CPU.
- Nous voulons également réaliser des simulations des équations complètes, à faible Rossby. Le coup de chaque simulation est estimé à 16 k heures CPU.

Nous estimons le coût de cette partie à 480 k heures CPU.

3. Dans un dernier temps, nous souhaiterions faire varier la rotation et la stratification, toujours à résolution modérée.

Nous estimons le coût de cette partie à 600 k heures CPU.

Nous souhaitons réaliser ces simulations conjointement sur la machine Azzurra et sur notre mésocentre (CRIMSON de l'Observatoire de la Côte d'Azur). Par conséquent, **nous demandons 500 k heures CPU sur la machine Azzurra.**

4 Plan de gestion de données

Pendant la phase de production, un volume de données sera généré, estimé à 5 To. Ces données correspondent aux champs de vitesse et de flottabilité, ainsi que les spectres spatiotemporels qui seront analysées en post-traitement. Une partie du post-traitement sera faite directement sur les machines de calcul, y compris la visualisation (si possible), et d'autres analyses numériques seront effectués localement dans notre mésocentre. Une fois que les données auront été traitées, ce volume pourra être réduit pour ne garder que les données les plus essentielles.

Références

- [Augier et al., 2021] Augier, P., Bolz-Tereick, C. F., Guelton, S., and Mohanan, A. V. (2021). Reducing the ecological impact of computing through education and Python compilers. *Nature Astronomy*, 5(4) :334–335.
- [Augier et al., 2019] Augier, P., Mohanan, A. V., and Bonamy, C. (2019). FluidDyn : A python open-source framework for research and teaching in fluid dynamics by simulations, experiments and data processing. *Journal of Open Research Software*, 7.

- [Brethouwer et al., 2007] Brethouwer, G., Billant, P., Lindborg, E., and Chomaz, J.-M. (2007). Scaling analysis and simulation of strongly stratified turbulent flows. *Journal of Fluid Mechanics*, 585 :343–368.
- [Caillol and Zeitlin, 2000] Caillol, P. and Zeitlin, V. (2000). Kinetic equations and stationary energy spectra of weakly nonlinear internal gravity waves. *Dynamics of Atmospheres and Oceans*, 32(2) :81–112.
- [Calpe Linares, 2020] Calpe Linares, M. (2020). *Numerical study of 2D stratified turbulence forced by internal gravity waves*. Theses, Université Grenoble Alpes [2020-....].
- [Davis et al., 2020] Davis, G., Jamin, T., Deleuze, J., Joubaud, S., and Dauxois, T. (2020). Succession of resonances to achieve internal wave turbulence. *Phys. Rev. Lett.*, 124 :204502.
- [Galtier, 2003] Galtier, S. (2003). Weak inertial-wave turbulence theory. *Phys. Rev. E*, 68 :015301.
- [Garrett and Munk, 1979] Garrett, C. and Munk, W. (1979). Internal waves in the ocean. *Annual Review of Fluid Mechanics*, 11(1) :339–369.
- [Gregg et al., 2018] Gregg, M., D’Asaro, E., Riley, J., and Kunze, E. (2018). Mixing efficiency in the ocean. *Annual Review of Marine Science*, 10(1) :443–473. PMID : 28934598.
- [Lelong and Riley, 1991] Lelong, M. P. and Riley, J. J. (1991). Internal wave—vortical mode interactions in strongly stratified flows. *Journal of Fluid Mechanics*, 232 :1–19.
- [Lvov et al., 2004] Lvov, Y. V., Polzin, K. L., and Tabak, E. G. (2004). Energy spectra of the ocean’s internal wave field : Theory and observations. *Phys. Rev. Lett.*, 92 :128501.
- [MacKinnon et al., 2017] MacKinnon, J. A., Zhao, Z., Whalen, C. B., Waterhouse, A. F., Trossman, D. S., Sun, O. M., Laurent, L. C. S., Simmons, H. L., Polzin, K., Pinkel, R., Pickering, A., Norton, N. J., Nash, J. D., Musgrave, R., Merchant, L. M., Melet, A. V., Mater, B., Legg, S., Large, W. G., Kunze, E., Klymak, J. M., Jochum, M., Jayne, S. R., Hallberg, R. W., Griffies, S. M., Diggs, S., Danabasoglu, G., Chassignet, E. P., Buijsman, M. C., Bryan, F. O., Briegleb, B. P., Barna, A., Arbic, B. K., Ansong, J. K., and Alford, M. H. (2017). Climate process team on internal wave-driven ocean mixing. *Bulletin of the American Meteorological Society*, 98(11) :2429–2454.
- [Maffioli, 2017] Maffioli, A. (2017). Vertical spectra of stratified turbulence at large horizontal scales. *Phys. Rev. Fluids*, 2 :104802.
- [Medvedev and Zeitlin, 2007] Medvedev, S. and Zeitlin, V. (2007). Turbulence of near-inertial waves in the continuously stratified fluid. *Physics Letters A*, 371(3) :221–227.
- [Mohan et al., 2019a] Mohan, A. V., Bonamy, C., and Augier, P. (2019a). FluidFFT : Common API (c++ and python) for fast fourier transform HPC libraries. *Journal of Open Research Software*, 7.
- [Mohan et al., 2019b] Mohan, A. V., Bonamy, C., Linares, M. C., and Augier, P. (2019b). FluidSim : Modular, Object-Oriented Python Package for High-Performance CFD Simulations. *Journal of Open Research Software*, 7.
- [Monsalve et al., 2020] Monsalve, E., Brunet, M., Gallet, B., and Cortet, P.-P. (2020). Quantitative experimental observation of weak inertial-wave turbulence. *Phys. Rev. Lett.*, 125 :254502.
- [Nazarenko, 2011] Nazarenko, S. (2011). *Wave Turbulence*. Lecture Notes in Physics. Springer Berlin Heidelberg.
- [Polzin and Lvov, 2011] Polzin, K. L. and Lvov, Y. V. (2011). Toward regional characterizations of the oceanic internal wavefield. *Reviews of Geophysics*, 49(4).
- [Pouquet et al., 2017] Pouquet, A., Marino, R., Mininni, P. D., and Rosenberg, D. (2017). Dual constant-flux energy cascades to both large scales and small scales. *Physics of Fluids*, 29(11) :111108.
- [Waite and Bartello, 2006] Waite, M. L. and Bartello, P. (2006). Stratified turbulence generated by internal gravity waves. *Journal of Fluid Mechanics*, 546 :313–339.
- [Zakharov et al., 2012] Zakharov, V., L’vov, V., and Falkovich, G. (2012). *Kolmogorov Spectra of Turbulence I : Wave Turbulence*. Springer Series in Nonlinear Dynamics. Springer Berlin Heidelberg.

Optique

EUR Spectrum

Comportement spatial de la propagation d'une pompe hautement multimode dans des amplificateurs et lasers à fibre

Responsable scientifique : Valérie Doya

Laboratoire : INPHYNI

Description :

Les lasers à haute puissance sont aujourd'hui utilisés dans une grande variété d'applications industrielles et environnementales. Afin d'améliorer leur robustesse et leur efficacité, ces systèmes sont maintenant majoritairement fibrés, c'est-à-dire que l'ensemble des éléments qui composent le laser est contenu dans une fibre optique. Le milieu amplificateur est obtenu par un dopage aux ions terres rares de la fibre et les miroirs (appelé réseaux de Bragg) qui constituent la cavité sont en général photo-inscrits dans le cœur de la fibre. Afin d'optimiser le schéma de pompage du laser, des réflecteurs sont en général employés pour recycler la pompe dans la cavité laser. L'un des enjeux actuels est de pouvoir les fabriquer directement dans la gaine de la fibre optique. L'aspect fortement multimode de la pompe se propageant dans la gaine rend l'interaction avec un réseau de Bragg non triviale. Les relations couramment employées dans le cadre d'un faisceau monomode ne sont plus valables et une connaissance fine du contenu modale de la pompe devient alors indispensable.

L'objectif de ce projet est donc de calculer numériquement l'évolution du signal de pompe au sein d'une fibre optique amplificatrice afin de pouvoir optimiser le design d'un réflecteur de pompe fibré. La géométrie du système et le schéma numérique employé nécessite actuellement un temps de calcul de plusieurs semaines empêchant toute étude paramétrique. Parallèlement à un travail d'optimisation du code, nous sollicitons les ressources dédiées du centre de calcul pour diminuer notablement le temps de calcul.

Ce projet s'inscrit dans le cadre d'une collaboration actuelle avec le Centre Optique Photonique et Laser (Université Laval, Québec) et TeraXion (entreprise québécoise spécialisée dans la fabrication de composants et lasers fibrés). Nous avons aussi amorcé une collaboration avec OptiWave, une entreprise canadienne qui édite un logiciel de calcul optique dédié à l'interaction laser/réseaux de Bragg.

Light scattering in disordered systems with correlations

Responsable scientifique : Geoffroy Aubry

Laboratoire : INPHYNI

Résultats obtenus grâce aux moyens informatiques attribués :

Yazhgur, P., Aubry, G. J., Froufe-Pérez, L. S., & Scheffold, F. (2022). Scattering phase delay and momentum transfer of light in disordered media. *Physical Review Research*, 4(2), 023235.

<https://doi.org/10.1103/PhysRevResearch.4.023235>

Description :

In this project, I plan to realize different kinds of numerical simulations in order to better understand the influence of correlated disordered dielectrics on the propagation of electromagnetic waves (light) in these systems.

Disordered dielectrics like clouds, foams, powders, etc. appear white because light propagate into it not in a straight line, like in vacuum, but get scattered, ie., changes its propagation direction much more than once in the medium. This explains the reason why the sun position can't be seen on a cloudy day, although the sun light still arrive to our eyes. In a cloud, there is, a priori, no correlation between the scatterers (i.e. the water droplets) positions: this medium is referred to as an "uncorrelated disordered system". In a "correlated disordered system", the positions of the scatterers are not completely independent from each other. The limit of this class of system are photonic crystals, where all the positions of the scatterers are linked to each other, and no disorder is present anymore.

Using correlated disorder, it is possible to tune light transport in order to observe a wide variety of regimes: from transparency, to light hindered to enter the medium (like in a photonic bandgap) including light diffusion or even light stuck in the medium, the so called Anderson localized regime.

In my previous work done during my postdocs in Konstanz (Germany) and Fribourg (Switzerland), I already did some experimental and numerical work on these kinds of systems (see selected references at the end of this paragraph). As a new permanent

researcher in the Institut de Physique de Nice, I plan to continue these studies that are of great interest for applications such as supercontinuum fiber lasers and for data processing using optical modules based on 2D silicon photonic crystal technology. Technological applications for 3D PBG bandgap materials offer great potential in lasing, wave-guiding, light-harvesting and even optical computers.

Selected references:

Aubry, G. J.; Schertel, L.; Chen, M.; Weyer, H.; Aegerter, C. M.; Polarz, S.; Cölfen, H. & Maret, G. Resonant transport and near-field effects in photonic glasses

Phys. Rev. A 96, 043871 (2017) <http://dx.doi.org/10.1103/PhysRevA.96.043871>

Aubry, G. J.; Froufe-Pérez, L. S.; Kuhl, U.; Legrand, O.; Scheffold, F. & Mortessagne, F.

Experimental Tuning of Transport Regimes in Hyperuniform Disordered Photonic Materials Phys. Rev. Lett. 125, 127402 (2020) <http://dx.doi.org/10.1103/PhysRevLett.125.127402>

Aeby, S.; Aubry, G. J.; Muller, N. & Scheffold, F. Scattering from controlled defects in woodpile photonic crystals Advanced Optical Materials, 9, 2001699 (2021)

<http://dx.doi.org/10.1002/adom.202001699>

Yazhgur, P.; Aubry, G. J.; Froufe, L. S. & Scheffold, F. Light scattering from colloidal aggregates on a hierarchy of length scales Opt. Express, 29, 14367 (2021)

<http://dx.doi.org/10.1364/OE.418735>

Dynamique des fluides

EUR Spectrum

Dynamique de polymères en turbulence et turbulence élastique

Responsable scientifique : Dario Vincenzi

Laboratoire : LJAD

Résultats obtenus grâce aux moyens informatiques attribués :

Caractérisation entropique de la transition coil-stretch dans des écoulements laminaires ou turbulents.

Collaboration avec Stefano Musacchio (Université de Turin, Italie) et Victor Steinberg (Weizmann Institute, Israël)

Le papier a été soumis à Physical Review Fluids et est disponible sur arXiv à l'adresse :

<https://arxiv.org/pdf/2112.01344.pdf>

Description :

Le projet est composé de deux volets. Le premier porte sur la dynamique lagrangienne de polymères élastiques dans un écoulement turbulent isotrope et sur la transition coil-stretch. Nous étudierons l'effet des interactions hydrodynamiques et de volume exclu. Le calcul numérique consiste en la simulation d'équations différentielles stochastiques et en l'analyse statistique des solutions.

Le deuxième volet porte sur la turbulence élastique, un régime chaotique que l'on observe dans les solutions de polymères visqueuses. Nous effectuerons des simulations numériques pseudo-spectrales du modèle FENE-P, qui décrit l'évolution couplée du tenseur de conformation des polymères et du champ de vitesse (donnée par les équations de Navier–Stokes). Les simulations seront d'abord effectuées en deux dimensions dans un carré avec conditions de bord périodiques. Le projet à plus long terme est de réaliser des simulations en trois dimensions.

Géophysique interne

EUR Spectrum

Dynamics of large earthquakes

Responsable scientifique : Jean-Paul Ampuero

Laboratoire : Géoazur

Description :

Earthquakes are a fundamental component of plate tectonics, one of the main processes that shape the surface of our planet, and one of the most damaging natural hazards.

Elucidating the physics of earthquakes remains a grand challenge. Thanks to recent advances in data gathering, computational seismology, laboratory rock mechanics and theoretical fault mechanics, the earthquake research community has made major advances towards answering first-order questions such as: How do earthquakes start? What determines the wide spectrum of fault slip behavior? What controls the speed and complexity of rupture propagation? How do seismic and aseismic slip interact? However, many key aspects of these questions still remain unsolved.

My team's research at Geoazur aims at answering these open questions by developing a fundamental understanding of the mechanical behavior of tectonic faults, across temporal and spatial scales spanning rupture dynamics, earthquake cycles, and the evolution of fault zone structure. Our vision is to achieve this goal by combining state-of-the-art theoretical modeling, high-performance computing and novel geophysical observation techniques. In this computational project, we aim at advancing our understanding of earthquake cycles and rupture dynamics on very large and segmented faults.

Our main recent achievement in earthquake theory was to establish an "equation of motion" for large earthquakes, whose rupture spans the whole seismogenic depth and then propagates horizontally. The distinctive feature of these ruptures is their large aspect ratios. We found through computational experiments that such elongated ruptures have radically different behavior than predicted by previous theories, which were limited to ruptures in 2D and to almost-circular ruptures in 3D. We thus developed a new theory of the evolution of elongated ruptures (Weng and Ampuero, *J. Geophys. Res.*, 2019), leading to an equation of motion and a "rupture potential", a quantity that

encapsulates the most important factors controlling rupture arrest. If we are able to estimate this rupture potential along a fault, we can map the fault areas that are most prone to generate a large earthquake. Motivated by results of our high-performance computer simulation experiments, we further developed the implications of the new theory to understand what controls the rupture speed of earthquakes (Weng and Ampuero, *Nature Geoscience*, 2020), a parameter that strongly affects their damaging potential. This work reconciled earthquake theory with the puzzling observation of earthquakes propagating at speeds that are forbidden in the previous 2D theory of dynamic fracture.

In this project, we propose to extend the theory of elongated ruptures in complementary directions that will allow us to extract more implications for real earthquakes:

- We will study the effects of heterogeneities of stress and strength on the dynamics of elongated ruptures. Our current theory accounts already for lateral heterogeneities, but we still need to consistently include depth-dependent heterogeneities. With this capability, we will study how systematic and stochastic heterogeneities affect the statistical properties of earthquakes, such as scaling laws and the variability of source time functions.
- We will extend the theory to account for variable rupture width in the theory of elongated ruptures. In long asperities surrounded by creep, dynamic rupture can penetrate inside the creeping region and thus the rupture width is not fixed. The idea we will pursue is to use 1D fracture mechanics problems on depth cross-sections to solve for the rupture width, then use that laterally variable width in the equation of motion of elongated ruptures. We will validate this approach by comparison to results of 3D dynamic rupture simulations.
- We will conduct 2.5D earthquake cycle simulations on long faults with heterogeneous frictional properties. We will use these simulations as a testbed to validate the developments above. We will further use the modified theory to gain insights on the predictability of earthquake rupture. In particular, we will search for practical proxies of the spatio-temporal evolution of the rupture potential which could be used to forecast the location, size and timing of the next large earthquake in our (idealized) earthquake cycle models. Those proxies are quantities such as the geodetic coupling rates.

• Large faults are divided into segments that can break separately or together. Because the number of broken segments controls the final earthquake size, it is essential to understand the mechanical interactions between fault segments. Can we relate the probability of multi-segment ruptures to observable fault properties? Can we distinguish between temporary and permanent barriers? Can we anticipate the size of the next large earthquake on a given fault? We will conduct a parametric study based on 2.5D earthquake cycle modeling to quantify the efficiency of fault barriers to stop ruptures, as a function of their length and frictional strength.

The 3D dynamic rupture simulations will use the SPECFEM3D software, in which our team leads the development of earthquake dynamics simulation capabilities. The 2.5D earthquake cycle simulations will use the SEM2DPACK software, fully developed by our team.

References:

H. Weng and J. P. Ampuero (2020), Continuum of earthquake rupture speeds enabled by oblique slip, *Nature Geo.*, 13, 817–821, doi:10.1038/s41561-020-00654-4

H. Weng and J. P. Ampuero (2019), The dynamics of elongated earthquake ruptures, *J. Geophys. Res.*, 124 (8), 8584–8610, doi: 10.1029/2019JB017684

Continental Transpression: Deciphering Deformation Digitally (CT3D)

Responsable scientifique : Guillaume Duclaux

Laboratoire : Géoazur

Résultats obtenus grâce aux moyens informatiques attribués :

Les modèles thermomécaniques que nous avons fait tourner sur Azzurra nous ont permis de tester le parallélisme de nouvelles options du code.

Le travail de thèse de Josselin Gremmel sur la tectonique transpressive progresse et l'an prochain ses premiers résultats devraient être publiés. Cette année, une contribution détaillant les résultats préliminaires des modélisations 3D réalisées sur Azzurra a été présentée par Josselin à l'EGU (European Geosciences Union) à Vienne (Autriche). La référence est ci-dessous :

- Gremmel, J., Duclaux, G., Corsini, M., Maillet, A., Jourdon, A., & Bascou, J. (2022). Structure, strain partitioning and exhumation mechanism during the late stage oblique tectonic evolution of the Variscan Tanneron massif (SE France) (No. EGU22-967). Copernicus Meetings.

Description :

Ce projet propose d'étudier le partitionnement de la déformation dans les systèmes continentaux en transpression à l'aide de modèles numériques thermomécaniques en trois dimensions (3D). Cette approche computationnelle nous permettra 1) de mieux comprendre la manière dont les fabriques des roches métamorphiques apparaissent et sont reliées les unes aux autres dans l'espace et dans le temps, et 2) de déterminer le niveau de couplage entre les différents niveaux structuraux de la croûte continentale dans les systèmes transpressifs. Dans les domaines orogéniques transpressifs anciens que nous connaissons, les roches ont subi une histoire de déformation polyphasée, résultant de la compétition des forces tectoniques et de volume évoluant au cours du temps, qui conduit au développement de fabriques finies complexes et variées à toutes les échelles. Dans les niveaux supérieurs de la croûte continentale on observe des systèmes de failles qui forment des structures en fleur, de zones de relais transpressifs et transtensionnels, ainsi que des bassins étroits dont la géométrie semble fortement influencée par les fabriques pré-existantes. Dans la croûte intermédiaire à profonde la déformation est marquée par le développement de grandes zones de cisaillement ductiles verticales limitant des dômes marqués par un aplatissement sub-horizontale et un étirement sub-

parallèle aux zones de cisaillement. Des plis d'axes obliques à la direction de raccourcissement principal se forment par fluage longitudinal synchrone de la transpression. Les questions que nous souhaitons adresser peuvent-être résumés comme suit : Comment ces fabriques dans les roches métamorphiques sont-elles reliées les unes aux autres dans l'espace et dans le temps ? Quel niveau de couplage existe entre les différents niveaux structuraux dans les systèmes orogéniques transpressifs ?

L'évolution de la nature et de l'orientation des structures au sein de la croûte sera modélisée à partir d'un code thermomécanique 3D (pTatin 3D : *May et al., 2014*³⁴ ; *Duclaux et al., 2019*³⁵) et analysée en post-processing à partir de scripts Python que nous avons développés. Ces travaux répondent également à un enjeu sociétal majeur et visent à améliorer notre compréhension du couplage mécanique entre les niveaux profonds des grandes zones de cisaillement ductiles et les systèmes de failles décrochantes et décro-chevauchantes sismogènes observées en surface dans les zones de convergence obliques.

³⁴ *Duclaux, G., Huismans, R. S., & May, D. A. (2020). Rotation, narrowing, and preferential reactivation of brittle structures during oblique rifting. Earth and Planetary Science Letters, 531, 115952.*

³⁵ *May, D. A., Brown, J., & Le Pourhiet, L. (2014, November). pTatin3D: High-performance methods for long-term lithospheric dynamics. In SC'14: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (pp. 274-284). IEEE.*

WIND (Waveform Inversion of Node Data)

Responsable scientifique : Stéphane Operto

Laboratoire : Geoazur

Résultats obtenus grâce aux moyens informatiques attribués :

Investigation d'une approche d'intelligence artificielle pour la résolution d'équations aux dérivés partielles (EDP) dans notre cas l'équation d'onde pour des applications d'imagerie sismique.

L'approche explorée est fondée sur du Physics Informed Neural Network (PINN).

Un article accepté dans Geophysics sous réserve de modifications majeures (en cours de traitement). Les améliorations demandées devraient être abordables si bien que la publication de cet article est envisageable.

Wave simulation in non-smooth media by PINN with quadratic neural network and PML condition

Y. Wu, H. Aghamiry, J. Ma and S. Operto, Geophysics, accepted pending major revisions.

Y. Wu a aussi présenté ses travaux lors du workshop organisé par Wave Complexity (10-12 Octobre) :

Yanqi Wu, Hossein S. Aghamiry, Stéphane Operto, Jianwei Ma. Wave simulation in non-smooth media by PINN with quadratic neural network and PML condition. Artificial Intelligence for Waves - AI4W 2022. https://wavecomplexity.univ-cotedazur.fr/events/ai4w2022/contribs/yanqi_wu/

Description :

Présentation générale du projet

Ce projet porte sur le développement d'une méthode d'imagerie sismique par inversion des formes d'ondes complètes (FWI: Full Waveform Inversion). La FWI consiste à reconstruire les propriétés mécaniques du sous-sol gouvernant la propagation des ondes en résolvant un problème d'optimisation sous contrainte par des approches locales (méthodes de gradient). Le problème direct résout l'équation d'onde, une équation aux dérivés partielles (EDP) linéaire dont la solution et le second membre sont respectivement le champ d'onde et la source. Les observables sont les enregistrements des champs d'onde aux capteurs. La FWI minimise une distance entre ces observables et leur analogues simulés numériquement pour mettre à jour les paramètres contenus dans les coefficients de l'EDP.

Les applications sismiques concernent (1) l'imagerie des bassins sédimentaires pour l'exploration pétrolière (projet WIND financé par un consortium industriel) et (2) l'imagerie de la croûte terrestre des zones de subduction (chantiers Geoazur). Ces deux applications sont effectuées à partir d'acquisitions de fond de mer mises en oeuvre avec des sources artificielles et des câbles sous-marins (OBC) ou des stations sous-marines multi-composantes (OBSs). En sismologie (sources passives), l'imagerie de la croûte et du manteau supérieur des Alpes à partir de téléseismes (projet LisAlps financé par l'ANR).

Nous implémentons la FWI dans les domaines temps-espace (codes GeoInv3D_fwkw et LITHOS) et fréquence-espace (code FFWI). Le problème direct y est respectivement décrit par un problème d'évolution avec conditions initiales et un problème de conditions aux limites nécessitant la résolution de systèmes d'équations linéaires creux à seconds membres multiples. Dans ce dernier cas, nous utilisons le solveur multifrontal MUMPS et un solveur hybride direct/itératif préconditionné par la méthode de décomposition en domaine ORAS (Optimized Restricted Additive Schwarz). Suivant les codes et les applications, nous utilisons des schémas de discrétisation fondés sur des différences finies, des éléments finis sur maillage tétraédrique et des éléments spectraux.

Les codes GeoInv3D_fwkw, LITHOS et FFWI sont parallélisés avec MPI et OpenMP et ont été exécutés en production sur le mésocentre CRIMSON, sur Occigen et sur Jean Zay. Ils sont basiquement écrits en Fortran 90 avec des ingénieries de C++. Nos travaux récents portent sur une reformulation de la FWI visant à étendre son régime linéaire via une méthode de Lagrangien augmenté et l'interfaçage de régularisations hybrides pour réduire son espace nul. Ces travaux ont été implémentés sous Matlab et doivent être transférés dans les trois codes pour tester les concepts sur des cas d'étude de grande dimension.

Demande sur Azzura

La présente demande portant spécifiquement sur l'accès à Azzura concerne deux aspects:

- Le premier concerne le portage progressif du code GeoInv3D_fwkw sur GPU.
- Le deuxième concerne une application d'intelligence artificielle fondée sur des "physics-informed neural networks". Dans le cadre de la thèse de Yanqi WU à Geoazur, nous souhaitons commencer à évaluer si la résolution de l'équation d'onde (le problème direct) et la mise à jour des paramètres constitutifs du milieu (problème inverse) peuvent être

effectués avec des approches de deep learning où la fonction cout est respectivement formée par la norme L2 de l'erreur sur l'équation d'onde (problème direct) et une fonction de pénalité ou de Lagrangien augmenté combinant la norme L2 de l'écart entre les données observées et simulées et la norme L2 de l'erreur sur l'équation d'onde. Des tests préliminaires ont été effectués sur les CPUs d'un serveur local de Geozur. Néanmoins, nous avons maintenant besoin d'accéder à des GPUs pour traiter des applications de dimension plus représentatives.

References sur PINN (Physics-Informed Neural Networks with applications to wave equation solution and seismic waveform inversion)

M. Raissi, P. Perdikaris and G. Karniadakis, Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, *Journal of Computational Physics*, 378, 686--707, 2019.

C. Song and T. Alkhalifah and U. B. Waheed, Solving the frequency-domain acoustic VTI wave equation using physics-informed neural networks, *Geophysical Journal International*, 225(2), 846--859, 2021.

C. Song and T. Alkhalifah, Wavefield reconstruction inversion via machine learned functions, *SEG Technical Program Expanded Abstracts 2020*, 1710-1714, 2020.

Study of the rheological controls on vertical sinking of subducting slabs through numerical modeling

Responsible scientifique : Karin Sigloch

Laboratoire : Geoazur

Résultats obtenus grâce aux moyens informatiques attribués :

1) We have a nice article in review in Journal of Geophysical Research, titled :

The effect of a weak asthenospheric layer on surface kinematics, subduction dynamics and slab morphology in the lower mantle

Nestor Cerpa, Karin Sigloch, Fanny Garel, Arnauld Heuret, D. Rhodri Davies, Mitchell G. Mihalynuk

Arxiv preprint here: <https://eartharxiv.org/repository/view/3230/>

2) The results were presented at the Annual Meeting of the European Geosciences Union in Vienna in May 2022: Subduction dynamics through the mantle transition zone in the presence of a weak asthenospheric layer

Nestor Cerpa, Karin Sigloch, Fanny Garel, Rhodri Davies, and Arnauld Heuret

Abstract here:

<https://meetingorganizer.copernicus.org/EGU22/EGU22-3822.html>

Description :

This is a numerical investigation of the modes in which the solid earth recycles material and heat between its surface and its deep interior, on geological time scales. Specifically, we study whether thermo-mechanical modeling of the subduction of tectonic plates into the earth's mantle can match recent geometric constraints on the deformed plate geometries in the Earth's mantle, as observed by geophysical imaging. We use the fluid-dynamics finite-element code Fluidity [Davies et al., 2011], which solves for the conservation of mass and momentum in 2-D geometries, with non-linear rheologies suitable for the study of rock deformation. The code uses an adaptive remeshing technique to tackle the strong viscosity contrasts that characterize geodynamical flows (up to 4 orders of magnitudes). Tectonic plates form in ocean basins from upwelling mantle rock, and spend several tens of millions of years cooling at the surface. Then these ocean plates dive back into the earth's mantle at subduction zones and slowly sink towards the core, cooling the ambient mantle in the process. This material cycle, which makes the earth the only habitable planet in the solar system, is thought to be driven by the "slab

pull" force, i.e., the excess density of plate margins that have entered a subduction zone relatively recently, and which keep pulling down after them the tectonic surface plate they are still attached to [Forsyth and Uyeda, 1975].

Although geodynamic conceptual and numerical models often show these slabs as remaining plate-like, thin sheets to great mantle depths, geophysical imaging (seismic 3-D tomography) is increasingly clear that from relatively shallow depths of only a few hundred kilometers, the slabs appear thickened to several times lithospheric thickness. This is a fundamental geometry difference that implies completely different rheological, advection and cooling characteristics for the earth's mantle. Presumably the thickening occurs through slab folding, as seen in Figs. 2&3, although these spatial details are currently not resolvable by geophysical imaging. In order to constrain the plausible range of rheologies and deformation styles, we add a number of observations from plate kinematic and from the land geology of accretionary mountain belts, which form the surface record of subduction.

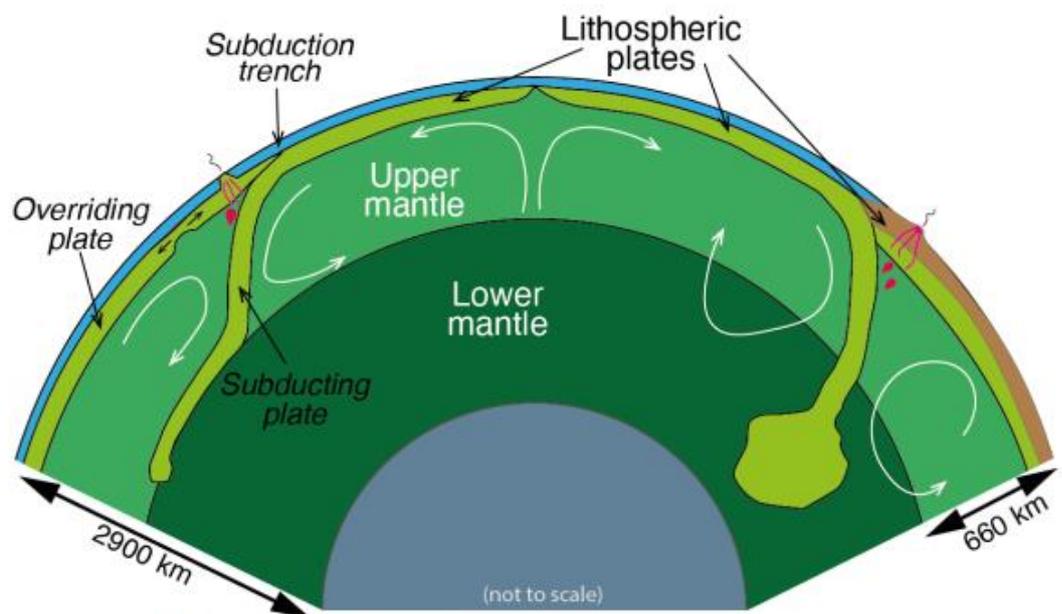


Figure 1 : Schematic view of two subduction zones.

Over the past decade, such "tomo-tectonic" analysis has revealed that many past subduction zones were quasi-stationary in an absolute reference frame [Sigloch and Mihalynuk, 2013; Mohammadzaheri et al., 2021]. The tomographic images also showed that stationary trenches are associated with wide "slab walls" in the lower mantle (400-700-km

wide), a special-case geometry that rather strongly points to slab thickening (folding?) at the mantle transition zone (Fig. 2). In contrast to these observations, state-of-the-art numerical models of subduction zone dynamics suggest that non-stationary trenches (slab-rollback and subduction-trench retreat) are the most common dynamical response of slab sinking in the mantle (fig. 3, [Garel et al., 2014]). Hence, a strong conceptual discrepancy has recently opened between observational constraints and current numerical geodynamic models, which has not been addressed quantitatively. Here we want to close this gap and numerically explore the mantle and slab properties that would result in thickened slabs under stationary trenches, while also reproducing realistically fast subduction rates.

This work is conducted as part of the PI's ERC Starting Grant "DEEPTIME" (639003). In May 2021, I moved from the University of Oxford to a Directeur de Recherche position at the CNRS and Geoazur. My ERC grant has moved with me to the CNRS, as has the postdoctoral associate who has been working on the sub-project presented here, Dr. Nestor Cerpa. With this request for computing time on Azzurra, we are hoping to replace the institutional computing resources that were available to us in Oxford, in order to successfully conclude and publish our investigation.

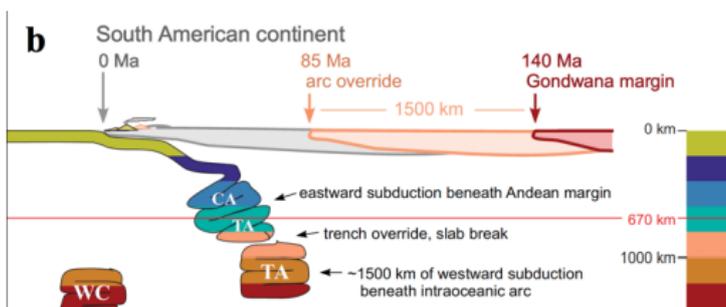


Figure 2 : Tomo-tectonic interpretation of the evolution of the South American subduction zone [from Mohammadzaheri et al. 2021].

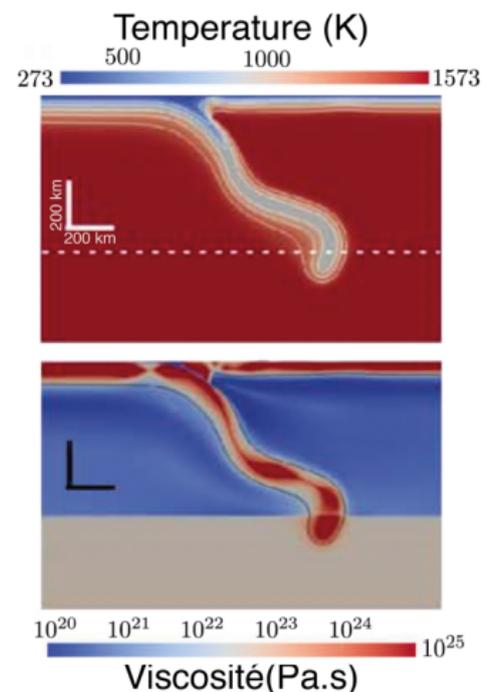


Figure 3 : Two-dimensional numerical model of a subduction zone [from Garel et al., 2014].

References

- Davies, D. R., Wilson, C. R., & Kramer, S. C. (2011). Fluidity: A fully unstructured anisotropic adaptive mesh computational modeling framework for geodynamics. *Geochemistry, Geophysics, Geosystems*, 12(6).
- Forsyth, D, and Uyeda, S. (1975). On the relative importance of the driving forces of plate motion. *Geophysical Journal International* 43, no. 1: 163-200.
- Garel, F., Goes, S., Davies, D. R., Davies, J. H., Kramer, S. C., & Wilson, C. R. (2014). Interaction of subducted slabs with the mantle transition - zone: A regime diagram from 2 - D thermo - mechanical models with a mobile trench and an overriding plate. *Geochemistry, Geophysics, Geosystems*, 15(5), 1739-1765.
- Mohammadzaheri, A., Sigloch, K., Hosseini, K., & Mihalynuk, M. G. (2021). Subducted Lithosphere under South America from Multi - frequency P - wave Tomography. *Journal of Geophysical Research: Solid Earth*, 126, e2020JB020704.
- Sigloch, K., & Mihalynuk, M. G. (2013). Intra-oceanic subduction shaped the assembly of Cordilleran North America. *Nature*, 496(7443), 50-56

Imaging high-frequency radiation of large earthquakes

Responsable scientifique : Anthony Sladen

Laboratoire : Géoazur

Projet en partenariat avec l'[Universidad de Chile \(FCSM\)](#).

Résultats obtenus grâce aux moyens informatiques attribués :

- Carolina Paz Gutiérrez Morales, Inversión bayesiana en altas frecuencias de la distribución de deslizamiento del terremoto de Tocopilla 2007, Mw 7.7. Rapport de master « Magíster en Ciencias, Mención Geofísica », Universidad de Chile (2022).
<https://repositorio.uchile.cl/handle/2250/187604>
- Imaging High-Frequency Radiation of Large Earthquakes with a Bayesian Inversion Method. Poster présenté par Carolina Paz Gutiérrez Morales à the IV Assembly of the Latin American and Caribbean Seismological Commission, Quito, Ecuador, Oct. 2022.

Description :

A project led by Carolina Gutierrez, research engineer at University of Chile (Santiago), and supervised by [Sergio Ruiz](#), Professor at University of Chile (Santiago), and [Anthony Sladen](#) CNRS researcher at University Côte d'Azur.

Characterizing the high-frequency distribution of earthquakes is critical for earthquake hazard assessment and rapid response

The high-frequency radiation generated during subduction earthquakes is a process that is still poorly understood in seismology. Because this type of emission causes the most damage in the near-field it is essential to understand its behavior to assess the seismic hazard in a specific region correctly. Multiple methodologies have been developed to model the strong-ground motion of large magnitude earthquakes, including the use of stochastic Green functions. This approach successfully simulates high frequencies (> 1 Hz) in the borehole and on the ground surface incorporating soil amplification transfer functions to account for the amplifying effect of the soil. Such methodology opens the way for improved and rapid assessment of earthquake damage, and a better understanding of the underlying physical processes.

Our modeling approach

Most studies of fault inversion use low-frequency data, leaving out the short-period emission of the rupture process. In this project, we use high-frequency acceleration records in the near-field, filtered between 0.05 and 20 Hz, to invert the rupture model for two large events, the 2007 Tocopilla earthquake (Mw 7.7) and the 2017 Valparaíso earthquake (Mw 6.9). Due to the high-frequency content of the data, it is difficult to simulate this type of motion with a deterministic approach, hence we use a stochastic simulation technique proposed by Otarola and Ruiz (2016), based on Boore's method from 1983. This method generates synthetic accelerograms for P, SV, and SH waves from a finite fault model where each sub-fault is a point source. For each point source, a window of Gaussian white noise is modulated following Brune's w source spectrum. Then in the time domain, they are rotated and added to obtain the time-series acceleration in the EW, NS, and UD components.

Solving the inverse problem

The inverse problem of determining the rupture model given the acceleration data is nonlinear and ill-conditioned, hence we use a Bayesian methodology to be able to explore the full parameter space and obtain the ensemble of probable models. To sample the space efficiently we use a TMCMC sampler.

Results obtained for the 2007 Tocopilla earthquake (Mw 7.7)

For the Tocopilla earthquake, we model four acceleration records from stations in boreholes, using a parametrization of two circular asperities for the source model. With the methodology described above, we infer the most probable rupture model given the observed data. The methodology is first validated with synthetic tests and then applied to the observed data of the Tocopilla earthquake. The results indicate that our methodology allows us to properly estimate a simple rupture consisting of two circular asperities with a similar radius of 40-50 km. The center of the first asperity was imposed at the hypocenter location, but the location of the second one is properly located beneath the Mejillones peninsula (Figure 1). The predicted accelerations generated from this model fit the data well, especially in the period range between 0.2 and 1 s (Figure 2).

Results obtained for the 2017 Valparaíso earthquake (Mw 6.9)

For the Valparaíso earthquake, we model records from four stations in boreholes and three stations on the ground surface, where we use soil amplification transfer functions based on the local velocity profile. We use a parametrization of a single circular asperity for the source model and infer the parameters that define the most probable rupture model given the data. The current results show a rupture model consisting of a circular asperity with a 14.5 km, fixed at the hypocenter

and with a maximum slip of 2.1 m (Figure 3). The predicted accelerations generated from this model show a better fit in the EW and UD components, reflected in the goodness of fit in Figure 4. For this event, we have more stations, which are located at different distances from the source. Achieving a good fit for the near-field and distant stations is challenging, considering the same high-frequency decay behavior. This can be partially solved by including 2 different gamma parameters to model the decay of high frequencies in Brune's spectral model. We are currently working on incorporating new approaches to improve the results presented.

What has been achieved thanks to Azzurra and the way forward

All the results shown here, as well as the synthetic tests, were obtained using Azzurra's resources. These have been essential in developing and completing my Master's thesis "High-frequency bayesian inversion of the slip distribution of the 2007 Tocopilla earthquake, Mw 7.7" at the University of Chile, and have allowed us to expand the analysis and apply this methodology to the Valparaíso earthquake.

An important aspect of the methodology of this project is the high computational cost of the forward model. For the two modeled earthquakes, simulating the acceleration data takes between 1-3 seconds. During the inverse problem, the sampler runs $\sim 10^4$ parallel chains to obtain the ensemble of models from the posterior distribution. At the end of the inversion, the forward model has been executed approximately 3×10^6 times, resulting in a total execution time of $\sim 2-3$ days. Nonetheless, this amount of chains is not enough to correctly represent the posterior distribution of rupture models. We need to solve the inverse problem with a larger amount of chains, closer to 10^6 to achieve a more robust solution.

Our objective is to publish the results of this project in an article and present it at Seismology conferences. We have submitted an abstract to present this project at the Latin American and Caribbean Seismological Commission IV Assembly, which will take place in Quito, Ecuador, from October 3rd to October 5th, 2022.

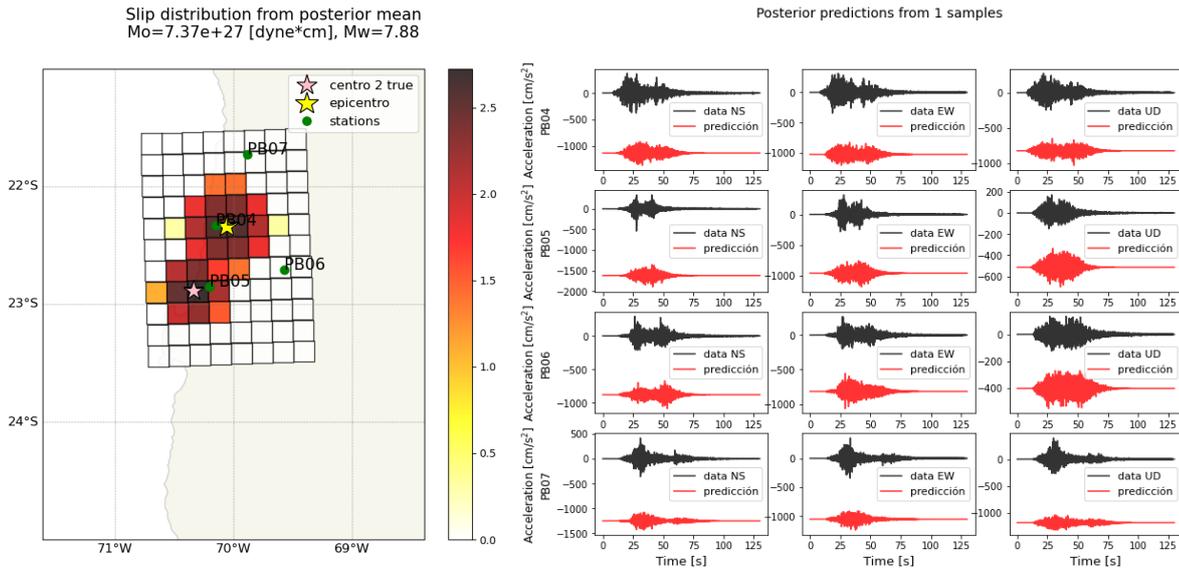


Figure 1: a) Estimated rupture model for the Tocopilla earthquake. The color scale shows the slip distribution for the two circular asperities. The north asperity is fixed in the hypocenter represented by the yellow star. The pink star is the estimated location of the center of the second asperity, right above the Mejillones peninsula. The estimated radii are 49 km and 38 km, and the maximum slips are 3 m and 2.6 m for the north and south asperity respectively. We estimate a rupture velocity of 3.4 km/s, a stress drop of 30 bars, and a gamma (high-frequency spectral decay in Brune's model) of 1.55. b) Comparison between the Tocopilla acceleration records in black and the acceleration predictions from the estimated rupture model in red, for each station and component.

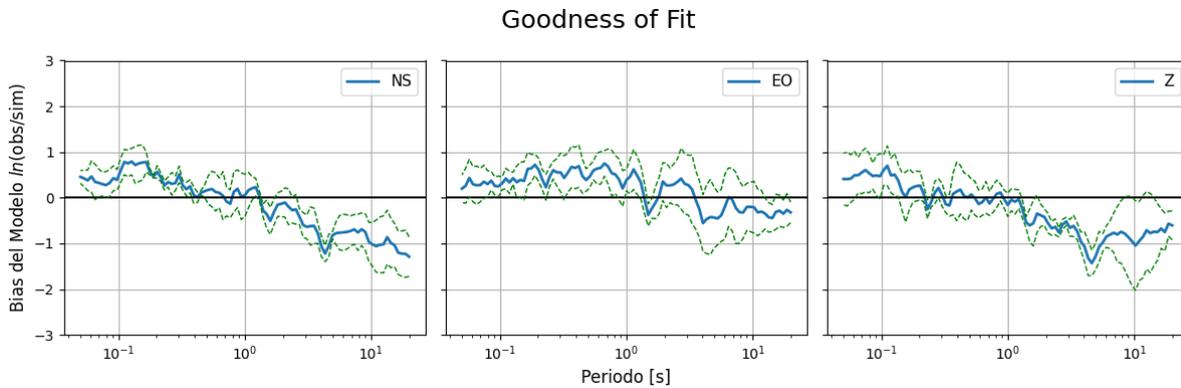


Figure 2: Goodness of fit (GOF) between the observed and predicted acceleration response spectra of 5% damping for the Tocopilla earthquake. The blue line shows the bias, and the region between the dashed lines shows the standard error.

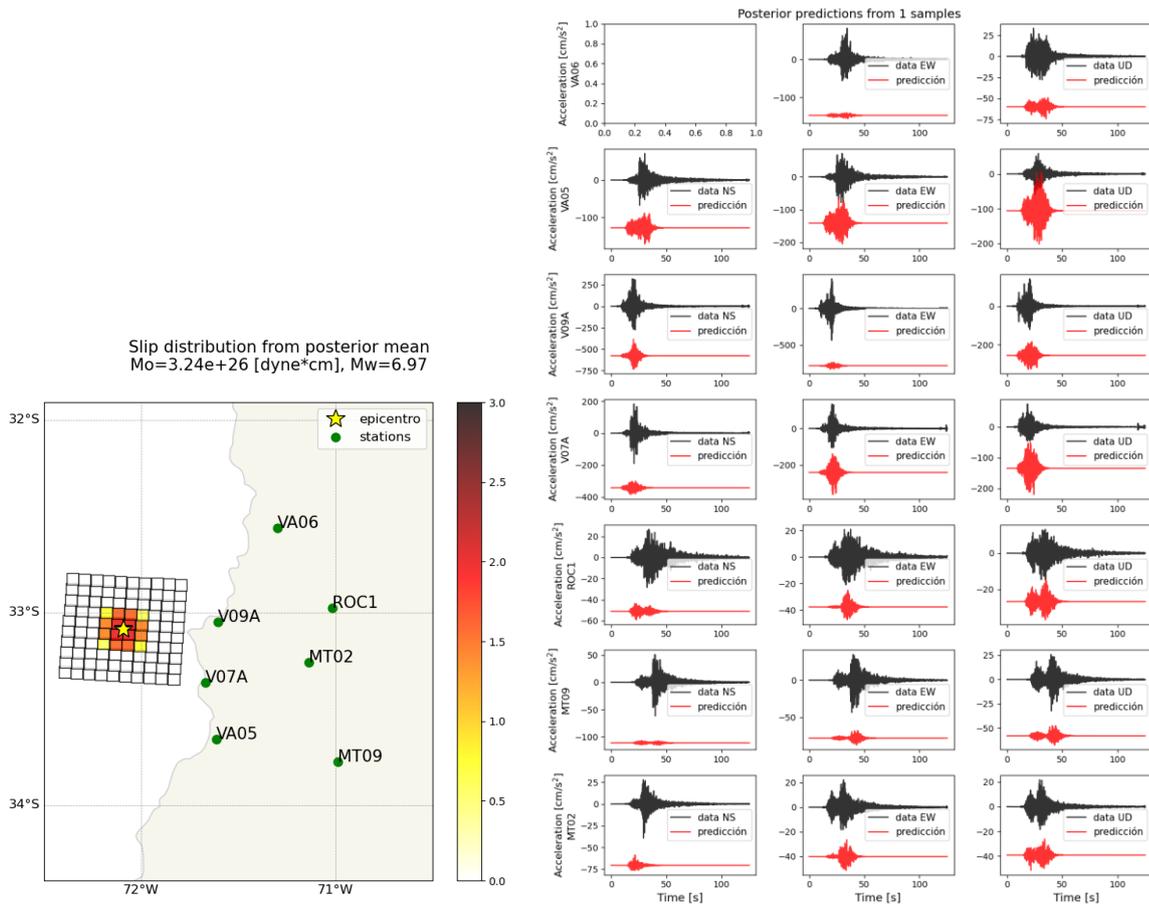


Figure 3: a) Estimated rupture model for the Valparaiso earthquake. The color scale shows the slip distribution for the circular asperity, which is fixed at the hypocenter of the event, represented by a yellow star. The estimated radius is 14.5 km, and the maximum slip is 2.1 m. We estimate a rupture velocity of 1.63 km/s, a stress drop of 31 bars, and a gamma (high-frequency spectral decay in Brune's model) of 1.85 for near stations (V09A, V07A, and VA05) and of 2.0 for far-away stations (VA06, ROC1, MT02, and MT09). b) Comparison between the Valparaíso acceleration records in black and the acceleration predictions from the estimated rupture model in red, for each station and component. The NS component is not well recorded for the VA06 station and it is not considered in the analysis.

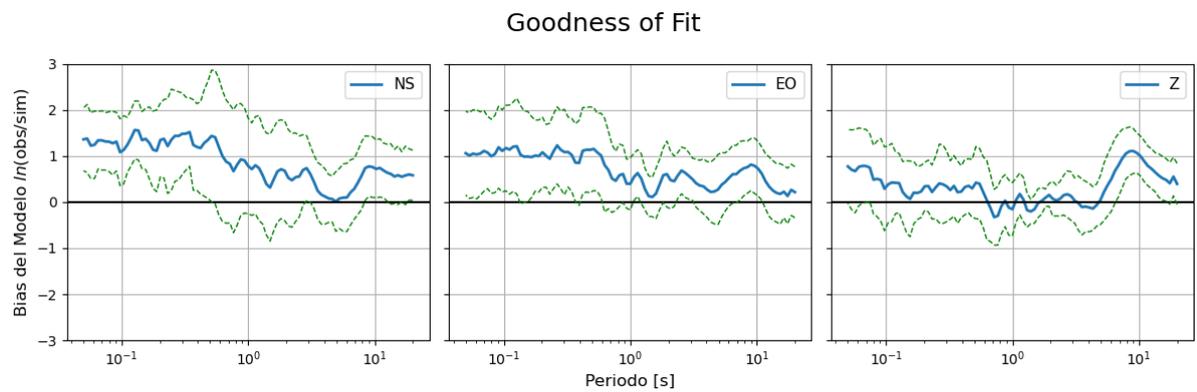


Figure 4: Goodness of fit (GOF) between the observed and predicted acceleration response spectra of 5% damping for the Valparaíso earthquake. The blue line shows the bias, and the region between the dashed lines shows the standard error.

Mathématiques, informatique

EUR DS4H, Spectrum

Contre-argumentation en langage naturel pour lutter contre la désinformation en ligne

Responsable scientifique : Serena Villata

Laboratoire : INRIA, I3S, CNRS, UCA, 3IA

Description :

Online social media platforms try to limit the virality of disinformation spread for example through content moderation. While these measures show some kind of effectiveness in limiting the diffusion of misleading or fake information, one big issue still remains unsolved: identifying disinformation and reporting its status to the users is not enough to counter it.

The goal of the PhD program is to address that urgent need by designing intelligent (semi-)automated ways to generate counter-arguments to fight the spread of disinformation online. The idea is to avoid the undesired effects that come with content moderation, such as overblocking, when dealing with disinformation online, and to directly intervene in the discussion (e.g., Twitter threads) with textual arguments that are meant to counter the fake content as soon as possible, and prevent it from further spreading. The idea is to obtain high quality counter-arguments while reducing efforts and supporting human fact-checkers in their everyday activities.

Descente d'échelle à très haute résolution pour la prévision de vent local - Dimensionnement de simulation de bassin de navigation sportive

Responsable scientifique : Mireille Bossy

Laboratoire : EPC CALISTO (Inria)

Résultats obtenus grâce aux moyens informatiques attribués :

Un résumé des résultats obtenues à partir de prévision météo fournies par nos partenaires (Sportrizer et Riskweathertech) est disponible ici : <https://windpos.Inria.fr/projects/avantage/>

Description :

Contexte scientifique

La prévision météorologique est un domaine d'expertise qui s'appuie sur un modèle numérique qui résout les équations de la physique et de la dynamique de l'atmosphère, alimenté par des données d'observation. Le modèle numérique produit une donnée de prévision pour des périodes de plusieurs jours à large échelle (25 à 50 km) sur l'ensemble du globe. Ensuite, une chaîne de traitement numérique dite de descente d'échelle va raffiner la résolution spatiale et temporelle de la prévision sur une région géographique en particulier.

Dans ce projet nous nous intéressons plus particulièrement à la prévision de condition de vent dans les basses couches de l'atmosphère, là où les effets de la turbulence atmosphérique sont les plus prépondérants et où la description des conditions de terrains (topographie et type de surface, urbaine, mer, végétation) peut apporter beaucoup plus de finesse et mieux résoudre la variabilité des conditions de vent prédites à l'échelle locale.

SDM-WindPoS est un logiciel développé dans l'EPC CALISTO qui met en œuvre une méthode innovante de simulation du vent qui s'appuyant sur la formulation lagrangienne des équations de la turbulence pour proposer un solveur numérique qui supporte des résolutions spatiales très fines. Ce solveur numérique peut être utilisé dans le domaine ou

sous-domaine pertinent pour l'analyse du vent, la dynamique au bord de ce domaine pouvant être extrapolée à partir d'une simulation météorologique de plus grande échelle.

SDM-WindPoS est notamment conçu pour être utilisé en association avec les sorties d'un solveur météorologique fournissant une prévision de vent dans les premières centaines de mètres de l'atmosphère, et seulement dans la zone où l'information locale de vent est décisive pour l'application visée. Des informations topographiques sont également introduites en entrée de la simulation pour enrichir la donnée de plus grande échelle, et ainsi saisir le plus finement possible les effets du relief et du terrain.

Description et objectifs

Le projet vise à caractériser les performances du modèle SDM-WindPoS sur de la prévision de vent à de très hautes résolutions spatiales, qui sont complètement innovantes dans une chaîne de prévision. Le domaine cible de cette descente d'échelle est le bassin de navigation sportive de la ville de Marseille, sur une plage horaire de quelques heures correspondant à la durée d'un événement sportif sur l'eau.

Différentes situations de vent, par secteur et force sont fournis par le modèle WRF en chaîne opérationnel de prévision. Le modèle WRF est descendu jusqu'à 1km de résolution sur la région de Marseille. Un run de WindPoS prend ensuite le relai de cette descente à quelques dizaines de mètres de résolution.

L'accès à Azzurra pour ce projet nous aidera à évaluer l'apport de SDM-windPoS dans la chaîne de prévision, via

- une validation qualitative des simulations (impact des reliefs naturels et urbains, sillages, recirculations, turbulence).
- une évaluation contre des mesures de vent (sous réserve de leur disponibilité dans la durée du projet).

Azzurra nous aidera également à dimensionner et à d'évaluer plus exhaustivement l'apport et l'impact des simulations emboîtées intermédiaires.

Complex Network Mining

Responsable scientifique : Claude Pasquier

Laboratoire : I3S

Résultats obtenus grâce aux moyens informatiques attribués :

Anna Diamant, stagiaire de M2 a utilisé les ressources de calculs pour faire tourner des pipelines bio-informatiques nécessitant beaucoup de RAM.

Vincent Guerlais, ingénieur, a utilisé Azzurra pour la détection de splicing alternatif dans les mRNAs. La grille de calcul a été utilisée pour tester et évaluer les performances de la méthode AMINE développée par l'équipe.

Rapport de Master 2 intitulé "Towards the identification of sequences involved in alternative splicing caused by the loss of LncRNA SNHG21" par Anna Diamant, 2022, stage de Master 2 SV (Sciences du Vivant) Biology, Informatics and Mathematics (UCA)

Nonoverlapping gene networks regulated by mating. Claude Pasquier, Alain Robichon. Life Science Alliance Nov 2021, 5 (2) e202101119; DOI: 10.26508/lisa.202101119

Description :

La modélisation des données sous forme de réseau est apparue dans diverses disciplines comme une manière unifiée de représenter des données relationnelles complexes. Formellement, ces réseaux complexes (que nous appelons réseaux multidimensionnels) sont des multigraphes dans lesquels les nœuds et les arêtes sont (multi-)étiquetés. L'analyse de ces réseaux est difficile car l'espace de recherche est beaucoup plus grand que pour les graphes étiquetés. Mais il est cependant nécessaire de disposer de méthodes efficaces qui peuvent aider à identifier les motifs cachés ou qui peuvent mettre en évidence des relations entre des attributs des nœuds. Le cœur de l'activité de recherche proposée se concentre sur l'analyse de ces réseaux complexes à des fins d'extraction d'informations. Le champ d'application privilégié de nos méthodes est la biologie moléculaire. Les traitements que nous sommes amenés à effectuer nécessitent d'avoir accès à beaucoup de RAM et/ou beaucoup de processeurs. Nous avons, ces dernières années, utilisé les ressources de CICCADA pour développer et exécuter des algorithmes

généétiques [1, 2], effectuer des tests de nos algorithmes selon diverses configurations [3] et analyser les résultats générés par des équipes de biologistes avec lesquelles nous collaborons [4, 5]. Les collaborations en cours [4, 5] vont se poursuivre au minimum jusqu'à la fin de l'année 2021. Des demandes de financement d'autres projets collaboratifs, qui nécessiteront des ressources de calculs, sont en cours d'évaluation (AAPG21 ANR CALDAMIR avec Frédéric Flouvat (Université de Nouvelle Calédonie), AAPG21 ANR FEST avec Bernard Mari (IPMC)).

[1] Pallez, D., Gardès, J., & Pasquier, C. (2017). Prediction of miRNA-disease associations using an evolutionary tuned latent semantic analysis. *Scientific reports*, 7(1), 1-13.

[2] Correa, L., Pallez, D., Tichit, L., Soriani, O., & Pasquier, C. (2019, December). Population-based meta-heuristic for active modules identification. In *Proceedings of the Tenth International Conference on Computational Systems-Biology and Bioinformatics* (pp. 1-8).

[3] Pasquier, C., Guerlais, V., Pallez, D., Rapetti-Mauss, R., Soriani, O. Identification of active module in interaction networks using node2vec network embedding. En préparation

[4] Collaboration avec Olivier Soriani (iBV) financée par l'Université Côte d'Azur et la fondation ARC

[5] projet GlioSplice, impliquant Mathieu Gabut (Centre de recherche en cancérologie de Lyon – CRCL), Thierry Virolle (Institut de Biologie valrose, Nice), Corinne Auge-Gouillou (Université de Tours) et financé par l'institut National de Cancer.

Simulations numérique de nanophotonique computationnelle avec le solveur Galerkin discontinue massivement parallèle DIOGENeS

Responsable scientifique : Stéphane Lanteri

Laboratoire : INRIA

Résultats obtenus grâce aux moyens informatiques attribués :

Les résultats numériques obtenus portent sur plusieurs problématiques en lien avec l'étude des interactions lumière-matière nanostructurée:

1./ Conception de métasurfaces pour des dispositifs optiques planaires (déflecteurs et métalentilles);

2./ Étude de nanostructures plasmoniques pour la détection.

Dans chaque cas, on exploite et combine différents ingrédients méthodologiques: des solveurs électromagnétiques basés sur une méthode élément fini de type Galerkin discontinu d'une part, et des optimiseurs numériques par apprentissage statistique.

Les résultats obtenus ont permis de consolider des collaborations avec des partenaires physiciens d'UCA (du CRHEA et d'INPHYNI).

Description :

Atlantis (<http://www-sop.inria.fr/atlantis/>) est une équipe-projet commune entre l'Inria, le CNRS et l'Université Côte d'Azur par son association avec le Laboratoire de Mathématiques J.A. Dieudonné (UMR 7351).

Ses activités de recherche se concentrent sur la conception, l'analyse, le développement et l'application de techniques de calcul avancées pour des systèmes d'équations aux dérivées partielles (EDP) modélisant des problèmes d'interaction lumière-matière à l'échelle du nanomètre. Le domaine correspondant est celui de la nanophotonique, aussi connue sous le nom de nano-optique, la nanophotonique vise à exploiter les interactions lumière-matière à des échelles nanométriques afin d'atteindre un haut niveau de contrôle de la lumière. La nanophotonique englobe une grande variété de sujets, notamment les métamatériaux, la plasmonique, l'imagerie haute résolution, la nanophotonique quantique et les matériaux photoniques fonctionnels.

C'est dans ce contexte que nous travaillons sur différentes problématiques et des applications en collaboration avec des chercheurs physiciens, notamment de laboratoires SNRS du périmètre UCA (notamment le CRHEA et INPHYNI) et des partenaires industriels (notamment, STMicroelectronics).

Les recherches de l'équipe projet ATLANTIS sont appuyées et soutenues par la suite logicielle DIOGENeS (<http://diogenes.inria.fr>). Cette suite logicielle comprend plusieurs solveurs numérique type élément fini basés sur la méthode Galerkin discontinue. Cette méthode, particulièrement adaptée au calcul parallèle, permet de simuler des configurations réalistes de structures nanophotoniques. Ces simulations peuvent permettre de valider des nouveaux modèles physiques pour la représentation des caractéristiques des matériaux à des échelles nanométriques, peuvent également permettre de valider de nouvelles méthodes numériques (schéma en temps particulier, etc..) ou bien encore de faire de l'optimisation paramétrique.

Numerical solution for a hand-held high-resolution medical scanner (MEDSCAN)

Responsable scientifique : Victorita Dolean

Laboratoire : LJAD

Résultats obtenus grâce aux moyens informatiques attribués :

Sahar Borzooei, Victorita Dolean, Claire Migliaccio and Pierre-Henri Tournier, "Solution of time-harmonic Maxwell's equations by a domain decomposition method based on PML transmission conditions" - Poster at SCEE 2022, The 14th International Conference on Scientific Computing in Electrical Engineering

Sahar Borzooei. A domain decomposition method with PML transmission conditions for Maxwell's equation
the 27th International Domain Decomposition Conference, DD27, in Prague, Czech Republic
(contributed talk)

Description :

PhD project MEDSCAN (BoostUrCareer program - cofunded by EU, IDEX, Région PACA)

Microwave tomography is a novel, early development stage way of imaging with a number of potentially attractive medical applications. With microwave tomography in the frequency range of several hundred MHz to several GHz, tissues are well differentiated and, consequentially can be imaged, based on differences in dielectric properties (i.e. contrast). It has been proven in a number of studies that dielectric properties of biological tissues are a strong indicator of their functional and pathological conditions. A very small data acquisition time (within 10-20 msec range) is another attractive feature of microwave tomography. The possibility of functional and fast imaging combined with decent spatial resolution (5-7 mm), cost efficiency and use of safe, very low dose non-ionizing radiation make microwave tomography very competitive with traditional imaging modalities, widely used in clinical practice, such as CT, MRI, PET, SPECT and ultrasound. Microwave imaging (MI) has attracted significant research interest in recent

times. By exposing tissues to low-level microwave incident field and capturing the scattered field by an array of antennas, the estimation of the dielectric properties of the tissues can be approximated and a diagnosis inferred. Physicians primarily rely on medical imaging systems such as computed tomography (CT) and magnetic resonance imaging (MRI) for confirmation of the diagnosis. There are, however, not fast, cost effective nor portable, and nor carried by first response paramedical teams. Furthermore, the increase use of CT scans has brought to light health concerns due to the ionizing radiation the technology relies upon. However, imaging the human tissues is an ambitious task; it is composed of multi-layered dielectrics of high loss.

From computational point of view, we expect a development of a parallel computing simulation tool which will be based on FreeFEM++ (and thus open source). The linear systems resulting from the discretization of the three-dimensional time harmonic Maxwell equations using continuous or discontinuous finite element methods designed on tetrahedral meshes are characterized by large sparse, complex coefficients and irregularly structured matrices. Classical preconditioned iterative methods (such as the GMRES Krylov method preconditioned by an incomplete LU factorization) generally behave poorly on these linear systems. A standard alternative solution strategy calls for parallel sparse direct solvers, but they generally fail to solve very large three-dimensional problems, due to the potentially huge memory requirements for these cases. Iterative methods can be used to overcome this memory problem. In the past years, a class of parallel solution methods namely two-level overlapping Schwarz methods already developed for some scalar partial differential equations was recently extended to Helmholtz and Maxwell equations in some particular cases. These equations are notoriously difficult, and more investigation is needed especially in heterogeneous cases.

There are two major challenges at the numerical level.

1. Developing a numerical tool for solving the forward problem to realize a realistic model of the data acquired by the microwave imaging system is novel and challenging. The modeling must have to take account accurately of the high heterogeneity and complexity of human tissues (skin, fat, muscle, veins and arteries...) for normal cases and for different possible pathology cases. The tissues are dissipative dielectric media in the microwave domain, i.e. they exhibit a complex permittivity. They are moreover dispersive (the

complex permittivity varies versus frequency). The wave/matter interaction must take also in account accurately of the field transmitted and received from imaging system (Evanescent Imaging Matrix).

2. Developing a new numerical method for solving the inverse problem able to reconstruct tomographic microwave images from experimental data (for different realistic phantoms models) acquired with the system.

Both points rely on the solution of large problems of millions of unknowns for which scalable methods from the state of the art need to be tested and others to be developed.

EURAD project

Responsable scientifique : Roland Masson

Laboratoire : LJAD

Description :

Coupling non-isothermal compositional liquid gas Darcy and free gas flows

Flow and transport processes in domains composed of a porous medium and an adjacent free-flow region appear in a wide range of industrial and environmental applications such as food processing, wood or paper production, salinization of agricultural land, prediction of convective heat and moisture transfer at exterior building surfaces, or also the study of the mass and energy exchanges at the interface between a nuclear waste disposal and the ventilation tunnels. Typically, in such processes, the porous medium initially saturated with the liquid phase is dried by suction in the neighborhood of the interface between the porous and free flows. To model such physical processes, one needs to account in the porous medium for the flow of the liquid and gas phases including the vaporization of the water component in the gas phase. In the free flow domain, a single phase gas free flow can be considered assuming that the liquid phase is instantaneously vaporized at the interface. This single phase gas free flow has to be compositional and non-isothermal to account for the change of the relative humidity and temperature in the tunnel which has a strong feedback on the liquid flow rate at the interface.

If many works have been performed to model and discretize the coupling of single phase Darcy and free flows, there is very little work on the coupling of a two phase gas liquid compositional Darcy flow with a single phase compositional free flow. Such a coupled model has been recently proposed in [3, 4] using proper matching conditions at the interface between the porous medium and the free flow regions. This model will be the starting point of our work and will be simplified taking into account the physical characteristics of our problem focusing on the drying processes at the interface between the nuclear waste repository and the ventilation excavated tunnels. To obtain our simplified model, we will assume that the longitudinal dimension of the free flow domain,

typically a tunnel, is large compared with its transversal dimensions. This allows to reduce the model in the tunnel to a 1D free flow coupled with the 3D flow in the porous medium. In the spirit of [6], we will extend to the non-isothermal case the approach proposed in [5].

The discretization of the 3D non-isothermal compositional liquid gas Darcy flow is based on the nodal Vertex Approximate Gradient scheme [1]. The coupling with the 1D free gas flow will be based on non-matching meshes between the 3D Darcy flow domain and the 1D domain in order to account for complex geometries of the tunnels. In order to allow for a modular implementation of the coupled liquid gas Darcy and free gas flows, a nonlinear Robin Robin domain decomposition algorithm based on the previous works [2, 7] will be designed.

The coupled liquid gas Darcy and free gas flows will be implemented in the open source parallel code ComPASS [1, 8] <http://www.anr-charms.org/page/compass-code>. The COMPASS parallelization is based on mesh partitioning and MPI communications. It uses the Petsc and Hypr libraries for the solution of the linear systems. Since the size of the mesh of the 1D free gas flow is negligible compared with the size of the mesh of the 3D Darcy flow, the domain decomposition approach will allow a simple parallel implementation of the coupled model. The code ComPASS <http://www.anr-charms.org/page/compass-code> is an open source parallel code initiated by LJAD-Inria and BRGM (Bureau de Recherches Géologiques et Minières – French Geological Survey) in 2015 [1, 8]. It is devoted to the simulation of multiphase non-isothermal Darcy flows and includes complex network of fractures/faults represented as interfaces of co-dimension one coupled to the surrounding matrix. The discretization is based on vertex and cell unknowns and is adapted to polyhedral meshes and heterogeneous media. The ComPASS code is co-developed since 2017 by the partners of the ANR CHARMS project including BRGM, LJAD-Inria, Storengy, la Maison de la Simulation and the Jacques Louis Lions laboratory.

References

[1] F. Xing, R. Masson, S. Lopez, Parallel numerical modeling of hybrid-dimensional compositional non-isothermal Darcy flows in fractured porous media, *Journal of Computational Physics*, 345, pp. 637-664, 2017, <https://hal.archives-ouvertes.fr/hal-01420361>

-
- [2] R. Masson, L. Trenty and Y. Zhang, Coupling compositional liquid gas Darcy and free gas flows at porous and free-flow domains interface, *Journal of Computational Physics*, 321, pp. 708-728, 2016.
- [3] Mosthaf, K., Baber, K., Flemisch, B., Helmig, R., Leijnse, A., Rybak, I. and Wohlmuth, B., A coupling concept for two-phase compositional porous-medium and single-phase compositional free flow, *Water Resources Research*, 47, 10, 2011.
- [4] Vanderborght, J., Fetzner, T., Mosthaf, K., Smits, K.M. and Helmig, R., Heat and water transport in soils and across the soil-atmosphere interface: 1. Theory and different model concepts, *Water Resources Research*, 53,2, pp. 1057-1079, 2017.
- [5] Brenner, K., Masson, R., Trenty, L., Zhang, Y., Coupling of a two phase gas liquid compositional 3D Darcy flow with a 1D compositional free gas flow, 50,5, pp. 1491-1522, 2016.
- [6] Beaudé, Laurence and Brenner, Konstantin and Lopez, Simon and Masson, Roland and Smaï, Farid, Non-isothermal compositional liquid gas Darcy flow: formulation, soil atmosphere boundary condition and application to high-energy geothermal simulations, *Computational Geosciences*, 23,3, pp. 443-470, 2019.
- [7] Birgler, Nabil and Masson, Roland and Trenty, Laurent, A domain decomposition method to couple nonisothermal compositional gas liquid Darcy and free gas flows, *Journal of Computational Physics*, 368, pp. 210-235, 2018.
- [8] Lopez, Simon and Masson, Roland and Beaudé, Laurence and Birgler, Nabil and Brenner, Konstantin and Kern, Michel and Smaï, Farid and Xing, Feng, Geothermal Modeling in Complex Geological Systems with the ComPASS Code, Stanford Geothermal Workshop 2018 - 43rd Workshop on Geothermal Reservoir Engineering, february 2018.

Magnétohydrodynamique

EUR Spectrum

Effet Dynamo et Instabilités Magnétohydrodynamiques

Responsable scientifique : Yannick Ponty

Laboratoire : Lagrange / OCA, CNRS

Résultats obtenus grâce aux moyens informatiques attribués :

Nous avons obtenu des nombreux résultats qui ont produit une publication mise en valeur par une mention « Editors' Suggestion » :

Paul M. Mannix, Yannick Ponty, and Florence Marcotte "Systematic Route to Subcritical Dynamo Branches" Phys. Rev. Lett. 129, 024502 – Published 5 July 2022

<https://doi.org/10.1103/PhysRevLett.129.024502>

Description :

Ce projet comporte trois volets, qui utilisent différentes techniques numériques et deux codes pseudo-spectraux. Ces projets seront conduits pendant au moins deux ans 2020 et 2021. Nous avons déjà obtenu du temps de calcul à GENCI (2020), notamment sur les machines Occigen (500 000 h/mono procs) et Curie (400 000 h/mono procs). La production de long et coûteux runs seront essentiellement effectués sur les clusters de GENCI. Le mésocentre Azzurra sera utilisé pour des tests et des simulations de plus faible consommation.

I) Effet dynamo dans le Von Karman (Yannick Ponty Lagrange)

II) Effet dynamo de fluctuation (Dario Vincenzi LJAD & Yannick Ponty)

III) Optimisation par méthodes adjointes pour la dynamo (Florence Marcotte INRIA/LJAD & Yannick Ponty).

Contexte et historique :

L'origine des champs magnétiques est abordée par l'effet dynamo, c'est à dire l'amplification spontanée d'un germe magnétique dans des fluides électriquement conducteurs en mouvement. Depuis de nombreuses années, plusieurs projets de dynamo expérimentales voient le jour tant en France qu'à l'étranger. Deux projets de dynamo expérimentales (Riga (novembre 1999 et juillet 2000) et Karlsruhe (décembre 1999) ont

obtenu un effet dynamo dans des écoulements hélicitaires très contraints dans des tubes, présentant un champ de vitesse avec très peu de fluctuations [1]. L'équipe française de Von Kàrmàn Sodium dont l'expérience est basée à Cadarache a essayé d'obtenir l'effet dynamo avec un écoulement très fluctuant et turbulent. Ils ont mené plusieurs campagnes d'expériences depuis les années 2000, où seuls les phénomènes d'induction avaient pu être étudiés. Mais en septembre 2006, le projet a obtenu ce qui semble être la première dynamo turbulente de l'histoire [2]. Une moisson de résultats expérimentaux a été obtenue malgré la difficulté des mesures dans les métaux liquides. Les simulations numériques qui ont accès à toutes les quantités physiques en même temps vont donc jouer un rôle important dans la compréhension de l'effet dynamo et de sa saturation [3].

L'avenir expérimental de la dynamo se trouve actuellement en Allemagne dans le centre Helmholtz zentrum à Dresden [4], où une expérience est en train de se construire dans un nouveau bâtiment. Un cylindre de 2m de diamètre rempli de sodium liquide sera forcé par un système de précession sur une énorme plateforme tournante. Ils espèrent obtenir un effet dynamo, dans un écoulement turbulent. En collaborant directement avec des équipes expérimentales et dans un souci permanent de comparer les simulations et les résultats expérimentaux, ce projet numérique commencé depuis 2002, a permis explorer et de comprendre les propriétés d'un tel système, et notamment l'influence de la turbulence sur le seuil critique de l'instabilité dynamo, la relation entre la topologie émergente à grande échelle et les fluctuations à petite échelle. Cette étude paramétrique, en nombre de Reynolds cinétique et magnétique nécessite des simulations temporelles longues et des résolutions spatiales élevées pour approcher les régimes turbulents, que seuls des codes parallèles peuvent atteindre.*

[1] Special issue : Magnetohydrodynamics (2002).

[2] R. Monchaux, M. Berhanu, M. Bourgoïn, M. Moulin, Ph. Odier, J.-F. Pinton, R. Volk, S. Fauve, N. Mordant, F. Pétrélis, A. Chiffaudel, F. Daviaud, B. Dubrulle, C. Gasquet, L. Marié, and F. Ravelet. "Generation of a Magnetic Field by Dynamo Action in a Turbulent Flow of Liquid Sodium" Phys. Rev. Lett. 98, 044502 (2007).

<http://dx.doi.org/10.1103/PhysRevLett.98.044502>

[3] M. Berhanu, R. Monchaux, S. Fauve, N. Mordant, F. Pétrélis, A. Chiffaudel, F. Daviaud, F. Ravelet, M. Bourgoïn, Ph. Odier, J.-F. Pinton, R. Volk, B. Dubrulle, L. Marie "Magnetic field

reversals in an experimental turbulent dynamo" *Europhys. Lett.* 77, 59001 (2007).
doi:10.1209/0295-5075/77/59001

[4] André Giesecke, Tobias Vogt, Thomas Gundrum, and Frank Stefani. "Nonlinear Large Scale Flow in a Precessing Cylinder and Its Ability To Drive Dynamo Action" *Phys. Rev. Lett.* 120, 024502 – Published 12 January 2018.

l) Effet dynamo dans le Von Karman.

{ Holger Homann & Yannick Ponty (Lagrange, OCA)}

Forçage par des disques en mouvement avec un perméabilité magnétique variable: Effet dynamo fluide-structure.

Dans cette approche nous simulons complètement le forçage des disques qui tournent en contra-rotation pour produire un écoulement turbulent proche des écoulements observés en laboratoire [2-3]. Nous avons obtenu des résultats similaires aux expériences dans des simulations purement hydrodynamiques. Un article sur des résultats hydrodynamiques de ce système est publié dans *New Journal of Physics* (La méthode de pénalisation est expliquée dans notre l'article publié). Nous avons introduit le champs magnétique pour étudier l'effet dynamo dans un tel système et le comparer aux dynamos trouvées dans l'expérience de VKS [2, 3]. Il s'agit surtout de modéliser l'effet de la perméabilité magnétique de la structure disques-pales sur le mode dynamo. Nous avons tous les ingrédients hydrodynamiques (champ moyen, tourbillons le long des pales). Il nous reste à explorer et à comprendre l'interaction fluide-structure.

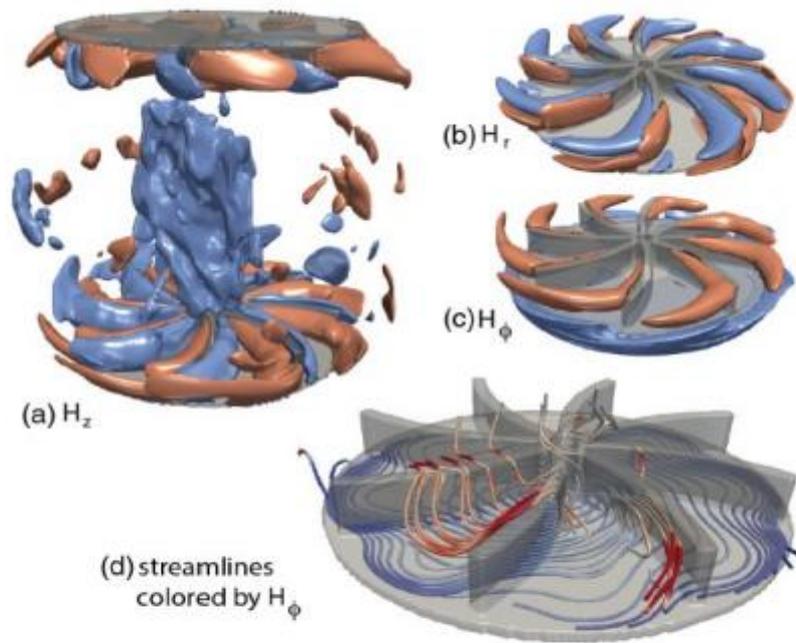


FIG. 4. (a) H_z , (b) H_r , (c) H_ϕ isoamplitude contours (red, positive value; blue, opposite negative value), and (d) magnetic streamlines colored by the H_ϕ value. $Re = 1500$, $Rm = 500$, and $\mu_r = 14$.

Fig 1 : Visualisation des composantes du champ magnétique H_z , H_r , H_ϕ et des lignes de champ autour de la structure solide du disque-pales, du mode dynamo ($m=0$).

En 2016-2019:

Nous avons pu faire une série de simulations où un effet dynamo est produit par notre von Karman numérique. Lors de cette série, nous avons varié la perméabilité magnétique de la structure solide de l'expérience (les disque-pales). Nous avons pu montrer que la géométrie des pâles et le saut de perméabilité joue un rôle moteur pour obtenir le mode dynamo ($m=0$) observé dans les expériences (voir Fig 1). Le saut de perméabilité magnétique pousse les lignes de champ vers le disque, qui devient un attracteur et un amplificateur de champ. Cet effet est fortement présent dans le régime de saturation. Notre article est publié dans Phys. Rev. Lett³⁶

³⁶ Dynamo Enhancement and Mode Selection Triggered by High Magnetic Permeability S. Kreuzahler, Y. Ponty, N. Plihon, H. Homann, and R. Grauer Phys. Rev. Lett. 119, 234501 – Published 6 December 2017 [PDF] Une illustration de cette publication a fait l'objet de la COVER PAGE du journal Phys. Rev. Lett. (PRL).

Pour 2021 et ...

- Nous devons continuer à explorer la dynamique non linéaire présente dans le régime de saturation (inversion erratique du champ magnétique) avec des simulations plus longues. Nous avons observé des inversions du champ magnétique, comme dans l'expérience VKS. Nous devons produire des simulations longues pour étudier la dynamique de ces inversions, leurs statistiques et les différents mécanismes associées (voir Fig 2).
- Nous pourrions aussi explorer une autre géométrie expérimentale qui n'a pas produit de dynamo (sphère) avec deux disques-pôles identiques à VKS.

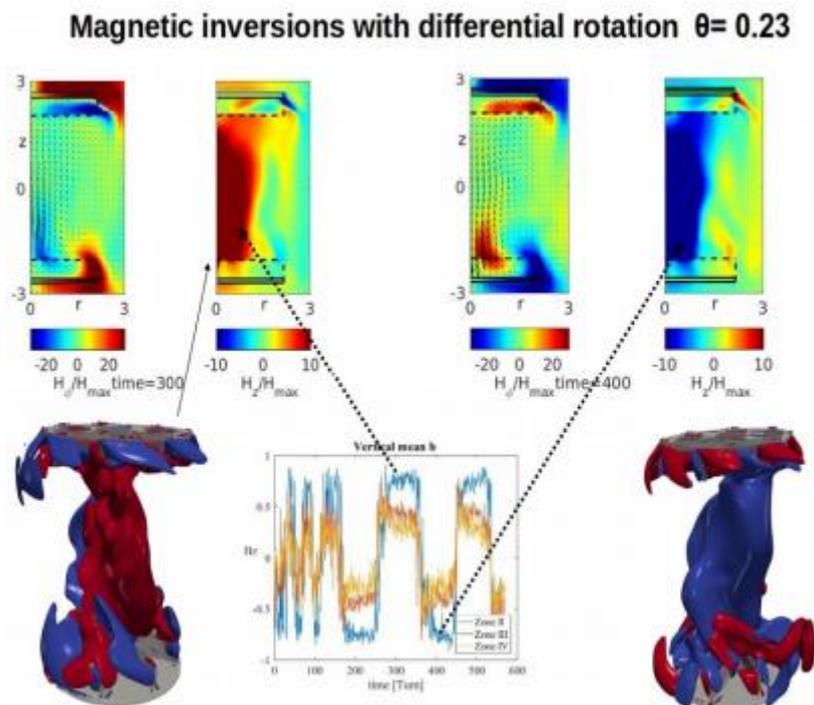


Fig 2: Inversion du champ magnétique dans une simulation où les deux disques ne tournent pas à la même vitesse : $\theta = (\Omega_2 - \Omega_1)/(\Omega_2 + \Omega_1)$.

II) Effet dynamo de fluctuation

{ Dario Vincenzi (LJAD) & Yannick Ponty (Lagrange, OCA)}

En collaboration avec Dario Vincenzi (Lab Dieudonné Nice), nous voulons explorer numériquement la dynamo de type Kasensev. Une contribution fondamentale à la théorie

de la dynamo turbulente (ou de fluctuation) a été donnée par Kazantsev³⁷ (1968), qui a étudié analytiquement la croissance de l'énergie magnétique dans un écoulement aléatoire incompressible, gaussien et à temps de corrélation très court. Le modèle de Kazantsev continue à revêtir une grande importance dans la théorie de la dynamo^{38 39}, car il est l'un des rares modèles de dynamo turbulente qui peut être traité analytiquement. Il a ainsi fait l'objet d'un grand nombre d'études et plusieurs généralisations du modèle ont été proposées afin de prendre en compte des propriétés d'un écoulement turbulent qui ne figuraient pas dans le modèle originaire. Même s'il est très utile pour une compréhension qualitative de l'effet dynamo turbulent, le modèle de Kazantsev est néanmoins assez limité à cause des hypothèses d'isotropie et de décorrélation temporelle. La compréhension de la dynamo turbulente dans des conditions réalistes se fonde donc nécessairement sur la simulation numérique et la comparaison avec les expériences de laboratoire. Ici nous proposons d'étudier numériquement différents écoulements de type Kraichnan soit avec un forçage delta corrélé en temps donc avec du bruit additif⁴⁰ soit une système PDE stochastique où l'écoulement (le bruit) est multiplicatif, où le schéma numérique prendra en compte de la stochasticité du système. De plus, l'effet de la compressibilité sera aussi étudié et comparé aux résultats analytiques⁴¹.

Ce projet utilisera le code Cubby. En utilisant la bibliothèque BOOST::RANDOM ou std::random, il y a une implémentation d'un contrôle de champ de vecteur aléatoire qui sont reproductibles avec un système de "seed" contrôlés et reproductibles. Cette propriété sera importante pour cette étude sur de la dynamo produite par des écoulements aléatoires et où le contrôle des séquences aléatoires seront importantes.

III) Optimisation par méthodes adjointes pour la dynamo

{Florence Marcotte (INRIA/LJAD) & Yannick Ponty(Lagrange, OCA)}.

³⁷ Kazantsev, AP (1968) Sov. Phys. JETP 26, 1031

³⁸ Rincon, F (2019) J. Plasma Phys. 85, 205850401

³⁹ D. Vincenzi, The Kraichnan–Kazantsev Dynamo J. Stat. Phys. 106, 1073-1091 (2002)

⁴⁰ A A Schekochihin et al 2007 New J. Phys. 9 300

⁴¹ M. Martins Afonso, D. Mitra & D. Vincenzi "Kazantsev dynamo in turbulent compressible flows" Proc. R. Soc. A 475, 20180591 (2019).

Un autre volet du projet concerne le développement et la mise en œuvre d'un code de contrôle optimal d'écoulements MHD incompressibles. Le Direct Adjoint Looping (contrôle optimal) permet aujourd'hui d'optimiser dans des espaces de grande dimension des systèmes fluides instationnaires et gouvernés par une dynamique entièrement non-linéaire. Fondé sur des méthodes adjointes, le DAL détermine itérativement la variable d'optimisation (par exemple un forçage) minimisant ou maximisant une quantité d'intérêt (fonctionnelle de coût) tout en prescrivant un certain nombre de contraintes (équations d'évolution, conditions de bord, etc) à l'aide de multiplicateurs de Lagrange⁴². En pratique, sa mise en œuvre consiste, en partant d'une première estimation arbitraire de la variable d'optimisation, à intégrer les équations d'évolution gouvernant le problème physique considéré (« problème direct ») à partir d'un instant initial jusqu'au temps cible T (où la quantité d'intérêt est évaluée), puis à remonter par une intégration en temps rétrograde du problème adjoint (gouvernant les multiplicateurs de Lagrange) à une condition permettant de mesurer la distance à l'optimalité. Une meilleure estimation de la variable d'optimisation est alors construite, et la procédure est répétée jusqu'à convergence de l'algorithme. Chaque itération (chaque passage dans la boucle DAL) représente en terme de coût de calcul l'équivalent de 2 simulations numériques directes (3 lorsque la mémoire requise nécessite l'utilisation de checkpointing). La méthode DAL a déjà été utilisée dans de nombreux contextes en hydrodynamique, notamment pour optimiser le mélange dans un fluide stratifié⁴³. L'application de la méthode DAL pour la première fois à des écoulements instationnaires gouvernés par le système MHD non-linéaire complet (équations de Navier-Stokes, équation d'induction et conditions solénoïdales) nécessite le développement et la validation des outils numériques appropriés. Elle permettra d'étudier les conditions optimales pour l'obtention d'une dynamo fluide (en terme d'abaissement du seuil de l'instabilité dynamo par un forçage approprié). Ce type d'étude est notamment motivé par l'intensification, au niveau international, de l'effort de recherche pour le développement de dynamos expérimentales, notamment à Dresde

⁴² D. Hill, Adjoint systems and their role in the receptivity problem for boundary layers, *J. Fluid Mech* (1995) 292.

⁴³ F. Marcotte & C.P. Caulfield, Optimal mixing in 2D stratified plane Poiseuille flow at finite Peclet and Richardson numbers. *J. Fluid Mech.* (2018) 853.

(projet DRESHDYN⁴⁴) ou à Madison (projet MPDX⁴⁵). Ce projet utilisera le code Cubby qui permet facilement d'implémenter de nouvelles physiques ou des systèmes d'équations. Sa construction modulaire est adaptée à la procédure DAL : DNS-Adjoint ou les équations adjointes changent à chaque changement de contraintes. L'algorithme d'optimisation nécessite de nombreuses boucles DNS-Adjoint pour converger, soit un cycle de 30/40 simulations pour une seule contrainte. Méthodes numériques: Les calculs seront effectués sur les codes «Latu» et «Cubby» qui sont deux codes pseudospectraux qui calculent des EPD dans un espace périodique. Le code «Latu» est un code pseudo-spectral parallélisé avec la bibliothèque MPI construit par Holger Homann (Laboratoire Lagrange, OCA). Il peut donc tourner sur toutes les plateformes. Il est écrit en C++ et utilise les transformées de Fourier (FFTW3) et P3DFFT. Ce code est efficace, il a été utilisé pour un projet PRACE en 2011 avec un scaling pour des runs en 4096^3 jusqu'à 64000 cores . Il permet d'étudier la dynamique des fluides, de MHD, de particules inertielles, modèles bifluide ou de traceur. Il y a aussi une implémentation des techniques de pénalisation pour des parois immersives et des objets en mouvement comme la rotation des disques à pales de VKS. Il est maintenu par un suiveur de version GIT. Récemment, il y a eu un effort récent pour le mettre le système de compilation sur CMAKE, et implémenter des tests automatiques. Le code « Cubby » est un code pseudo-spectral parallélisé avec MPI et pthread. Il y a donc une double parallélisation . Il a été originellement construit par Yannick Ponty (Laboratoire Lagrange OCA) et ensuite il a été re-designé par un IR Alain Miniussi (OCA) en C++ (design pattern), de façon extrêmement modulaire. Il a un très bon scaling. Il est maintenu par un suiveur de version Subversion, et va passer prochainement sous GIT. Sa construction modulaire permet l'implémentation d'un nouveau système de PDE ou de schéma numérique très facilement. Il utilise CMAKE comme système de compilation, avec plus de 150 tests automatiques unitaires ou complets.

⁴⁴ F. Stefani et al., Dresdyn - a new facility for mhd experiments with liquid sodium, *Magnetohydrodynamics* (2014) 48

⁴⁵ C. Cooper et al., The Madison plasma dynamo experiment: a facility for studying laboratory plasma astrophysics, *Physics of Plasmas* (2014) 21

Contrôle optimal de dynamos sous-critiques

Responsable scientifique : Florence Marcotte

Laboratoire : LJAD / Inria

Résultats obtenus grâce aux moyens informatiques attribués :

Article:

- A systematic route to subcritical dynamo branches, Mannix, P.M., Ponty, Y. & Marcotte, F., Physical Review Letters 129, 024502 (2022).

HAL Id : hal-03862091, version 1

ARXIV : 2112.11376

DOI : 10.1103/physrevlett.129.024502

Preprint:

- Discrete adjoint-based control: A robust gradient descent procedure for optimisation with PDE and norm constraints (<https://arxiv.org/pdf/2210.17194.pdf>)

Mannix, P.M., Skene, C.S., Auroux, D. & Marcotte, F.

Posters:

- Graines minimales de transition dynamo. 25e Rencontre du Non-Linéaire, Université Paris Cité, 29 – 31 mars 2022, P. M. Mannix, Y. Ponty, Florence Marcotte, présenté par Florence Marcotte.

- Advances in geophysical and astrophysical turbulence TURW05, Isaac Newton Institute, University of Cambridge, 16-20 mai 2022, présenté par Paul Mannix

Conférences :

- Easy paths to subcritical dynamo branches. Isaac Newton Institute for Mathematical Sciences, University of Leeds, Fluid flow and magnetic field generation in fluids and plasmas - theory and laboratory experiments, 17 octobre 2022 - présentation orale (Yannick Ponty)

- Minimal seeds for subcritical dynamos. Isaac Newton Institute for Mathematical Sciences, University of Cambridge, Dynamos in planets and stars: similarities and differences, 14 septembre 2022 - séminaire invité (Florence Marcotte)

- Identification systématique de branches dynamo sous-critiques. P.M. Mannix, Y. Ponty, F. Marcotte, 25ème Congrès Français de Mécanique, Nantes, 29 août-2 septembre 2022 - présentation orale (Paul Mannix)

- Identification of subcritical dynamos and minimal seeds. Festival de Théorie d'Aix-en-Provence, Chaos Control, Feedback & Model Reduction, 4 juillet 2022 - séminaire invité (Florence Marcotte)

- A systematic approach to identify subcritical dynamo instabilities. Courant Institute of Mathematical Sciences, New York University, Magneto-Fluid Dynamics Seminar, 7 décembre 2021
- séminaire invité (Florence Marcotte)

Résumé publié sur le site de la MSI : <https://msi.univ-cotedazur.fr/actualites/graines-de-dynamo>

Description du projet :

Le projet vise à développer une méthode d'analyse de stabilité non-linéaire d'écoulements magnétohydrodynamiques (MHD) fondée sur le contrôle optimal, en considérant dans un premier temps le cas particulier des instabilités dynamos sous-critiques. Le code de contrôle optimal développé dans le cadre de ce projet s'appuie sur un code déjà existant, validé, massivement parallèle et librement distribué (DEDALUS) pour intégrer les équations de la MHD incompressible instationnaire en 3D (équations de Navier-Stokes + équation d'induction + contraintes solénoïdales pour le champ de vitesse et le champ magnétique) en temps avant, ainsi que leurs équations adjointes en temps rétrograde. Ce code permettra d'étudier les conditions optimales pour l'obtention d'une instabilité dynamo fluide dans un domaine à géométrie cylindrique, en particulier dans un écoulement linéairement stable. Ce type d'étude est notamment motivé par l'intensification, au niveau international, de l'effort de recherche pour le développement de dynamos expérimentales, notamment à Dresde (projet DRESHDYN) ou à Madison (projet MPDX). Il doit permettre à terme de caractériser, dans un contexte plus général, la stabilité non-linéaire d'autres types d'écoulements MHD (applications en fusion à confinement magnétique notamment).

Astrophysique

EUR Spectrum

Red supergiant and AGB atmosphere and winds using 3D radiative hydrodynamics simulations

Responsable scientifique : Andrea Chiavassa

Laboratoire : Lagrange / OCA

Description en page suivante.

Red supergiant and AGB atmosphere and winds

In the last years it has become possible to produce multi-dimensional radiative hydrodynamics (RHD) simulations of the gas flow in the surface layers including the effect of radiation. The energy exchange between the gas and the radiation field is critical as it determines the temperature stratification and is ultimately what drives the convective motion. 3D simulations are ab initio, time-dependent, multi-dimensional, and not-local. These simulations are of two kinds: (i) *box-in-a-star* models (for main sequence stars) cover only a small section of the surface layers of the deep convection zone, and the numerical box includes some fixed number of convective cells, which are large enough to not constrain the cells by the horizontal (cyclic) boundaries; (ii) *star-in-a-box* models (Fig., left panel) cover the whole convective envelope of the star and are used to model RSG and AGB stars. The transition where the box-in-a-star models become inadequate to represent the star occurs around $\log(g) \sim -1$, when the convective structures become too large (e.g. Chiavassa et al. 2010) and the influence of sphericity becomes important. The numerical code we use and for which we have a solid expertise is CO5BOLD-code (Freitag et al. 2012). This code is state-of-the-art for the *star-in-a-box* simulations. CO5BOLD is computer-time demanding. However, with the increased computational power, it has become possible to compute numerical grids of RSG and AGB stars (Fig., right panel). The computational time varies with stellar parameters, the numerical resolution, the physical ingredients included, the numerical approaches and the directive of parallelization taken into account. Until recently, CO5BOLD' calculations were performed in parallel using Open Multi-Processing (OpenMP) that limits the shared memory and the number of computing threads to a single server. Our most recent version of the code allows us to tackle a hybrid parallelised approach (i.e. Message Passing Interface - MPI - and OpenMP) for the calculations that increases the speed by a factor of 8 at least. In the end, the computation of a simulation depends on the numerical resolution required, the treatment of the radiative transfer, and the temporal follow up. Nowadays, it takes about 15 days to obtain a relaxed simulation on 320 cores (8 nodes with 2 CPU, 20 cores/CPU).

The presence of a magnetic field in stars is intimately linked to the convection across the stellar photosphere. Convection generates a field and at the same time increases the atmospheric velocities and the chromospheric temperatures. As a consequence, the overall structure of the stellar atmosphere is affected. For example, in the case of RSGs, local dynamos are expected to appear in relation with the large convective cells, as shown using CO5BOLD with their first toy models (Freitag et al. 2002). The numerical scheme used for the solution of equations of Magneto-HydroDynamics (MHD) in CO5BOLD is different from the one employed for the case of pure RHD case (see Freitag et al. 2012 for more details). Recently, CO5BOLD has been used to calculate the first Magneto-HydroDynamics simulations of evolved stars (Fig. 2, central panel), demonstrating the feasibility of this kind of approach.

We will work on the computation/development/analysis of a new grid of (at least 20) RSG and AGB simulations including the magnetic field. The already existent 3D grid will be used as starting simulations for the new grid. Since the WP aim at covering the parameters space of the observations (Fig., right panel), it will be very important to make stringent choices of the initial parameters and numerical resolution. This aspect will be tackled in detail working on: (i) higher numerical resolution to resolve small-scale features at the photospheric surface, (ii) large temporal coverage, (iii) characterisation of the effects of magnetic field in the outer layers, (iv) simulations including rotation (already undertaken in Freitag et al. 2017 for AGBs and by Chiavassa for RSGs, work in progress) will also be performed to study the enhancement of the dipolar component of the magnetic field.

These simulations are of paramount importance because they will provide our knowledge basis on mechanisms launching evolved stars winds, assuming that combined effects of magneto-acoustic shocks, convection, radiation pressure on gas can create the required conditions. The grid across the different stellar parameters will be crucial to puzzling out how the mass-loss rate and other wind properties depend on the stellar evolution.

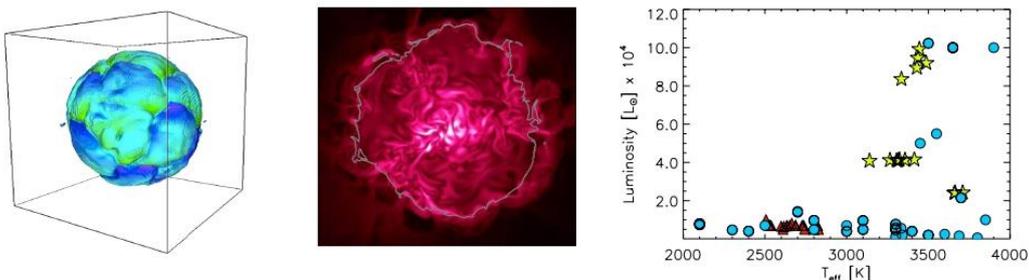


Fig. Left panel: Example of *star-in-a-box* simulation. Central panel: 3D simulation of an RSG with magnetic field (Freitag, private communication). The image displays the magnetic field strength in the interior, increasing towards the core. The approximate position of the surface (indicated with the gray line) shows the emergence of many small magnetic loops, which develop in the outer layers. Right panel: actual grid of RSG (yellow stars) and AGB (red triangles) simulations compared to the actual sample (blue circles) of our observations

Couplage turbulent gaz-grains solides dans les disques protoplanétaires

Responsable scientifique : Héloïse Méheut

Laboratoire : Lagrange / OCA

Résultats obtenus grâce aux moyens informatiques attribués :

Les moyens informatiques ont permis d'effectuer des simulations bidimensionnelles de la turbulence des disques protoplanétaires couplée à la dynamique de la phase solide. Ces résultats ont été soumis aux European Physical Journal Plus (EPJ Plus). Dans la suite du projet, nous nous orientons vers des simulations tridimensionnelles, donc plus lourdes.

Description :

Etat de l'art

Les disques protoplanétaires sont des pouponnières de planètes, en permettant leur formation à partir de la poussière interstellaire. Il s'agit de systèmes particulièrement complexes composés de gaz stratifié, cisailé et en rotation autour d'une étoile jeune et de solides de tailles allant de la centaine de microns jusqu'à celles des planètes. Les observations récentes (ALMA, SPHERE) de disques protoplanétaires ont montré la présence de structures complexes dans la distribution de poussière (spirales, croissants, asymétries). Leur interprétation nécessite de mieux comprendre la dynamique des solides dans ces disques gazeux. Cette problématique rejoint des questions encore ouvertes pour la compréhension de la formation des planètes. D'une part, le transport dans les disques protoplanétaires ne peut pas être expliqué exclusivement par la turbulence magnéto-rotationnelle, mais nécessite d'autres effets. D'autre part, il est nécessaire de mieux comprendre les mécanismes conduisant à une croissance rapide des planétésimaux. Le couplage dynamique entre la poussière et la turbulence du gaz constitue une réponse possible à ces questions. Par exemple, l'instabilité d'entraînement (streaming instability) permet d'expliquer en partie une agrégation rapide. Elle résulte de la rétroaction de la poussière sur le gaz et se matérialise par une excitation locale de la dynamique et l'apparition de filaments avec de très fortes concentrations de solides. Bien que de nombreuses études aient porté sur le développement macroscopique de cette instabilité,

les propriétés de la turbulence générée lors du couplage entre le gaz et la poussière, ainsi que son impact sur la concentration de poussière et le transport dans les disques, restent à ce jour mal compris. Aussi, l'effondrement gravitationnel de ces concentrations de solides est encore mal connu, du fait des limites des méthodes numériques à ce jour utilisées. Ce projet s'appuiera sur les résultats obtenus lors du stage d'Anand Kapoor financé par l'EUR Incise qui a permis de démontrer la faisabilité de ce projet.

OBJECTIFS

Le but de ce projet est d'apporter de nouvelles perspectives à ce problème en proposant une approche novatrice, transdisciplinaire de l'étude du couplage entre la dynamique de la poussière et la turbulence du gaz porteur. Nous développerons des outils de pointe pour la simulation directe du fluide et d'une phase solide qui sera modélisée soit par un grand nombre de particules ponctuelles, soit par une approche hydrodynamique. La collaboration interdisciplinaire entre l'OCA et l'INRIA est ici cruciale pour étudier la turbulence avec les outils théoriques et numériques utilisés en mécanique des fluides et ses applications industrielles, et ainsi débloquent un verrou scientifique dans la compréhension de la formation des planètes.

Les travaux proposés s'articulent en trois objectifs principaux :

1. Étude de la dynamique et de la concentration des particules solides dans un fluide turbulent cisailé et en rotation, en fonction de la fraction massique de particules solides et de leur taille. Les particules sont déjà prises en compte dans notre code, et cet objectif ne nécessitera pas de nouveaux développements.
2. Quantification des modulations de la turbulence incompressible par ces petites particules. Cet objectif nécessite d'implémenter dans le code existant l'effet retour des particules sur le fluide. Ces développements seront achevés dans les mois qui viennent.
3. Extension au cas d'écoulements faiblement compressibles. L'idée est de reproduire les conditions propres aux disques protoplanétaires. Il sera alors possible d'étudier les différents mécanismes de concentration des solides ainsi que leur effet sur le transport dans les disques.

Le but de cette demande est de s'attaquer au premier objectif. Les deux autres objectifs, à plus long terme, seront l'objet de demandes de temps complémentaires sur les moyens de calcul régionaux ou nationaux.

OUTILS

Ce projet utilisera notre code spectral développé en C++ pour l'étude de la turbulence d'un fluide cisailé. Il s'agit d'un code à mémoire partagée (utilisant FFTW3 et POSIX threads), ce qui est adapté pour cette étude bi-dimensionnelle avec une charge mémoire légère. Nous prévoyons de faire une étude de paramètres (vitesse de rotation et temps d'arrêt des particules) en deux dimensions, sur des grilles de taille 2048^2 et avec 106 particules. Nous envisageons d'effectuer 36 simulations correspondant à 6 valeurs de vitesse de rotation et 6 valeurs de temps d'arrêt. L'écoulement doit être intégré sur plusieurs dizaines de temps de retournement pour garantir des statistiques suffisamment convergées. Le temps de calcul d'une simulation est estimé à environ 5000 hCPU.

Construction of the Bayesian next generation of planetary ephemerides, B-INPOP

Responsable scientifique : Agnès Fienga

Laboratoire : Geoazur

Résultats obtenus grâce aux moyens informatiques attribués :

Grâce aux calculs faits sur Azzurra, nous avons pu simuler l'apport des données Bepi-Colombo de radio-science à la construction des éphémérides planétaires INPOP. Nous avons pu simuler de nouvelles contraintes sur la masse du graviton et sur les théories de type dilaton dans le système solaire.

Une communication à un colloque international (Rencontres de Moriond 2022) et à un workshop Bepi-Colombo (Bepi-Colombo Geodesy and Geophysics Working Group Meeting) ont été effectuées et un poster dans une école de statistique (Bayesian Deep Learning for Cosmology and Time Domain Astrophysics.) a été produit. Une première publication est en cours d'écriture pour un envoi aux éditeurs pour la fin de l'année.

56th Rencontres de Moriond "Gravitation", January 30th - February 6th 2022 :

Tests of GR with planetary ephemerides with BepiColombo simulations

Agnès Fienga (Géoazur, Observatoire de la Côte d'Azur)

https://moriond.in2p3.fr/2022/Gravitation/transparencies/02_tuesday/01_morning/04_fienga.pdf

4th BepiColombo GGWG Virtual Meeting (21st January 2022) : Agnes Fienga, "INPOP Planetary Ephemerides and BepiColombo : from Orbit Improvements to GR and Alternative Theories tests"

<https://www.cosmos.esa.int/web/bepicolombo-ggwg>

Bayesian planetary numerical ephemerides B-INPOP with MCMC. Vincenzo Mariani. Poster pour Bayesian deep learning for cosmology and time-domain astrophysics 2022 (20/06 - 24/06 2022).

Description du projet pages suivantes.

CONSTRUCTION OF THE BAYESIAN NEXT GENERATION OF PLANETARY EPHEMERIDES, B-INPOP

Supervisor: A. Fienga, Géoazur
agnes.fienga@oca.eu

Co-fundings: ERC AstroGeo (J. Laskar, IMCCE, Observatoire de Paris)
jacques.laskar@imcce.fr

Abstract

This project aims at initiating paths for the future development of planetary ephemerides. With the income of new very accurate and very numerous data from space missions such as ExoMars, Bepi-Colombo and JUICE, the maintenance and the development towards new scientific applications of INPOP planetary ephemerides will be very challenging. The INPOP planetary ephemerides are worldwide famous and used planetary ephemerides allowing, beside the regular delivery of positions and velocities of about 14,000 objects of the solar system (major and minor planets as well as the Earth and Moon orbital and rotational states) to common users by also space agencies (such as CNES and ESA), to test general relativity and alternative theories of gravitation and to constraint the mass distribution in the solar system by determining masses for main belt and Kuiper belt objects and consequently the possible existence of an unobserved planet P9. INPOP is also used as an important benchmark for long-term ephemerides dating geological samples. INPOP is currently built using about 200,000 observations from sources very different in technics but also in accuracies and on the basis of a classic least-square inversion. However because of the important number of unknowns in the dynamical modeling (i.e. about 800 parameters for the planetary orbits, 102 for the Moon libration and 42,000 for the minor planets) the inversion of such a system is non-linear and, since (Fienga *et al.*, 2015), new approaches have been tested for improving the inversion (Fienga *et al.*, 2020a). In this project, we propose to investigate new directions starting from algorithms developed in the frame of artificial intelligence and deep learning. For example we plan to investigate the Metropolis Algorithm and the Highly Parallelized Neighborhood Algorithm (Sambridge, 1999a) for improving the determination of about 350 main belt asteroid masses that perturb heavily the inner planet orbits. These perturbations are the most limitative factor for the accuracy of the modern planetary ephemerides (Kuchynka *et al.*, 2010) but are also a important source of information for measuring the masses of these poorly known objects (only about 50 asteroids have their masses known in precision). By improving the characterization of these masses we will improve the accuracy of the Mercury orbit (important for the Bepi-Colombo mission (Iess *et al.*, 2021)), the orbit of Mars (fundamental for ExoMars missions) but also for the Earth (which is a critical point for paleoclimatology and the ERC AstroGéo). We will also apply these algorithms for constraining the distribution of mass in the Kuiper belt. We will do this application together with a densification and re-analysis of old observations for Uranus and Neptune which are very sensitive to the perturbations induced by these big but far objects. This aspect is also of crucial interest for the ERC AstroGéo as an important keypoint for the improvement of long-term ephemerides. Finally, as in (Fienga *et al.*, 2020b; Bernus *et al.*, 2019; 2020), we will apply these AI approaches to test alternative theories of gravitation.

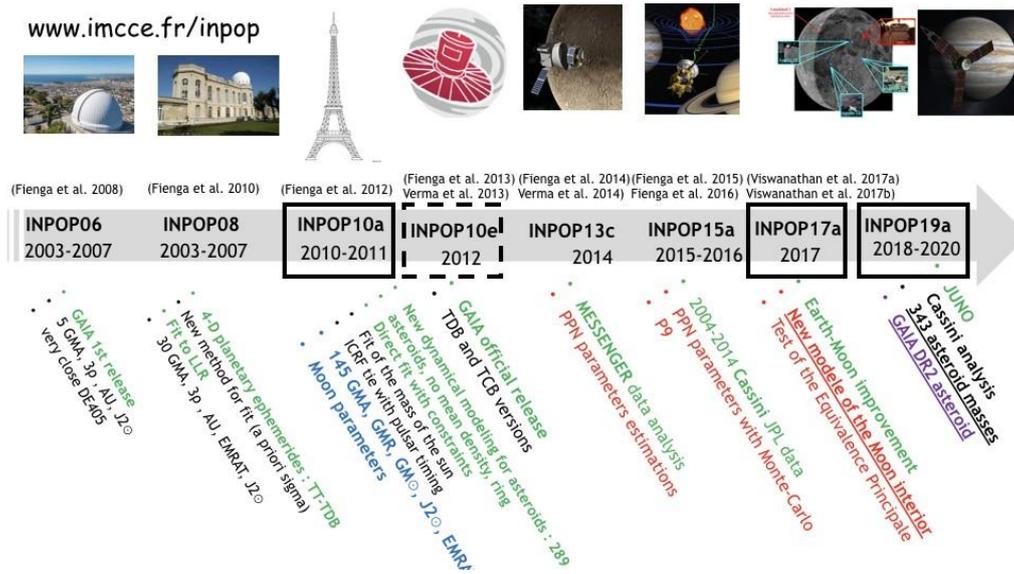


Figure 1: INPOP planetary, lunar and asteroid ephemerides Evolution

1 INTRODUCTION

The INPOP (Intégrateur Numérique Planétaire de l’Observatoire de Paris) planetary ephemerides have started to be built in 2003 by J. Laskar and A. Fienga and their team. With the years, INPOP has become a worldwide standard acknowledged by the community (Archinal *et al.*, 2018; Hilton *et al.*, 2016). On Figure 1, are summarized the different evolutions of the INPOP planetary ephemerides: from the first 4-D position, velocity and time-scale ephemerides in 2003 (INPOP06 and its version delivered for the Gaia DPAC, INPOP10e) and the improved Earth-Moon modeling in 2017 (INPOP17a) and its applications in selenophysics (Viswanathan *et al.*, 2019) to the first partially derivative-free inversion algorithms applied to fundamental physics testings and asteroid mass determinations (Fienga *et al.*, 2015; Bernus *et al.*, 2019; 2020; Fienga *et al.*, 2020a) and the implementation of the almost 2 millions of Gaia solar system object observations in INPOP19a (Deram *et al.*, 2021).

The planetary ephemerides are used for:

- Providing very accurate positions, velocities for 14,000 objects of the solar system (major and minor planets as well as the Earth and Moon orbital and rotational states and relativistic time-scales) to common users by also space agencies, such as CNES and ESA. The INPOP planetary ephemerides are used by ESA for the navigation and the data analysis of the missions Gaia, Bepi-Colombo and JUICE. The INPOP planetary ephemerides are accessible to the whole public on the IMCCE website www.imcce.fr/inpop.
- Enhancing the scientific return of space missions such as Mars orbiters (ExoMars, MEX, MRO...), Venus Express (VEX) but also Cassini and Juno missions. For example, with INPOP and the Cassini data, the mass of the Kuiper belt and constraint the existence of unknown and still undiscovered planet P9 have been estimated (Di Ruscio *et al.*, 2020; Fienga *et al.*, 2020b). VEX data were also used to improve mass estimations of main belt asteroids together with Mars orbiters navigation (Verma *et al.*, 2013) as well as solar physics parameters.
- Characterizing solar physics parameters such as the sun oblateness ((Fienga *et al.*, 2015)), the electronic densities of the fast and slow winds ((Verma *et al.*, 2013))
- Testing laws of gravitation in the solar system. Since 2015, INPOP has been intensively used for providing constraints on possible violation of General Relativity (Verma *et al.*, 2014; Fienga *et al.*, 2015) or for testing alternative theories of gravitation in the solar system (Bernus *et al.*, 2019; 2020).
- Estimating the mass distribution in the solar system with the introduction of the latest observations obtained with the Cassini and Juno missions ((Di Ruscio *et al.*, 2020)). With (Fienga *et al.*, 2020a), 100 main belt asteroid masses have been estimated with an accuracy better than 30%. This result doubles the number of asteroid masses obtained with such accuracy.
- Calibrating the datation of geological samples for paleoclimatologic studies. This latest aspect is at the center of the ERC AstroGeo. Indeed, for dating geologic samples, a long-term ephemeris LaX is used for estimating the Earth

insolation (Laskar *et al.*, 2011) and LaX is calibrated in using INPOP. Furthermore as it has been demonstrated in

(Laskar, 1999), uncertainties on present ephemeris can have an important impact on its long-term extension, LaX. So INPOP improvements will impact LaX and geological datation.

g) Limiting the source of uncertainties for the detection of Cosmological Background (CMB) gravitational waves with pulsar timing arrays. The CMB gravitation waves are different from the one detected by Virgo/Ligo as they are issued from the first ages of the Universe (expansion). As it has been demonstrated (Caballero *et al.*, 2018; Vallisneri *et al.*, 2020), the positions and velocities uncertainties of giant planet orbits are a limiting factor in the bayesian detection of CMB signatures in all pulsar timing arrays. Collaborations between the INPOP team and the European Pulsar Timing array (EPTA) and the US one (Nanograv) have started since 2019 on the question of the assessment of the planetary ephemerides accuracies.

2 STATE-OF-THE-ART

The regularly improved accuracies of the spacecraft navigation or radio experiment (below 1 m for Mars orbiters, about 1 cm for Bepi-Colombo, 20 m for Cassini and Juno missions) as well as the increasing number of data (i.e. 50,000 Mars orbiters observations and about 2 millions for the Gaia DR2) make the inversion of this complex dynamical modeling counting for up to 42,000 unknowns very challenging. When for the case of the construction of the 14,000 main belt asteroids, efficient linear algebra algorithm has been implemented successfully (Deram *et al.*, 2021), some cases remain complex to handle in the frame of the classic least-square minimization.

A good example is the problem of the main belt asteroid perturbations on the inner planets. As one can see on Figure 2, the main belt gathers hundreds of thousand of objects of different sizes and of unknown masses. The projection of these object perturbations on the inner planet orbits along the line-of-observations (i.e. the earth-mars distance as on Fig. 2) is non linear and induces correlated displacements that can reach several hundreds of meters, far above the present accuracies of the Mars orbiter observations (80 cm over 10 years). The individual contributions are then very difficult to disentangle and the determination of the masses becomes impossible without introducing specific algorithms. In 2020 (Fienga *et al.*, 2020a), we introduced a Monte Carlo Least-Square (MCLS) algorithm as a first attempt to combine Monte Carlo exploration of possible density constraints extracted from spectroscopic surveys and Least-Square minimization. We obtained 100 asteroid masses over 350 with an accuracy better than 30% but 250 asteroid masses are still not estimated accurately. Furthermore one can question the selection of 350 asteroids proposed by (Williams, 1984) as the most efficient representation of the asteroid perturbations.

The same type of approach combining Monte Carlo (MC) or Genetic Algorithm (GA) and least-square minimizations has been used for the estimation of highly correlated parameters imposed by alternative theories of gravitation (Fienga *et al.*, 2015; Bernus *et al.*, 2020) or for the localisation of the unknown and still undiscovered planet P9 (Fienga *et al.*, 2020b).

The good side of these mixed methods is the extension of the parameter domain exploration via the MC or the GA with a limited cost of time as these random walks are operated for a reduced number of parameters (the parameters of interest such as the angular positions and distances of P9) but not for all the INPOP parameters. The inconvenient of such methods is the use of least-square minimization for a non-linear model which can lead to local minima even if associated with a wider exploration of the parameter domain. Furthermore, up to now, the analysis of the obtained results were operated in using frequentist paradigm (χ^2 likelihood).

The goal of the proposed PhD is to develop a complete bayesian approach, never applied to the construction of ephemerides, in order to test new ways for disentangling signatures and exploring wider range of parameters in two major fields of application: the distribution of mass in the solar system and tests of alternative theories of gravitation.

3 DESCRIPTION OF THE PROJECT

The goal of this project is to create a new planetary ephemerides called B-INPOP, which will be the bayesian version of INPOP. The basic dynamical modeling will remain the same but the least-squares adjustment will be associated or totally replaced by a derivative-free approach. Several strategies will be investigated.

1) At the beginning, the student will have to investigate the case of the Post Parameterized Newtonian (PPN) parameters β and γ in the frame of the Bepi-Colombo simulations. This work is currently under development. Results have already been obtained with a frequentist method (χ^2 likelihood). PDFs of the problem have been obtained and a

preliminary computation of the bayesian integrals has been done. From there, the student will have to implement a Metropolis algorithm (Tarantola, 1987) and see how efficient it could be in comparison with the results obtained with the likelihood. In a second step highly parallelized Neighborhood Algorithm (Sambridge, 1999a;b) will also be tested. Mixed strategies (MC for the two PPN parameters and least-square minimization for the rest of the parameters) or fully random samplings (priors for all the INPOP parameters, from planetary orbit initial conditions, oblateness and mass of the Sun, Earth-Moon mass ratio to asteroid masses but also PPN) will be tested and compared to the likelihood

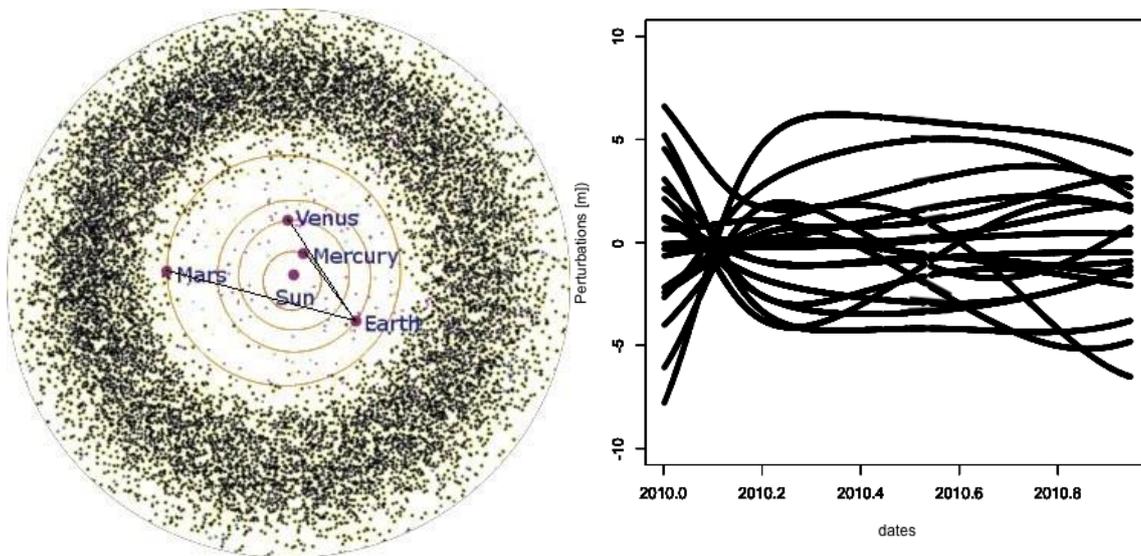


Figure 2: Main belt perturbations on inner planer orbits. On the right-hand side, is plotted an example of perturbations induced by 20 asteroids on the Earth-Mars distances over 1 year.

results. In terms of time of computation, the random sampling is faster than the mixed strategy (1 h for the MCLS versus 10 mn for the full random sampling when running on 40 processors Intel Xeon 2.3GHz) but requires a more important number of runs. A first publication should come out from this first step.

2) In a second step, a specific focus will be address the determination of the asteroid masses. This problematic has already been experimented with a MCLS approach (Fienga *et al.*, 2020a) but in this project the student will be asked to generalize the results that he or she will obtain with the PPN parameter determination to a more complex problem as the number of parameters to consider will be of about 350. Using the method developed by (Kuchynka and Folkner, 2013), the student will investigate if additional asteroids are not required in the model and will complete the INPOP dynamical modeling accordingly. Tests involving additional ring implementation will also be considered as a way to disentangle more easily the point-mass object signatures from signatures induced by objects too small to be detected. The same approach will be implemented and tested for the Kuiper belt objects. This task will be associated with an update of the Uranus and Neptune old optical observations based on a new analysis of their satellites observations, taking into account the Gaia catalog to provide improved positions for the planets. At least a second publication should come out from the step 2.

3) Additional applications to other types of theories of gravitation such as the dilaton theory would also be investi- gated.

4) During the PhD a Monte Carlo Markov Chain will be implemented in order to assess the planetary ephemerides accuracies. The assessment will be firstly compared to the INPOP covariance matrix estimated with the partial deriva- tives and distributed to the Pulsar timing array communities (EPTA

and Nanograv). Publications in collaboration are planned.

We also expect the student to propose other strategies that can emerge from an intensive work of bibliography and discussion with UCA colleagues (i.e. Géoazur, Lagrange, Ecole des mines) using derivative-free inversion methods. In terms of computational power, we are currently using the PSL mesocentre but we will start using the UCA OPAL cluster this fall.

4 EXPECTED OUTPUTS

With the implementation of such bayesian approach and the creation of B-INPOP, we expect three major breakthroughs in planetology, fundamental physics and paleoclimatology:

1) With the determination of 350 asteroid masses, at least as this list could be extended, we will obtain the first lot of consistent mass determinations of this size. In collaboration with M. Delbo (Lagrange, OCA) we will combine the obtained masses with volumes extracted from the MP3C database for the 350 main belt bodies among which we have

planetesimals (expected 170) and collisionally generated asteroids. From the comparison of body bulk density to the composition and the mineralogical analogues we will infer the porosity, which informs about the internal structure. We will use composition and internal structure of planetesimals to constrain the properties of the different stages of their formations and we will infer properties of the protoplanetary disk collisional environments.

2) With the determination of the B-INPOP mass distribution in the Kuiper Belt, besides the new input for the model of solar system formation, we will improve significantly the accuracy on the Uranus and Neptune orbits which are key elements for the paleoclimatological datation. The mass distribution in the Kuiper belt plays an important role for the localisation of the whole solar system as it can induce a significant displacement on the solar system barycenter (SSB) relative to the sun. A displacement of about 100 km of the SSB (this is what was obtained with INPOP20a and the latest update of the Kuiper belt mass) leads to a modification of the Earth isolation of several degrees. The topic is directly associated with the ERC AstroGo, co-funder of this project, its aim being the improvement of the Astronomical Solutions over Geological Time.

3) With the B-INPOP implementation and its bayesian approaches for testing GR and alternative theories, we will increase the scientific return of the Bepi-Colombo mission by considering complexed theories such as the dilaton and SME (Standard Model Extension) theories. Based on recent simulations not yet published, we also expect to gain a factor 35 for the determination of the PPN parameter γ and about a factor 4 on PPN β thanks to the combination of Bepi-Colombo radio-science measurements and to derivative-free technics.

4) A better assessment of the accuracy of the INPOP planetary ephemerides, specifically for the outer planet orbits, will also help for a better description of the pulsar timing sensitivity to the planetary orbits and for a better description of the planetary priors into the CMB gravitational wave detections.

In conclusion, this project aims at developing B-INPOP, the new generation of INPOP planetary ephemerides, a worldwide used and ERC granted ephemeris, that has shown its important impact in a wide field of topics from planetology, fundamental physics, paleoclimatology. With the new approaches that are proposed to be investigated in this project, B-INPOP will increase the scientific return of ESA missions such as Bepi-Colombo. It will participate at a new understanding of the stages of the solar system formation as well as a better calibration of geological datation, a mandatory step for the long-term studies in Earth climate evolution. Finally it will contribute to the detection of the relic CMB gravitational waves. The risks for this project are very low because of the well-established INPOP team and its strong collaboration in the field and a clear program of publication.

5 TECHNICAL ASSESSMENTS

The technical assessments for this project are based on the configuration used on the MesoPSL cluster since 2016. In the least-square configuration, each job is composed by the iteration of the following loop: the numerical integration (parallelized) of the planet and asteroid equations of motions over 1 century (10 mn), the computation of the residuals by comparison with observations (3 mn), the construction of the linearized Newton-Gauss equation (1 mn) and its

inversion (3 mn). The stop of the iteration is given by the convergency of the χ^2 to a minimum. It usually takes 4 hours and each solution occupies, after removing auxiliary files, 150 Mb. The analysis of the results are very fast, based on R routines. In the bayesian configuration, the inversion of the jacobian matrix and the iterations are not necessary, leading to a job duration of about 10 mn and a size of the solution of less than 15 Mb. After analysis, only few Kb remains per solutions. In term of memory, 100 runs in the least square configurations use less than 32 Gb, the matrix inversion consuming the maximum of the memory. With the bayesian approach, the use is even less and should not be a limiting factor. A first step of development and construction of the bayesian approach will be done on our local cluster in Goazur (machine gpm with 72 nodes and 32 Gb of memory). An installation on Azzura will then be operated. By considering 4 cores of 32 nodes each, we will launch 160 runs per 10mn for the bayesian configuration. In asking for 400,000 hours this year over 4 cores, we guarantee 2,4 millions of runs for the bayesian approach which should be enough for the installation and the validation of the procedure as well as the first operational runs of the project. In the next three years, we should continue to work on the same basis.

REFERENCES

- A. Fienga, J. Laskar, P. Exertier, H. Manche, and M. Gastineau, *Celestial Mechanics and Dynamical Astronomy* **123**, 325 (2015).
- A. Fienga, C. Avdellidou, and J. Hanus, *MNRAS* **492**, 589 (2020a).
- M. Sambridge, *Geophysical Journal International* **138**, 479 (1999a).
- P. Kuchynka, J. Laskar, A. Fienga, and H. Manche, *A&A* **514**, A96 (2010), arXiv:1004.3119 [astro-ph.EP] .
- B. A. Archinal, C. H. Acton, M. F. A'Hearn, and al., *Celestial Mechanics and Dynamical Astronomy* **130**, 22 (2018).
- J. L. Hilton, C. Acton, J.-E. Arlot, and al., *Transactions of the International Astronomical Union, Series A* **29A**, 22 (2016).
- V. Viswanathan, N. Rambaux, A. Fienga, J. Laskar, and M. Gastineau, *Geophys. Res. Lett.* **46**, 7295 (2019), arXiv:1903.07205 [physics.geo-ph] .
- P. Deram, A. Fienga, A. K. Verma, M. Gastineau, and J. Laskar, *Celestial Mechanics and Dynamical Astronomy* (2021).
- A. Di Ruscio, A. Fienga, D. Durante, L. Iess, J. Laskar, and M. Gastineau, *A&A* **640**, A7 (2020).
- A. K. Verma, A. Fienga, J. Laskar, K. Issautier, H. Manche, and M. Gastineau, *A&A* **550**, A124 (2013), arXiv:1206.5667 [physics.space-ph] .
- A. K. Verma, A. Fienga, J. Laskar, H. Manche, and M. Gastineau, *A&A* **561**, A115 (2014), arXiv:1306.5569 [astro-ph.EP] .
- L. Iess, S. W. Asmar, P. Cappuccio, G. Cascioli, and al., *Space Sci. Rev.* **217**, 21 (2021).
- A. Fienga, A. Di Ruscio, L. Bernus, P. Deram, D. Durante, J. Laskar, and L. Iess, *A&A* **640**, A6 (2020b).
- L. Bernus, O. Minazzoli, A. Fienga, M. Gastineau, J. Laskar, and P. Deram, *Phys. Rev. Lett.* **123**, 161103 (2019), arXiv:1901.04307 [gr-qc] .
- L. Bernus, O. Minazzoli, A. Fienga, M. Gastineau, J. Laskar, P. Deram, and A. Di Ruscio, *Phys. Rev. D* **102**, 021501 (2020), arXiv:2006.12304 [gr-qc] .
- J. Laskar, A. Fienga, M. Gastineau, and H. Manche, *A&A* **532**, A89 (2011), arXiv:1103.1084 [astro-ph.EP] .
- J. Laskar, *Philosophical Transactions of the Royal Society of London Series A* **357**, 1735 (1999).
- R. N. Caballero, Y.J. Guo, K. J. Lee, and al., *MNRAS* **481**, 5501 (2018), arXiv:1809.10744 [astro-ph.EP] .
- M. Vallisneri, S. R. Taylor, J. Simon, and al., *ApJ* **893**, 112 (2020), arXiv:2001.00595 [astro-ph.HE] .
- J. G. Williams, *Icarus* **57**, 1 (1984).
- A. Tarantola, *Inverse problem theory: methods for data fitting and model parameter estimation*, edited by Elsevier (1987).
- M. Sambridge, *Geophysical Journal International* **138**, 727 (1999b).
- P. Kuchynka and W. M. Folkner, *Icarus* **222**, 243 (2013).

Groupe prospective IA du laboratoire Lagrange

Responsable scientifique : Sylvain Rousseau

Laboratoire : Lagrange

Description :

Nous formons un petit groupe d'ingénieurs informaticien au sein du laboratoire Lagrange, encouragés par la direction à acquérir un savoir-faire dans différentes méthodes de Machine Learning (dont les réseaux de neurones), afin d'être appliquées à des problématiques actuelles émergeant dans notre domaine (traitement du signal en astronomie/astrophysique).

En plus d'une plateforme où l'on pourra mettre en pratique notre expérience des codes acquis lors de nos formation en IA, l'utilisation d'Azzurra nous permettrait d'aborder des problèmes d'astrophysique exigeants en terme de puissance de calcul. Par exemple, l'extraction et la reconnaissance de signal dans les données des "data challenges" des futurs satellites LISA (détection d'onde gravitationnelle) nécessiteront le développement de méthodes "deep learning" sur de gros volumes de données.

Biologie, santé

EUR Life

Cibler l'hétérogénéité de réponse cellulaire aux anticancéreux pour augmenter leur efficacité thérapeutique (brevet Phénotype-Avancé Génotype)

Responsable scientifique : Jérémie Roux

Laboratoire : IRCAN

Description :

The project is about the development of an algorithmic framework of tumor cells dynamic analyses, integrating machine learning and signal processing.

The objective is to develop a new algorithmic architecture integrating deep learning, and signal processing to predict drug response phenotypes of each cell forming a tumor clone, from dynamic imaging data acquired in the laboratory. To study these partial responses, the evolution of tumor cells treated with cancer drugs is imaged at regular intervals by microscopy. The sequence of images produced makes it possible to observe changes in cell properties (shape, radiometry), their division, and their phenotypic response (cell death or survival). Cells are plated on 96 well plates than imaged with the DV microscope. The output for one experiment, are images from up to 48 hours of acquisition of up to 40 sequences. Each sequence is a stack of 3200 images (about 1 GB): it is composed of up to 4 channels and 800 frames per channel (the number of channels can be increased depending on the experiment).

The aim will be to create a framework for semi-automatic or automatic dynamic analysis of these image sequences in order to understand, and therefore predict, the phenotypic responses to cancer treatments of each cell constituting a tumor clone.

The framework is currently under development on python using recent packages of data representation, image processing, signal processing, and machine learning (Tensorflow, keras, sklearn).

These analysis processes will make it possible to better understand the molecular origin of the variability of cellular responses to cancer drugs, and to highlight new therapeutic targets and biomarkers of sensitivity to the drugs tested.

Etude de la contribution d'une protéine de liaison aux ARNs conservée dans la localisation des ARNs synaptiques et recherche systématique de ses cibles

Responsable scientifique : Florence Besse

Laboratoire : IBV

Description :

Le projet consiste à analyser différents jeux de données de séquençage :

- RNA-Seq sur des échantillons de synaptosomes de drosophiles sauvages et mutants
- iCLIP sur des protéines de liaison aux ARNs sauvages et mutantes
- HyperTRIBE sur des protéines de liaison aux ARNs sauvages et mutantes

Les ressources Azzurra sont essentielles à l'alignement des séquences sur les génomes de références.

Automatic cell type annotation for cell atlas construction

Responsible scientifique : Christophe Bécavin

Laboratoire : IPMC

Description :

Chronic respiratory diseases (COPD, fibrosis, asthma, cystic fibrosis, ciliopathies) affect several hundred million people worldwide and the COVID pandemic has recently drawn world attention on the impact of infectious respiratory diseases. The tissue destruction that is associated with these diseases is often followed by a strong tissue remodeling of the airway tract, one of body's first line of defense. To better understand the importance of the different resident cell types in these processes and determine useful new targets for developing the therapeutical arsenal, we are developing single-cell analyses.

Single-cell technologies are becoming essential in many research projects to analyze complex biological systems. The power of these approaches, which allow the determination of cell composition and delineation of complex transcriptional regulatory networks in cellular subsets, involves a large panel of numerical methods to integrate properly many different biological modalities (information about RNA expression, splicing, editing, proteins, epigenome, genome, spatiotemporal expression). Many questions are emerging when working such massive data to optimize properly the integration of several separate datasets and be able to make comparisons between different experimental situations. Deep learning methods show great promise to automate classification or dimensionality reduction. We plan to use them in order to build the cellular atlas of the respiratory tree, after identification of the most appropriate deep network structure able to identify the different cell types.

The research project aims to:

- (1) determine the most suitable model of deep neural network to automate cell classification;
- (2) develop a visualization algorithm for dynamically annotate the cell types, linked to the deep learning model;

(3) use these approaches for creating a lung cellular atlas that integrates in vivo and in vitro datasets, and document the variations observed under specific phenotypic/genotypic situations.

Exploring associations between serum cytokine levels and behavior abnormalities in children

Responsable scientifique : Nicolas Glaichenhaus

Laboratoire : Institut de Pharmacologie Moléculaire et Cellulaire

Résultats obtenus grâce aux moyens informatiques attribués :

Martinuzzi, E., S. Barbosa, P. Courtet, E. Olie, S. Guillaume, E. C. Ibrahim, D. Daoudlarian, L. Davidovic, N. Glaichenhaus and R. Belzeaux (2021). "Towards a blood test to differentiate bipolar disorder and major depressive disorder during a major depressive episode." *Brain Behav Immun Health*, Mar10;13:100232; <https://doi.org/10.1016/j.bbih.2021.100232>

Barbosa, S., O. Khalfallah, A. Forhan, C. Galera, B. Heude, N. Glaichenhaus and L. Davidovic (2020). "Serum cytokines associated with behavior: A cross-sectional study in 5-year-old children." *Brain Behav Immun*, Jul;87:377-387. <https://doi.org/10.1016/j.bbi.2020.01.005>

Barbosa, S., O. Khalfallah, A. Forhan, C. Galera, B. Heude, N. Glaichenhaus and L. Davidovic (2020). "Immune activity at birth and later psychopathology in childhood." *Brain Behav Immun Health* 8: 100141. <https://doi.org/10.1016/j.bbih.2020.100141>

Association entre cytokines à la naissance et troubles émotionnels chez l'enfant à 5 ans : réplication et validation. Stage de Benjamin Lapijover à l'IPMC en 2022 (étudiant Master 2 SV BIM à UCA).

Description :

Les cytokines sont de petites protéines solubles produites par les cellules du système immunitaire. A ce titre, elles jouent un rôle déterminant dans la défense de l'organisme contre les agents pathogènes. Toutefois, ces molécules peuvent également perturber le fonctionnement du cerveau et provoquer des anomalies comportementales telles que l'hyperactivité ou des difficultés d'interactions sociales. Afin de mieux comprendre le rôle des cytokines sur le fonctionnement du cerveau, nous avons eu accès à des données cliniques et des prélèvements biologiques qui ont été collectés dans le cadre de la cohorte EDEN, une cohorte française généraliste dont l'objectif est l'étude des

déterminants pré et post natals précoces du développement psychomoteur et de la santé de l'enfant. Ainsi, dans le cadre de la cohorte EDEN, plus de 2000 mères ont été recrutées pendant leur grossesse et suivies, ainsi que leur enfant, pendant plus de 5 ans.

De très nombreuses données psychosociales ont été collectées pendant la grossesse et après la naissance de l'enfant parmi lesquelles, la consommation de tabac et d'alcool de la mère pendant la grossesse, les niveaux d'éducation du père et de la mère, le revenu du foyer, le nombre de frères et de sœurs, les antécédents médicaux des parents, les évènements de santé entre la naissance et l'âge de 5 ans, le poids et la taille de l'enfant. De plus, les enfants ont été évalués à l'âge de 5 ans par un(e) psychologue ce qui a permis de définir, pour chacun d'entre eux, 5 scores cliniques mesurant les performances de l'enfant dans 5 dimensions comportementales: les problèmes comportementaux, les problèmes d'interactions avec les autres, les problèmes émotionnels, l'hyperactivité et les problèmes d'inattention, et le comportement pro-social.

Parallèlement au recueil des données psychosociales, l'étude EDEN a permis de récolter des échantillons de sang qui ont été prélevés chez la mère pendant la grossesse, à la naissance dans le cordon ombilical, et chez l'enfant à l'âge de 5 ans. Ces échantillons ont ensuite été analysés pour la concentration de 33 cytokines différentes.

L'objectif de notre projet est d'effectuer des études d'association de prédiction, afin d'identifier les variables cliniques et biologiques associées aux différentes dimensions comportementales étudiées dans la cohorte EDEN.

Méthode(s) utilisée(s)

Nous utiliserons 3 méthodes/algorithmes de régression pour comparer les résultats :

1- La régression logistique standard dans la cadre des modèles linéaires généralisés qui utilise l'estimateur usuel des moindres carrés.

2- Elastic Net est une méthode de contraction des coefficients de régression qui permet d'ajuster des modèles de régression logistique pénalisée. Elastic Net est normalement

utilisé dans des problèmes de grande dimension ($n \ll p$), un cas où des méthodes plus classiques ne fonctionnent pas. Elastic Net utilise des algorithmes consommateurs de temps de calcul et de stockage. En pratique, Elastic Net est testé de manière itérative pour différentes valeurs des paramètres λ et α de manière à identifier une solution optimale par validation croisée et une mesure de la performance par calcul de l'aire sous la courbe (AUC).

3 – Algorithmes de Gradient boosting (XGBoost, etc) pour créer des modèles prédictifs. L'objectif principal de Gradient Boost est de minimiser la fonction de perte en ajoutant des apprenants faibles à l'aide d'un algorithme d'optimisation de la descente de gradient. L'optimisation des paramètres de l'algorithme peut prendre un certain temps même si celui-ci a été spécialement conçu pour améliorer la vitesse et les performances.

Exécuter ces algorithmes dans un PC prend du temps car la parallélisation est réduite à quelques cœurs et les performances générales de la machine seront réduites.

Hi-C-Genomes

Responsable scientifique : Etienne Danchin

Laboratoire : Institut Sophia Agrobiotech - UMR INRAE - UCA - CNRS

Description :

La reproduction sexuée est ultra-majoritaire chez les animaux si bien que les rares espèces capables de se reproduire en l'absence de reproduction sexuée (par parthénogénèse) sont considérées comme des curiosités et des impasses évolutives. En effet, en l'absence de reproduction sexuée, les animaux ne sont pas capables de réaliser de brassage génétique et donc d'échanger des allèles favorables. Certains de ces animaux ont même perdu la capacité de réaliser la méiose et donc à éliminer via recombinaison les allèles délétères. Ces animaux dits strictement parthénogénétiques sont supposés avoir un désavantage sélectif et adaptatif en comparaison de leurs congénères capables de se reproduire sexuellement.

Dans notre laboratoire, nous travaillons sur des vers ravageurs de cultures responsables de dégâts considérables à l'échelle mondiale. Ces vers appartiennent au genre *Meloidogyne* dans lequel on retrouve à la fois des espèces à reproduction sexuée et d'autres asexuée obligatoire. Etonnamment, les espèces causant le plus de dégâts ont une reproduction asexuée obligatoire et ne font pas de méiose. Ces parthénogénétiques mitotiques sont capables d'attaquer une plus grande variété de plantes et ont une plus grande aire de répartition géographique que leurs congénères se reproduisant sexuellement. Ces observations semblent donc contradictoires avec la théorie qui voudrait que les asexués soient désavantagés et s'adaptent moins bien que les sexués.

Comment les animaux sont capables d'évoluer, de persister et de s'adapter en l'absence de sexe reste très peu connu et constitue une énigme dans le domaine de l'évolution. Pour en apprendre plus, nous avons réalisé le séquençage du génome d'une de ces espèces,

Meloidoyne incognita en 2008 à l'aide de techniques de séquençage de 1ère génération⁴⁶. Ces travaux nous ont permis de découvrir que le génome avait une structure dupliquée particulière mais celui-ci étant incomplet il ne nous a pas été possible de lier cette structure à une capacité à survivre sans sexe. Plus récemment, grâce à l'emploi de techniques de séquençage de 2ème génération nous avons pu améliorer ce premier génome et séquencer ceux de deux autres espèces du genre Meloidogyne. Ces travaux ont montré que ces espèces sont polyploïdes et ont subi des événements d'hybridations⁴⁷.

Grâce à l'obtention d'un grand projet de séquençage France Génomique, nommé ALPAGA et dont je suis coordinateur, nous avons pu générer des données génomiques avec des technologies de séquençage d'ADN de 2ème et 3ème génération. Ces données nous ont permis de réaliser un bon considérable dans l'assemblage de ces génomes, passant de génomes fragmentés en dizaines de milliers de morceaux à des génomes fragmentés en centaines de morceaux. Ces fragments moins nombreux et plus grands nous ont offert un niveau de résolution inégalé sur la structure du génome.

Cependant, nous sommes encore loin d'avoir assemblé les génomes de ces différentes espèces à un niveau de résolution chromosomique puisque ces organismes possèdent 40 à 50 chromosomes.

De ce fait, notre connaissance de l'arrangement et de l'organisation des gènes le long de ces génomes reste encore limitée par cette fragmentation du génome. Récemment des techniques dites 3-C pour chromosome capture conformation et leurs dérivés de type Hi-C (pour high-throughput) permettent d'obtenir des informations de contact dans la chromatine et grâce à ces informations des outils bioinformatiques sont capables d'assembler les génomes à un niveau de résolution chromosomique⁴⁸. Ce bond en avant technologique a permis d'obtenir un assemblage de génome à l'échelle des chromosomes

⁴⁶ Abad, P. et al. Genome sequence of the metazoan plant-parasitic nematode *Meloidogyne incognita*. *Nat. Biotechnol.* 26, 909–915 (2008).

⁴⁷ Blanc-Mathieu, R. et al. Hybridization and polyploidy enable genomic plasticity without sex in the most devastating plant-parasitic nematodes. *PLOS Genet.* 13, e1006777 (2017).

⁴⁸ Ghurye, J. & Pop, M. Modern technologies and algorithms for scaffolding assembled genomes. *PLOS Comput. Biol.* 15, e1006994 (2019).

pour un nématode parasite de plante⁴⁹ et permet d'espérer des résultats similaires chez les nématodes du genre *Meloidogyne* se reproduisant par parthénogénèse.

L'une des méthodes d'assemblage les plus performantes basée sur les informations de contact dans la chromatine Hi-C est la méthode Graal⁵⁰, implémentée dans le programme instaGraal⁵¹. Ce programme requiert l'utilisation intensive de GPU via la technologie cuda core sur cartes NVIDIA.

L'objectif de notre demande est de pouvoir accéder à des ressources GPU NVIDIA afin d'utiliser le programme instaGraal pour assembler les génomes de différentes espèces de nématodes présentant des modes de reproduction différents. Nous ne disposons pas de ressources de type GPU dans notre laboratoire et n'avons pas non plus accès à de telles ressources à l'extérieur.

A ce jour nous disposons de données Hi-C pour les génomes de 3 espèces, *M. incognita*, *M. javanica* et *M. arenaria*, et des données vont être générées dans les prochaines semaines pour 2 autres espèces. Nous espérons, grâce à ces données Hi-C et à l'utilisation d'instaGraal sur des cartes GPU NVIDIA pouvoir reconstruire les différents génomes à un niveau de résolution chromosomique. Ceci nous permettra d'enfin résoudre la structure de ces génomes, leur histoire évolutive et l'arrangement de la suite des gènes le long des chromosomes.

⁴⁹ Lian, Y. et al. Chromosome-level reference genome of X12, a highly virulent race of the soybean cyst nematode *Heterodera glycines*. *Mol. Ecol. Resour.* 19, 1637–1646 (2019).

⁵⁰ Marie-Nelly, H. et al. High-quality genome (re)assembly using chromosomal contact data. *Nat. Commun.* 5, (2014).

⁵¹ Baudry, L. et al. instaGRAAL: chromosome-level quality scaffolding of genomes using a proximity ligation-based scaffolder. *Genome Biol.* 21, 148 (2020).

Droit, économie

EUR ELMI, LexSociété

Estimation de modèles complexes en économie et finance à base de simulations

Responsable scientifique : Veronika Czellar

Laboratoire : INSTAR, SKEMA

Résultats obtenus grâce aux moyens informatiques attribués :

Un papier soumis au Journal of Econometrics fin 2022 :

Bastien Buchwalter, Veronika Czellar, Engin Iyidogan. Multifractal Cryptocurrencies.

Description :

Le but du projet est de développer et d'utiliser des méthodes à base de simulations qui construisent des modèles capables de reproduire les données et de les prédire. Les applications en économie et finance sont particulièrement gourmandes en calculs car les données financières utilisées sont de grande taille, et ces méthodes à base de simulations nécessitent la simulation de pseudo-données de taille similaire, puis la minimisation de la distance entre les pseudo-données simulées et les vraies données.

OCR chinois

Responsable scientifique : Frédéric Constant

Laboratoire : ERMES, UFR de Droit

Description :

Les membres du groupe de recherche/du projet Numerica Sinologica se sont réunis à partir du constat partagé de la nécessité de développer des outils de reconnaissance de l'écriture manuscrite et xylographique spécifiques au chinois classique, et plus largement aux langues asiatiques utilisant des sinogrammes, comme préalable au développement des humanités numériques dans les études sur l'histoire et les langues de l'Asie orientale. Les chercheurs associés au projet sont membres des trois principales unités de recherche sur la Chine impériale en France, le CCJ (UMR 8173 Chine, Corée, Japon), le CRCAO (Centre de recherche sur les civilisations de l'Asie orientale, UMR 8155) et l'IAO (Institut d'Asie orientale, UMR 5062). Au-delà des laboratoires concernés, des centres de recherche et des bibliothèques en Europe nous ont également fait part de leur intérêt et de leur soutien à notre projet.

Les implications du développement de modèles de reconnaissance de caractères fiables pour notre discipline sont très larges et permettraient à la sinologie de véritablement prendre le tournant des humanités numériques et de rattraper son retard par rapport à d'autres champs des sciences humaines et sociales. Outre la constitution de bases de données, l'analyse quantitative, la fouille de textes ou encore le traitement automatique du langage, de tels outils permettraient de diffuser auprès du grand public des collections de documents dont certaines ne sont pour le moment accessibles que via la souscription d'abonnements auprès d'organismes privés. Au-delà de l'innovation technologique, il y a donc un véritable enjeu scientifique et de diffusion du savoir. Notre projet s'inscrit par ailleurs dans une démarche de science ouverte, et les codes informatiques seront rendus publics sans aucune restriction (par exemple sur GitHub).

La reconnaissance automatique de l'écriture manuscrite du chinois fait face à des contraintes particulières qui expliquent pourquoi, alors que des plateformes

(Transkribus, eScriptorium, Calfa Vision) existent déjà pour les écritures de type alphabétique, rien de tel n'est pour le moment accessible à la communauté des sinologues. Cette situation est due tout d'abord aux spécificités du chinois, pour lequel il est difficile de transposer les méthodes déjà utilisées pour d'autres langues sans les adapter. Le chinois est composé de milliers de signes ce qui nécessite une puissance de calcul beaucoup plus importante que pour une langue alphabétique ou de développer de nouvelles approches afin d'entraîner des modèles performants. Les outils existants ont en outre été conçus pour des écritures horizontales, ce qui oblige à développer des modèles de reconnaissance des zones de texte et lignes de base adaptés aux écritures verticales. Le contexte académique est également peu favorable, la reconnaissance des textes en chinois classiques n'a pour le moment pas été jugée prioritaire en France ou en Europe, et les chercheurs chinois n'ont de leur côté pas encore développé de solutions efficaces et accessibles librement à la communauté des chercheurs. C'est ainsi que les projets existants offrent soit des résultats limités, soit des solutions propriétaires et extrêmement coûteuses, dont il est par ailleurs difficile d'évaluer les performances.

Notre groupe s'est constitué en mai 2021, en collaboration avec le projet eScriptorium, dirigé par Daniel STOEKL (EPHE), qui a développé une infrastructure de reconnaissance optique des langues à partir de Kraken, un logiciel open-source d'OCR basé sur l'intelligence artificielle et développé pour les langues écritures non-latines. Nous avons, grâce à leurs outils, entraîné tout d'abord des modèles de segmentation et d'analyse de la mise en page des textes chinois afin d'identifier les zones de texte à reconnaître dans les documents analysés. Nous avons également constitué un jeu de données de correspondances entre des images et leur transcription suffisamment large pour couvrir l'ensemble des caractères chinois usuels dans les sources classiques. Cette étape préalable à l'entraînement d'un premier modèle a pu être financée par des contributions financières du CCJ, du CRCAO et de l'IAO. Au cours des prochaines semaines, les ingénieurs d'eScriptorium modifieront Kraken afin de l'adapter aux spécificités du chinois, et nous serons alors en mesure d'entraîner un premier modèle. Fort de ce modèle nous serons en mesure, en fonction de son efficacité, de démontrer la viabilité de notre démarche et notre capacité à développer un outil performant.

En l'état, les infrastructures existantes ne permettront cependant pas à des utilisateurs inexpérimentés (c'est-à-dire ne maîtrisant pas les langages informatiques) d'utiliser ce modèle pour obtenir la transcription d'un corpus d'images donné, encore moins de

l'adapter pour un usage spécifique, en raison par exemple des caractéristiques propres à un type particulier de sources (variété des écritures manuscrites et xylographiques). Nous souhaitons donc aller plus loin et offrir à toute personne la possibilité d'utiliser sans contraintes techniques ces outils (comme le fait par exemple Transkribus et son outil de saisie assistée par ordinateur permettant l'entraînement de modèles par l'utilisateur). Notre objectif est ainsi de nous organiser pour réunir les fonds nécessaires pour: 1. finaliser les phases d'entraînement nécessaires au perfectionnement d'un premier modèle ; 2. proposer un outil permettant la reconnaissance de caractères chinois anciens et le développement de modèles adaptés aux besoins des chercheurs sans qu'il soit nécessaire d'avoir des compétences informatiques particulières. Cet outil pourrait prendre la forme soit d'un portail en ligne offrant à la fois une interface facile d'accès et une puissance de calcul suffisante, soit d'une suite logicielle pouvant être installée sur un poste de travail et réutilisée au sein des laboratoires à l'aide d'un support technique classique.

Projets industriels

K-EPSILON, dont le siège est basé à Sophia Antipolis, bénéficie d'un accès aux ressources de calcul d'Azzurra pour ses différents projets de R&D (voir RA 2020-21).

K-EPSILON est un leader dans la simulation de l'hydrodynamique et de l'aérodynamique pour les applications marines et offshore.

Fondée en 2003 et pionnière en CFD, K-EPSILON propose une nouvelle génération de solutions sur mesure, des études numériques spécifiques, pour simuler et prédire le comportement de systèmes complexes en milieu marin et offshore.

Résultats obtenus grâce aux moyens informatiques attribués :

Faye, Antoine & Carrere, N & Sacher, Matthieu & Hauville, F & Nême, A. (2022).

Determination of the equivalent structural properties of a composite hydrofoil. 25ème Congrès Français de Mécanique. Nantes, 20 août au 2 septembre 2022

https://www.researchgate.net/publication/363406203_Determination_of_the_equivalent_structural_properties_of_a_composite_hydrofoil

Corentin Lothode, Jules Poncin, Didier Lemosse, David Gross, Eduardo Souza de Cursi, Investigation of blade-mast fluid-structure interaction of a tidal turbine,

Ocean Engineering, Volume 261, 2022, 112046, ISSN 0029-8018,

<https://doi.org/10.1016/j.oceaneng.2022.112046>

OPAL

Azzurra fait partie de l'infrastructure OPAL, dont une présentation peut être consultée [sur cette page](#).

Au cours de l'année 2022, **137** utilisateurs ont pu bénéficier d'une accréditation OPAL **par l'université** (+8,73 % par rapport à 2021), parmi lesquels :

- **77** utilisateurs disposant d'un compte sur Azzurra pour un projet en cours de validité.
 - **2** d'entre eux ont aussi demandé un compte sur le cluster Licallo (OCA)
 - **11** d'entre eux ont aussi demandé un compte sur le cluster NEF (Inria)
- **60** utilisateurs sans compte sur Azzurra, dans le but de calculer sur un autre cluster OPAL (tous sur NEF). Ce sont **4** projets de recherche et **1** cours, décrits en page suivante.

20 personnels accrédités par une autre entité (12 par OCA et 6 par Inria) disposent quant à eux d'un compte sur Azzurra.

L'accréditation OPAL donnée par UCA est valable un an, renouvelable après [envoi d'un rapport des publications de l'année écoulée](#).

**Projets accrédités OPAL par l'université ne calculant pas
sur Azzurra**

Design de metasurfaces pour système flashlidar

Responsable scientifique : Stéphane Lanteri

Laboratoire : Inria

Plateforme OPAL : NEF (Inria)

Nombre d'accrédité : 1

Résultats obtenus grâce aux moyens informatiques attribués :

Enzo Isnard (stagiaire alternant MAM5 INUM 2021-2022) a utilisé le cluster NEF et présenté un poster à la conférence "2nd Colloquium on Physics and Applications of Metasurfaces", Florence, 18-22 juillet 2022.

Description :

Les metasurfaces sont des composants optiques planaires constitués d'une assemblée de structures sub-longueur d'onde réparties sur une grille dont la période est également inférieure à la longueur d'onde d'illumination. Ces structures ont des tailles, des formes et des orientations différentes, afin que l'onde qui traverse la metasurface puisse accumuler une phase qui n'est pas la même en tout point de l'interface. La modulation du profil de phase permet ainsi de contrôler le comportement de la lumière en sortie. Selon l'application souhaitée, il suffit d'adapter ce profil afin de fabriquer une lentille pour focaliser la lumière, ou un élément optique diffractif pour projeter des images holographiques, ou pour ajuster la réflexion, la réfraction, la polarisation de l'onde, ... Ce projet a pour but de concevoir et de fabriquer des metasurfaces qui seront intégrées à des dispositifs d'imagerie LIDAR (light detection and ranging), le tout, en ayant les efficacités les plus élevées possibles. Il s'inscrit dans le cadre d'une collaboration en cours entre l'équipe de Patrice Genevet au CRHEA et l'équipe Atlantis du Centre de Recherche Inria d'Université Côte d'Azur, avec Stéphane Lanteri comme correspondant. L'accès à ces ressources de calcul nous permettra de réaliser des simulations électrodynamiques type FDTD (Finite Difference Time Domain), DGTD (Discontinuous

Galerkin Time Domain) ou bien encore GDM (Green Dyadic Method) pour déterminer les propriétés optiques des structures qui constituent les métasurfaces. Ces simulations sont très coûteuses en temps de calcul et en mémoire RAM, cependant elles peuvent être fortement accélérées avec la parallélisation des codes. D'autre part, nous envisageons d'utiliser une approche de deep learning pour accélérer la prédiction des propriétés optiques de nos composants micrométriques. La possibilité d'accéder à un grand nombre de nœuds GPU, à disposition au centre de calcul Nef, serait un réel atout pour accroître la vitesse d'apprentissage des réseaux neuronaux.

Seismic imaging of the Alps (LISALPS project - ANR)

Responsable scientifique : Stéphane Operto

Laboratoire : Geoazur

Plateforme OPAL : NEF (Inria)

Nombre d'accrédités : 2

Description :

Scientific objective:

The goal of this project is to develop new 3D quantitative models of the crustal and upper mantle of the Alps from the teleseismic data of the AlpArray experiment (<http://www.alparray.ethz.ch/en/home>) by Full Waveform Inversion (FWI). The Alp Array experiment is an European initiative that allowed for the deployment of a dense array of broadband seismometers covering the entire Alpine Chain. In this framework, our goal is to develop high-resolution models of the Alpine chain by FWI. More precisely, we first want to build a new model of the crust and upper mantle of the Alpine chain with a wavelength-scale resolution and reconstruct several classes of parameters such as P and S anisotropic wavespeeds, density and attenuation. In a second step we will focus on a more focused target centred on the Ligurian knot in the western Alps by pushing the inversion at higher frequencies.

Methods:

FWI is a seismic imaging technique which is formulated mathematically as a partial-differential (PDE)-constrained optimization problem, which aims at reconstructing the constitutive properties of the Earth's interior from seismic wave measurements made at the Earth's surface. The constraint is the elastodynamic equations, which control the propagation of waves in the frame of linear elasticity. The parameters describing the Earth's properties are contained in the coefficients of the elastodynamic equations and are estimated by fitting the seismic measurements with the simulated counterparts. Due to the size of the data and parameter spaces, this optimization problem is solved with

local approaches where the gradient of the objective function can be computed efficiently with the adjoint state method. In the framework of the teleseismic configuration, the target to be imaged is located below the station layout, while the source (the teleseismic event) is located thousand km away from the target. To deal with this configuration, seismic wave simulation is subdivided in two steps. First we simulate the 3D wavefields from the event to the edges of the target in a simplified axisymmetric Earth with four 2D simulations (AXISEM software). Then, we use these wavefields at the edges of the 3D heterogeneous target to simulate waves inside it with the spectral element method (This two step approach is referred to as wavefield injection technique). Once the forward problem is solved we compute the data misfit at receiver and build the gradient of the objective function by correlation between the incident wavefields and the adjoint wavefields computed by time-reversed modeling of the data residuals. The Hessian can be taken into account with limited-memory quasi-Newton method (L-GFGS) or second-order adjoint state method and linear conjugate gradient solver.

Codes:

To tackle this application, we have 3D FWI codes that work on CPUs and GPUs. The codes are parallelized by domain decomposition of the unstructured hexahedral meshes discretizing the Earth's model and right-hand sides distribution (the teleseismic sources, keeping in mind that the image of the Earth is built by summation of the contribution of each source). Ideally a MPI process is assigned to each subdomain and each right-hand side. On GPU, wavefield simulation in a domain is speedup by multithreading. In a pure CPU configuration, we may use OpenMP when available cores are available for this third level of parallelism.

For running the code to image the entire Alpine chain, we have estimated that we need 2 500 000 scalar hours on CPUs. Depending on the CPU generation, we expect a speedup of at least one order of magnitude compared to CPU time (As an example, running SPECSEM3D code on GPU generates a speedup of 25). In terms of storage, we need 30Tb to store the AXISEM solutions at the edges, the datasets, and the FWI results.

This research will be performed in the framework of the LISALPS project, which is funded by ANR. It involves three partners Geoazur, LMA and ISTERre. The PI is Stephane Operto. A PhD thesis and a postdoc position at Geoazur were funded by ANR. The PhD thesis is performed by Najmieh Mohammadi at Geoazur. The postdoc position has been

attributed to Stephen Beller, who is an expert on FWI, teleseismic imaging and high performance computing. The PhD thesis is mainly supervised by Stephen Beller and Vadim Monteiller from LMA.

CPER Ubiquarium

Responsable scientifique : Michel Riveill

Laboratoire : I3S

Plateforme OPAL : NEF (Inria)

Nombre d'accrédités : 16

Résultats obtenus grâce aux moyens informatiques attribués :

Thèses :

- Détection et caractérisation des moments saillants pour les résumés automatiques par Laura Melissa Sanabria Rosas (2021)
- Détection des effets indésirables des médicaments dans les notes cliniques par Edson Alejandro Florez Suarez (2020)

Stages :

- Correction orthographique et annotations de rapport médicaux par Syrine RADHOUANE (équipe MAASAI - INRIA).
- Prédiction du score d'imputabilité associé à une déclaration d'effet indésirable par Gustavo GAVANZO (MSc IDEX Data Science and AI)
- Segmentation sémantique neuromorphique de scènes en temps réel par Dalia Hareb. Master of Science in Informatics at Grenoble (MoSIG), spécialité Intelligence artificielle et Data science à l'ENSIMAG (Ecole Supérieure d'Informatique et de Mathématiques Appliquées de Grenoble)

Description :

Dans le cadre du Cadre du CPER 2014-2020, les acteurs du campus SophiaTech se sont regroupés afin de poursuivre l'équipement du campus. Une demande pour un budget global de 3.100 K€ a été ainsi validée autour de 4 projets :

- RANDOM :Plateforme de mesure de Rayonnements électromagnétique d'ANTennes et de Diffraction jusqu'aux Ondes Millimétriques, d'objets communicants, de capteurs et de

systèmes radar- aménagement de la chambre anéchoïde dont les aspects bâtiments ont été réalisés dans le cadre du précédent CPER

- PLATON : PLATeforme ouverte hybride pour la cONception et l'expérimentation des réseaux du futur
- Véhicule autonome en déplacement urbain
- Ubiquarium : Espace d'expérimentation pour l'informatique ambiante et les objets connectés

Le projet qui nous concerne ici, est Ubiquarium et regroupe principalement 2 aspects :

- Un ensemble de matériels de nature diverse pour équiper la salle d'expérimentation domotique du même nom,
- L'achat de gros serveurs de calcul (5 serveurs Power EDGE T640) pour le traitement massif de données (big data) en vue notamment de leur exploitation dans le cadre des études sur l'intelligence artificielle.

C'est évidemment, ce dernier volet qui nous intéresse ici. En accord avec les différentes tutelles et pour des raisons spécifiques (climatisation, administration, maintenance...). Ce matériel a été installé dans les locaux Inria comme différents nœuds du cluster NEF.

L'objet cette demande est de permettre aux membres de l'équipe SPARKS de l'IS3 d'avoir accès de manière privilégiée selon la gestion des nœuds au sein de NEF à ce matériel en vue d'y mener les travaux prévus. Nous rappelons que tous ceux qui ont un accès à NEF (via Inria ou via OPAL) peuvent bénéficier de la puissance de calcul offerte par ces différents nœuds.

- Les membres des EPC Inria-IS3 ont déjà accès à NEF mais pas à ces nœuds de calcul de manière prioritaire. Cela se fera via la mise en place d'un groupe d'utilisateur SPARKS au sein de NEF.

Voici la liste des travaux prévus pour cette première tranche :

- Analyse de rapport médicaux (Correction orthographique, Annotation, Extraction de relations)
- Extraction non supervisée de relations entre entités nommées - Application à l'analyse de rapports médicaux

-
- Analyse d'image et interprétation des décisions
 - Correction orthographique et annotations de rapport médicaux
 - Prédiction du score d'imputabilité associé à une déclaration d'effet indésirable
 - Detection and classification of the most important actions in videos, using Deep Learning and multimodal features
 - Réseaux de neurones à spikes appliqués à des données événementielles, selon des apprentissages non supervisés, par renforcement ou bien encore par imitation
 - Réseaux de neurones à spikes appliqués à des données événementielles, selon des apprentissages non supervisés, par renforcement ou bien encore par imitation
 - Aide et conseil aux utilisateurs et métaanalyse pour la recommandation de workflows d'apprentissage automatique (suite du projet ROCKFlows)
 - Building a recommender system to suggest learning resources to user of e-learning platforms, based on information extracted from Knowledge Graphs
 - Étude des méthodes d'analyse des données sur les MOOCs en informatique pour l'amélioration de leur contenu et l'augmentation de leur taux de réussite et conception d'une chaîne de traitement des données en collaboration avec la plateforme FUN (France Université Numérique) pour transformer les données liées au processus d'apprentissage vers un format adapté pour l'analyse et l'ouverture de ses données

SpikingAI (NeuroMod project)

Responsable scientifique : Gilles Scarella

Laboratoire : LJAD

Plateforme OPAL : NEF (Inria)

Nombre d'accrédités : 16

Résultats obtenus grâce aux moyens informatiques attribués :

Validation de nouveaux modèles d'apprentissage d'expériences en neurocognition (publications en cours de rédaction) et de réseaux de neurones artificiels (1 rapport de stage et publications en cours de rédaction).

Description du projet :

Research projects at NeuroMod institute aim at gathering together computer scientists and mathematicians with biologists and cognitivists. On the long term, the SpikingAI project aims at developing new algorithms for understanding the learning dynamics of rodents and humans.

The data set used by these algorithms consists of the bio-cognitive data collected on rodents navigating in a maze and eeg data collected on humans during cognitive tasks. At behavioral level, the actions granularity (paths, turns, choices etc.) used by the rodents or participants to take their decisions is automatically inferred from animal position data [MZ17] or recorded answers. At neuronal level, the functional connectivity between the neurons is analyzed statistically [RMB21] from the neuron spike trains collected in vivo inside the brain of the animals or from the eeg data recorded on humans. Finally, the relationship between behavioral and neuronal data is studied statistically and computationally, In particular, new machine learning algorithms are developed to integrate neuronal data.

We need High Performance Computing (HPC) for:

- Running large scale simulations of the electrical activity in neuronal networks [MMR20].

-
- Running the statistical analysis approach upon: the neuronal spike trains (generated by the large scale simulations, or collected on the animals) or the eeg data and behaviors.
 - Running the machine learning algorithm at both behavioral and neuronal levels.
 - Running new theoretical/artificial neuronal networks aiming, on the long term, to reproduce the learning activity in vivo.

References

[MZ17] Alexandre Muzy and Bernard P. Zeigler. "Activity-based Credit Assignment (ACA) Heuristic for Simulation-based Stochastic Search in a Hierarchical Model-base of Systems". In: IEEE Systems Journal 11.4 (2017), pp. 1916–1927. doi: 10.1109/JSYST.2014.2342534. url: <https://hal.archives-ouvertes.fr/hal-01315156>

[MMR20] Cyrille Mascart, Alexandre Muzy, and Patricia Reynaud-Bouret. "Discrete event simulation of point processes: A computational complexity analysis on sparse graphs". working paper or preprint. Jan.2020. url: <https://hal.archives-ouvertes.fr/hal-02455422>

[RMB21] Patricia Reynaud-Bouret, Alexandre Muzy, and Ingrid Bethus. "Towards a mathematical definition of functional connectivity". In: Comptes Rendus de l'Académie des Sciences - Series I – Mathematics (2021). url: <https://hal.archives-ouvertes.fr/hal-03093516>

MSc IDEX Data Science and AI

Responsable pédagogique : Giovanni Neglia

Plateforme OPAL : NEF (Inria)

Nombre d'accrédités : 26

Description :

Dans le cadre du cours "Optimization for Data Science" du master UCA Data Science and AI, nous permettons aux étudiants d'utiliser le cluster NEF pour implémenter et tester des algorithmes d'entraînement (e.g. avec PyTorch).