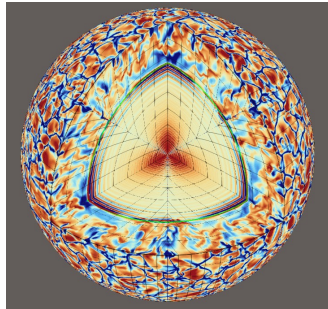


Allan Sacha Brun
(IRFU/DAP)

Diversité des simulations HPC
dans COAST
(planètes, étoiles, MIS, Galaxies, Univers)

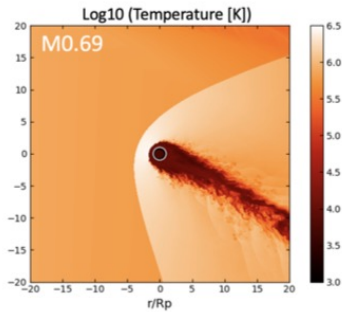
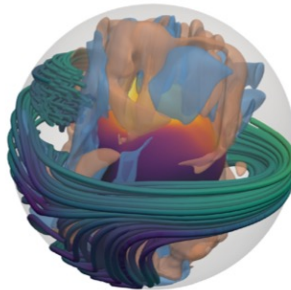


ρ [M_☉/baryon]

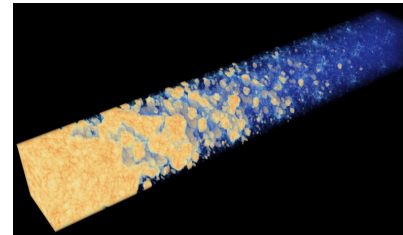
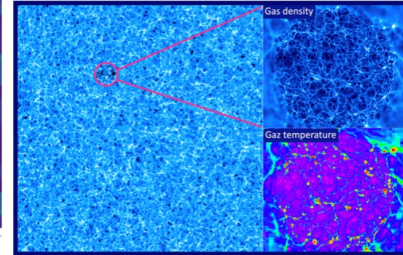
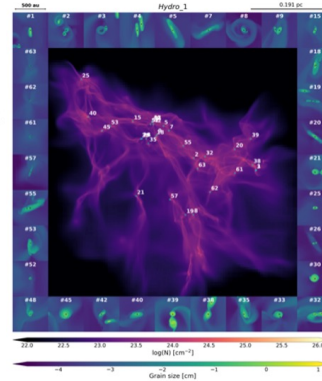
17
12.75
8.5

L2-0A

z
x y



8700 km



*HPC: High Performance Computing; *COAST: programme transverse IRFU

*MIS: Milieu InterStellaire

Précédente 2020-2025: Production et exploitation des données Big-Data, IA (Défi 13), Partage et accès ouvert (Défi 14) Simulations numériques (Défi 17)

Prospective 2025-2030: HPC/Calculs => Moyens Prioritaires (50+ réponses)

Une demande (HPC/IA/HPDA) a aussi été faite aux OSU (moyens/personnels)

⇒ Une CID 55 (sciences et données) sections 17 et 6 a été créée

⇒ Evolution des Instituts, INSU va fusionner avec IN2P3 en 2027 ⇔ IRFU!

⇒ Une Action Spécifique Numérique est en cours (Y. Dubois responsable). Elle a défini un nouveau service national (SNO): le code-objet (plutôt que le code-moyen)

*CID: commission interdisciplinaire (sections) du CNRS

Science Vision & Infrastructure Roadmap 2020-2030

The new European roadmap for Astronomy

The key astronomical questions which may be answered in the next twenty years by a combination of observations, **simulations**, laboratory experiments, interpretation and theory.

5 working groups (no CEA members)

Origin and Evolution of the Universe

Formation and Evolution of Galaxies

Formation and Evolution of Stars

Formation and Evolution of Planetary Systems

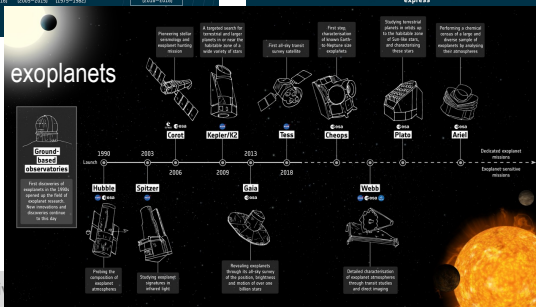
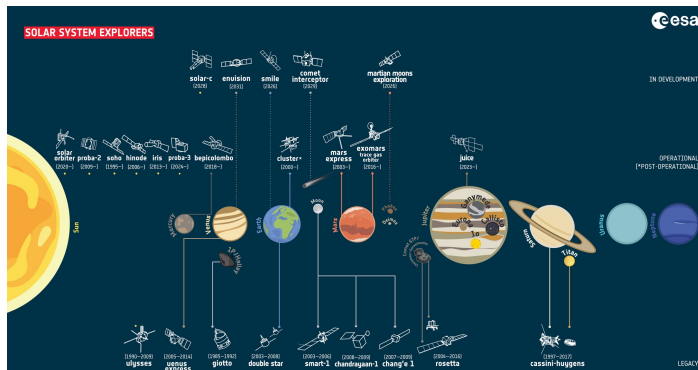
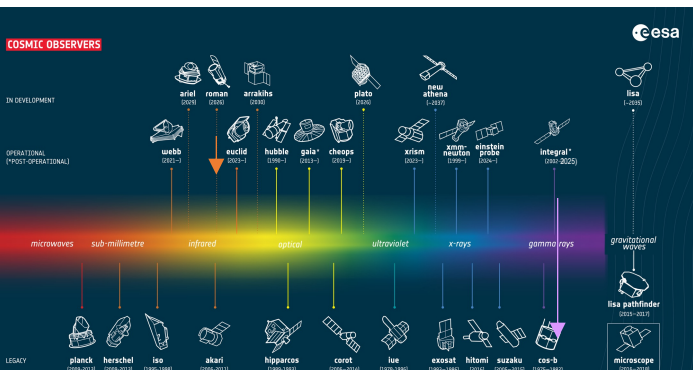
Understanding the solar system and conditions for life

Observatoire Virtuel

the development of theory and computing capacity must go hand in hand with that of observational facilities. Systematic archiving of properly calibrated observational data in standardised, internationally recognised formats will preserve this precious information obtained with public funds for future use by other researchers, creating a Virtual Observatory. The Virtual Observatory will enable new kinds of multi-wavelength science and present new challenges to the way that results of theoretical models are presented and compared with real data. The Roadmap therefore proposes continued **development of the Virtual Observatory**, including recommending planning for the provision of a public, VO-compliant archive with any new facility.

At the heart of the Cosmic Vision programme are four big questions: What are the conditions for planet formation and the emergence of life? How does the Solar System work? What are the fundamental physical laws of the Universe? How did the Universe originate and what is it made of?

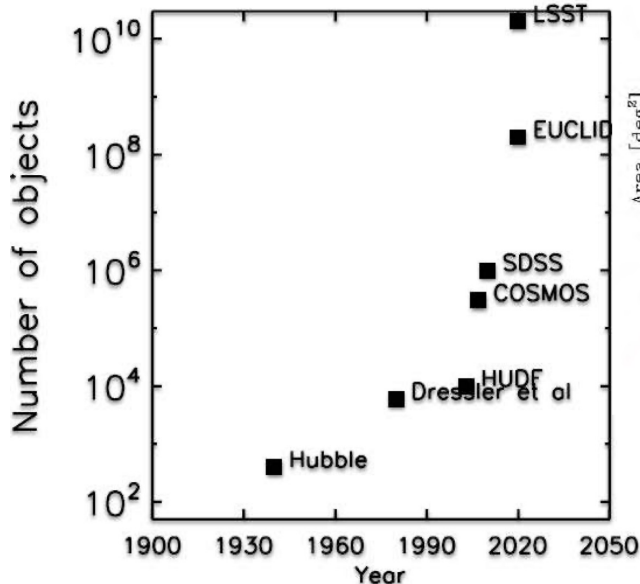
But what do we want to be studying in 20 or even 50 years from now? We have to start planning now for the missions we want to launch **two or more decades from now** => **Voyage 2050**



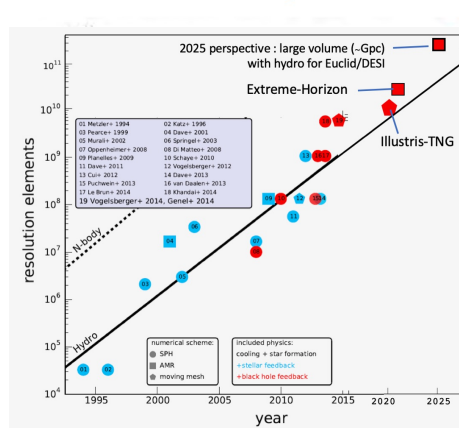
- 1) Exploitation scientifique des missions ESA Cosmic Vision (Solar Orbiter, Euclid, JWST, PLATO, Ariel, Athena), futures (Voyage 2050; VIGIL) et des missions Horizon 2000 tjrs operationnelles (XMM, SoHO, ~~Integral~~). Hors ESA: SVOM, FERMI
- 2) Exploitation scientifique des telescopes sols (ELT/Metis, CFHT/Megacam, LSST, eRosita, Alma, CTA, SKA (en 2026 la France est devenue 14eme état membre))
- 3) Projet COAST, exascales ready code (hybridation GPU/CPU/ARM), **code-objet**
- 4) HPDA/IA/Machine Learning – methodes avancées en traitement du signal
- 5) Assimilation de données en astrophysique, temps réels, alertes et prévisions, **couplage modèle-données**
- 6) En lien avec données simu et ou instruments: **base de données ouvertes**. Galactica pour les simu en cours, gestions des données de nos grands programmes (hors cadre des data centers/segments sols). Approche OVA, Fair (principes FAIR : Facile à trouver, Accessible, Interopérable, Réutilisable - modulo que r (si reproduisible) n'est pas tjrs possible si évènement ponctuel ou observations rares), open data science

Nos moyens informatiques locaux, le réseau, les simulations et l'analyses doivent **anticiper les besoins à venir** aussi bien **sols**, que **spatiaux** que **théoriques**

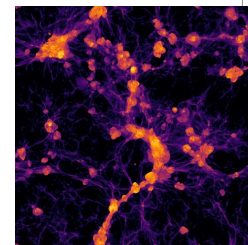
Les « surveys » astro



IDEM pour les SIMULATIONS



Taille des simulations
Nombre de particules
ou d'éléments de résolution
+ processus physiques



OBSERVATOIRE VIRTUEL

Données scientifiques des grands surveys:
doublement tous les 18 mois !

CDS: images: 350 TB en 2020 → ~ 2-3 PB en 2025

⇒ 10+ PB en 2030

→ Hiérarchie des données, Big data technologie
(HiPS Hierarchical Progressive survey)

→ Le Cloud EOSC

Approche FAIR

Depuis 2018, et le plan national, la Science Ouverte est favorisée/encouragée

<https://www.enseignementsup-recherche.gouv.fr/pid39205/science-ouverte.html>

Un exemple à l'IAP vu sur le site du Ministère



Qu'y a-t-il entre les galaxies, ou pourquoi étudier les « vides » cosmiques ?

« VIDE » a permis de construire des bases de données ouvertes de vides cosmiques

Guilhem Lavaux

Chargé de recherche à l'IAP, labellisé Domaine d'Intérêt Majeur par la Région Île-de-France, Institut d'Astrophysique de Paris



Open database au CEA?

La **science ouverte** a pour objectif de faire sortir la recherche financée sur fonds publics **du cadre confiné des bases de données fermées**. Elle réduit les efforts dupliqués dans la collecte, la création, le transfert et la réutilisation du matériel **scientifique**. Elle augmente ainsi l'efficacité de la recherche.

<http://www.cosmicvoids.net>

https://bitbucket.org/cosmicvoids/vide_public/src/master/

▪ *Enjeux communs* aux diverses échelles astro dans **COAST** :

- Passage à l'exascale, nouvelles architectures => besoin de nouveaux codes hydrodynamiques à maillage adaptatif, à la fois performant et portable.
- L'enjeu n'est pas juste de calculer plus vite, mais aussi de profiter des nouvelles opportunités offertes (temps-réel HPDA+HPC, grands volumes de mémoire etc) tout en optimisant le coût énergétique.

▪ *Spécificités propres* aux divers objets astrophysiques :

- Cosmologie : particules non-collisionnelles, gravité, etc
- Formation stellaire : auto-gravité, etc
- Physique stellaire – SPML : multi-régime, des intérieurs aux vents stellaires, etc
- Objets compacts : multi-fluide

⇒ Besoins de couches logicielles très spécifiques aux problématiques astro à ajouter aux bases de codes génériques

⇒ **Métier de modélisateur astrophysique au DAp** (largement > 1 poste pour exploiter l'exascale) complémentaire des compétences générales du DEDIP.

NB : L'INSU reconnaît le besoin de ce métier dans les labo d'astro (nouvelle ANO « **code-objet** », CID55)

Codes // encore utilisés dans COAST: ASH, RAMSES, PLUTO, MAGIC => Dyablo (ginéa/whole sun)?

▪ *Enjeux communs* aux diverses échelles astro dans **COAST** :

- Passage à l'exascale et au delà, nouvelles architectures => besoin de nouveaux codes MHD à maillage adaptatif, à la fois performant et portable =< Dyablo
- L'enjeu n'est pas juste de calculer plus vite, mais aussi de profiter des nouvelles opportunités offertes (temps-réel HPDA+HPC, grands volumes de mémoire etc) tout en optimisant le coût énergétique.

▪ *Nouvelles opportunités à exploiter* :

- Space Weather : possible avec les perspectives de quasi temps-réel HPDA+HPC à l'exascale couplé à l'assimilation de données (DA)
- Cosmologie : besoin de modélisation de très grands volumes physiques, possible avec les volumes mémoires de l'exascale

=> Opportunités pour le DAP et l'IRFU à saisir avec des profils forts (CDI).

▪ *Visée de codes communautaires* :

- Nécessite des échanges d'expertise avec la communauté
Démarche à soutenir via un « COAST fellowship » récurrent.

- *Mise à disposition* des données spatiales, sol, simulations :

Outils existants (base Galactica) mais besoin de :

- stockages : maintien récurrent des stockages locaux, mutualisation
- réseaux : actuellement un bottleneck majeur
- ouverture : par exemple pourra-t-on mettre à disposition des

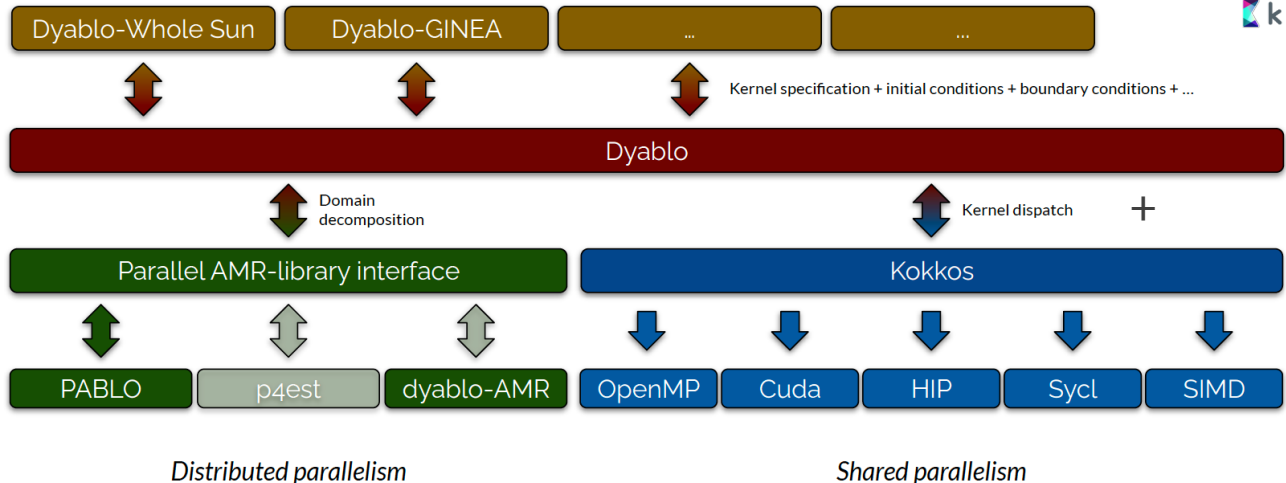
données produites et/ou stockées au TGCC ?

- *Besoin de performance réseaux* :

- Actuellement le déplacement des données de simulation est difficile, à l'exascale et au delà (x10 d'ici 2030) il sera impossible sans évolution.
- Données spatiales impossibles à déplacer dans des temps réalistes via les réseaux actuels.
- La problématique doit être résolue face à l'évolution des moyens de calcul

- *Besoin de spécificité* : une grande partie des enjeux HPC/HPDA/IA repose sur des métiers alliant technicité et forte spécificité astrophysique, dans les labos d'astrophysique.

Dyablo: a C++ high-portable performance AMR framework



Dyablo's open source software stack

- Written in C++, uses external libraries
- Kokkos + MPI parallelism
- Newer parallel algorithms

Supports Exascale Hardware

We are on github !

- Look us up at <https://github.com/Dyablo-HPC/Dyablo>

FEATURES IN DYABLO

ADDRESSING THE SIMULATION NEEDS FOR THE COMMUNITY IN ASTRO CFD

Multi-physics simulations

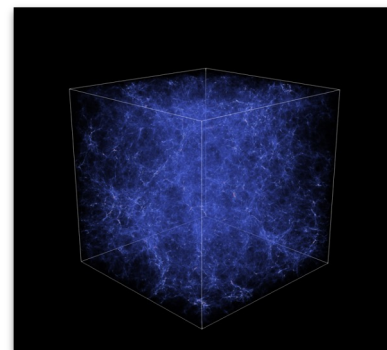
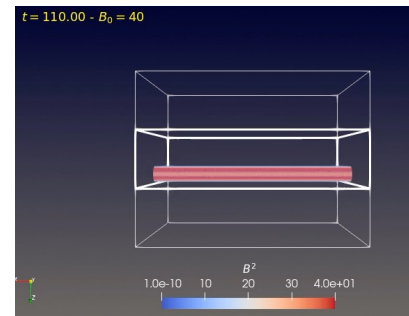
- Finite-volumes method
- Hydrodynamics / MHD / RHD
- Self-Gravity
- Particles
- Diffusion terms: Thermal conduction, Viscosity
- Cosmology: Comoving coordinates

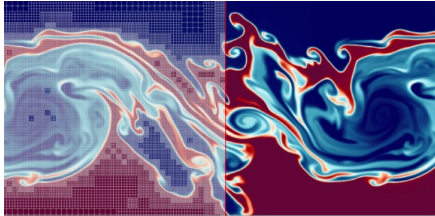
Adaptive Mesh Refinement (AMR)

- Wide range of time/space scales in same simulation
- New geometry module
- Block based approach to AMR

Massively parallel simulations :

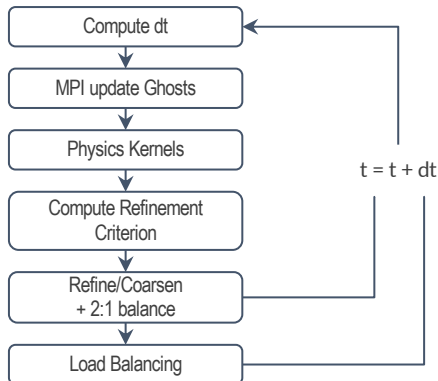
- Shared-memory parallelism with *Kokkos* (CPU, GPU, ...)
- Distributed parallelism with MPI





Adaptive Mesh Refinement

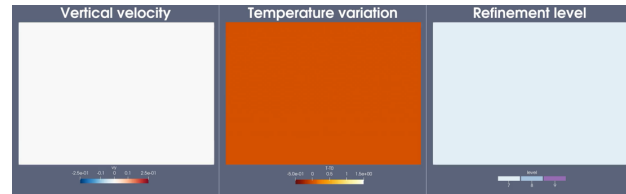
- More resolution in regions of interest
 - Octree-based AMR mesh (*cartesian AMR*)
 - *Dynamic mesh changing at every timestep*
- AMR cycle may be costly, access patterns are random



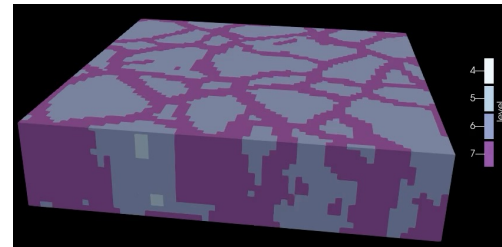
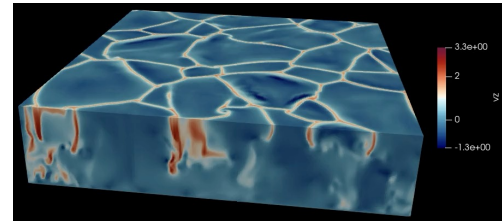
Block-based AMR :

- Every octant is refined in 2 in each direction
- The leaves of the tree are regular cartesian subgrids

2D



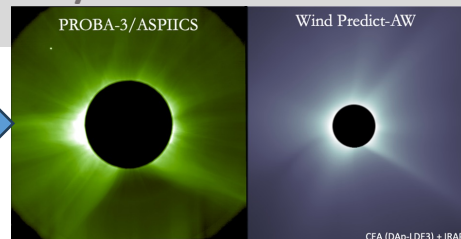
3D



Refined mesh for a convection run in
Dyablo Whole Sun
(Brun, Strugarek, Finley, Delorme)

Brun, Perri, Strugarek (DAp/LDE3) **ERC Exomagnet**

Prévision de l'éclipse du 12 aout 2026, code PLUTO →



- Exposer vos défis scientifiques sur du court/moyen/long terme
 - À 3 ans Couplage de codes en météo Espace (End to End)
 - 3- 5 ans Dyablo Whole Sun Solaire local et global Exascale + base de données de simulations
 - 8-10 ans code mature 3D AMR MHD+RT+MultiRegimes modelisant de l'intérieur à la couronne et vent solaire, capable d'utiliser les plateformes post-exascale (dépôt 2026: ERC+ Unlocking the Sun)
- Associer et identifier les verrous scientifiques et techniques pour identifier/lister:

Les verrous auxquels vous faites face

- Hybridation de l'AMR, Couplage profond intérieur-extérieur du Soleil, multi-régime (beta et Mach number => compressible+ All Mach), Pb typique de la turbulence: Reynolds max est dimensionné par la résolution, parallélisme en temps, transfert de rayonnement (short characteristics), well balanced

Les outils numériques dont vous avez besoin pour lever ces verrous

- Nouveau code exascale-ready multi-régime (Dyablo Whole Sun). Visualisation 3D/holographique

Les impacts sur votre institut

- Humains, financiers, autre? Il faudrait pérenniser cela avec l'embauche d'un expert en météorologie de l'espace pour l'assimilation de données, prévision et alertes, le couplage modèle-données (DA)

- Comment constituer cette liste de besoins:

- 1 expert en météorologie de l'espace (temps réels, assimilation de données, prévision), relations Soleil-Terre, SOLARCAST, données Solar Orbiter, Vigil (2031)
- numéricien astrophysique (voir plus loin)

*DA: data assimilation

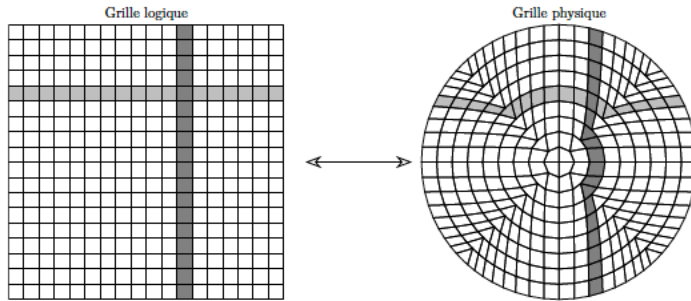
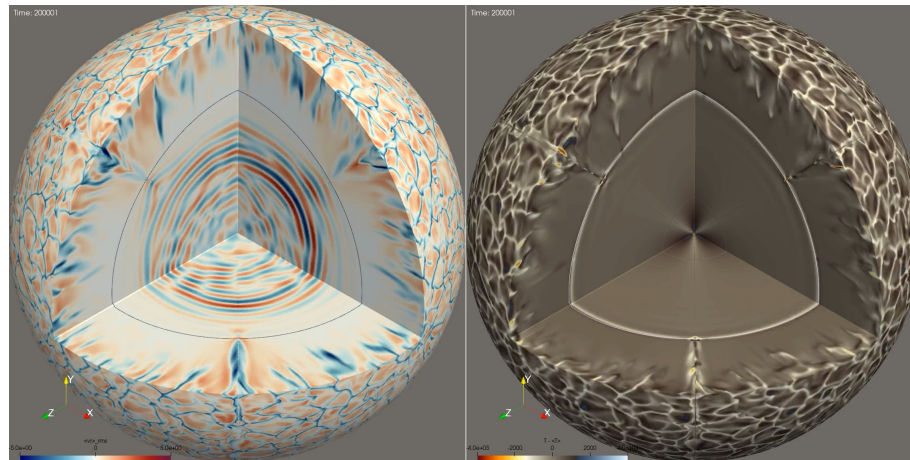


FIGURE 4.5 – Grille radiale

ERC Whole Sun



Brun, Strugarek (Dap/LDE3)
Delorme (Dedip)

Thèse de G. Doebele

Acoustic and Gravity waves

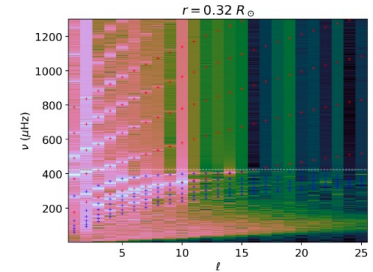


FIGURE 6.12 – Diagramme (t, ν) du spectre de puissance de la vitesse radiale à $r = 0.32 R_{\odot}$. Les points superposés correspondent aux modes propres théoriques calculés avec GYRE (Townsend et Teitler, 2013) (modes p en rouge, modes g en bleu). La ligne pointillée blanche indique la fréquence de Brunt-Väisälä correspondant à ce rayon.

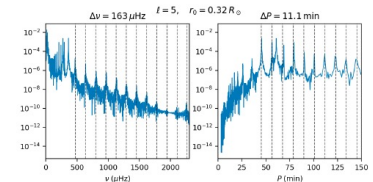


FIGURE 6.13 – Coupe du spectre de puissance à $r = 0.32 R_{\odot}$ pour $l = 5$, en fréquence (à gauche) et en période (à droite). On mesure un espacement régulier en fréquence $\Delta \nu = 163 \mu\text{Hz}$ pour les modes p et un espacement en période $\Delta P = 11.1 \text{ min}$ pour les modes g .

Guilet, Bugli, Foglizzo, Raynaud (DAP/LMPA)

- Exposer vos défis scientifiques sur du court/moyen/long terme

Objectif à 1 an : Perfectionner le modèle de dynamo décrivant l'ensemble de la proto étoile à neutrons (notamment couplage des zones de stratification stable et instable)

Objectif à 5 ans : calculer la signature multi-messenger de la formation d'un magnétar et l'explosion de supernova associée, simulation à plus long terme de l'évolution de la proto-étoile à neutrons, améliorer le lien avec les observations de GRBs et SLSNe

Objectif à 10 ans : développement de simulations ab-initio cohérentes décrivant l'ensemble des propriétés des explosions extrêmes, ce qui nécessite probablement un modèle sous-maille de l'amplification du champ magnétique qui sera auparavant testé avec les modèles de dynamo de proto-étoile à neutrons

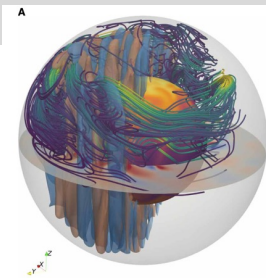
Verrous : manque dans la description de l'évolution de la proto-étoile à neutrons, manque de couplage entre les simulations locales et globales (par exemple à l'aide d'un modèle sous-maille décrivant les petites échelles)

Outils numériques :

-> Développement de MagIC : GPU, amélioration de la parallélisation MPI, traitement du transport de neutrinos, évolution du background, écart à la géométrie sphérique

-> explorer la piste du deep learning pour obtenir des modèles sous maille de turbulence

Besoins : Soutien informatique pour avoir des moyens performants et fiables de stockage et de traitement des données et - Numéricien astrophysique



Code PLUTO

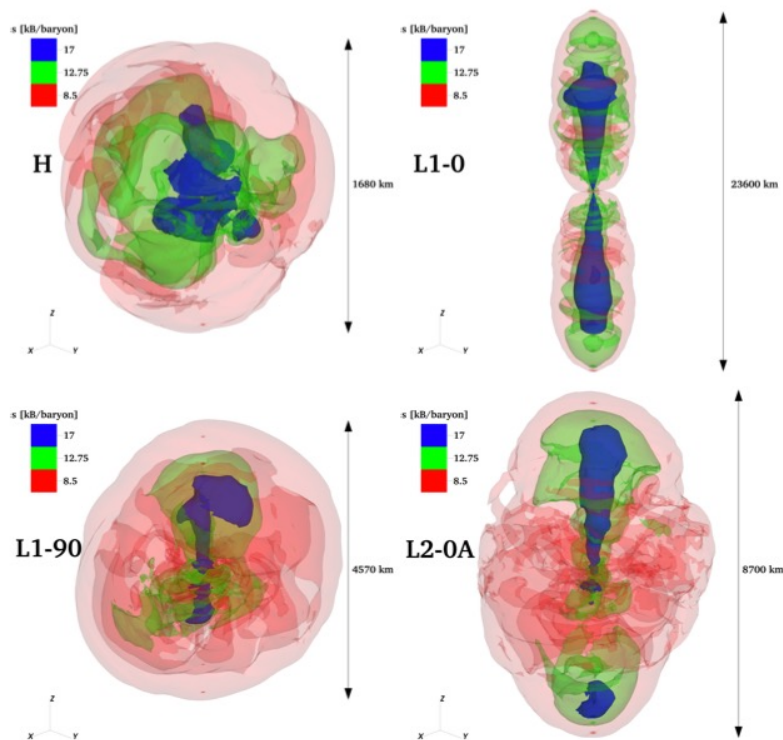
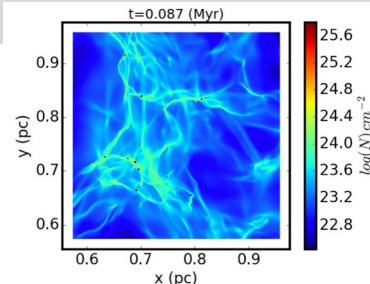
Erc blackJet
(M. Bugli)

Figure 2. Volume rendering of the specific entropy for CCSN models with different initial magnetic field configurations: no magnetic field (top left panel), aligned dipole (top right), equatorial dipole (lower left panel), and aligned quadrupole (lower right panel) at $t = 410$ ms p.b. [36]

Hennbelle, Lebeuilly, Gonzales (DAP/LMPA)

ERC Ecogal



■ Objectif global :

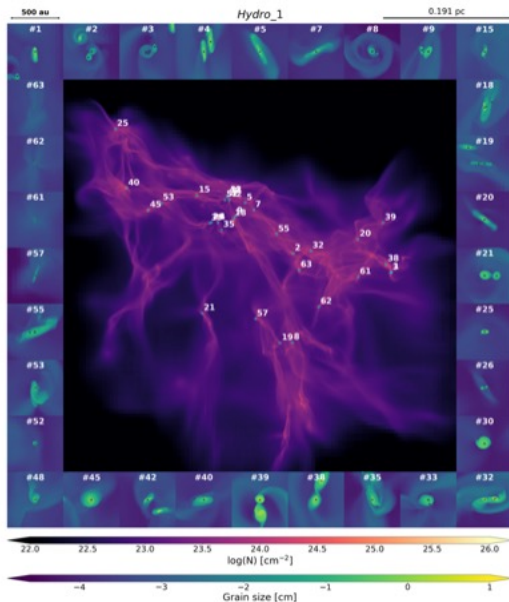
- Comprendre l'(in)efficacité de la formation d'étoiles, allant des l'échelle des galaxies (Cordis) aux étoiles individuelles (Hennebelle, Lebeuilly).
- Impact sur l'évolution globale des populations de galaxies, IMF des étoiles.
- Liens observations forts (ALMA, JWST, Euclid-legacy, E-ELT...)

■ Simulations HPC : Milieux à très fort contrastes de densités (hauts Mach, chocs).

- Evolution très rapides des structures sur de grands nombres de temps dynamiques, et très multi-échelle (développement de “zooms” imbriqués)
- Les raffinements adaptatifs sont indispensables (cout mémoire) mais sont aussi très irréguliers et rapidement évolutifs (GPU-isation complexe)

■ Enjeux/besoin : L'exascale (surtout GPU) est un cas complexe compte-tenu de ces contraintes. Effort de prospective pour identifier les bons prototypes de codes dans ce cadre

P. Hennebelle, U. Lebreuilly



Modèle **RAMSES** résolvant les grandes échelles des nuages moléculaires et les petites échelles des disques protoplanétaires en prenant en compte la croissance des grains

~1000 CPUs/modèle (~1MCPUh/modèle)

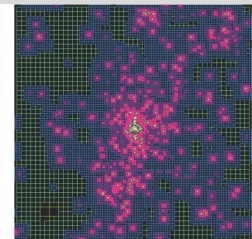
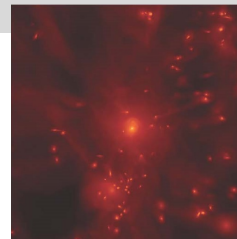
Prochaines générations de modèles

DYABLO : nouvelle branche ISM (Ph.D Solange Druelle)

RAMSES I → **II** : simulations haute résolution de disques (Ph.D Ronan Meunier)

Cordis, Elbaz, Daddi (DAp/LCEG)

Lien avec Cosmo stat: Stark, Farrenz, Lanusse, Kilbinger et DPhP



■ Contexte :

Les grandes simulations actuelles (e.g. “Euclid Flagship”) sont “matière noire seule”
=> biais incontrôlé de $\sim 5\sigma$ pour les expériences type Euclid (d’après Schneider, Teyssier et al. 2019 eux-mêmes)

Les simulations multi-physique (hydro, feedback, etc...) couvrent des échelles 10 fois trop petites ($< 1 \text{ deg}^2$ sur le ciel $\ll$ échelles du weak shear)

■ Opportunité :

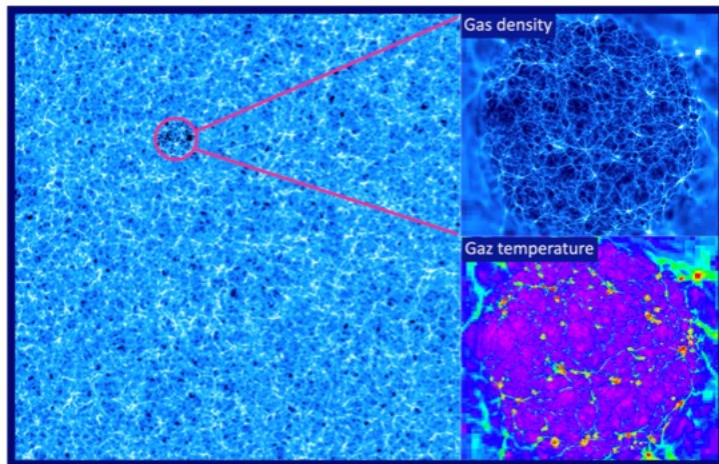
- L’exascale va permettre des simulations grand-volume multi-physiques sans dépasser la mémoire des supercalculateurs
- Pour ces applications de recherche d’effets à grande échelle, les raffinements AMR sont peu profonds et peu évolutifs => nos codes AMR sont déjà adaptés pour ce cas d’étude / au bon moment par rapport au spatial

=> Opportunité pour l’institut mais peu de FTE dédiées aux simulations cosmologiques AMR de grand volume

Cosmological simulations with baryons (RAMSES)

2. 100 Mpc/h Zoom-in simulations in a Gpc box

Goal : investigate how cosmic variance affect baryonic effects on cosmological observables

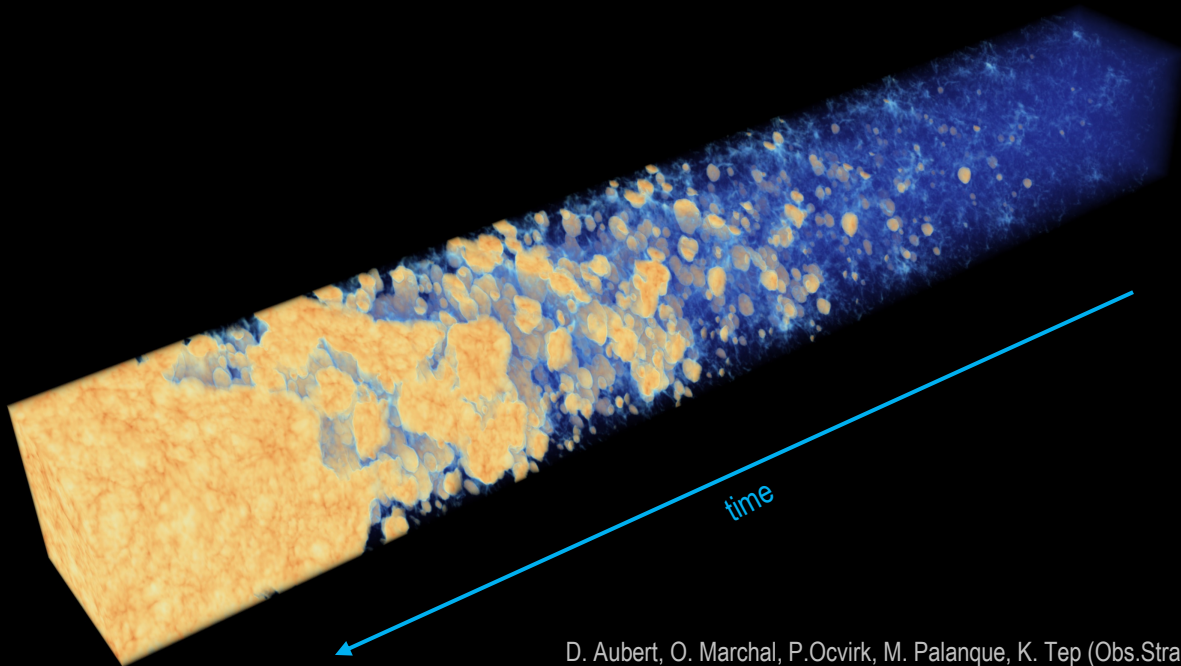


ERC COLIBRI.
S. Codis, A. Barthelemy,
E.Aycoberry, @DAP/LCEG

2

10-100 milliards de particules. Resolution de base: 1024^3 ou 2048^3 , 17 niveaux AMR (réduits à 9 selon besoin)

$1024^3 + 4$ niveaux AMR, 10^9 particules

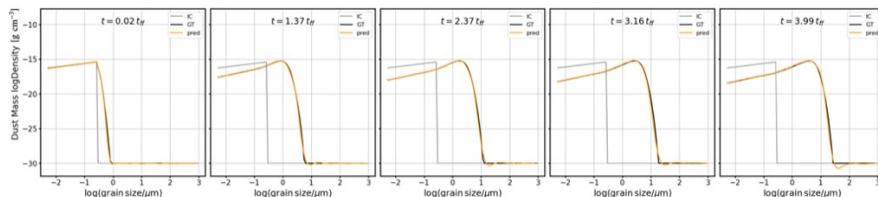


D. Aubert, O. Marchal, P. Ocavirk, M. Palanque, K. Tep (Obs. Strasbourg).
M.A. Breton, M. Delorme, A. Durocher (CEA). C. Cadiou (IAP).

En **bleu** la matière noire, en **orange** le gaz ionisé. On observe la réionisation de l'univers avec l'allumage progressif des premières étoiles et galaxies.

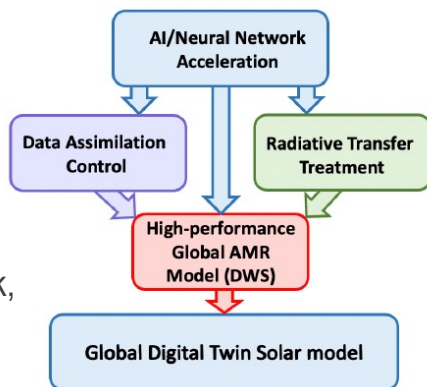
Réseaux de neurones entraînés pour modéliser la croissance des grains

Objectif Aller au-delà du traitement simplifié de la coagulation et de la fragmentation des poussières dans des modèles 3D complets de disques protoplanétaires

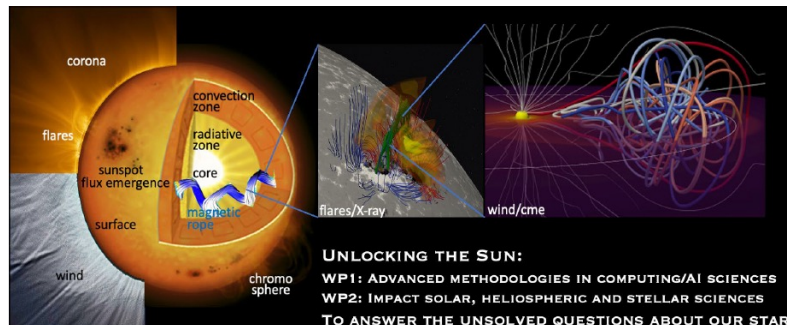
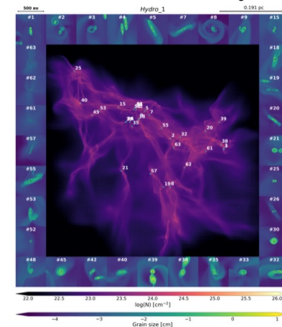


Collaboration: LMPA, LFEMI & Univ Heidelberg

ERC+
(Brun,
Strugarek,
Perri)

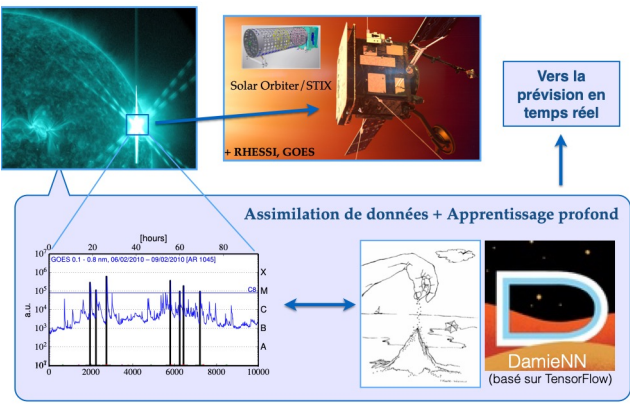


Hennebelle, Lebreuilly



UNLOCKING THE SUN:

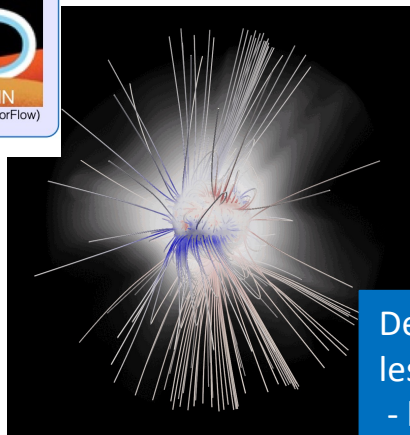
WP1: ADVANCED METHODOLOGIES IN COMPUTING/AI SCIENCES
WP2: IMPACT SOLAR, HELIOSPHERIC AND STELLAR SCIENCES
TO ANSWER THE UNSOLVED QUESTIONS ABOUT OUR STAR



SolarCast

Portail ESA: swe.ssa.esa.int

- **Concept original** pour résoudre un problème de fond (prévision des éruptions les + intenses, Vent solaires et high speed streams)
- Cycle solaire de 11ans
- Résultats **très encourageants** sur des données synthétiques et réelles



DeepLearning **essentiel** pour les deux verrous:

- besoin hardware (GPU)
- besoin RH pour aider au développement des pipelines Data assimilation SOLAR CAST (Flare, Solar and Wind Predict)

Verrous majeurs à débloquent:

- Accélérer l'assimilation de données pour fonctionner en temps réel
- Renforcer l'efficacité du pipeline sur des données observées

Projet COAST: Mieux coupler COAST + DYABLO + GALACTICA

Meso-Machine : jouvence Alfvén

Visualisation: Maintenir les moyens locaux pour analyses et visualisation
des gros jeux de données astrophysiques (simu ou obs)
machine gpus $\Leftrightarrow$ jouvence anais, storm, ecogal

Couplage: HPC + IA (PINNs) + Assimilation de données

RH: - Renforcer coté DEDIP les experts en HPC, parallélisme massif
- Continuer de recruter des astrophysiciens numériques pour maintenir
notre savoir faire et haute visibilité (source de bcp d'ERC)
- Créer un COAST fellowship