

GRANDS CHALLENGES

JOLIOT-CURIE



LA SIMULATION

**Comprendre
le monde**

**Au cœur
de l'innovation**

**Mieux connaître
le corps humain**

Édition juin 2019



GRANDS JOLIOT-CURIE CHALLENGES

Édition juin 2019



Christine Ménaché
Responsable du TGCC



Philippe Lauocat
PDG de GENCI

Édito

Le calcul intensif ne cesse d'élargir ses champs d'actions et ses modes d'utilisation si bien que **les chercheurs, qu'ils soient académiques ou industriels, ont des besoins en constante augmentation, aussi bien en terme de puissance de calcul qu'en terme de traitement de données.** Ainsi l'association du traitement rapide et de l'analyse des données à la simulation numérique a récemment décuplé le champ des possibles dans des domaines comme l'évolution du climat, l'astrophysique, le séquençage du génome et la médecine personnalisée, ou le cycle de vie complet d'un produit industriel...

Dans ce contexte, il est d'autant plus important de maîtriser les technologies matérielles et logicielles qui constituent le cœur d'un centre de calcul à haute performance et de proposer de nouveaux services adaptés à l'évolution de plus en plus rapide des besoins. Les équipes de GENCI et du CEA/TGCC s'y emploient en permanence. **L'accès à de nouveaux outils d'analyse et de visualisation de données, de logiciels issus de l'Intelligence Artificielle (IA) pour la mise en œuvre de réseaux de neurones profonds, ou à des ressources virtualisées ont été ainsi mis en œuvre.** Elles travaillent également au développement d'outils de métrologie pour mieux comprendre l'utilisation des ressources (CPU, mémoire, réseau...), identifier les goulots d'étranglement sur les codes et aider les utilisateurs dans l'optimisation de leurs applications à différents niveaux.

Pour répondre à la croissance des besoins des communautés scientifiques, GENCI a investi dans l'achat d'un nouveau supercalculateur dénommé « Joliot-Curie », Irene pour ses utilisateurs, qui remplace le supercalculateur « Curie » installé au TGCC en 2011. D'une puissance crête de 9.4 Pflop/s, ce supercalculateur a été acquis auprès du constructeur ATOS/Bull à l'issue d'une procédure de dialogue compétitif, menée par une équipe intégrée GENCI-CEA/TGCC. Autour d'une architecture équilibrée (capacité de calcul, capacité mémoire, débits réseau et entrées-sorties), il dispose de deux partitions de calcul complémentaires et des ressources pour l'analyse et la visualisation distante des données. La première partition, généraliste, est composée d'environ 80 000 cœurs de calculs Intel Skylake cadencés à 2.7 GHz, interconnectés par un réseau InfiniBand. La seconde plus novatrice, comprend des processeurs many-coeurs d'Intel (KNL) dotés de mémoire rapide et est équipée du tout nouveau réseau d'interconnexion développé par Atos : le BXI, « Bull eXascale Interconnect ». L'arrivée de ce supercalculateur a été accompagnée par le renouvellement de l'environnement de stockage global du TGCC et un renforcement de sa sécurité.

L'installation et la mise en service du supercalculateur Joliot-Curie ont été réalisées avec le soutien de GENCI, par les équipes du CEA et d'Atos, démontrant leur savoir-faire dans la mise en œuvre, dans les délais prévus de grandes infrastructures de calcul. Une période particulière de 3 mois, dédiée au



démarrage et à la mise au point de la machine, a été proposée et réservée à quelques utilisateurs, acceptant par principe des conditions pas complètement stabilisées et optimisées. Cette organisation simultanée de la période de « rodage » pendant la mise en production du nouveau supercalculateur est également une démonstration de l'implication des équipes du centre et de GENCI pour aider au maximum les scientifiques. Les utilisateurs sélectionnés ont alors eu **l'occasion – unique – de pouvoir accéder à des ressources de calculs pouvant aller jusqu'au supercalculateur tout entier, permettant la réalisation de simulations ou de traitements exceptionnels, appelés « Grands Challenges »**. La présence pendant cette période, des équipes du centre de calcul (spécialistes systèmes et applicatifs) et des experts du constructeur informatique, mobilisés pour résoudre les difficultés éventuelles de démarrage de la machine, est alors un atout supplémentaire : la collaboration étroite qui s'établit entre les équipes de chercheurs et ces spécialistes permet bien souvent d'optimiser la mise au point des logiciels de simulation et de régler les problèmes de mise en production du supercalculateur.

Ainsi sont réalisées des simulations numériques qui poussent aux limites non seulement les capacités de la machine mais aussi les logiciels de simulation eux-mêmes et l'environnement informatique pour en exploiter les résultats. Elles sont ainsi essentielles pour franchir les changements d'échelle, aussi bien applicatifs que scientifiques, rendus possibles par les avancées technologiques dans le domaine du HPC.

Après un appel à candidatures lancé par GENCI auprès des différents Comités Thématiques représentant les disciplines utilisatrices, **15 Grands Challenges ont donc été retenus et ont partagé le supercalculateur entre mai et juillet 2018** pour la partition Skylake et entre juin et septembre 2018 pour la partition KNL. Côté recherche académique, la majorité des domaines utilisateurs était représentée. Il s'agit du climat et de l'environnement, des écoulements non réactifs, des écoulements réactifs ou multiphasiques, de l'astrophysique et géophysique, de la physique des plasmas, de l'informatique et algorithmique, de la biologie, de la dynamique moléculaire et des propriétés des matériaux. Côté recherche ouverte industrielle, deux grands challenges étaient portés par les groupes Safran et Renault.

La diversité et l'intérêt de l'utilisation du calcul intensif pour soutenir la modélisation par simulation numérique a une nouvelle fois été mise en évidence **par les résultats exceptionnels présentés dans ce numéro, parfois mêmes des « premières mondiales », obtenus par ces pionniers du calcul.** Cet apport du calcul à la théorie et à la recherche ouverte, activité de plus en plus stratégique au service de la compétitivité scientifique et économique, est une pièce de plus à verser au dossier de l'intérêt manifesté par 28 États membres pour le déploiement d'un grand programme de machines de classe Exascale, qui devrait alors permettre des percées et des découvertes scientifiques d'encore plus grande ampleur.

A l'issue de cette phase de « Grands Challenges » la machine Joliot-Curie a été mise en production afin de répondre à la fois aux besoins nationaux mais aussi à la tenue de l'engagement français dans l'infrastructure européenne de recherche PRACE, avec dans les deux cas une forte attente des communautés scientifiques.

Fin 2019 GENCI procèdera à un renforcement très conséquent des capacités du supercalculateur Joliot-Curie avec plus d'un doublement de la puissance de calcul prévue.

C'est donc avec un grand plaisir et une grande fierté que **le TGCC et GENCI s'associent pour féliciter tous les acteurs de ces succès.** Forts de l'expérience aujourd'hui acquise, nous les invitons maintenant à préparer la nouvelle étape de dimension européenne qui conduira à la mise en service et à l'utilisation de supercalculateurs de classe pré-exaflopique vers début 2021, puis exaflopique en 2022-2023 dans le cadre de l'initiative européenne EuroHPC mise en place à la fin 2018.



JOLIOT-CURIE

Le nouveau supercalculateur au TGCC CEA pour la recherche ouverte scientifique et industrielle

De 9,4 PFlop/s à 22 PFlop/s en 2020



Atos/BULL SEQUANA system X1000 Une architecture équilibrée avec 2 partitions complémentaires

SKL :

- 1 656 noeuds fins bi-processeurs Intel Skylake 8168 à 2,7 GHz avec 24 coeurs par processeur, soit au total 79 488 coeurs de calcul et une puissance de 6,86 PFlop/s
- 192 Go de mémoire DDR4 / noeud
- Réseau d'interconnexion Infiniband EDR.

KNL :

- 828 noeuds manycore Intel KNL 7250 à 1,4 GHz avec 68 coeurs par processeur, soit au total 56 304 coeurs pour une puissance de 2,52 PFlop/s
- 96 Go de mémoire DDR4 + 16 Go de mémoire MCDRAM /noeud
- Réseau d'interconnexion Atos-BULL BXI.

Accès à 500 GB/s à un système de fichiers parallèle Lustre multi-niveaux

DÈS 2020 Extension des capacités globales à 22 PFlop/s Nouvelles partitions de calcul & post-traitement/IA

Manycore :

- 2292 noeuds de calcul bi processeurs AMD Rome Epyc à 2,5 GHz avec 64 coeurs par processeur soit un total de 293 376 coeurs de calcul et une puissance de 11,75 PFlop/s
 - 256 Go de mémoire DDR4 / noeud,
 - Réseau d'interconnexion Infiniband HDR100.

Exploratoire ARM :

- 60 noeuds de calcul bi processeurs ARM Marvell ThunderX3 de prochaine génération (à partir de 2H 2020)
- 256 Go de mémoire DDR4 / noeud,
- Réseau d'interconnexion Atos-BULL BXI.

Post-traitement / IA :

- 32 noeuds hybrides avec par noeud 2 processeurs Intel CascadeLake 20 coeurs 2.1 GHz et 4 GPU nVIDIA V100, soit un total de 128 GPU pour une puissance de 1,13 PFlop/s.





C215 alias Christian Guémy
Artiste / Pochoiriste / Street artiste

Le mot de C215

Un visage humain à la recherche scientifique

GENCI m'a confié la mission de réaliser deux portraits, d'Irène et Frédéric Joliot-Curie pour le nouveau supercalculateur Joliot-Curie.

Ces deux scientifiques sont pour moi toute la gloire de la Science française. J'ai beaucoup d'admiration pour ce couple de Nobel, mais c'est aussi une admiration d'une manière plus large pour les grands scientifiques français.

J'ai découvert un jour cette phrase magnifique : « Juché sur les épaules des géants, on peut voir beaucoup plus loin ». Les grands instruments de la science, que sont notamment les supercalculateurs et tous ces chercheurs, ne sont-ils pas ces géants qui nous permettent de comprendre le monde ?

Ce qui m'a séduit plus particulièrement dans cette mission, c'est de pouvoir peindre pour un lieu assez inattendu, le Très Grand Centre de Calcul, et, en tant qu'artiste, d'intervenir sur un supercalculateur, un concentré de technologies qui a une dimension scientifique aussi intense. A partir de ces deux portraits et de la peinture de ce couple, je souhaite donner un visage humain à la Recherche scientifique.

Si l'on voit le supercalculateur, on imagine une machine extrêmement froide et surpuissante. Alors qu'on oublie souvent que, derrière, il y a toute la Recherche scientifique et ce sont des êtres faits de chair et de sang, animés par des passions, par des sentiments.

C'est toute ma passion pour les Femmes et Hommes scientifiques que j'ai mise dans ce projet.



Sommaire

CT1

*Monter en résolution
dans la modélisation globale
du climat*

page 8



CT2

*HYPERBOLA : Dynamique
d'une couche limite hypersonique
de gaz dense*

page 18



*Simulation numérique d'un
allumage par pré-chambre
dans un moteur à combustion interne*

page 28



*Passage à l'échelle des méthodes
haute-fidélité pour les turbomachines*

page 38



*Séparer les ondes et les tourbillons
dans un milieu turbulent
en présence de stratification*

page 54



CT4

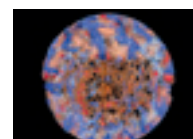
*Simuler les inversions
du champ magnétique terrestre*

page 60



*Les mystères
de la tachocline solaire*

page 66



CT5

*Simulation d'accélération
d'électrons en configuration
double-étage dans le cadre
du laser CILEX-APOLLON*

page 74



*GOSPEL – Gyrokinetic Outer layer
Simulations with Penalised
Extraction at the Limiter*

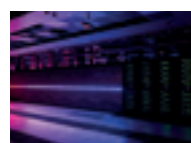
page 86



CT6

*Grand Challenge MPM-1000 :
Adaptation de maillage anisotrope
massivement parallèle appliquée
à des structures complexes*

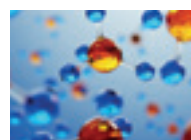
page 96



CT7

*Modelling a vesicle-based
nano-reactor for confined
enzymatic activity*

page 108



Optimisation du code Tinker-HP

page 114



CT8

*Oxygen Evolution Reaction
at the (110)-Co₃O₄ surface
by collective contact matrix
metadynamics DFT-MD simulations*

page 120



CT9

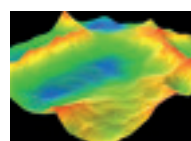
*Recherche de structures dans
des hydrures sous haute pression*

page 124



*Phases exotiques du dioxyde de carbone
fluide à conditions géologiques*

page 132

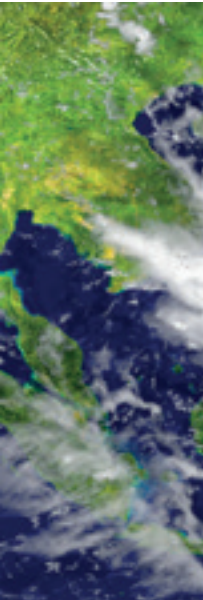


Équipe

Abdoul Khadre TRAORE, (LMD/IPSL)
 Joséfine GHATTAS, (IPSL/CNRS)
 Yann MEURDESIOIF, (LSCE/IPSL/CEA)
 Frédéric HOURDIN, (LMD/IPSL/CNRS)
 Thomas DUBOS, (LMD/IPSL/X)

Ont contribué au projet

Laurent FAIRHEAD, (LMD/IPSL/CNRS)
 Ehouarn MILLOUR, (LMD/IPSL/CNRS)
 Arnaud CAUBEL, (LSCE/IPSL/CEA)
 Olivier BOUCHER, (IPSL/CNRS)
 Marie-Alice FOUJOLS, (IPSL/CNRS)
 Sébastien DENVIL, (IPSL/CNRS)
 Jérôme SERVONNAT, (LSCE/IPSL/CEA)
 Olivier MARTI, (LSCE/IPSL/CEA)
 Sonia LABETOUILLE, (IPSL/CNRS)
 Venkatramani BALAJI, (LSCE/IPSL/CEA)



Monter en résolution dans la modélisation globale du climat

16 millions d'heures

■ 1. Contexte

La modélisation joue un rôle central dans la science climatique car elle permet de réaliser l'équivalent numérique d'expériences de laboratoires, y compris dans des mondes hypothétiques. Une limitation des simulations climatiques globales à l'Institut Pierre-Simon Laplace (IPSL) et ailleurs est leur résolution spatiale typique de 100-200km. Augmenter cette résolution doit permettre de mieux simuler certains aspects du système climatique tels que les oscillations El-Niño, le cycle de l'eau, la couverture neigeuse, et d'acquérir de meilleurs statistiques d'événements extrêmes.

Ce Grand Challenge a permis à l'IPSL de réaliser sa première simulation climatique globale à haute résolution, scientifiquement exploitable. Le modèle à haute résolution produit un climat global comparable, voire sur certains points, meilleur que le modèle à résolution standard, tout en capturant mieux certains extrêmes, notamment les pluies intenses.

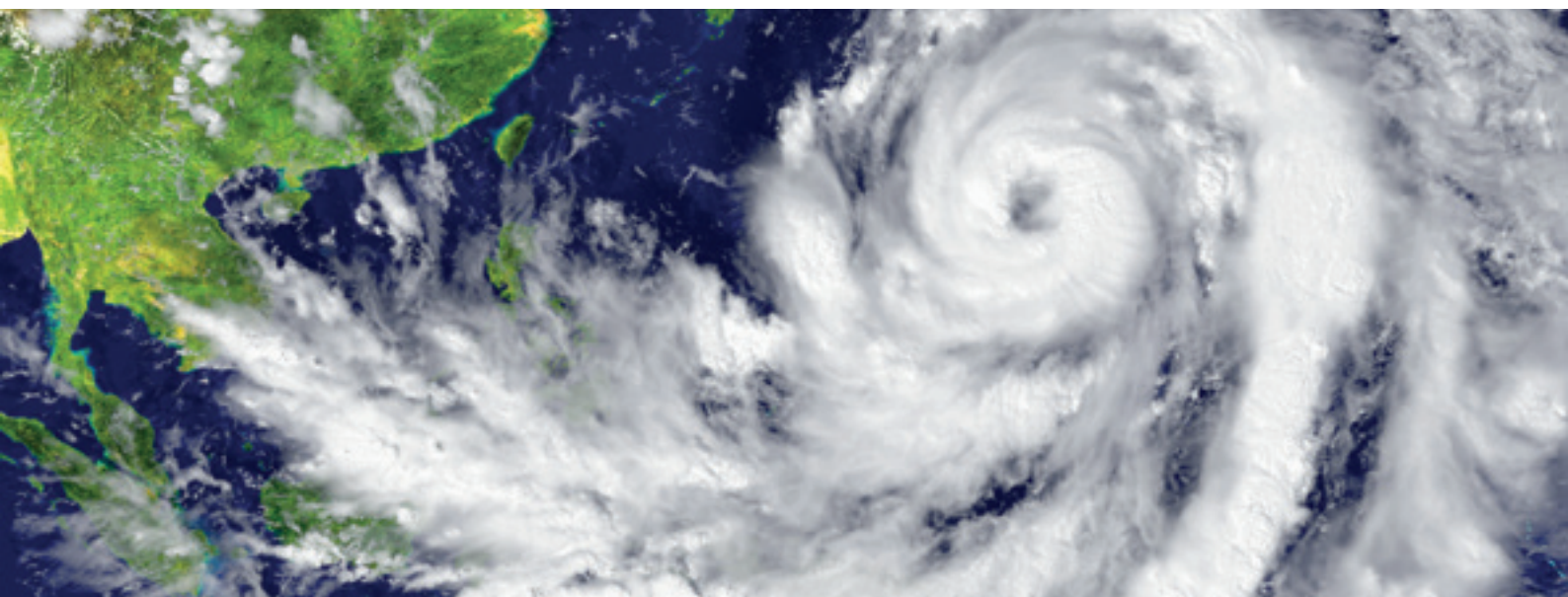
Au-delà de ce jalon hautement symbolique, qui constitue à la fois un exercice de faisabilité de simulations exceptionnelles en termes de volume de calculs et de données, et un travail à visée scientifique, ce Grand Challenge a permis d'avancer significativement vers la mise en production d'un solveur atmosphérique de nouvelle génération permettant un meilleur temps de retour de ces simulations et de simulations à encore plus haute résolution.

1.1 La modélisation globale du climat

La modélisation numérique du climat s'est développée à partir des années 60, à la suite de l'avènement de la prévision numérique du temps. La modélisation joue un rôle central dans la science climatique. Son but n'est pas seulement de reproduire ou d'anticiper des événements

du monde réel : **la modélisation permet de réaliser l'équivalent numérique d'expériences de laboratoires, y compris dans des mondes hypothétiques.** C'est en répétant ces expériences et en variant leurs conditions (composition atmosphérique, géographie, paramètres solaires et orbitaux), qu'on peut identifier des causalités,





apporter des explications aux phénomènes, évaluer les conséquences futures de nos actions et leur incertitude.

Dans son principe cette modélisation est simple : l'océan, l'atmosphère, sont des fluides mis en mouvement par la gravité, la rotation de la planète et des différences de densité, celles-ci résultant du chauffage dû au rayonnement solaire et terrestre (infrarouge). Mais elle est intrinsèquement complexe dans sa mise en oeuvre, d'une part en raison du grand nombre de processus physiques actifs (rayonnement, turbulence, changements de phase de l'eau, nucléation et coalescence de gouttes ...), d'autre part en raison de la gamme considérable d'échelles spatiales et temporelles mises en jeu. Aussi un modèle de climat doit faire de nombreux compromis dans les processus physiques pris en compte et dans la gamme d'échelles couvertes [2]. Ce dernier compromis est fortement contraint par les ressources de calcul mobilisables.

Le nombre d'études potentielles à réaliser sur des sorties de simulations climatiques fait qu'on privilégie la génération de très grands volumes de données pouvant être réinterprétés par la suite non seulement par les personnes ayant réalisé les simulations mais aussi par des équipes

complètement extérieures utilisant les sorties de différents modèles pour leurs recherches. Aussi **les résultats des simulations doivent être produits selon des protocoles standardisés très exigeants**, permettant leur mise à disposition internationale. Ces dernières années, l'IPSL a joué un rôle majeur sur ces aspects en développant le logiciel **XIOS**, partagé à présent pour les deux modèles de climat français (celui de l'IPSL et celui de Météo-France) et en s'investissant dans la distribution des données via les portails ESGF (*Earth System Grid Federation*). Un défi considérable de ce projet est de s'astreindre à ces protocoles pour des volumes de données multipliés par 10 voire 100.

1.2 Limites de la modélisation globale du climat

Pour le modèle climatique de l'IPSL la résolution courante du modèle atmosphérique est d'environ 200 km en moyenne sur le globe. Bien qu'une augmentation de la résolution ne se traduise pas automatiquement par une amélioration de la climatologie produite, **un certain nombre de phénomènes climatiques sont mieux simulés lorsque la résolution spatiale augmente** :



oscillations du phénomène El-Niño, ondes de gravité tropicales, cycle de l'eau et couverture neigeuse, zone de convergence intertropicale.

Par ailleurs l'étude d'événements extrêmes et potentiellement dévastateurs tels que les **cyclones ou les précipitations intenses, de la possible modification de leur fréquence et de leur intensité** en réponse au réchauffement global, nécessite des simulations avec des résolutions spatiales de quelques dizaines de km et suffisamment longues (plusieurs centaines d'années) pour accumuler des statistiques fiables. Compte tenu du coût numérique très élevé, il n'existe à ce jour que peu d'équipes de recherche qui ont eu l'opportunité de mener des simulations climatiques globales très longues et à des résolutions spatiales inférieures à 50 km.

1.3 Modélisation globale du climat à l'IPSL —

Le modèle de climat de l'institut Pierre-Simon Laplace dont les composantes atmosphérique (LMDZ) et surface (ORCHIDEE) sont développées conjointement au Laboratoire de Météorologie Dynamique (LMD) et au Laboratoire des Sciences du Climat et de l'Environnement (LSCE) est l'un des modèles les plus avancés d'Europe pour étudier les phénomènes climatiques. Il est largement utilisé dans les laboratoires de l'IPSL et au-delà pour des études de processus sur les climats passés, actuel et futur. Il contribue également fortement aux projets CMIP d'intercomparaison des modèles globaux servant en partie de base aux rapports du GIEC (Groupe d'Experts Intergouvernemental sur l'Evolution du Climat) sur le changement climatique.

La composante atmosphérique LMDZ du modèle, comme tous les modèles de circulation générale, intègre et résout explicitement les équations de la mécanique des fluides sur la grille du modèle.

D'autres phénomènes ne relèvent pas de la mécanique des fluides (rayonnement) ou sont d'échelle plus petite que celle de la maille (nuages, orages, etc.). Ces phénomènes sont représentés au travers de "paramétrisations physiques". La frontière entre solveur dynamique et paramétrisations peut changer en fonction de la résolution. Ce n'est pas le cas entre 500km et environ 10km, mais en-dessous d'une dizaine de km on entre dans ce qu'on appelle la « zone grise » de la convection orageuse (au sens où elle peut être en partie paramétrée mais aussi en partie résolue explicitement).

1.4 Objectifs scientifiques

Les objectifs scientifiques du projet s'étagent en fonction des paliers envisagés dans la montée en résolution. Ces objectifs ne seront pleinement réalisés qu'à moyen terme. En effet contrairement peut-être à d'autres projets de ce grand défi, la modélisation du climat n'obéit pas à une logique "one-shot" où une simulation exceptionnelle permettrait de répondre à une question brûlante. Les simulations exceptionnelles de ce grand défi ont vocation à préfigurer les simulations standard de demain, si elles tiennent leurs promesses en termes scientifiques, ce qui prendra du temps et nécessitera de répéter ces simulations ou des variantes.

Le premier objectif correspond à une montée en résolution de la configuration standard du modèle jusqu'à une résolution de 50 km. Cette augmentation implique un facteur 16 sur le volume de données produites. Ce défi technique doit être relevé pour démontrer la faisabilité des simulations en "conditions réelles" et assurer leur exploitation scientifique ultérieure. Les simulations permettront premièrement d'évaluer si une montée résolution spatiale suffit à éliminer certains défauts systématiques dans la climatologie moyenne du modèle de l'IPSL. Au-delà des moyennes, cette résolution pourrait améliorer la capacité du modèle à représenter certains événements extrêmes : pluies intenses, tempêtes et peut-être cyclones tropicaux.

Le second objectif correspond à une montée en résolution supplémentaire d'un facteur 2 horizontalement (25 km) avec également éventuellement 30 % de points supplémentaires sur la verticale. **Le supplément de résolution verticale pourrait améliorer la représentation des nuages ; le supplément de résolution horizontale doit améliorer le réalisme des tempêtes**, et permettre à des cyclones tropicaux statistiquement corrects de se former.

Enfin le dernier objectif correspond à une transition vers le nouveau solveur dynamique DYNAMICO interfacé à la physique terrestre LMDZ, à des résolutions équivalentes à des fins de comparaison, et idéalement à des résolutions inaccessibles à LMDZ (12 et 8 km). Ces dernières simulations très coûteuses, peuvent être réalisées sur des durées de quelques années seulement. À résolution équivalente à celle de LMDZ, ces simulations permettront d'évaluer la capacité éventuelle du nouveau cœur dynamique à mieux représenter la circulation atmosphérique grâce à la meilleure isotropie de son maillage et à l'absence de filtrage longitudinal dans les



hautes latitudes. On s'attachera en particulier aux régions polaires où des oscillations numériques apparaissent dans des configurations récentes du modèle longitude-latitude, après avoir amélioré (mais en même temps rendu moins diffusive) la paramétrisation des couches limites très fortement stratifiées en température.

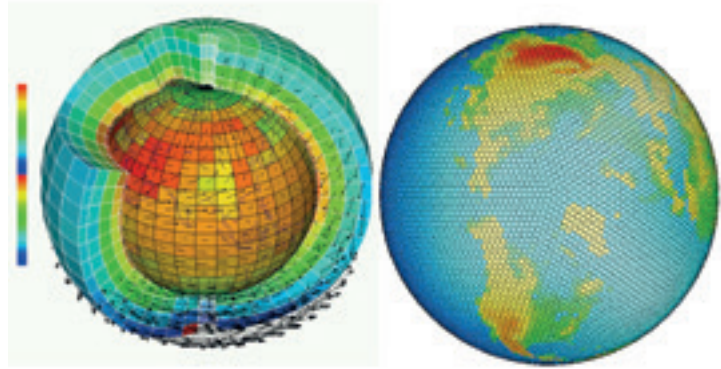
2. Code, moyens de calcul, simulation / méthode numérique

2.1 Scalabilité : limites et progrès

À coût total donné, dépendant fondamentalement des éléments physiques pris en compte par le modèle, un élément-clé de la faisabilité de simulations à haute résolution est le parallélisme massif, qui permet d'obtenir un temps de retour de plusieurs années simulées par jour écoulé (SYPD).

Dans sa version actuelle, le modèle LMDZ utilise pour son solveur dynamique un maillage régulier longitude-latitude qui présente des avantages numériques mais nécessite près des pôles des opérations de filtrage globales dans une direction du maillage, ce qui limite la capacité de parallélisation par découpage de domaines. En exploitant au maximum cette capacité par une parallélisation hybride MPI/OpenMP, il est néanmoins possible depuis peu d'effectuer des simulations à 50 km de résolution sur plusieurs milliers de cœurs à quelques SYPD.

Pour dépasser cette limite de scalabilité (extensibilité des performances), un nouveau solveur hydrodynamique appelé DYNAMICO a été développé conjointement par le LMD et le LSCE. Afin de contourner les limitations intrinsèques des grilles longitude-latitude, DYNAMICO est formulé sur un maillage sphérique icosaédrique-hexagonal. Une fois résolues les questions numériques posées par ce maillage géométriquement irrégulier, il permet une scalabilité bien meilleure par découpage en de nombreux domaines locaux topologiquement réguliers. Le jeu de paramétrisations physiques de LMDZ a été récemment interfacé avec ce nouveau cœur dynamique afin de répondre à des questions inaccessibles avec le maillage longitude-latitude de LMDZ et de viser à long terme des résolutions globales kilométriques. La scalabilité n'est plus



Gauche : grille latitude-longitude du solveur hydrodynamique de LMD-Z

Droite : maillage icosaédrique-hexagonal du solveur DYNAMICO

limitée que par la nécessité de confier à chaque cœur une quantité minimale de travail utile à accomplir. La limite de scalabilité est atteinte lorsque chaque cœur calcule sur quelques dizaines de colonnes atmosphériques, soit quelques milliers de points de la grille tridimensionnelle.

Enfin, avec ou sans DYNAMICO, un point-clé de la scalabilité du modèle est la gestion parallèle et asynchrone des entrées-sorties via XIOS, qui permet en théorie d'éliminer la perte de performance liée aux lectures/écritures sur disque.

2.2 LMDZ-ORCHIDEE

Le code LMDZ se décompose en 2 parties distinctes : (1) le cœur dynamique chargé de résoudre explicitement les équations de Navier-Stokes sur la sphère en milieu tournant et (2) les équations spécifiques de « la physique » qui résolvent les phénomènes et processus physiques au sein d'une même colonne atmosphérique. Dans la partie « physique », aucun échange n'est nécessaire entre les différentes colonnes d'atmosphère, ce qui rend cette partie parfaitement scalable MPI/OpenMP. Cette partie est commune à toutes les configurations mises en oeuvre dans le projet et reprend les choix effectués pour les simulations CMIP6 de référence du modèle de Système Terre IPSL-CM.

Les communications entre processus MPI sont concentrées dans la partie dynamique. Malgré les limites de parallélisation du maillage régulier longitude-latitude dû au filtrage en longitude aux pôles, de grands efforts ont été consentis pour améliorer l'efficacité du code afin de



pouvoir tourner sur des milliers de processeurs à haute résolution, notamment en exploitant le parallélisme OpenMP qui permet de distribuer la dimension verticale sur plusieurs threads.

Le passage d'une résolution de 200km à une résolution 50 km permet de quadrupler le nombre de processus MPI, et le passage à 25 km permettra un doublement supplémentaire. Le passage à une résolution verticale plus élevée permet d'exploiter pleinement des threads OpenMP disponibles.

2.3 DYNAMICO

Le développement récent du nouveau cœur dynamique « DYNAMICO » améliore significativement l'efficacité du parallélisme du modèle [3]. En effet le passage de 200km à 50km permet maintenant de multiplier par 16 le nombre de processus MPI (au lieu de 4 pour LMDZ) et le passage de 50km à 25km permet

un quadruplement supplémentaire (au lieu d'un doublement). À coût total de calcul inchangé, cette augmentation du parallélisme permet d'obtenir un gain en temps de restitution d'un facteur 20 à 40 à haute résolution. Ce gain ne se fait pas au détriment de l'efficacité sur chaque cœur. En effet les données sont organisées de façon semi-structurée en accolant 2 à 2 les triangles principaux de l'icosaèdre. Les tuiles ainsi formées sont

topologiquement cartésiennes ce qui favorise la vectorisation : sur machine vectorielle NEC SX8, DYNAMICO est vectorisé à 99.67%, permettant d'atteindre 60% de la puissance crête. Ceci garantit une exploitation des unités arithmétiques vectorielles des générations de processeurs scalaires actuels et futurs. Sur ces architectures, la principale limitation résulte de la faible intensité arithmétique inhérente au code, qui sollicite intensément l'accès à la mémoire.

2.4 XIOS

Pour les tâches d'entrées/sorties, des processus MPI (les serveurs) sont exclusivement dédiés et les données sont envoyées de manière asynchrone des clients vers les serveurs,

permettant ainsi de recouvrir les transferts et les écritures par du calcul. En utilisant les buffers adéquats, XIOS permet de lisser les pics I/O et d'envoyer un flux de données constant au système de fichier tout au long de la simulation, sans impacter les performances du modèle « client ». En outre, XIOS utilise les I/O parallèles qui permettent, d'une part d'agréger la bande passante I/O et d'autre part d'obtenir des fichiers d'une seule pièce ne nécessitant pas de reconstruction en post-traitement. Le format de sortie est NETCDF4, s'appuyant sur HDF5.

XIOS a démontré sur le code océanique NEMO tournant à 16 000 cœurs de calcul sa capacité à soutenir en continu 3.6 Go/s. Sur une simulation restant un mois en machine, une telle bande passante génère 10 Po de données. XIOS un élément-clé de la chaîne de performance de DYNAMICO-LMDZ-ORCHIDEE, et ce d'autant plus à haute résolution.

2.5 Configurations

Sur maillage latitude-longitude, le nombre de processus MPI est limité à la moitié du nombre de points de grille dans la direction méridienne. Une résolution de 50km correspond à ½ degré de latitude, donc à 180 processus MPI (360 MPI à 25km). Le nombre de threads OpenMP est limité par le nombre de points de grille dans la direction verticale : en-dessous d'une dizaine de couches par thread l'efficacité baisse et il est inutile d'aller en-dessous de 3-4 couches par processeur. Avec 79 niveaux verticaux on peut donc utiliser 8-12 threads efficacement sur Joliot-Curie, dont chaque processeur possède 24 cœurs. Un léger gain en temps de retour est possible en poussant à 24 threads, au prix cependant d'un coût total plus élevé. Les ordres de grandeur sont donc :

- 50 km : 180 MPI x 8-24 OpenMP = 1440 à 4320 cœurs
- 25 km : 360 MPI x 8-24 OpenMP = 2880 à 8640 cœurs

Sur maillage icosaédrique-hexagonal, le nombre de processus MPI est limité par un équilibre à trouver entre calcul utile et opérations 'inutiles' (communication de halo, calculs supplémentaires sur les halos). Un bon compromis est d'environ 400 colonnes d'atmosphère (20x20) entourées d'un halo de 80 colonnes (20%), voire 256 colonnes d'atmosphère (16x16) entourées d'un halo de 64 colonnes (25%). Le nombre de threads OpenMP est limité comme précédemment. Une résolution de 50km correspond à environ 256 000 colonnes

À coût total de calcul inchangé, cette augmentation du parallélisme permet d'obtenir un gain en temps de restitution d'un facteur 20 à 40 à haute résolution.



d'atmosphère organisées en 10 tuiles de 160x160. Les ordres de grandeur sont donc :

- 50 km : 640-1000 MPI x 8-24 OpenMP = 5000 à 24 000 coeurs
- 25 km : 2560-4000 MPI x 8-24 OpenMP = 10000 à 96 000 coeurs
- 12 km : 10240 MPI x 8 OpenMP = 80 000 coeurs

Afin de limiter le nombre de processus MPI, un niveau de parallélisme OpenMP sur l'horizontale peut être exploité, réduisant le nombre de processus MPI et augmentant d'autant le nombre de threads par processus.

2.6 Moyens de calcul et simulation

Durant le Grand Challenge proprement dit (avril-juillet), nos efforts se sont concentrés sur la configuration à 50km du modèle "standard" LMDZ-OR formulé en longitude-latitude, avec l'objectif de réaliser une simulation exploitable scientifiquement, et donc strictement conforme au protocole CMIP6, particulièrement exigeant en termes de volume de données à produire. Nous avons pu ainsi réaliser une simulation dite "historique" répondant aux forçages (températures océaniques, teneur en CO₂ et aérosols, éruptions volcaniques...) effectivement survenus pendant les années 1950 à 2014.

En parallèle, nous avons mené des études de faisabilité sur des configurations icosaédriques à diverses résolutions (jusqu'à 8km) et divers nombres de coeurs (jusqu'à 23 000). Ces simulations étaient très courtes et allégées en termes de sorties et/ou de contenu du modèle.

2.7 Performance

D'un point de vue scientifique, la **performance se mesure de façon absolue en termes d'années simulée par jour (SYPD), quelques SYPD étant le minimum vital** pour la modélisation de quelques décennies du climat moderne. Cette performance doit être mesurée sur une configuration complète pour faire sens. En utilisant 12 threads pour chacun des 180 processus MPI, ainsi que 44 processus MPI pour XIOS, LMDZ-OR à la résolution de 50km a simulé **3,5 années par jour**. C'est un résultat très satisfaisant, mais qui démontre également que LMDZ-OR ne peut pas atteindre une performance exploitable au-delà de cette résolution. En effet à une résolution de 25km, il ne serait possible d'utiliser que 360 processus MPI pour une taille de grille quadruplée et un pas de temps

divisé par deux ; la performance tomberait alors en-dessous de 1 SYPD.

D'un point de vue informatique, il était aussi important de comparer la performance réalisée sur Joliot-Curie à celle obtenue sur son prédécesseur Curie, dans la mesure du possible. Pour les configurations qui ont pu être exécutées à l'identique sur les deux machines, le ratio de performance observé était très variable, de 1 à 2, avec en général **un gain d'environ 30% en faveur de Joliot-Curie**. Ce chiffre peut paraître décevant dans la mesure où l'architecture de calcul de Joliot-Curie permet en principe un doublement des performances. Cependant, il ne constitue pas entièrement une surprise car la structure intrinsèque des calculs réalisés par les **modèles** de climat nécessite à la fois un bon débit arithmétique et une bonne bande passante mémoire, celle-ci ne suivant pas nécessairement l'augmentation de celui-là. Un travail spécifique d'optimisation de nos applications en lien avec les équipes du TGCC, d'ATOS et d'Intel a été lancé en ce sens.

2.8 Environnement technique

L'environnement de compilation, de calcul et de stockage était d'emblée très similaire à celui, familier, du calculateur Curie, ce qui a permis une entrée en matière rapide.

Un certain nombre de problèmes ont été rapidement résolus après signalement (accès aux dépôts externes de codes source ...) tandis que d'autres ont perduré plus longtemps. Citons notamment, ce qui n'était pas en soi inattendu sur une machine à peine ouverte et dans un contexte de Grands Challenges (visant aussi à stabiliser la machine avant sa mise en production), une certaine instabilité dans l'exécution : noeuds parfois défectueux, blocages difficiles à reproduire à grand nombre de coeurs de calculs. Ces problèmes ont progressivement disparu, soit suite à nos échanges avec l'assistance, soit "spontanément", suite probablement à des mises à jour du système tout au long de la période du Grands Challenges.



3. Observations et résultats scientifiques

Les simulations présentées ici suivent le protocole HighResmip de CMIP6 [4]. Il s'agit de simulations de la période actuelle avec le modèle atmosphérique LMDZ-OR. Afin de discuter l'apport de la haute résolution (HR), il est utile de les comparer à des simulations effectuées à basse résolution (LR). Dans le cadre de CMIP6 ont notamment été réalisées des simulations LR selon un protocole similaire à HighResMIP, dit AMIP. Ces deux protocoles diffèrent néanmoins sur quelques points (températures océaniques, glace de mer, végétation, ...). De ce fait, on ne peut pas comparer directement les simulations AMIP-LR et HighResmip-HR, et il faut donc effectuer une simulation complémentaire avec le protocole HighResmip mais la grille LR. Il y aurait beaucoup à dire de ces comparaisons. Ici, nous nous contentons de discuter l'apport de la haute résolution sur la climatologie moyenne d'une part, et sur la répartition des pluies d'autre part.

Sur la figure 2 on compare en moyenne annuelle et pour la décennie 1989-1990 les deux simulations, HighResmip-LR à gauche et HighResmip-HR à droite. La comparaison des deux colonnes montre une amélioration significative de la climatologie du modèle avec la montée en résolution entre les configurations LR et HR. On observe notamment, dans la troposphère (entre la surface et 21 km, axe de droite), une diminution systématique des biais de vent zonal (en haut). Le dipôle entre des biais négatifs dans les moyennes latitudes (verts) et positifs entre les tropiques (jaunes) visible en configuration LR correspondent à un positionnement des jets d'Ouest (contours positifs) trop proche de l'équateur. La diminution de ces biais correspond à un déplacement de ces jets vers les pôles, en meilleur accord avec l'observation.

À ce déplacement des jets est associée une diminution des biais froids (ligne du milieu) et humide (ligne du bas) des moyennes latitudes dans la troposphère. La réduction du biais froid est particulièrement marquée à la tropopause dans les hautes latitudes, où elle est sans doute liée à une diminution de la diffusion numérique qui a tendance à amener trop de vapeur d'eau dans cette région en configuration LR, aboutissant à des biais froids plus forts que -6 K au-dessus de l'Antarctique. Cette interprétation est corroborée par le léger biais positif en humidité dans la version LR dans ces hautes altitudes vers 60S et

60 à 90N qui se transforme en biais positif dans la simulation HR.

On analyse à présent les changements de distribution de précipitations. La figure 3 (voir page 14) compare les sorties des simulations LR et HR à des observations obtenues sur un pluviomètre sur un site d'observation intensive du réseau AMMA-Catch, situé au Mali. On voit sur l'évolution des humidité spécifiques (traits fins) l'alternance entre un hiver très sec, correspondant à l'advection d'air depuis le Sahara par l'Harmattan, et une saison humide (de mai à octobre) au cours de laquelle se réalise l'intégralité des précipitations (barres).

Les deux simulations parviennent à reproduire dans les grandes lignes cette saisonnalité. Elles ont des défauts en commun comme une sous-estimation des maxima de températures (courbes épaisses) en mai, qui fait que la décroissance des températures observée pendant la saison des pluies est moins marquée. **Un point saillant est la meilleure capacité de la simulation HR à reproduire des valeurs élevées de précipitations et à concentrer l'essentiel des précipitations sur la période juin-septembre.**

4. Perspectives

D'ores et déjà, il semble acquis que les simulations à haute résolution présentent une climatologie satisfaisante. Les analyses présentées ici sont très préliminaires et superficielles, et plusieurs mois seront nécessaires pour évaluer en profondeur les améliorations - et les défauts - des simulations à haute résolution par rapport aux simulations à résolution « standard ».

L'essentiel des résultats scientifiques obtenus durant le Grand Challenge l'ont été avec la formulation longitude-latitude de LMDZ. Notre projet initial incluait également la réalisation de simulations avec la formulation icosaédrique, plus scalable mais moins mûre techniquement. Sur ce point, nous en sommes restés à des essais techniques, non exploitables scientifiquement, mais qui confirment la meilleure scalabilité de DYNAMICO par rapport à LMDZ, sa relativement bonne adéquation à l'architecture de calcul de Joliot-Curie, et sa capacité à atteindre des résolutions encore plus élevées dans des conditions permettant une exploitation scientifique.



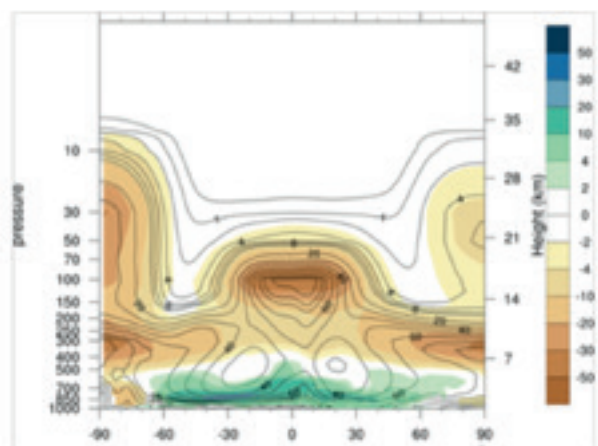
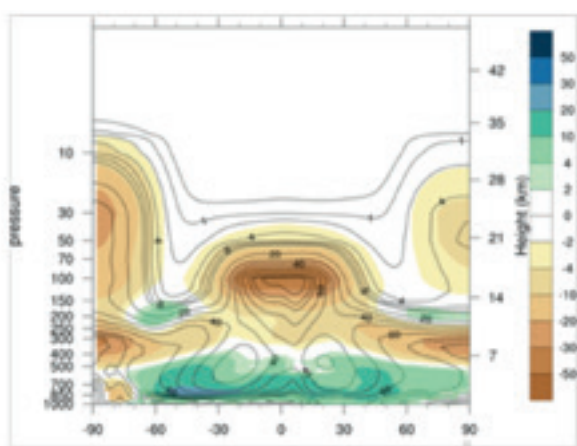
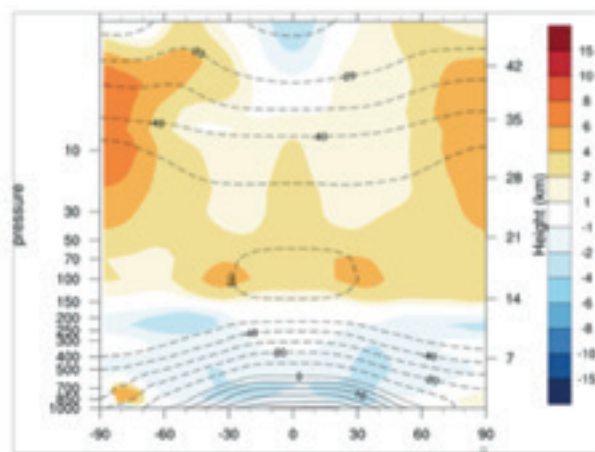
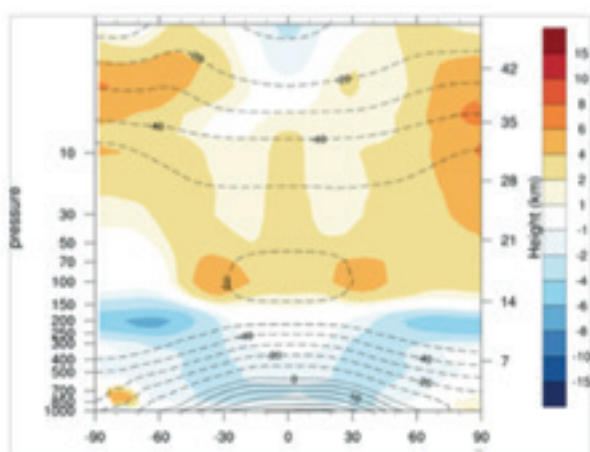
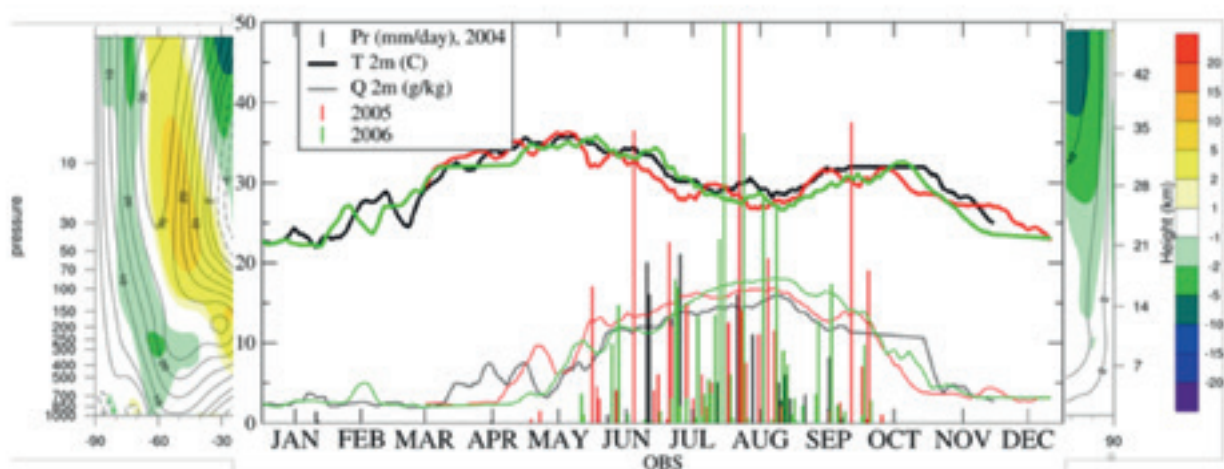


Figure 2 : Moyenne zonale et annuelle de la température (K, en haut), du vent zonal (m/s, au milieu) et de l'humidité relative (% en bas) dans des simulations HighResmip-LR (à gauche) et HighResmip-HR (à droite). Sur chaque graphe, l'axe des ordonnées montre à gauche la pression en hPa et à droite l'altitude correspondante. Les contours correspondent aux champs bruts et les aplats de couleurs à la différences par rapport aux observations (réanalyses ERA Interim).



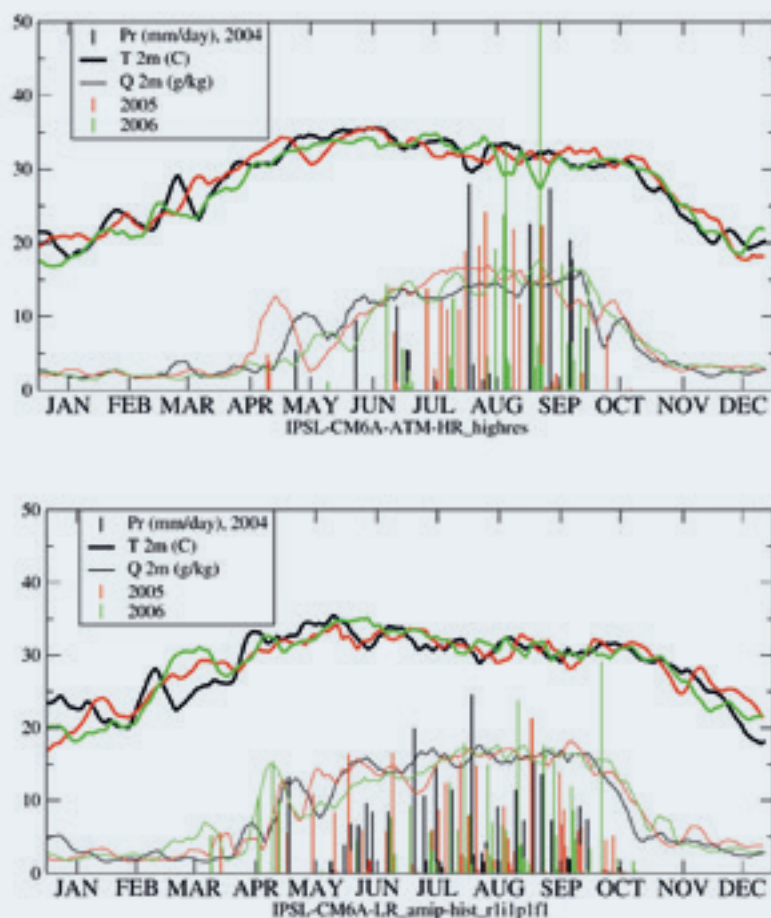


Figure 3 : Evolution sur 3 années consécutives (2004, noir, 2005, rouge, 2006, vert) des températures de l'air à 2m (courbes épaisses), de l'humidité spécifique (courbes fines) et de la précipitations (barres verticales) à la station d'Agoufou au Mali, pour les observations (en haut), la simulation LR (au milieu) et la simulation HR (en bas)

Nous avons poursuivi ce travail après le Grand Challenge, dans le cadre standard de l'ouverture de Joliot-Curie aux projets scientifiques. Des simulations de plusieurs années à basse résolution ont été réalisées à analysées à des fins de validation de DYNAMICO-LMDZ, ainsi que des simulations courtes à 50 km et 25 km de résolution. Ces dernières montrent en particulier l'émergence de tempêtes de moyennes latitudes intenses et de cyclones tropicaux, comme attendu. D'ici à l'été 2019, nous devrions avoir répété le protocole HighResMIP, cette fois-ci avec DYNAMICO et à des résolutions de 50km et 25km. Il nous tarde de pouvoir analyser ces simulations.

■ 5. Conclusions

La simulation du climat à haute résolution, en raison de son coût et de sa lourdeur, a vocation à compléter plutôt qu'à remplacer la simulation à « basse » résolution. Sur certains aspects spécifiques du climat, notamment les événements extrêmes, les simulations à haute résolution sont potentiellement à haute valeur ajoutée, voire incontournables.

Ce Grand Challenge a permis à l'IPSL de réaliser sa première simulation climatique globale à haute résolution, scientifiquement exploitable. Au-delà de ce jalon hautement symbolique, il a permis d'avancer significativement vers la mise en production du modèle DYNAMICO-LMDZ, qui permet un meilleur temps de retour de ces simulations et de simulations à encore plus haute résolution.



■ 6. Remerciements

Nous remercions le projet ANR Convergence et le projet ANR HEAT qui ont préparé et soutenu l'effort nécessaire pour mener à bien ce Grand Challenge. Nous tenons à remercier l'assistance du TGCC pour son écoute et sa réactivité durant le Grand Challenge, et pour les améliorations significatives apportées à l'environnement Joliot-Curie.

■ 7. Bibilographie

[1] IPCC (2013), Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

[2] F. Hourdin et al. *The Art and Science of Climate Model Tuning*. Bulletin of the American Meteorological Society, 98:589-602, March 2017

[3] T. Dubos et al.. *DYNAMICO-1.0, an icosahedral hydrostatic dynamical core designed for consistency and versatility*. Geoscientific Model Development, 8:3131-3150, October 2015.

[4] R. J. Haarsma et al. *High Resolution Model Intercomparison Project (HighResMIP v1.0) for CMIP6* Geosci. Model Dev., 9, 4185-4208, 2016



Équipe Aérodynamique et Aéroacoustique
du laboratoire DynFluid

Paola CINNELLA, Professeur,
Laboratoire DynFluid, ENSAM

Xavier GLOERFELT, Professeur,
Laboratoire DynFluid, ENSAM

Francesco GRASSO, Professeur,
Laboratoire DynFluid, CNAM

Donatella PASSIATORE, Stagiaire,
Laboratoire DynFluid, ENSAM

Luca SCIACOVELLI, Maître de Conférences,
Laboratoire DynFluid, ENSAM



HYPERBOLA

Dynamique d'une couche limite hypersonique de gaz dense

8 millions d'heures

Le projet HYPERBOLA (*Dynamics of transitional and turbulent HYPERsonic BOundary LAyers of dense gases*) vise à étudier la dynamique de couches limites turbulentes hypersoniques (Mach 6) en comparant un gaz parfait (air) et un gaz dense (PP11) représentatif des gaz utilisés dans les cycles de Rankine Organiques. En effet, l'écoulement de couche limite en gaz dense n'a jamais été étudié et, en l'absence d'expériences réelles capables de nous renseigner sur sa dynamique, **le Grand Challenge sur Joliot-Curie nous donne l'opportunité de réaliser pour la première fois une base de données haute-fidélité par simulation numérique directe**. Les effets originaux de gaz dense sont notables pour les régimes hautement compressibles. C'est pourquoi une vitesse hypersonique a été choisie pour notre étude. Nous avons observé plusieurs différences remarquables par rapport au comportement d'un gaz parfait : la thermodynamique complexe du PP11 et notamment sa grande capacité thermique altère profondément les mécanismes de transition et le développement de la couche limite. Ces modifications ont un impact notable sur la génération de la traînée et les transferts thermiques.

1. Contexte

Au cours des dernières années, la communauté scientifique a accordé une attention particulière aux écoulements de gaz compressibles ayant un comportement thermodynamique complexe (par rapport aux gaz parfaits). En effet, de nombreux problèmes en physique et en sciences de l'ingénieur sont caractérisés par des conditions thermodynamiques dans lesquelles

l'approximation de gaz parfait n'est plus valable. Nous nous intéressons plus spécifiquement aux « écoulements de gaz denses », c'est-à-dire des gaz caractérisés par des molécules complexes et opérant à des pressions et des températures du même ordre de grandeur que le point critique liquide / vapeur. Dans des conditions thermodynamiques correspondant à la région dite « de gaz dense », certaines vapeurs organiques présentent une variation non classique de la vitesse du son avec la pression. Les phénomènes





les plus frappants sont attendus pour les fluides dits de Bethe-Zel'dovich-Thompson (BZT) qui présentent, dans leur phase vapeur, une région où la dérivée fondamentale de la dynamique des gaz (une mesure de la concavité des isentropes dans un diagramme de Clapeyron), est négative. Dans cette région, des ondes non classiques peuvent se former dans les régimes d'écoulement supersoniques et transsoniques, comme par exemple des chocs de détente, des ondes mixtes ou des chocs soniques. Le domaine d'application de ces gaz complexes est très vaste. On peut citer les souffleries à grand nombre de Reynolds, les groupes auxiliaires de puissance des avions, les chambres de combustion à haute pression, ou les systèmes de conversion d'énergie. Une application attrayante est représentée par la conversion d'énergie à partir de sources renouvelables ou la récupération de chaleur fatale au moyen de cycles de Rankine organiques (ORC). De tels cycles utilisent des fluides organiques lourds pour la production de l'énergie électrique, notamment à partir de sources solaires, géothermiques ou de biomasse. De même, les ORC peuvent être utilisés pour récupérer la chaleur fatale des moteurs utilisés dans les transports terrestres (camions et trains)

ou navals. Une revue des applications des ORC peut être trouvée dans [Colonna].

Compte tenu des conditions de fonctionnement de ces gaz (hautes pression et température), il est difficile de mener des campagnes expérimentales ; d'autant plus que la moindre intrusion d'un dispositif de mesure dans l'écoulement peut provoquer une décomposition du gaz. En l'absence de mesures expérimentales détaillées et fiables pour des écoulements turbulents de gaz denses, la simulation numérique directe (DNS) constitue un outil puissant pour étudier la dynamique de ces gaz et générer des bases de données haute-fidélité qui permettront de guider le développement de modèles.

Le laboratoire DynFluid étudie les gaz denses depuis plusieurs années, à la fois en termes de compréhension physique de leur dynamique qu'en termes de développement d'outils de simulation et d'optimisation. L'objectif général est d'améliorer à la fois les connaissances théoriques sur ces écoulements et leurs applications. À cette fin, des études numériques DNS d'écoulements turbulents compressibles de gaz denses sont réalisées pour des configurations de complexité croissante : turbulence homogène et isotrope



[Sciacovelli2016,2017b], canal plan turbulent [Sciacovelli2017a]. L'écoulement canonique le plus étudié dans la communauté de mécanique des fluides est sans doute la couche limite, c'est à dire la fine région d'écoulement proche de la paroi où la vitesse décroît de sa valeur externe jusqu'à zéro en raison de la viscosité du fluide. Etudier la couche limite est fondamental pour comprendre les mécanismes de génération de la traînée autour d'un corps se déplaçant dans un fluide. De même, pour un écoulement conduisant la chaleur, une couche limite thermique existe près de la paroi et pilote les transferts de chaleur.

Le projet Grand Challenge HYPERBOLA sur Joliot-Curie nous donne l'opportunité de réaliser pour la première fois une base de données haute-fidélité par simulation numérique directe. Etant donné que les effets originaux de gaz dense sont notables pour les régimes hautement compressibles, une vitesse hypersonique a été choisie pour notre étude. On va donc comparer une couche limite transitionnelle et turbulente se développant sur une plaque plane pour un gaz parfait (air) et un gaz dense. Nous étudions en particulier le PP11 (Perfluoroperhydrophénanthrène, formule chimique $C_{14}F_{24}$), un gaz qui a été souvent utilisé dans la littérature des gaz denses et notamment dans nos précédentes études, car il possède une large région BZT.

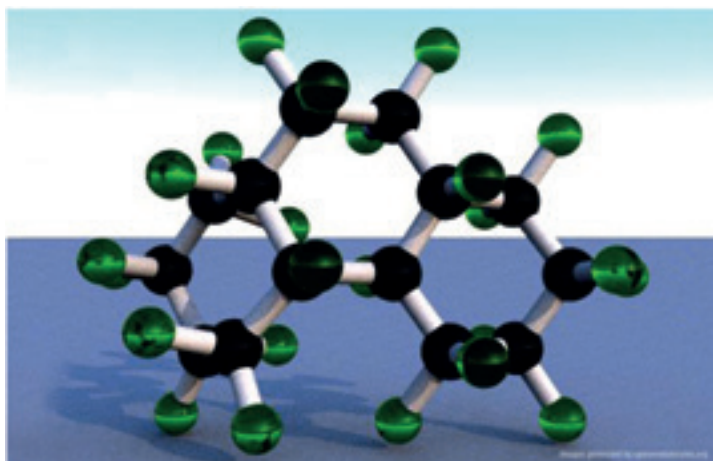


Figure 1 : molécule de PP11 (reproduite du site http://f2chemicals.com/flutec_pp11.html)

L'objectif du Grand Challenge HYPERBOLA est donc d'effectuer deux DNS d'une couche limite turbulente à Mach 6 pour un gaz parfait et dense. Ces simulations vont permettre de :

- Fournir une base de données DNS pour un écoulement hypersonique de couche limite d'un gaz dense pour la première fois dans la littérature;
- Comprendre les différences entre les gaz parfaits et denses (dynamique des structures à grande échelle, rayonnement des ondes de Mach, ...);
- Clarifier les mécanismes de transition laminaire/turbulent dans des conditions hypersoniques [Mack] (rôle des modes acoustiques, différences entre les gaz parfaits et les gaz denses, ...)

2. Code de calcul et simulations sur Joliot-Curie/Skylake

2.1 Code MUSICA2

Ce projet est réalisé avec le code de simulation MUSICA2 (*MUltiscale Solver In Computational Aerodynamics and Aeroacoustics*), développé au laboratoire DynFluid et écrit en Fortran. Ce code résout les équations de Navier-Stokes compressibles sur des grilles structurées [Cohen]. Les équations discrétisées sont intégrées en temps en utilisant un schéma de Runge-Kutta à quatre étapes à faible stockage. Les dérivées spatiales sont discrétisées à l'aide d'un schéma aux différences finies du dixième ordre pour les termes convectifs et du quatrième ordre pour les termes visqueux. À la dernière sous-étape du Runge-Kutta, une dissipation artificielle sélective du neuvième ordre est incorporée afin d'éliminer les oscillations maille à maille. Une viscosité artificielle non linéaire du premier ordre est activée en présence des chocs.

La parallélisation est basée sur la bibliothèque MPI. Afin de limiter les contentions des communications lorsque plus de 1000 cœurs sont utilisés, des communications non bloquantes ont été mises en place (ISEND/IRECV). Les entrées-sorties sont réalisées en MPI-IO pour mieux utiliser les capacités de MPI sur de grands nombres de cœurs et pour pouvoir redémarrer avec un nombre différent de cœurs.



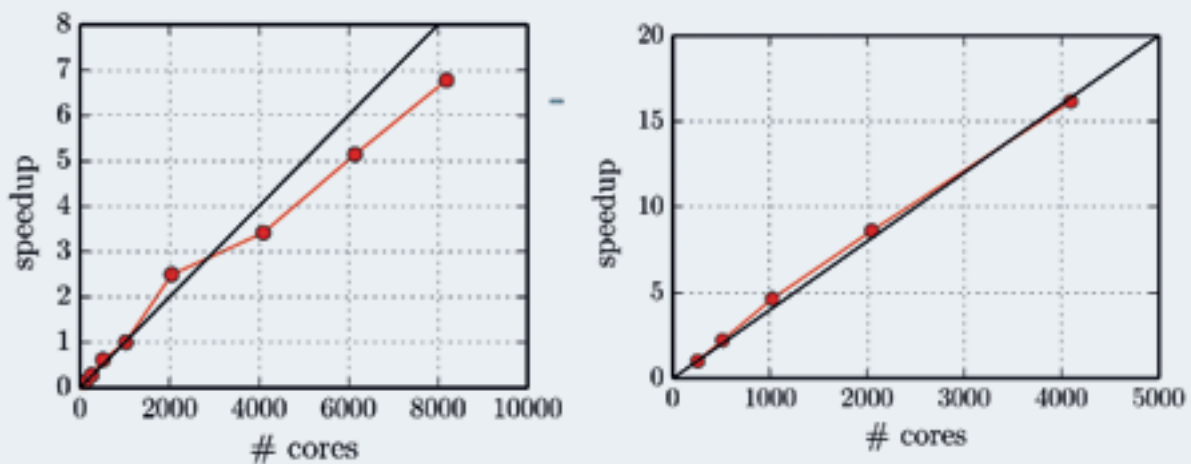


Figure 2 : Scalabilité faible (gauche) et forte (droite) du code MUSICA2. Les tests de scalabilité faible sont basés sur des blocs de 90x50x90 points et le speed-up est normalisé par la valeur correspondant à 1024 cœurs. Pour la scalabilité forte les n cœurs découpent un maillage de 921.6 M de points et le speed-up est normalisé par rapport à 256 cœurs.

La scalabilité parallèle du code a été étudiée sur plusieurs machines du GENCI dans le cadre de projets précédents, entre autres sur Curie et Turing. Des tests de scalabilité faible et forte ont été réalisés sur un domaine similaire à celui considéré pour la présente étude. Comme montré sur la figure 2, les performances de scalabilité sont satisfaisantes, la perte d'efficacité par rapport au comportement idéal étant très limitée, de 8 à 8192 cœurs. Ce comportement a été validé jusqu'à 32768 cœurs sur la machine BlueGene/Q Turing. Nous montrons sur la même figure les performances du solveur en termes de scalabilité forte, c'est-à-dire en maintenant la taille totale du problème constante (environ un milliard de points). Une très bonne scalabilité est obtenue.

2.2 Configurations simulées

Le point de fonctionnement pour le gaz parfait a été choisi afin d'avoir un point de comparaison avec le calcul de Franko et Lele [Franko]. Les température et pression d'arrêt sont $T_t = 533$ K et $P_t = 37.74$ atm. La paroi est isotherme et la température est maintenue égale à la température adiabatique en régime laminaire.

La loi de viscosité choisie est la loi de puissance et la transition laminaire/turbulent est réalisée par une condition d'aspiration/soufflage à la paroi identique à celle utilisée dans [Franko]. La longueur de notre maillage est environ trois fois plus grande que [Franko] afin d'étudier les statistiques pour une couche limite pleinement turbulente. Le maillage cartésien final est de $7700 \times 300 \times 400$ (924 millions points) avec un découpage MPI de $154 \times 6 \times 8$ (7392) cœurs.

Pour les calculs de couche limite avec un gaz dense (ici PP11), nous utilisons le même code avec une équation d'état plus complexe (la loi de Martin et Hou) et les lois de Chung et al. pour les propriétés de transport. Le point de fonctionnement correspond à une température externe de 646.8 K, soit $0.994 T_c$ (température critique) et une masse volumique de $348.4 \text{ kg} / \text{m}^3$, soit un volume massique $1.8 v_c$ (volume massique critique). La paroi est isotherme avec une température de 663.2 K. La valeur de la dérivée fondamentale de la dynamique Γ est de -0.21 à l'extérieur de la couche limite. Nous avons choisi d'exciter le mode supersonique le plus instable par aspiration-soufflage. Ce calcul comporte



14336 x 320 x 280 (1.3 milliards) points avec un découpage MPI de 256 x 8 x 8 (16384) cœurs.

2.3 Portage sur la partition Skylake et déroulement des calculs

L'optimisation de notre code MUSICA2 sur Joliot-Curie/Skylake a été réalisée grâce au soutien de l'équipe du TGCC. Avant l'ouverture du Grand Challenge, de nombreux tests préliminaires sur Curie (Sandybridge, partition thin-nodes) ont permis de valider et de mesurer les performances de la nouvelle version de notre code, prenant en compte des modèles thermodynamiques complexes. Le code a ensuite été testé sur Joliot-Curie (partition Skylake) avec différentes versions/options du compilateur Intel et différentes souches MPI. Le gain en performance par rapport à Curie est limité (au mieux, un gain de performance de 1.2 avec l'option -xCORE-

AVX512). Un travail spécifique, notamment sur la vectorisation, sera nécessaire afin de tirer parti au mieux des processeurs Skylake.

Nous avons d'abord réalisé plusieurs classes de calcul pour le cas « gaz parfait » (maillage de 924 millions de points sur 7392 cœurs). Une classe correspond à 100000 itérations du code et dure 18.3 heures (soit 5.28 s/point/itération). Après une phase transitoire, plusieurs plans de coupe du domaine de

calcul sont enregistrés afin de mener une analyse fine des champs turbulents. Ces entrées-sorties intensives utilisant MPI-IO ont mis au jour un problème important de fuites mémoire avec la version 2.0.2 d'OpenMPI (positionnée par défaut). Beaucoup de tests ont alors été réalisés (options MCA, versions OpenMPI, IntelMPI, ...). Nous avons ainsi noté que la version 1.8.8 résout en partie les fuites mémoire mais s'avère instable avec la commande `mpi_run` et non compatible avec `ccc_mprun`. Un contournement a été mis en place avec l'aide du support TGCC pour pouvoir utiliser `ccc_mprun`, qui a permis de réaliser quelques classes de calcul. Finalement la version 2.0.2 a été corrigée par l'équipe technique du TGCC début juin 2018 et de nouveaux tests ont été réalisés

: le problème de fuites mémoire est corrigé si l'on utilise la commande `ccc_mprun`. Le calcul DNS Mach 6 pour le gaz parfait aura consommé environ 2 Mh Skylake. Le stockage nécessaire est d'environ 15 To.

Une autre difficulté est apparue lorsque nous avons préparé le lancement de la simulation « couche limite de gaz dense », non pas liée à un aspect technique mais à la physique (encore inconnue) de cet écoulement. Nous avons fait de nombreux essais pour mettre au point la condition d'aspiration/soufflage (position / fréquence / amplitude) afin de faire transitionner la couche limite laminaire vers un état turbulent, mais sans succès. Nous avons alors écrit un code de similitude pour la couche limite laminaire avec équation d'état quelconque et généralisé un code de stabilité linéaire locale aux gaz denses afin de déterminer les ondes instables. Le mode, de type Tollmien-Schlichting, que nous utilisons pour provoquer la transition dans le cas « gaz parfait » s'avère toujours stable en « gaz dense ». Nous avons trouvé un mode instable à plus haute fréquence, qui a été identifié dans certains écoulements de couches limites fortement refroidis sous le qualificatif de mode supersonique. Les calculs pour des couches limites avec le gaz dense représentent environ 2 Mh Skylake sur Joliot Curie. Une classe du calcul DNS final (maillage de 1.3 milliards de points sur 16384 cœurs) dure environ 22h pour 100000 itérations (soit 9.9 s/point/itération, presque deux fois plus grand que le gaz parfait). Malheureusement, le temps de mise au point des ondes instables en gaz dense et les problèmes techniques (fuite mémoire lors des I/O avec OpenMPI; crashes et arrêt de la machine; surcharge très importante en juin pour la partition Skylake) n'a pas permis d'utiliser la totalité des heures allouées pour le Grand Challenge (un peu plus de 4 Mh sur 8 Mh). Une allocation GENCI A5 de 14 Mh sur la machine Turing/IDRIS (BlueGene/Q) a permis de poursuivre le calcul (la machine BlueGene est environ 4 fois plus lente que Joliot-Curie/Skylake).

Un travail spécifique, notamment sur la vectorisation, sera nécessaire afin de tirer parti au mieux des processeurs Skylake



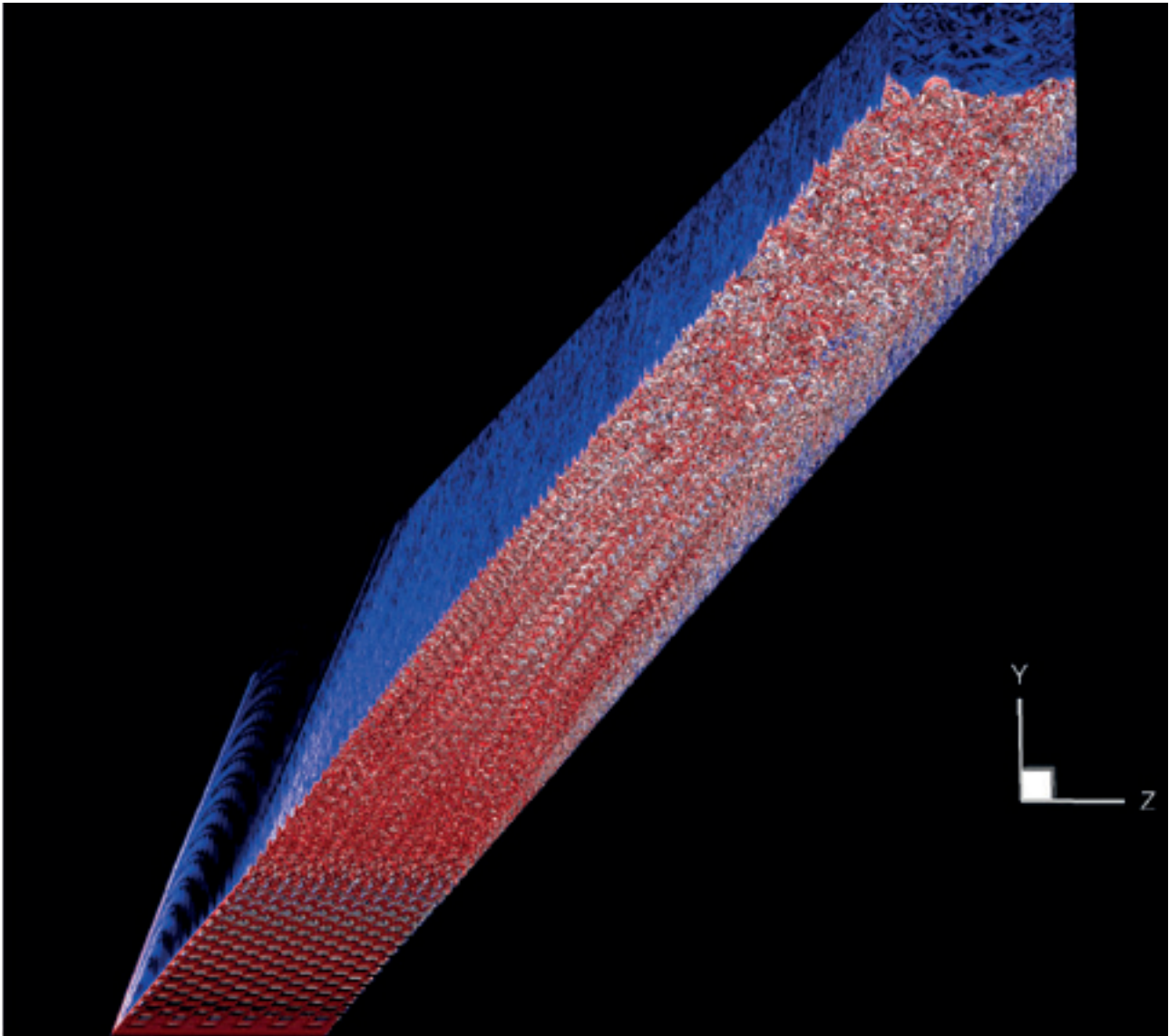


Figure 3 : Vue instantanée d'ensemble d'un critère tourbillonnaire (critère Q coloré par la vitesse longitudinale) qui montre le développement de la couche limite à Mach 6 en gaz parfait de l'état laminaire à l'état pleinement turbulent.

3. Observations et résultats scientifiques

La figure 3 nous donne un aperçu de l'écoulement de couche limite qui se développe sur la plaque pour le cas du gaz parfait. De l'amont (en bas à gauche) vers l'aval (en haut à droite), nous observons : le choc de bord d'attaque au début de la plaque, les perturbations introduites par l'aspiration-soufflage localisée juste après ce choc, le développement des instabilités linéaires (réseau en damier résultant des modes obliques d'angles opposés), l'irruption de la turbulence avec

un caractère transitionnel (stries cohérentes) qui se maintient jusqu'à environ la moitié du domaine de calcul et, enfin, l'état pleinement turbulent de la couche limite. A l'extérieur de la couche limite, les fluctuations (en bleu sur la figure) représentent le rayonnement acoustique de la couche limite, dominé par des ondes de Mach pour cette vitesse élevée. Le premier tiers du domaine de calcul correspond aux mêmes conditions que la DNS de [Franko] et nous avons donc pu comparer directement certaines quantités qui valident parfaitement notre simulation.

Pour le gaz dense, les analyses préliminaires pour l'écoulement laminaire (basées sur la



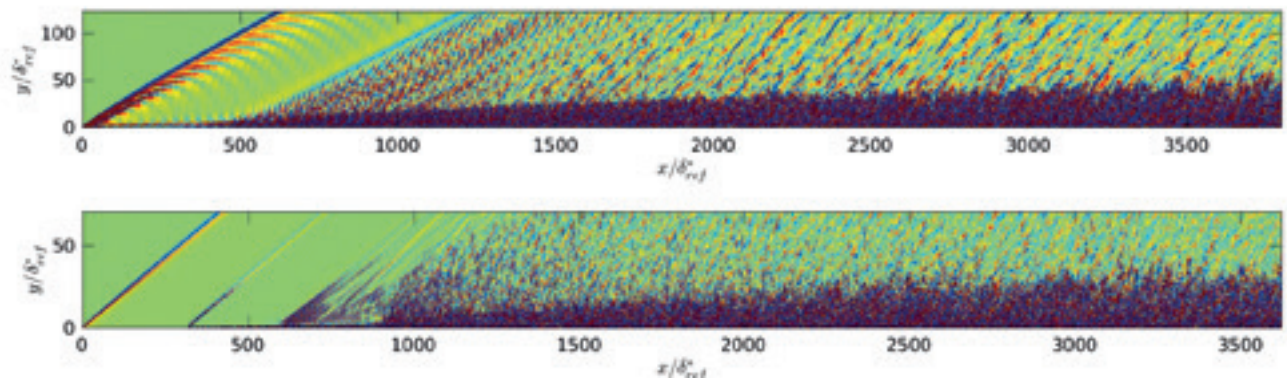


Figure 4 : Vue instantanée dans le plan médian du champ de divergence de la vitesse pour le gaz parfait (en haut) et le gaz dense (en bas) mettant en évidence les effets de compressibilité (l'échelle verticale a été dilatée par un facteur 3 pour la visualisation).

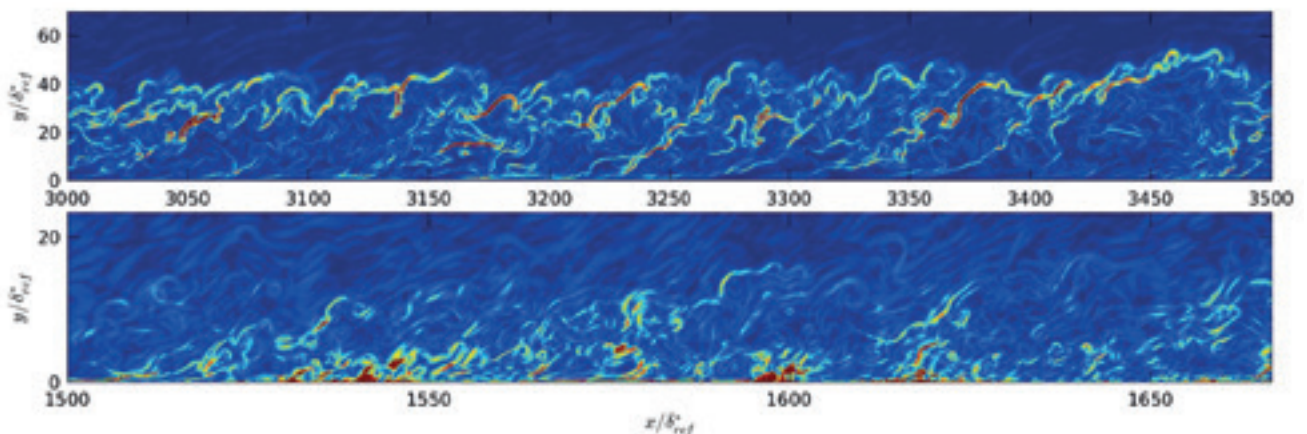


Figure 5 : Comparaison des gradients de masse volumique pour le gaz parfait (en haut) et le gaz dense (en bas).

solution de similitude) montrent qu'en raison du fort découplage des couches limites dynamique et thermique, les profils de vitesse restent proches de ceux du régime incompressible jusqu'à des valeurs très élevées du nombre de Mach externe. Par ailleurs, les variations de température sont limitées à moins de 3%, tandis que les propriétés de transport présentent une variation opposée à celle observée dans les gaz parfaits (par exemple, la viscosité suit une tendance similaire à la masse volumique et diminue à l'approche de la paroi).

Deuxièmement, l'analyse de la stabilité linéaire a permis d'identifier l'influence des effets de gaz dense sur les scénarios de transition modale. La stabilité des couches limites de gaz caractérisés par des chaleurs spécifiques élevées est profondément modifiée par rapport à celle des couches limites de gaz parfaits pour un même nombre de Mach et rappelle le comportement des

couches limites supersoniques et hypersoniques fortement refroidies.

Enfin, les résultats obtenus pour la région d'écoulement pleinement turbulent montrent que les statistiques turbulentes pour le gaz dense restent plus proches du régime incompressible que pour le gaz parfait. Ceci est dû au moindre réchauffement par frottement dans ce gaz doté d'une grande capacité thermique. De ce fait, la couche limite est beaucoup plus fine mais elle atteint des nombres de Reynolds beaucoup plus élevés pour une même longueur de plaque, compte tenu de la forte densité du gaz. Malgré des profils moyens relativement proches du régime incompressible, de forts effets de compressibilité sont observés (voir figure 4). En particulier, les structures turbulentes de la partie externe (supersonique) de la couche limite peuvent induire des *shocklets*, y compris non



classiques (comme les *shocklets* de détente). De façon générale, les structures de détente sont favorisées tandis que celles de compression sont affaiblies et confinées dans une région plus proche de la paroi pour le gaz dense. Cet effet est visible sur la figure 5 qui montre les gradients de masse volumique. Pour le gaz parfait, des *shocklets* intenses sont observés près de la frontière externe de la couche limite. Pour le gaz dense, les forts gradients apparaissent en proche paroi et les *shocklets* sont faibles près de la frontière externe. Globalement, des différences importantes sont observées au niveau du coefficient de frottement pariétal et des transferts thermiques, ce qui peut avoir un impact sur la validité des modèles mis en oeuvre pour le design de systèmes utilisant des gaz denses.

4. Observations et résultats scientifiques

Le Grand Challenge sur Joliot-Curie a été une belle opportunité pour notre équipe d'effectuer une avancée importante dans l'étude de la dynamique turbulente des gaz denses. **La mise à disposition d'une telle puissance de calcul ainsi que l'implication forte de l'équipe de support du TGCC, nous a permis de franchir une nouvelle étape vers des simulations ambitieuses et démonstratives.** Le cas d'étude est un écoulement canonique classique dans la communauté de mécanique des fluides : le développement d'une couche limite sur une plaque de l'état laminaire vers l'état turbulent. La configuration choisie, à savoir l'utilisation d'un gaz non-idéal dit « dense », est moins classique puisqu'aucune campagne expérimentale n'existe à ce jour. Ainsi l'expérience numérique se substitue à l'expérience physique pour nous donner accès à la dynamique transitionnelle et turbulente d'une couche limite de gaz dense. Une vitesse hypersonique (Mach 6) a été retenue pour notre étude afin d'exacerber les différences observées entre un écoulement de gaz parfait et de gaz dense.

Les faits marquants de ces expériences numériques sont :

- L'état laminaire de la couche limite de gaz dense reste plus proche du régime incompressible en raison du découplage entre les couches limites dynamique et thermique. En

particulier, la couche limite épaissit moins vite mais, paradoxalement, elle atteint rapidement des nombres de Reynolds élevés de par sa grande densité ;

- Malgré les grands nombres de Reynolds, la transition modale de la couche limite est plus difficile. Le mode de type Tollmien-Schlichting est stable pour le gaz dense à Mach 6 et seul le mode supersonique peut être amplifié. Ce mode est haute fréquence donc peu efficace mais possède la particularité de rayonner spontanément des ondes acoustiques ;

- A l'état turbulent, les statistiques sont proches du régime incompressible en gaz dense contrairement au gaz parfait. De façon concomitante, la couche limite turbulente de gaz dense développe des structures fortement compressibles, donnant lieu à des *shocklets*. En raison du caractère BZT du fluide choisi, ces derniers peuvent être de nature non classique.

Au niveau technique, le portage du code MUSICA2 sur Joliot-Curie/Skylake, son analyse de performance et l'ensemble des échanges qui ont permis de tester le nouvel environnement de calcul constituent également une avancée significative pour notre équipe.

L'analyse aérodynamique et acoustique détaillée de cette base de données reste à finaliser. Elle permettra à terme de mieux comprendre la dynamique de la turbulence pariétale pour des écoulements de gaz dense, afin par exemple de guider le développement de modèles de turbulence. Ces nouvelles simulations ouvrent ainsi la voie vers de nouvelles expériences numériques visant à décrire et à mieux comprendre des configurations d'intérêt applicatif plus tangible et notamment des écoulements supersoniques dans les turbines ORC.

5. Remerciements

Nous remercions le président du CT2a « fluides non réactifs » pour avoir soutenu notre projet. Nous tenons à remercier vivement l'ensemble de l'équipe du TGCC pour son soutien et son aide dans la résolution des problèmes techniques.



■ 6. Bibilograhie

[Colonna] Colonna, P., Casati, E., Trapp, C., Mathijssen, T., Larjola, J., Turunen-Saaresti, T. & Uusitalo, A. (2015) Organic Rankine cycle power systems: from the concept to current technology, applications, and an outlook to the future. *Journal of Engineering for Gas Turbines and Power* 137 (10), 100801.

[Sciacovelli2016] Sciacovelli, L., Cinnella, P., Content, C. & Grasso, F. (2016) Dense gas effects in inviscid homogeneous isotropic turbulence. *Journal of Fluid Mechanics* 800, 140–179.

[Sciacovelli2017a] Sciacovelli, L., Cinnella, P. & Gloerfelt, X. (2017) Direct numerical simulations of supersonic turbulent channel flows of dense gases. *Journal of Fluid Mechanics* 821, 153–199.

[Sciacovelli2017b] Sciacovelli, L., Cinnella, P. & Grasso, F. (2017) Small-scale dynamics of dense gas compressible homogeneous isotropic turbulence. *Journal of Fluid Mechanics* 825, 515–549.

[Mack] Mack, L.M. (1984) Boundary-layer linear stability theory. Tech. Rep.. California Institute of Technology - Jet Propulsion Laboratory.

[Cohen] Cohen, E. & Gloerfelt, X. (2018) Influence of pressure gradients on wall pressure beneath a turbulent boundary layer. *Journal of Fluid Mechanics* 838, 715–758.

[Franko] Franko, K.J. & Lele, S.K. (2013) Breakdown mechanisms and heat transfer overshoot in hypersonic zero pressure gradient boundary layers. *Journal of Fluid Mechanics* 730, 491–532.

■ Équipe projet



De gauche à droite

Francesco Grasso, Xavier Gloerfelt, Paola Cinnella, Donatella Passiatore et Luca Sciacovelli. Au centre, notre mascotte, Axel (qui a vécu le projet de l'intérieur).





Équipe

Quentin MALÉ (A,B)
 Gabriel STAFFELBACH (A)
 Antony MISDARIIS (A)
 Frédéric RAVET (B)
 Olivier VERMOREL (A)
 Thierry POINSOT (A,C)

A - CERFACS, 42 Avenue Gaspard Coriolis,
 31057 Toulouse Cedex 01
 B - Renault, Centre Technique de Lardy,
 1 Allée Cornuel, 91510 Lardy
 C - IMFT, 2 Allée du Professeur Camille Soula,
 31400 Toulouse

Simulation numérique d'un allumage par pré-chambre dans un moteur à combustion interne

8 millions d'heures

1. Introduction

1.1 Contexte

L'augmentation fulgurante de l'utilisation de l'énergie via la combustion ces trente dernières années contraint les industriels à concevoir des systèmes répondant à de nombreuses exigences en termes de consommation, d'émissions polluantes ou encore de nuisances sonores. Alors que le parc automobile mondial ne cesse d'augmenter, les émissions des véhicules participent à l'augmentation de la pollution de l'air et de la concentration de gaz à effet de serre. Face à ce constat, les motoristes recherchent des innovations qui pourraient permettre d'améliorer le rendement et de réduire les émissions polluantes des moteurs à combustion interne. La combustion pauvre homogène est l'une des voies les plus prometteuses pour cela [6]. Elle bénéficie d'une meilleure efficacité grâce à des pertes par pompage plus faibles, à des pertes de chaleur aux parois moindres et à un rapport de capacités thermiques plus élevé en comparaison avec la combustion stoechiométrique. Par ailleurs, **la combustion pauvre homogène produit de faibles émissions polluantes**. Cependant, une vitesse de combustion plus faible et un allumage plus difficile

empêchent les moteurs à allumage commandé de fonctionner en mélange pauvre homogène [20, 15]. L'allumage par bougie d'un mélange pauvre entraîne une combustion erratique, des ratés d'allumage et des combustions incomplètes. Ainsi, brûler un mélange pauvre homogène représente un défi majeur, nécessitant une rupture technologique afin d'enflammer le mélange et de produire une combustion rapide. Étudié expérimentalement [13, 8, 9, 2, 21], le système d'allumage par pré-chambre semble pouvoir ouvrir la voie à la combustion pauvre dans les moteurs à combustion interne. Ce système réduit les instabilités de combustion en induisant un allumage fiable. De plus, les multiples sources d'allumage et l'injection de turbulence impliquent une combustion rapide de la charge principale éliminant les cycles de combustion incomplets et permettant d'augmenter le taux de compression sans induire de cliquetis, augmentant ainsi à nouveau l'efficacité.





1.2 L'allumage par pré-chambre de combustion

L'allumage par pré-chambre consiste à allumer une charge auxiliaire dans une pré-chambre reliée à la chambre de combustion principale par de multiples conduits. La combustion à l'intérieur de la pré-chambre crée une augmentation de la pression due à la dilatation thermique. La différence de pression entre les chambres crée un flux de masse qui force la flamme et les gaz brûlés à pénétrer dans la chambre principale, enflammant la charge principale par un couplage complexe de mécanique des fluides, d'effets thermiques et de cinétique chimique.

1.3 Objectifs

La recherche et le développement d'une telle technologie impliquent une bonne compréhension des phénomènes physiques et de la structure des flammes turbulentes générées. Les simulations de type Large Eddy Simulation (LES) ont le potentiel de prédiction adéquat pour capter les phénomènes instationnaires primordiaux sur ce type d'application. Aujourd'hui cependant, la physique de ce type de technologie a été peu étudiée à l'aide de LES sur des configurations

réalistes de par la complexité et le coût de tels calculs. Le travail réalisé à l'aide de l'allocation d'heures de calcul «Grand Challenge» se propose de répondre à ce manque de données. La LES est ici utilisée comme outil afin d'étudier le comportement d'un système d'allumage par pré-chambre sur un moteur et un point de fonctionnement du Groupe Renault.

1.4 Défi technique

Le calcul d'une telle configuration nécessite une résolution spatiale très fine. En effet, la combustion se déroulant à haute pression et à haute température, les épaisseurs de flamme sont de l'ordre de la dizaine de micromètres. Par ailleurs, les conduits sont d'un diamètre de l'ordre du millimètre ce qui implique de très petites tailles de cellules en complément d'une loi de paroi afin de résoudre correctement les pertes de charge. Les hautes vitesses atteintes dans les conduits combinées aux faibles tailles de cellules impliquent un pas de temps très faible, de l'ordre de la nanoseconde. La cinétique chimique nécessaire à la bonne description du mécanisme d'allumage par gaz brûlés implique le transport d'une vingtaine d'espèces et la résolution de plus



de 150 réactions. Cette combinaison de difficultés nécessite de nombreuses heures de calcul et l'utilisation d'un code massivement parallèle. En outre, la déformation de la géométrie dû au mouvement du piston implique l'utilisation de méthodes numériques spécifiques.

2. Configuration numérique

Le solveur utilisé dans ce travail est un code explicite massivement parallèle qui résout les équations de Navier-Stokes compressibles multi-espèces appelé AVBP [22, 10]. Pour résoudre l'ensemble des équations LES sur maillage non structuré, un schéma numérique explicite

en deux étapes Taylor-Galerkin, est utilisé [4]. Une description complète des équations de la LES peut être trouvée dans [19]. Le modèle de turbulence de sous-maillages WALE [18] est utilisé pour modéliser les échelles turbulentes qui ne sont pas résolues. L'interaction chimie-turbulence est modélisée grâce à la formulation dynamique du modèle de flamme épaissie. Le comportement de la couche limite aux parois est modélisé avec une loi linéaire ou logarithmique en fonction de la distance à la paroi. L'allumage par bougie est modélisé par le modèle de dépôt d'énergie décrit par [14]. La description de la chimie est réalisée à l'aide d'une chimie analytiquement réduite (ARC de l'anglais Analytically Reduced Chemistry) en raison de l'importance de la chimie dans le processus d'allumage par gaz brûlés [7, 1, 23].

2.1. Caractéristiques du maillage

Afin d'optimiser le coût de calcul associé à la taille du maillage, trois configurations sont utilisées : AERO, COMB^{pc} et COMB^{mc}. Les tailles caractéristiques associées aux zones du maillage présentées sont présentées au sein de la Figure 1. La configuration AERO est utilisée de la fermeture des soupapes d'admission jusqu'à 4 degrés vilebrequin avant l'allumage par bougie de la pré-chambre ; COMB^{mc} est utilisée à partir de 4 degrés vilebrequin avant l'allumage par bougie jusqu'à l'instant où les gaz brûlés évacuent la pré-chambre ; COMB^{mc} est utilisée dès l'instant où les gaz brûlés évacuent la pré-chambre jusqu'à la fin de la combustion.

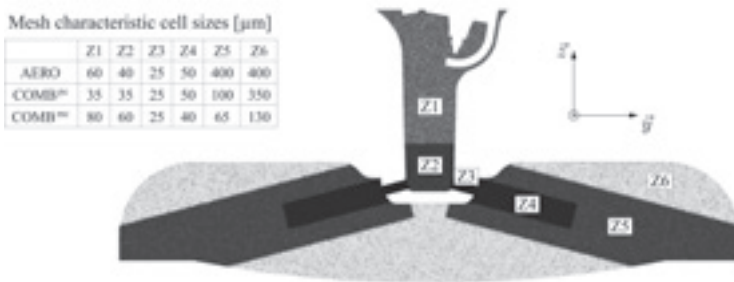


Figure 1 : Zones du maillage.

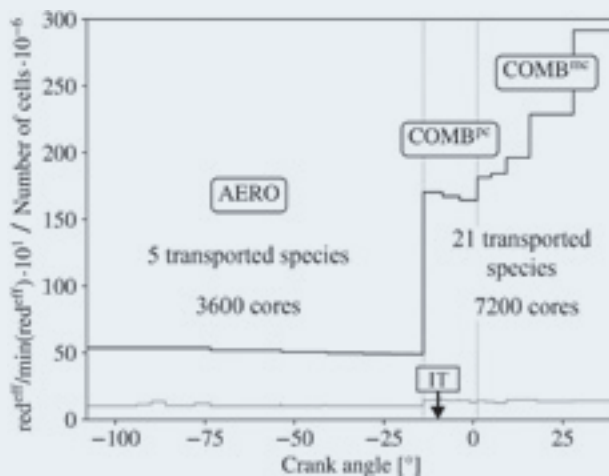


Figure 2 : Evolution du nombre de cellules et de l'efficacité réduite.

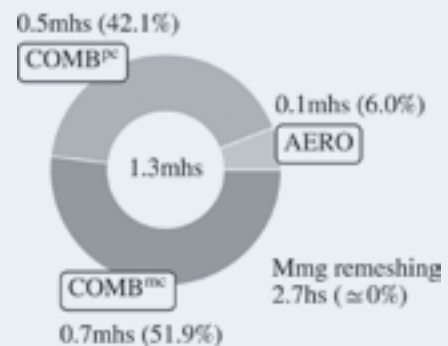


Figure 3 : Temps de calcul divisé en 3 parties : AERO, COMB^{pc} et COMB^{mc}.

2.2. Maillage mobile et remaillage automatique

La résolution des équations de Navier-Stokes sur maillage mobile est obtenue en utilisant une méthode de type Arbitrary Eulerian Lagrangian (ALE) [12]. En raison de la déformation du maillage, les cellules sont de plus en plus déformées au fil du temps jusqu'à ce que les erreurs numériques deviennent trop importantes pour être négligées. Dans ce cas, la méthode ALE peut être utilisée avec des techniques de remaillage et d'interpolation pour changer le maillage au cours de la simulation [17]. La méthode développée au cours de ce « Grand Challenge » utilise la librairie Mmg [5] afin de remailler le domaine de façon automatique. Lorsque qu'une cellule dans le domaine atteint un rapport d'aspect supérieur à 2 [16], la bibliothèque Mmg est appelée afin de remailler automatiquement le domaine de calcul. La solution courante est interpolée sur le nouveau maillage avant de continuer la simulation. Le surcoût associé au remaillage est négligeable en comparaison des coûts de simulation (Figure 3).

2.3. Temps et efficacité du calcul

La Figure 3 affiche le coût de calcul pour les 3 configurations de calcul. Afin d'optimiser les coûts de calcul associés aux équations résolues, la configuration AERO est effectuée en utilisant seulement 5 espèces : C_3H_8 , O_2 , N_2 , CO_2 et H_2O , sans résoudre les réactions chimiques. Grâce à cette stratégie combinée à un nombre relativement faible de cellules, le calcul AERO ne représente que 6 % du coût total de calcul, malgré la durée élevée du calcul en degrés vilebrequin. La Figure 2 présente l'efficacité réduite, définie par

$$red^{eff} = \frac{n_{CPU} \cdot t_{CPU}}{n_{it} \cdot n_{cell} \cdot n_{eq}}$$

où n_{CPU} est le nombre de processus de calcul utilisés, t_{CPU} est le temps du calcul, n_{it} est le nombre d'itérations, n_{cell} est le nombre de cellules et n_{eq} est le nombre d'équations résolues. Cet indice donne un aperçu de l'efficacité du calcul et doit varier le moins possible suite à l'augmentation de n_{cell} et n_{eq} . La faible variation de red^{eff} lors du calcul montre le bon comportement du solveur numérique suite au changement de n_{cells} et de n_{eq} . La légère augmentation de red^{eff} entre les configurations AERO et COMB est due au calcul des réactions chimiques.

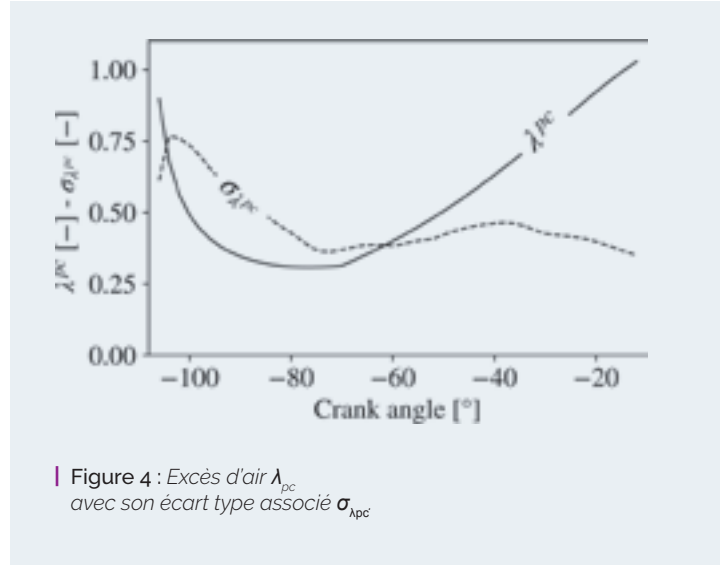


Figure 4 : Excès d'air λ_{pc} avec son écart type associé $\sigma_{\lambda_{pc}}$

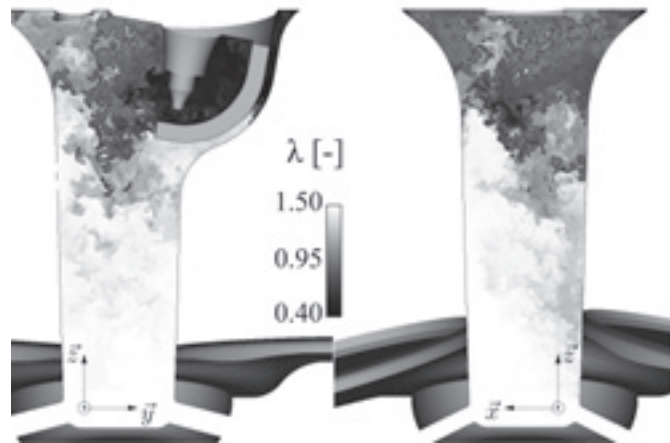


Figure 5 : Coupe médiane de la pré-chambre à 10.5 degrés vilebrequin colorée par l'excès d'air.

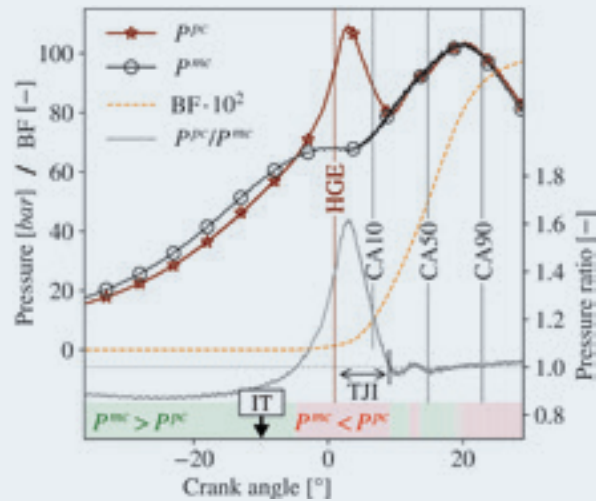
3. Point de fonctionnement

La LES est initialisée à partir d'une solution issue d'une simulation préliminaire de type RANS (Reynolds Average Navier Stokes) à la fermeture des soupapes d'admission. Elle se termine à la fin de la combustion de la chambre principale. Les parois sont maintenues à une température constante de 420 K. Le propane est utilisé comme carburant.

La pression, la température et l'excès d'air λ à l'admission sont de respectivement 2.6 bars, 373 K, 1.5. L'injection gazeuse de propane dans la pré-



Figure 6 : Pression dans la pré-chambre et dans la chambre principale, rapport de pression entre chambre principale et pré-chambre et fraction brûlée en fonction du degré vilebrequin.



chambre commence immédiatement après la fermeture des soupapes d'admission. L'allumage dans la pré-chambre est commandé à -10 degrés vilebrequin.

4. Remplissage de la pré-chambre

La pré-chambre est remplie à la fois par des gaz frais provenant de la chambre principale à un excès d'air égal à $\lambda = 1.5$ et par du carburant gazeux provenant d'une injection effectuée directement dans la pré-chambre. Pendant la phase de compression, les gaz frais se propagent à travers les conduits en raison de la différence de pression entre la pré-chambre et la chambre principale due à la compression par le piston. L'injection de carburant depuis la fermeture des soupapes d'admission jusqu'à un angle vilebrequin de 70.0 permet d'obtenir un mélange stœchiométrique à l'intérieur de la pré-chambre au moment de l'allumage par bougie.

L'excès d'air dans la pré-chambre λ_{pc} diminue significativement lors de l'injection de carburant, puis augmente en raison de la dilution par les gaz frais provenant de la chambre principale pour atteindre $\lambda_{pc} = 1$ au moment de l'allumage par bougie de la pré-chambre (Figure 4). L'écart-type de l'excès d'air dans la pré-chambre $\sigma_{\lambda_{pc}}$ présente des valeurs élevées significatives, ce qui signifie

une forte ségrégation du mélange. La Figure 5 illustre cette ségrégation avec un mélange riche situé en haut de la pré-chambre, confiné par le flux de gaz frais provenant des conduits liant pré-chambre et chambre principale, situés en bas de la pré-chambre. Ces résultats montrent que le mélange peut être considérablement amélioré afin de produire une charge plus homogène au sein de la pré-chambre.

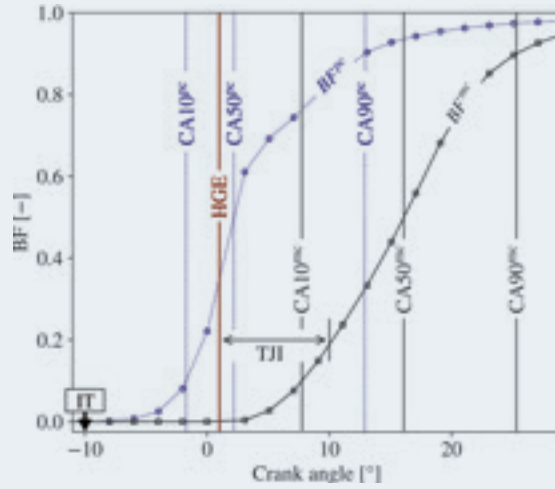
5. Aérothermochimie

5.1. Comportement général

L'allumage par bougie de la pré-chambre se produit à un angle vilebrequin de 10, 0. La Figure 6 montre l'évolution de la pression à l'intérieur de la pré-chambre et de la chambre principale, ainsi que le rapport de pression entre les chambres et la fraction brûlée globale, définie comme la quantité de carburant brûlée. CAx représente l'angle de vilebrequin auquel x % de carburant a été brûlé. Les fractions brûlées à l'intérieur de la pré-chambre et de la chambre principale sont indiquées sur la Figure 7. Suite à l'allumage de la pré-chambre (IT de l'anglais « Ignition Timing »), après un délai dû à la cinétique chimique et à la croissance du noyau de flamme, la pression augmente rapidement en raison de la combustion dans le petit volume semi-confiné de la pré-chambre, ce qui a deux



Figure 7 : Fraction brûlée en fonction de l'angle vilebrequin.



conséquences. Premièrement, la vitesse de consommation augmente avec la compression des gaz frais. Deuxièmement, les gaz frais sont expulsés de la pré-chambre par les conduits. Ces deux phénomènes produisent une augmentation rapide de la fraction brûlée à l'intérieur de la pré-chambre jusqu'à environ CA50°. Par la suite, la combustion dans la pré-chambre est ralentie par les pertes de chaleur aux parois ainsi que par la combustion des poches de mélange riche dues à la ségrégation du mélange. Le rapport de pression entre les chambres atteint une valeur maximale de 1.61, ce qui produit de vigoureux jets turbulents. La première éjection de gaz brûlés, indiquée par HGE (de l'anglais « Hot Gases Ejection »), se produit à un angle vilebrequin de 1.0 soit 11.0 après l'allumage par bougie de la pré-chambre. Les jets turbulents de gaz brûlés enflamment la chambre principale pendant la phase TJI (de l'anglais « Turbulent Jet Ignition »).

5.2. Propriétés des jets

La Figure 8 présente l'évolution de grandeurs caractéristiques intégrées à travers une surface normale aux conduits, située à la sortie des conduits (côté chambre principale).

Toute moyenne surfacique d'une variable ξ s'écrit

$$\langle \xi \rangle = \frac{1}{S} \iint_S \xi dS \quad (2)$$

où S est la surface d'intégration. Le débit d'enthalpie totale est défini comme

$$\dot{H}_t = \iint_S \rho h_t \vec{u} \cdot \vec{n} dS \quad (3)$$

où ρ est la densité, h_t est l'enthalpie totale par unité de masse, u est la vitesse de l'écoulement, n est la normale à la surface S . Le nombre de Reynolds Re est calculé en utilisant la valeur moyenne de la densité ρ , la vitesse normale U_n et la température T pour le calcul de la viscosité. Le nombre de Reynolds atteint des valeurs allant jusqu'à

$4 \cdot 10^5$ lors de l'éjection des gaz frais et chute rapidement lors de l'éjection des gaz brûlés en raison de l'augmentation de la viscosité avec la température. Cependant, l'ordre de grandeur est toujours de 10^5 , ce qui classe les jets comme turbulents. Bien que le rapport de pression entre les chambres atteigne une valeur élevée (Figure 6), il n'atteint pas le rapport de pression critique d'amorçage du conduit estimé à 1, 85 avec un rapport de capacité calorifique constant $\gamma = 1, 33$. Le nombre de Mach moyen normal

$\langle M_n \rangle$ atteint des valeurs allant jusqu'à 0,75, alors que la vitesse normale moyenne dépasse 600 m/s. L'allumage par gaz chauds commence lorsque les premiers produits de combustion traversent un des conduits et se termine lorsque le rapport de pression entre les chambres divisées devient proche de l'unité (débit massique proche de zéro). La durée de



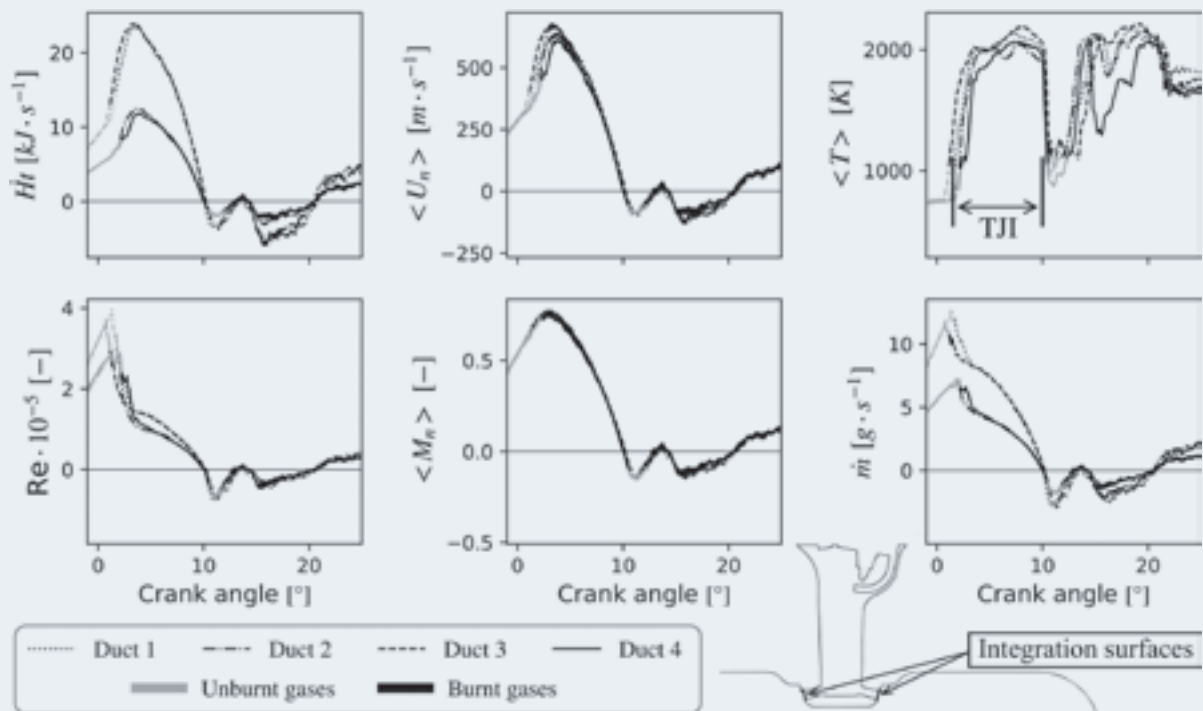


Figure 8 : Évolution du débit d'enthalpie total, de la vitesse normale moyenne, de la température moyenne, du nombre de Reynolds, du nombre de Mach et du débit massique à travers les conduits.

l'allumage par jet de gaz chaud (TJI) est d'environ 9 degrés vilebrequin (Figure 8).

5.3 Allumage de la chambre principale —

Avant l'éjection des gaz brûlés, les jets turbulents de gaz frais agissent comme une source de turbulence pour la chambre principale. Par la suite, les jets turbulents de gaz brûlés agissent également comme source de chaleur, et source de radicaux de combustion. Une fois la chambre principale allumée, les jets assistent la combustion jusqu'à l'équilibre de pression entre les chambres.

La Figure 9 montre la topologie de la flamme lors de l'allumage de la chambre principale (3 degrés vilebrequin) et lorsqu'une flamme réussit à se développer dans la chambre principale (6 degrés vilebrequin).

— L'allumage de la chambre principale se fait dans un régime de combustion dispersée [23]. En raison de la forte intensité de la turbulence, le mélange est rapide, les gaz brûlés sont rapidement dilués par le mélange

frais environnant et le taux de réaction global est principalement limité par la chimie.

— Plus tard, une flamme réussit à se développer et à former un front de flamme, plissé par la turbulence. Il existe encore des zones de mélange intense réactives. Les jets soutiennent encore la combustion dans la chambre principale en fournissant des gaz brûlés et une source de turbulence jusqu'à ce que la combustion dans la chambre principale rétablisse un équilibre de pression entre les chambres.

La Figure 10 montre plusieurs quantités intégrées à travers une surface normale aux conduits, situées à la sortie des conduits (côté chambre principale). La moyenne surfacique est calculée via l'Equation 2. La fraction massique des produits intermédiaires de combustion est définie comme :

$$Y_I = 1 - (Y_{C_3H_8} + Y_{O_2} + Y_{CO_2} + Y_{H_2O} + Y_{N_2}) \quad (4)$$

L'éjection des gaz brûlés peut être divisée en trois phases annotées sur la Figure 10 :

— 1) des produits intermédiaires de combustion sont expulsés à cause de la convection de



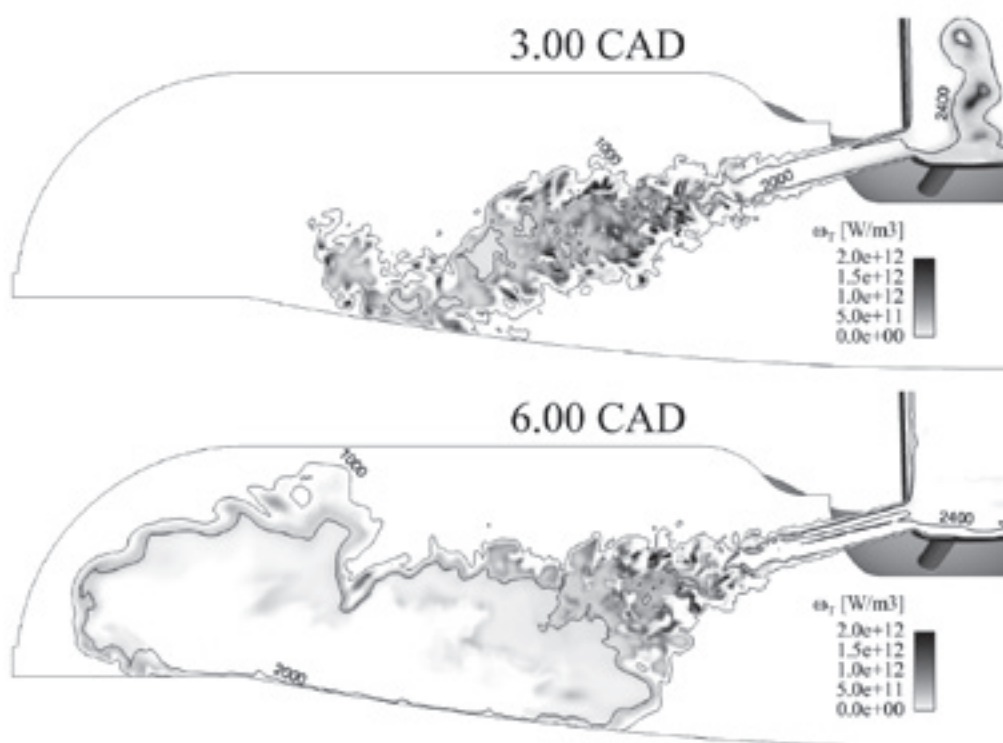


Figure 9 : Coupe plane normale à l'axe x colorée par le taux de dégagement de chaleur avec trois iso-niveaux de température.

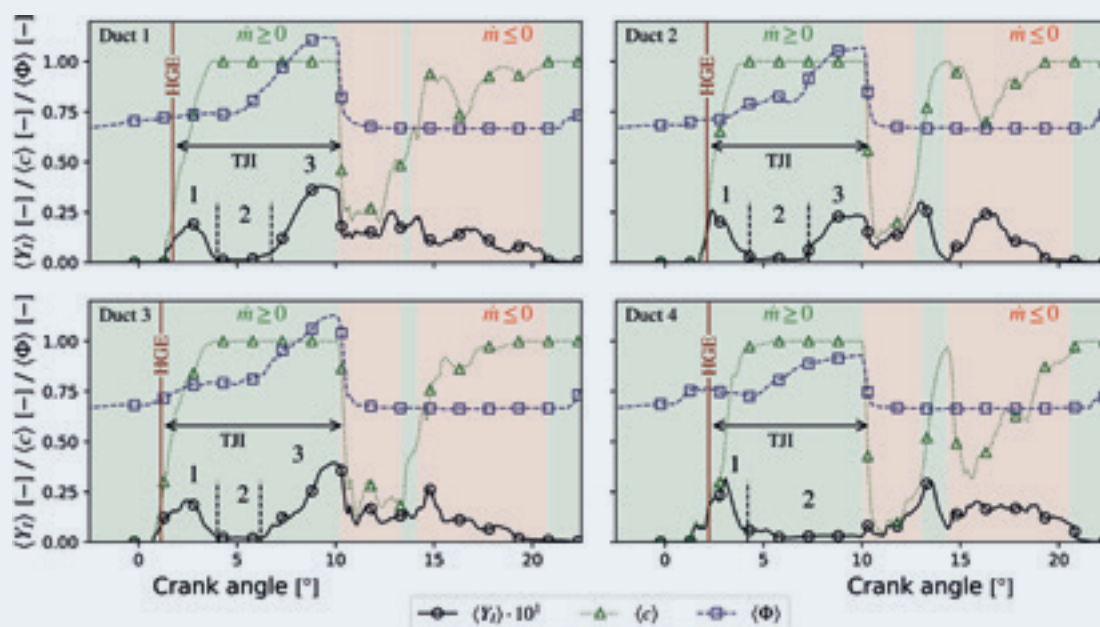


Figure 10 : Moyenne surfacique des fractions massiques, de la variable de progrès et de la richesse à travers les conduits.



fronts de flamme à travers les conduits entre 1 et 4 degrés vilebrequin où $0.1 < c < 0.9$.

— 2) des gaz brûlés pauvres produits par une combustion complète sont éjectés entre environ 4 et 7 degrés vilebrequin selon le conduit.

— 3) le mélange brûlé riche préalablement emprisonné dans la partie supérieure de la pré-chambre atteint le fond de la pré-chambre et est évacué par les conduits. Lorsque la pression dans la chambre principale devient supérieure à la pression dans la pré-chambre, un retour d'écoulement se produit, l'allumage par jet (TJI) est considéré comme terminé.

■ 6. Conclusion

L'allocation d'heures de calcul «Grands Challenges» a permis de réaliser à la fois un défi technique et scientifique. En effet, un calcul de type LES en configuration industrielle comprenant une grande discrétisation spatiale, de nombreuses espèces transportées, un maillage mobile et une stratégie de remaillage automatique a pu être réalisée en parallèle sur plus de 7 000 processus de calcul. Scientifiquement, la base de données générée est très riche : elle a donné lieu à une publication [24] et sera exploitée tout au long du **processus de recherche mené par le CERFACS et le Groupe Renault**. En plus de permettre la compréhension détaillée des phénomènes physiques dont un aperçu est donné dans ce document, elle permet également de fournir des références et des ordres de grandeurs utiles pour la construction de cas académiques plus simples, représentatifs de systèmes d'allumage par pré-chambre.

■ 7. Remerciements

Les résultats des travaux de développement menés dans le cadre du projet EAGLE et certains résultats sont publiés dans Mueller, C., B. Morcinkowski, C. Schermus, K. Habermann, T. Uhlmann, «Development of a pre-chamber for spark ignition engines in vehicle applications», 4^e Conference Internationale "Ignition Systems for Gasoline Engines", Dec. 2018

En outre, les auteurs remercient le projet EAGLE pour la publication des données relatives à ce travail sur les systèmes d'allumage par pré-chambre.

Le projet EAGLE a reçu un financement du programme de recherche et d'innovation, Horizon 2020 de l'Union européenne, au titre de la convention de subvention n° 724084.

■ 8. Références

- [1] P M Allison, P de Oliveira, A Glusti, and E Mastorakos. *Fuel*, 230, 274–281, 2018.
- [2] W P Attard, E Toulson, A Huisjen, X Chen, G Zhu, and H Schock. *SAE 2012 World Congress & Exhibition*, 1 :2012–01–0823–16, 2012.
- [3] S Biswas, S Tanvir, H Wang, and L Qiao. *Applied Thermal Engineering*, 106(C) :925–937, 2016.
- [4] O Colin and M Rudgyard. *Journal of Computational Physics*, 162(2) :338–371, 2000.
- [5] C Dapogny, C Dobrzynski, and P Frey. *Journal of Computational Physics*, 262(C) : 358–378, 2014.
- [6] Derek Dunn-Rankin and Peter Therkelsen, editors. *Lean Combustion - Technology and Control*. Academic Press, second edition edition, 2016.
- [7] Q Fei, A Shah, H Zhi-wei, P Li-na, P Tunestal, and B Xue-Song. *Combustion and Flame*, 188 :1–10, 2017.
- [8] F A R Filho, J G C Baêta, A F Teixeira, R M Valle, and J L F de Souza. *Energy Conversion and Management*, 121(C) :251–271, 2016.
- [9] F A R Filho, A F Teixeira, R B R da Costa, J G C Baêta, and R M Valle. *SAE Technical Paper Series*, pages 1–9, 2016.
- [10] L Y M Gicquel, N Gourdain, J-F Boussuge, H Deniau, G Staffelbach, P Wolf, and T Poinso. *Comptes Rendus Mécanique*, 339(2-3) :104–124, 2011.



[11] J O Hirschfelder, C F Curtiss, R Byron Bird, and M G Mayer. *Molecular theory of gases and liquids*, volume 26. Wiley New York, 1954.

[12] C W Hirt, A A Amsden, and J L Cook. *Journal of Computational Physics*, 135 :203–216, 1997.

[13] A Jamrozik. *Journal of Mechanical Science and Technology*, 29(5) :2269–2278, May 2015.

[14] G Lacaze, E Richardson, and T Poinot. *Combustion and Flame*, 156(10) :1993– 2009, 2009.

[15] H Li, G A Karim, and A Sohrabi. *Journal of Engineering for Gas Turbines and Power*, 131(1) :012801–7, 2009.

[16] V Moureau. PhD thesis, 2004.

[17] V Moureau, G Lartigue, Y Sommerer, C Angelberger, O Colin, and T Poinot. *Journal of Computational Physics*, 202(2) :710–736, 2005.

[18] F Nicoud and F Ducros. *Flow, Turbulence and Combustion*, 62(3) :183–200, 1999.

[19] L Y M Gicquel G Staffelbach T Poinot, G Staffelbach, and T Poinot. *Progress in Energy and Combustion Science*, 38(6) :782–817, 2012.

[20] A A. Quader. SAE International, 1976.

[21] R P Roethlisberger. *Applied Thermal Engineering*, pages 1–13, 2002.

[22] T Schoenfeld. *AIAA Journal*, 37(11) :1378–1385, 1999.

[23] S Yamaguchi, N Ohiwa, and T Hasegawa. *Combustion and Flame*, 59 :177–187, 1985.

[24] Q. Malé, G. Staffelbach, O. Vermorel, A. Misdariis, F. Ravet and T. Poinot. *Flow Turbulence and Combustion*, 2019.



Équipe

Julien CARMONA, Doctorant CERFACS

Safran Aircraft Engines

Jérôme DOMBARD, Chercheur sénior CERFACS

Ghislain LARTIGUE, Ingénieur de recherche CNRS, CORIA

Renaud MERCIER, Ingénieur de recherche, Safran Tech

Antony MISDARIIS, Chercheur sénior CERFACS

Vincent MOUREAU Chargé de recherche CNRS, CORIA

Dimitrios PAPADOGIANNIS, Ingénieur de recherche

Safran Tech

Eleonore RIBER, Chercheur sénior CERFACS

Gabriel STAFFELBACH, Chercheur sénior CERFACS

Safran Tech

CERFACS

CORIA

Passage à l'échelle des méthodes haute-fidélité pour les turbomachines

25 millions d'heures

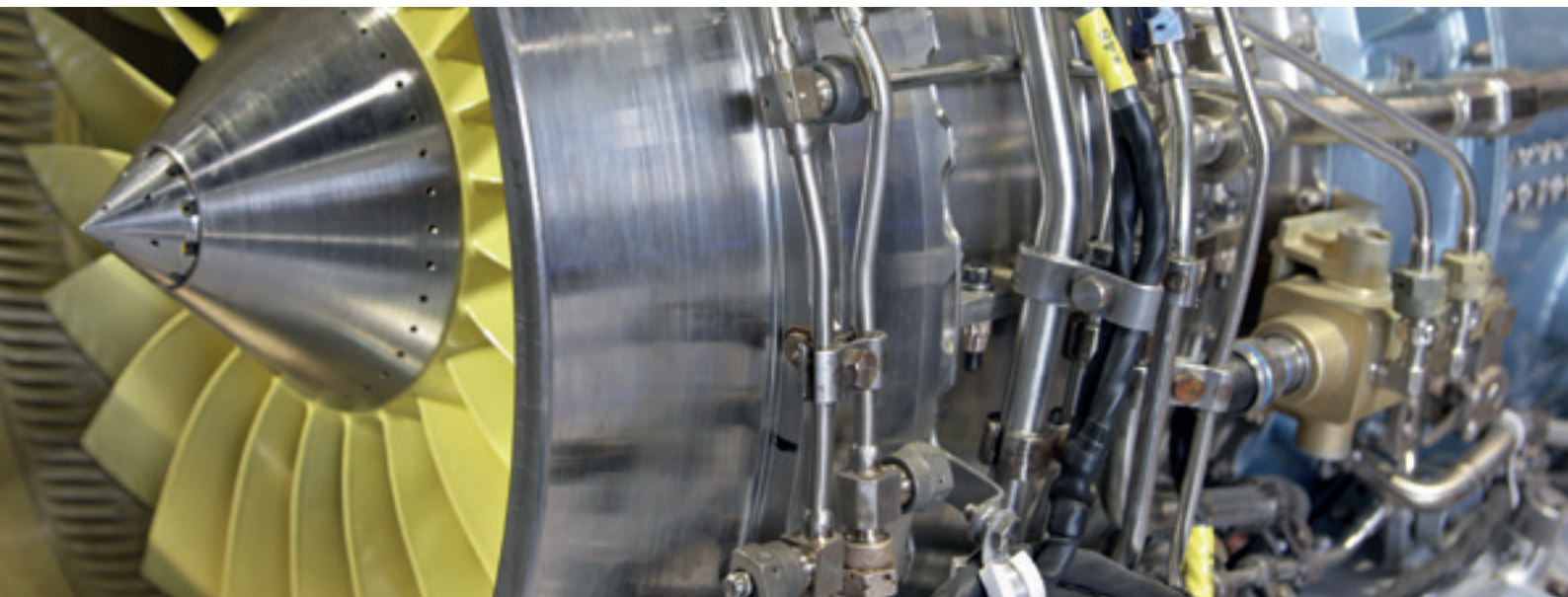
1. Introduction

Aujourd'hui, les industriels se basent de plus en plus sur les simulations numériques pour prédire les écoulements dans les moteurs aéronautiques. Il y a deux approches utilisées quotidiennement pour les différents composants d'un moteur.

La première, peu coûteuse, consiste à résoudre des équations stationnaires avec le formalisme **Reynolds Averaged Navier Stokes (RANS)**, en combinaison avec un modèle de turbulence, typiquement pour les écoulements autour des compresseurs et des turbines qui sont caractérisés par des nombres de Reynolds élevés. La seconde approche appelée **Simulation aux Grandes Échelles (Large Eddy Simulations (LES) en anglais)**, beaucoup plus coûteuse, est instationnaire et permet d'atteindre de hauts niveaux de fidélité. Dans le contexte industriel actuel, cette approche est utilisée pour la chambre de combustion qui présente des nombres de Reynolds plus faibles rendant ainsi possibles ces simulations. Bien que les deux approches soient bien matures et calibrées, des difficultés persistent :

- Pour les composants tournants (compresseurs et turbines), les simulations RANS et les corrélations qui y sont associées montrent des limites pour la conception et le design des turbomachines modernes et leur étude sur toute leur plage de fonctionnement. En effet, les écoulements autour des compresseurs et turbines sont caractérisés par divers phénomènes hautement instationnaires, comme des décollements, des instabilités tournantes/pompage (pour les compresseurs), de la transition laminaire - turbulente, des chocs (et leurs interactions avec les couches limites), des écoulements secondaires etc. Ces instationnarités de l'écoulement doivent être simulées pour **une grande gamme d'échelles spatio-temporelles**. Leur modélisation complète et générique est très difficile dans une approche RANS, ce qui peut conduire à des conceptions moins efficaces..





• Dans les chambres de combustion, bien que l'approche haute-fidélité LES soit aujourd'hui répandue dans l'industrie, les calculs restent simplifiés par rapport à la réalité, notamment pour l'injection et l'atomisation du carburant liquide : soit le carburant est supposé pré-éaporé et est injecté directement sous forme gazeuse, soit un modèle phénoménologique d'injection est utilisé, faisant l'hypothèse de gouttes sphériques dont les tailles et les vitesses sont présumées. Cependant, l'atomisation primaire du carburant, phénomène complexe principalement piloté par le champ aérodynamique, est connue pour impacter fortement la formation du mélange fuel-air, et par conséquent sa combustion ainsi que la formation des polluants. **L'accès difficile à des simulations de haute-fidélité pour l'atomisation primaire des jets carburant est aujourd'hui un verrou majeur** pour la meilleure prédiction de l'efficacité et de l'empreinte environnementale du moteur.

L'objectif global de ce Grand Challenge était d'évaluer la capacité de passage à l'échelle d'un grand nombre de cœurs de deux méthodes de calcul haute-fidélité sur la machine Joliot-Curie : Le calcul LES appliqué à deux compresseurs et la prédiction LES de l'atomisation carburant sur des configurations de types nappe et jet en écoulements transverses.



2. Wall-Modeled LES du compresseur basse pression DREAM

2.1 Configuration

Le compresseur axial choisi pour cette étude est le compresseur basse pression DREAM [6]. Ce compresseur, comprenant 1,5 étages (Stator – Rotor – Stator), est représentatif de compresseurs axiaux industriels et il est accompagné par des mesures expérimentales détaillées (mesures stationnaires et instationnaires). De plus, des calculs RANS sont disponibles. Cela permet de valider de manière fine l'approche LES pour les compresseurs axiaux et de comparer avec les calculs stationnaires RANS sur ce type de configurations.

Le coût élevé de l'approche LES pour les turbomachines ne permet pas de simuler le compresseur complet (calcul 360 degrés). La difficulté est que c'est fréquemment impossible de trouver un secteur périodique réduit parce que les concepteurs choisissent des nombres d'aubes par roue qui n'ont pas de commun diviseur. Pour dépasser cette difficulté, un secteur angulaire réduit est retrouvé en changeant très légèrement le nombre d'aubes de la roue mobile ; une aube est enlevée et les autres

sont modifiées par homothétie afin de garder la solidité, la charge aérodynamique et le débit inchangé. Cela nous permet d'avoir un domaine avec seulement 4 aubes d'entrée (ou *Inlet Guide Vane* (IGV) en anglais), 3 rotors et 4 stators.

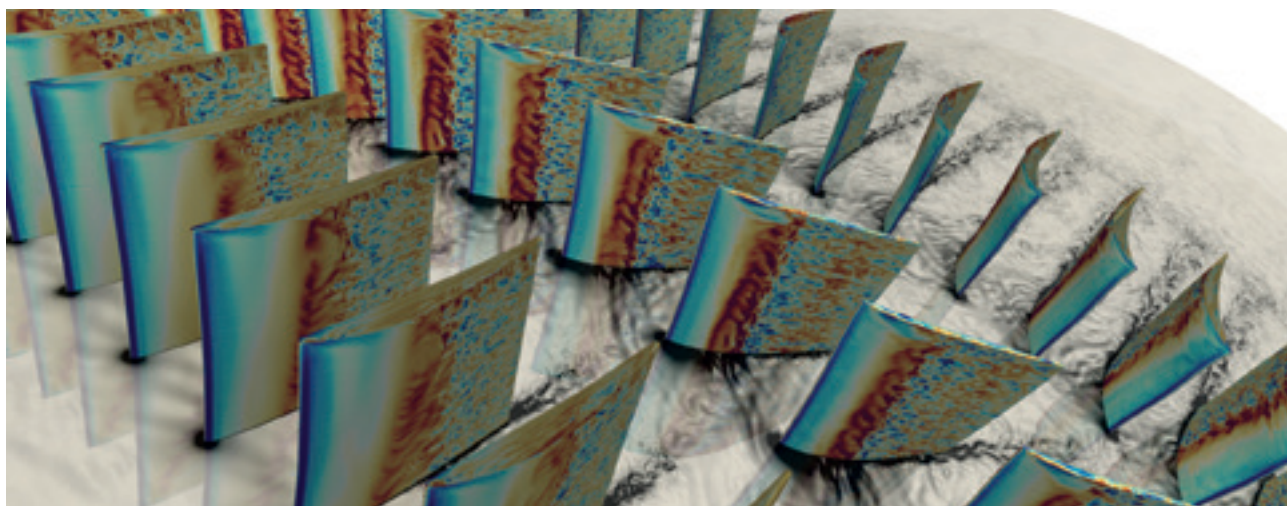
2.2 Setup numérique

Le maillage généré pour cette étude est un maillage multiéléments avec 10 couches de prismes autour des aubes (pour améliorer la résolution de l'écoulement dans la couche limite) et des tétraèdres dans le reste du domaine. La hauteur de la première couche des prismes a été $y^+ \approx 10 - 15$ ce qui a été montré suffisant pour les calculs WMLES. **Le maillage complet contient 480 millions de cellules.**

Le calcul a été effectué avec le code **TurboAVBP**, une version étendue du code AVBP qui est aujourd'hui une référence en matière de LES de turbomachines. Ce code utilise une méthode de volume/éléments finis cell-vertex et une discrétisation explicite en temps. Le schéma numérique utilisé dans cette étude est le schéma Lax-Wendroff, d'ordre 2 en temps et en espace. Les conditions aux limites suivent le formalisme NSCBC pour éviter les réflexions d'ondes acoustiques. A l'entrée, un profil de pression totale et de la température totale ainsi que les angles de l'écoulement sont imposés. A la sortie du domaine, la pression statique vise une pression moyenne avec une condition aux limites qui tend vers un équilibre radial.

Deux points de fonctionnement ont été calculés pendant le Grand Challenge, le point nominal et un point proche pompage. Le coût du calcul pour un tour complet du rotor est environ 300 000 heures CPU. Environ 10 tours

Figure 1 : *Frottement autour des aubes et gradient de la densité à 50% d'hauteur fonction du degré vilebrequin.*



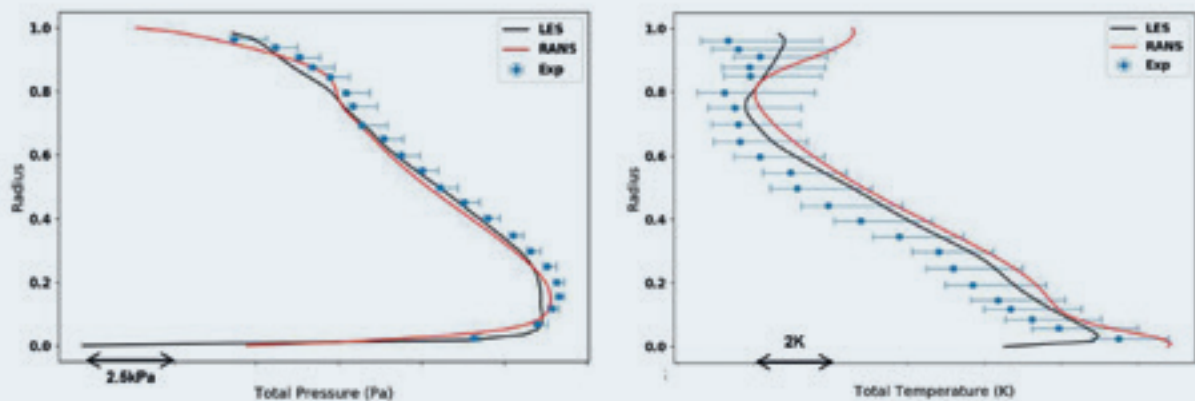


Figure 2 : Profils radiaux de la pression totale (gauche) et de la température totale (droite) derrière le dernier stator pour le point nominal

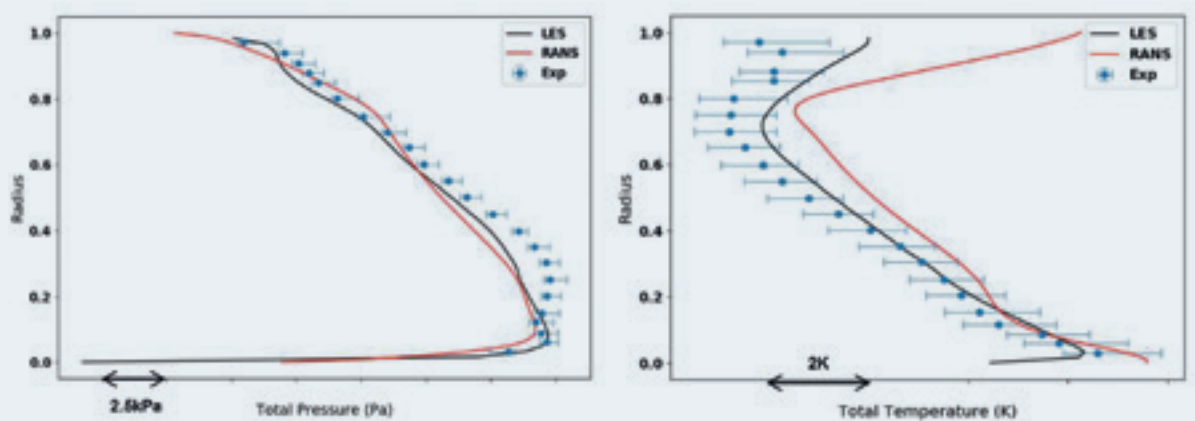


Figure 3 : Profils radiaux de la pression totale (gauche) et de la température totale (droite) derrière le dernier stator pour le point proche pompage.

ont été nécessaires pour chaque point de fonctionnement afin d'évacuer les transitoires et obtenir des champs moyennés, soient environ 3Mh utilisées.

2.3 Résultats

La Figure 1 montre le frottement autour des aubes et le gradient de la densité à 50% de hauteur pour le point nominal à un instant donné. La structure de l'écoulement prédit par l'approche LES est révélée : il y a de la transition-laminaire turbulent de l'écoulement dans les couches limites, de la séparation de l'écoulement, des interactions rotor-stator, des structures turbulentes etc. Cette richesse sur les prédictions est le point différenciant de l'approche

LES qui est instationnaire et capable de résoudre les grandes échelles de la turbulence.

Pour évaluer les prédictions du calcul LES de façon plus quantitative, les profils radiaux de la pression totale et de la température totale derrière le dernier stator, moyennés azimutalement, sont illustrés sur les Figures 2 et 3 pour le point nominal et proche pompage respectivement. Ils sont comparés avec les mesures expérimentales (avec leurs barres d'incertitude) ainsi qu'avec les calculs stationnaires RANS.

Sur le point nominal, Fig. 2, les résultats du calcul LES sont en très bon accord avec les mesures. Quelques différences existent proche du moyeu, probablement liés aux écoulements secondaires qui se développent au pied de l'aube. Il y a également un bon accord avec les calculs



RANS, qui sont d'habitude très fiables sur le point nominal. Sur le point proche pompage, Fig. 3, le bon accord entre le calcul LES et les mesures continue. Par contre, les prédictions du calcul RANS se dégradent légèrement, notamment sur la température totale. Ces différences illustrent bien les difficultés que les calculs stationnaires peuvent avoir sur des points de fonctionnement «off-design» où l'hypothèse de la stationnarité ainsi que les modèles de turbulence classiques atteignent leurs limites.

2.4 Perspectives

Avec les résultats de ce calcul, une publication est en cours de rédaction. De plus, le calcul proche pompage va servir comme point initial pour investiguer comment le compresseur entre en pompage et ainsi approfondir sur les phénomènes physiques complexes qui y interviennent. Cela est nécessaire afin de pouvoir concevoir des compresseurs toujours plus efficaces et avec des marges d'opérabilité suffisantes.

3. Simulation aux grandes échelles d'un banc turbine haute-pression d'un moteur d'hélicoptère

Le code de calcul de simulation aux grandes échelles Turbo AVBP a montré récemment sa capacité à calculer de manière satisfaisante les performances de turbomachines représentatives de configurations industrielles de plus en plus complexes - du compresseur à la turbine haute pression [Wang et al., 2013, Papadogiannis et al., 2014, 2016, Dombard et al., 2018, Duchaine et al., 2017] - avec toujours plus d'effets technologiques calculés (cavité sous stator, purges, refroidissement pas film etc.). Au-delà des performances au point de fonctionnement nominal, et comme présenté dans la section précédente Sec.2, le principal apport de la LES face au RANS est d'être relativement prédictif sur des points « *off-design* », de fournir une compréhension plus fine de la physique grâce à la dimension temporelle et d'être plus précis

avec moins de modélisation empirique. En ce sens, la LES peut aujourd'hui être considérée comme un outil complémentaire au RANS et intéressante pour le dessin de turbomachines si on en accepte le coût calcul.

L'objectif principal de cette étude est de revenir à une configuration plus simple, sans effets technologiques, instrumentée plus finement, afin de pouvoir comparer les différentes méthodes numériques. Le second objectif est de quantifier l'apport prédictif de ces méthodes dans des turbines où la roue haute pression (RHP) est de plus en plus chargée aérodynamiquement, que ce soit par le dessin de la RHP, le taux de détente de la turbine ou la vitesse de rotation de l'anneau. Un des nombres sans dimension qui caractérise cette charge lors du dessin de la RHP est le nombre de Zweifel (noté Zw), rapport de la corde sur la cambrure de la pale. Plus il est supérieur à un et plus la RHP est la charge aérodynamique de la RHP est grande. Cette étude propose donc de faire varier le nombre de Zweifel autour du point nominal à $Zw=0.95$ en l'augmentant à $Zw=1.09$ puis $Zw=1.26$. Les trois dessins de RHP calculés sont présentés Fig.4 à droite. Trois taux de pression et vitesses de rotation différentes ont aussi été calculés pour le point de dessin ($Zw=0.95$) à tout paramètres égaux par ailleurs. La base de données ainsi générée consiste donc en sept calculs distincts.

3.1 Configuration

La configuration choisie, présentée Fig.4 à gauche, est une portion sectorielle d'un banc d'une turbine HP comprenant deux distributeurs haute pression (DHP) et cinq RHP, instrumenté à Safran Helicopter Engines. Le secteur angulaire calculé est de 40 degrés et est à périodicité machine, au sens qu'aucune modification des aubages n'a été réalisée pour être axi-périodique.

Les données d'essais sont disponibles au plan d'entrée du domaine, au plan aval DHP et aval RHP.

3.2 Mise en données numérique

Le maillage utilisé dans cette étude consiste en un ensemble de **120 millions de tétraèdres** de l'ordre de la centaine. Cet ordre de grandeur est celui utilisé dans les études précédentes et a déjà montré un bon accord avec l'expérience lorsqu'une loi de paroi classique est utilisée [Duchaine et al., 2017]. Le code de calcul est Turbo-AVBP, qui est une instanciation du code



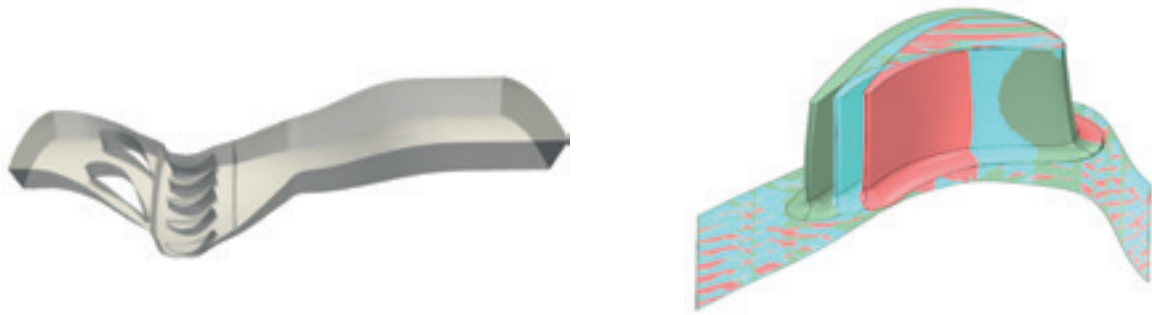


Figure 4 : Gauche: Domaine numérique calculé (2 distributeurs HP - 5 roues HP).
Droite: Vue de côté des trois roues HP calculées dans cette étude en fonction du nombre de Zweifel
(verte: $Zw = 0.95$, bleue: $Zw = 1.09$ et rouge: $Zw = 1.26$)

AVBP développé au CERFACS et à l'IFPen [Schoenfeld and Rudgyard, 1999], massivement parallèle et non structuré. Le schéma convectif est le schéma d'ordre deux en espace et en temps Lax-Wendroff, le modèle de sous maillage est WALE [Nicoud and Ducros, 1999]. Les conditions aux limites sont de type NSCBC [Poinsot and Lele, 1992] avec la pression totale, température totale et angles imposés en OD en entrée [Odier et al., 2018] et une contre pression statique en sortie qui dégénère vers le bon équilibre radial [Granet et al., 2010, Koupper et al., 2015]. L'ensemble des murs sont modélisés avec des lois de parois glissantes classiques (log-law) et les murs tournants utilisent la méthode ALE [Moureau et al., 2005].

Le banc est à froid, sans point chaud aval et il a été vérifié que la cartographie en entrée est relativement homogène, ce qui valide la valeur OD utilisée. De la turbulence est injectée avec une intensité de turbulence de 5% et une échelle intégrale de l'ordre d'un quart de la hauteur de veine. Ces quantités ont été corroborées à l'aide d'un calcul LES de l'ensemble du banc amont (non montré ici).

Un tour de machine coûte environ 30 kh CPU et environ 20 tours de machines ont été réalisés par calcul, soit une consommation d'environ 600kh CPU par calcul. La consommation totale pour la base de données (7 calculs) est donc d'environ 3 millions d'heures CPU une fois le setup numérique trouvé. Le nombre moyen de coeurs utilisé par calcul était de l'ordre de 4000, valeur jusqu'à laquelle un bon scaling a été observé.

3.3 Résultats

Une illustration de la grande complexité de l'écoulement est présentée Fig.5 avec le gradient de densité à 50% de hauteur de veine. Aux conditions choisies, le DHP est amorcé et l'écoulement devient supersonique au col, i.e le débit d'air qui passe dans l'étage de turbine est fixé par la pression totale, température totale, le gamma du fluide et la surface débitante. On peut observer une très forte interaction entre le sillage du DHP, le réseau de choc en aval du col du DHP et le choc fort au bord d'attaque de la RHP qui remonte en amont. Le débit et donc le réseau de choc dans les canaux de RHP varient donc en fonction de l'interférence géométrique entre le DHP et la RHP, ce qui peut être visible avec des sillages et chocs plus ou moins forts aux bords de fuite des RHP dans les différents canaux.

On voit ici tout l'intérêt d'avoir une méthode numérique instationnaire avec maillage glissant



Figure 5 : Gradient de densité à 50% de hauteur de vane - Zoom autour de la roue HP.



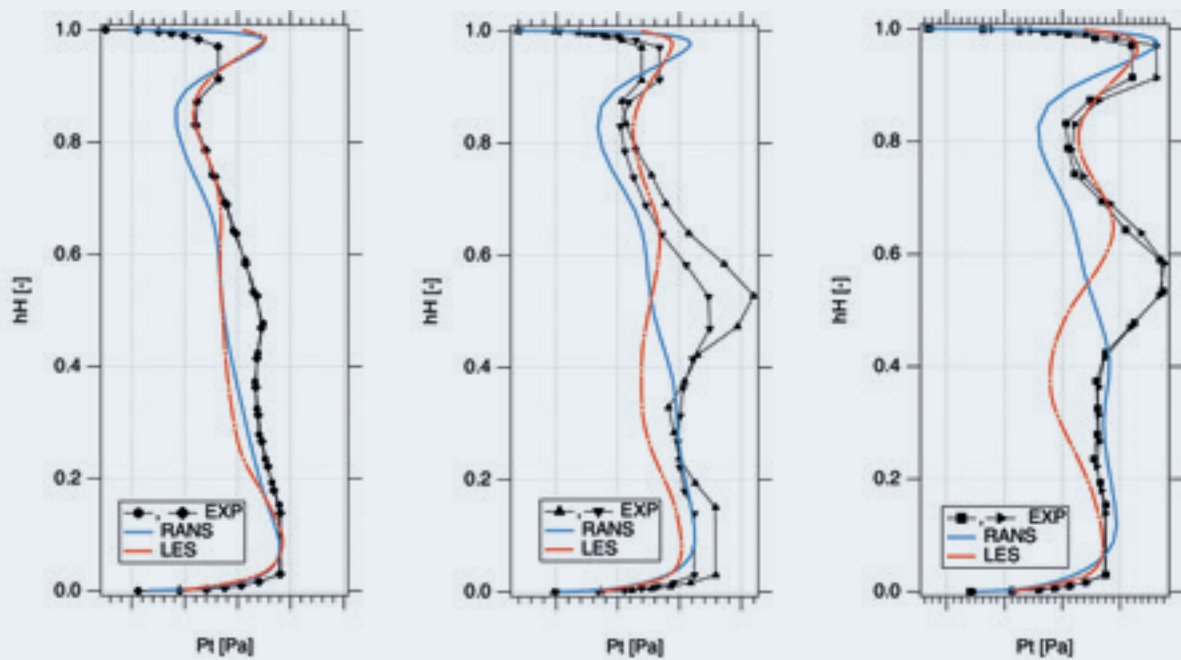


Figure 6 : Profils de pression totale au plan d'instrumentation en fonction du nombre de Zweifel (0.95 ; 1.09 ; 1.26). EXP (symboles), LES (ligne rouge) et RANS (ligne bleue).

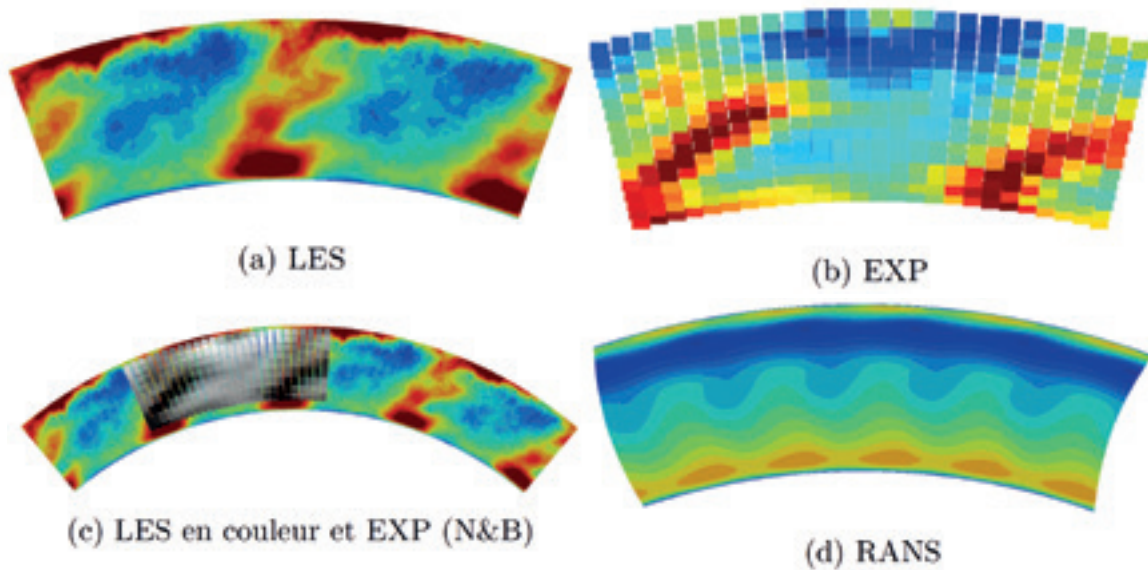


Figure 7 : Cartographies de pression totale au plan d'instrumentation aval roue HP (même échelle de couleur) pour $Zw = 0.95$. La cartographie exp. est surimposée à celle LES Fig. 7c

[Wang et al., 2014]. (i.e sans plan de mélange), où il y a une continuité spatiale et temporelle aux interfaces stator et rotor. Ce n'est par définition pas le cas pour les méthodes stationnaires RANS.

Une première comparaison des méthodes numériques face aux données expérimentales est proposée Fig.6 avec les profils radiaux la pression totale, moyennés en azimuth, tracés

en fonction de la hauteur de veine pour les trois nombres de Zweifel considérés. A noter que la même procédure de post traitement (même type de moyenne métier) a été utilisée pour les simulations numériques et les expériences. Ce point est déterminant pour comparer les données de manière pertinente. Au point de dessin, $Zw=0.95$, à la fois le RANS et la LES se comparent



favorablement aux données expérimentales. La LES prédit un peu mieux le profil en tête ($hH > 0.7$) que le RANS alors que c'est l'inverse en pied pour $hH < 0.4$. Cette observation reste vraie pour les trois nombres de Zweifel, Fig.6a à Fig.6c. Il peut néanmoins être observé qu'à la fois la forme à trois bosses et l'allure (la pression totale se «redresse» en tête) sont mieux captés par la LES que le RANS, même si le deuxième pic est capté un peu trop haut par la LES par rapport aux expériences à environ 65% de hauteur de veine.

Ces différents résultats sont concluants et valident l'approche LES.

Une possible explication peut être donnée en observant les cartographies 2D de pression totales Fig.7 acquises au plan d'instrumentation, où sont issues les profils 1D analysés précédemment. Tout d'abord, on peut noter sur la cartographie expérimentale (Fig.7b) que la trace la plus importante correspond aux sillages DHP (au nombre de deux). Ce n'est pas le cas pour la cartographie RANS (Fig.7d) où seules les traces des tourbillons de passage en pied et du jeu RNS en tête (plus faiblement) sont présentes. Cette observation est expliquée par la présence du plan de mélange entre les DHP et RHP qui moyenne le signal en azimuth. C'est donc normal que les traces des DHP n'apparaissent pas. A l'opposé, la LES capte raisonnablement bien à la fois la bonne topologie et les bons niveaux de pression totale, même si les sillages semblent plus anguleux et puissants sur la première moitié de veine pour l'expérience. C'est sans doute pour cette raison que le profil 1D Fig.6a a une pression totale moins important dans cette zone, i.e que c'est dû à un effet de moyenne azimuthale.

3.4 Perspectives

La première perspective est le dépouillement du reste de la base de données avec les calculs réalisés aux deux autres taux de détente et vitesse de rotation.

Un papier sera écrit pour valoriser ce travail une fois la base de données complète analysée et comparée aux données expérimentales et aux différentes méthodes numériques.

4. Simulation aux grandes échelles d'une configuration académique d'atomisation de type airblast

Bien que l'approche haute-fidélité LES soit aujourd'hui répandue dans l'industrie pour la modélisation des chambres de combustion, l'injection et l'atomisation du carburant liquide y sont simplifiées : soit le carburant est supposé pré-évaporé et est injecté directement sous forme gazeuse, soit un modèle phénoménologique d'injection est utilisé pour prédire les tailles et les vitesses des gouttes en faisant l'hypothèse de leur sphéricité. Or, l'atomisation primaire du carburant, phénomène complexe principalement piloté par le champ aérodynamique, est connue pour impacter fortement la formation du mélange fuel-air, et par conséquent sa combustion ainsi que la création des polluants. Ceci fait de la simulation de l'atomisation primaire du carburant un verrou majeur pour améliorer les calculs hautes fidélités et ainsi mieux prédire l'efficacité et l'empreinte environnementale des moteurs.

Dans la présente étude, on se proposait d'avancer vers la simulation des injecteurs de type airblast. Dans ce type d'injecteur (fig. 8), le spray primaire issu d'un injecteur pressure swirl impacte les parois d'un diffuseur. Un film liquide mince est ainsi créé. Il ruisselle jusqu'au bout de la lèvre du diffuseur

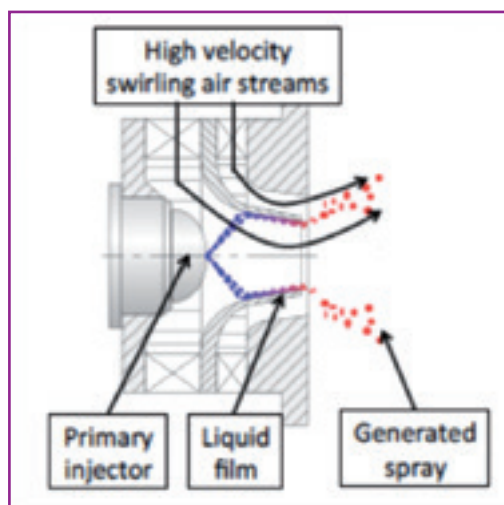


Figure 8 : Schéma de principe d'un injecteur de type Airblast. [Chaussonnet, 2014].



où il s'atomise. Le spray ainsi formé est alors très différent du spray primaire. Ces injecteurs ont des performances très stables (taille de gouttes, angle de spray...) dans une large gamme de condition de fonctionnement (pression, température). Ceci les

rend très intéressants pour la stabilité et la qualité de la combustion.

Les injecteurs airblast sont pour l'heure modélisés à travers des approches phénoménologiques comme celle proposée dans la thèse de G. Chaussonnet [Chaussonnet, 2014] et établie à partir de la configuration académique d'atomisation en bout de lèvre du KIT (Karlsruhe Institute of Technology) [Gepperth, 2012]. Cette configuration expérimentale est représentative des phénomènes d'atomisation dans les injecteurs aéronautiques de type airblast.

L'objectif de ce Grand Challenge était donc de

s'appuyer sur les récents développements dans le solveur SPRAY de YALES2 pour calculer l'atomisation en bout de lèvre avec la méthode de suivi d'interface et de comparer les résultats aux données expérimentales du KIT (en particulier la distribution de gouttes obtenue après l'atomisation

La simulation de l'atomisation primaire du carburant, un verrou majeur pour améliorer les calculs hautes fidélités et ainsi mieux prédire l'efficacité et l'empreinte environnementale des moteurs.

4.1 Setup numérique

Différents maillages ont été testés pour cette étude. Dans tous les cas, seule une tranche de la géométrie de 5 mm d'épaisseur a été calculée afin de pouvoir porter l'effort de raffinement sur la zone de développement du film liquide et d'atomisation primaire en gardant une taille de maillage raisonnable. **Les tailles de mailles dans ces zones sont de 60 microns pour le maillage le plus grossier (7M de tétraèdres) et de 20 microns pour le maillage de plus fin (97M de tétraèdres).**

Les calculs ont été réalisés avec le solveur SPRAY du code YALES2. Ce code permet de résoudre des écoulements incompressibles sur des maillages non-structurés par une méthode volume finie. Le schéma numérique choisi dans cette étude est le schéma TFV4A d'ordre 4 en espace et en temps. A l'entrée du domaine, le débit du gaz (air à 298K) est imposé afin d'obtenir comme dans l'expérience une vitesse débitante de 70m/s au niveau de la plaque. Le liquide est, quant-à-lui, injecté via une fente placée directement sur la plaque 44.5mm avant le bord d'atomisation. Le liquide choisi est le Shellsol D70 utilisé dans les expériences de Gepperth [Gepperth, 2012] et est injecté à un débit de 1,925 g/s (point de fonctionnement $A_f = 50\text{mm}^2/\text{s}$ dans [Gepperth, 2012]). Le temps de calcul consommé pour cette étude est d'environ 2 millions d'heures.

4.2 Résultats

La figure 12 présente une visualisation expérimentale du film liquide et de son atomisation en bout de plaque.

Il convient, tout d'abord, de remarquer les instabilités 2D de type Kelvin-Helmoltz présentes à la surface du film. Comme l'illustre la figure 10, qui montre l'évolution de la surface du film liquide au court du temps, Ces instabilités ne sont pas prédites par le code. En effet, des erreurs numériques à l'interface perturbent fortement le développement des instabilités. Ces problèmes numériques sont dus à la méthode d'estimation de la courbure de l'interface de YALES2.

Ensuite, comme le met en évidence la figure 13 (zoom sur l'interface liquide/gaz au niveau du bord de plaque), l'atomisation se fait principalement suivant le régime de «bag break-up» : après une phase d'accumulation de liquide au niveau du bord d'atomisation, l'interface s'amincit et en se déchirant forme des ligaments qui par la suite deviendront des gouttes. La figure 11, qui présente une

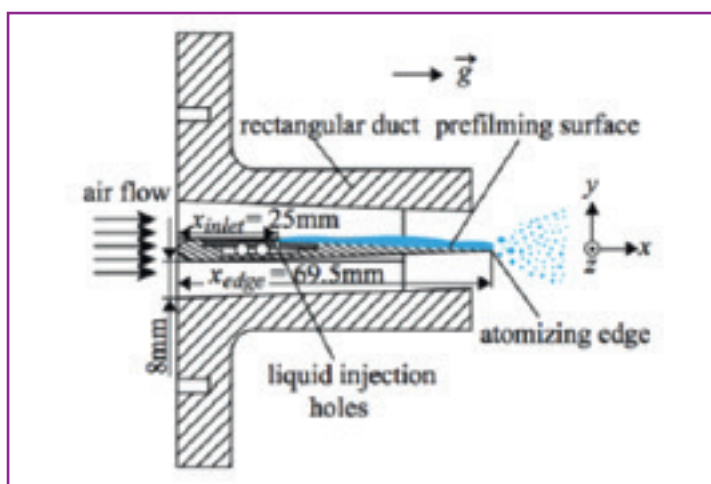


Figure 9 : Schéma de l'expérience réalisée au KIT. [Chaussonnet, 2014] primaire).



Figure 10 : Snapshots du mouvement de l'interface liquide/gaz - vue de côté.

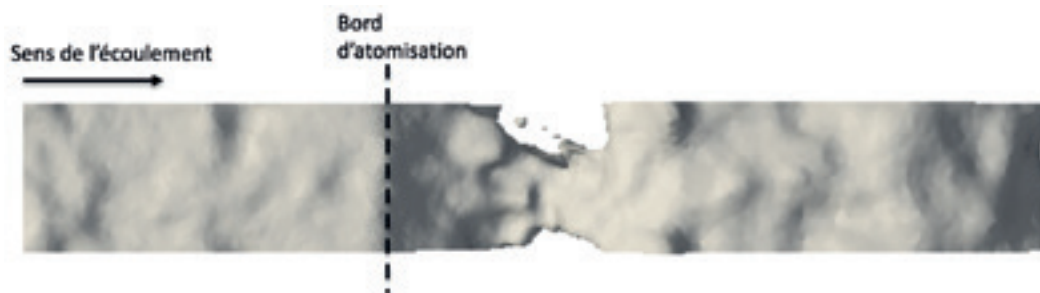


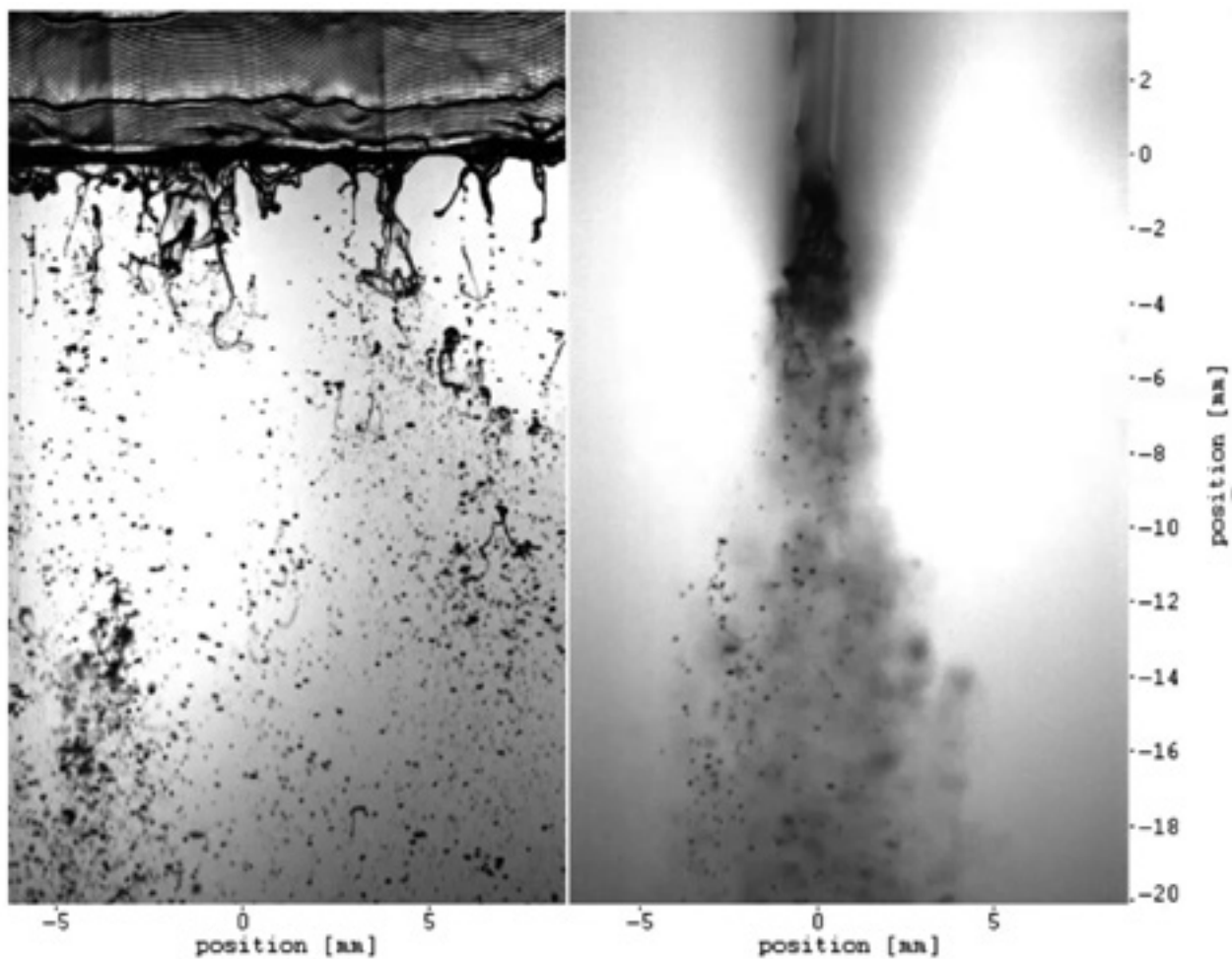
Figure 11 : Snapshot l'interface liquide/gaz au bord d'atomisation - vue du dessus.

visualisation vue du dessus de l'interface liquide/gaz au niveau du bord d'atomisation, montre que le code YALES2 ne permet pas de prédire non plus l'atomisation. En effet, il y a bien une accumulation de liquide en bout de plaque mais elle est bien supérieure aux expériences. La séparation du liquide se fait ensuite par blocs sans aucune correspondance avec les données expérimentales. Ce phénomène peut en partie s'expliquer par la réinitialisation du profil tangente-hyperbolique qui modélise l'interface dans la méthode Level-Set [Desjardin et al., 2008]. Cette réinitialisation a pour but de contrer la diffusion numérique du profil mais elle crée aussi mécaniquement une tension de surface numérique qui perturbe le processus d'atomisation.

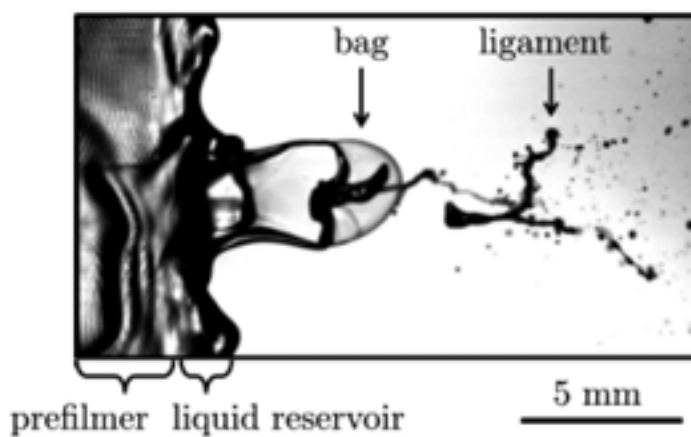
4.3 Conclusion de l'étude

Même si cette étude ne permet pas de retrouver les résultats expérimentaux, il convient de rappeler que la réalisation de simulations d'atomisation primaire est pour l'heure réservée à des codes cartésiens où les configurations sont de fait très loin des réalités industrielles. L'intérêt majeur de cette étude a été de mettre en évidence les faiblesses actuelles du solveur SPRAY de YALES2 et de définir des axes de travail pour aller vers la simulation de véritables injecteurs de type Airblast.





| Figure 12 : Vue expérimentale du film et de l'atomisation en bout de plaque [Chaussonnet, 2014].



| Figure 13 : Phénomène d'atomisation en bord de plaque [Chaussonnet, 2015].



5. Passage à l'échelle d'une méthode d'adaptation de maillage dynamique pour le calcul LES de l'atomisation carburant

Une approche attractive pour pallier les difficultés exposées en Section 4 avec le solveur SPRAY de YALES2 consiste à assurer une résolution en maillage constante et très fine autour de l'interface liquide-gaz en tout temps de la simulation. Dans cet optique, un couplage entre le solveur de SPRAY et une adaptation dynamique du maillage a récemment été développée par Leparoux et al. [2018]. Cette communication scientifique propose notamment une description détaillée de la méthodologie de couplage entre le solveur SPRAY de YALES2 et l'adaptation de maillage parallèle. La configuration considérée ici est celle d'un jet injecté dans un écoulement transverse mesurée par Ragucci et al. [2007]. Le domaine fluide considéré est représenté dans la Figure 14. Le point de fonctionnement #1, détaillé dans le Tableau 1, est sélectionné pour cette étude. Le maillage initial, utilisé comme support de l'adaptation de maillage dynamique locale, est représenté en Figure 15.

Pour ce point de fonctionnement, Leparoux et al. [2018] ont réalisé une simulation de l'atomisation du jet carburant avec une résolution de l'interface $\Delta_x = 20\mu\text{m}$ et $\Delta_z = 10\mu\text{m}$. **Les tailles de maillages obtenues, une fois le jet totalement développé, étaient alors de $N = 130.0 \times 10^6$ cellules et $N = 1.0 \times 10^9$ cellules, respectivement.** Ces calculs ont été conduits sur environ 500 et 5000 cœurs, respectivement. La Figure 16 montre une vue instantanée de l'interface liquide-gaz pour le calcul $\Delta_x = 10\mu\text{m}$.

L'objectif de cette partie du Grand Challenge Joliot-Curie était de repartir du calcul $\Delta_x = 10\mu\text{m}$ afin d'évaluer plus profondément l'impact de la taille de maille à l'interface sur la distribution de gouttes et éventuellement conclure sur une résolution minimale assurant la convergence.

Une première tentative a consisté à réduire la taille de maille à l'interface à $\Delta_x = 5\mu\text{m}$. Le nombre de mailles attendu était alors d'environ

	#1		#2
Nozzle diameter, D (mm)	0,5		0,5
Gas velocity, u_g (m/s)	37,0		50,0
Liquid velocity, u_l (m/s)	17,0		43,0
Gas density, ρ_g (kg/m ³)	11,59		11,59
Liquid density, ρ_l (kg/m ³)	790,0		790,0
Momentum flux ratio, q (-)	14,2		49,6
Aerodynamic Weber number, We_a (-)	63,5		406,2
Ohnesorge number, Oh (-)	0,0146	47	0,0146

Tableau 1 : Conditions Opératoires

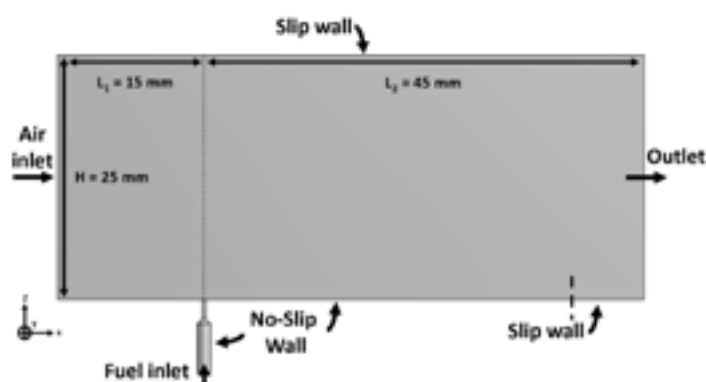
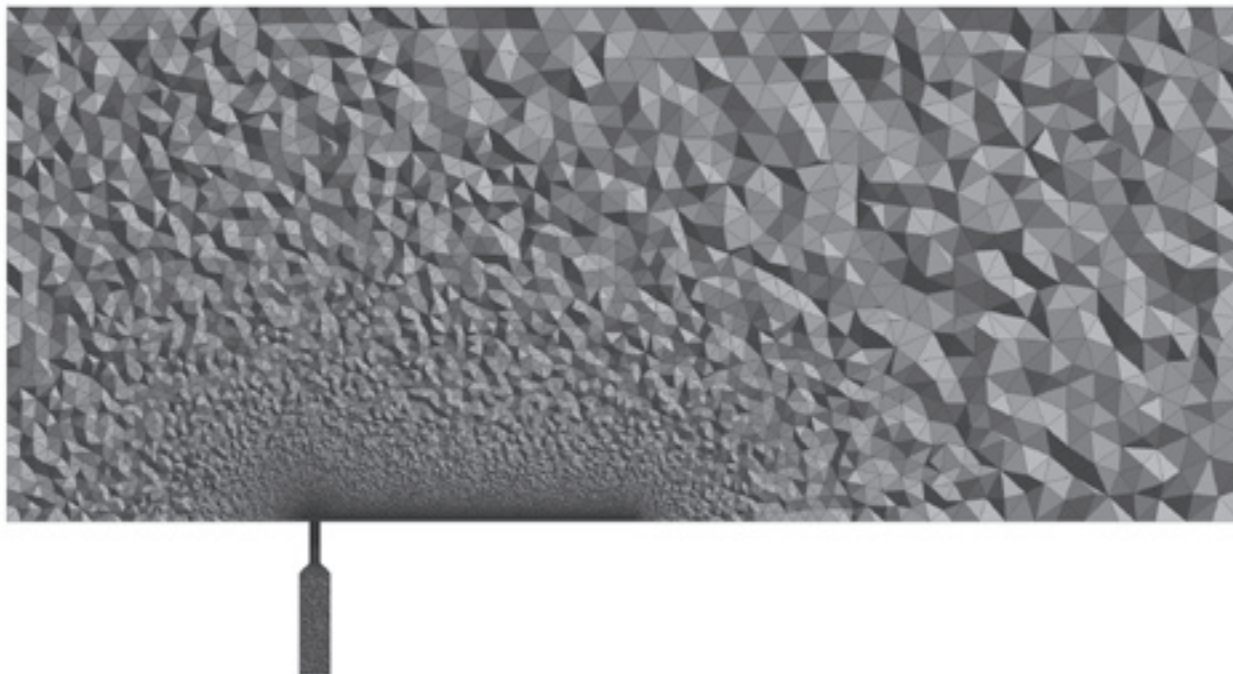


Figure 14 : Géométrie de la configuration étudiée.

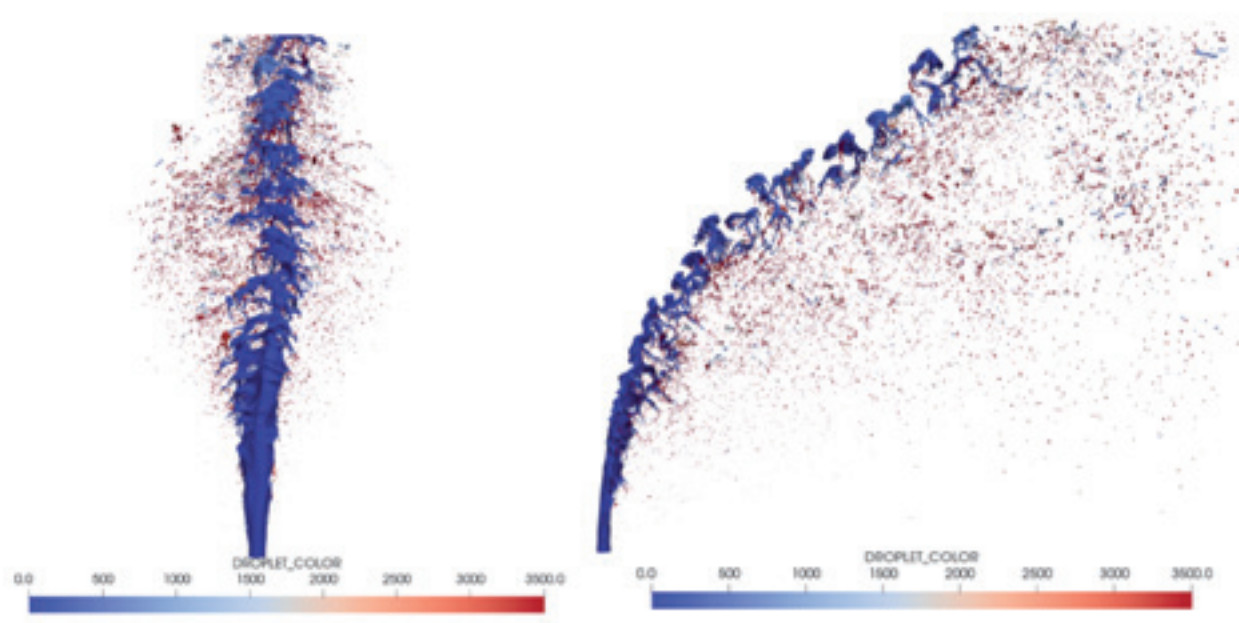
$N = 4.0 \cdot 10^9$ pour un nombre de cœurs skylake cible à 20000. La méthodologie de calcul proposée dans [Leparoux et al., 2018] s'est montrée robuste au changement de taille de maille cible. En revanche, nous avons constaté une dégradation catastrophique des performances de la boucle temporelle. En effet, les communications point à point de l'adaptation de maillage dynamique ainsi que les communications collectives du solveur linéaire (équation de poisson sur la pression) ont montré des performances très ralenties. Pour des nombres de cœurs à 20000, il était alors impossible de poursuivre les calculs vu la lenteur de la boucle. Un grand nombre d'investigations détaillées ont été réalisées en collaboration étroite avec le support du TGCC que nous souhaitons remercier chaleureusement ici. Ces investigations se sont étalées depuis l'ouverture du Grand Challenge jusqu'à sa fermeture après plusieurs extension de quota. Quelques points marquants sont ressortis des investigations:

- L'augmentation du nombre de cœurs par nœuds par rapport à la configuration machine précédente (machine CCRT Cobalt





| Figure 15 : Maillage initial.



| Figure 16 : Vue instationnaire de l'interface.



avec processeurs Intel Broadwell) semble pénaliser les communications collectives pour les calculs sur un grand nombre de cœurs (>5000). Les méthodes numériques nécessitant un grand nombre de ces communications s'en voient pénalisées fortement pour leur passage à l'échelle.

- Les approches permettant la désynchronisation des communications collectives [Ghysels and Vanroose, 2014] lors des itérations du solveur linéaire nous ont permis d'améliorer les performances pour des nombres de cœurs inférieurs à 10 000. Ces améliorations semblent repousser légèrement le problème mais ne permettent pas de le résoudre au-delà de 10 000 cœurs.

- L'ensemble des tests et investigations effectuées l'ont été sous le paradigme de parallélisme full-MPI. Le **paradigme hybride entre mémoire partagée et distribuée semble être une piste intéressante pour adresser le passage à l'échelle sur les nouvelles machines de type *Many Integrated Cores*.**

6. Conclusion générale du projet

Ce projet Grand Challenge est un succès ! Les différentes configurations visant des calculs sur des nombres de cœurs inférieurs à 5 000 ont pu être calculées avec d'excellentes performances. Les premiers résultats ont été présentés dans ce rapport et les dépouillements se poursuivent. Du point de vue HPC, les différentes tentatives de passage à l'échelle ont permis **d'identifier les marges de progression informatiques et de commencer à évaluer plusieurs solutions** telles que la désynchronisation des communications, le paradigme hybride OpenMP-MPI, etc.

7. Remerciements

Le Très Grand Centre de Calcul (TGCC) est remercié pour la fourniture des heures CPU consommées lors de ce projet. L'équipe support est aussi chaleureusement remerciée pour son implication et son efficacité.

8. Bibliographie

[1] G. Chaussonnet. *Modeling of liquid film and breakup phenomena in Large-Eddy Simulations of aeroengines fueled by airblast atomizers*. PhD thesis, INP Toulouse, 2014.

[2] G. Chaussonnet. A new phenomenological model to predict drop size distribution in large-eddy simulations of airblast atomizers. *International Journal of Multiphase Flow*, 2015.

[3] Jérôme Dombard, Florent Duchaine, Laurent Gicquel, Gabriel Staffelbach, Nicolas Buffaz, and Isabelle Trébinjac. Large eddy simulations in a transonic centrifugal compressor. In *ASME Turbo Expo 2018: Turbomachinery Technical Conference and Exposition*. American Society of Mechanical Engineers, 2018.

[4] F. Duchaine, J. Dombard, LYM Gicquel, and C. Koupper. On the importance of inlet boundary conditions for aerothermal predictions of turbine stages with large eddy simulation. *Computers & Fluids*, 154:60–73, 2017.

[5] S. Gepperth. *International Conference on Ligament and Droplet Characteristics in Prefilming Airblast Atomization (ICLASS)*, 2012.

[6] P. Ghysels and W. Vanroose. Hiding global synchronization latency in the preconditioned conjugate gradient algorithm. *Parallel Computing*, 40(7):224 – 238, 2014. ISSN 0167-8191. 7th Workshop on Parallel Matrix Algorithms and Applications.

[7] V. Granet, O. Vermorel, T. Leonard, L. Gicquel, and T. Poinso. Comparison of nonreflecting outlet boundary conditions for compressible solvers on unstructured grids. *AIAA journal*, 48(10):2348–2364, 2010.

[8] C. Koupper, L. Gicquel, F. Duchaine, and G. Bonneau. Advanced combustor exit plane temperature diagnostics based on large eddy simulations. *Flow, Turbulence and Combustion*, 95(1):79–96, 2015.

[9] Julien Leparoux, Renaud Mercier, Vincent Moureau, and Haris Musaefendic. Primary atomization simulation applied to a jet in crossflow aeronautical injector with dynamic



mesh adaptation. *International Conference on Liquid Atomization and Spray Systems (ICLASS)*, 2018.

[10]. V Moureau, G Lartigue, Y Sommerer, C Angelberger, O Colin, and Thierry Poinsot. Numerical methods for unsteady compressible multi-component reacting flows on fixed and moving grids. *Journal of Computational Physics*, 202(2):710–736, 2005.

[11]. Franck Nicoud and Frédéric Ducros. Subgrid-scale stress modelling based on the square of the velocity gradient tensor. *Flow, turbulence and Combustion*, 62(3):183–200, 1999.

[12]. V. Moureau O. Desjardin and H. Pitsch. An accurate conservative level set / ghost fluid method for simulating turbulent atomization. *Journal of Computational Physics*, 227: 8395–8416, 2008.

[13]. Nicolas Odier, Marlène Sanjosé, Laurent Gicquel, Thierry Poinsot, Stéphane Moreau, and Florent Duchaine. A characteristic inlet boundary condition for compressible, turbulent, multispecies turbomachinery flows. *Computers & Fluids*, 2018.

[14]. Dimitrios Papadogiannis, Florent Duchaine, Frédéric Sicot, Laurent Gicquel, Gaofeng Wang, and Stéphane Moreau. Large eddy simulation of a high-pressure turbine stage: Effects of sub-grid scale modeling and mesh resolution. *ASME Paper No. GT2014-25786*, 2014.

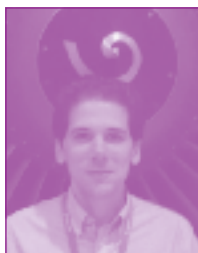
[15]. Dimitrios Papadogiannis, Florent Duchaine, Laurent Gicquel, Gaofeng Wang, and Stéphane Moreau. Effects of subgrid scale modeling on the deterministic and stochastic turbulent energetic distribution in large-eddy simulations of a high-pressure turbine stage. *Journal of Turbomachinery*, 138(9):091005, 2016.

[16]. T J Poinsot and SK Lele. Boundary conditions for direct simulations of compressible viscous flows. *Journal of Computational Physics*, 101(1):104–129, 1992.

[17]. R. Ragucci, A. Bellofiore, and A. Cavaliere. Trajectory and momentum coherence breakdown of a liquid jet in high-density air cross-flow. 17:47–70, 2007.

[18]. Thilo Schoenfeld and Michael Rudgyard. Steady and unsteady flow simulations using the hybrid flow solver avbp. *AIAA journal*, 37(11):1378–1385, 1999.

■ Équipe projet



De gauche à droite
Dimitrios PAPADOGIANNIS
Renaud MERCIER
Ghislain LARTIGUE



Jérôme DOMBARD
Julien CARMONA
Vincent MOUREAU



[19]. Gaofeng Wang, Dimitrios Papadogiannis, Florent Duchaine, Nicolas Gourdain, and Laurent YM Gicquel. Towards massively parallel large eddy simulation of turbine stages. *ASME Paper No. GT2013-94852*, 2013.

[20]. Gaofeng Wang, Florent Duchaine, Dimitrios Papadogiannis, Ignacio Duran, Stephane Moreau, and Laurent Y M Gicquel. An overset grid method for large eddy simulation of turbomachinery stages. *Journal of Computational Physics*, 274:333–355 October 2014.



Équipe

Andrea MAFFIOLI, Post-doctorant - Ecole Centrale de Lyon / LMFA

Alexandre DELACHE, Enseignant Chercheur

Université Jean-Monnet de Saint-Etienne / LMFA

Fabien GODEFERD, Directeur de recherche CNRS

Ecole Centrale de Lyon / LMFA



Séparer les ondes et les tourbillons dans un milieu turbulent en présence de stratification

10 millions d'heures

Notre recherche concerne l'importance des ondes dans un régime turbulent stratifié, régime que l'on retrouve par exemple dans l'océan ou l'atmosphère. Nous avons opté pour un Grand Challenge sur Joliot-Curie car nous avons pu avoir un très grand nombre d'heures de calcul (10 millions d'heures) sur un court laps de temps. Nous avons pu alors concentrer et coordonner nos efforts afin de produire des simulations à hautes résolutions et obtenir des régimes d'écoulement hors de notre portée en si peu de temps. Les résultats nous ont permis de répondre à la question concernant la partition énergétique entre ondes et vortex dans un régime turbulent stratifié. Nous avons trouvé la signature des ondes internes dans ces régimes turbulents : la relation de dispersion valable sous hypothèse linéaire, a pu être extraite pour des régimes fortement non linéaires (turbulents). D'autre part, nous avons montré que les ondes contiennent au moins 30 % de l'énergie totale, proportion qui est largement sur-évalué par la décomposition classique de Riley ce qui aura un impact sur les paramétrisations turbulentes océaniques et atmosphériques.

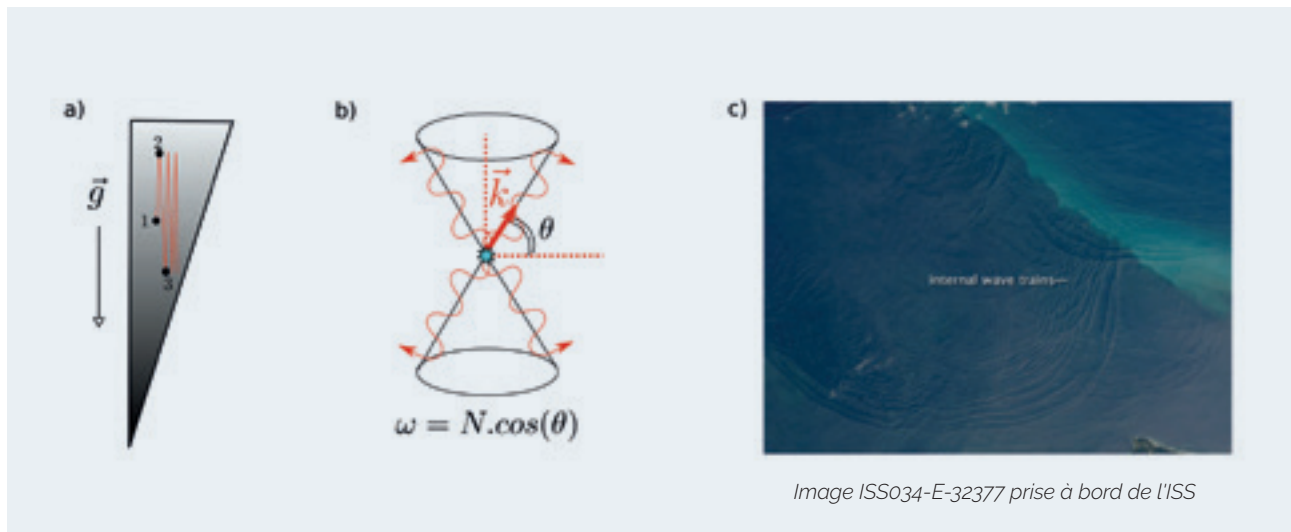
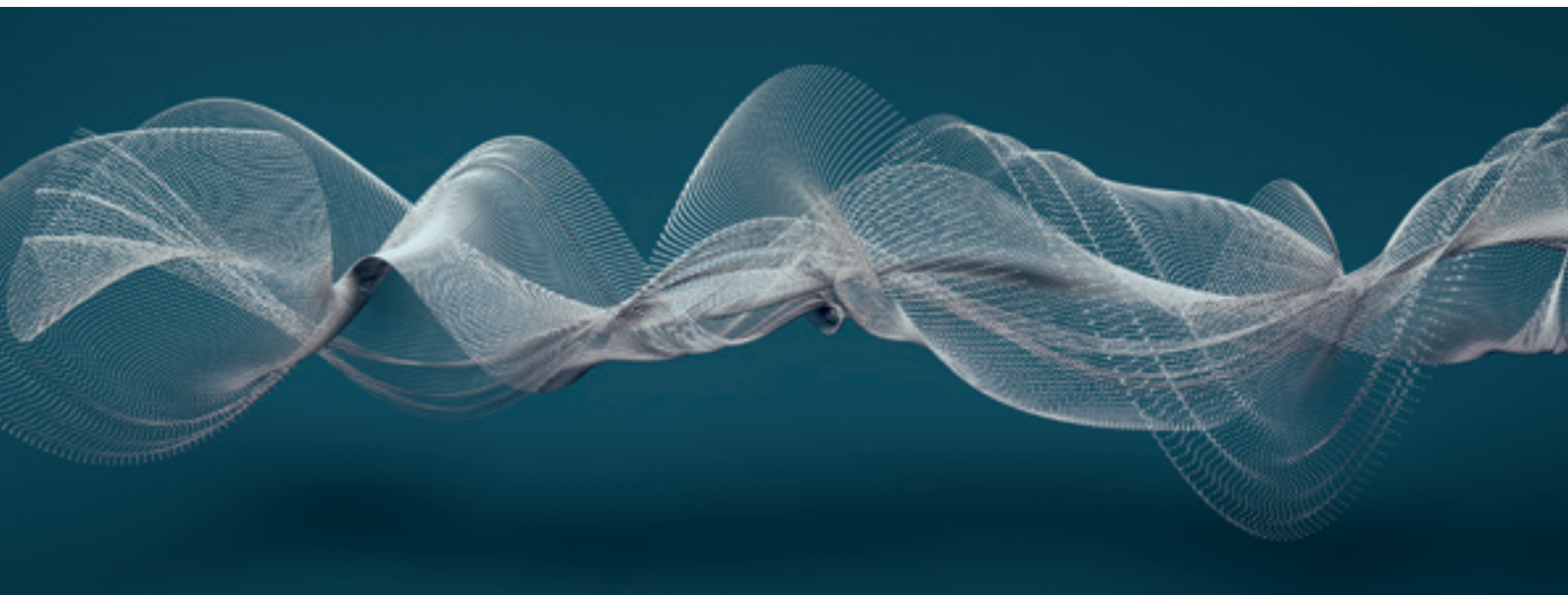
1. Contexte

1.1 Ondes et turbulence dans les écoulements géophysiques.

L'océan ou l'atmosphère sont le siège d'une multitude de phénomènes physiques spécifiques mais à l'intérieur de ces deux milieux, on rencontre régulièrement un mélange d'ondes internes dues à la gravité avec un état turbulent du milieu fluide. Ces deux phénomènes permettent de brasser les énergies à la surface de la terre, il est donc crucial de comprendre l'interaction entre eux.

Les ondes internes nécessitent une stratification stable du milieu fluide, c'est à dire que les couches de fluide sont de plus en plus lourdes avec la profondeur. Cela peut être due par exemple, à des différences de température ou de salinité. Si une particule fluide en équilibre dans une couche de fluide, est déplacée dans une couche de fluide plus légère (de 1 à 2 sur la figure 1 a) alors cette particule finit par retomber sous l'effet de la gravité. Or par inertie, elle retombe plus bas que sa position initiale (de 2 à 3 sur la figure 1 a). Ce faisant, elle est se retrouve dans une couche de fluide plus lourde et finit par





remonter plus haut que sa position initiale (de 3 à 2 sur la figure 1 a). Cela engendre des oscillations verticales qui sont à l'origine des ondes internes. Ces ondes sont caractérisées par une relation de dispersion reliant la fréquence de cette onde avec la fréquence de Brunt-Väisälä N , liée à la stratification moyenne, et l'angle entre le nombre d'onde et l'horizontal (figure 1 b). Sous certaines conditions, il est possible d'observer ces ondes internes à la surface des eaux (figure 1 c).

En présence de turbulence, les ondes internes s'étalent sur une certaine gamme d'échelle dépendant de la stratification ainsi que du taux de turbulence [1]. Suivant cette gamme, le déferlement de ces ondes internes à l'intérieur du milieu engendre des « vagues internes » qui mélangent plus ou moins bien les couches de densités. Ces échelles sont trop petites pour être entièrement résolues et sont donc paramétrées dans les modèles océaniques et atmosphériques.



1.2 Objectifs scientifiques

L'objectif de ce grand défi est de répondre à une question de longue date en matière de turbulence : quelle est l'influence des ondes de gravité internes sur la dynamique de la turbulence fortement stratifiée ?

Plus précisément, notre travail cherche à comprendre et à quantifier l'importance des ondes et des vortex dans les milieux stablement stratifiés soumis à un état turbulent. En effet, on utilise classiquement une décomposition onde/vortex basée sur le travail de Riley et al. [2] mais cette décomposition est limitée car elle ne prend uniquement en compte que l'espace ou la caractéristique principale d'une onde est de se développer en temps et en espace : une analyse spatio-temporelle est alors nécessaire en effectuant une transformée de Fourier en temps et en espace des champs 3D. Les résultats qui en découlent, pourront aider à formuler des paramétrages plus physiques et finalement plus précis à utiliser dans les modèles comme les modèles de circulation générale (MCG) pour prédire le climat.

numérique directe des équations de Navier-Stokes des grandes échelles jusqu'aux plus petites échelles, sans modèles. C'est un code Fortran et MPI. L'espace physique est découpé entre chaque processus grâce à une grille 2D cartésienne de taille $M_1 \times M_2$ (décomposition de type « pencil »). La sortie des fichiers s'effectue par la librairie HDF5. Il tourne régulièrement sur tous types de plateformes : des machines régionales (FLMSN) aux machines nationales OCCIGEN (CINES) ou TURING (IDRIS). Il permet de simuler les écoulements soumis à la stratification, à la rotation, au champ magnétique ainsi qu'aux effets non-newtoniens. La présence de paroi peut être utilisée grâce à une méthode de pénalisation. Il est développé au sein du laboratoire LMFA depuis une vingtaine d'années et il a subi deux évolutions majeures en 2006 et 2011 afin de le rendre massivement parallèle. Jusqu'à présent, une trentaine de personnes l'ont régulièrement utilisée.

2. Code, moyens de calcul, simulation / méthode numérique

2.1 Code

Notre code de calcul est basé sur une méthode spectrale standard utilisant les FFT grâce aux librairies FFTW et P3DFFT (librairie scalabilisée jusqu'à 65536 coeurs). Il permet la simulation

2.2 Configuration

Pour répondre aux enjeux scientifiques, nous avons entrepris des simulations numériques à haute résolution spatiale et temporelle, dans des configurations idéalisées (boîtes plus ou moins allongées) afin de simplifier les différents phénomènes. Nous avons simulé les équations de Navier-Stokes d'un milieu fluide stablement stratifié soumis à un écoulement turbulent. Ces simulations ont été entrepris à très hautes résolutions spatiales et sur des temps longs afin de réaliser une analyse spatio-temporelle et récupérer la part des ondes dans le magma turbulent. Ces deux étapes de calculs

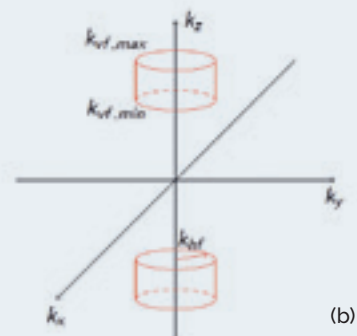
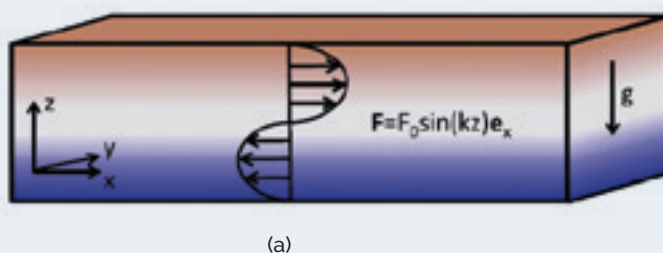


Figure 2 : a) écoulement de Kolmogorov (schéma issue de [4])
b) forçage stochastique dans l'espace de Fourier concentré sur la surface des cylindres.

demandent d'importantes ressources de calcul et de stockage. Pour forcer l'écoulement à être dans un état turbulent statistiquement stationnaire, nous avons utilisé deux types de forçage :

- un écoulement de Kolmogorov jusqu'à 512x2048x512 points de grille.
- un écoulement soumis à un nouveau forçage stochastique jusqu'à 2048x2048x2048 points de grille.

D'une part, l'écoulement de Kolmogorov a été introduit par Kolmogorov vers les années 1960 comme modèle pour étudier le passage à la turbulence dans les fluides newtoniens. Il est constitué d'un écoulement parallèle à profil sinusoïdal (voir figure 2 a). Cet écoulement est perpendiculaire à la gravité et produit un fort cisaillement capable de déstabiliser le champ de densité et ainsi de produire des ondes internes. D'autre part, l'écoulement soumis à un forçage stochastique a également été mis en place, il consiste à injecter continuellement de l'énergie dans des modes de Fourier spatiaux à grande échelle mais de manière à ne pas nourrir des modes « parasites » qui pompent l'énergie sans engendrer d'état turbulent (voir figure 2 b). La difficulté est de bien régler les paramètres des simulations : il faut à la fois obtenir un important effet turbulent tout en ayant une forte stratification.

2.3 Moyens de calcul, simulation et performance

Dans le contexte de ces Grands Challenges, nous avons opté pour un découpage sur une grille cartésienne M1xM2 :

- pour un écoulement de Kolmogorov : 64x32=2048 processus
- pour un écoulement soumis à un nouveau forçage stochastique : 64x64=4096 processus

Typiquement, les calculs ont été effectués sur 2048 coeurs (soit 86 nœuds de calcul Skylake) avec un maximum de 4096 coeurs (soit 171 nœuds de calcul Skylake). La scalabilité du coeur du code (P3DFFT) a été validée jusqu'à 8192 coeurs sur Joliot-Curie.

2.4 Environnement technique

Le support applicatif du TGCC a été très réactif, il a rapidement installé la librairie P3DFFT avec nos options de compilation. Nous avons affronté un certain nombre de problèmes techniques (non démarrage du job ou calcul très lent). Après de nombreux échanges, nous avons eu une amélioration des performances qui a permis la production de nos simulations

3. Observations et résultats scientifiques

Une illustration de la perturbation du champ de densité de l'écoulement de Kolmogorov est donnée sur la figure 3. Nous distinguons l'évolution de la perturbation au cours du temps. Elle s'organise par couche et on distingue l'évolution turbulente des instabilités primaire et secondaire à grande échelle. Dans le cas du forçage stochastique, nous illustrons la perturbation du champ de densité sur la figure 4. L'écoulement y apparaît moins structuré mais avec beaucoup plus de tourbillons.

Lorsque l'écoulement devient statistiquement stationnaire, il est alors possible d'effectuer une analyse spatio-temporelle. Sur la figure 5 a) nous avons reporté l'énergie potentielle obtenue, en fonction de la fréquence et de l'angle theta des nombres d'onde. Nous observons que cette

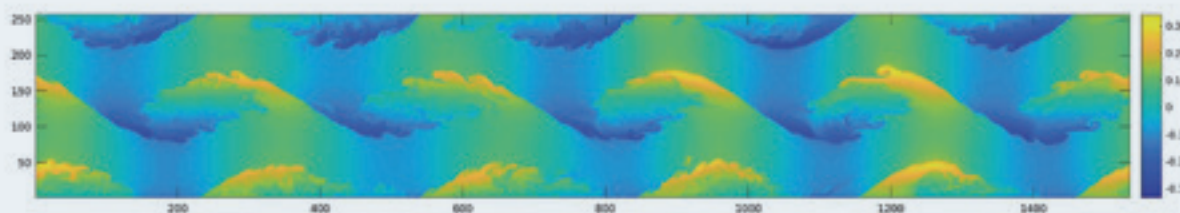


Figure 3 : Coupe verticale de la perturbation de la densité dans le cas d'un écoulement de Kolmogorov : il représente la déstabilisation du champ de densité avec les instabilités primaire et secondaire.



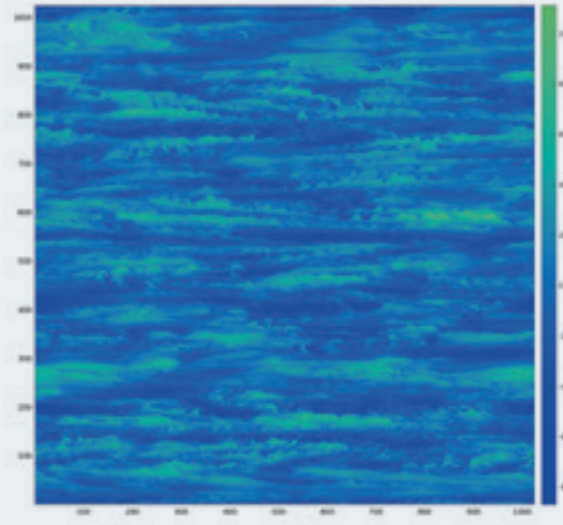


Figure 4 : Coupe verticale de la perturbation de la densité dans le cas d'un écoulement forcé stochastiquement.

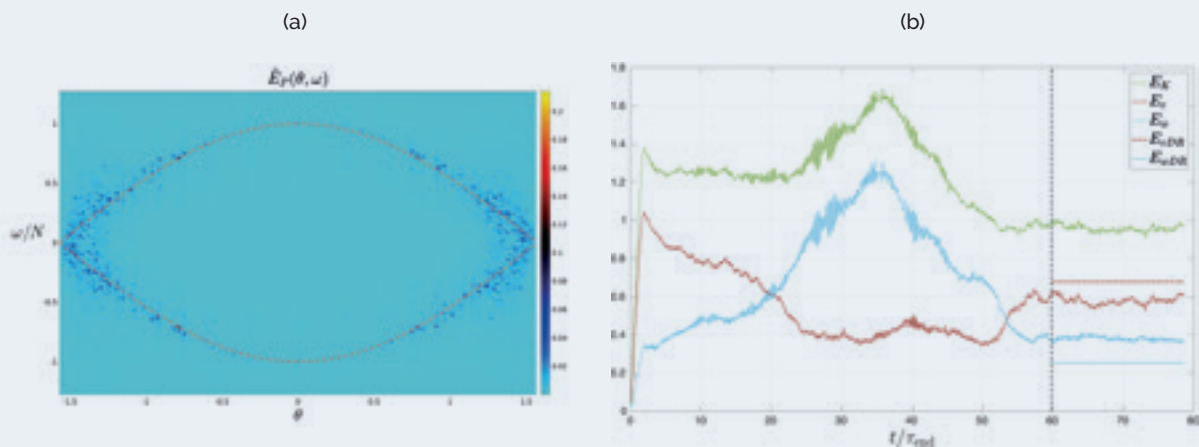


Figure 5 : a) énergie potentielle en fonction de la fréquence et de l'angle du nombre d'onde par rapport à l'horizontal (point) et la relation de dispersion idéale.
b) évolution temporelle des différentes formes d'énergie (v=vortex, w=wave).

énergie potentielle s'organise autour de la relation de dispersion alors que cet écoulement est turbulent. Cette relation a déjà été observée pour des écoulements stratifiés [3] mais pas lorsque l'écoulement est pleinement turbulent, contrairement à nos simulations. Cela indique la robustesse de cette relation de dispersion.

D'autre part sur la figure 5 b) l'évolution en temps de l'énergie cinétique totale est décomposée en une partie onde (w) et vortex (v) suivant la décomposition de Riley et al. [2] (trait plein). Par ailleurs, en effectuant une analyse

spatio-temporelle sur l'énergie cinétique, il est possible de filtrer l'énergie cinétique autour de la relation de dispersion : on extrait alors la part de l'énergie cinétique dans les ondes de la part de l'énergie cinétique dans les vortex (trait en pointillé, constant en temps). On observe alors que les ondes contiennent au moins 30 % de l'énergie totale et qu'avec la décomposition de Riley et al. [2], il y a une sur-estimation de 50 % de l'énergie dans les ondes..



■ 4. Perspectives

La robustesse de la relation de dispersion dans le cas d'une turbulence en fluide stratifié, nous permet d'envisager le même type d'analyse mais dans le cas d'un fluide soumis à la rotation qui développe des ondes d'inertie au lieu d'ondes internes. D'autre part, nos simulations montrent qu'il y a moins d'énergie dans les ondes que prévue par les décompositions classiques. Nous espérons que ces résultats physiques pourront être pris en compte dans les futurs modèles de l'océan ou de l'atmosphère.

■ 5. Conclusions

Les simulations sur Joliot-Curie ont été cruciales pour mettre en évidence une nouvelle partition énergétique entre ondes et vortex dans la cadre de turbulence stratifié par rapport aux analyses classiques.

■ 6. Remerciements

Nous remercions les organismes qui ont apporté leur soutien au projet (GENCI et FLMSN). Ce travail a été réalisé grâce au soutien financier du LABEX iMUST (ANR-10-LABX-0064) de l'Université de Lyon, dans le cadre du programme «Investissements d'Avenir» (ANR-11-IDEX-0007) géré par l'Agence Nationale de la Recherche (ANR).

■ 7. Bibliographie

- [1] Ozmidov, R. V. On the turbulent exchange in a stably stratified ocean. *Izv. Acad. Sci. USSR. Atmos. Oceanic Phys.*, 1965, vol. 1, p. 861-871.
- [2] Riley, J. J., Metcalfe, R. W., & Weissman, M. A. "Direct numerical simulations of homogeneous turbulence in density stratified fluids" In *AIP Conference Proceedings* (Vol. 76, No. 1, pp. 79-112). AIP.(1981), December
- [3] Le Reun T., Favier B. & Le Bars M., Parametric instability and wave turbulence driven by tidal excitation of internal waves, *J. Fluid Mech.*, in press (2018).
- [4] Garaud P., Gallet B. & Bischoff T., The stability of stratified spatially periodic shear flows at low Péclet number, *Phys. Fluids*, Vol. 27, 084104, 2015.

■ Équipe projet



Alexandre
DELACHE



Andrea
MAFFIOLI

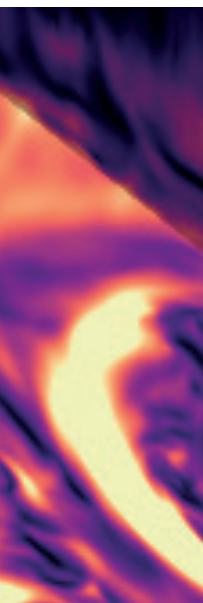


Équipe

Nathanaël SCHAEFFER, *Chargé de Recherche CNRS, ISTerre, Univ. Grenoble Alpes.*

Ont contribué au projet

Alexandre Fournier (IPGP)
Thomas Gastine (IPGP)



Simuler les inversions du champ magnétique terrestre

23 millions d'heures

Au cours de son histoire, les pôles magnétiques de la Terre se sont inversés à de nombreuses reprises, le nord devenant le sud et vice-versa. Ces inversions du champ magnétique terrestre restent mystérieuses, bien que des simulations du noyau terrestre existent depuis 1995. **D'un côté nous avons des simulations numériques de géodynamo qui produisent des inversions, mais qui opèrent avec des viscosités trop fortes et des champs magnétiques trop faibles pour être pertinentes pour le cas terrestre.** D'un autre côté nous avons des simulations beaucoup plus réalistes du noyau terrestre et de son champ magnétique, en particulier en terme de viscosité et d'intensité du champ magnétique, mais celles-ci ne s'inversent pas car il est impossible de faire des simulations assez longues. L'opportunité d'un **Grand Challenge sur Joliot-Curie nous a permis de réaliser des simulations plus réalistes et qui présentent des inversions.** La compréhension théorique de ces dernières va s'en trouver transformée, et nous pourrions à terme prévoir les conséquences d'une inversion réaliste sur la morphologie du champ magnétique et sur notre société.

1. Contexte

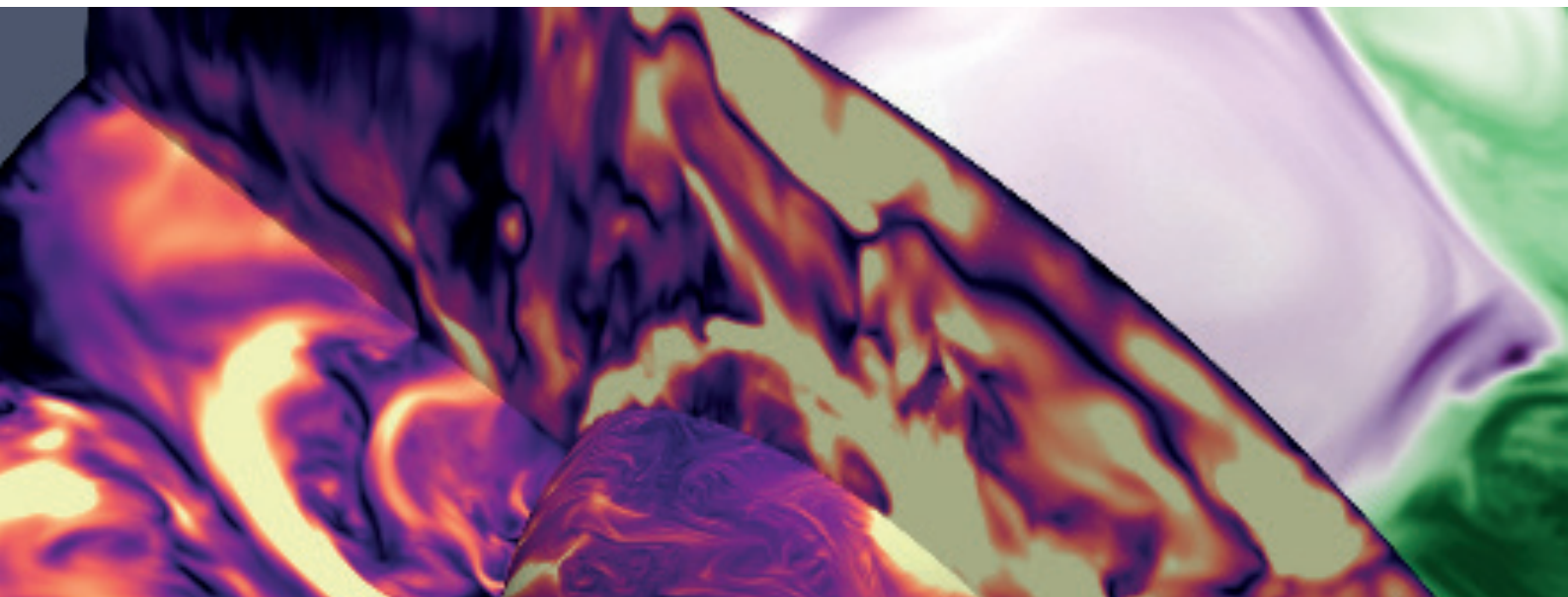
1.1 Les inversions du champ magnétique terrestre

Au cours de son histoire, les pôles magnétiques de la Terre se sont inversés à de nombreuses reprises, le nord devenant le sud et vice-versa. Ces inversions du champ magnétique terrestre restent mystérieuses. Elles se produisent de manière irrégulière, avec un champ magnétique stable pendant 40 millions

d'années, suivi de renversements espacés en moyenne d'environ 250 000 ans depuis 25 millions d'années (voir Illustration 1).

La simulation réaliste et la compréhension de ces phénomènes extraordinaires permettraient de contraindre davantage les paramètres physiques du noyau terrestre, et sont une étape clé pour une éventuelle prévision de tels phénomènes et de leurs implications qui reste pour l'instant basée sur des modèles dynamiques de basse dimensionnalité.





1.2 Simulations existantes d'inversion du champ magnétique

Dès 1995, une simulation du noyau terrestre a produit une inversion de champ magnétique [Glatzmaier & Roberts 1995]. Depuis, des études systématiques de paramètres [Kutzner & Christensen 2002] ou des simulations individuelles [Takahashi, Matsushima & Honkura 2007; Sheiko, Finlay & Jackson 2016; par exemple] ont également produit des inversions de champ magnétique.

Cependant, toutes ces simulations opèrent dans des régimes de paramètres peu réalistes, soit par excès de viscosité, soit à cause d'un champ magnétique beaucoup trop faible comparé à celui de la Terre. D'un autre côté les simulations actuelles les plus réalistes [Schaeffer, Jault, Nataf & Fournier 2017] ne produisent pas d'inversion par manque de temps de simulation.

D'autre part, et c'est fondamental, les lois d'échelles qui ont été proposées à partir des inversions de champ magnétique obtenues dans les simulations sont incompatibles avec ce que nous connaissons du noyau terrestre et de sa physique. En effet, les simulations suggèrent que pour obtenir des inversions, les vitesses dans le noyau doivent être supérieures de plusieurs

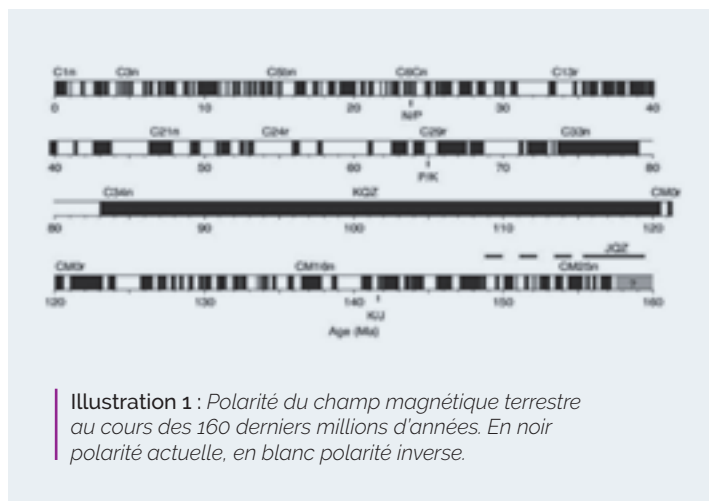


Illustration 1 : Polarité du champ magnétique terrestre au cours des 160 derniers millions d'années. En noir polarité actuelle, en blanc polarité inverse.

ordres de grandeur par rapport à ce qui est déduit des variations du champ magnétique.

1.3 Objectifs scientifiques

Par ce Grand Challenge, nous avons souhaité pousser les simulations produisant des champs magnétiques vers des régimes plus réalistes pour comprendre ce qui produit une inversion de champ magnétique, et en prévoir les conséquences.



2. Le code XSHELLS pour simuler le noyau terrestre

2.1 XSHELLS, un code haute performance pour les noyaux planétaires.

Le code XSHELLS¹ que nous avons développé, simule en 3D dans une sphère le champ magnétique et l'écoulement généré par de la convection thermique [Schaeffer et al. 2017], voir illustration 2.

XSHELLS discrétise les équations de Navier-Stokes via des différences finies dans la direction radiale, et utilise les harmoniques sphériques dans les autres directions. Cette approche reste à ce jour la plus efficace en géométrie sphérique [Matsui et al. 2016].

La transformée en harmonique sphérique reste la partie la plus gourmande en temps de calcul. Elle est effectuée par la très performante bibliothèque SHTns [Schaeffer 2013] qui a été optimisée avec succès pour les instructions AVX-512 du processeurs Intel Skylake ou Intel KNL. La parallélisation utilise à la fois MPI et OpenMP.

2.2 Moyens de calcul et simulation.

Un total de 23 millions d'heures de calcul ont été utilisées sur plusieurs simulations pour comprendre les inversions du champ magnétique terrestre. Afin de simuler des temps géologiques (plusieurs millions d'années au moins), nous devons limiter la résolution de nos simulations à environ 50 millions de points.

2.3 Performance .

Nous avons réalisé une étude de performance en fonction du nombre de noeuds utilisés sur les deux partitions de la machine (KNL et Skylake). Les performances initiales étaient décevantes. Nous avons travaillé sur les points durs et identifiés que la parallélisation OpenMP ne se comportait pas bien sur KNL. Nous avons donc implémenté une parallélisation OpenMP différente, basés sur les tâches («OpenMP tasks») et qui permettait d'obtenir de bien meilleures performances sur KNL (mais pas sur Skylake).

Finalement, nous avons obtenu des performances satisfaisantes sur KNL, bien que le scaling ne soit pas optimal, mais plutôt décevantes sur Skylake, à cette résolution modérée (voir illustration 3). Nous avons également constaté que c'est avec 4 processeurs

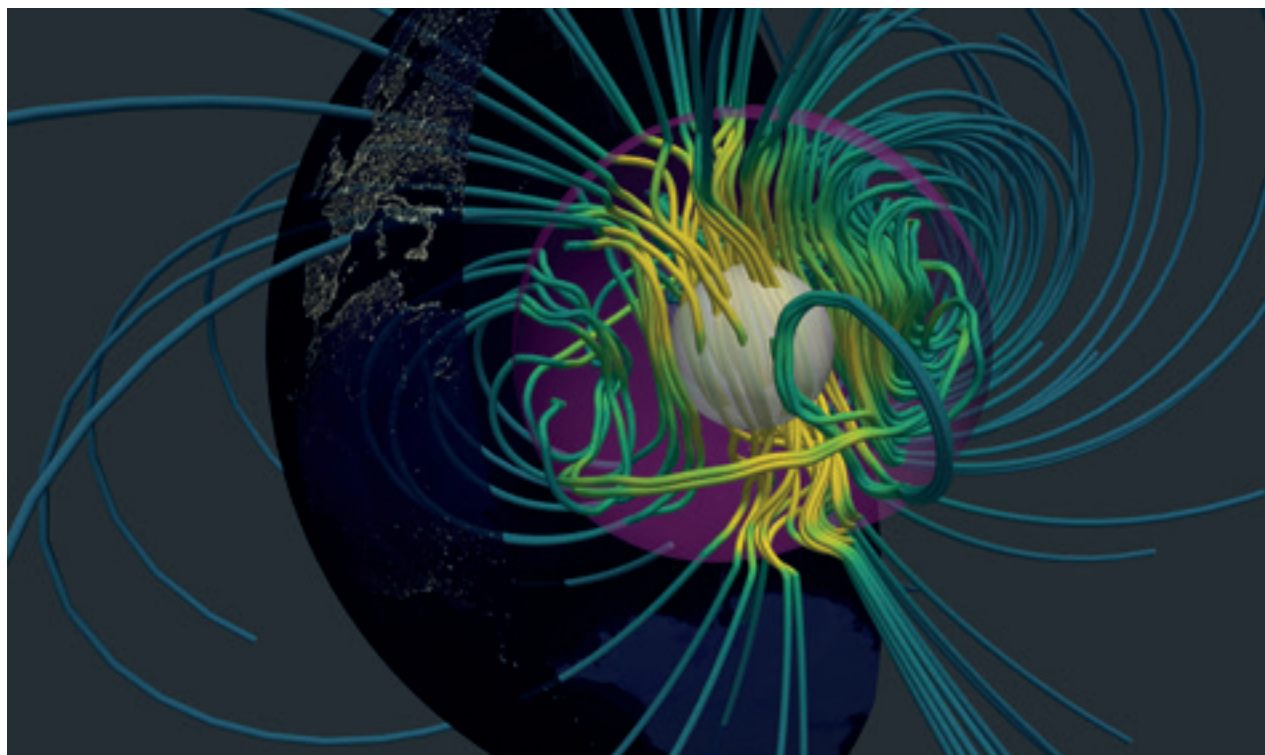


Illustration 2 : Exemple de champ magnétique obtenu par une simulation XSHELLS du noyau terrestre.



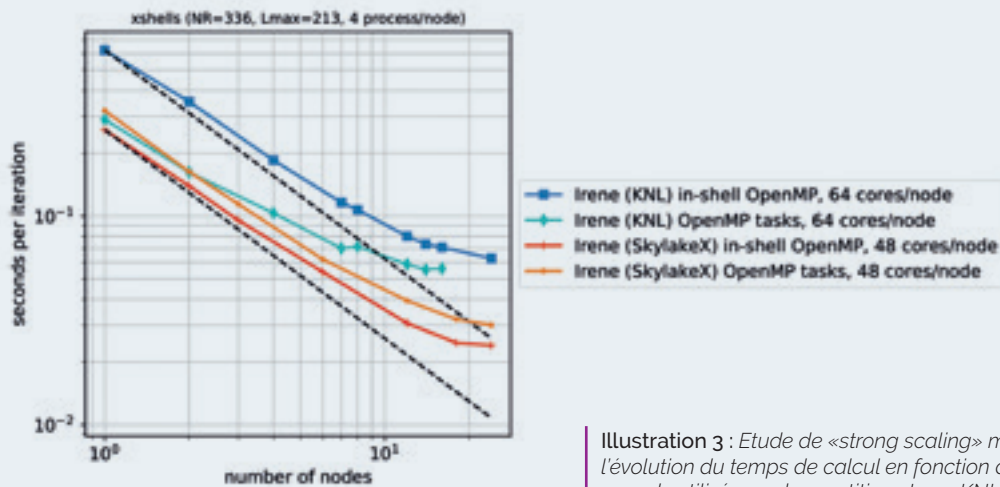


Illustration 3 : Etude de «strong scaling» montrant l'évolution du temps de calcul en fonction du nombre de noeuds utilisés, sur les partitions Irene KNL et Skylake.

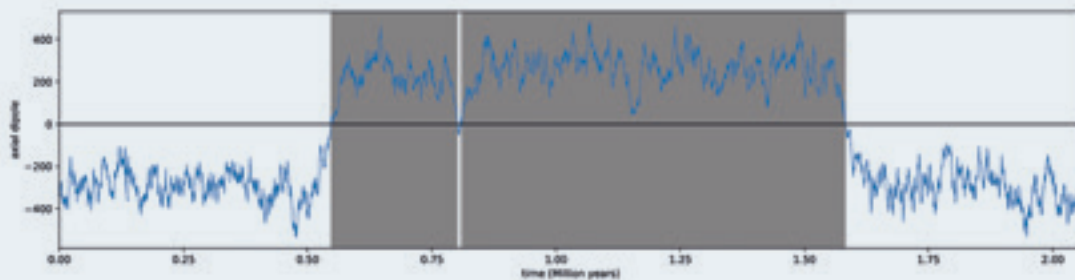


Illustration 4 : Intensité et direction du dipôle axial (positif vers le nord, négatif vers le sud) dans notre simulations phare.

MPI par noeud que le code était le plus performant, que ce soit sur Skylake ou sur KNL.

Pour la simulation phare de ce Grand Challenge, nous avons donc surtout utilisé la partition KNL, avec 7 noeuds (soit 448 coeurs effectifs) pendant plus de 3 mois.

3. Observations et résultats scientifiques

Pour la première fois, nous avons observé des inversions de champ magnétique dans un régime de champ magnétique fort, c'est à dire que l'énergie magnétique du système est supérieure à son énergie cinétique. De plus, ces inversions

n'obéissent plus aux mêmes lois que les inversions classiques en champ faible. Ce n'est plus le nombre de Rossby local (une mesure de l'effet de la force de Coriolis sur les mouvements de petite échelle), mais le nombre de Reynolds magnétique qui semble contrôler les inversions de champ magnétique et sa morphologie. C'est un **résultat important que nous avons pu obtenir grâce à ce Grand Challenge.**

Notre simulation phare possède une énergie magnétique plus de 3 fois supérieure à l'énergie cinétique. Ceci implique une influence majeure des forces magnétiques sur la dynamique du système. En un peu plus de 2 millions d'années simulées, nous observons 2 inversions de champ magnétique, et 1 évènement qu'on appelle « excursion » au cours duquel la partie dipolaire du champ magnétique s'effondre (voir Illustrations 4 et 5).



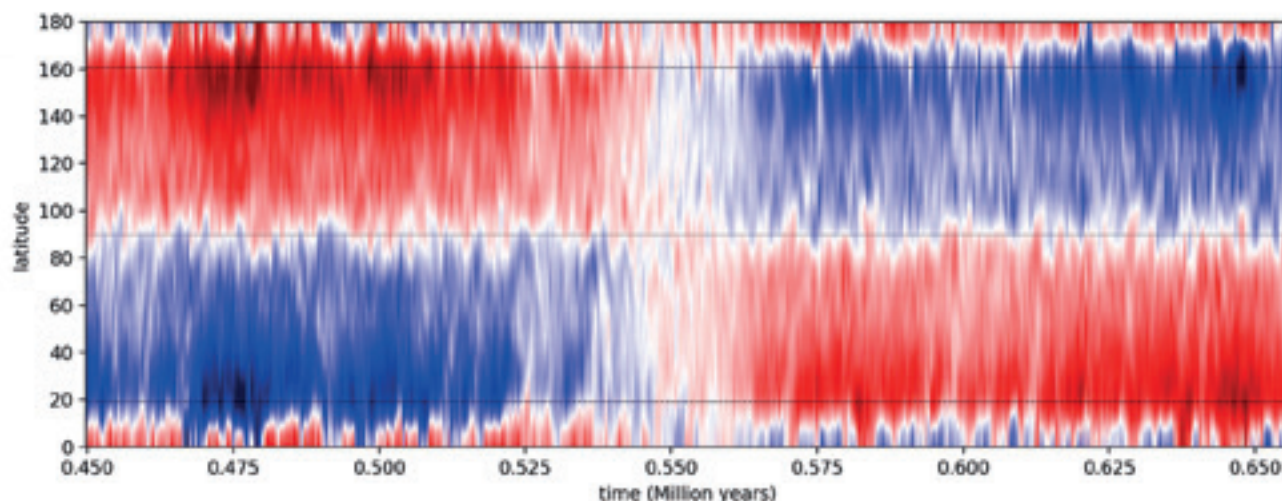


Illustration 5 : Intensité et direction du dipôle axial (positif vers le nord, négatif vers le sud) dans notre simulations phare.

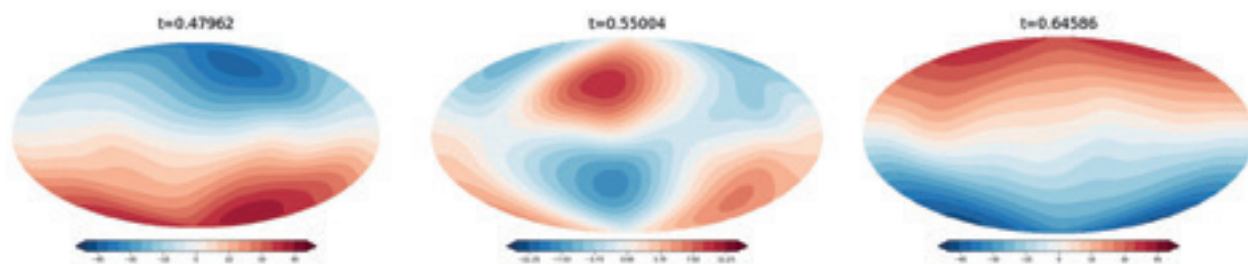


Illustration 6 : Intensité et direction du dipôle axial (positif vers le nord, négatif vers le sud) dans notre simulations phare.

■ 4. Perspectives

Les simulations vont être comparées finement aux données géomagnétiques. Nous allons utiliser les simulations pour ajuster les modèles de renversements existants, car les données paléomagnétiques sont incomplètes pour ce faire. Il nous faudra aussi essayer de décortiquer le mécanisme d'inversion, en examinant les champs magnétiques et cinétiques. Les observations préliminaires des simulations semblent indiquer que ces renversements seraient des événements extrêmes au sens statistique, c'est-à-dire qu'ils ont une probabilité faible d'arriver mais restent décrits par la même physique que les fluctuations plus faibles.

Nous allons également essayer de répondre à la question de l'existence de précurseurs aux inversions de champ magnétique. Enfin, nous pourrions documenter précisément la morphologie du champ magnétique lors d'une inversion, et évaluer l'impact sur notre société, en particulier sur la vulnérabilité des satellites.

■ 5. Conclusions

Les simulations sur Joliot-Curie ont permis d'explorer un régime de paramètres inédits pour les inversions de champ magnétique : celui des champs magnétiques forts. La puissance de calcul de la machine, couplée à des optimisations importantes du code de calcul ont permis cette prouesse.

Jusqu'à présent, les simulations du noyau terrestre produisaient soit des champs magnétiques forts, soit des inversions mais jamais les deux ensembles.

Nous avons donc démontré que les inversions peuvent survenir dans le régime à champ magnétique fort. De plus, le critère classique d'inversion indépendant du champ magnétique doit être revu pour tenir compte de ce nouveau régime pour lequel il ne s'applique pas.



■ 6. Bibliographie

Glatzmaier, G. and Roberts, P. (1995). *A three-dimensional self-consistent computer simulation of a geomagnetic field reversal*, Nature 377 : [203-209].

Kutzner, C. and Christensen, U. (2002). *From stable dipolar towards reversing numerical dynamos*, Physics of the Earth and Planetary Interiors 131 : 29-45.

Matsui, H.; Heien, E.; Aubert, J.; Aurnou, J. M.; Avery, M.; Brown, B.; Buffett, B.; Busse, F.; Christensen, U.; Davies, C.; Featherstone, N.; Gastine, T.; Glatzmaier, G.; Gubbins, D.; Guermond, J.-L.; Hayashi, Y.-Y.; Hollerbach, R.; Hwang, L.; Jackson, A.; Jones, C.; Jiang, W.; Kellogg, L.; Kuang, W.; Landeau, M.; Marti, P.; Olson, P.; Ribeiro, A.; Sasaki, Y.; Schaeffer, N.; Simitev, R.; Sheyko, A.; Silva, L.; Stanley, S.; Takahashi, F.; Takehiro, S.; Wicht, J. and Willis, A. (2016). *Performance benchmarks for a next generation numerical dynamo model*, Geochemistry, Geophysics, Geosystems 17 : 1586-1607.

Schaeffer, N. (2013). *Efficient spherical harmonic transforms aimed at pseudospectral numerical simulations*, Geochemistry, Geophysics, Geosystems 14 : 751-758.

Schaeffer, N.; Jault, D.; Nataf, H.-C. and Fournier, A. (2017). *Turbulent geodynamo simulations: a leap towards Earth's core*, Geophysical Journal International 211 : 1-29.

Sheyko, A.; Finlay, C. C. and Jackson, A. (2016). *Magnetic reversals from planetary dynamo waves*, Nature 539 : 551-554.

Takahashi, F.; Matsushima, M. and Honkura, Y. (2007). *A numerical study on magnetic polarity transition in an MHD dynamo model*, Earth Planets Space 59 : 665-673.

■ Équipe projet



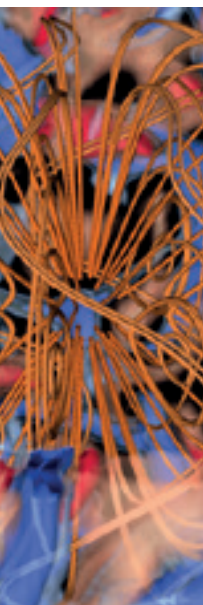
Nathanaël
SCHAEFFER



Équipe

Allan Sacha BRUN, Astrophysicien et Chef
du LDE3 au Département d'Astrophysique,
CEA Paris-Saclay

Antoine STRUGAREK, Astrophysicien au
Département d'Astrophysique, CEA Paris-Saclay



Les mystères de la tachocline solaire

15 millions d'heures

L'objectif du grand challenge est de modéliser le magnétisme et la dynamique turbulente solaire interne, notamment afin de mieux comprendre comme se forme une couche essentielle en son sein: la tachocline. L'enjeu est ici de comprendre la dynamique et la stabilité de cette couche mince (moins de 5% du rayon solaire) et très turbulente (grande gamme d'échelles spatiales et temporelles mises en jeu) qui agit comme une barrière de transport de moment cinétique dans l'intérieur solaire et est un ingrédient central de nombreux modèles du magnétisme cyclique de notre étoile. Ces questions sont aujourd'hui dans le contexte du lancement en 2020 de la sonde de l'ESA Solar Orbiter qui cherche à comprendre l'origine du magnétisme solaire et son impact sur son environnement. En effet, le magnétisme du Soleil impacte aujourd'hui notre société technologique, notamment par ses manifestations éruptives très intenses qui peuvent endommager les satellites en orbite autour de la Terre.

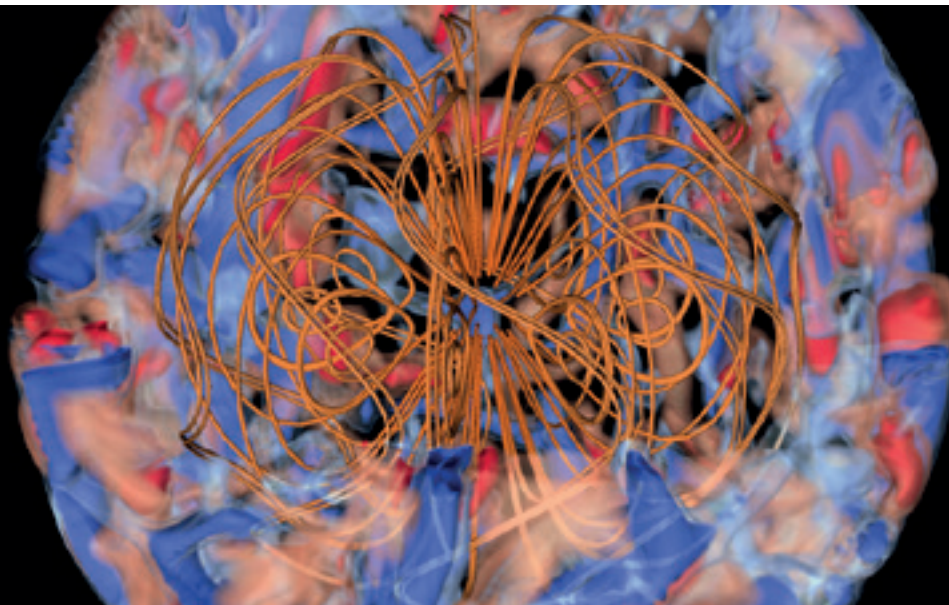
1. Contexte

1.1 La tachocline solaire, un défi de modélisation magnéto-hydro-dynamique

La tachocline (voir Fig. 1) est une coquille sphérique fortement et stablement stratifiée de l'intérieur solaire, animée de cisaillements de vitesse de grande échelle à la fois en latitude et en rayon (dans la direction de la stratification). A ces écoulements s'ajoute la pénétration de

mouvements convectifs turbulents provenant de la région convective au-dessus d'elle, qui génèrent en son sein un transport de moment cinétique mais excitent également des ondes de gravité. La tachocline est aussi couplée au champ magnétique dynamo généré dans la zone convective au-dessus d'elle, et peut-être à un hypothétique champ magnétique fossile enfoui en dessous dans la zone radiative. La complexité de cette couche mince rend sa modélisation particulièrement difficile, et nous ne





Rendu 3D de la vitesse radiale (en bleu flots descendants) et des lignes de champ magnétique (en orange) dans une simulation haute performance de l'intérieur solaire (intérieur radiatif et enveloppe convective) en utilisant le code parallèle ASH sur Joliot-Curie. Cette simulation permet de caractériser la dynamique et le double confinement de la zone de fort cisaillement à la base de la zone convective (i.e. la tachocline) où le champ magnétique cyclique du Soleil est organisé (Strugarek & Brun 2019).

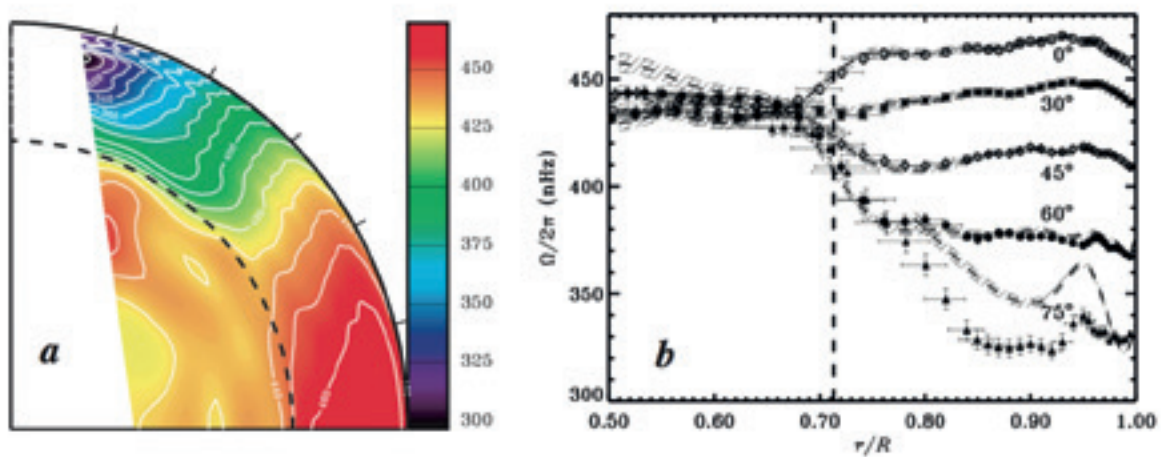


Figure 1 : Profil de rotation interne du Soleil tel que vu par l'héliosismologie [1]. A gauche, le profil de rotation est montré en coupe méridienne. On voit que notre étoile possède une rotation différentielle en latitude dans la zone convective, au-dessus du trait pointillé noir, avec un équateur rapide (rouge) et un pôle lent (bleu). La zone radiative est en rotation solide à environ 430-440 nHz, comme on peut le voir sur le panneau de droite sur les coupes 1D à différentes latitudes. La tachocline est la fine couche radiale séparant rotationnellement les deux régions et régimes de rotation (cisailé en surface, et solide en profondeur).



comprenons aujourd'hui pas comment une telle couche peut rester aussi mince dans l'intérieur de notre étoile. Sa modélisation présente de plus un intérêt majeur car le cisaillement présent dans la tachocline est un ingrédient essentiel des nombreux modèles du magnétisme cyclique du Soleil. Le magnétisme du Soleil impacte aujourd'hui notre société technologique, notamment par ses manifestations éruptives très intenses qui peuvent endommager les satellites en orbite autour de la Terre.

1.2 Objectifs scientifiques

L'objectif du grand challenge est de modéliser pour la première fois la tachocline solaire dans un régime de paramètres magnétohydrodynamiques réalistes en 3D, sous l'influence de la rotation et de la turbulence convective à très grand nombre de Reynolds et faible nombre de Prandtl. L'enjeu est ici de comprendre la dynamique et la stabilité de cette couche mince qui agit comme une barrière de transport de moment cinétique dans l'intérieur solaire. En effet, nous savons aujourd'hui grâce à l'héliosismologie (Fig. 1) que son épaisseur ne dépasse pas 5% du rayon solaire. Or, on pourrait s'attendre à l'âge du Soleil qu'elle ait diffusée pour en couvrir plus de 30% ! Nous allons ainsi ici étudier la possibilité de confiner la tachocline sur sa faible épaisseur à l'aide d'un champ magnétique d'origine fossile enfouit dans l'intérieur radiatif du Soleil [2]. Ce scénario, qui répond à la problématique fondamentale du confinement de la tachocline

[3], est particulièrement ambitieux à modéliser car il nécessite de coupler une large gamme d'échelles spatio-temporelles, allant de la turbulence MHD aux temps d'évolution séculaire de notre Soleil.

Ce grand challenge se base sur une étude préliminaire de la tachocline solaire publiée dans [5] (voir Fig. 2). Avec les progrès du code parallèle ASH et de la puissance de calcul offerte par ce grand challenge, **nous proposons de revisiter ce travail pour le placer dans un cadre beaucoup plus réaliste de très faible nombre de Prandtl (ratio de la viscosité cinématique sur la dissipation thermique, inférieur à 10^{-3} dans le Soleil), qui est un régime difficile à atteindre numériquement.**

2. Code, moyens de calcul, simulation / méthode numérique

2.1 Le code Anelastic Spherical Harmonics (ASH)

Le code global pseudo-spectral ASH est écrit en Fortran et résout dans une coquille sphérique tridimensionnelle en rotation les équations du mouvement, de continuité, d'énergie et d'induction dans l'approximation anélastique filtrant les ondes sonores. L'objectif du grand challenge est de modéliser le

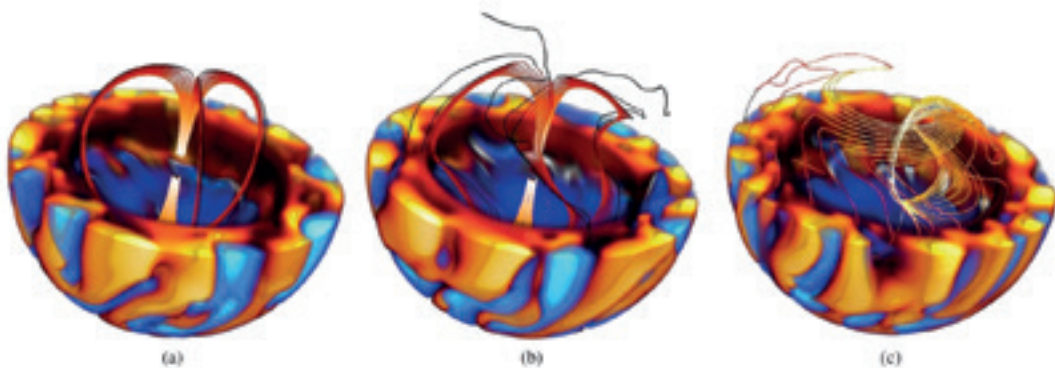


Figure 2 : Test du scénario de confinement magnétique de la tachocline solaire [5]. (a) Condition initiale du scénario : un champ magnétique dipolaire fossile (lignes orange-noir) est confiné à l'intérieur du Soleil. (b)-(c) Ce champ magnétique diffuse vers l'extérieur et interagit avec la zone convective turbulente (structures bleu/orange) en passant à travers la tachocline, rendant son existence difficile à réconcilier avec les observations du Soleil. Cette étude nécessite cependant d'être revisitée avec un modèle moins dissipatif grâce aux progrès en scalabilité du code ASH et des machines de calcul GENCI.



magnétisme et la dynamique turbulente solaire interne, notamment afin de mieux comprendre comme se forme une couche essentielle en son sein: la tachocline. L'enjeu est ici de comprendre la dynamique et la stabilité de cette couche mince (moins de 5% du rayon solaire) et très turbulente (grande gamme d'échelles spatiales et temporelles mises en jeu) qui agit comme une barrière de transport de moment cinétique dans l'intérieur solaire et est un ingrédient central de nombreux modèles du magnétisme cyclique de notre étoile. Ces questions sont aujourd'hui dans le contexte du lancement en 2020 de la sonde de l'ESA Solar Orbiter qui cherche à comprendre l'origine du magnétisme solaire et son impact sur son environnement. En effet, le magnétisme du Soleil impacte aujourd'hui notre société technologique, notamment par ses manifestations éruptives très intenses qui peuvent endommager les satellites en orbite autour de la Terre. Les variables thermodynamiques, les trois composantes du flux de masse et du champ magnétique sont projetées sur les harmoniques sphériques pour leur structure horizontale (avec N_θ points latitudinaux et $N_\phi = 2 \times N_\theta$ points longitudinaux). Une méthode de différences finies d'ordre 4 est utilisée dans la direction radiale. La discrétisation temporelle est semi-implicite, utilisant un Crank-Nicholson implicite d'ordre 2 pour les termes linéaires et un Adams-Bashforth explicite d'ordre 2 pour les termes non-linéaires, les forces de Coriolis et de Lorentz. Le code ASH utilise une approche de type LES (*Large Eddy Simulation*) avec comme traitement sous maille (SGS), soit des diffusivités effectives turbulentes ou une hyperviscosité, ainsi que la possibilité d'introduire des formulations plus complexes basées sur différentes méthodes de fermeture des équations du mouvement. La communication inter-processeur utilise MPI (Message Passing Interface), cf. [6]. Le code passe à l'échelle correctement jusqu'à 30k cœurs (cf Fig. 3). Il dispose *bien sûr* d'un système de sauvegarde/reprise. L'utilisation de routines numériques spécifiques machine maximise l'efficacité, notamment la librairie FFTW, que nous avons utilisée pour ce Grand Challenge (soit en tant que librairie indépendante, soit par la librairie Intel MKL avec le compilateur Intel). Le code peut être compilé sans perte de performances avec des compilateurs fortran gcc ou Intel. Enfin, ASH a notamment participé aux codes benchmark de l'IDRIS pour l'achat des nouvelles plateformes (Ada et Turing) en 2012.

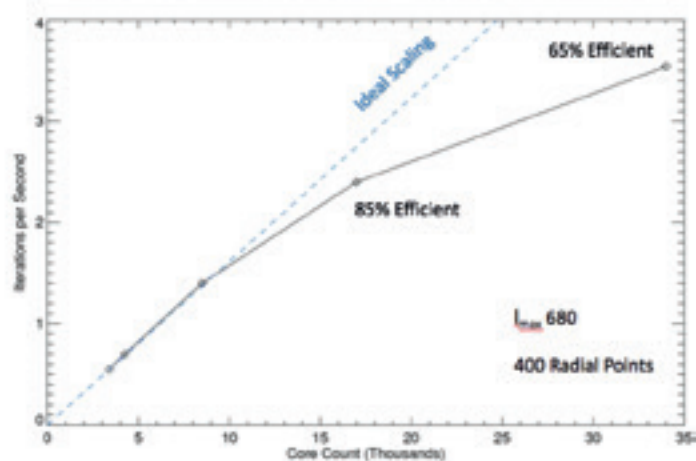


Figure 3 : Scaling du code ASH jusqu'à 30000 cœurs.

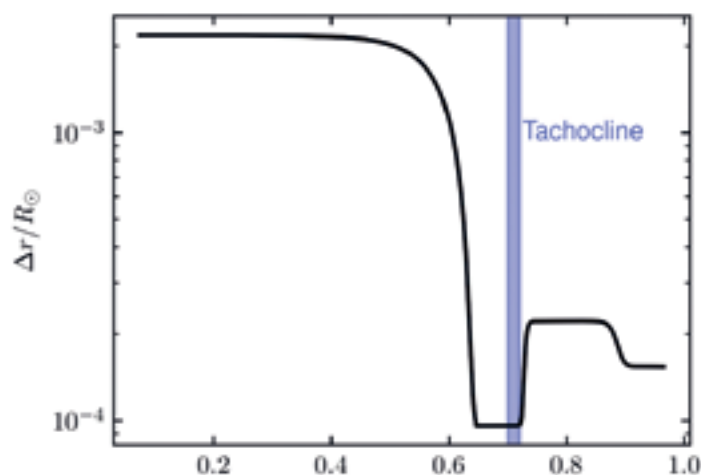


Figure 4 : Résolution radiale utilisée pour ce Grand Challenge. Quatre domaines ont été définis pour résoudre de façon adéquate la tachocline, la zone convective, et la surface, tout en conservant une résolution modérée dans le cœur radiatif de l'étoile.

2.2 Moyens de calcul, performances et simulation

Nous avons utilisé pour ce Grand Challenge la partition Skylake de la machine Joliot Curie. Nous avons développé un modèle numérique à 1.3 milliards de points de grille pour ce projet, avec une résolution de $N_r \times N_\theta \times N_\phi = 2485 \times 512 \times 1024$. Nous avons obtenu un scaling excellent jusqu'à 8192 cœurs, configuration que nous avons utilisée pour ce Grand Challenge. La répartition des processeurs s'est faite avec 204 cœurs dans la direction radiale, 40 cœurs dans la direction horizontale, et 32 cœurs pour les I/O. Aucune pression de mémoire ne s'est présentée, avec un



besoin moyen de mémoire vive d'environ 640 Go pour chaque calcul.

Nous avons construit notre projet Grand Challenge en deux étapes. En premier lieu, un modèle hydrodynamique de l'intérieur solaire a dû être développé pour traiter de façon réaliste la tachocline solaire. Pour cela, nous avons dû travailler à nombre de Prandtl très faible dans la partie basse de la tachocline et avons réussi à descendre jusqu'à $Pr = \nu/\kappa = 2.10^{-6}$ (voir section 3). Pour cela, nous avons construit une grille radiale non-uniforme afin de maximiser la résolution au niveau de la tachocline (voir Fig. 4).

Cette première étape s'est relevée bien plus ardue qu'anticipée, et a consommé la majeure partie des ressources du projet. La deuxième étape a consisté en l'introduction d'un champ magnétique enfoui dans l'intérieur de la zone radiative, afin de tester le scénario de confinement de la tachocline dans un régime dynamique réaliste (voir section 3). Après l'introduction du champ magnétique, le pas de temps de notre modèle s'est stabilisé à 64 secondes. Nous avons pu pour cette étape évoluer le modèle sur environ 535,000 itérations, correspondant à environ 500 jours d'évolution physique soit environ 16 temps caractéristiques de la turbulence. Nous avons initialement planifié d'évoluer cette étape sur environ 10^7 itérations, ce qui nécessiterait aujourd'hui environ 2 mois de plus de calculs sur 8192 cœurs sur Joliot Curie/skylake.

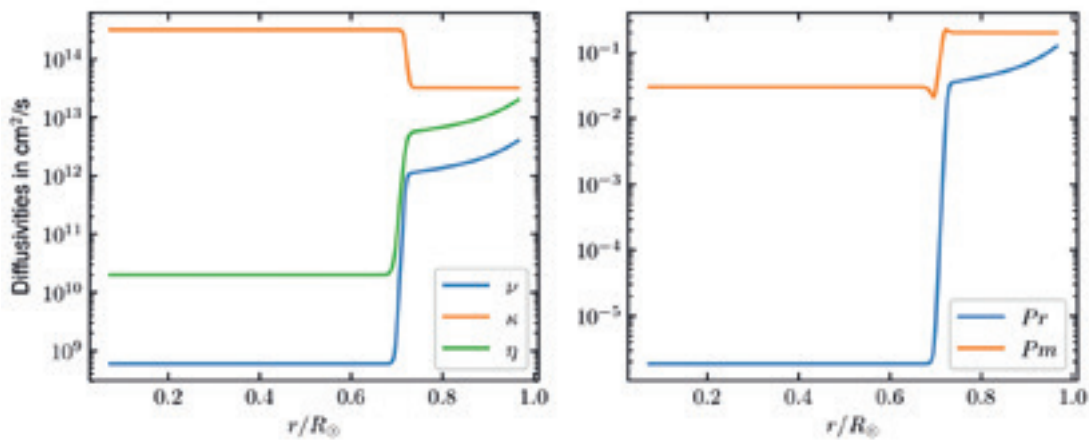
3. Observations et résultats scientifiques

3.1 Un modèle hydrodynamique réaliste de la tachocline solaire

Comprendre le Soleil en 3D dynamique est une tâche compliquée à cause de son très haut degré de turbulence et la complexité de sa structure interne. En particulier, la tachocline solaire est difficile à modéliser (voir section 1), dû au fait qu'elle est sujette à des phénomènes magneto-hydrodynamique sur des temps courts (typiquement de l'ordre du mois) qui ont un effet sur des temps longs (typiquement de l'ordre du siècle, voir plus) sur son évolution. Le paramètre essentiel pour modéliser de façon réaliste la tachocline est le paramètre σ [7] :

$$\sigma = \sqrt{Pr \frac{N^2}{\Omega^2}}$$

où Pr est le nombre de Prandtl, N la fréquence de Brunt-Väisälä, et Ω est le taux de rotation du Soleil. Ce paramètre détermine comment la rotation différentielle de la tachocline s'enfonce dans la zone radiative au cours du temps [3]. Si $\sigma > 1$, la rotation différentielle s'enfonce simplement de façon visqueuse. C'est dans ce régime que l'étude précédente [5] a été menée. Si $\sigma < 1$, qui est le régime solaire, la rotation différentielle s'enfonce en étant accompagnée d'un écoulement méridien contrôlé par la dissipation thermique du milieu. On parle alors



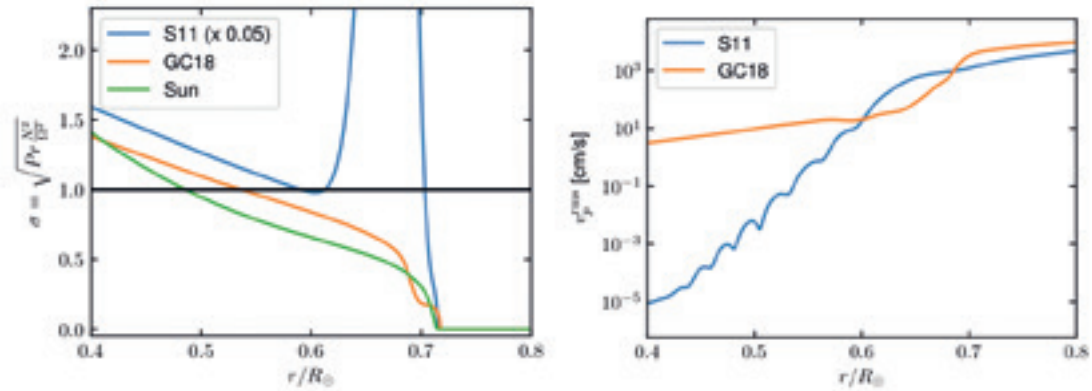


Figure 5 : Gauche : Profil de σ solaire (vert), du modèle Grand Challenge (orange), et l'ancienne étude [5] (bleu). Droite : Profils radiaux de la vitesse poloidale moyenne sous la tachocline pour le modèle [5] (en bleu) et le modèle Grand Challenge (orange).

de propagation radiative de la tachocline. Ce régime est particulièrement difficile à atteindre par la simulation numérique car il nécessite, en présence d'une stratification (N) et d'une rotation (Ω) solaire, de travailler à très faible nombre de Prandtl.

Afin d'atteindre un faible nombre de Prandtl dans la tachocline, nous avons considéré pour ce projet des profils de diffusivité qui varient avec la profondeur de l'étoile. Nous montrons les profils de viscosité (ν en bleu), dissipation thermique (κ , en orange), et dissipation ohmique (η , en vert) dans le panneau gauche de la Figure 5, ainsi que les profils radiaux des nombres de Prandtl Pr (en bleu) et Prandtl magnétique $Pm = \nu / \eta$ (en orange) dans le panneau droit. On observe qu'à la base de la zone convective, le nombre de Prandtl chute effectivement drastiquement.

Le profil de σ permis par ces profils extrêmes de dissipation est montré dans le panneau de gauche de la figure 6. Le profil solaire est indiqué en vert, tandis que le profil réalisé dans notre simulation numérique est montré en orange. On voit bien que, sous la tachocline, la quantité σ se maintient bien en dessous de 1 jusqu'à environ 0.55 rayon solaire. À l'inverse, le profil de σ utilisé dans notre étude précédente [5] (en bleu et multiplié par 0.05 pour être visible) était largement supérieur à 1 juste en dessous de la tachocline. Ce changement de régime s'accompagne, comme prévu par la théorie, par un changement de régime de l'écoulement méridien dans et sous

la tachocline. Le panneau de gauche montre une mesure de cet écoulement via la vitesse moyenne poloidale 'rms' (root-mean-square). On voit que dans notre ancien modèle, la vitesse poloidale chute fortement sous la tachocline (en bleu), tandis qu'elle se maintient à un niveau non-négligeable dans le modèle avec $\sigma < 1$. C'est cet écoulement méridien qui permet d'envisager le fonctionnement du scénario de confinement magnétique de la tachocline, comme nous allons le détailler à présent.

3.2 Vers un confinement magnétique ? —

Une fois le modèle hydrodynamique présenté en 3.2 stabilisé, la tachocline s'épaissit au-delà de 5% du rayon solaire sur un temps de diffusion thermique. Pour contrer cet épaississement, qui serait en désaccord avec ce que nous observons aujourd'hui dans le Soleil (voir Fig. 1), nous testons le scénario de confinement magnétique proposé par [2]. Ce scénario fait appel au phénomène de « double confinement ». Un champ magnétique enfouit dans l'intérieur radiatif supprime par la force de Lorentz le cisaillement de la rotation différentielle qui s'enfonce lentement dans l'étoile : premier confinement. Cependant, un tel champ magnétique diffuse naturellement vers l'extérieur ce qui met en péril le premier confinement. L'écoulement méridien mis en évidence en 3.1 intervient alors pour empêcher théoriquement ce champ de diffuser jusqu'à la zone convective où il pourrait connecter



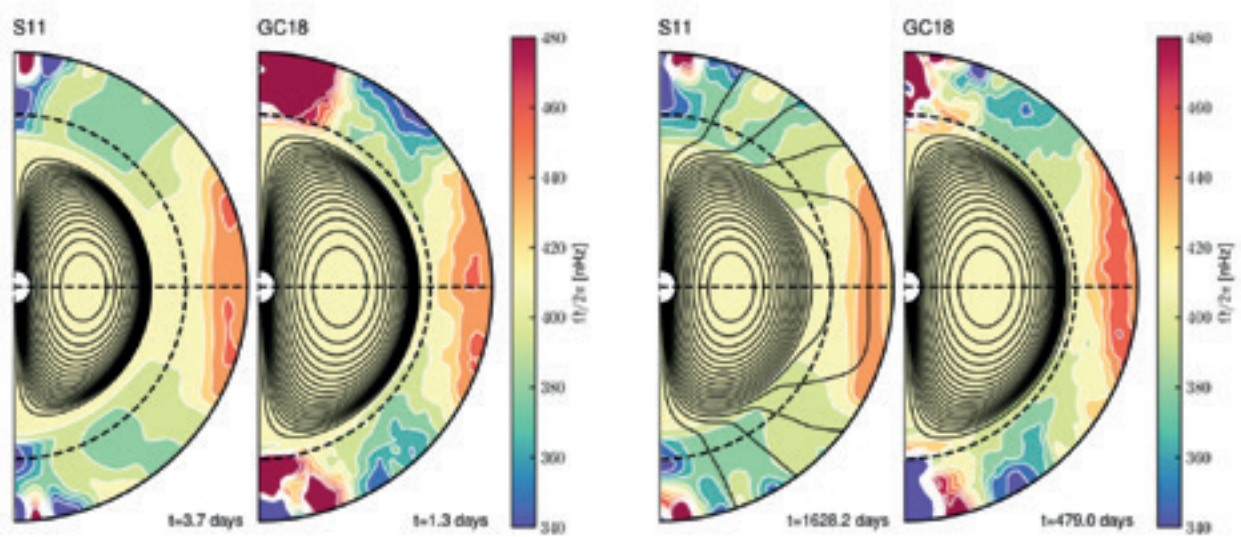


Figure 7 : Gauche : Initialisation du champ magnétique (lignes noires) enfouit dans la zone radiative solaire, pour le modèle S11 [5] à gauche et le modèle GC18 à droite. La rotation différentielle instantanée est montrée par la table de couleur dans les deux panneaux, avec une rotation rapide en rouge et lente en bleu. Droite : Évolution du champ magnétique enfouit après plusieurs temps dynamiques dans les deux modèles.

les deux zones du Soleil et fournir un moyen d'homogénéisation du profil de rotation [5] : c'est le deuxième confinement.

Nous montrons en Fig. 7 (panneaux de gauche) l'initialisation d'un tel champ magnétique fossile (lignes noires) dans nos simulations pour le modèle [5] et pour l'étude de ce Grand Challenge (GC18). Notons immédiatement que la tachocline est beaucoup moins visqueuse (et donc beaucoup moins épaisse) dans notre nouvelle étude comparée à l'étude précédente [5]. De plus, afin d'accélérer au maximum cette expérience numérique, nous avons confiné à l'instant initial le champ magnétique bien plus proche de la tachocline pour éviter de calculer la lente diffusion du champ enfouit avant son interaction avec les écoulements qui animent la zone convective supérieure. Sur ces figures nous pouvons aussi observer la rotation différentielle instantanée de nos modèles, qui montre des fluctuations temporelles substantielles par rapport au profil de rotation moyen de notre étoile (Fig. 1).

Après plusieurs temps de retournement convectif (temps dynamique caractéristique de l'ordre du mois), nous montrons les mêmes coupes méridiennes dans le panneau de droite pour des temps d'évolution équivalents dans les deux modèles. Nous observons que sur ces temps dynamiques, une tachocline réaliste

permet effectivement ce phénomène de double confinement : le lignes de champ magnétique (en noir) du modèle GC18 ne s'ouvre pas à travers la tachocline comme dans le modèle S11, et ainsi ne connectent pas les deux zones radiatives et convective. Ces résultats préliminaires (voir section suivante) sont ainsi très prometteurs et questionnent l'existence d'un tel champ magnétique au plus profond de notre étoile.

4. Perspectives

Le coût numérique de ces simulations (voir section 2) ne nous a pas permis d'explorer ce mécanisme de double confinement sur les temps séculaires (diffusifs) d'intérêt pour une étoile comme le Soleil. Bien qu'un confinement semble possible dans le modèle GC18, nous observons une dissipation lente mais cependant non-négligeable de la composante azimutale du champ depuis la base de la tachocline vers la zone convective externe. Il apparaît donc nécessaire à présent de poursuivre cette étude pour valider la possibilité de ce scénario de confinement dans notre étoile sur les échelles de temps longues. Nous comptons pour cela poursuivre cette étude avec des futures allocation GENCI/PRACE, ce qui nous permettra



de proposer pour la première fois l'étude de l'intérieur solaire et de sa fine tachocline dans toute sa globalité en 3D, dans un régime de paramètres solaire. Cette étude est notamment essentielle dans le contexte du lancement de la sonde Solar Orbiter en février 2020.

■ 5. Conclusions

Nous avons développé pour ce grand challenge un modèle ambitieux, global et réaliste de l'intérieur solaire avec une attention particulière à la tachocline, couche à la minceur inexplicée de l'intérieur solaire. Nous avons poussé le code ASH dans ses limites pour atteindre les nombres de Prandtl extrêmement faibles de la zone radiative du Soleil. **Le calculateur Joliot-Curie via ce Grand Challenge a été crucial pour mettre en place une telle simulation**, fortement coûteuse en ressources numériques tant par le grand nombre de points de mailles nécessaires (13 milliards) que par la longue évolution temporelle requise pour évaluer l'épaississement (ou le confinement) lent de la tachocline. Les résultats préliminaires très prometteurs présentés ici doivent maintenant être poursuivis pour percer définitivement le mystère du magnétisme solaire interne et du rôle de la tachocline dans l'organisation du cycle de 11 ans.

■ 6. Bibliographie

[1] Thompson M., Christensen-Dalsgaard J., Miesch M., Toomre J., Annual Review of Astronomy & Astrophysics 2003, 41, 599 : The internal rotation of the Sun.

[2] Gough D. & McIntyre M., Nature 1998, 394, 755 : Inevitability of a magnetic field in the Sun's radiative interior

[3] Spiegel E. A. & Zahn J.-P., Astronomy & Astrophysics 1992, 265, 106 : The solar tachocline

[4] Barnabé R., Strugarek A., Charbonneau P., Brun A.S., Zahn J.P., Astronomy & Astrophysics 2017, 601 : Confinement of the solar tachocline by a cyclic magnetic field

[5] Strugarek A., Brun A.S., Zahn J.P., Astronomy and Astrophysics 2011, 532, A34 : Magnetic confinement of the solar tachocline : II. Coupling to a convection zone

[6] Clune T. C., Elliott J., Miesch M., Toomre J., Glatzmaier G. A., Parallel Computing 1999, 25, 4, 361-380 : Computational aspects of a code to study rotating turbulent convection in spherical shells

[7] Wood T. & Garaud P., The Astrophysical Journal 2011, 738, 1, 47 : The Sun's meridional circulation and interior magnetic field

■ Glossaire

• **Nombre de Reynold [Re]** : Nombre sans dimension mesurant l'importance relative des forces d'inertie d'un écoulement par rapport aux forces visqueuses.

• **Nombre de Prandtl [Pr]** : Nombre sans dimension mesurant l'importance relative de la dissipation thermique d'un milieu par rapport aux forces visqueuses.

• **Fréquence de Brünt-Väisälä [N]** : Dans un milieu stablement stratifié, c'est la fréquence à laquelle une particule fluide déplacée perpendiculairement à la stratification oscille autour de son point d'équilibre.

■ Équipe projet



Allan Sacha
BRUN



Antoine
STRUGAREK



Équipe

Mathieu LOBET, Ingénieur de recherche CEA, Maison de la Simulation, Saclay
 Julien DEROUILLAT, Ingénieur de recherche CEA, Maison de la Simulation, Saclay
 Arnaud BECK, Ingénieur de recherche CNRS,
 Laboratoire Leprince-Ringuet, Ecole polytechnique
 Francesco MASSIMO, Post doctorant,
 Laboratoire Leprince-Ringuet, Ecole polytechnique

Simulation d'accélération d'électrons en configuration double-étage dans le cadre du laser CILEX-APOLLON

10 millions d'heures

Ce grand challenge avait pour but scientifique la **la simulation de l'accélération d'électrons par un laser dans une configuration dite à double-étage grâce au code Smilei. Smilei est un code Particle-In-Cell massivement parallèle, ouvert et collaboratif** qui fait l'objet d'une politique de développement soutenue de la part du LLR, du LULI et de la MdLS. Nous avons découpé ce grand challenge suivant deux aspects que sont la physique et l'informatique. D'un point de vue physique, nous avons pour objectif de démontrer la faisabilité d'une étude 3D complètement résolue pour une configuration multi-étage, de mettre en lumière, le cas échéant, les différences de propagation laser et d'accélération avec la méthode d'enveloppe et de participer au **dimensionnement et la compréhension des futures expériences sur le laser APOLLON**. D'un point de vue développement, nous voulions finaliser l'optimisation générale du code, étudier les performances sur un grand nombre de nœuds de calcul CPU de dernière génération. Dans ce contexte et fort d'une expérience solide avec les grands centres nationaux français, ce Grand Challenge sur le système de dernière génération Joliot-Curie équipé de processeurs Intel Skylake nous est apparu comme une **opportunité de finaliser ces projets scientifiques et informatiques**.

1. Contexte

1.1 Accélération d'électrons par laser dans une configuration double-étage

L'accélération d'électrons par sillage laser [3] consiste à créer une onde plasma dans le sillage d'un laser se propageant dans un gaz et à injecter des électrons dans cette onde. Différentes techniques d'injection existent mais, dans tous les cas, les électrons injectés dans l'onde plasma

vont, à la manière de surfeurs, pouvoir transférer une partie de l'énergie de celle-ci vers leur propre énergie cinétique. Ce faisant, les électrons atteignent quasi-instantanément des énergies relativistes et le paquet d'électrons injectés devient un faisceau d'électrons relativistes. Ce phénomène est illustré par les simulations effectuées dans la section 2.

La technique de l'accélération à deux étages consiste à guider ce faisceau d'électrons, appelé "injecteur", dans le sillage d'un second laser



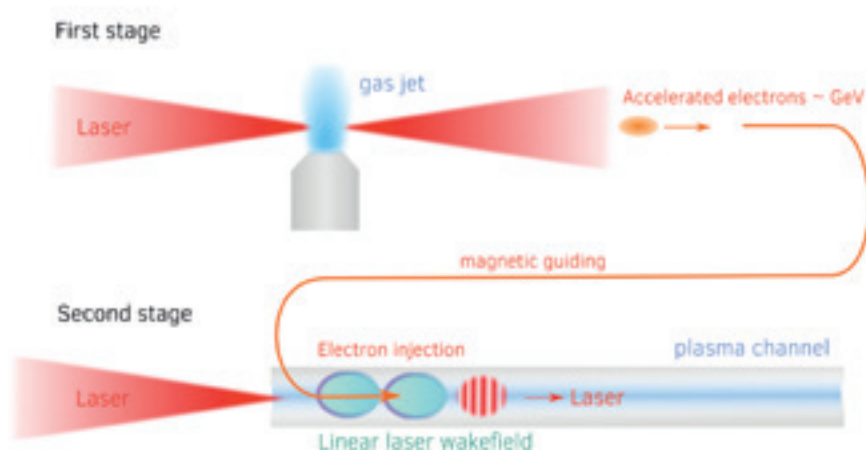
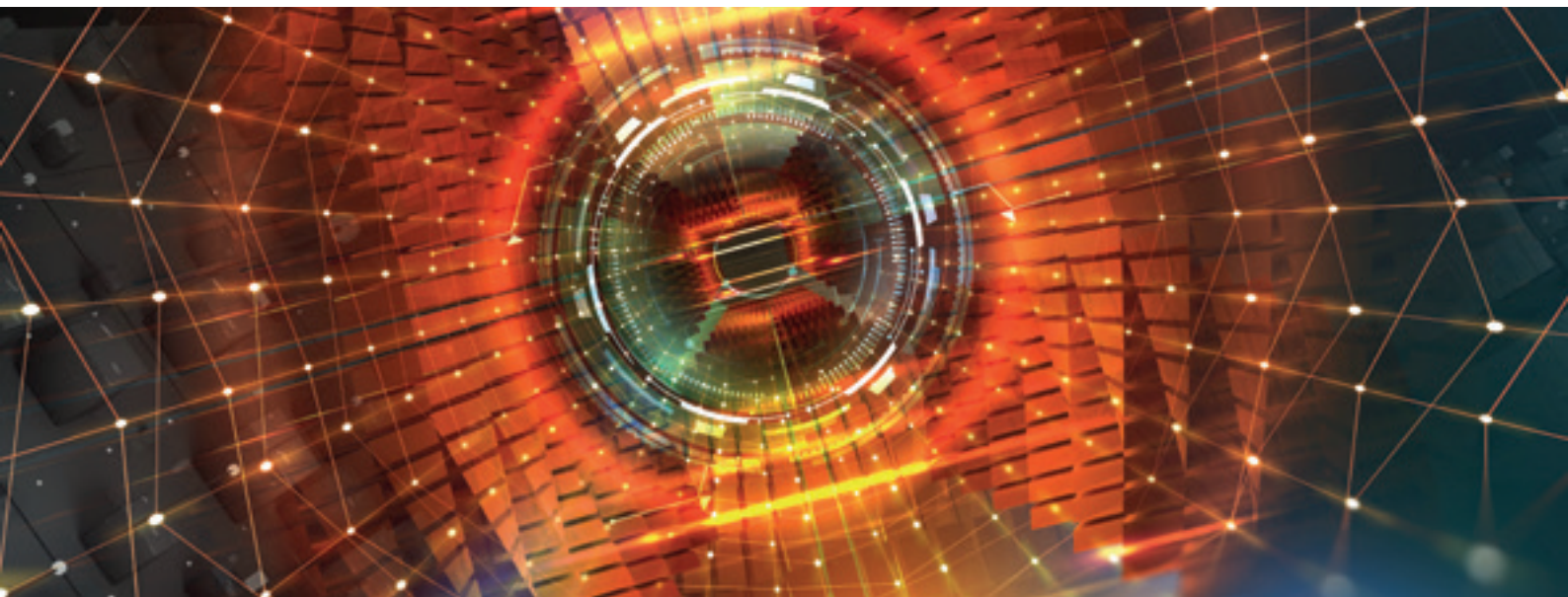


Figure 1 : Schéma de l'accélération d'électrons à deux étages.

afin de l'accélérer à des énergies encore plus élevées. La technique est illustrée par le schéma ci-dessous. La simulation de ce second étage commence donc avec un injecteur déjà formé qui se propage à la suite d'un laser de puissance. Initialement dans le vide, le laser et l'injecteur entrent dans un plasma. Le laser forme une onde plasma dans son sillage. L'interaction de cette onde avec l'injecteur va permettre d'augmenter très fortement l'énergie des électrons qui le composent.

1.2 Adaptation à l'évolution des processeurs et des machines de calcul modernes

Depuis une dizaine d'années, les processeurs que nous utilisons pour le calcul scientifique sont devenus multi-cœurs et donc parallèles dans le but de prolonger la loi de Moore. La puissance de calcul par cœur est restée quant à elle un peu près constante moyennant des changements dans l'architecture et les jeux d'instructions principalement pour des raisons de dissipation thermique. La diminution de la finesse de gravure



(et donc l'augmentation du nombre de transistors) est jusqu'à maintenant toujours possible bien qu'une forme de convergence soit de plus en plus palpable. Cela nous a permis il y a quelques années de rentrer dans l'air des processeurs dits many-core capables de concentrer dans une même puce, un nombre très important de cœurs de calcul (jusqu'à 72 pour les processeurs Intel KNL). Les processeurs les plus récents se situent entre deux avec un nombre de l'ordre de la vingtaine de cœurs. C'est notamment le cas des processeurs de la gamme Intel Skylake EP qui équipent le calculateur Joliot-Curie. En parallèle de l'augmentation du nombre de

cœurs, les constructeurs, et en particulier Intel, ont poussé l'extension des capacités vectoriels SIMD (Single Instruction Multiple Data) des cœurs de calcul. La vectorisation SIMD est la capacité pour un pipeline de traiter non plus un élément (traitement scalaire) mais un vecteur de données pour le même coût en terme de cycle d'horloge. Les jeux d'instructions les plus récents (AVX512 chez Intel) permettent l'addition et la multiplication de vecteur de 8 flottants à double précision en quelques cycles. Cette capacité

s'est néanmoins développée au détriment de la fréquence d'horloge. Les cœurs KNL sont par exemple bridés à 1.4 GHz peu importe la capacité du vecteur exploité. Sur Skylake, cette contrainte a été en partie effacée en permettant au processeur d'adapter sa fréquence (2.4, 2, 1.5 GHz) en fonction de la taille de vecteur utilisée (2, 4 ou 8 flottants double précision par exemple).

Pour exploiter pleinement les super-calculateurs équipés de ces processeurs, il est nécessaire de tenir compte des différents niveaux de parallélisation à savoir un parallélisme distribué entre sockets, un parallélisme partagé entre les cœurs d'un même socket et la vectorisation. La **vectorisation est un point critique pour beaucoup de codes**. Les processeurs actuels sont toujours capables de faire tourner les jeux d'instructions plus anciens. Tout code peut être compilé et exécuté sur un

Intel Skylake ou KNL par exemple. En revanche, la vectorisation n'est pas toujours fonctionnelle sur des algorithmes codés pour des cœurs scalaires. Les conséquences peuvent être dramatiques. Sans effort d'adaptation des algorithmes les plus exigeants, les codes peuvent s'exécuter finalement moins vite que sur des processeurs plus anciens.

Des efforts constants sont menés pour permettre au code Smilei de fonctionner sur les derniers modèles de processeurs de manière optimale. Il y a quelques années, des développements novateurs (utilisation d'OpenMP, meilleure utilisation des caches, équilibrage de charge) ont permis d'adapter le code au niveau de parallélisme partagé inter-nœud. Ces développements ont été testés grâce aux précédents Grands Challenges. Les efforts les plus récents ont porté sur la vectorisation des sections de code les plus gourmandes en temps. Les codes Particle-In-Cell ont le désavantage d'être particulièrement sensibles à la vectorisation. Certaines étapes sont tout simplement impossible à vectoriser dans leurs implémentations classiques.

Ces efforts sont menés en étroite collaboration entre les ingénieurs HPC de la Maison de la Simulation et les équipes de développement scientifique du LULI et du LLR. Ce travail s'inscrit dans une volonté plus large d'améliorer les connaissances et les expériences sur nouvelles architectures afin d'en faire profiter ensuite d'autres codes de calcul.

1.3 Objectifs scientifiques et numériques —

1.3.1. Objectifs scientifiques. La simulation numérique du deuxième étage est extrêmement coûteuse car le temps d'interaction entre le faisceau et l'onde de sillage est de deux ordres de grandeurs plus important que lors du premier étage. Afin de permettre des études paramétriques, il est donc important de mettre en oeuvre des modèles réduits qui permettent de réduire considérablement ce temps de calcul.

C'est ainsi que nous avons développé le modèle d'enveloppe décrit en détail dans la section 2. L'objectif scientifique de ce grand challenge est de mettre à l'épreuve ce modèle d'enveloppe dans des conditions réalistes en le comparant à une simulation laser complète qui serait par ailleurs trop coûteuse pour être effectuée. Ce grand challenge permet donc de valider et d'évaluer les limites du modèle d'enveloppe tel qu'il est implémenté dans Smilei.

L'objectif scientifique de ce grand challenge est de mettre à l'épreuve ce modèle d'enveloppe dans des conditions réalistes en le comparant à une simulation laser complète.



Ces résultats sont critiques car ils définiront le cadre de toutes les études à venir qui utiliseront ce modèle.

1.3.2. Objectifs numériques. Au niveau numérique, nous avons pour objectif de finaliser la vectorisation du code et de tester nos optimisations sur des cas de production et sur une machine récente où la vectorisation joue un rôle important.

2. Code, moyens de calcul, simulation / méthode numérique

2.1 Nom du code / projet

Le code *Particle-In-Cell* (PIC) utilisé, Smilei [1,2], est un code majoritairement C++. Il est hybride avec une parallélisation MPI et OpenMP. Smilei utilise également Python pour lire et interpréter ses fichiers d'entrée. Python peut aussi être utilisé pour définir ses propres fonctions de diagnostic.

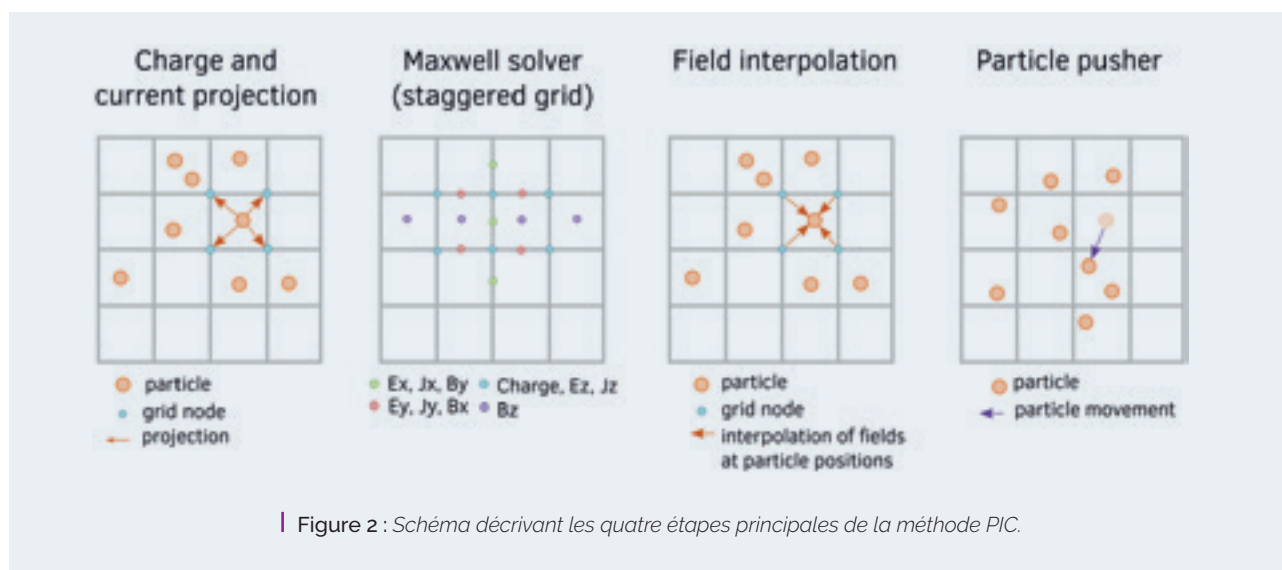
Les méthodes dites *Particle-In-Cell* doivent leur nom à la façon dont les fonctions de distribution en vitesse et en espace décrivant la matière (plasma) sont discrétisées. Elles sont ici décomposées sous forme de macro-particules. Ces macro-particules peuvent être vues comme des groupes de particules réelles de même espèce ayant les mêmes propriétés

cinétiques. Ce formalisme permet la simulation de larges configurations 2D et 3D (matière dense de plusieurs μm de volume sur plusieurs picosecondes) à moindre coût de calcul. Elle s'oppose aux méthodes Monte-Carlo (simulation directe des particules) beaucoup plus chères et les méthodes de type Vlasov plus précises dans certaines conditions mais coûteuses en 3D.

L'interaction électromagnétique entre les particules, ou avec un champ externe, est permise par la discrétisation des équations de Maxwell et la mise en place d'un processus de projection/interpolation avec ces grilles. Dans Smilei, les équations de Maxwell peuvent être discrétisées suivant une décomposition cartésienne par un schéma aux différences finies centré (méthode de Yee et ses dérivées modernes pour améliorer la dispersions numériques) ou par l'utilisation de méthodes pseudo-spectrales (plus coûteuses mais largement plus précises car sans dispersion numérique).

Smilei utilise un parallélisme hybride. Une première décomposition de son domaine (grilles de champs et particules) en sous-domaines est utilisée pour MPI. Au sein de chaque sous-domaine MPI, la charge est de nouveau subdivisée sous forme de très petites portions de domaine appelée patch. La décomposition par patch a plusieurs buts :

- Permettre une parallélisation hybride OpenMP au sein des domaines MPI
- Améliorer la localité mémoire par l'optimisation des caches
- Permettre la mise en place d'une méthode efficace d'équilibrage de charge par échange de patches entre domaines MPI



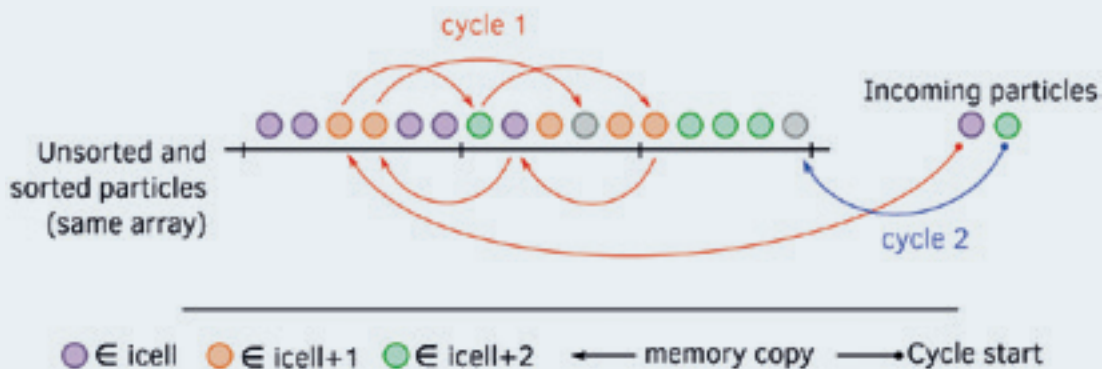


Figure 3 : Schéma décrivant le système de permutations régissant le tri par cycle.

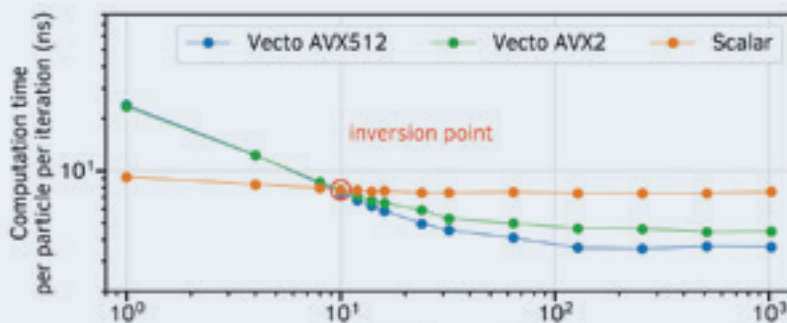


Figure 4 : Temps de simulation par particule et par itération en fonction du nombre de particules par cellule en utilisant les opérateurs scalaires, vectoriels compilés en AVX2 et vectoriels en AVX512 pour un cas de plasma homogène simulé sur un nœud Skylake.

2.2 Nouvelles implémentations

2.2.1. Optimisation et vectorisation. Avoir des particules bien triées en fonction de la cellule dans laquelle elles sont localisées permet de construire plus facilement des algorithmes vectoriels efficaces sur ces particules. Une nouvelle méthode de tri a tout d'abord été implémentée pour faciliter, par la suite, la vectorisation de nombreux opérateurs dans le code. Cette méthode dite par cycle, cherche à minimiser le nombre de déplacements en mémoire en effectuant le maximum de permutations successives. Le tri s'effectue par cellule de maillage comme schématiquement montrée par la figure 3. Cette méthode de tri remplace une autre méthode couramment utilisée dite de tri comptage (ou counting sort). Cette première méthode, bien que très simple à

implémenter, implique une duplication complète de la mémoire et un nombre de copie égal au nombre d'éléments à trier. Elle était utilisée avant l'introduction des algorithmes vectoriels sans être systématiquement utilisée. Le tri par cycle est en fait bien plus efficace lorsque le niveau de désordre est faible. Ce tri étant amené à être effectué à chaque itération temporelle, cette condition est assurée (d'autant que les particules ne peuvent se déplacer de plus d'une cellule par pas de temps). La méthode est décrite avec beaucoup de détails dans l'article soumis [4]. En plus de permettre une meilleure vectorisation, le tri a des effets bénéfiques sur l'utilisation des caches mémoires.

Il est habituel de décomposer la méthode PIC électromagnétique en 4 étapes. Elles sont



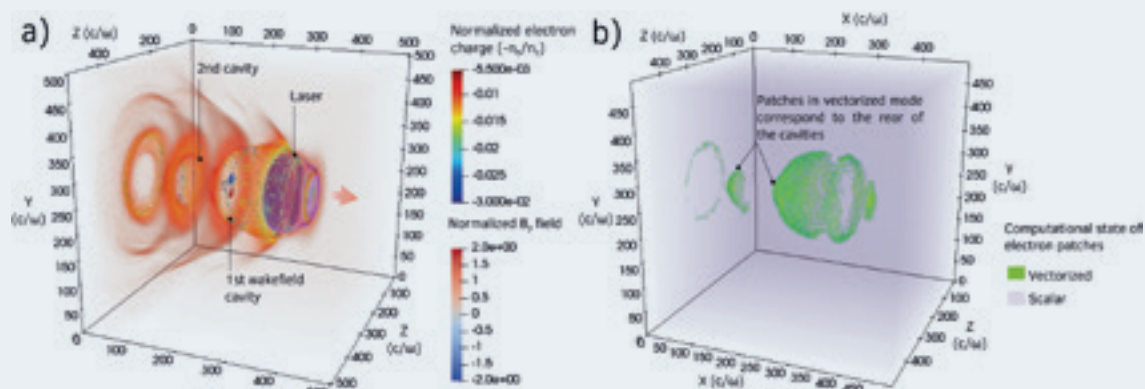


Figure 5 : a) Densité de charge électronique (normalisée) au cours de la simulation. b) État vectoriel ou scalaire des patches pour l'espèce électron pour le même instant que la densité de charge électronique.

explicitées sur la figure 2. Les deux étapes les plus coûteuses et les plus difficiles à vectoriser sont celles où se déroulent une interpolation entre les grilles et les particules soient la déposition de courant et l'interpolation des champs. Ces deux étapes ont été vectorisées en tenant compte de l'avantage d'un tri par cycle. Les détails des algorithmes utilisés sont encore dans l'article soumis [4].

En faisant tourner des simulations de référence avec les opérateurs vectoriels et les opérateurs scalaires, nous nous sommes rendus compte que les opérateurs scalaires restaient très efficaces lorsqu'il y avait peu de particules par maille. Pour un cas de plasma homogène thermalisé, une étude paramétrique a été réalisée en fonction du nombre de particules par cellule. Les résultats sont tracés sur la figure 4.

En dessous d'une dizaine de particules par cellule, il est plus intéressant d'utiliser les algorithmes scalaires. Cela s'explique par la complexité des algorithmes vectoriels (utilisation de buffers temporaires, décomposition des boucles...) qui n'est compensées qu'à partir d'un certain nombre d'éléments à traiter. Dans la pratique, même en menant une simulation avec beaucoup de particules par maille, il arrive dans certaines configurations que le plasma évolue au point de ne laisser qu'un faible nombre de particules par cellule à traiter dans certaines zones. L'idée d'une vectorisation adaptative présentée par la suite a germé au début du grand challenge.

Le concept de la vectorisation adaptative consiste à pouvoir s'adapter en temps et

localement à la variation de particules par maille pour être le plus performant. Le concept a été implémenté avec succès au niveau de la granularité la plus basse du code, à savoir le patch. Un patch est une petite région du domaine et une sous-décomposition des régions MPI possédant ses propres particules et grilles en champ, traité par un thread OpenMP. Sur un cas de plasma homogène (comme utilisé dans l'étude paramétrique), la vectorisation adaptative permet dans n'importe quel cas de toujours bénéficier des meilleurs opérateurs. Les détails de la méthode sont explicités dans l'article soumis [4]. La vectorisation adaptative a été testée sur des cas de production sur le calculateur Joliot-Curie grâce à l'allocation du grand challenge.

Parmi les cas physiques de production étudiés figure l'accélération d'électron par onde de sillage. Ce cas est illustré par la figure 5a. Dans cette simulation, un laser se propage dans un plasma sous-dense d'hydrogène (issu d'un jet de gaz). Lors de sa propagation, le laser génère dans son sillage une onde plasma appelée onde de sillage. Cette onde se compose de cavités successives quasiment vides entourées par une bulle électronique de forte densité. Au sein d'une cavité (parfois appelée bulle), le champ électrique longitudinal (dans la direction x) est accélérateur pour les électrons dans la partie arrière et décélérateur dans la partie avant. Les champs magnétiques peuvent également être focalisant. Des électrons issus de l'enveloppe peuvent être injectés depuis l'arrière de la bulle et se retrouver dans la partie accélératrice. Ils peuvent aussi être artificiellement injectés. Ces



électrons peuvent successivement former un faisceau compact et dense d'énergie relativement élevé (pouvant atteindre plusieurs GeV dans les meilleures configurations). Ces électrons peuvent alors être extraits pour diverses applications.

Dans le plasma ambiant, le nombre de particules est suffisamment faible pour que le mode scalaire soit utilisé. L'enveloppe et en particulier l'arrière des cavités sont beaucoup plus denses et par conséquent riches en particules. Ces zones bénéficient des opérateurs vectorisés qui sont sélectionnés automatiquement par le mode adaptatif. Cela est parfaitement illustré par la figure 5b. Les zones vertes correspondent aux patches vectoriels alors que les zones restées en violet fonctionnent en scalaire. Sur les cas de production testés, cette méthode apporte des speed-ups mesurés allant de x1.3 à x2 sur le traitement des particules.

2.2.2. Modèle d'enveloppe

L'hypothèse du modèle d'enveloppe, à savoir une forme du potentiel vecteur du laser donnée par une enveloppe complexe variant lentement $\tilde{A}$ et modulée par des oscillations rapides à la fréquence du laser $\omega_0 = k_0 c$, peut s'écrire :

$$\tilde{A}(x, t) = \text{Re}[\tilde{A} \exp(ik_0(x - ct))]$$

En introduisant cette expression dans l'équation de d'Alembert, qui décrit l'évolution du potentiel vecteur, on obtient l'équation d'enveloppe dans les coordonnées du laboratoire qui est résolue dans Smilei.

Les équations de Maxwell ainsi que l'équation du mouvement des particules sont moyennées en temps sur un cycle optique du laser. On

obtient ainsi un jeu d'équations modifiées qui ne sont plus contraintes par l'extrême résolution spatiale et temporelle du laser et qui sont donc beaucoup moins coûteuses à résoudre. En revanche, une équation supplémentaire doit être résolue à chaque pas de temps : l'équation d'enveloppe qui décrit l'évolution du laser. La boucle PIC modifiée est décrite dans le schéma ci-dessous.

2.3 Configuration

2.3.1. Simulation d'accélération d'électrons pour tester la vectorisation adaptative (Fig. 5)

Pour la simulation présentée en Fig. 5 qui a permis de tester la vectorisation sur un cas de production, nous avons utilisé un domaine de dimension $503 \times 503 \times 503 (c/\omega_0)^3$ où c est la vitesse de la lumière dans le vide et ω_0 la fréquence laser. Le domaine a une discrétisation de $1280 \times 320 \times 320$ cellules ce qui correspond à un pas d'espace $dx = 0.39 c/\omega_0$ et $dy = dz = 0.157 c/\omega_0$. Le pas de temps respecte la CFL avec $dt = 0.31 \omega_0^{-1}$. On utilise une fenêtre glissante qui voyage avec le laser pour couvrir une longue distance d'interaction. La simulation a été réalisée avec plusieurs quantités de particules par cellule : 4, 8 et 16. La reconfiguration pour le mode adaptatif est effectuée toutes les 50 itérations.

2.3.2. Simulation d'accélération d'électrons pour tester la vectorisation adaptative (Fig. 5)

La simulation de production standard avec le laser complet est effectuée sur $7168 \times 256 \times 256$ cellules. Les cellules sont de tailles

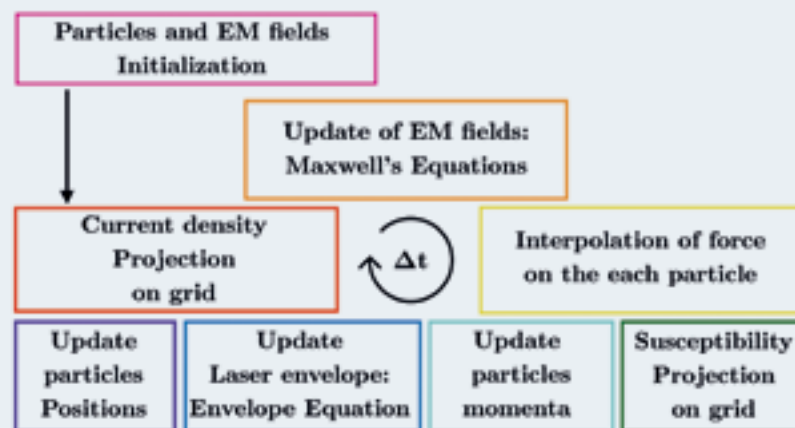


Figure 6 : Boucle temporelle d'un code PIC enveloppe montrant les opérations effectuées à chaque itération.

$0.196 \times \lambda_0 \times \lambda_0 (c/\omega)^3$. Pour respecter la CFL le pas de temps dt s'accorde au pas d'espace longitudinal dx ainsi $dt = 0.195 dx$. On utilise 8 particules par cellules.

Pour la simulation d'enveloppe, la résolution longitudinale est 16 fois moindre et la résolution temporelle 13 fois moindre. Les autres paramètres sont identiques.

2.4 Moyens de calcul et simulation

2.4.1. Simulation d'accélération d'électrons pour tester la vectorisation adaptative (Fig. 5).

Pour cette simulation, nous utilisons 48 nœuds Intel Skylake soit 96 processeurs. Cela correspond à 2304 cœurs. Nous utilisons 1 processus MPI par socket et 24 threads OpenMP par processus MPI. Le code a été compilé avec le compilateur Intel 18. On a utilisé IntelMPI pour les communications MPI.

2.4.2. Simulation de production laser complet V.S. modèle d'enveloppe.

La simulation en laser complet utilise 900 nœuds de la partition Skylake de Joliot-Curie soit 43200 cœurs dans la même configuration hybride que la simulation précédente. **La simulation avec modèle d'enveloppe, beaucoup moins gourmande n'utilise que 128 nœuds soit 6144 cœurs toujours dans la même configuration hybride.** Il est notable que la simulation en modèle d'enveloppe est 16 fois moins volumineuse mais n'utilise que 7 fois moins de nœuds. Les nœuds sont donc en fait environ deux fois moins chargés ce qui est généralement défavorable à la scalabilité des applications étudiées (effet de strong scaling).

2.4.3. Passage à l'échelle faible et fort

Les études de passage à l'échelle ont été menées en début de projet, le code n'intégrait alors pas toutes les optimisations concernant la vectorisation. Nous utilisons ici un plasma homogène thermalisé qui a pour intérêt d'être un cas parfaitement équilibré au niveau de la charge. Au début du grand challenge, nous utilisons la bibliothèque OpenMPI. Il s'agissait d'une version multi-threaded, supportant le niveau MPI_THREAD_MULTIPLE requis par le code, en dehors des bibliothèques par défaut. Le compilateur utilisé était Intel. Les résultats présentés ici reflètent donc l'état du code et de la machine au début du grand challenge. Ils sont donc à prendre avec du recul. Il serait intéressant de refaire la même étude avec

la configuration stabilisée que nous utilisons maintenant à savoir le compilateur Intel 18 et IntelMPI. Pour les études de scalabilités faibles et fortes présentées par la suite, les deux cas présentés utilisent différentes répartitions MPI/OMP :

- Cas hybride à un processus MPI par socket et 24 threads OpenMP par processus MPI
- Cas MPI complet ce qui signifie 48 processus par nœud et 1 OMP par processus MPI.

2.5 Performance

2.5.1. Simulation d'accélération d'électrons pour tester la vectorisation adaptative (Fig. 5)

La vectorisation adaptative apporte des speed-ups mesurés allant de x1.3 à x2 sur le traitement des particules en fonction des cas. Le bénéfice est optimal lorsque la simulation passe d'un mode à un autre (de scalaire à vectorisé par exemple) de manière marquée sur l'ensemble du domaine.

2.5.2. Simulation de production laser complet V.S. modèle d'enveloppe

La simulation de production de référence sur 900 nœuds a pu être poussée jusqu'au temps $t = 2.05 \times 10^5 \omega_0^{-1}$. Pour cela, 46 heures de calcul soit un total d'environ 2 Mh cœurs ont été nécessaires.

Pour arriver à ce même temps avec le modèle d'enveloppe, 7H de calcul auront été suffisantes pour un total de 43 kH de calcul. On observe donc un gain en temps de calcul d'un facteur environ 46 et cela dans des conditions d'échelle a priori défavorable au cas enveloppe pour lequel les processus MPI étaient moins chargés.

2.5.3 Etudes de passage à l'échelle

Une première étude de scalabilité faible a été menée à la fois pour tester les performances du code sur la machine fraîchement disponible et aussi pour évaluer la machine en elle-même. Les temps obtenus sont donnés sur la figure 7 pour les deux configurations MPI/OpenMP. L'efficacité est donnée sur la figure 8. La courbe en temps

Le concept de la vectorisation adaptative consiste à pouvoir s'adapter en temps et localement à la variation de particules par maille pour être le plus performant.



montre qu'en dessous de 256 nœuds pour notre cas, l'utilisation unique de MPI reste le plus performant. L'hybride devient intéressant lorsque le nombre de nœuds devient très important. Les communications MPI finissent par devenir trop lourdes par rapport au coût supplémentaire induit par l'hybride. Cela se démontre en étudiant les courbes de scalabilité pour chaque section du code (non montrées ici). Les étapes consacrées aux échanges MPI des grilles maxwelliennes voient l'efficacité de scalabilité s'écrouler avec le nombre de nœuds. Elles sont responsables en majeure partie de la chute au niveau du code complet. Le seuil est atteint plus vite en utilisant MPI uniquement. Une efficacité de 70 % est atteinte sur quasiment toute la machine en hybride complet. Elle chute à 50 % en utilisant uniquement MPI.

Des études plus récentes ont montré que, sur quelques nœuds, l'optimal était atteint avec une configuration entre les deux présentées. Ces

études n'ont pas été confirmées sur un grand nombre de nœuds.

A la suite de l'étude en scalabilité faible, une étude de scalabilité forte a été menée. Comme pour la section précédente, deux configurations ont été étudiées. La première est le cas entièrement MPI et la deuxième le cas hybride complet (un processus MPI par nœud et 24 threads OpenMP par MPI). Les temps sont montrés sur la figure 9. L'efficacité est quant à elle affichée en figure 10. Le cas entièrement MPI présente une très bonne efficacité jusqu'à 256 nœuds et décroche ensuite pour tomber à 50 % à 1024 nœuds. Le mode hybride est beaucoup plus surprenant. L'efficacité décroît jusqu'à 50 % à 128 nœuds avant de remonter au dessus de 100 % par la suite. Si cela est observable sur les détails de scalabilité dans les étapes de communication MPI, la raison n'est pas connue. L'étude demande à être refaite avec la version actuelle du calculateur.

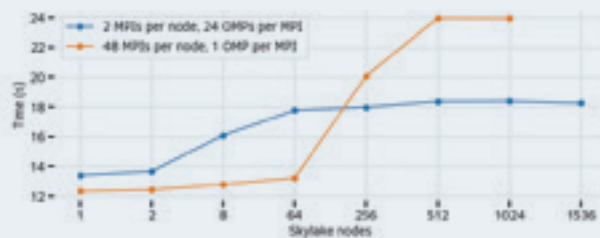


Figure 7 : Étude de scalabilité faible : Temps de simulation en fonction du nombre de nœuds utilisés pour un cas entièrement MPI et un cas hybride à 1 processus MPI par socket et 24 threads OMP par processus MPI.

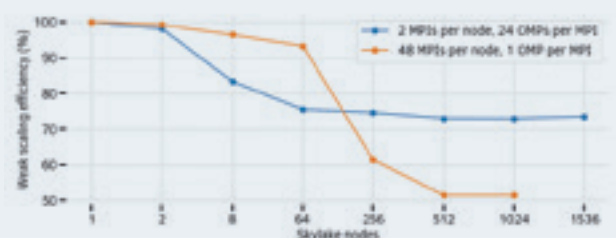


Figure 8 : Étude de scalabilité faible : Efficacité en fonction du nombre de nœuds utilisés pour un cas entièrement MPI et un cas hybride à 1 processus MPI par socket et 24 threads OMP par processus MPI.

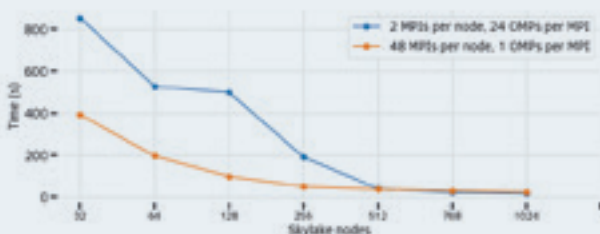


Figure 9 : Étude de scalabilité forte : Temps de simulation en fonction du nombre de nœuds utilisés pour un cas entièrement MPI et un cas hybride à 1 processus MPI par socket et 24 threads OMP par processus MPI.

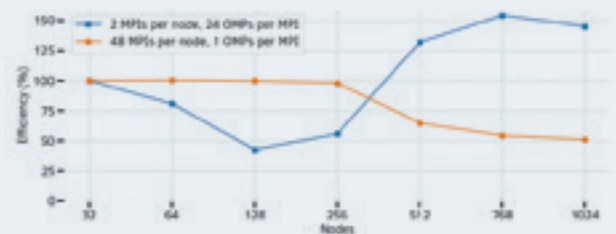


Figure 10 : Étude de scalabilité forte : courbe d'efficacité de en fonction du nombre de nœuds utilisés pour un cas entièrement MPI et un cas hybride à 1 processus MPI par socket et 24 threads OMP par processus MPI.



2.6 Environnement technique

Problèmes rencontrés :

- Instabilités de la bibliothèque OpenMPI supportant le MPI_THREAD_MULTIPLE proposée sur la partition Skylake. Les simulations ont au final été menées dans un environnement IntelMPI, il a notamment été nécessaire d'installer une bibliothèque HDF5 parallèle compatible avec cette bibliothèque MPI.
- Deadlock lors de l'initialisation des simulations en raison de l'utilisation par le code d'un interpréteur Python mettant en jeu le système de fichiers NFS. Le système de fichiers a été stabilisé par le TGCC au début du mois de juin.

3. Observations et résultats scientifiques

Après des validations sur des cas tests petits et simples, un cas de production relativement réaliste a été abordé. Il s'agit de l'étude d'un second étage d'accélération par sillage laser tel qu'il est imaginé pour le moment dans la future installation Cilex [5] autour du laser Apollon. Les paramètres du laser et du plasma sont donc tout à fait réalistes. Seuls les paramètres du faisceau injecté ont été idéalisés pour le moment car nous ne disposons pas encore de données complète sur ce point.

Ci-dessous sur la figure 11, 12, 13 et 14, des représentations dans le plan de polarisation du laser ainsi que sur l'axe du champ électrique accélérateur et de la densité électronique après 3 mm de propagation. Sur chacune des représentations, des comparaisons entre laser

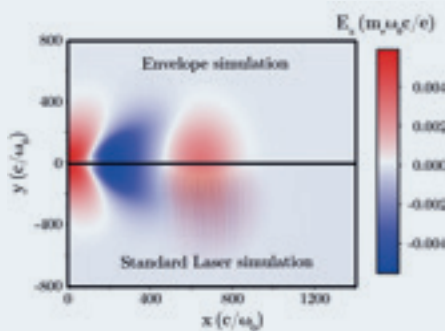


Figure 11 : Champ électrique longitudinal E_x après 3 mm de propagation.

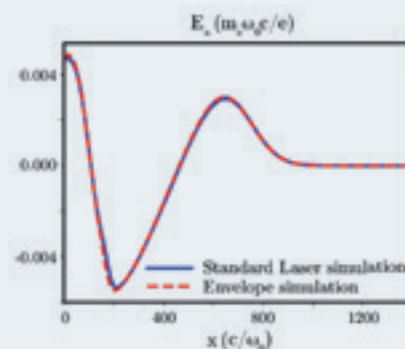


Figure 12 : Champ électrique longitudinal normalisé sur l'axe de propagation après 3 mm de propagation.

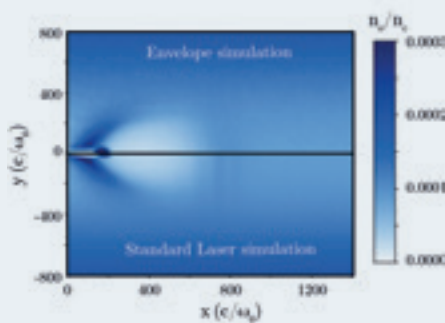


Figure 13 : Densité de charge électronique normalisée après 3 mm de propagation.

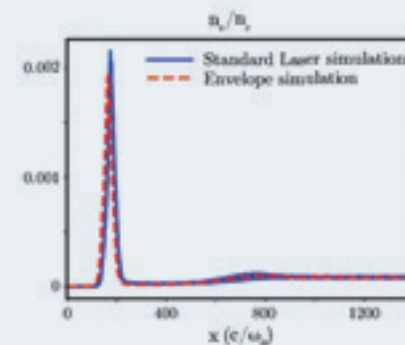


Figure 14 : Densité de charge électronique normalisée sur l'axe de propagation après 3 mm de propagation.



standard et modèle d'enveloppe sont faites. On note un excellent accord alors même que les simulations d'enveloppe ont nécessité 20 fois moins de ressources de calcul pour produire ces résultats sans faire de compromis sur la géométrie complètement 3D du problème. On note néanmoins un très léger déphasage du faisceau accéléré.

Dans des études prolongées sur une distance 5 fois plus importante, ce déphasage a tendance à s'accroître et finit par créer de petites différences dans l'énergie même du faisceau qui ne voit pas exactement le même champ accélérateur dans les deux simulations. Cela dit, la simulation "laser standard" est elle aussi soumise à de la dispersion numérique. Des études plus poussées utilisant des solveurs de Maxwell sans dispersion sont donc nécessaires pour départager les deux simulations sur ce point précis.

■ 4. Perspectives

Le modèle d'enveloppe a désormais fait ses preuves pour les faibles distances de propagation. Nous allons pouvoir l'exploiter pour des simulations d'injecteurs inférieures à 3 mm. La plupart de ces simulations sont faites sur 1 ou 2 mm donc le domaine de validité du modèle est très largement respecté. Le facteur 20 en temps de calcul va rendre possible certaines études paramétriques notamment sur les durées des impulsions lasers et les densités de plasmas en régime d'auto-injections.

Ces résultats motivent également certains développements plus approfondis pouvant dorénavant se baser sur le modèle d'enveloppe. Il est notamment question d'inclure un modèle d'ionisation et aussi d'adapter ce modèle à la géométrie cylindrique pour accélérer davantage les simulations.

Enfin la question de la précision du modèle d'enveloppe sur les grandes distances n'est pas encore complètement éclaircie. Celle-ci devra être adressée par des études complémentaires avec des solveurs de type spectraux réputés beaucoup moins sensibles à la dispersion numérique.

■ 5. Conclusions

Le grand challenge a permis d'effectuer des simulations de grande ampleur tout en continuant à améliorer le code. Dans un premier temps, les efforts de vectorisation commencés avant le grand challenge ont été finalisés et ajustés pendant cette période. Les opérateurs vectorisés et la méthode de vectorisation adaptative ont pu être testés sur la dernière architecture Intel Skylake. **Surtout, le grand challenge nous a donné la capacité de tester ces optimisations sur des cas de production exploitant jusqu'à 2300 cœurs.** Le sujet principal du grand challenge portait sur les simulations d'accélération d'électrons en configuration double-étage. En parallèle des efforts d'optimisation, un modèle d'enveloppe a été implémenté dans le code. Ce type de modèle est adapté à ce type de scénario physique. Il permet surtout de réduire considérablement le temps de calcul nécessaire par rapport à des simulations avec un laser complet. Il donne par conséquent accès à des temps et des domaines de simulation beaucoup plus grands. Le modèle d'enveloppe a été validé avec succès sur la simulation du deuxième étage. La comparaison avec une simulation à laser complet effectuée sur 72h avec deux reprises sur 43200 cœurs a pu être faite grâce aux ressources du grand challenge.

■ 6. Bibliographie

- [1] - J. Derouillat, et al. "SMILEI : A collaborative, open-source, multi-purpose particle-in-cell code for plasma simulation", In Computer Physics Communications, Volume 222, 2018, Pages 351-373, ISSN 0010-4655.
- [2] - A. Beck, et al. "Load management strategy for Particle-In-Cell simulations in high energy particle acceleration", In Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, Volume 829, 2016, Pages 418-421, ISSN 0168-9002.
- [3] - T. Tajima, J.M. Dawson, " LAser electron accelerator", Phys. Rev. Lett. 43 (4), 1979, p. 267.



[4] - A. Beck, J. Derouillat, M. Lobet et al., Adaptive SIMD optimizations in particle-in-cell codes with fine-grain particle sorting, Computer Physics Communications (2019), <https://doi.org/10.1016/j.cpc.2019.05.001>.

[5] - B. Cros et al., Nucl. Instr. and Meth. in Phys. A, 740, 2014, p. 27.

■ Équipe projet



De gauche à droite :
M. Lobet,
A. Specka, I. Zenzemi,
A. Beck, J. Derouillat,
F. Massimo.



Philippe GHENDRIH, CEA/Direction de la Recherche Fondamentale/Institut de Recherche sur la Fusion par confinement Magnétique

Pour l'équipe GYSELA

Elisabetta CASCHERA, Guilhem DIF-PRADALIER, Peter DONNEL, Xavier GARBET, Philippe GHENDRIH, Virginie GRANDGIRARD, Guillaume LATU, Chantal PASSERON, Yanick SARAZIN



GOSPEL – Gyrokinetic Outer layer Simulations with Penalised Extraction at the Limiter

24 millions d'heures

1. Contexte

ITER, le chemin vers la fusion thermonucléaire par confinement magnétique.

La maîtrise de l'énergie thermonucléaire est une des grandes aventures scientifiques qui s'offrent à l'humanité. Elle combine un enjeu majeur, celui de produire de l'énergie, ainsi que de nombreux défis interdépendants, défis d'organisation au niveau mondial, défis technologiques, ainsi que défis de la recherche fondamentale notamment sur la turbulence et l'auto-organisation des systèmes hors-équilibres. Cette source d'énergie a beaucoup d'attrait mais se révèle difficile à mettre en œuvre. La communauté internationale a choisi de se rejoindre dans la construction d'un dispositif expérimental exceptionnel : le tokamak ITER. Le verrou qui doit être surmonté est l'obtention d'un régime de réactions quasiment auto entretenues. ITER est un dispositif expérimental avec une configuration magnétique torique appelée tokamak. En support à ce défi scientifique et technologique, le CEA, la Fédération de Recherche pour la Fusion par Confinement Magnétique, et les partenaires européens, chinois et japonais, ont modifié le tokamak TORE SUPRA (aussi localisé à Cadarache) pour réaliser un dispositif

permettant de mieux soutenir la conception et la construction d'ITER. Ce nouveau dispositif expérimental s'appelle WEST. L'enjeu des expériences est de réaliser les conditions requises pour que la matière, dite à l'état plasma, atteigne des températures supérieures à 10 millions de Kelvin. Maintenir de telles températures sur Terre, sur des échelles de l'ordre du mètre, suppose la mise en œuvre d'une **capacité d'isolation thermique exceptionnelle. Le champ magnétique intense joue ce rôle et assure le confinement du plasma.** L'organisation géométrique du champ magnétique est un élément crucial du transport de la chaleur dans ces plasmas. A la frontière du plasma, elle détermine l'empreinte du plasma sur les éléments solides. Dans cette étude, nous considérons un élément solide de géométrie relativement simple, le limiteur. La couche limite de plasma qui se développe à cette interface est appelée SOL (Scrape-Off Layer). Elle entoure le volume torique du plasma central et se caractérise par une largeur faible. La physique de la SOL est un enjeu majeur du programme de WEST, et, si sa largeur est trop faible, elle constituera une contrainte opérationnelle dure pour ITER. Le rôle de la SOL est double, d'une part sa largeur définit la surface d'interaction entre le plasma et



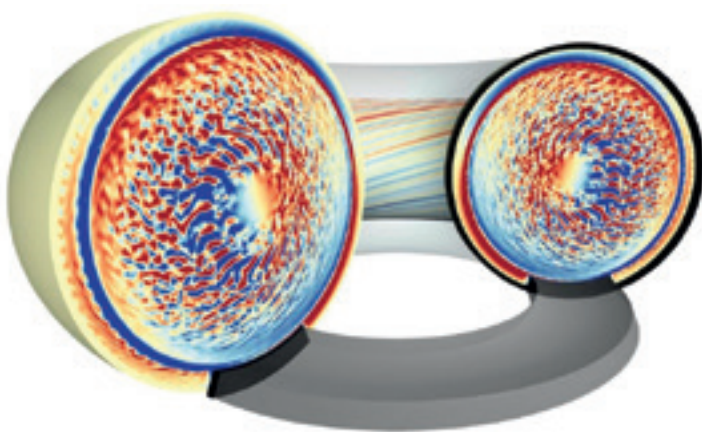


Figure 1 : Image reconstruite à partir des simulations GYSELA donnant une carte du champ de potentiel électrique à un instant donné. Deux coupes sont effectuées : (i) le tore est ouvert permettant de visualiser la structure des structures de potentiel, ces dernières sont analogues aux vortex pour la turbulence des fluides neutres, (ii) une zone pénalisée (zone grisée) permet de créer un puits de chaleur semblable à celui réalisé expérimentalement.

les éléments de paroi, d'autre part, ses propriétés de transport constituent les conditions limites pour le transport turbulent dans la région centrale. Pour ces deux raisons, la SOL est un enjeu scientifique très important et aura un rôle significatif sur la performance du dispositif. Pour ITER, le régime de référence, déterminé empiriquement à partir des expériences existantes, est caractérisé par une barrière de transport renforçant le confinement à l'interface avec la SOL. Alors que la largeur de la SOL résulte a priori d'un transport turbulent,

la barrière de transport est caractérisée par une baisse significative du transport turbulent. Il n'existe pas à ce jour de compréhension complète de ce régime. Le seuil d'apparition, et en conséquence le contrôle, ne sont connus qu'empiriquement à partir des dispositifs existants et opérants à performance bien moindre que celle visée pour ITER. La largeur de la couche limite fait aussi l'objet de recherche avec récemment des résultats contradictoires entre la projection à partir d'une loi empirique et une simulation réalisée



pour ITER. L'enjeu de la recherche fondamentale en arrière-plan de ces deux problèmes, au cœur du programme scientifique d'ITER, est celui de la turbulence et son auto-organisation dans le contexte des plasmas de fusion, systèmes hors-équilibres par excellence.

Les expériences dans ITER doivent commencer progressivement pour atteindre après une dizaine d'années, vers 2035, le niveau nominal de performance. La construction et la mise en œuvre des très nombreux sous-systèmes de

haute technologie sont au centre de l'activité actuelle. À terme, la stratégie expérimentale doit combiner une prise de risque nécessaire à la progression scientifique la plus efficace et l'intégrité d'un dispositif unique dans un environnement nucléaire. Dans ce contexte, les simulations, celles basées sur des modèles empiriques comme celles dites de premier principe joueront sans nul doute un rôle stratégique aussi bien pour orienter le programme expérimental que pour analyser les mesures. Le but est non seulement d'atteindre les conditions de fusion

thermonucléaires d'isotopes de l'hydrogène et des températures supérieures à celles du soleil, plus de cent millions de degrés, mais aussi de préparer les étapes suivantes pour la réalisation de réacteurs destinés à la production d'énergie.

Les simulations fondées sur des modèles empiriques, comme celles dites de premier principe, joueront un rôle stratégique aussi bien pour orienter le programme expérimental que pour analyser les mesures.

2. GYSELA en support à WEST et à ITER

La performance des plasmas de fusion thermonucléaire est déterminée par le transport turbulent qui se développe dans ce milieu très loin de l'équilibre thermodynamique. Pour atteindre les conditions de fusion thermonucléaire et en conséquence améliorer le confinement de l'énergie, **l'alternative est l'augmentation de la taille du dispositif ou une maîtrise accrue de la turbulence.** Dans cette logique ITER est le

plus grand et le plus complexe des dispositifs de confinement magnétique jamais construit. Améliorer la performance pendant son opération impose de progresser dans le contrôle du transport turbulent. Ce point de vue s'est construit sur une base empirique complétée par l'analyse théorique et les simulations. En effet, les mesures dans les plasmas sont très difficiles, aussi bien du point de vue de la réalisation que de l'interprétation. La turbulence est connue essentiellement par des mesures de fluctuations, souvent en un seul point pour une expérience donnée. La connaissance impose ainsi de combiner des bases de données expérimentales et une grille de lecture théorique. Les simulations sont nécessaires pour rendre compte de cette physique non-linéaire, si possible de la manière la plus quantitative possible. Plusieurs approches sont suivies, en premier lieu, comme pour la turbulence des fluides neutres, des modèles basés sur des équations de type Navier Stokes sont utilisées. En raison de la faible densité du plasma, des modèles cinétiques semblent plus appropriés. Le principal avantage de ces derniers est de rendre compte de manière correcte de phénomènes cinétiques, comme les collisions, et d'offrir un cadre théorique respectant les symétries du problème dans la limite des champs magnétiques très intenses. Le coût est de passer de l'étude 6 champs 3D en représentation fluide à typiquement 16 384 champs 3D en tenant compte des niveaux de vitesse de la représentation cinétique. Pour les deux approches, nos équipes privilégient une approche globale du transport, dans un régime de forçage par un flux imposé de l'extérieur dans le but de réduire au maximum les paramètres libres des équations et d'offrir une description la plus proche possible des premiers principes.

Le code GYSELA est utilisé dans cette perspective, par le biais d'une approche cinétique et en mettant l'accent sur les processus d'auto-organisation entre les différentes échelles : de la taille de la machine, aux échelles microscopiques caractérisant les trajectoires des particules. Pour être cohérent, le modèle doit traiter de manière correcte la couche limite qui se développe entre le plasma et le voisinage de la paroi du réacteur. Le problème est difficile et peu de codes ont la possibilité de traiter cette interface, cette géométrie étant en général exclue lors de la conception du code. L'objet du grand challenge GOSPEL est de mettre en œuvre une technique de pénalisation qui permet de résoudre cette difficulté dans GYSELA et ainsi de modéliser cette interface. Les équations sont modifiées dans la région



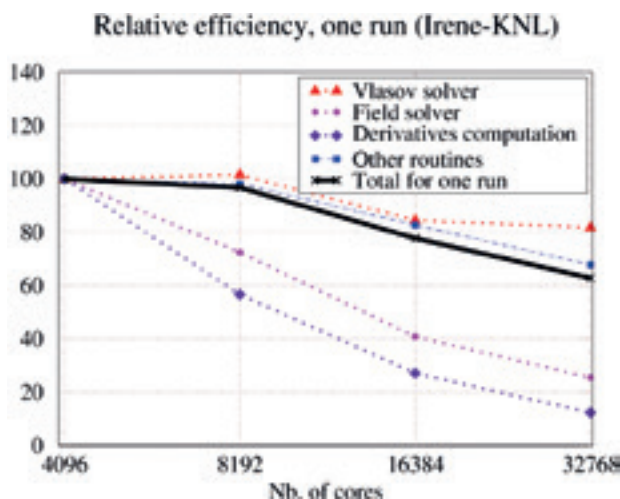


Figure 2 : Étude de performance dite de scalabilité (strong scaling) réalisée avec GYSELA sur JOLIOT-CURIE, partition KNL, dans le cadre du grand challenge GOSPEL. Efficacité relative pour un benchmark de strong scaling avec un maillage de 34 milliards de points.

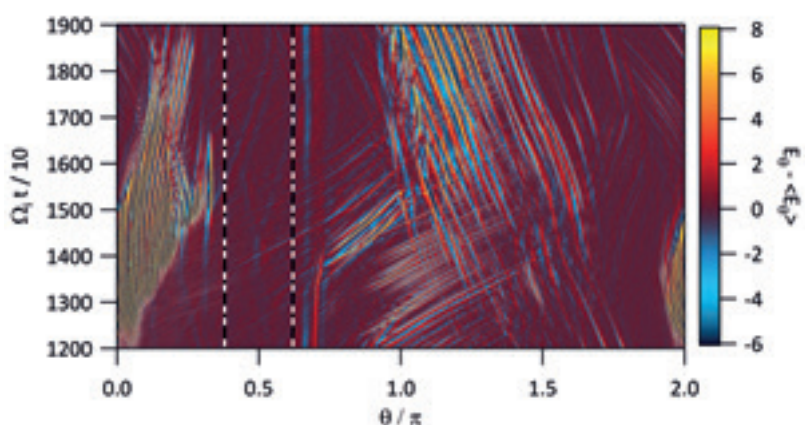


Figure 3 : Structure à petite échelle du champ électrique en fonction de l'angle -approximativement transverse à la direction du champ magnétique ainsi qu'à la direction radiale- et du temps. Les fluctuations constituant le drapé de filaments se développent à des échelles très petites. La région délimitée par les traits verticaux blancs et noirs correspond à celle du limiteur. L'inclinaison des filaments dans cette figure est inversement proportionnelle à la vitesse de propagation des fluctuations. Elle dépend de la localisation de l'instabilité et du signe du champ électrique dans la direction radiale. Ce dernier s'inverse au voisinage du maximum de cisaillement. Sur la figure, on observe la coexistence de directions de propagation opposées de ces structures filamentaires. Cela correspondrait à des mécanismes de génération de part et d'autre du point de champ électrique nul.

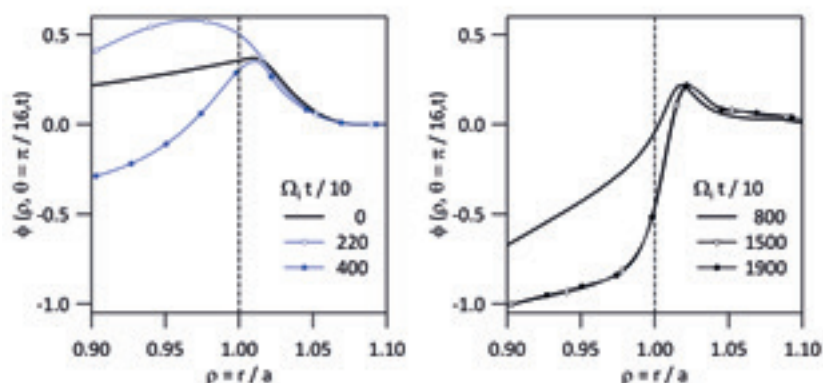


Figure 4 : Deux séquences de l'évolution du profil radial du potentiel électrique. La ligne verticale à $\rho = 1$, délimite la SOL à droite, du plasma central à gauche. Dans la SOL, identifiée par un masque, la contrainte de pénalisation fait tendre le potentiel électrique vers le profil initial—courbe en trait noir— voir la figure de gauche correspondant au début de la simulation. La courbe du potentiel s'inverse ce qui correspond à un changement du signe du champ électrique. Le potentiel électrique se creuse dans la région centrale. Un régime quasi-stationnaire est atteint dans le dernier quart de cette simulation, figure de droite. Le maximum de cisaillement, donné par la dérivée seconde du potentiel, est localisé au voisinage $\rho = 1$, la couche limite se développe comme attendu au voisinage de la séparatrice entre la SOL et le plasma central.



correspondant au solide. On introduit en particulier dans ce volume un point froid qui agit comme un bain thermique, absorbant toute la chaleur transportée dans ce volume. Cette approche permet de réaliser des simulations comprenant la SOL. C'est l'objet du grand challenge GOSPEL « Gyrokinetic Outer layer Simulations with Penalised Extraction at the Limiter ». Cependant, s'agissant d'un modèle simplifié, nous n'avons pas accès à l'ensemble de la physique.

■ 3. Objectifs

Des résultats contradictoires ont été obtenus sur la physique de la couche limite et en particulier sa largeur. Cette dernière est un élément crucial pour l'opération des expériences de fusion. Pour ITER, la projection à partir des résultats actuels couvre toutes les possibilités, y compris des conditions d'opérations très difficiles. L'image physique utilisée en appui à la loi empirique rendant compte de la base de données repose sur un régime sans turbulence, tout au moins à l'échelle ionique qui nous intéresse, et dominé par les effets conjoints des collisions et des écoulements à grande échelle imposés par les variations spatiales du champ magnétique, variation en direction est en amplitude à l'échelle du tokamak. Notre objectif était triple, i) explorer la physique introduite par la pénalisation, ii) étudier la validité de l'interprétation proposée pour la largeur de la SOL, iii) étudier l'interdépendance de la turbulence au centre du plasma et celle se développant dans la SOL. Un quatrième thème s'est surimposé à ces objectifs, celui des barrières de transport.

■ 4. Code de calcul GYSELA

Cette étape dans le développement du code GYSELA est importante notamment dans la perspective de l'évolution du code vers l'Exascale.

Le code de calcul GYSELA est un code de simulation de la turbulence plasma, cette dernière déterminant le transport de chaleur des ions, voire des électrons dans les cas encore exploratoires où cette espèce est prise en compte. Le code est adapté aux architectures massivement parallèles et fait l'objet d'une recherche d'amélioration continue dans le cadre de collaborations

nationales et internationales. Le **modèle pour le plasma est cinétique, c'est-à-dire qu'il traite l'espace des phases, a priori à 6 dimensions sous forme de convection en coordonnées de position et en vitesse sous l'influence de collisions et en présence d'un champ de vitesse turbulent associé au champ électrique. Ce dernier est déterminé en utilisant les équations de Maxwell.** Une séparation d'échelle, cohérente avec l'intensité du champ magnétique présent dans les tokamaks, **permet de réduire la dimension de l'espace à 5 dimensions. Cette réduction, cruciale pour la réalisation de simulations numériques avec les puissances de calcul actuellement disponibles, se traduit par une plus grande complexité des équations.** En effet, dans cette approche, les particules sont transformées en boucles de courant dont la taille est finie et comparable, en général plus grande, à celle du maillage. On introduit ainsi un caractère non-local dans la physique et une nouvelle contrainte sur le numérique. Parmi les signatures de ce changement de modèle il convient de souligner que les nouvelles particules portent une polarisation et une magnétisation. Le code permet de déterminer la fonction de distribution 5D des ions dans le cadre formel appelé gyrocinétique. Un schéma semi-lagrangien réalise l'avancement en temps. Il est original par rapport à la plupart des schémas utilisés dans les codes concurrents, *Particle In Cell* (PIC) ou Eulérien. En simplifiant, on peut considérer que l'on étudie la déformation du maillage lors des convections dans l'espace des phases et que l'on projette alors la fonction de distribution attachée à ce maillage sur un maillage régulier. On combine ainsi les propriétés de suivi des trajectoires des codes PIC avec l'utilisation d'une grille fixe comme dans le cadre des codes Eulériens. Le code **GYSELA a déjà été testé sur 450 000 cœurs**, sur l'un des plus gros calculateurs européens (http://www.fz-juelich.de/ias/jsc/EN/Expertise/High-Q-Club/_node.html), avec une bonne efficacité. La simulation de base utilisée comme point de référence pour ce projet est une simulation de 272 109 points de maillage utilisant quelques milliers de cœurs pour un temps de calcul d'un mois environ. Différentes versions des équations et de la pénalisation ont été utilisées pendant le grand challenge, en particulier un cas axisymétrique ou une symétrie par rapport à l'axe de la machine est imposée au plasma ainsi qu'un cas annulaire (sans simulation de la région la plus centrale du tore) afin de pouvoir augmenter le raffinement de maillage dans la région périphérique.



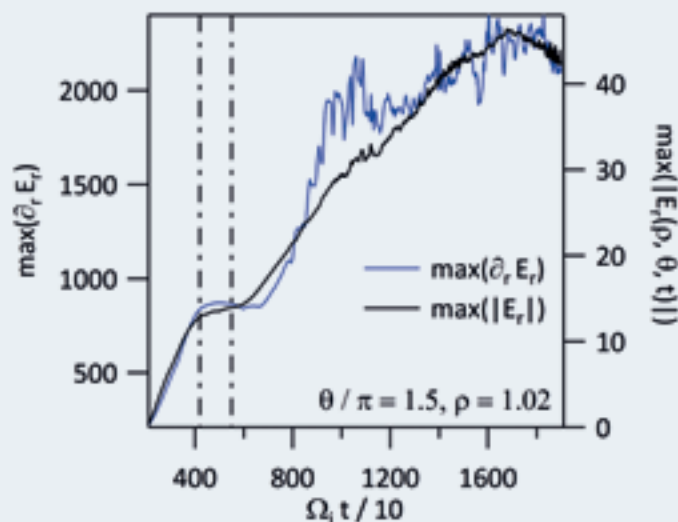


Figure 5 : Évolution de deux indicateurs caractérisant la couche limite, d'une part l'amplitude du cisaillement du champ électrique, échelle logarithmique de gauche et courbe en trait bleu, d'autre part le minimum de champ électrique, échelle linéaire à droite, courbe en trait noir. Les deux tracés présentent un plateau que l'on peut attribuer à une phase de saturation de l'instabilité de Kelvin-Helmholtz, avant que le développement de la couche limite ne reprenne et sature dans le dernier quart de la simulation.

5. Portage et optimisation de GYSELA sur architecture manycore

Depuis une dizaine d'années, les contraintes énergétiques ont limité l'augmentation des fréquences d'horloge chez les fondeurs de processeurs et ont conduit à l'**ère du multicore puis du manycore**. Cela se traduit par une augmentation du parallélisme et du nombre de cœurs au sein de chaque nœud de calcul des supercalculateurs. Dans ce cadre, la technologie de l'*hyperthreading* permet à deux threads d'être actifs et de réaliser des calculs sur un même cœur. Un certain nombre de ressources se retrouvent ainsi partagées : bus internes, registres, unités de calculs, mémoires cache. En 2009, on employait typiquement 8 threads par nœud. Mais avec l'architecture KNL, en considérant des nœuds à 64 cœurs, c'est maintenant 128 threads qui sont actifs dans le code GYSELA. Cette approche permet de maximiser le nombre d'instructions en calcul flottant qui sont traitées à chaque cycle horloge, et l'approche est beaucoup plus efficace que dans les générations

précédentes de processeurs. En affectant deux *threads* sur un même cœur, cela permet de proposer deux flux de travaux indépendants qui pourront plus facilement occuper les unités de calcul disponibles. De plus, les unités de calcul sont vectorielles, c'est à dire qu'elles sont capables de réaliser des additions, multiplications sur des petits vecteurs de 8 flottants double précision sur KNL. Dans GYSELA, comme dans d'autres codes, cette capacité de traitement sur des petits vecteurs augmente les capacités de calcul de manière significative. Mais elle n'est effective que si le compilateur possède les qualités nécessaires pour transformer le code source en une série d'opérations vectorielles. Les versions 2017 et 2018 des compilateurs d'INTEL sont caractérisées par un net progrès dans le traitement de cette tâche difficile. Cependant, le programmeur doit mettre en forme le code source pour que la transformation par le compilateur soit la plus efficace possible. C'est ainsi que de nombreux codes, dont GYSELA, ont été remaniés afin d'assister le compilateur dans la détection des opérations vectorielles les plus performantes. Ce remodelage du code augmente significativement le débit des opérations de calcul flottant. Par ailleurs, le nombre élevé de cœurs au sein de chaque nœud nous a conduits



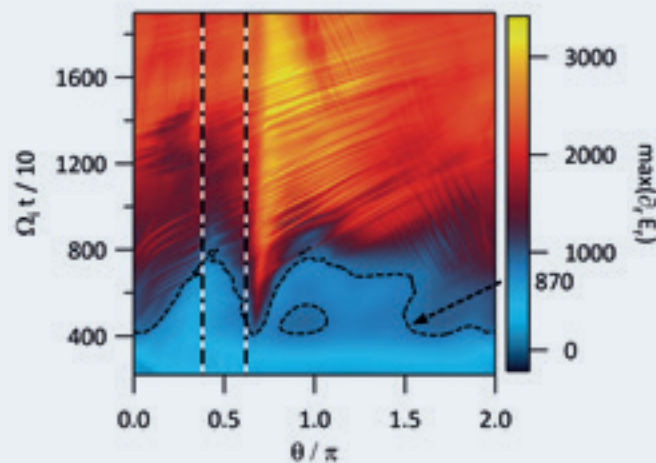


Figure 6 : Maximum du cisaillement du champ électrique, et donc de la vitesse de dérive électrique en fonction de la position angulaire et du temps. Les traits verticaux, tirets blancs et noirs, indiquent la position angulaire du limiteur. L'orientation des filaments suggère que le maximum de cisaillement est situé avant le point d'inversion du champ électrique en se déplaçant vers la SOL. Le contour noir correspond à une valeur du cisaillement qui semble correspondre au seuil pour l'instabilité Kelvin-Helmholtz.

à rechercher puis à trouver plus de parallélisme à grain fin en utilisant une plus grande variété de fonctionnalités du paradigme de parallélisation d'OpenMP. Cette révision de la parallélisation des « nids de boucles » a permis d'occuper les cœurs de calculs avec un meilleur équilibrage de charge et réduire l'inoccupation des cœurs. Une révision de certaines utilisations combinées de MPI et OpenMP ont aussi permis d'éviter les synchronisations globales qui ralentissent les *threads hyperthreadés*. Ces actions ont permis de gagner en temps d'exécution. L'ensemble de ces travaux de vectorisation et d'équilibrage de charge ont été efficaces à la fois pour l'architecture manycore Intel KNL, mais aussi pour les architectures x86 Intel Broadwell et Skylake. **Les accélérations s'échelonnent entre 1.7x et 3x selon l'architecture considérée.**

En 2017, le code GYSELA avait été porté et optimisé sur la partition KNL de Marconi (<http://www.hpc.cineca.it/hardware/marconi>) dans le cadre d'un projet EUROfusion. Le retour d'expérience a montré que GYSELA avait la capacité de tourner sur cette architecture. Néanmoins cela nécessite de tenir compte d'une augmentation du temps de calcul nécessaire par rapport à l'architecture INTEL Skylake, mais aussi de considérer un nombre plus important de cœurs pour une simulation de taille donnée. Sur la base de ces résultats, le grand challenge GOSPEL sur la partition KNL de Irène Joliot-Curie

était pertinent et ne devait pas conduire à des difficultés informatiques majeures.

6. Mise en œuvre de GYSELA sur JOLIOT-CURIE/KNL

Durant les mois d'avril-mai-juin, nous avons interagi avec le support pour une mise en place progressive d'un système de fichiers parallèles opérationnels, d'une bibliothèque MPI en état de marche sur un grand nombre de nœuds, et enfin d'une partition de calcul pleinement fonctionnelle. La **partition KNL devient véritablement disponible à la production en juillet**. Cela nous pose d'emblée un problème d'organisation car le grand challenge principalement était prévu initialement en mai-juin. Apparaissent également des **problèmes de stabilité de la machine lorsque nous augmentons le nombre de nœuds mobilisés**. Pour surmonter les difficultés, un partenariat est mis en place en interaction directe avec quelques personnes de l'équipe support plutôt que l'utilisation habituelle de tickets-emails proposée par la hotline. De plus, des réservations exceptionnelles de 512 nœuds sur des plages de temporelles planifiées permettent de résoudre efficacement les problèmes se produisant



à large échelle. La disponibilité d'un grand nombre de nœuds opérationnels permet, fin juillet, de réaliser un test de strong scaling sur une configuration comprenant 34 milliards de points de maillage jusqu'à 512 nœuds. Les résultats indiquent un bon comportement général des performances à large échelle. L'efficacité relative est de 63% à près de 33 000 cœurs.

Pendant les deux dernières semaines du grand challenge, une stabilisation relative de la partition permet de progresser et de dégager les objectifs et les paramètres des simulations complémentaires que nous réalisons dans le cadre de PRACE et d'EUROfusion. A l'arrêt du grand challenge, le programme scientifique de GOSPEL est engagé, mais nous sommes très loin des ambitions initiales, notamment pour les plus grands maillages envisagés dans le projet. Les simulations emblématiques réalisées avec GYSELA restent ainsi celles de grands challenges de la période 2014-2015. Le programme scientifique initié avec GOSPEL a été poursuivi avec les autres moyens HPC auxquels nous pouvons avoir accès. Les résultats présentés ici ne font pas la distinction entre les ressources utilisées.

7. Observations et résultats scientifiques

Le grand challenge GOSPEL est le premier test d'une méthode de pénalisation avec le code gyrocinétique GYSELA. C'est un point de passage pour réaliser des simulations se rapprochant des conditions expérimentales. Notre objectif est d'étudier le transport de chaleur ionique, de la source située au centre du tokamak jusqu'à la périphérie où l'énergie touchant la paroi est absorbée. Ce dernier mécanisme est mis en œuvre dans notre méthode de pénalisation en imposant un bain thermique froid dans la région correspondant au limiteur. Les électrons demandant des ressources numériques trop importantes pour être simulés, nous utilisons en général une réponse approximative. Le deuxième volet de la pénalisation consiste à adapter la réponse des électrons dans la SOL pour retrouver le comportement attendu. L'application de la pénalisation se fait dans les régions souhaitées par un système de masques. Une gradation dans la mise en œuvre de la pénalisation a été suivie en ajoutant par étape les processus de transport qui

gouvernent le transfert de la chaleur. L'élément le plus important est le transport dans la direction parallèle où les effets cinétiques dominent : l'introduction d'un point froid génère un front froid qui se propage dans la direction parallèle au champ magnétique. Nous avons pu vérifier que la propagation de ce front est absolument conforme aux prédictions. Dans un deuxième temps nous avons complété le mouvement des particules par une vitesse de dérive imposée par les variations spatiales du champ magnétique à l'échelle du tokamak ainsi que par les processus de collisions. Un champ électrique se développe alors car les réponses des électrons et des ions sont différentes. Le deuxième volet de la pénalisation a ainsi été vérifié avec l'apparition d'une inversion du champ électrique dans la direction radiale donc transverse au champ magnétique. Ces deux effets se traduisent par le développement d'écoulements, le premier dans la direction parallèle est transitoire, alors que le second, un écoulement appelé zonal, transverse au champ magnétique et à la direction radiale, est caractérisé par de fortes variations de son amplitude. Celles-ci déterminent ainsi un cisaillement de ce champ de vitesse. L'effet non anticipé résultant de ce

cisaillement est une instabilité de type Kelvin-Helmholtz à très petite échelle. Ces résultats sont toujours à l'étude et doivent donc être considérés comme préliminaires. L'étude de cette instabilité est difficile en raison de l'inhomogénéité spatiale de la couche de vitesse cisailée et son évolution pendant l'instabilité. Le phénomène de couche limite associé demande également une très bonne résolution. La conjugaison de ces deux facteurs conduit à une certaine zoologie des régimes de transport dont certains sembleraient être liés à la résolution. Dans une troisième étape nous avons réalisés des simulations complètes. Dans les conditions retenues, un transport turbulent de chaleur domine dans la région centrale du plasma. A la périphérie, on observe alors deux régimes : un régime où une turbulence de type Kelvin-Helmholtz se développe au voisinage de la couche limite, ainsi qu'un régime où cette deuxième instabilité n'est pas observée de telle

La déstabilisation de la zone de cisaillement de l'écoulement zonal est un élément nouveau dans la physique de la couche limite du plasma.



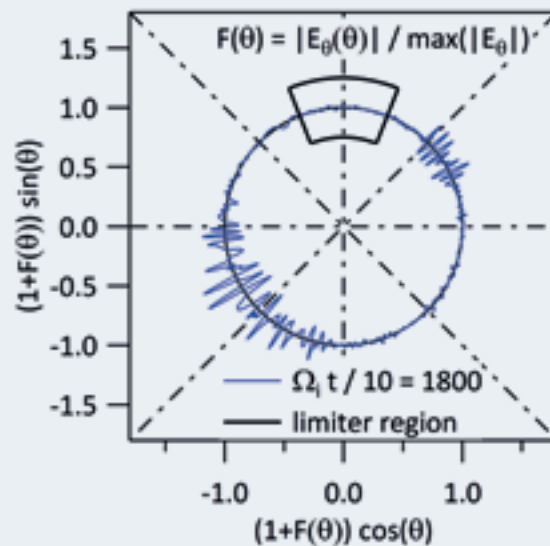


Figure 7 : Représentation de l'intensité du champ électrique en fonction de la variable angulaire. Dans cette représentation le rayon est une fonction $F(\theta)$ du champ électrique pour l'angle choisi. L'instabilité à cet instant de la simulation est localisée à $\theta < \pi/4$, et à $\theta \approx 5\pi/4$. La zone où le limiteur est localisé est indiquée par les deux arcs en tracés noirs.

sorte que l'écoulement zonal fortement cisailé se maintient et crée une barrière de transport. Dans le premier cas, le transport turbulent se propage à l'ensemble du volume, alors que dans le second cas la barrière de transport semble se maintenir de manière stable. L'exploration de ces deux régimes est rendue difficile par les aspects de résolution et de transitoire, sans compter le questionnement sur l'impact du modèle de pénalisation sur ces régimes. Néanmoins, le point fort de ces résultats est la convergence vers des régimes de transport observés expérimentalement à l'exception de celui que l'on souhaite éviter dans ITER, le développement d'une turbulence forte dans tout le volume avec un fort transport de chaleur dissipant la couche de cisaillement de l'écoulement zonal.

La déstabilisation de la zone de cisaillement de l'écoulement zonal conduisant à une turbulence de type Kelvin-Helmholtz est un élément nouveau dans la physique de la couche limite du plasma. Ce résultat du grand challenge est une contribution originale au débat scientifique et à la recherche sur l'extension de cette couche limite et son rôle dans la performance de confinement global du plasma. Plus généralement ces résultats confortent notre stratégie basée sur une approche globale des propriétés de transport

dans les plasmas de tokamak. L'importance des transitoires et la possibilité d'états multiples sont des effets que nous attendions au regard de la richesse des résultats expérimentaux. Nos efforts se poursuivent donc pour comprendre les résultats et caractériser les régimes observés. La résolution élevée qui est indispensable pour ces études a un impact fort sur les ressources HPC qui sont mobilisées. Par ailleurs, nous souhaitons poursuivre les changements dans le code GYSELA pour modifier le modèle de pénalisation et étudier une physique de la couche limite qui sera comparable à celle de la modélisation «fluides» également réalisée à l'IRFM avec le code TOKAM3X.

8. Conclusion

Le grand challenge GOSPEL définit de nouvelles perspectives dans le programme de simulations avec le code GYSELA. Les résultats, toujours en cours d'analyse, et **complétés par d'autres simulations sur d'autres calculateurs ont permis d'obtenir des résultats originaux sur la physique des couches limites des plasmas dans les dispositifs de confinement magnétique.**



Le grand challenge GOSPEL n'aurait pas pu se dérouler dans l'implication et l'aide des équipes de support du TGCC. D'autres ressources de calcul ont été utilisées, notamment dans le cadre d'un projet PRACE et de projets EUROfusion. Ce travail a bénéficié d'une aide du gouvernement français au titre du Programme Investissements d'Avenir, Initiative d'Excellence d'Aix-Marseille Université - A*MIDEX.

9. Bibliographie et références

G. Laval et al., 2007, La Fusion Nucléaire : de la recherche fondamentale à la production d'énergie ?, Rapport scientifique et technique n°26 de l'Académie des sciences, EDP Sciences, <https://www.edp-open.org/images/stories/books/fullldl/rapport6.pdf>

Jean Jacquino, 2013, ITER et les recherches sur la fusion par confinement magnétique, Vol. 32, page 19, Reflets de la physique, DOI: <https://doi.org/10.1051/refdp/201332019>

E.J. Doyle, W.A. Houlberg, Y.Kamada, V. Mukhovatov, T.H. Osborne, A. Polevoi, G. Bateman, J.W.Connor, J.G. Cordey, T. Fujita, X. Garbet, T.S. Hahm, L.D. Horton, A.E. Hubbard, F. Imbeaux, F.Jenko, J.E. Kinsey, Y. Kishimoto, J. Li, T.C. Luce, Y. Martin, M. Ossipenko, V. Parail, A. Peeters, T.L.Rhodes, J.E. Rice, C.M. Roach, V. Rozhansky, F. Ryter, G. Saibene, R. Sartori, A.C.C. Sips, J.A. Snipes, M. Sugihara, E.J. Synakowski, H. Takenaga, T. Takizuka, K. Thomsen, M.R. Wade, H.R. Wilson, ITPA Transport Physics Topical Group and ITPA Confinement Database and Modelling Topical Group and ITPA Pedestal and Edge Topical Group, 2007, Chapter 2: Plasma confinement and transport, Nuclear Fusion, Vol. 47, pages S18, <http://stacks.iop.org/0029-5515/47/i=6/a=S02>

V. Grandgirard, J. Abiteboul, J. Bigot, T. Cartier-Michaud, N. Crouseilles, G. Dif-Pradalier, Ch. Ehrlacher, D. Esteve, X. Garbet, Ph. Ghendrih, G. Latu, M. Mehrenberger, C. Norscini, Ch. Passeron, F. Rozar, Y. Sarazin, E. Sonnendrücker, A. Strugarek, D. Zarzoso, 2016, A 5D gyrokinetic full-f global semi-Lagrangian code for flux-driven ion turbulence simulations, Computer Physics Communications, Vol. 207, pages 35-68, DOI : <https://doi.org/10.1016/j.cpc.2016.05.007>

F. Rozar, G. Latu, J. Roman, V. Grandgirard, 2015, Achieving Memory Scalability in the GYSELA Code to Fit Exascale, Concurrency Computat.: Pract. Exper., Vol. 27, pages 994-1009. doi: [10.1002/cpe.3429](https://doi.org/10.1002/cpe.3429)

G. Latu, N. Crouseilles, V. Grandgirard, E. Sonnendrücker, 2007, Gyrokinetic Semi-Lagrangian Parallel Simulation using a Hybrid OpenMP/MPI programming Recent Advances in Parallel Virtual Machine and Message Passing Interface, Springer, Vol. 4757, pages 356-364 <https://hal.archives-ouvertes.fr/hal-00605748>

X. Garbet, Y. Idomura, L. Villard, T.H. Watanabe, 2010, Gyrokinetic simulations of turbulent transport, Vol. 50, page 043002, DOI <https://doi.org/10.1088/0029-5515/50/4/043002>

■ Équipe projet



Philippe
GHENDRIH



L'équipe GYSELA au CEA Cadarache en octobre 2017



Équipe

Hugues DIGONNET, Chargé de recherche à École Centrale de Nantes, Institut de Calcul Intensif



Grand Challenge MPM-1000 : Adaptation de maillage anisotrope massivement parallèle appliquée à des structures complexes

6 millions d'heures

Ce Grand Challenge a pour but principal de combiner adaptation de maillage anisotrope et calcul massivement parallèle sur le nouveau calculateur du TGCC: Joliot-Curie. Dans ce contexte, nous avons eu l'occasion d'avoir accès à la quasi-totalité de la machine et ainsi de pouvoir exécuter des calculs dimensionnants. Cette combinaison nous a ainsi permis de générer **un maillage de plus de 1 000 milliards d'éléments en utilisant 65 536 cœurs et environ 200 To de mémoire, que ce soit en 2D ou 3D**. L'utilisation d'un estimateur d'erreur anisotrope nous a également permis de construire un maillage adapté à une microstructure composite tridimensionnelle comportant 3 millions de fibres. Ce maillage adapté comporte, au final, 3 milliards de nœuds pour 18 milliards d'éléments, pour une précision équivalente à celle d'un maillage homogène à 5 000 milliards d'éléments. L'utilisation du logiciel Visit nous a également permis de produire des images de très haute résolution (16 Giga-pixels), une image étant distribuée en 256 sous-images de 8192x8192 pixels, pour un volume total, une fois l'image compressée, de 8Go.

■ 1. Contexte

Le numérique est aujourd'hui présent dans pratiquement tous les domaines et plus particulièrement dans le domaine de la simulation. La simulation numérique reste actuellement l'un des plus gros consommateurs en puissance de calcul et reste toujours limitée par celle-ci pour obtenir des calculs toujours plus précis. Cette précision est obtenue par l'enrichissement des

modèles physiques, une meilleure discrétisation spatiale et/ou temporelle pour pouvoir capturer des phénomènes très locaux dans des vastes domaines de calcul.

Les supercalculateurs permettent d'accéder à une puissance de calcul considérable: supérieure de plusieurs ordres de grandeur à nos ordinateurs personnels. Si l'utilisation du calcul parallèle est, aujourd'hui, un outil indispensable dans la réduction des temps de calcul et donne la possibilité de réaliser, en un temps donné, des





calculs toujours plus importants et précis, elle n'est qu'une optimisation parmi d'autres pour permettre de calculer plus vite. Une autre manière de réduire les temps de calcul dans des simulations par éléments finis est l'utilisation de l'adaptation de maillage, qui permet de réduire considérablement la taille des maillages utilisés pour une précision donnée. Elle permet, à l'aide d'un estimateur d'erreur, de concentrer les mailles là où les phénomènes ont eu lieu et, au contraire, de les élargir là où il ne se passe rien. Pour une précision donnée, les maillages ainsi générés sont beaucoup plus petits qu'une grille homogène et réduisent d'autant les temps de calcul. L'utilisation d'un estimateur d'erreur anisotrope permet de réduire encore davantage les maillages, en offrant la possibilité d'avoir des tailles de maille différentes en fonction de la direction spatiale.

Ce grand challenge a pour vocation de montrer qu'il est possible de combiner ces deux optimisations (calcul parallèle et adaptation de maillage) dans le but de **rendre possible des calculs inenvisageables sinon, et ceci à l'échelle d'un supercalculateur de la taille de Joliot-Curie : utilisation de 65 536 cœurs et maillage à plus de 1 000 milliards d'éléments.**

■ 2. Portage et difficultés rencontrées

Si aucun souci de portage n'a été rencontré lors du portage de l'application sur Joliot-Curie, pas mal d'instabilités ont été rencontrées les premières semaines, principalement à grande échelle. À l'échelle de la totalité du supercalculateur, il a été difficile d'obtenir une configuration sans aucun problème sur les nœuds de calcul.

La principale difficulté rencontrée a été liée à l'utilisation de la bibliothèque OpenMPI : lors de communications sur un nombre relativement important de cœurs (>1024), des erreurs de précision numérique et/ou de crash lors de l'appel à `MPI_Wait()` apparaissent. Le problème n'a pas été totalement reproductible (1 fois sur 16 sur 1 024 cœurs mais quasiment tout le temps au-dessus de 16 384 cœurs). Ceci a été résolu lors du Grand Challenge en utilisant la bibliothèque IntelMpi. Cependant, la durée d'initialisation des premières communications est très longue avec IntelMpi (exponentiel avec le nombre de cœurs, et de l'ordre de 30 minutes sur 65 536 cœurs). Un contournement du problème a été trouvé récemment, avec l'ajout d'un appel `MPI_Barrier()` dans une fonction (sans influence sur les temps de calcul).



■ 3. Adaptation de maillage isotrope / homogène

Cette première batterie de tests consiste à tester l'extensibilité de notre procédure d'adaptation de maillage à l'échelle du supercalculateur Joliot-Curie (jusqu'à 65 536 cœurs) et de faire une comparaison en termes de performance avec son prédécesseur, le calculateur Curie. Le but est de rééditer la génération d'un maillage de taille record à plus de 1 000 milliards d'éléments (aussi bien en 2D qu'en 3D), qui avait déjà été réalisé sur Curie en 2017. Si les temps de calcul sont réduits sur Joliot-Curie, la taille des maillages générés sont les mêmes que sur Curie, la partition Skylake de Joliot-Curie ayant sensiblement le même nombre de cœurs que la partition noeuds fins de Curie (79 488 pour Joliot-Curie vs 80 640 pour Curie) et de mémoire RAM (310 To pour Joliot-Curie vs 315 To pour Curie).

3.1 Scalabilité de l'application

Cet accès privilégié à Joliot-Curie lors des Grands Challenges a été pour nous l'**occasion de tester l'extensibilité jusqu'à la quasi-totalité du supercalculateur**, que ce soit en nombre de cœurs (65 536 / 79 488 cœurs) mais également pour la mémoire RAM utilisée (environ 200 / 310 To). Des tests en scalabilité faible ont été menés de 1 à 65 536 cœurs, que ce soit dans le cas de maillages 2D ou 3D. Les tests se font donc à charge de travail constante par cœur, en augmentant la taille

globale linéairement en fonction du nombre de cœurs utilisés. Dans le cas optimal (complexité algorithmique linéaire et accélération idéale), le temps de calcul devrait lui rester constant peu importe le nombre de cœurs utilisés.

3.2 En 2D

Pour le cas 2D, la charge de travail constante consiste à passer d'un maillage initial à environ 4 millions d'éléments par cœur à un maillage final quatre fois plus gros, d'environ 16 millions d'éléments. Les tests ont été effectués sur : 1, 4, 16, 64, 256, 1 024, 4 096, 16 384 et 65 536 cœurs.

Le tableau 1 et l'illustration 1 présentent les temps d'exécution en secondes pour chacune des exécutions, en distinguant les temps du processus d'adaptation, d'entrée incluant le surcoût de l'initialisation d'IntelMPI, et les temps de sortie au format TB (un fichier binaire par cœur).

Cette série de tests, de 1 à 65 536 cœurs, a été un succès et nous a permis de générer un maillage final à plus de 1 000 milliards d'éléments. Plusieurs enseignements sont à tirer de ces temps :

1. pour la phase d'adaptation (hors I/O et initialisation d'IntelMPI), la **scalabilité de l'application est bonne**. On observe une augmentation notable du temps de calcul entre 1 et 64 cœurs, qui peuvent s'expliquer par la montée en charge à l'intérieur d'un nœud de calcul à 48 cœurs : boost en fréquence, plus de mémoire cache/cœur lorsque le nœud n'est pas complètement rempli. Une fois ce palier

Nb coeurs	Temps adaptation (s)	Temps entrée + init intelMPI (s)	Temps sortie (s)	Temps total (s)
1	428	0.27	0.71	429
4	488	1.74	0.67	490
16	740	0.43	0.75	741
64	886	1.10	1.16	888
256	897	2.12	1.22	900
1 024	964	3.50	4.10	971
4 096	1061	10.0	14.8	1086
16 384	1100	118	51.0	1269
65 536	1280	1482	235	3097

Tableau 1 : Test de scalabilité faible (charge de travail constante par cœur) pour l'adaptation de maillage dans le cas 2D. Exécution lancée en utilisant de 1 à 65 536 cœurs de Joliot Curie SKL.



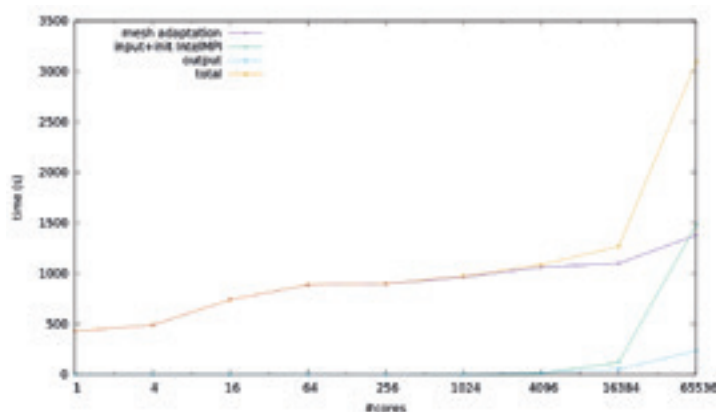


Illustration 1 : Évolution du temps de calcul des phases d'adaptation, de lecture + initialisation d'IntelMPI, d'écriture et total lors du test de scalabilité faible. Idéalement, les temps devraient rester constants. On observe une assez forte augmentation de 1 à 64 cœurs (montée en charge à l'intérieur d'un nœud de calcul) puis un plateau entre 64 et 4 096 cœurs, suivi d'une nouvelle dégradation au-dessus, principalement pour l'initialisation d'IntelMPI et les écritures et dans une moindre mesure de la phase d'adaptation.

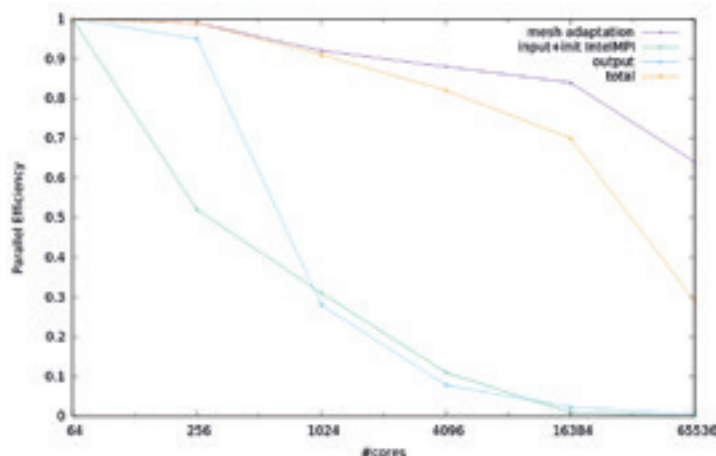


Illustration 2 : Évolution de l'efficacité parallèle des différentes phases lors des tests de scalabilité 2D avec comme référence le cas lancé sur 64 cœurs. Si la procédure d'adaptation se comporte presque parfaitement jusqu'à 16 384 cœurs (>0.9) avec une légère dégradation sur 65 536 cœurs (0.7), les entrées-sorties et l'initialisation d'IntelMPI chutent rapidement au-dessus de 256 cœurs. Pour l'écriture, ces mauvaises performances représentent quand même un débit de 130Go/s, soit assez proche des limites hardware du calculateur qui n'est pas forcément dimensionné pour une utilisation mono-utilisateur.

Nb cœurs	Temps adaptation (s)	Temps entrée + init intelMPI (s)	Temps sortie (s)	Temps total (s)
16	5447	2.73	1.05	5450
128	6661	3.04	1.64	6667
1 024	7440	3.34	3.98	7447
8 192	8794	35.0	52.9	8881
65 536 (*)	19515	1489	326	21330

Tableau 2 : Test de scalabilité faible (charge de travail constante par cœur) pour l'adaptation de maillage dans le cas 3D. Exécution lancée en utilisant de 16 à 65 536 cœurs de Joliot-Curie . (*) le cas sur 65 536 cœurs a eu des problèmes lors de l'exécution, des barrettes mémoire étaient défectueuses sur un des nœuds, d'où une diminution de la bande passante et un temps de calcul multiplié par environ deux



atteint, les temps augmentent beaucoup plus lentement et confèrent une bonne efficacité parallèle à l'application, et ceci de 64 à 65 536 cœurs (efficacité de 0.64) ;

2. **l'initialisation d'IntelMPI devient vraiment pénalisante et prépondérante sur les entrées à très grande échelle** (>4 096 cœurs), finissant avec un coût de l'ordre d'une demi-heure sur 65 536 cœurs ;

3. les sorties parallèles au format pvtu sont parfaitement scalables jusqu'à 256 cœurs avec un temps d'écriture légèrement variable, mais assez constant (<10s). Au-delà de 256 cœurs, les temps d'écriture augmentent assez fortement. Sur le cas à 65 536 cœurs, les écritures représentent environ 3-4 minutes, pour un volume de 30 To, soit **une écriture à quand même 130 Go/s en moyenne**.

L'illustration 2 montre les efficacités parallèles des différentes phases, entre 64 et 65 536 cœurs. On observe un bon comportement pour la phase d'adaptation et une forte détérioration à grande échelle pour l'écriture et l'initialisation d'IntelMPI (limitation principalement liée au hardware et/ou software).

3.3 En 3D

Dans le cas 3D, le test de scalabilité consiste à passer d'un maillage initial d'environ 2.25 millions d'éléments par cœur à un maillage final huit fois plus gros (environ 18 millions d'éléments). Les tests ont été effectués sur 16, 128, 1 024, 8 192 et 65 536 cœurs. Des cas sur 1 et 2 cœurs ont été utilisés pour générer le maillage d'entrée sur 16

cœurs, mais ils ne vérifient pas la contrainte d'une charge de travail constante par cœur et ne sont donc pas présentés ici.

Le tableau 2 présente les temps d'exécution (en secondes) pour chacun des exécutions, en distinguant les temps du processus d'adaptation, les temps d'entrée incluant le surcoût de l'initialisation d'IntelMPI et les temps de sortie au format TB (un fichier binaire par cœur).

À l'exception des problèmes sur le cas à 65 536 cœurs, les remarques sont les mêmes que pour le cas 2D, mais le coût de l'adaptation de maillage est beaucoup plus important avec, respectivement, 2h en 3D et 16 minutes en 2D pour l'exécution lancée sur 1 024 cœurs.

3.4 Une comparaison avec Curie

Ces tests de scalabilité nous ont également permis de faire une comparaison avec le supercalculateur Curie que Joliot-Curie remplace. Comme déjà précisé, la partition Skylake est comparable en termes de taille à la partition noeuds fins de Curie. Les calculs exécutés sur Joliot-Curie ne peuvent donc pas être plus imposants que ceux réalisés sur Curie d'où, en ce qui me concerne, une légère déception de ce côté-là.

Nous présentons ici une comparaison entre Joliot-Curie et Curie, effectuée sur deux tests :

- un cas de taille moyenne, exécuté sur le même nombre de cœurs (16 384 cœurs). Le cas test consiste à générer un maillage final comportant 32 milliards de noeuds et 54 milliards d'éléments. Ce cas n'est pas le cas

	Cas 1		Cas 2	
	Curie	Joliot-Curie	Curie	Joliot-Curie
Nb coeurs	16 384	16 384	19 968	16 384
Taille maillage (nbe)	64x109	64x109	256x109	256x109
Temps adaptation (s)	473	277 (x1.71)	1844	1100 (x2.04)
Temps entrée + Init IntelMPI (s)	9.92	114 (x0.034)	44.9	118 (x0.46)
Temps sortie (s)	33.3	2.26 (x14.7)	180	51 (x4.3)
Temps total (s)	816	393 (x2.07)	2069	1269 (x1.99)

Tableau 3 : Comparaison des temps d'exécution entre Joliot-Curie et son prédécesseur Curie. A nombre de cœurs équivalent, Joliot-Curie est environ deux fois plus rapide que Curie pour notre application. Les gains semblent encore plus importants pour les I/O, mais le contexte n'est pas le même (Grand Challenge vs Production). Enfin, IntelMPI a un coût d'initialisation très important, d'environ 100 secondes sur 16 384 cœurs.



le plus gros qui soit passé sur 16 384 cœurs, mais il est plus représentatif de la charge par cœur généralement utilisé ;

- un cas de taille maximale (cf. section 1.3), qui a été exécuté sur un nombre différent de cœurs : 19 968 sur Curie et 16 384 sur Joliot-Curie. Ce cas est quatre fois plus gros que le cas 1, et a la particularité d'utiliser la quasi-totalité de la mémoire. Pour ce cas, le calcul de l'accélération a été modifié pour tenir compte de la différence en nombre de cœurs utilisés.

$$A = (t_{curie} * \frac{19968}{16384}) / t_{joliot-curie}$$

Le tableau 3 présente les temps obtenus sur Joliot-Curie et sur Curie.

Mis à part le problème d'initialisation d'IntelMPI, qui ajoute un surcoût d'au moins 100 secondes lors de la phase de lecture, les autres étapes sont beaucoup plus rapides sur Joliot-Curie : **d'un facteur 1.71 à 2.04 pour la procédure d'adaptation de maillage ; entre 4.3 et 15 fois plus rapide pour l'écriture des résultats.**

Au final, ceci fait une **accélération globale voisine de 2** à nombre de cœurs équivalent. Ces résultats ont été obtenus sans modification du code source et sont très satisfaisants si l'on fait exception du surcoût de l'initialisation d'IntelMPI.

3.5 Un cas record à 1000 milliards d'éléments

L'utilisation d'IntelMPI nous a permis d'utiliser la quasi-totalité de Joliot Curie, avec 65 536 cœurs, pour obtenir un maillage final à plus de 1000 milliards d'éléments, et ceci aussi bien en 2D qu'en 3D. Quelques difficultés cependant lors de ces exécutions :

- des problèmes de crash mémoire lors d'appels à MPI_Waitall() avec OpenMPI au-dessus de 1000 cœurs. Crashes pas forcément

reproductibles et également assez aléatoires (dépendance aux nœuds utilisés lors des runs ? Problème de communication non bloquante ?).

Un contournement du problème a été mis en place après ces Grands Challenges par l'ajout d'un appel à MPI_Barrier() dans le code, mais la cause du crash reste indéterminée ;

- un temps d'initialisation avec IntelMPI très important, avec la première communication qui coûte environ 1/2 heure sur 65 536 cœurs ;
- à cette échelle (65 536 cœurs), nous avons souvent, parmi les nœuds de calcul engagés dans l'exécution, des nœuds avec certains problèmes, d'où des temps de calcul pas forcément très réalistes. Par exemple, le cas 3D a eu son temps environ doublé suite à un module mémoire d'un des nœuds défaillant lors de l'exécution.

Le tableau 4 présente les caractéristiques des deux maillages (2D et 3D), générés sur 65 536 cœurs.

3.6 Les sorties sur disque

Lors de cette campagne nous avons également pu tester différents types et stratégies d'écriture pour nos résultats. L'utilisation d'un très grand nombre de cœurs entraîne la génération d'un grand volume de données. La stratégie d'écriture initiale, qui consiste à écrire un fichier binaire par cœur (.TB), génère également un très grand nombre de fichiers (inode). Dans cette optique, une version utilisant MPI-IO (.TGB) a été implémentée et consiste à créer un seul très gros fichier. Cependant, les performances de cette approche ainsi que la taille du fichier généré, la rendent inutilisable à très grande échelle (temps d'écriture très long et fichier de plusieurs dizaines de To). Finalement, une approche hybride a été

	2D	3D
Nb cœurs	65 536	65 536
Nb noeuds	509 720 612 451	200 503 473 793
Nb éléments	1 015 438 749 726	1 192 367 638 380
Nb voisins min/max/moy	1/18/6.09	2/38/16.9
Qualité min/max/moy	0.49/1/0.93	0.27/1/0/83

Tableau 4 : Caractéristiques des deux maillages (2D et 3D), à 1 000 milliards d'éléments, générés en utilisant 65 536 cœurs et 200 To de RAM du supercalculateur Joliot-Curie. On note un nombre de domaines voisins les uns des autres raisonnable (une moyenne de 6 en 2D et 17 en 3D, pour un maximum de, respectivement, 18 et 38) et une très bonne qualité (normalisée) des maillages générés (une moyenne de 0.93 en 2D et 0.83 en 3D pour une qualité minimale de, respectivement, 0.49 et 0.27).



développée : elle consiste à générer un gros fichier MPI-IO par groupe de N ($N=256$) cœurs (.DTGB). Cette approche combine les avantages des deux autres méthodes : temps d'écriture contenu, réduction du nombre d'inodes utilisés et fichier de taille importante sans être trop gros.

Le tableau 5 montre les caractéristiques de ces deux types de sortie lors de l'exécution du cas 2D sur 65 536 cœurs. Si le temps d'écriture est plus petit avec la stratégie d'un fichier par cœur (.TB), le temps d'écriture avec la version distribuée de MPI-IO (.DTGB) reste convenable (2,5 fois plus lent) mais a l'avantage de ne générer que 256 fichiers de grande taille (120Go), directement transférables sur l'espace de stockage sur le store (stockage sur bandes), là où la version classique utilise 256 fois plus d'inodes. Bien entendu, la taille du groupe (N) est un paramètre ajustable/optimisable entre 1 (version classique) et le nombre de cœurs utilisé (version MPI-IO globale avec un seul fichier).

de la position dans le domaine). L'utilisation d'un maillage anisotrope hétérogène permet en effet de réduire de façon très importante la taille des maillages utilisés pour une précision donnée.

Le cas test choisi pour cette partie consiste donc à adapter anisotropiquement un maillage pour capturer une microstructure fibreuse complexe constituée de plusieurs millions de fibres. Ces maillages peuvent, par la suite, être utilisés pour calculer des écoulements au travers de cette dernière. Le processus d'adaptation consiste, en partant d'un maillage homogène isotrope, à le modifier itérativement en un maillage anisotrope hétérogène adapté à la micro-structure, processus itératif entre les phases d'immersion de la géométrie, de calcul d'une métrique et une passe d'adaptation de maillage. L'immersion signifie le calcul, sur chaque nœud du maillage de fond, de la fonction de phase de la géométrie qui représentera, dans notre microstructure composite, le renfort.

■ 4. Adaptation de maillage anisotrope-hétérogène

La partie précédente montre la bonne scalabilité de notre procédure d'adaptation de maillage parallèle et ceci jusqu'à l'utilisation de la quasi-totalité du supercalculateur Joliot-Curie (65 536 cœurs et 200 To de RAM utilisés). Cependant, la génération d'un maillage isotrope et homogène (même taille de maille dans chaque direction de l'espace et sur tout le domaine) ne met pas vraiment en avant l'apport que procure l'utilisation de l'adaptation de maillage.

4.1 La problématique

Pour mieux illustrer le bénéfice de l'adaptation de maillage, nous nous plaçons dans le contexte d'un maillage anisotrope (taille de maille différente suivant les directions de l'espace) et hétérogène (taille de maille différente en fonction

4.2 Microstructure sur 9216 cœurs

Le cas présenté ici consiste à immerger une microstructure de plus de 3 millions de fibres dans un cube. Il est intéressant de noter qu'une plaque classique injectée dans un matériau polymère renforcé par des fibres contiendrait environ une centaine de millions de fibres.

Le calcul a été réalisé sur 9 216 cœurs. En partant d'un maillage initial de 5 millions de nœuds et 30 millions d'éléments, nous avons généré, au bout de 25 itérations, un maillage final adapté comportant 3 milliards de nœuds et 18 milliards d'éléments (ce qui correspond à environ 1000 nœuds par fibre). Ce processus itératif n'a pas un coût constant par itération, puisque la taille du maillage traité varie au cours du calcul. L'illustration 3 montre l'évolution, en nombre de nœuds, du maillage ainsi que celle du temps de calcul.

La durée totale du processus d'adaptation a été de 17 heures pour les 25 itérations. L'illustration 4 montre l'occupation mémoire sur un des nœuds au cours des 10 dernières itérations.

Type de sortie	Temps (s)	Volume	Taille fichier	Nb inodes
.TB	235	30 To	480 Mo	65 536
.DTGB	624	30 To	120 Go	256

Tableau 5 : Comparaison de deux stratégies d'écriture des résultats : un fichier par cœurs (.TB) ou un gros fichier MPI-IO par groupe de 256 cœurs. Comparaison faite sur le cas record d'un maillage à 1 000 milliards d'éléments exécuté sur 65 536 cœurs.



■ 5. Visualisation

Un autre point important abordé lors de ce Grand Challenge était la possibilité de visualiser de si gros maillages. En effet, les fichiers résultat au format ".pvtu" représentent un volume conséquent (environ 1 Go par cœur) et leur visualisation à l'aide du logiciel Visit nécessite d'utiliser presque autant de ressources que lors du calcul. La résolution des images générées représente également un challenge car une très haute résolution est nécessaire pour conserver une vue d'ensemble, tout en préservant la possibilité de regarder des détails.

5.1 La problématique

Aux vues de la taille des données traitées, l'utilisation des logiciels de visualisation nécessite d'utiliser quasiment autant de ressources que lors de la génération de maillages. Cette contrainte nous oblige à ne pas effectuer ce travail de visualisation en interactif, mais de considérer cette phase comme un calcul standard à part entière. Pour cela, nous utilisons Visit en mode batch. L'objectif est de générer des images qui pourront, par la suite, être analysées "offline", mais ces dernières doivent être de très grande résolution pour permettre de regarder des détails dans une vue d'ensemble.

Les difficultés avec les images en haute résolution sont : d'une part, Visit via MESA est limité à une résolution maximale de 32768 x 32768 pixels ; d'autre part, à cette échelle, l'empreinte mémoire utilisée par ces dernières est très importante et multipliée par le nombre de cœurs, ce qui nous oblige à utiliser les nœuds de calcul en dépeuplé. Pour ces raisons, la stratégie mise en place consiste à plutôt générer une famille d'images de 8192 x 8192 pixels qui, si elles sont rassemblées, correspondent à une seule image de très haute résolution. Nous avons également fait le choix de générer plusieurs qualités d'image d'une vue d'ensemble en une seule image et jusqu'à 256 images pour la plus grande résolution.

Remarque : lors de ce Grand Challenge, la soumission de Visit en parallèle a été réalisée en utilisant directement le gestionnaire de batch SLURM, sans passer par ccc_msub. C'est effectivement le script python de Visit qui démarre lui-même les processus MPI et, même s'il connaît les commandes srun et sbatch, il ne reconnaît pas le ccc_msub.

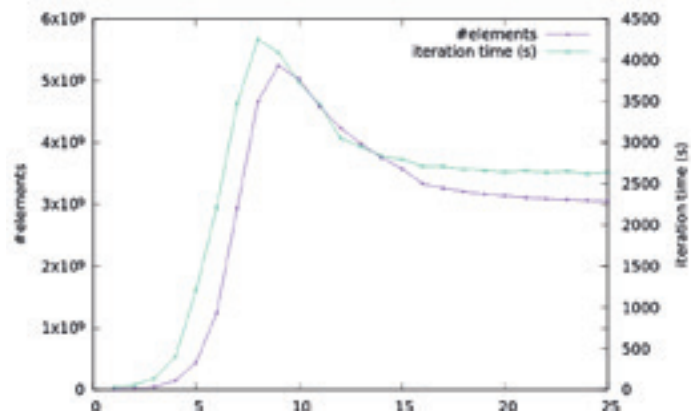


Illustration 3 : Evolution de la taille du maillage (axe de gauche) et du coût de calcul (axe de droite) durant les 25 itérations de la procédure d'adaptation de maillage anisotrope, réalisée sur 9 216 cœurs.

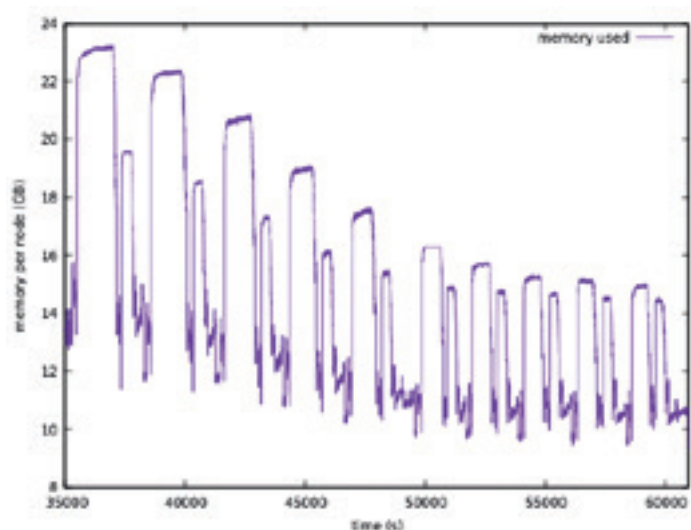


Illustration 4 : Evolution de l'utilisation mémoire sur l'un des 192 nœuds de calcul lors des dix dernières itérations d'adaptation de maillage anisotrope-hétérogène. Le maillage étant équitablement distribué entre les cœurs, ce graphe est représentatif également des autres nœuds. On remarque que cette occupation est directement liée à la taille du maillage (cf illustration 3) et reste bien en dessous des 192 Go de RAM par nœud pour ce cas-là.



5.2 Des images en très haute définition, à 16 Giga-pixels

En appliquant cette stratégie, il nous a été possible de générer des images en très haute résolution : 131072x131072 pixels, soit 16 Giga-pixels, qui sont distribuées en 256 images de 8192x8192 pixels, soit 64 Mega-pixels. À cette résolution, il est possible de distinguer individuellement les éléments du maillage, même sur des maillages à plusieurs milliards d'éléments.

Un script python a également été développé pour permettre une navigation dans ces images distribuées (à la Google Maps) de façon à pouvoir les exploiter, sur un ordinateur classique, sans être obligé de charger une image occupant plus de 8 Go sur le disque (et environ 50 Go de RAM une fois décompressée).

Les images suivantes (illustration 5) représentent une vue d'ensemble du maillage et l'une des 256 images qui aboutissent à l'image finale de 16 Giga-pixels (131072 x 131072 pixels), pour les deux types d'adaptation de maillage réalisés : homogène-isotrope et anisotrope-hétérogène. L'image a été générée en utilisant 8 192 cœurs de Joliot-Curie pendant une dizaine d'heures. Nous avons également réalisé des vidéos illustrant la navigation dans de telles images.

- générer des maillages 2D et 3D à plus de 1000 milliards d'éléments en utilisant la quasi-totalité la partition Skylake de Joliot-Curie (65 536 cœurs pour environ 200 To de RAM) ;
- tester le processus itératif d'adaptation de maillage anisotrope et hétérogène autour d'une microstructure complexe comportant plusieurs millions de fibres. Calcul exécuté sur 9 216 cœurs pour un maillage adapté de 3 milliards de nœuds et 18 milliards d'éléments, pour l'immersion de 3 millions de fibres ;
- produire des images en très haute résolution (16 Giga-pixels, soit 131072x131072 pixels) qui nous permettent de visualiser des maillages de très grande taille, tout en gardant tous les détails. Ces images ont été réalisées à l'aide du logiciel Visit, installé durant cette période sur Joliot-Curie, mais nécessitent d'utiliser presque autant de ressources que la création des maillages.

Tout ceci ne s'est pas fait sans difficultés. Plus particulièrement, pour les cas qui ont nécessité beaucoup de cœurs :

- des soucis d'hardware et/ou de software sur quelques nœuds particuliers ont engendré des crash ou de forts ralentissements lorsque ces nœuds ont été sélectionnés (probabilité proche de 1 sur 65 536 cœurs) ;
- un temps d'initialisation d'IntelMPI qui a explosé au-dessus de 4 096 cœurs (environ 30 minutes sur 65 536 cœurs).

Mais le principal problème a été l'impossibilité d'utiliser OpenMPI au-dessus de 16 384 cœurs, sans avoir des problèmes lors de l'exécution : des problèmes de précision numérique (semblaient cependant liés aux communications), mais également des "segfault" dans un appel à MPI_Wait(). Ces problèmes étaient assez aléatoires et pas toujours reproductibles. À l'heure actuelle, la raison n'est toujours pas connue mais il semblerait que ceci soit dû à une interversion de messages asynchrones (réception d'un message en lieu et place d'un autre). Un contournement du problème a été réalisé après la fin de ces Challenges en introduisant un appel à MPI_Barrier dans une fonction, contournement qui ne détériore pas les performances.

- Plusieurs perspectives s'offrent suite à ce travail :
- dans un premier temps, il serait intéressant de pouvoir tester nos algorithmes sur un nombre encore plus important de cœurs et pouvoir générer des maillages de plus grande taille de façon à poursuivre l'extensibilité de notre code vers les machines Exascale ;

6. Conclusions et perspectives

Cette session Grand Challenges a été l'occasion d'avoir un accès privilégié au nouveau supercalculateur Joliot-Curie du TGCC. Ces accès, juste après la mise en route de telles infrastructures, sont toujours spéciaux : il est en effet assez facile lors de cette phase d'avoir accès à la quasi intégralité du calculateur pour un seul utilisateur (chose plus difficile à réaliser lors de la mise en production, à l'heureuse exception des périodes de maintenance). Mais, ceci signifie aussi souvent une machine moins stable (problèmes d'hardware et software levés et corrigés au fil de l'eau par les équipes techniques). Ceci a pour conséquence principale de renforcer les liens entre le personnel des centres de calculs et leurs utilisateurs.

En ce qui concerne mon Grand Challenge, le succès fut au rendez-vous. Les objectifs fixés au début ont été atteints :



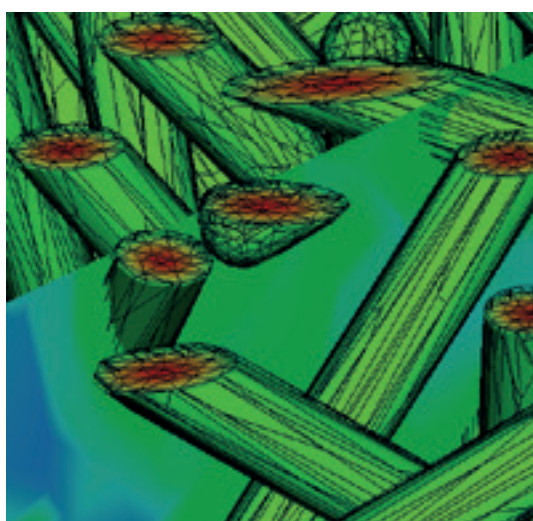
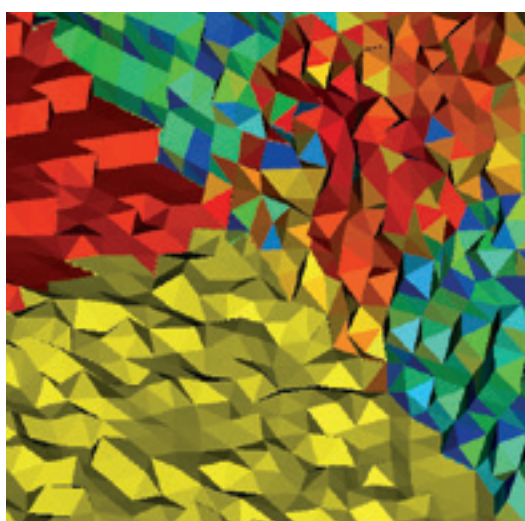
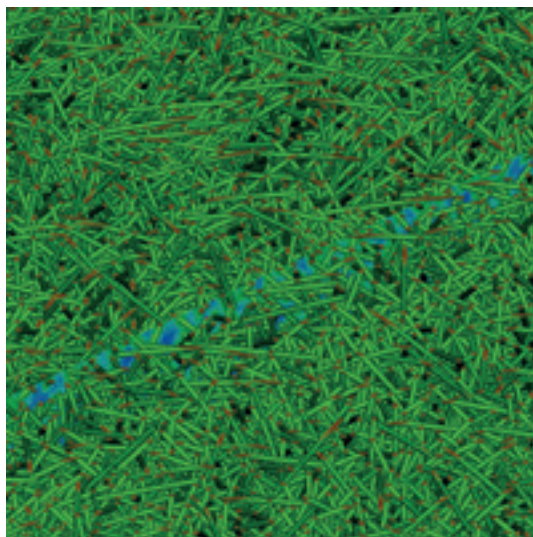
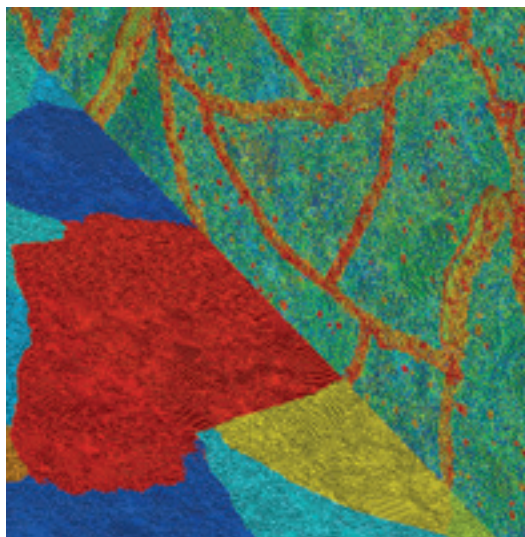
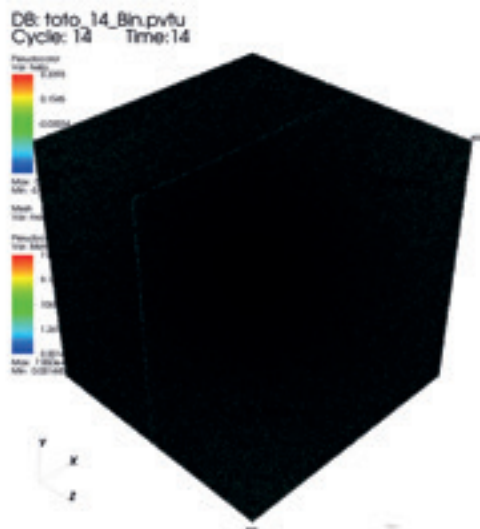
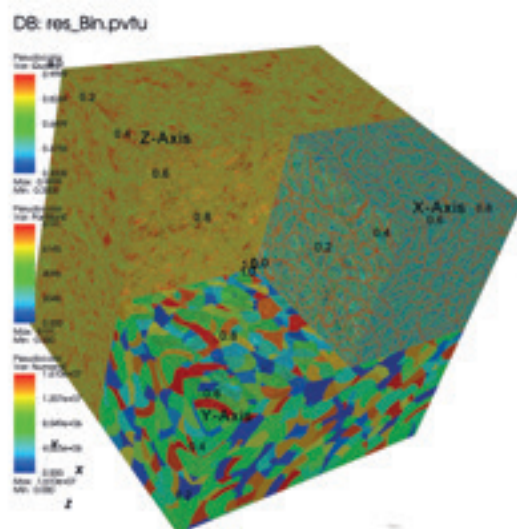


Illustration 5 : Visualisation des résultats à l'aide du logiciel VisIt. Images de très haute définition 16 Giga-pixels distribué en 256 sous images de 8192x8192 pixels. Les deux colonnes représentent les deux types d'adaptations effectuées : à gauche homogène-isotrope et à droite hétérogène-anisotrope. Les trois lignes représentent différents niveaux de zoom : une vue d'ensemble (zoom x1), une des 256 sous images (zoom x16) et un détail à taille "réelle" où un pixel de l'image égale un pixel de l'écran : 512x512 pixels (zoom x256).



- de même, il serait intéressant d'évaluer les performances de notre code sur d'autres types d'architectures: KNL, ARM voir GPU et/ou avec plusieurs niveaux de parallélisme thread, OpenMP, MPI. Des tests sont actuellement menés sur la partition KNL de Joliot- Curie ;
- il serait également envisageable de générer un maillage anisotrope adapté à une structure complexe comportant 100 millions de fibres qui représente environ le nombre de fibres présent dans une plaque injectée. **Cet objectif semble envisageable via un prochain projet PRACE avec un ordinateur environ deux fois plus puissant que Joliot- Curie ;**
- enfin, d'autres actions sont également à mener dans le domaine des solveurs, comme l'implémentation d'un solveur multi-grilles ou l'augmentation de l'ordre d'interpolation des éléments, toujours dans le but de réduire les temps de calcul pour une précision donnée, mais avec la volonté de cumuler toutes ces optimisations et non de les opposer.

■ 7. Remerciements

Je remercie chaleureusement GENCI et le TGCC pour m'avoir autorisé à participer à ces Grands Challenges sur Joliot-Curie. Un très grand merci également à l'équipe support du TGCC pour m'avoir accompagné lors de cette période, que ce soit sur les soucis techniques rencontrés ou l'installation de nouveaux logiciels tels que Visit. Un merci plus particulier à Loïc Madies et Vincent Ducrot qui ont été mes interlocuteurs privilégiés.

■ 8. Bibliographie

- [1] Hugues Digonnet, Thierry Coupez, Patrice Laure, and Luisa Silva. *Massively parallel anisotropic mesh adaptation*. The International Journal of High Performance Computing Applications, 0(0) :1094342017693906, 2017.
- [2] Luisa Silva, Thierry Coupez, and Hugues Digonnet. *Massively parallel mesh adaptation and linear system solution for multiphase flows*. International Journal of Computational Fluid Dynamics, 30(6) :431–436, 2016.
- [3] T. Laurencin, L. Orgeas, P.J.J. Dumont, S. Rolland du Roscoat, P. Laure, S. Le Corre, L. Silva, R. Mokso, and M. Terrien. *3d real-time and in situ characterisation of fibre kinematics in dilute non-newtonian fibre suspensions during confined and lubricated compression flow*. Composites Science and Technology, 134 :258 – 266, 2016.

■ Équipe projet



Hugues
DIGONNET





Équipe

Stepan TIMR, Astrid BRANDNER
and Fabio STERPONE

Laboratoire de Biochimie Théorique,
IBPC, CNRS UPR9080



Modelling a vesicle-based nano-reactor for confined enzymatic activity

4,8 millions d'heures

■ 1. Introduction

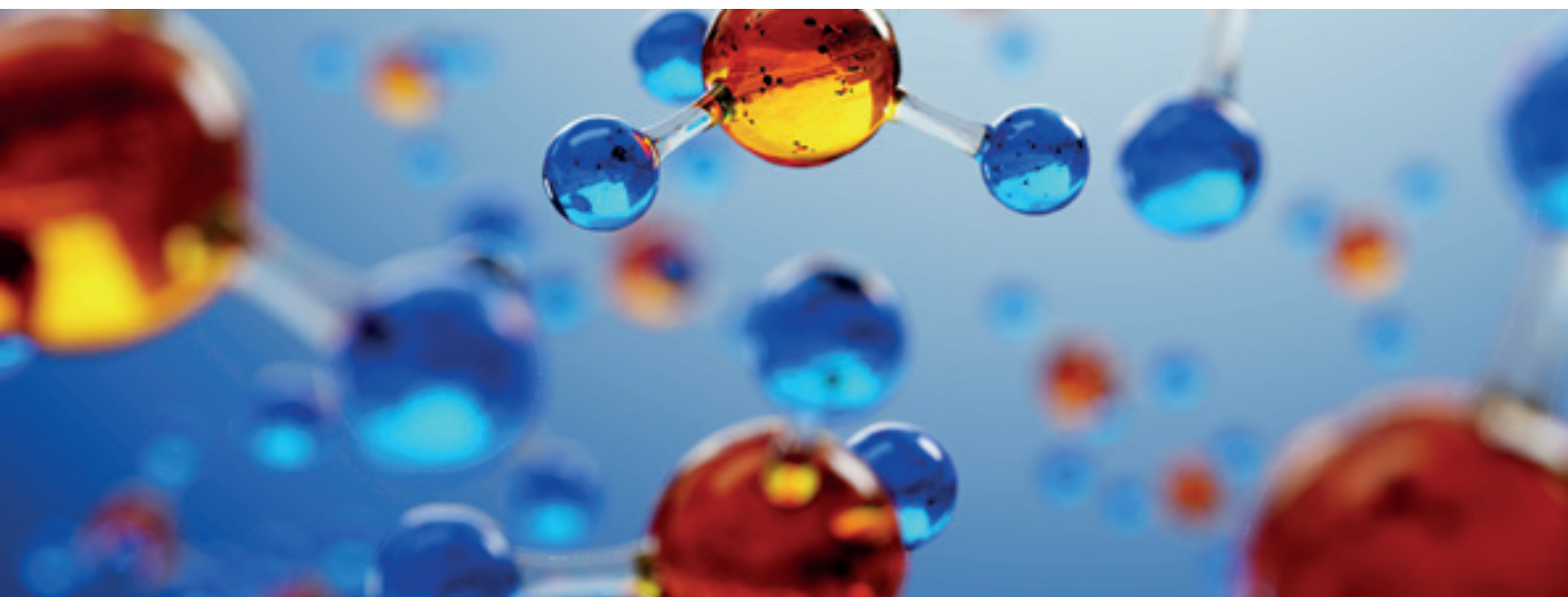
Vesicles filled with enzymes and substrates are frequently used to investigate biochemical reactions in confined conditions.[1] These biological nanoreactors can serve as simplified experimental models of compartments and organelles in living cells.[1,2] Moreover, novel vesicular constructs encapsulating therapeutic enzymes are being designed to replace malfunctioning organelles in the human body or to deliver drugs into cells. [2,3] For rational design of such constructs, it is **critical to understand how enzymatic reactions proceed when confined to crowded microscopic volumes**, i.e., in conditions which are very different from those in standard biochemical assays.

Several experimental works [4-6] demonstrated that confinement can significantly enhance the activity of enzymes. For instance, α -chymotrypsin enzymes confined in vesicles of varying sizes were found to exhibit increased catalytic efficiencies for decreasing vesicle radii (see Figure 1).[5] Moreover, one of the studies [6] pointed out a synergetic effect of confinement and macromolecular crowding on enzyme activity.

For an enzymatic reaction to take place, the enzyme must first meet a substrate molecule. Therefore, the rate of an enzymatic reaction can be strongly dependent on the probability of encounter between the substrates and the enzymes. This probability can be significantly modified by the confined and crowded conditions inside a nanoreactor. Thus, a possible explanation for the enhanced reactivity inside vesicular nanoreactors suggests an important role of the inner surface of the confining membrane, which might locally increase the chance of substrate-protein encounter. While this hypothesis has been supported by highly simplified theoretical models [7,8], a detailed molecular model which would provide an atomistic or near-atomistic picture of the diffusion in a vesicular nanoreactor has been lacking.

Here we employed a novel multi-scale computational approach [9] to simulate **the first detailed model of a vesicular nanoreactor, inspired by experimental constructs** [4,5]. In our approach we **combine the molecular dynamics and the Lattice Boltzmann techniques (LBMD)** to simulate molecular systems based on an

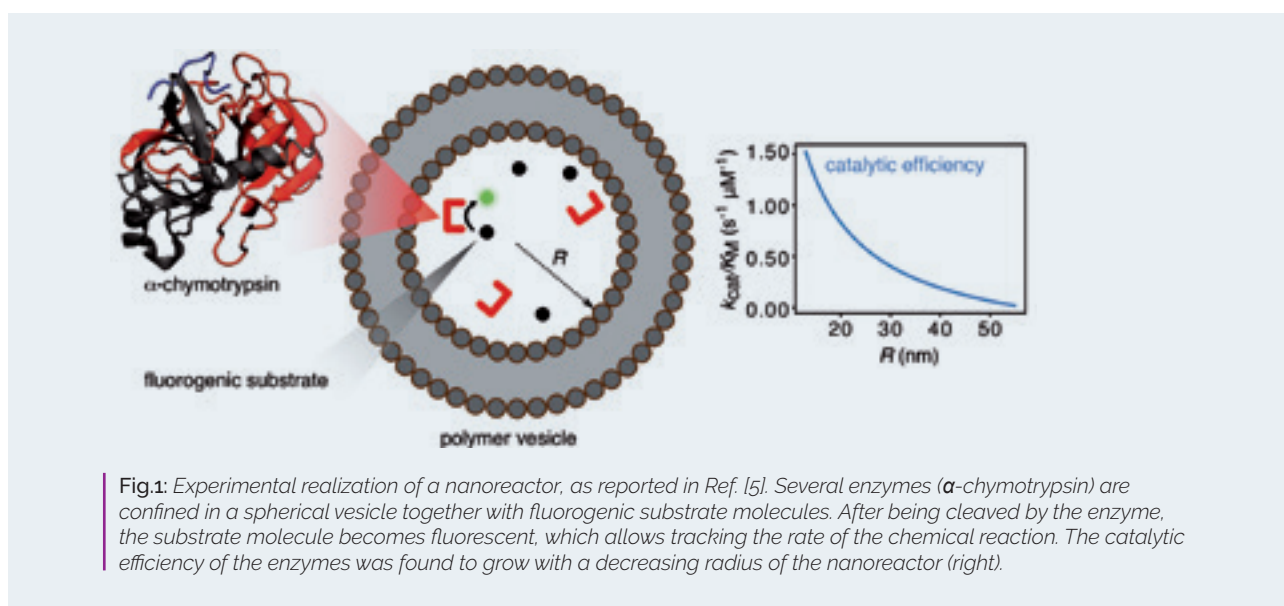




implicit solvent coarse-grained representation but including naturally hydrodynamic interactions.

We selected α -chymotrypsin as the enzyme to be placed into the nanoreactor. The nanoreactor had a form of a POPC lipid vesicle with the inner

radius of 12 nm placed in a cubic simulation box of size $L=41$ nm. The vesicle was filled with 5 α -chymotrypsin molecules in the presence of 6 protein crowders (bovine serum albumin - BSA), corresponding to a crowder concentration of 92 g/L, accompanied by 100 oligopeptide



molecules (alanyl-phenylalanyl-alanine), representing the substrate for the enzymatic reaction. Such a system would count 7 million atoms in the all-atom resolution, a size that is still prohibitive computationally, and for which a multi-scale methodology is preferable.

While in this report we will present the results only for this large final system, we stress that a large part of the allocation of the Grand Challenge was used to simulate a set of intermediate constructs designed to test several aspects of the methodology and of the system components, i.e. the role of attractive vs repulsive interactions between the membrane surface and the proteins, the effect of surface dynamics, the effect of the underlying methodology to generate stochastic dynamics.

■ 2. Method

In LBMD, the particle based dynamics is coupled to a kinetic representation of the solvent, simulated via the Lattice Boltzmann (LB) technique [10]. This technique allows including naturally hydrodynamics in implicit solvent coarse-grained molecular dynamics of biomolecules. The coupling between the particles and the solvent has the form of a Stokes-like drag force acting on a per-particle basis:

$$\vec{F}_i^D = -\gamma(\vec{v}_i - \vec{u}_i)$$

where $\vec{v}_i$ is the i -th particle velocity, $\vec{u}_i$ is the fluid velocity $\vec{u}$ smeared over a finite extension of the i -th particle, and γ is the frictional coupling, an adjustable parameter in the methodology.

The drag force adds up to the usual conservative forces derived from the Hamiltonian of the system, $\vec{F}_i^C = -\vec{\nabla}_i U(\{r\})$ and to a random white noise, $\vec{F}_i^R$, that represents thermal fluctuations. The reader can find more technical details in Ref. [9,11].

In our simulations, the LB evolution was solved by using the BGK (Bhatnagar-Gross-Krook) collisional operator, with a lattice grid spacing of ~ 3 Å (corresponding to a total of 137^3 grid points), sufficient to resolve local hydrodynamic interactions for macromolecular systems [9,11], and water kinematic viscosity ν_0 . The molecular and LB dynamics were evolved synchronously using a physical time-step of 10 fs, a value defined by the properties of the molecular force

fields. For each protein type the friction was optimised in order to reproduce the diffusivity in dilute conditions while for the lipid the diffusivity in a membrane.

■ 3. Model

Since for the problem considered in this challenge the internal flexibility of the proteins is not a necessary ingredient, the proteins were modelled using an elastic-network (EN) based on the coarse-grained model OPEP[12]. The EN allowed reducing the number of degrees of freedom and using a large time-step, e.g. 10 fs. The lipids in the vesicle were instead modelled using the Dry Martini force field [13]. We introduced a first tentative coupling among the OPEP and Dry Martini model by using existing potentials in the OPEP force field.

■ 4. MuPhy

All the simulations were performed using the code MuPhy [14], mainly developed by S. Melchionna (CNR, Rome), where several molecular models for bio-systems (proteins and lipids) were implemented by our group. MuPhy is a code for multi-scale simulations that integrates the LB with particle molecular dynamics and that can be used to simulate very disparate systems like proteins, DNA, polymer solutions, red-blood cells and bio-fluids in general. The code exhibits nearly ideal scalability for the LB component while the implicit-solvent modelling of sparse particle systems suffers from load-unbalancing, **a problem that thanks to the Grand Challenge we are now solving.**

■ 5. Results

First of all, our 300 ns simulation featuring a detailed model of a vesicular nanoreactor (see Fig. 2) allowed us to characterize the spatial distributions of the different molecules inside the nanoreactor. We observed that both the protein crowders as well as the substrate molecules were concentrated near the lipid membrane (Fig. 3A), with the large protein crowders diffusing laterally along the membrane surface. The



enzymes were also likely to interact with the lipid bilayer. However, rather than being in direct contact with the membrane surface, some of the enzymes preferred to bind to the crowder, which brought them somewhat closer to the center of the vesicle. The shape of the vesicle remained stable in the course of the simulation, with the maximum deviations not exceeding 4-6% of the vesicle size in each dimension.

For all the three molecular species, we observed a slowdown in diffusion compared to dilute conditions without confinement. At a short timescale (1-5 ns), this slowdown amounted to a 35 % decrease in the diffusion coefficient for the enzymes, to a 49% reduction for the protein crowders, and to a 39% decrease for the substrate molecules. At a longer timescale (10-50 ns), the slowdown in diffusion became even more pronounced (67%, 55%, and 57% for the enzymes, crowders, and substrates, respectively), starting to reflect the finite size of the vesicle interior and the interactions with the lipid membrane. At this timescale, the diffusion coefficient of the enzymes practically matched that of the significantly larger crowders (see Fig. 3B), as a consequence of their mutual interactions in the vicinity of the membrane. Thus, our simulation shows how the membrane surface in combination with the crowded environment modulates the diffusion coefficients inside a nanoreactor.

We have also calculated the probability that the enzymes are in contact with a substrate molecule

as a function of the distance from the center of the vesicle (Fig. 3C). The results demonstrate that enzymes located closer to the confining membrane are more likely to interact with the substrate and thus also more likely to catalyze the enzymatic reaction. This finding supports the hypothesis that the inner membrane surface plays an important role in controlling the reactivity of the nanoreactor. For this reason, tuning the physicochemical properties of the inner membrane surface seems to be an effective tool to design more efficient nanoreactor constructs.

In overall, the **results of our simulation confirm the capacity of the LBMD technique to provide a detailed molecular view of protein diffusion in large and realistic systems consisting of many biomolecules surrounded by a lipid membrane**. Therefore, the results **open the way for simulations of entire organelles or therapeutic constructs** to capture the complex biological processes occurring inside them.

The Grand Challenge allowed us also to individuate key parts of the MuPhy engine that need to be improved in view of simulations of larger systems that would require massive resources. The first problem concerns the paradigm of parallelisation. For systems where the fluid and particle components are not uniformly distributed in space, the standard domain decomposition approach exhibits

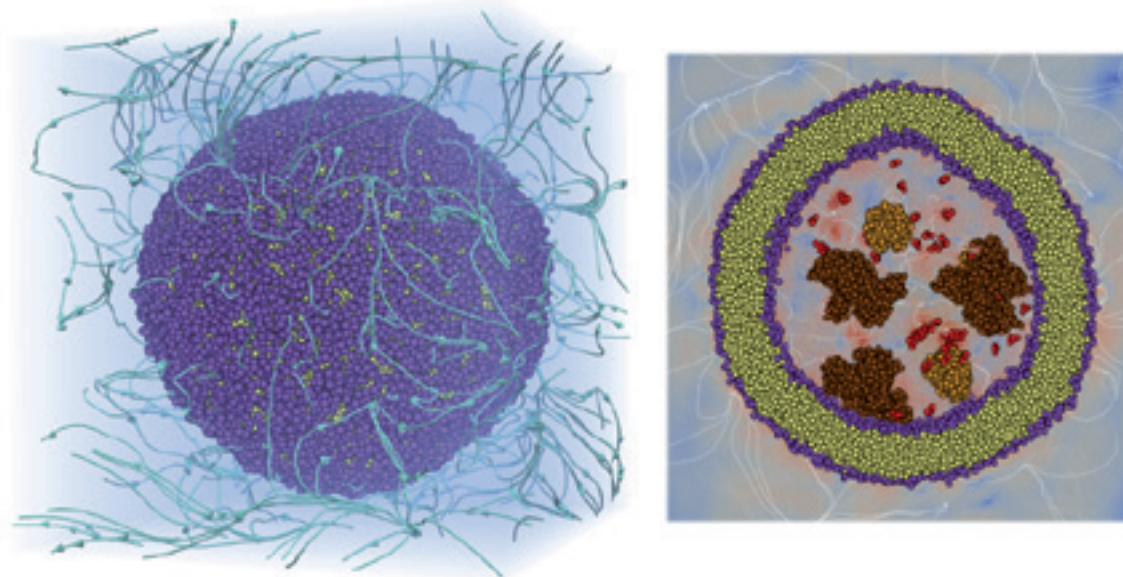


Fig.2: Snapshots from the LBMD simulation of a vesicular nanoreactor. (Left) View from the outside displaying fluid streamlines around the lipid vesicle. (Right) Cross section showing a cut through the lipid membrane (violet and yellow), the enzymes (orange), the crowder (brown), and the substrate molecules (red).



non-ideal scalability and alternatives must be considered, including multi-grid schemes, or non-uniform partitions. Secondly, the calculation of the forces for the particle component could benefit from in-depth recoding in order to exploit the potentiality of machine vectorisation.

Finally, we acknowledge the strong support from CINES HPC and GENCI team for the opportunity it gave us to move a step forward in our methodological development. The results of this challenge will be soon reported in a publication [15].

6. Bibliography

- [1] Kuechler, A.; Yoshimoto, M.; Luginbuhl, S.; Mavelli, F.; Walde, P. Enzymatic reactions in confined environments. *Nature Nanotechnology* 2016, 11, 409–420.
- [2] Godoy-Gallardo, M.; York-Duran, M. J.; Hosta-Rigau, L. Recent Progress in Micro/Nanoreactors toward the Creation of Artificial Organelles. *Advanced Healthcare Materials* 2017, 1700917, 1700917.
- [3] Bunker, A.; Magarkar, A.; Viitala, T. Rational design of liposomal drug delivery systems, a review: Combined experimental and computational studies of lipid membranes, liposomes and their PEGylation. *Biochimica et Biophysica Acta - Biomembranes* 2016, 1858, 2334–2352.
- [4] Chen, Q.; Schönherr, H.; Vancso, G. J. Block-copolymer vesicles as nanoreactors for enzymatic reactions. *Small* 2009, 5, 1436–1445.
- [5] Chen, Q.; Rausch, K. G.; Schönherr, H.; Vancso, G. J. α -chymotrypsin-Catalyzed Reaction Confined in Block-Copolymer Vesicles. *Chem. Phys. Chem.* 2010, 11, 3534–3540.
- [6] Baumann, P.; Spulber, M.; Fischer, O.; Car, A.; Meier, W. Investigation of Horseradish Peroxidase Kinetics in an "Organelle-Like" Environment. *Small* 2017, 13, 10–12.
- [7] Chiu, D. T.; Wilson, C. F.; Karlsson, A.; Danielsson, A.; Lundqvist, A.; Stromberg, A.; Ryttsen, F.; Davidson, M.; Nordholm, S.; Orwar, O. et al. Manipulating the biochemical nanoenvironment around single molecules contained within vesicles. *Chem. Phys.* 1999, 247, 133–139.
- [8] Chow, E.; Skolnick, J. Effects of confinement on models of intracellular macromolecular dynamics. *Proc. Natl. Acad. Sci. USA* 2015, 112, 14846–14851.

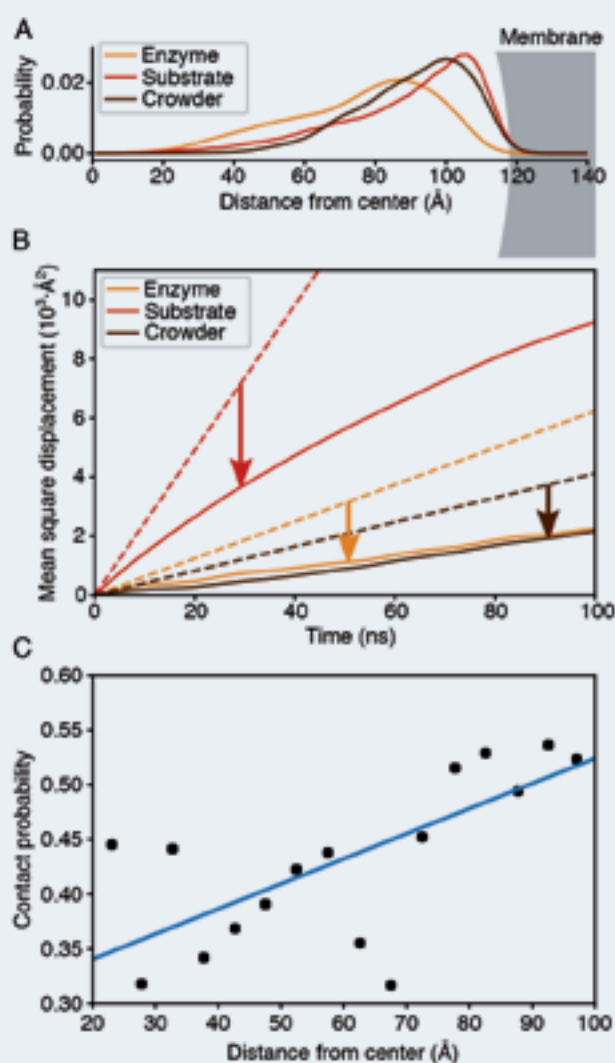


Fig.3: A) Spatial distributions of molecules inside the nanoreactor. (B) Mean square displacement of the different molecular species, with the dashed lines showing a comparison with dilute conditions in the absence of confinement. (C) The probability that an enzyme is in contact with a substrate molecule as a function of the displacement of the enzyme's center of mass from the center of the vesicle.



[9] Sterpone, F.; Derreumaux, P.; Melchionna, S. Protein Simulations in Fluids: Coupling the OPEP Coarse-Grained Force Field with Hydrodynamics. *J. Chem. Theory Comput.* 2015, 11, 1843–1853.

[10] Succi, S. *The Lattice Boltzmann Equation for Fluid Dynamics and Beyond*; Clarendon Press, 2001.

[11] Chiricotto, M.; Melchionna, S.; Derreumaux, P.; Sterpone, F. Hydrodynamic effects on b-amyloid (16–22) peptide aggregation. *J. Chem. Phys.* 2016, 145, 035102.

[12] Chebaro, Y.; Pasquali, S.; Derreumaux, P. The Coarse-Grained OPEP Force Field for Non-Amyloid and Amyloid Proteins. *J. Phys. Chem. B* 2012, 116, 8741–8752.

[13] Arnarez, C.; Uusitalo, J. J.; Masman, M. F.; Ingolfsson, H. I.; De Jong, D. H.; Melo, M. N.; Periolo, X.; De Vries, A. H.; Marrink, S. J. Dry martini, a coarse-grained force field for lipid membrane simulations with implicit solvent. *J. Chem. Theory. Comput.* 2015, 11, 260–275.

[14] Bernaschi, M.; Melchionna, S.; Succi, S.; Fyta, M.; Kaxiras, E.; Sircar, J. MUPHY: A parallel MULTi PHYsics/scale code for high performance bio-fluidic simulations. *Comput. Phys. Comm.* 2009, 180, 1495–1502.

[15] Brandner, A.; Timr, S.; Melchionna, S.; Derreumaux, P.; Baaden, M.; Sterpone, F. Modelling Lipids Systems in Fluids: Coupling the Dry Martini Force Field with Lattice Boltzmann Molecular Dynamics.

■ Équipe projet



Stepan
TIMR



Astrid
BRANDNER



Fabio
STERPONE



Équipe

Jean-Philip PIQUEMAL, Laboratoire de Chimie Théorique,
UMR 7616 CNRS, Sorbonne Université, Paris.



Optimisation du code Tinker-HP

15 millions d'heures

Les différents sous-projets réalisés dans le cadre de ce Grand Challenge GENCI ont été réalisés grâce au logiciel Tinker-HP [1], logiciel de dynamique moléculaire de nouvelle génération dédié aux méthodes de champs de forces polarisables. Comme proposé dans le projet déposé, nos efforts se sont concentrés autour de deux axes : i) optimisation des performances du code et, ii) applications à grande échelle de la dynamique moléculaire de champs de forces polarisables.

■ 1. Optimisation du code Tinker-HP

Cette tâche initiale a visé à tirer le meilleur parti de la nouvelle architecture disponible sur la machine Irène Joliot-Curie et a pour cela été scindée en deux sous parties :

1.1 Optimisation algorithmique : l'importance des pas de simulations « longs » en dynamique moléculaire

Notre accès à la machine Irène, s'est ouvert alors que nous étions en pleine recherche d'une amélioration algorithmique permettant d'augmenter la taille des pas en temps (*timestep*) utilisés en dynamique moléculaire. En effet, en pratique, afin de pouvoir s'attaquer à de très

grands systèmes d'intérêt biologique, la façon la plus simple pour accélérer la dynamique moléculaire est d'utiliser des pas de temps les plus grands possibles. Dans ce contexte, les approches à pas multiples (*multi-timestep*) ont émergé, mais le pas maximum utilisable est rapidement limité par des effets de résonance qui dégradent très rapidement la validité physique de la simulation. Certains auteurs ont suggéré qu'il était possible de surmonter ces effets à l'aide de dynamiques modifiées capables d'échantillonner correctement la mesure, mais ces solutions, bien que capables de fortement accélérer les simulations, modifient toutes les propriétés dynamiques (par exemple la vitesse de diffusion dans un solvant). Notre travail a visé à offrir cette accélération tout en préservant la qualité des dynamiques moléculaires et garantissant la





précision des propriétés statiques et dynamiques obtenues par simulation. Cette recherche s'est faite dans un contexte particulier, celui des champs de forces polarisables qui nécessitent un coût de calcul allant de 5 à 10 fois celui de la dynamique moléculaire par champ de forces classiques, non-polarisables et sont, de facto, très gourmands en heures de calcul. La mise au point d'un tel algorithme serait très important pour GENCI permettant de sauver des millions d'heures de calculs qui pourraient être ainsi dédiées à des simulations beaucoup plus précises car « polarisables » de systèmes biologiques de plus grande taille sur des échelles de temps plus grandes. Au final, **ce projet a été un succès** car nous avons vu **développer une nouvelle approche « multi-timestep » basée sur la dynamique de Langevin permettant d'accélérer la dynamique moléculaire de dipôles polarisables tout en préservant l'efficacité de l'échantillonnage**. Le nouvel intégrateur ainsi développé et nommé BAOAB-RESPA1 permet un gain très important (facteur 7 par rapport aux approches standards avec un pas de 1fs) en temps de calcul et permet de se rapprocher des performances des champs de forces classiques tout en conservant des simulations polarisables classiques.

L'ensemble de ces résultats a été publié dans la revue à haut impact *Journal of Physical Chemistry Letters*.

Publication

Pushing the limits of Multiple-Timestep Strategies for Polarizable Point Dipole Molecular Dynamics. L. Lagardère, F. Aviat, J.-P. Piquemal, *J. Phys. Chem. Lett.*, 2019, 10, 2593–2599 DOI: 10.1021/acs.jpcclett.9b00901

1.2 Optimisation HPC de Tinker-HP : AVX 512 et collaboration avec Intel

Cette partie « informatique HPC » du travail a visé à développer avec le concours de la société Intel (collaboration Intel : Alejandro Duran), une version de Tinker-HP complètement dédiée aux processeurs Skylake et permettant d'utiliser au mieux les jeux d'instructions de vectorisation (AVX512). Le projet initial était de définir 2 versions du logiciel, l'une dédiée à la partition Xeon-Phi, l'autre à la partition Skylake mais l'abandon par Intel de la famille de processeurs Xeon-Phi nous a forcé à arrêter prématurément le développement de la version Xeon-Phi pour nous concentrer sur la partition Skylake



(partant du principe que la vectorisation Skylake serait aussi une source de gains importants sur les Xeon-Phi). Cette optimisation a donné ses premiers fruits à mi-challenge et permet dorénavant une accélération des toutes nos méthodes avec un gain moyen de 35% du directement à la vectorisation, gain pouvant aller encore plus loin pour les très gros systèmes biologiques. Le nouveau logiciel est ainsi complètement vectorisé et tire partie pleinement de l'architecture de Joliot-Curie offrant des

taux de vectorisation AVX512 allant jusqu'à 85%. La Figure 1 montre le gain en performance de la vectorisation sur un système de grande taille : le virus Satellite Mosaïque du Tabac (STMV, 1066628 atomes) traité avec le champ de forces AMOEBA [2] pour lequel nous avons poussé Joliot-Curie jusqu'à 16000 cœurs. Sur de

gros systèmes très gourmands en heures de calculs, la vectorisation AVX512 activée en mode parallélisme massif sur des processeurs de types Skylake permet un gain de 59% par rapport à la version Skylake vectorisé.

Nous avons bien sûr combiné ces avancées de vectorisation avec les avancées algorithmiques portant la production par jour du STMV à près de 2ns/jour (à comparer avec les 0,45 ns/jour maximum obtenue initialement sur la machine Occigen du CINES avec divers algorithmes d'accélération). En pratique entre accélération algorithmique et optimisation HPC, **le code Tinker-HP a accéléré de l'ordre d'un ordre de grandeur (X 10 !) par rapport à nos meilleurs résultats obtenus au GENCI avant l'arrivée de Joliot-Curie.**

Ces travaux qui seront co-publiés avec Intel (embargo jusqu'à publication), sont en cours d'évaluation par une revue scientifique.

Publication

Raising the Performance of the Tinker-HP Molecular Modeling Package on Intel's HPC Architectures: a Living Review [Article v1.0]. L. H. Jolly, A. Duran, J. W. Ponder, P. Y. Ren, L. Lagardère, J.-P. Piquemal, 2019, soumis

2. Applications

Ces nouvelles avancées ont été utilisées pour commencer à démontrer la faisabilité à grande échelle de simulations de dynamiques moléculaires utilisant des champs de forces polarisables.

Pour cela, nous avons comparé des simulations de type *Steered Molecular Dynamics* (SMD) utilisant des champs de forces polarisables (AMOEBA) et non polarisables (AMBER, CHARMM, OPLS). L'approche SMD est une technique puissante capable d'accélérer les événements rares dans les simulations de dynamique moléculaire en appliquant une force extérieure à un ensemble d'atomes choisis. Malgré le fait que cette approche génère des simulations hors-équilibre, la dynamique SMD reste capable de reconstruire les propriétés d'équilibre tels que le Potentiel de Force Moyenne (PMF). Pour permettre ces études, la méthodologie SMD a été mis en place dans le cadre massivement parallèle du logiciel Tinker-HP permettant des simulations polarisables et non polarisables de grosses protéines. Pour valider cette nouvelle implémentation, nous avons comparé tout d'abord les résultats SMD de Tinker-HP à la littérature (NAMD). Les tests ont été réalisés sur trois systèmes de référence différents : le deca M-A-alanine (112 atomes), l'ubiquitine (9737 atomes) et le complexe CD2CD58 (97594 atomes). Pour chacune des simulations, les PMF ont été reconstruits et comparés en termes de barrière d'énergie libre et de fluctuations des liaisons hydrogènes au fil du temps. Comme on pouvait s'y attendre, de fortes divergences sont remarquées entre les modèles polarisables et non-polarisable lorsque considérés dans le vide où les champs de forces classiques, alors que les résultats sont plus comparables lorsqu'un environnement aqueux est ajouté (systèmes solvatés). Toutefois, pour le plus grand des systèmes, c'est-à-dire le CD2CD58, de fortes différences liées à la modélisation d'un pont salin ont été remarqués lorsque des champs de forces classiques sont utilisés ce qui met en lumière l'importance de prendre en compte la polarisation dans de tels calculs. En deux mots, les calculs sur le complexe CD2CD58 visaient à briser le pont salin et à mesurer les différences d'énergies entre les états avant et après la rupture mécanique.

Nous avons pu montrer que la rupture s'opérait selon 2 mécanismes différents selon que le champ de force AMOEBA ou CHARMM étaient



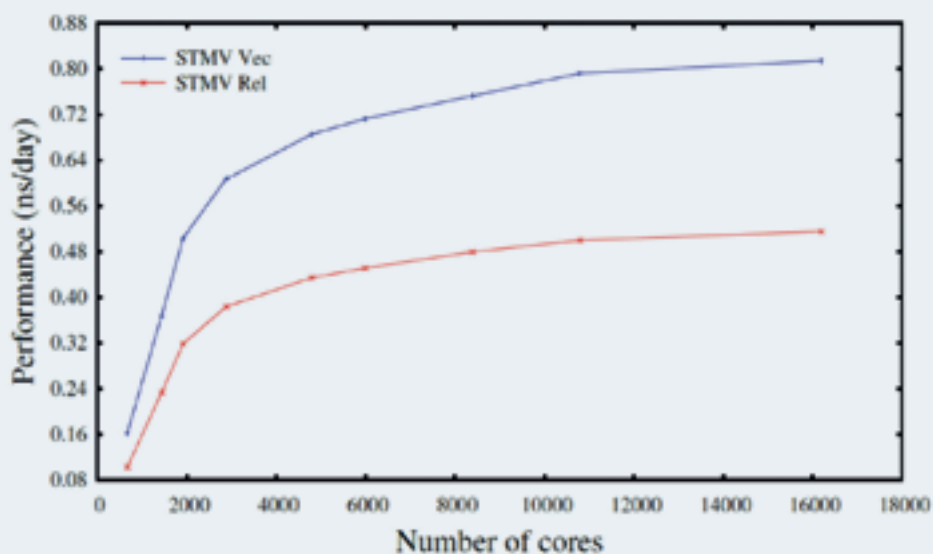


Figure 1 : production (ns/jour) du logiciel Tinker-HP avec le champ de forces AMOEBA

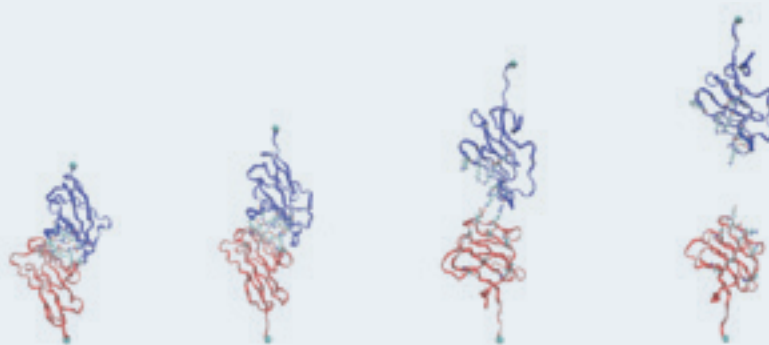


Figure 2 : destruction du point salin du complexe CD2-CD58

utilisés, les valeurs de PMF d'AMOEBA étant en bien meilleur accord avec l'expérience. Ces expériences numériques de grande taille sont inédites et très importantes car elles permettent de comparer et de mieux comprendre la physique des modèles de champs de forces employés afin de dégager des méthodes prédictives. Notre travail ouvre de larges perspectives en biologie structurale pour l'utilisation de méthodes de champs de forces polarisables.

Ce travail a donné lieu à une publication dans un journal à fort facteur d'impact.

Publication

Massively parallel implementation of Steered Molecular Dynamics in Tinker-HP: polarizable versus non-polarizable simulations. F. Célerse, L. Lagardère, E. Derat, J.-P. Piquemal, *J. Chem. Theory. Comput.* 2019, en ligne.

DOI: 10.1021/acs.jctc.9b00199

Raising the Performance of the Tinker-HP Molecular Modeling Package on Intel's HPC Architectures: a Living Review [Article v1.0].

L. H. Jolly, A. Duran, L. Lagardère, J. W. Ponder, P. Y. Ren, J.-P. Piquemal, 2019, soumis, preprint disponible sur ArXiv

<https://arxiv.org/abs/1906.01211>



■ 3. Références

[1] Tinker-HP: a Massively Parallel Molecular Dynamics Package for Multiscale Simulations of Large Complex Systems with Advanced Polarizable Force Fields.

L. Lagardère, L.-H. Jolly, F. Lipparini, F. Aviat, B. Stamm, Z. F. Jing, M. Harger, H. Torabifard, G. A. Cisneros, M. J. Schnieders, N. Gresh, Y. Maday, P. Ren, J. W. Ponder, J.-P. Piquemal, *Chem. Sci.*, 2018, 9, 956–972 (Open Access), DOI: 10.1039/C7SC04531J

[2] a) Polarizable Atomic Multipole-Based AMOEBA Force Field for Proteins.

Y. Shi, Z. Xia, J. Zhang, R. Best, C. Wu, J. W. Ponder, P. Ren, J. *Chem. Theory Comput.*, 2013, 9, 4046–4063; DOI: 10.1021/ct4003702; b) AMOEBA Polarizable Atomic Multipole Force Field for Nucleic Acids. C. Zhang, C. Lu, Z. Jing, C. Wu, J.-P. Piquemal, J. W. Ponder, P. Ren, J. *Chem. Theory. Comput.*, 2018, 14, 2084–2108 DOI: 10.1021/acs.jctc.7b01169





Équipe

Fabrizio CREAZZO, Giuseppe CASSONE, Daria RUTH GALIMBERTI,
Simone PEZZOTTI, Jiri SPONER and Marie-Pierre GAIGEOT

LAMBE UMR8587,

Université d'Evry val d'Essonne,

Université Paris Saclay,

91025 Evry, France

Oxygen Evolution Reaction at the (110)-Co₃O₄ surface by collective contact matrix metadynamics DFT-MD simulations

600 000 heures

The need for environmental friendly and large scale sustainable energy conversion and storage have triggered renewed interest into the electrochemical and photo-electrochemical splitting of water into hydrogen and oxygen¹, where a key challenge in this field is the development of economically viable electro-catalyst materials for the oxygen evolution reaction (OER) $2\text{H}_2\text{O} \rightarrow \text{O}_2 + 4\text{e}^- + 4\text{H}^+$. Transition metal oxide catalysts, in particular cobalt (hydr)oxide phases, have received large attention because they meet these requirements.²⁻⁴ The spinel Co₃O₄ magnetic semiconductor is one promising anode material for the electrochemical OER and is the material we are investigating with theoretical simulations to that end, in collaboration with electro-chemical experiments performed in the Allongue's group at the Ecole Polytechnique.

Most of the existing first-principles calculations published in the literature on the water splitting on the Co₃O₄ oxide catalyst (and other materials) are "surface science" calculations⁵⁻⁸, consisting in static DFT calculations without considering the presence of the electrolytic water solution in contact with the oxide material at finite temperature. We have recently revealed the structural composition and speciation of the

(110)-Co₃O₄ surface put in contact with liquid water, how the water is organized at the interface with the oxide material, and how/why water can be found in undisciplined behaviors, by DFT-based molecular dynamics simulations (DFT-MD).⁹ These simulations were done through the GENCI for access to the French national supercomputers.

Within the GENCI Challenges on the new TGCC Joliot Curie machine, we have been one step further and have started more challenging DFT-MD simulations to reveal the mechanisms and associated energetics of the OER at gas phase and aqueous (110)-Co₃O₄ surfaces. To that end, biased DFT-MD simulations have been performed within the framework of novel "contact matrix" biased metadynamics¹⁰, and we present hereafter some of the results obtained at the surface in contact with the air where one water molecule splits into O₂. In particular, our investigation reveals that the hydroxylated $\mu\text{-oxo-Co}^{3+}$ surface sites are active catalytic sites leading to low energy OER paths, with a free energy barrier for the reaction to occur at roughly 0.25 eV. See Figure 1b. This is the reference energy to observe the water splitting at the (110)-Co₃O₄ surface in contact with the air.





The mechanism for the water splitting found at the (110)- Co_3O_4 surface in contact with the air is shown in figure 1a. The multiple-steps reaction starts with the water attack and its dissociation at the μ -oxo- Co^{3+} surface site to form an active

radical group $\text{Co}^{3+}\text{-OOH}$, while the dissociated hydrogen of the water is protonating a nearby $\text{Co}^{3+}\text{-O}$ surface site. Once the active radical group $\text{Co}^{3+}\text{-OOH}$ is formed, our biased DFT-MD simulations show a proton jumping from this

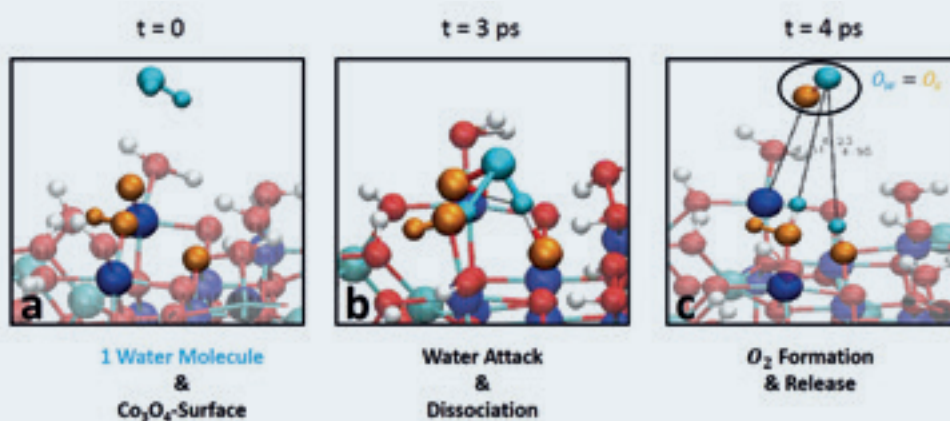


Figure 1a: Collective contact matrix metadynamics DFT-MD simulations showing the OER at the (110)- Co_3O_4 cobalt oxide surface in contact with the air. The water that is splitting at the surface is colored in blue and the surface sites involved into the whole chemical reaction process are colored in orange.



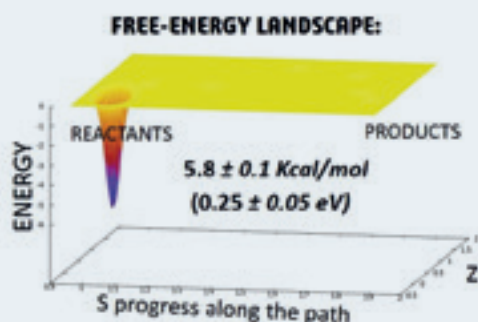


Figure 1b: Associated free energy landscape of the water splitting at the (110)-Co₃O₄ cobalt oxide surface in contact with the air. The 0.25 eV barrier has to be converted into a potential to apply in electrochemical conditions.

group to its neighboring $\mu_2\text{O}$ surface inner-site, leaving $\text{Co}^{3+}\text{-OO}$ behind. Roughly one picosecond after the initial water attack at the cobalt oxide surface, a double-bonded O_2 desorbs from the $\mu\text{-oxo-Co}^{3+}$ surface site, leading to the expected O_2 formation in the OER.

The same DFT-MD metadynamics for the water splitting are currently running for the (110)-Co₃O₄ surface in contact with liquid water, in order to compare the involved energetics and hence understand at which conditions electrochemical experiments would need to be performed for observing the water splitting at this semiconductor oxide. More DFT-MD are running by including electrolytes and 'real' electrochemical conditions within the theoretical set-up, in order to compare energetics and mechanisms of the OER in more realistic theoretical electrochemical conditions.

These are essential first steps into getting a microscopic knowledge into the OER at cobalt-oxide

resources from GENCI-France Grant No. 072484 (CINES/IDRIS/TGCC) and GENCI Challenge TGCC-Joliot Curie.

This work is part of a collaborative consortium within the LABEX CHARMA3T "Excellence Laboratory" program with Professor Ph. Allongue and F. Maroun at the Ecole Polytechnique, University Paris-Saclay.

Bibliography

- [1] Lewis, N. M. and Nocera, D. G., PNAS, 2006, 103, 15729.
- [12] Kanan, M. W. and Nocera, D. G., Science, 2008, 321, 1072-1075.
- [3] Deng, X. and Tüysüz, H., ACS Catalysis, 2014, 4, 3701-3714.
- [4] Suen, N. T., Hung S. F., Quan, Q., Zhang, N., Xu, Y. J., Chen, H. M., Chemical Society Reviews, 2017, 46, 337-365.
- [5] Wu X., Selloni A., Phys. Rev. B, 2011, 83.
- [6] Chen J. and Selloni A., J. Phys. Chem. Lett., 2012, 3.
- [7] Chen J. and Selloni A., J. Phys. Chem. C, 2013, 117.
- [8] Selcuk S. and Selloni A., J. Phys. Chem. C, 2015, 119.
- [9] F. Creazzo, D. R. Galimberti, S. Pezzotti, and M. P. Gaigeot, J. Chem. Phys., 2019, 150, 4. Chosen as Feature Article by JCP, Chosen as 2018 JCP Editor's Choice, and highlighted by AIP
SciLight at <https://aip.scitation.org/doi/full/10.1063/1.5085400>
- [10] Cassone G., Sponer J., Sponer J. E., Pietrucci, F., Saitta, A. M., Saija F., Chemical Communications, 2018, 54(26).

Acknowledgements

This work was performed under Grant No. LABEX CHARMA3T 11-LABX-0039/ANR-11-IDEX-0003-02 "Excellence Laboratory" program of the University Paris-Saclay and Grant No. ANR DYNWIN ANR-14-CE35-0011-01 (ANR Agence Nationale de la Recherche), and using HPC





Équipe

Jean-Baptiste CHARRAUD,
doctorant, CEA/DAM

Marc TORRENT, chercheur CEA/DAM

Grégory GENESTE, chercheur CEA/DAM



Recherche de structures dans des hydrures sous haute pression

10 millions d'heures

Le but du projet est de mettre en place la méthode de recherche aléatoire de structures, couplée au code de calcul *ab initio* ABINIT. ABINIT est un logiciel scientifique simulant les propriétés physiques des matériaux à partir du comportement des électrons. Il utilise pour ce faire des algorithmes s'appuyant sur des équations de la physique quantique. L'algorithme de recherche aléatoire de structures consiste en une **génération aléatoire de structures initiales, suivie d'une optimisation structurale**. Il a d'abord été mis en œuvre et testé sur des hydrures de fer-aluminium sous pression (dans le domaine des 50-100 GPa), a permis de proposer des structures cristallines possibles, et de montrer que la stœchiométrie en hydrogène augmente avec la pression dans ces systèmes.

■ 1. Contexte

1.1 Les hydrures sous pression : une voie possible vers l'obtention de supraconducteurs conventionnels à haute température critique.

L'étude des hydrures sous haute pression est un domaine de recherche en plein essor. Il résulte de l'idée initiale, formulée par N. Ashcroft en 2004 [1], que **certaines propriétés de l'hydrogène métallique pourraient être obtenues dans des hydrures à des pressions plus basses que celles nécessaires à l'obtention de cet hypothétique composé**. L'hydrogène est supposé devenir métallique autour de 450-500 GPa, une **pression**

très difficilement atteignable avec les moyens expérimentaux actuels, mais la théorie prédit que ce système devrait être un supraconducteur conventionnel avec une température critique très élevée, peut-être supérieure à la température ambiante [2,3]. Depuis 2004, de nombreuses équipes, suivant l'intuition de Ashcroft, cherchent à vérifier si cette propriété exceptionnelle pourrait se manifester à des pressions plus basses – et donc plus facilement atteignables en laboratoire – dans des hydrures sous pression (typiquement dans le domaine des 100-200 GPa).

Un hydrure est l'association de l'hydrogène avec un élément moins électronégatif (typiquement un métal) : l'hydrogène y est chargé négativement. A pression ambiante,





de nombreux hydrures existent, par exemple des hydrures d'alcalins (ex : NaH) ou d'alcalino-terreux (ex : MgH_2). Ces systèmes sont par ailleurs pressentis comme possibles voies de stockage solide de l'hydrogène. Les nombreuses études, depuis 2004, de ces systèmes sous pression ont montré que la concentration en hydrogène a tendance à augmenter dans les hydrures lorsque la pression augmente, donnant naissance à des stœchiométries inédites. Prenons l'exemple de l'hydruration du fer. A pression ambiante, le fer et l'hydrogène ne réagissent pas et ne forment pas de composé. A une pression de quelques GPa, un premier hydrure de formule FeH se forme, puis à plus haute pression, FeH_2 et FeH_3 . Enfin, à une pression de 130 GPa, un composé de formule FeH_5 apparaît [4]. Ces systèmes très riches en hydrogène, stables sous haute pression uniquement, et dans lesquels l'hydrogène adopte une forme atomique (et non moléculaire) sont appelés « polyhydrides » ou « superhydrides ».

Les recherches actuelles se focalisent sur la famille des hydrures de terres rares, qui sous pression, forment des composés exceptionnellement riches en hydrogène (ex : LaH_{10}), qui adoptent des structures cristallines particulières qualifiées de « clathrates

d'hydrogène » [5] : les atomes d'hydrogène forment des sortes de cages autour des atomes métalliques, de sorte qu'un réseau d'hydrogène, possédant des points communs forts avec l'hydrogène métallique, apparaît. Ces systèmes semblent aujourd'hui les plus prometteurs pour l'obtention d'un supraconducteur conventionnel à haute température critique.

1.2 Prédiction des structures cristallines —

Dans cette recherche, le calcul *ab initio* basé sur la Théorie de la Fonctionnelle de la Densité (DFT) joue un rôle très important. La DFT est une théorie permettant de calculer l'énergie et la densité électronique de l'état fondamental d'un gaz d'électrons plongé dans un potentiel externe (le potentiel électrostatique engendré par les noyaux). Elle rend ainsi possible la simulation des propriétés des matériaux avec une très grande précision. D'une part, la DFT permet d'identifier les systèmes les plus propices à la formation de superhydrides et de prédire leur possible supraconductivité, ce qui permet d'aiguiller les expérimentateurs dans la bonne voie. Ainsi, c'est la DFT qui a permis de proposer les hydrures de terres rares comme de très bons candidats à



l'obtention de supraconducteurs haute T_c . D'autre part, les expériences usuelles seules ne sont pas capables de caractériser complètement la structure des hydrures sous pression – et donc de déterminer sans ambiguïté leur stœchiométrie en hydrogène – car la diffraction des rayons X ne permet pas de « voir » l'hydrogène (les éléments de faible numéro atomique possèdent peu d'électrons et diffusent très peu les rayons X). Seule la DFT est capable de fournir une telle information, et de décrire un large éventail de propriétés et de prédire certaines quantités (ex : enthalpies de formation de hydrures, volume des composés en fonction de la pression, modes de vibration, magnétisme, etc).

Néanmoins, pour être efficace, la DFT doit être couplée à une méthode permettant de prédire à l'avance la structure cristalline du composé. En d'autres termes, la DFT doit être vue dans ce contexte comme une brique essentielle d'un algorithme de prédiction de structures cristallines. La méthode de recherche aléatoire de structures, proposée par C. J. Pickard et R. J. Needs [6], répond donc à cet objectif. Elle a permis de prédire la structure de nombreux composés sous pression. Par exemple, cette méthode a permis d'élucider la structure de la phase III de l'hydrogène dense. La recherche aléatoire de structures consiste à générer, en utilisant un processus aléatoire, des structures de départ. Ce processus peut notamment s'appuyer sur des biais issus d'arguments physiques ou de données expérimentales connues. Ces configurations initiales qui correspondent à des « images » du matériau étudié sont alors données en entrée d'un « algorithme d'optimisation structurale ». Cet algorithme utilise la DFT pour calculer et minimiser l'énergie de ces images en changeant les positions des atomes et la maille cristalline. Plus une « image » est trouvée stable en DFT, plus grande est la probabilité qu'elle apparaisse dans les expériences. Si la maille primitive du cristal ne contient pas un nombre trop élevé d'atomes (typiquement jusqu'à 20-30 atomes), on finit par tomber sur la structure la plus stable, qui généralement possède un bassin d'attraction assez large dans l'espace des configurations.

1.3 Objectifs scientifiques

L'objectif de ce projet est (i) la mise en œuvre de la méthode de recherche aléatoire sur un hydrure sous pression, et (ii) son implantation directe au sein du code de calcul DFT ABINIT. Après un certain nombre de tests de la méthode,

nous avons choisi de démarrer sur des hydrures de fer-aluminium de formule FeAlH_x et Fe_3AlH_x .

2. Code, moyens de calcul, simulation / méthode numérique

2.1 Nom du code / projet

Nous utilisons le code de calcul DFT ABINIT (<http://www.abinit.org>), qui est en grande partie développé dans notre laboratoire.

ABINIT est une suite logicielle *open-source* (licence GNU-GPL) – d'environ 800 000 lignes de code Fortran – développée dans le cadre d'un projet collaboratif international à grande visibilité, incluant des contributions de plusieurs laboratoires, essentiellement européens et nord-américains. ABINIT est utilisé intensivement, tant pour les travaux de recherche et de développement méthodologique, que pour la production de données en Physique de la Matière Condensée, à l'échelle microscopique. Il s'agit d'un code basé sur les « premiers principes » (modèles *ab initio*); les calculs sont effectués sans aucun ajustement sur des données expérimentales. Les systèmes, constitués de noyaux atomiques et d'électrons, sont étudiés à l'échelle microscopique, dans le formalisme de la DFT. Plusieurs types de calculs peuvent être effectués, notamment rechercher les propriétés structurales et thermodynamiques des matériaux, ainsi que leur réponse à une sollicitation. Il est possible d'effectuer des calculs statiques mais aussi de la dynamique moléculaire ou de relaxer des structures atomiques.

Le cœur d'un code de calcul dans le formalisme DFT est un algorithme itératif de diagonalisation. Il s'agit de déterminer les vecteurs propres d'un opérateur (dont nous n'avons pas de représentation matricielle) associées aux plus petites valeurs propres.

Dans le cadre de l'étude présentée ici, nous avons utilisé la version 8.10 de ABINIT. Cette version a été spécifiquement adaptée aux supercalculateurs à architecture *many-core*, et plus spécifiquement aux processeurs de type KNL de Intel, qui sont ceux qui ont été utilisés ici. En effet, le code tire parti efficacement d'un parallélisme hybride *MPI-openMP*, notamment grâce à une implémentation « *coarse grained* » du parallélisme multi-tâche – qui a demandé



une restructuration importante du code – , afin de fournir une quantité de travail importante aux différentes tâches (*threads*).

Enfin, notons que, pour optimiser les performances du code sur un plus grand nombre de supercalculateurs, nous avons *externalisé* au maximum les opérations mathématiques « élémentaires » en faisant appels à des bibliothèques externes, que ce soit celles fournies par les constructeurs (ici ATOS-Bull) ou des libraires spécifiquement développées pour les calculs à l'échelle atomique.

2.2 Configuration

La version du code utilisée ici a été installée par le support informatique du TGCC. ABINIT est un code disponible depuis de nombreuses années sur les supercalculateurs du TGCC et sa mise en œuvre se fait sans aucune difficulté technique.

Comme mentionné précédemment, lors d'une exécution, le code **passé la majorité du temps de calcul dans des bibliothèques mathématiques externes, ici principalement la librairie « Math Kernel Library » (MKL) de Intel**, car elle donne accès aux fonctions BLAS (*Basic Linear Algebra Subprograms*), Lapack (*Linear Algebra Package*) et FFT (*Fast Fourier Transform*). Il a donc fallu porter une attention particulière, lors de l'édition des liens, pour garantir que la librairie MKL serait bien utilisée en mode « *multi-thread* », tout en conservant ses capacités en termes de parallélisme distribué. Les équipes du centre de calcul avaient mis à disposition tous les « modules » nécessaires à cette compilation.

L'étape de validation de l'installation d'ABINIT s'est déroulée sans difficulté particulière. Grâce à la batterie de tests automatiques du code (environ 1000 tests automatiques de non-régression), nous avons conclu rapidement qu'aucune difficulté numérique ne se manifestait, sur les types de calcul que nous souhaitions effectuer.

Les tests de performances se sont avérés tout à fait concluants, ABINIT ayant déjà fait l'objet d'un portage sur architecture KNL. Ces performances se sont révélées tout à fait conformes aux prédictions, que ce soit pour les tests de *scalabilité* forte, pour les accès mémoire ou les entrées/sorties.

2.3 Moyens de calcul et simulation

La partition « KNL » du supercalculateur Joliot-Curie a été exclusivement utilisée pour ce projet.

L'essentiel du temps de calcul disponible (**10 millions d'heures**) a été utilisé pour la mise en œuvre de la méthode de recherche aléatoire sur les hydrures de fer-aluminium sous pression. Pour chaque stœchiométrie, une centaine de structures aléatoires ont été générées, en utilisant comme biais uniquement une contrainte de distance minimum entre atomes proches. Les mailles simulées contenaient typiquement 2 formules unités de FeAlH_x et une formule unité de Fe₃AlH_x. Puis des optimisations structurales ont été réalisées dans chaque cas, avec ABINIT, en démarrant de la configuration générée aléatoirement. Cela représente donc une centaine d'optimisations structurales pour chaque stœchiométrie.

Le projet a donc nécessité le lancement de multiples instances indépendantes de ABINIT, environ une centaine par recherche de structure. Chacune de ces instances était exécutée sur plusieurs nœuds KNL du supercalculateur et avait pour tâche de réaliser l'optimisation structurale d'une cellule de simulation tirée au hasard. Chaque optimisation structurale a nécessité entre 10000 et 25000 heures CPU (temps variable d'une instance à l'autre, voir ci-dessous).

Pour une grande part, le parallélisme utilisé ici était parfaitement efficace (« embarrassingly parallel ») puisqu'il s'agissait de traiter en parallèle de très nombreuses cellules de simulation indépendantes, chaque calcul indépendant nécessitant l'utilisation de plusieurs processus MPI et tâches *openMP*.

Ce parallélisme dit « sur les images », alors qu'il n'est pas consommateur de communications, demande cependant une attention particulière car la charge de calcul est inégalement distribuée entre les instances parallèles. En effet, même si toutes les cellules de simulation doivent traiter exactement le même nombre d'atomes et d'électrons, chacune d'entre elles a un problème très différent des autres à traiter. Dans chaque cellule, les atomes sont distribués de manière très différente et le nombre de pas nécessaire à la relaxation des forces entre les atomes peut être très différent d'une cellule à l'autre. Il se pose donc un réel problème d'optimisation de la charge de calcul.

2.4 Environnement technique

Les équipes du TGCC ont fourni le « module » permettant de charger ABINIT et son environnement, tel que présenté dans la section 2.3. La librairie MKL s'est montrée directement utilisable et performante.



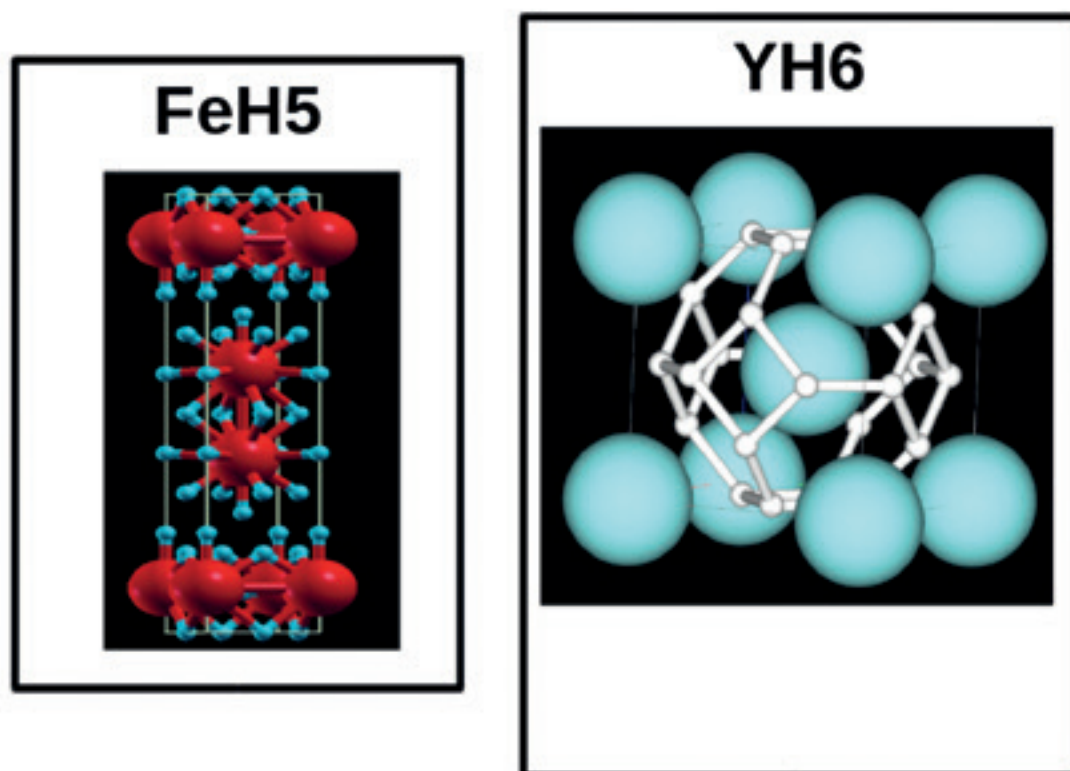


Figure 1 : image des structures de FeH_5 (à gauche) et de YH_6 (à droite)

Notons cependant que l'installation de l'une des librairies utilisées par ABINIT a nécessité un travail de portage supplémentaire. Il s'agit de la librairie ELPA (*Eigenvalue Solvers for Petaflop-Applications*) [7] qui est une librairie de diagonalisation de matrice spécifiquement développée pour les codes de simulation de matériaux à l'échelle atomique. La version KNL de la librairie ELPA présentait des instabilités numériques qu'il a fallu corriger, en échangeant, à la fois, avec les développeurs de la librairie, et les équipes du TGCC.

Les performances du réseau MPI mesurées n'ont pas montré d'impact très fort de la localisation des échanges de messages (intra-nœud ou inter-nœud). Aucun effet NUMA (*Non-Uniform Memory Access*) n'a été détecté sur les nœuds KNL, ce qui était un résultat plutôt attendu.

La maîtrise du code ABINIT par les équipes du centre de calcul, associée à la parfaite connaissance du code par notre équipe – qui participe à son développement -- a permis un démarrage très rapide du projet.

3. Observations et résultats scientifiques

Avant de démarrer sur les hydrures de fer-aluminium, nous avons choisi de tester la méthode sur deux systèmes connus : (i) FeH_5 , (ii) les hydrures d'yttrium, en particulier YH_6 .

FeH_5 est un superhydrure de fer dont l'existence a récemment été démontrée dans notre équipe. Des structures aléatoires contenant 2 formules unités de FeH_5 (12 atomes) ont donc été générées et la bonne structure a été retrouvée très rapidement, au bout de 100 tirages. De même, la structure du superhydrure YH_6 , dont l'existence a été prédite théoriquement en 2015 [8], est retrouvée très rapidement par la méthode de recherche aléatoire. Les deux structures en question sont représentées sur la figure 1.

Ensuite, nous avons étudié la série des FeAlH_x , pour des stoechiométries allant de $x=2$ à $x=8$. Les recherches ont été faites à une pression de 50 GPa pour les plus faibles stoechiométries en hydrogène, et 100 GPa pour les plus élevées. Pour chaque valeur de x , la structure la plus stable trouvée est retenue, puis son équation d'état est calculée en faisant varier la pression. En-dessous d'une vingtaine de GPa, et en excès



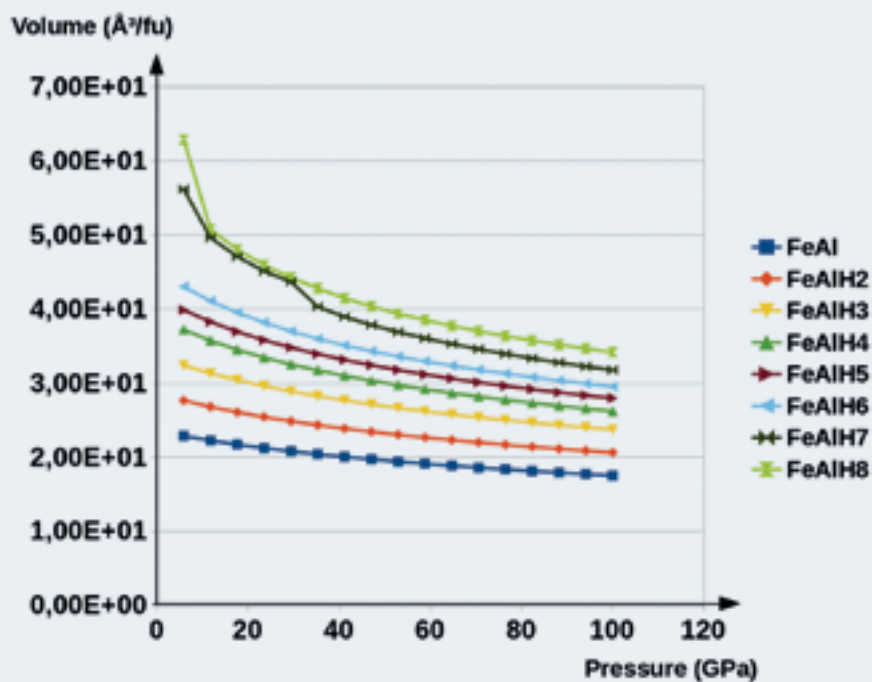


Figure 2 : volume calculé en fonction de la pression pour les différents hydrures de fer-aluminium FeAlH_x .

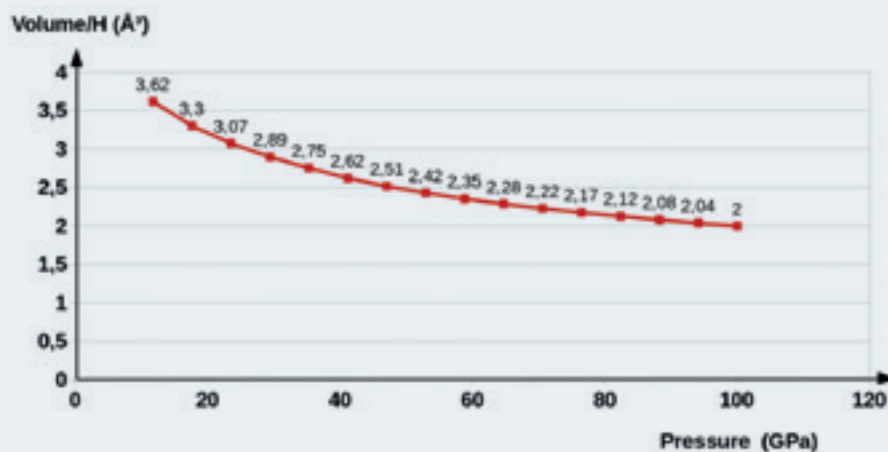


Figure 3 : incrément de volume par incorporation d'un atome H par formule-unité, en fonction de la pression, dans les FeAlH_x .

d'hydrogène, l'hydrure FeAlH_4 est trouvé stable, suivi par FeAlH_5 jusqu'à ~ 115 GPa. Au-delà, FeAlH_6 devient l'hydrure le plus stable. Nous confirmons ainsi une tendance observée dans de nombreux hydrures sous pression : la stoechiométrie en hydrogène tend à augmenter avec la pression

(la pression stabilise des composés de plus en plus riches en hydrogène).

Par ailleurs, le tracé des équations d'état (volume en fonction de la pression : figure 2) montre que l'incorporation d'un atome d'hydrogène supplémentaire (par formule



unité) se traduit, à une pression donnée, par une augmentation de volume, qui dépend peu de la stœchiométrie.

Il est donc possible d'extraire cet incrément de volume par atome d'hydrogène incorporé, et d'observer comment il évolue avec la pression (figure 3). Cette information est essentielle car elle peut permettre aux expérimentateurs de trouver la stœchiométrie en hydrogène des composés qu'ils synthétisent uniquement sur la base du volume de la maille, qui peut, dans une certaine mesure, être accessible par diffraction des rayons X.

Les valeurs trouvées sont en bon accord avec celles observées sur d'autres hydrures comme les hydrures de fer [4].

Des résultats comparables ont été obtenus sur la série des Fe_3AlH_x .

4. Conclusions et perspectives

Le présent projet nous a permis de mettre en place la méthode de recherche aléatoire de structures. Dans la recherche effectuée par notre équipe, **une telle compétence est indispensable**. Nous poursuivons actuellement ce projet avec la recherche de structure d'hydrures d'yttrium, d'hydrures d'autres métaux de transition, et d'hydrures ternaires (autres que les FeAlH_x) contenant du fer. **Cette recherche a pour but l'identification de composés susceptibles de posséder des propriétés supraconductrices à haute température critique et pouvant permettre un stockage de l'hydrogène en phase solide à très haute densité**. Nous espérons que ces travaux pourront contribuer à mettre en évidence de nouveaux matériaux avec de telles propriétés

5. Remerciements

Le programme XCrySDen [9] a été utilisé pour réaliser la figure 1. Le programme Agate a été utilisé pour réaliser la figure 1 ainsi que la vidéo montrant l'optimisation structurale aboutissant à la structure de YH_6 .

6. Bibliographie

- [1] N. W. Ashcroft, Phys. Rev. Letters 92, 187002 (2004).
- [2] N. W. Ashcroft, Phys. Rev. Letters 21, 1748 (1968).
- [3] J. M. McMahon, D. M. Ceperley, Phys. Rev. B 84, 144515 (2011).
- [4] C. M. Pépin, G. Geneste, A. Dewaele, M. Mezouar, P. Loubeyre, Science 357, 382 (2017).
- [5] F. Peng, Y. Sun, C. J. Pickard, R. J. Needs, Q. Wu, Y. Ma, Phys. Rev. Letters 119, 107001 (2017).
- [6] C. Pickard, R. J. Needs, J. Phys. Cond. Matt. 23, 053201 (2011).
- [7] <http://elpa.mpcdf.mpg.de>
- [8] Y. Li, J. Hao, H. Liu, J. S. Tse, Y. Wang, Y. Ma, Sci. Reports 5, 9948 (2015)
- [9] A. Kokalj, Comp. Mater. Sci. 28, 155 (2003).





Équipe

Mathieu MOOG, *Doctorant*Fabio PIETRUCI, *MCF*A. MARCO SAITTA, *PREX*Institut de Minéralogie, Physique des Matériaux
et de Cosmochimie (IMPMC),
Sorbonne Université, CNRS, MNHN, IRD, Paris

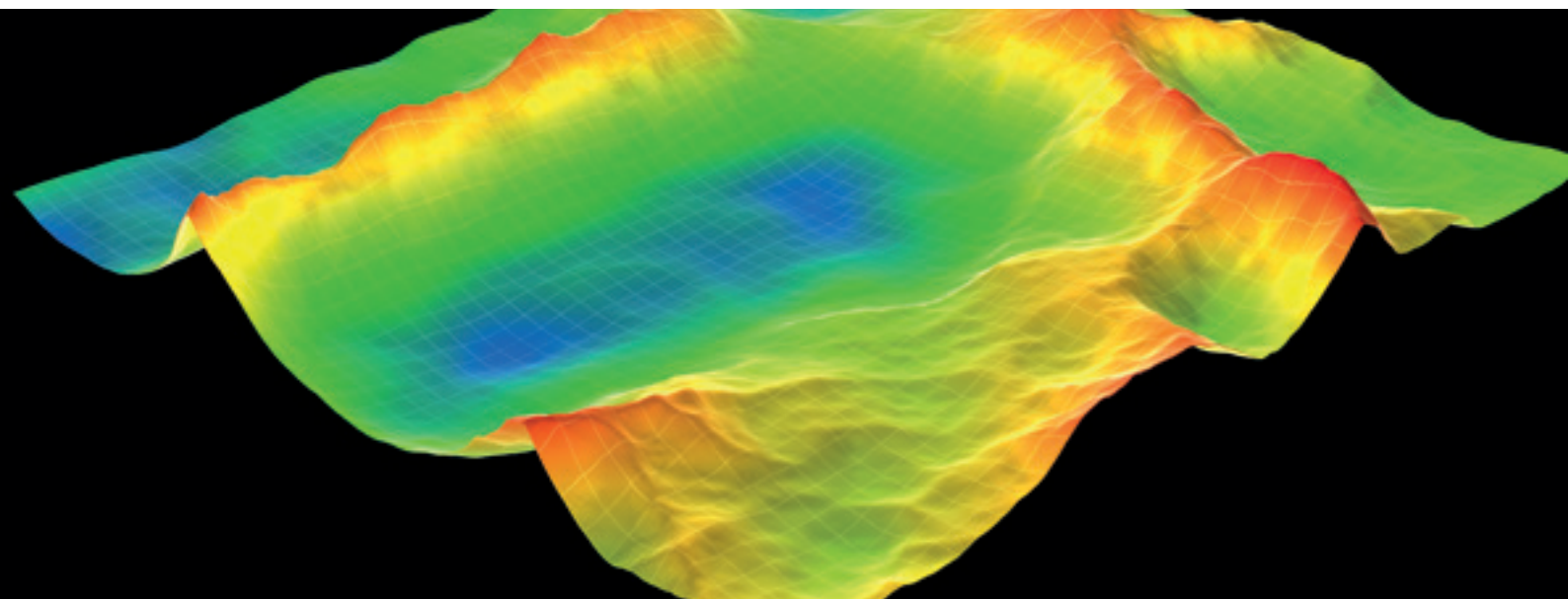
Phases exotiques du dioxyde de carbone fluide à conditions géologiques

5 millions d'heures

Principalement connu et étudié pour ses effets sur le climat, **le dioxyde de carbone (CO₂) est un composant très important du manteau terrestre**. Il s'y forme, dans un état de fluide, à des conditions extrêmes de pression et de température lors des phénomènes de subduction de couches contenant des matériaux organiques, et il s'en **échappe de façon importante lors des phénomènes volcaniques et par les rejets des failles océaniques**. Sa présence dans le manteau a un impact fort sur ses propriétés et son comportement : il constitue une étape importante du cycle du carbone en profondeur, et, mélangé notamment avec l'eau, il constitue un composé très important dans le transport géologique. Il a été suggéré que sa présence en profondeur avait un effet sur les phénomènes de répliques des tremblements de terre. Le CO₂ réagit naturellement avec les autres constituants du manteau, ce qui pourrait notamment affecter ses propriétés de transport de certains matériaux lourds dans le manteau. Pour toutes ses raisons **le dioxyde de carbone en condition géologique est un objet d'étude important pour la compréhension du manteau terrestre**.

Ainsi et pour ces raisons, le CO₂ a été un objet important d'étude expérimentale pour les physiciens s'intéressant aux phénomènes se produisant à conditions extrêmes de pression et température. Un grand nombre d'études expérimentales, complétées par quelques études théoriques, ont ainsi pu mettre en évidence l'existence de 7 phases cristallines moléculaires, ainsi que deux phases cristallines polymériques et une phase amorphe polymérique. Cependant, ces études se sont heurtées à des limitations expérimentales importantes et n'ont pas permis de bien comprendre comment le fluide moléculaire, qui est la phase présente dans les conditions typiques des couches les plus superficielles du manteau, se transformait sous compression. Afin de guider les expériences et de permettre d'avoir une idée du comportement de ce fluide, deux études théoriques récentes se sont penchées sur le sujet et ont proposé deux comportements principaux : d'abord une transition liquide-liquide du premier ordre autour de 50 GPa, et potentiellement dans ces mêmes conditions, une possible formation de la phase solide cristalline V. Si ces calculs, de type statique-énergétique, ont donc produit des prédictions intéressantes, ils





souffrent de deux limitations importantes pour le dioxyde de carbone en condition géothermique : les simulations étaient très courtes (du fait de leur coût et de la volonté de couvrir le plus possible le diagramme de phase) et étaient réalisées dans des conditions de températures bien supérieures ($\sim 3000\text{ K}$) à celles de la géothermie (autour de $2000\text{ K} < T < 2500\text{ K}$).

Ici nous, tout en partant de ces résultats théoriques précédents, nous étendons largement l'étude et découvrons des nouveaux résultats inattendus. **Nous nous concentrons essentiellement sur des conditions physiques proches de la géothermie ($40\text{--}70\text{ GPa}$ et $2000\text{--}2500\text{ K}$) et nous effectuons des simulations longues (100 ps) afin d'accumuler des statistiques suffisantes pour étudier les divers mécanismes de réaction et transformations qui ont lieu dans ses conditions extrêmes.**

Dans ce contexte, l'obtention d'heures dans le contexte du Grand Challenge nous a permis de faire des simulations sur des boîtes de simulations beaucoup plus grosses que celles qui ont été utilisées dans l'ensemble des travaux précédents (108 molécules de CO_2 soit 324 atomes). La taille de la boîte est un aspect crucial

pour une réelle extrapolation macroscopique des phénomènes observés : une **cellule de simulation trop petite induit des effets de bord capables de fausser complètement les résultats obtenus**. Ceci est très important ici, comme l'un des enjeux de l'étude est l'observation et la caractérisation de la polymérisation du liquide moléculaire, phénomène pour lequel les effets de bord risquent justement d'être cruciaux. Du point de vue numérique, cette avancée en termes de description physique du système s'accompagne d'une augmentation importante du coût des calculs avec le nombre d'électrons simulés, qu'il serait prohibitif dans le cadre des allocations standards (**une seule simulation de 10ps pouvant équivaloir à plusieurs centaines d'heures de calculs**).

Grâce au Grand Challenge et aux heures allouées dans le contexte des allocations GENCI standard nous avons pu mettre en évidence quatre régimes différents pour le dioxyde de carbone fluide en condition géothermique :

- A basse température et pression, le dioxyde de carbone est un fluide moléculaire classique, avec une double liaison $\text{C}=\text{O}$ forte.



- A plus haute pression et température ($P > 30 \text{ GPa}$ et $T \sim 2000 \text{ K}$) le fluide moléculaire évolue progressivement pour devenir un fluide moléculaire réactif où les molécules de CO_2 peuvent réagir entre elles pour former des dimères éphémères avec des échanges d'oxygènes entre les diverses molécules.
- En dessus de 50 GPa , le fluide moléculaire se transforme rapidement en un fluide polymérique très réactif et dominée par des unités CO_3 planaires, qui forment des chaînes longues.
- Dans les conditions de pression les plus extrêmes ($> 70 \text{ GPa}$), le dioxyde de carbone présente un comportement très similaire à la phase amorphe polymériques existant à des pressions similaires mais à plus basse température ($< 500 \text{ K}$). Ceci laisse à croire qu'en condition géologique, on observerait alors formation de la phase cristalline polymérique V qui est plus favorable, du point de vue énergétique, à la phase amorphe.

L'ensemble de ces résultats fait actuellement l'objet de la rédaction d'un article scientifique.







GENCI adresse ses remerciements aux équipes du CEA, aux ingénieurs et aux chercheurs qui ont contribué à la rédaction de cette Revue Grands Challenges.

Directeur de la publication : Philippe LAVOCAT - Coordination : Séverine Saint-Hubert
Conception et réalisation : Backdrop - Impression : Welcom
Photos : © CEA, © AdobeStock, DR.
GENCI, 6bis, rue Auguste Vitu - 75015 Paris - France. Tél. : +33 1 42 50 04 15
© GENCI - juin 2019.



