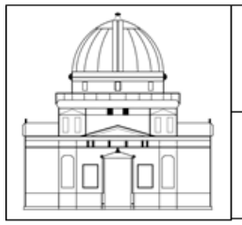


Observatoire **astronomique** de Strasbourg | ObAS

First Simulations of the Epoch of Reionization with Dyablo

Dominique Aubert, Grenoble, June 2026

On behalf of many collaborators : O. Marchal, P. Ocvirk, M. Palanque, K. Tep., N. Diet (ExaSKAle ANR/Obs. Strasbourg), P. Fernique, T. Boch (CDS), E. Thélie (U. Texas), J. Hiegel (IPHC), M. Andres-Breton, M. Delorme, A. Durocher (CEA), C. Cadiou (IAP), J. Chardin (CFCAL), M. Petrault, J. Sorce (Cristal Lille), S. Mesquita (LPSC)

D. Aubert, O. Marchal, K. Tep, P. Ocvirk, M. Palanque
ExaSKAle ANR Project
DYABLO EoR Simulation 1024x1024 x 4000 frequency channels

Neutral

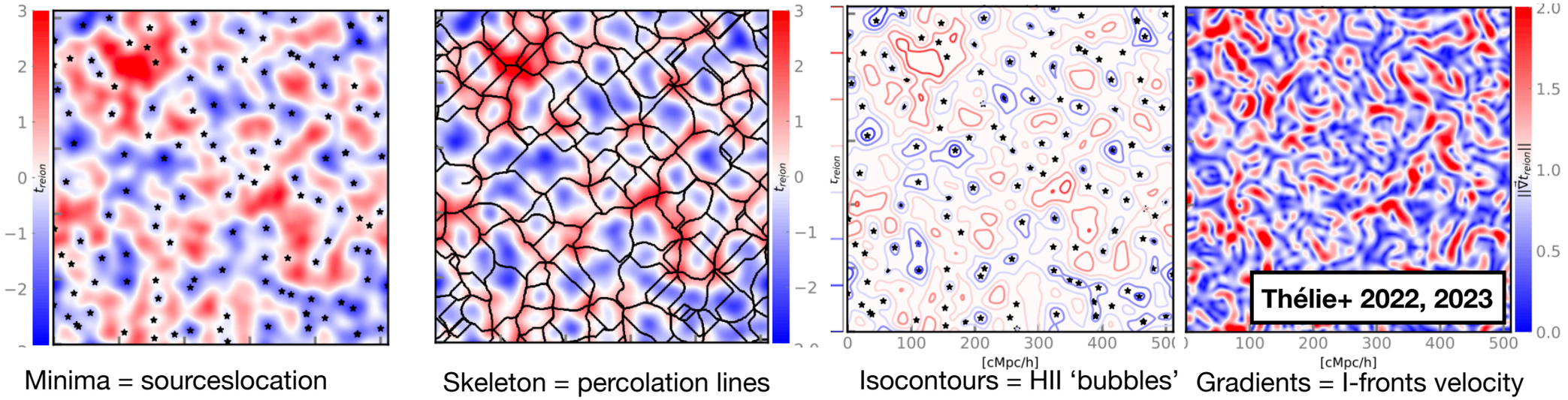
z= 11
t= 400 Myrs

Ionized

300 cMpc

z= 5.5
t = 1.1 Gyrs

The evolving geometry of ionization patches (SKA), driven by the rise of the first astrophysical sources (JWST) during the Epoch of Reionization (EoR).



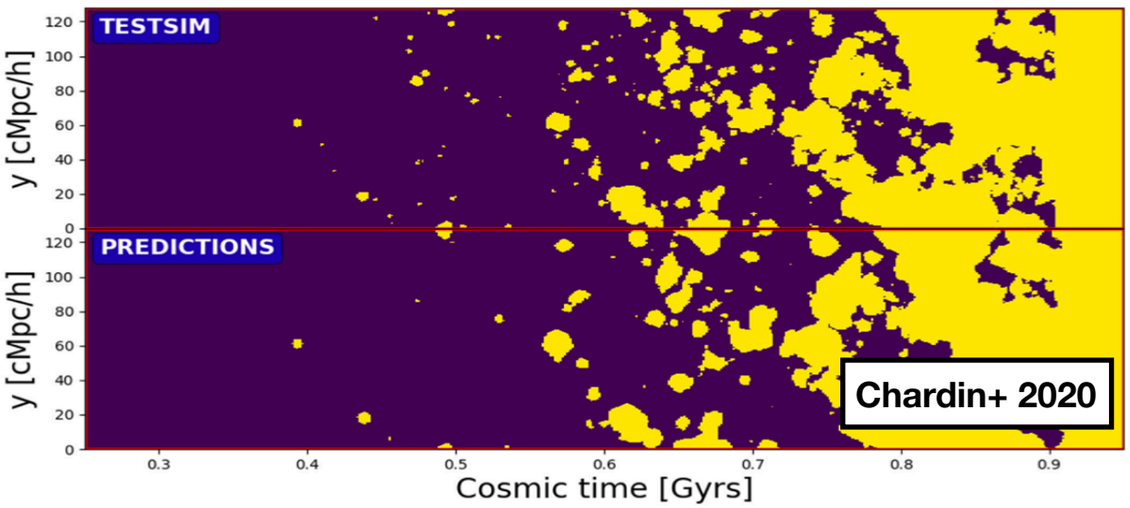
Minima = sourceslocation

Skeleton = percolation lines

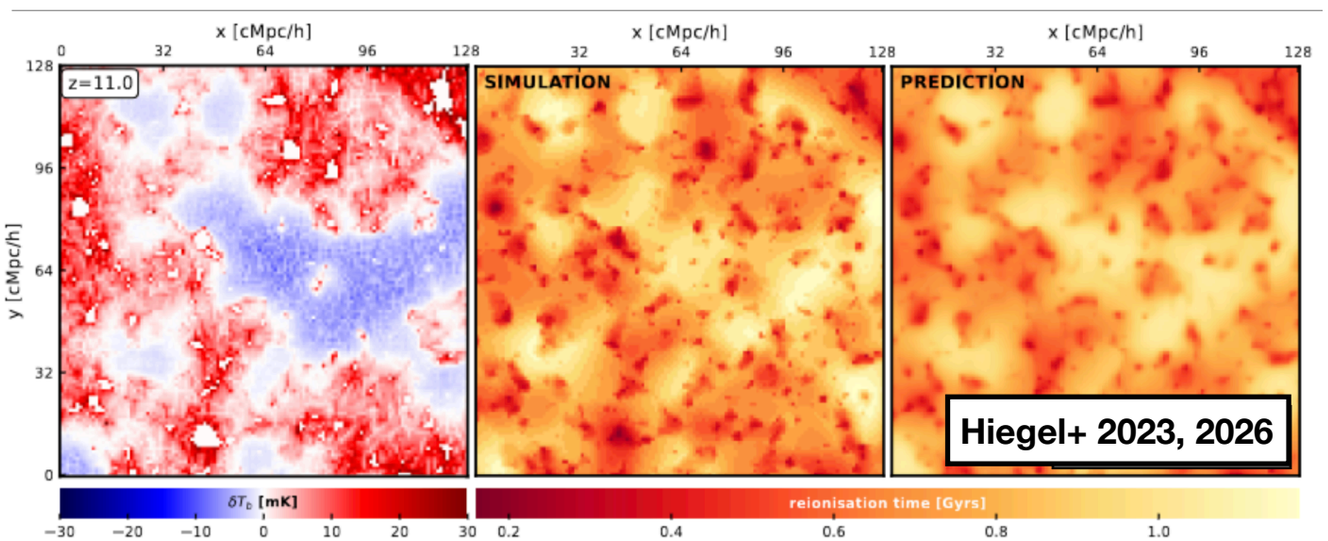
Isocontours = HII ‘bubbles’

Gradients = I-fronts velocity

The **evolving geometry** of ionized regions can be analyzed via **‘reionization times’ field and their topology**

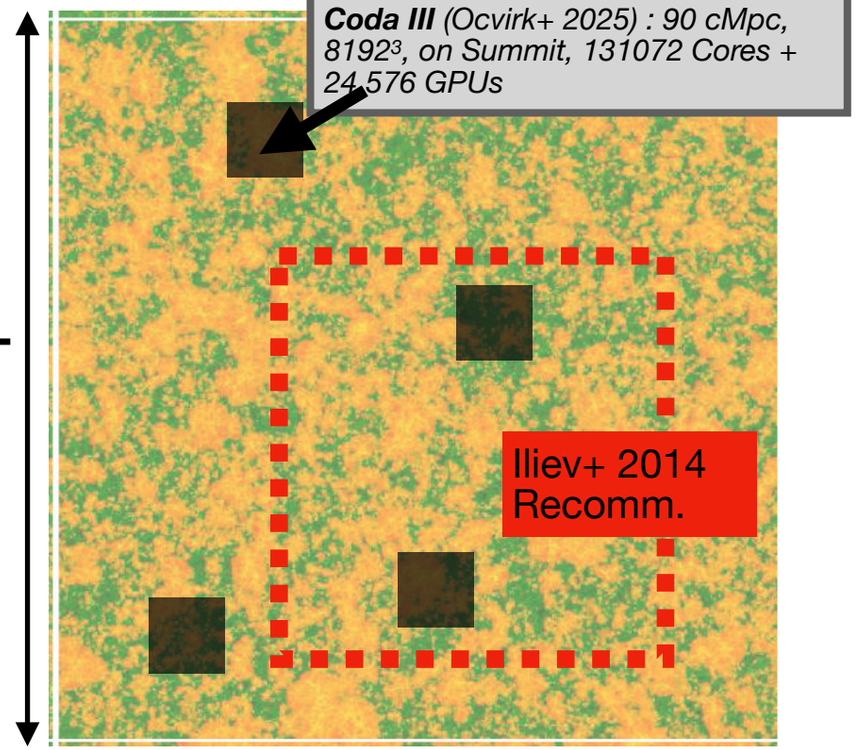


Emulation of reionization histories and spatial distribution using Convolutional Neural networks trained on 1024^3 + 3 AMR levels EMMA simulations (128 Jean-Zay GPUs)



Reionization time fields can be, in principle, reconstructed from ML post-processing of 21cm radio observation using e.g. **SKA**, **thus adding a temporal dimension to a radio image.**

EoR Simulations : Numerical Challenge



Let's say we need a simulated volume of $(250\mathrm{Mpc})^3$ that is able to resolve a $10^8 M_\odot$ DM halo with ~100 dark matter particles.The mass of a DM particle should be $m_{\mathrm{DM}} \sim 10^6 M_\odot$

Therefore such a simulation should have $N_{\mathrm{part}} \sim 8192^3 \sim 5.5 \times 10^{11}$ DM particles. That's a lot, especially if we add the same number of resolution elements for hydrodynamics and RT.

Conservatively, we therefore need at least 15 numbers per resolution element, that's around 70 TB (double precision) of 'information' to handle. Such amount requires 10 000s of CPUs to run.

Large Volume + High resolution

Ideal

Large Volume + Low Resolution → Statistically Representative but Sources/absorbers are missing

Small Volume + High Resolution → Small scale absorbers and small sources are included but lack of representativity

Dyablo

- **Massively parallel** for Exascale supercomputers
- **Hybrid computing** using local shared memory parallelism : many-cores CPU (à la OpenMP), GPUs (from any vendors) or possibly any kind of current and future computing devices.
- Capable of and optimized for **evolution** : aiming for 20+ years of usage
- **Separation of concerns** between computer scientists (infrastructure, framework, parallelisation) and astrophysicists (physics modules, production features, analysis)
- **« Modern » Software development** philosophy (versioning, branch management, continuous integration, documentation, support)

Main Developers : **A. Durocher & M. Delorme** (DEDIP/CEA) building upon first developments made by **P. Kestener** (DEDIP/CEA)

with contributions from :

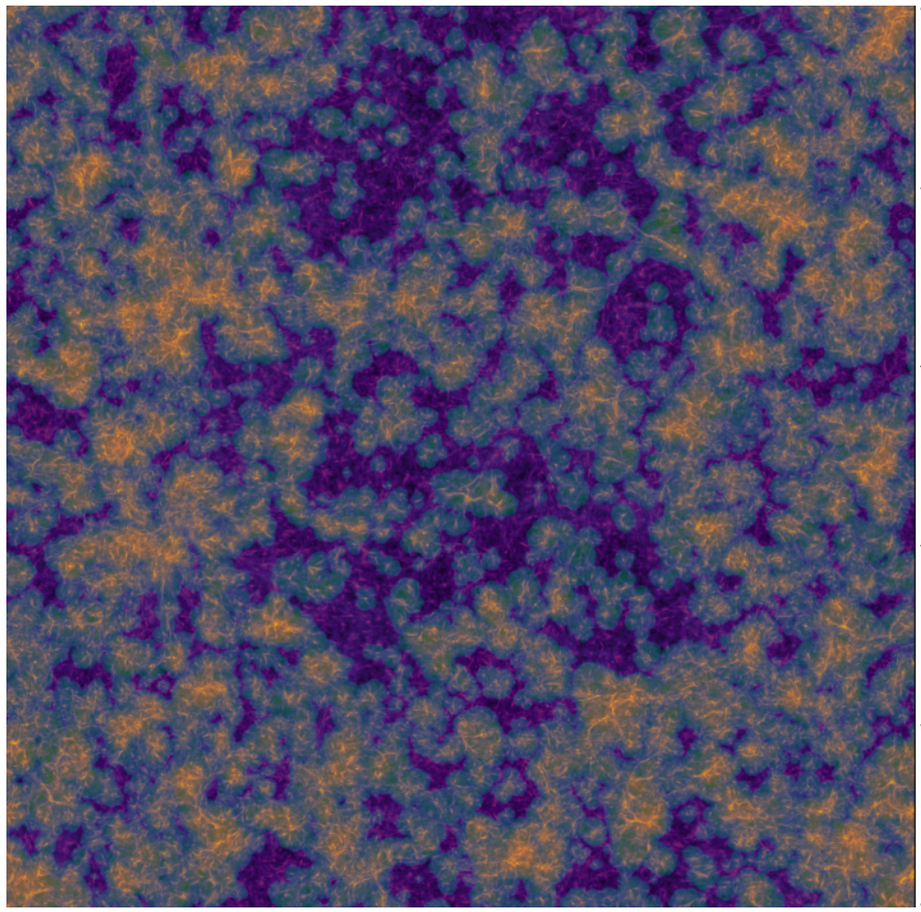
- **A.S Brun, M. Delorme & Collaborators**: (Solar Physics, MHD, ISM), CEA-Saclay
- **D. Aubert, O. Marchal, P. Ocvirk, M. Palanque, K. Tep** (Cosmology + Radiative Transfer), Strasbourg
- **M. Doebele, M.A. Breton** (Geometry, Gravity), CEA-Saclay
- **P. Tremblin** (MHD, Gravity), Maison de la Simulation
- **M. Petrault, J. Sorce** (Subgrid Physics), Cristal, Lille
- **C. Cadiou, A. Storck (**Subgrid physics), IAP Paris
- **S. Han, Y Dubois** (Cosmic Rays), IAP, Paris
- **L. Sewanou, B. Commerçon** (Dust), CRAL, Lyon)

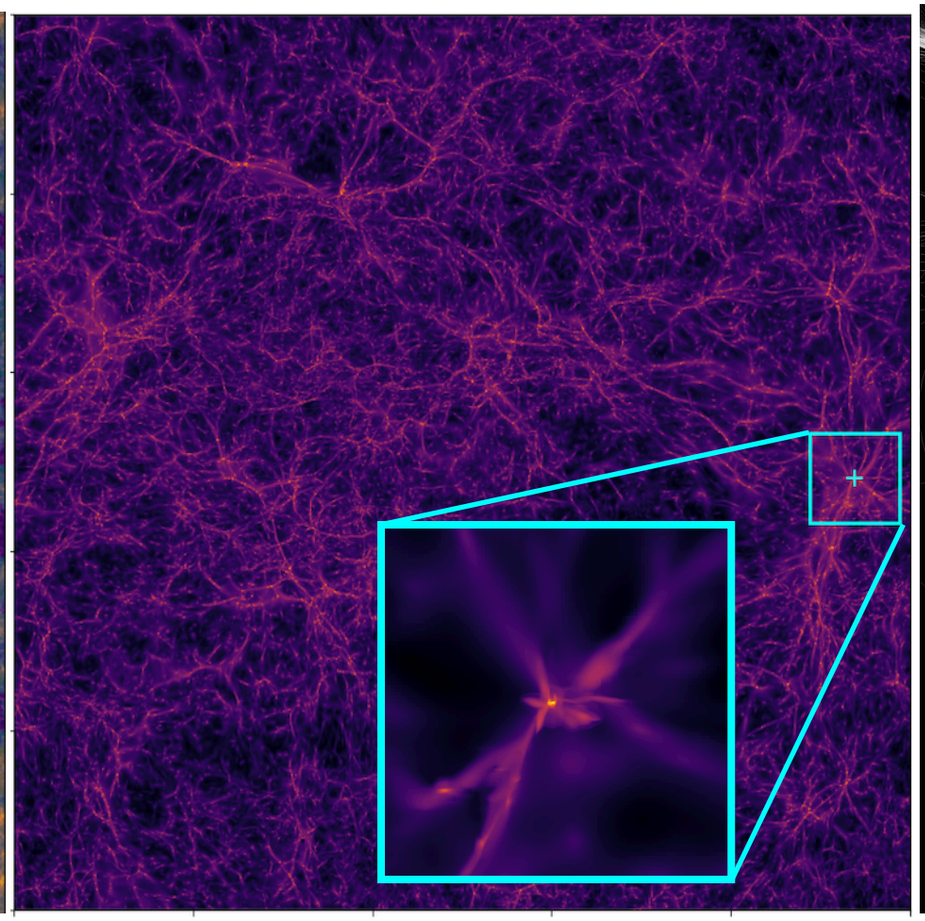
kokkos/**kokkos**

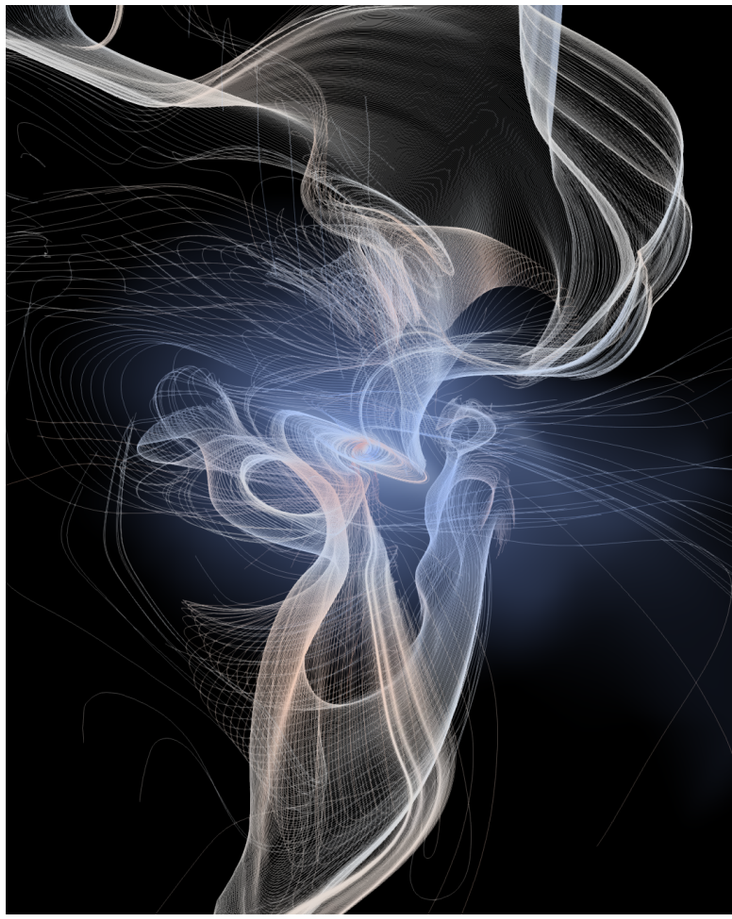
Kokkos C++ Performance Portability Programming Ecosystem: The Programming Model - Parallel Execution and Memory Abstraction

MPI

Dyablo Cosmo for Reionization

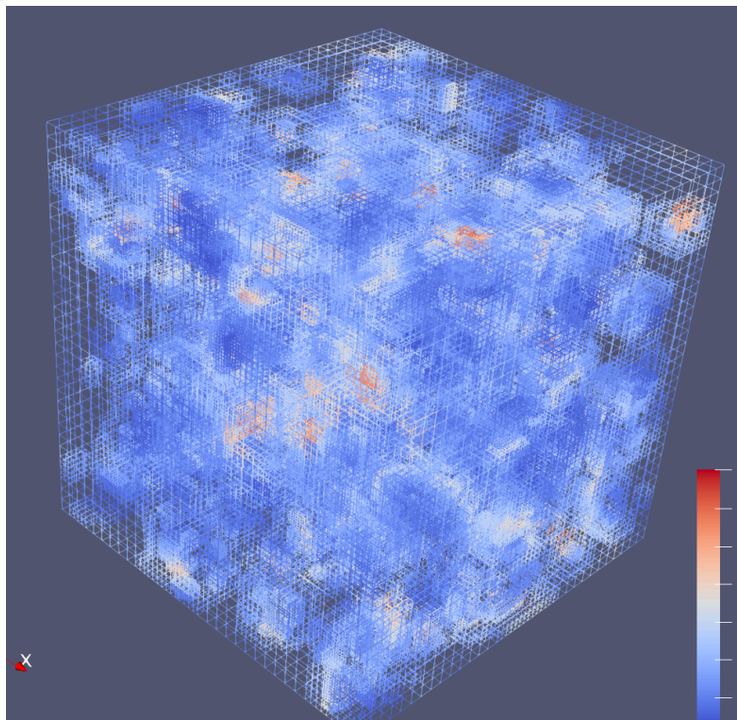






Current physics features

- Block-Based Adaptive Mesh Refinement
- grid-based hydrodynamics
- dark matter particle-based collisionless dynamics
- multigrid solver for self-gravity, from gas+particles
- multi-frequency M1 radiative transfer
- supercomoving coordinates to account for physics in an expanding Universe
- quasi-lagrangian refinement for cosmology
- cosmological initial conditions ingestion (grafic format) + Zoom
- star formation + feedback
- H/He equilibrium cooling, H out-of-equilibrium thermochemistry

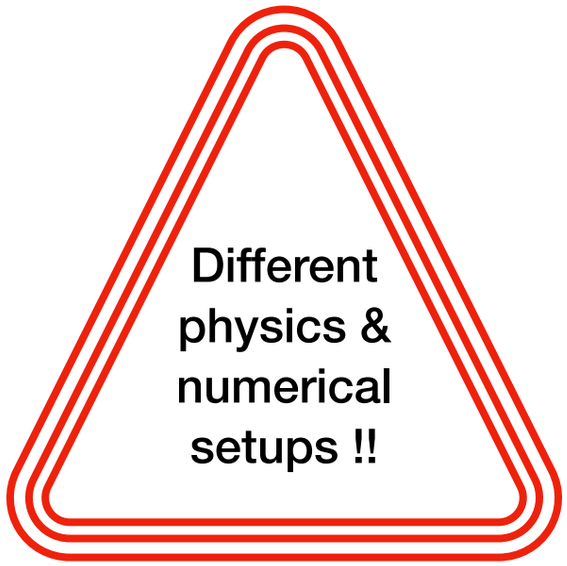


For **EoR simulations**

(Radiative Hydrodynamics, H out-of-equilibrium thermochemistry, star formation, stellar feedback)

On our local server in Strasbourg:

- Large scale simulation for SKA forecast on 1 single Nvidia H100 GPU
  - (300 cMpc - 512)^{3,} c=0.1, z= 5.6, ~1.4 GPU hours
  - restitution times similar to semi-analytical codes such as 21cmfast.



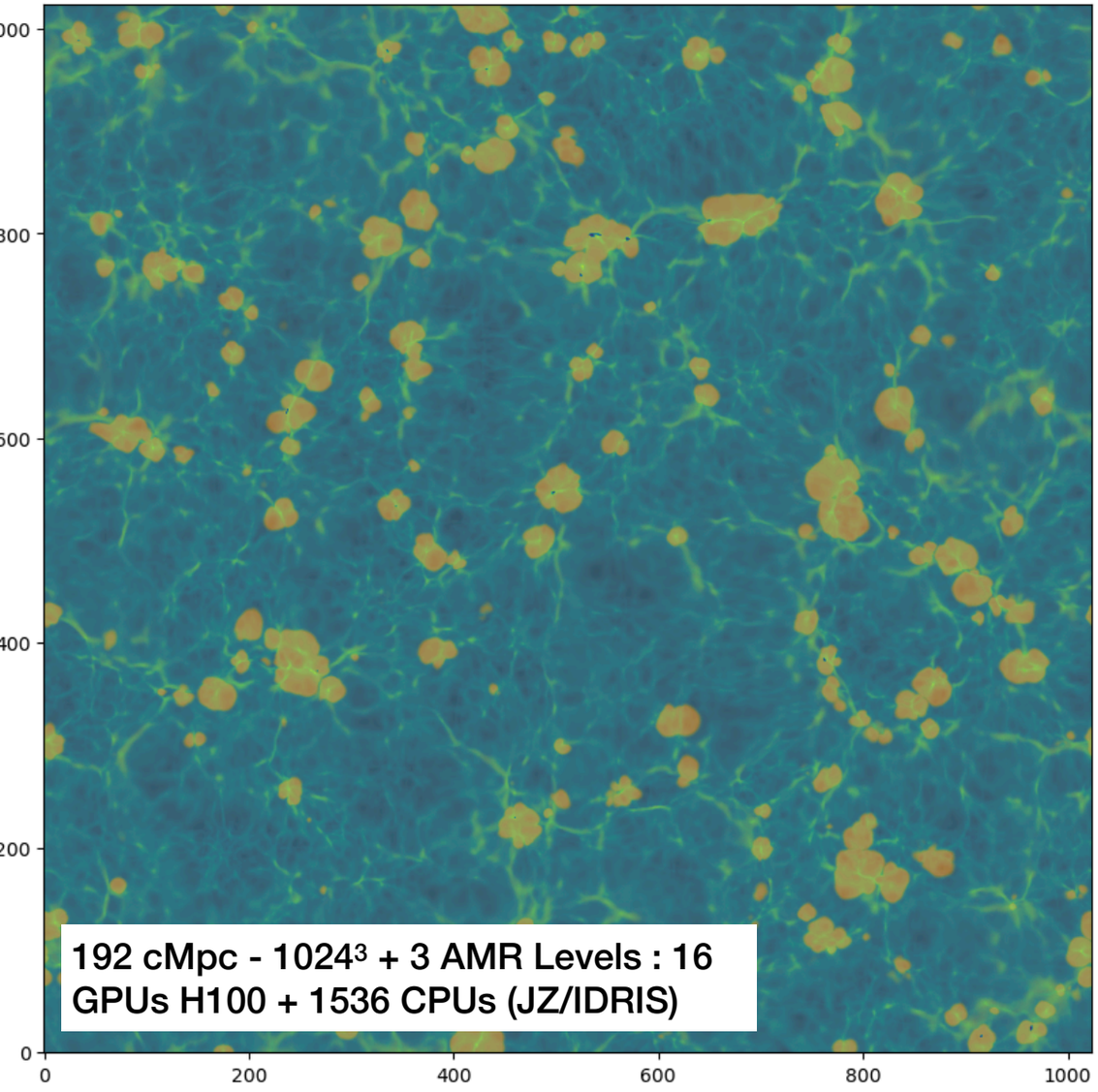
On JZ :

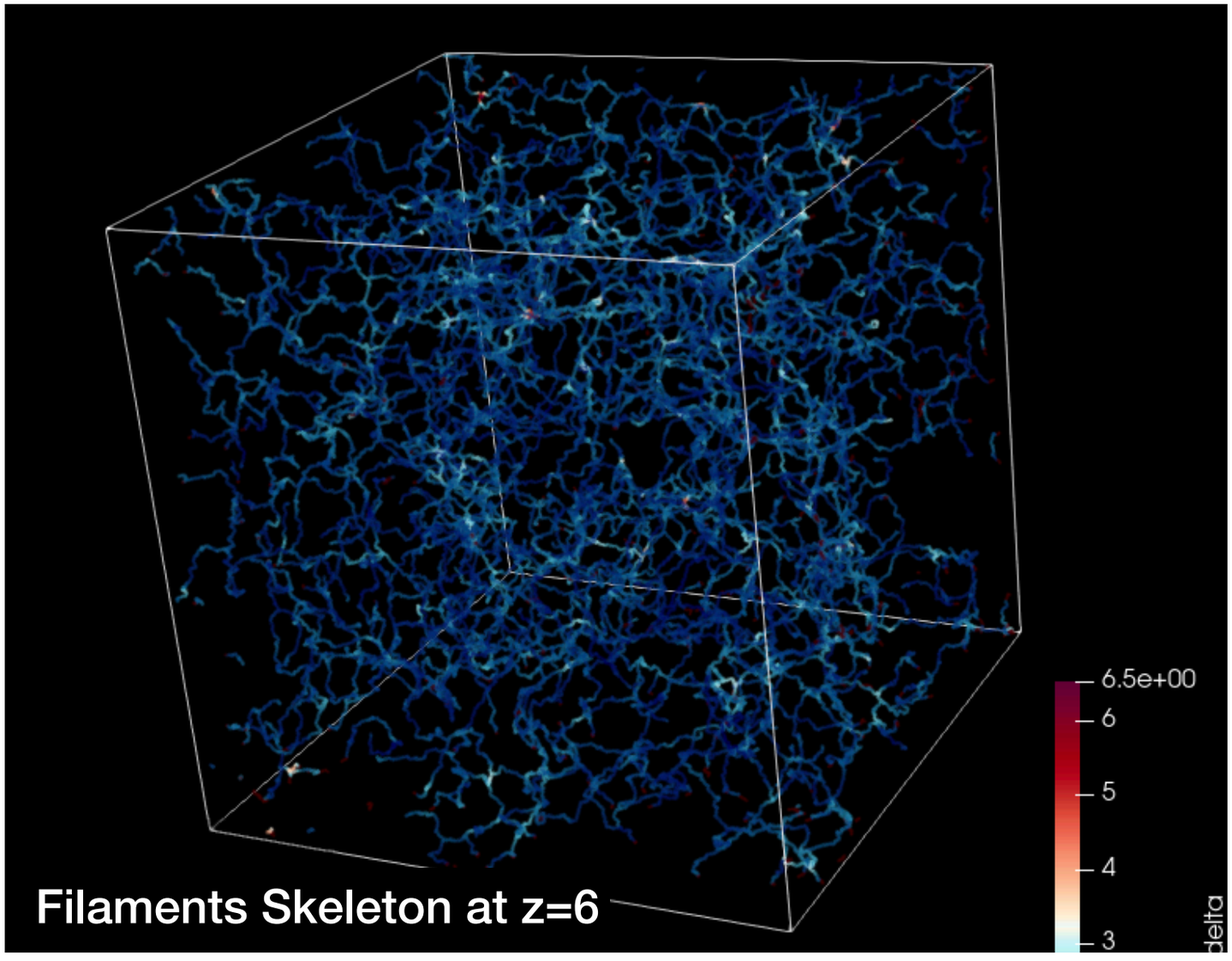
- Small scales simulations for galaxy formation during the EoR
  - 4 x Nivida H100 GPUs (12 cMpc - 512)^3, c=0.1, z= 5.6, ~ 220 GPU hours
  - 16 x Nivida H100 GPUs (24 cMpc - 1024)^3 c=0.1, z= 5.6, ~ 2500 GPU hours

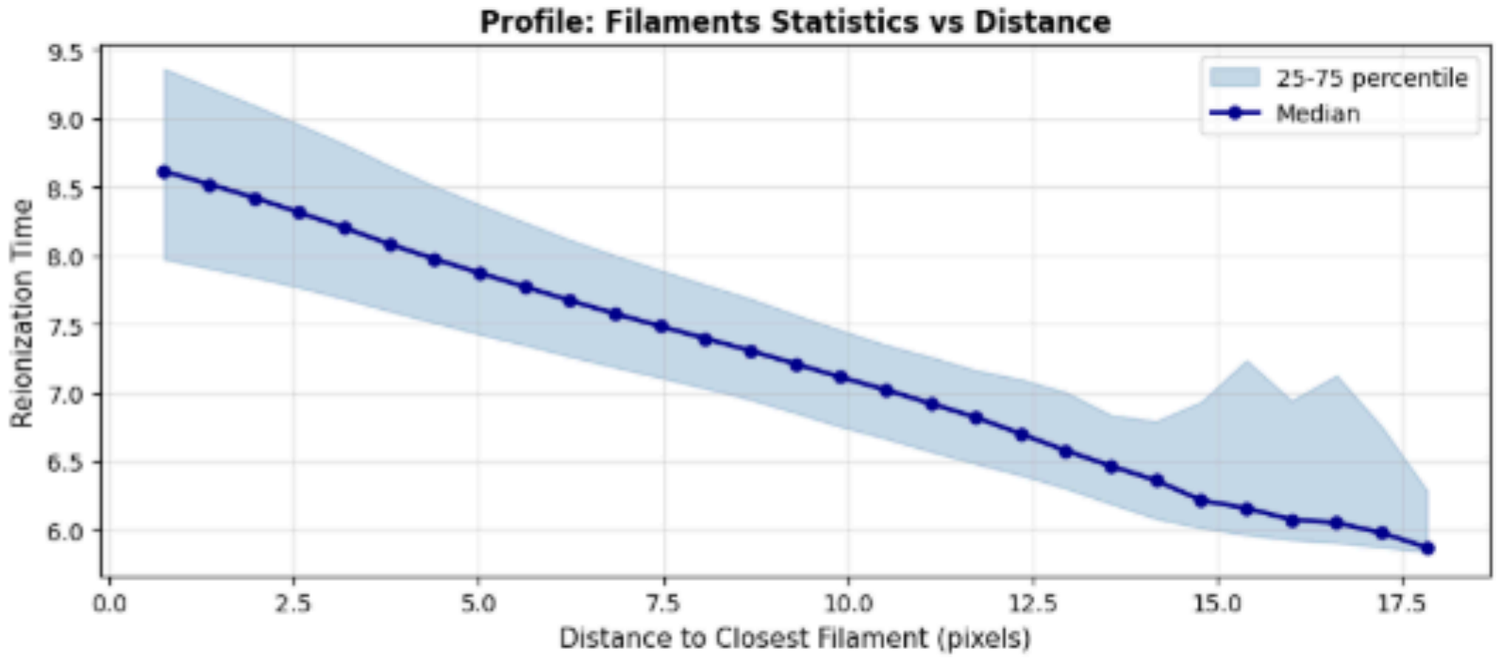
For the moment, overall scaling (both strong and weak) is limited by the multigrid poisson solver. Optimizations are under way. Not worried.

Meanwhile hyperbolic solvers scale (strong and weak) perfectly on admittedly small parallel configurations (<32 H100 GPUs).

Overall though : surprised by the single-GPU performance and how science-ready product can be produced with « small configurations ». Almost limited by our ability to ingest products.

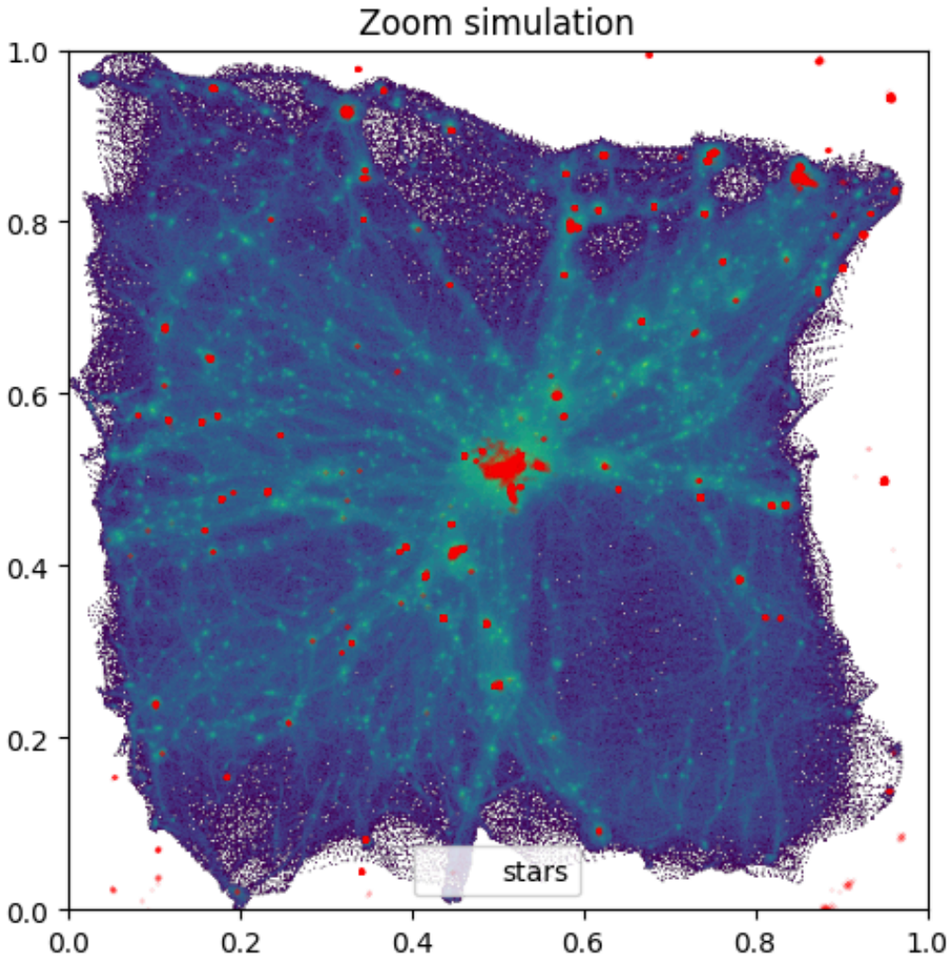


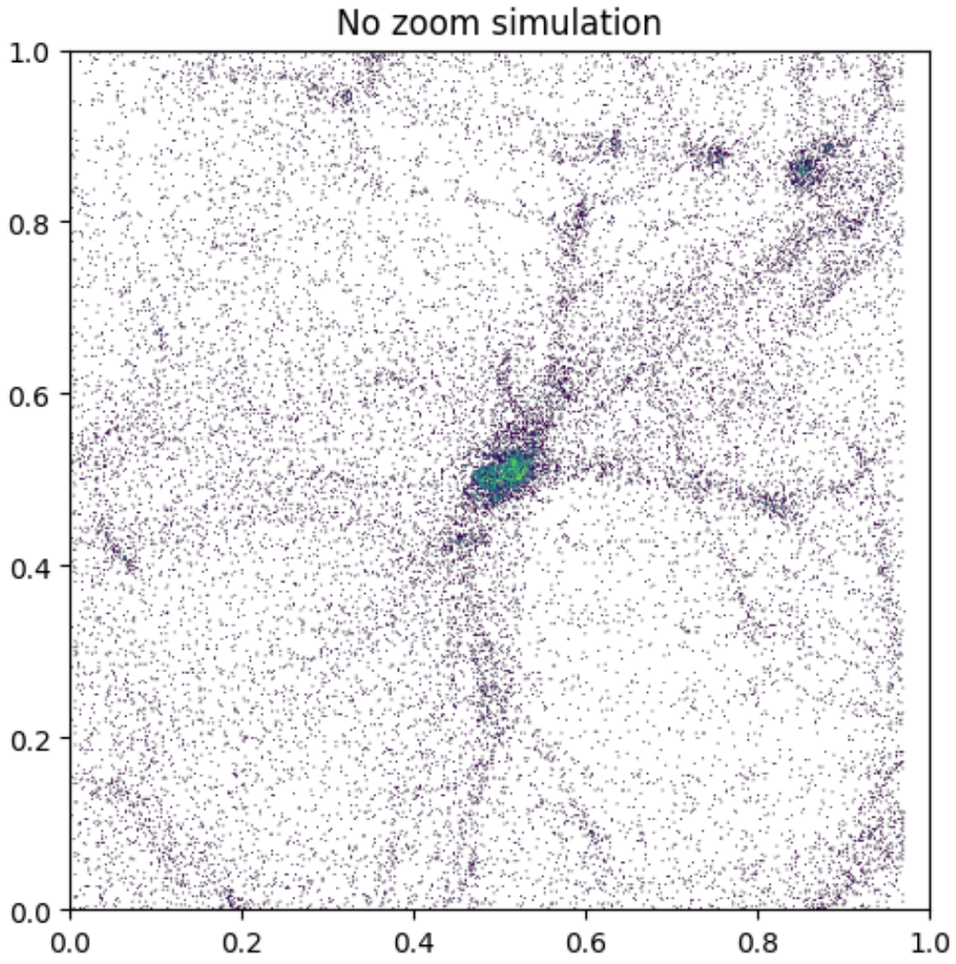


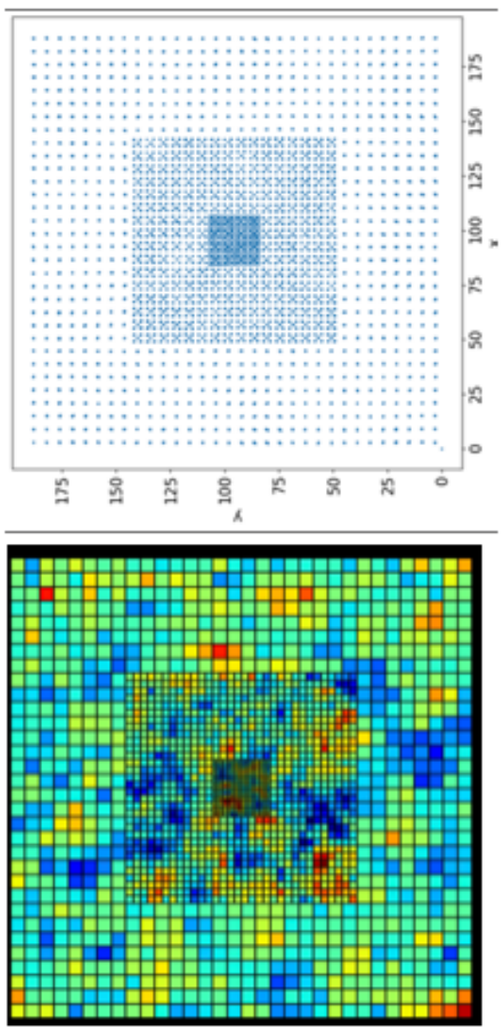


Example of EoR Simulation ->
Filamentary regions reionize earlier +
The closer you are to a filament the earlier you reionize

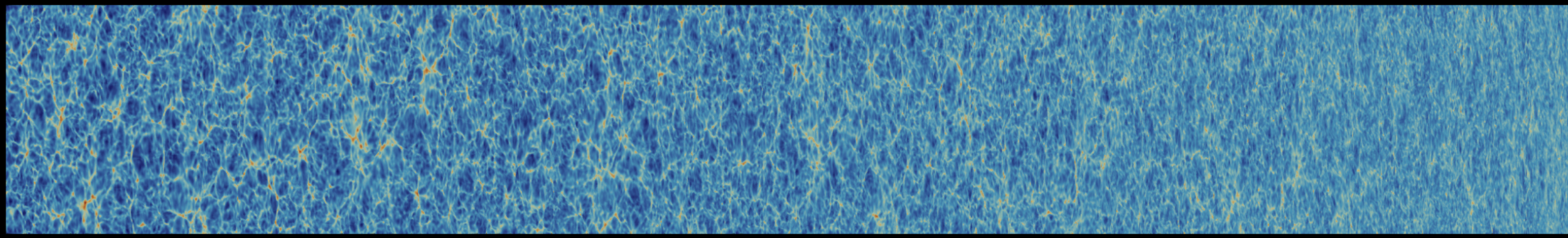
Zoom Simulations

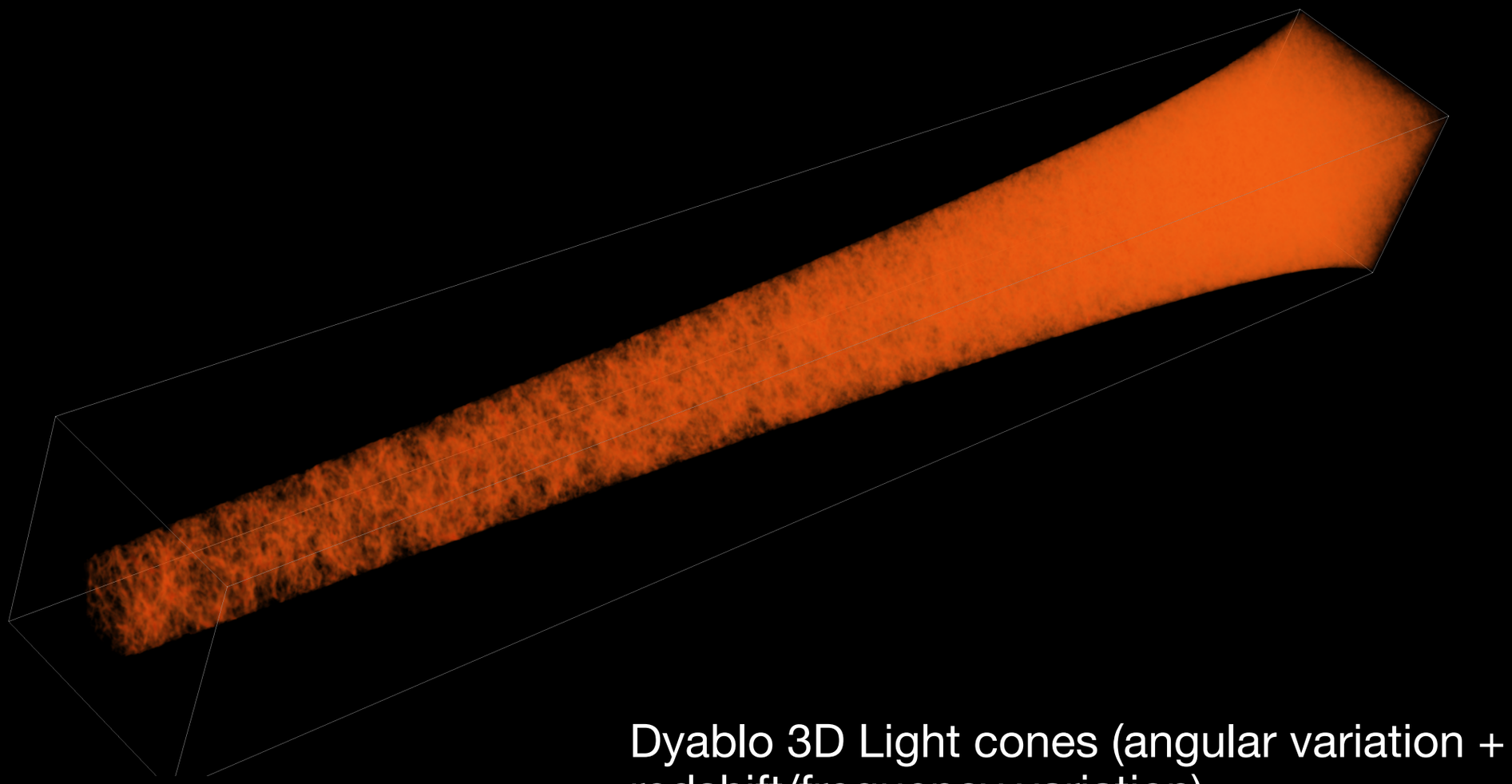






When performing zoom simulations, we specify the hierarchy of directories from which to extract the data at different resolution, as well as the location/extent of each resolution zone. Particles with different masses are also generated. Refinement is restricted to the most zoomed-in region



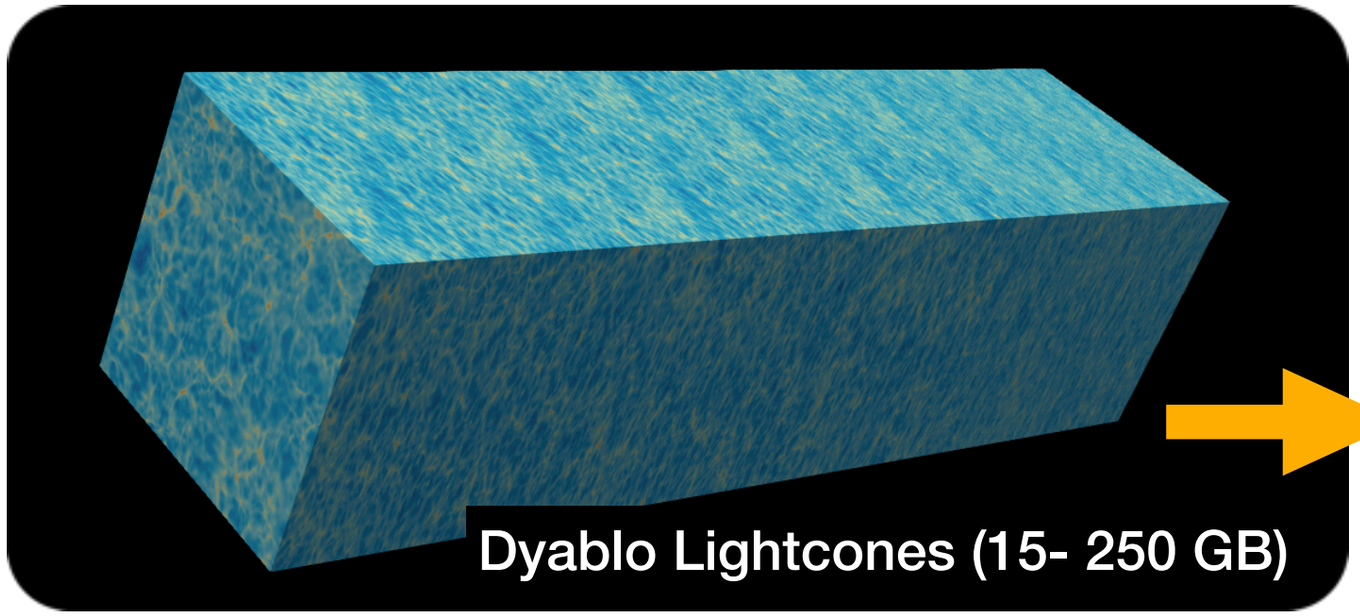


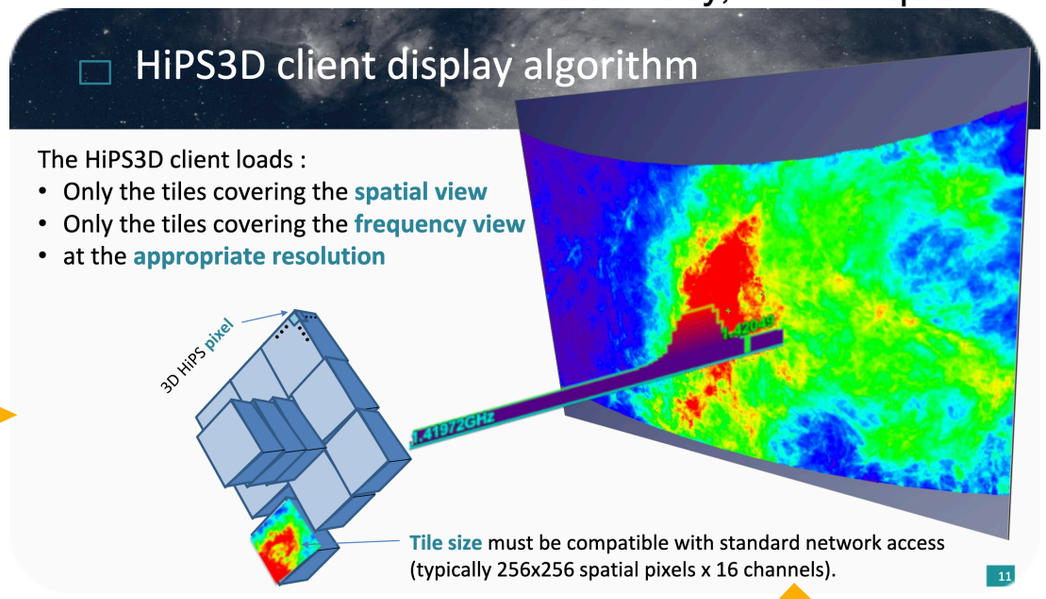
Dyablo 3D Light cones (angular variation + line-of-sight temporal/ redshift/frequency variation)

=> SKA-science forecast of 21cm radio emission

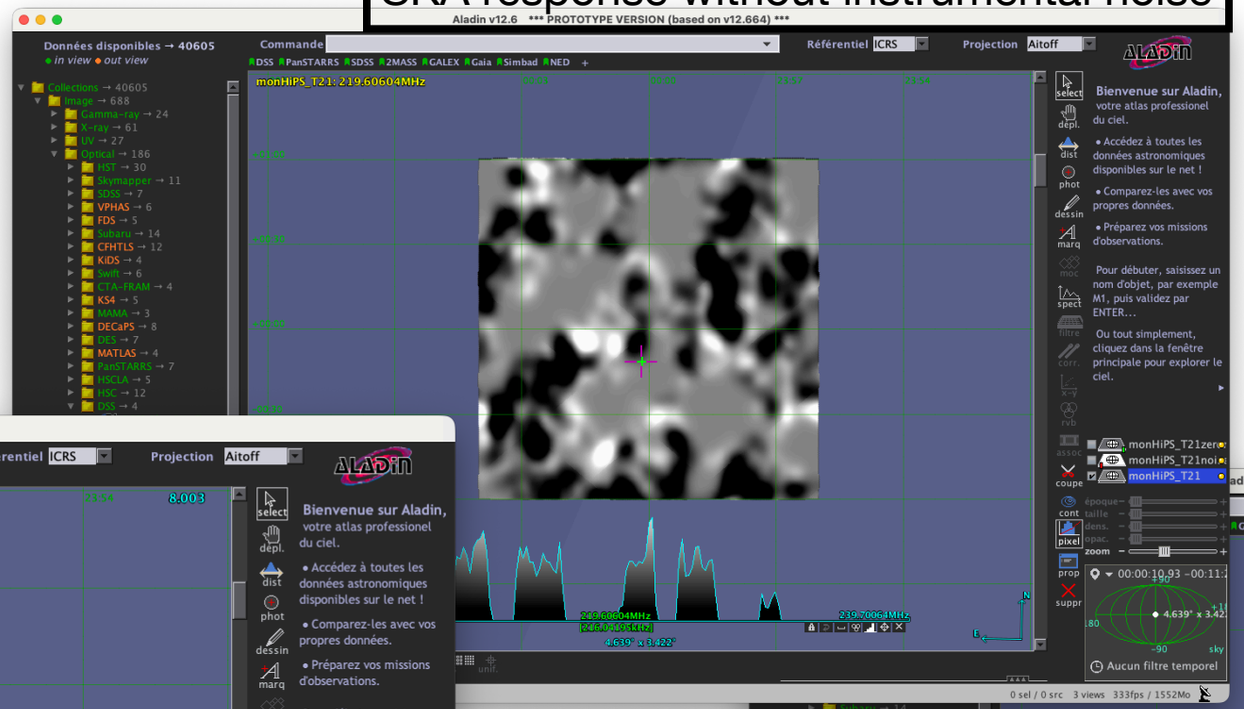
=> Mock surveys for massive dataset remote exploration developed at CDS

Courtesy, P. Fernique

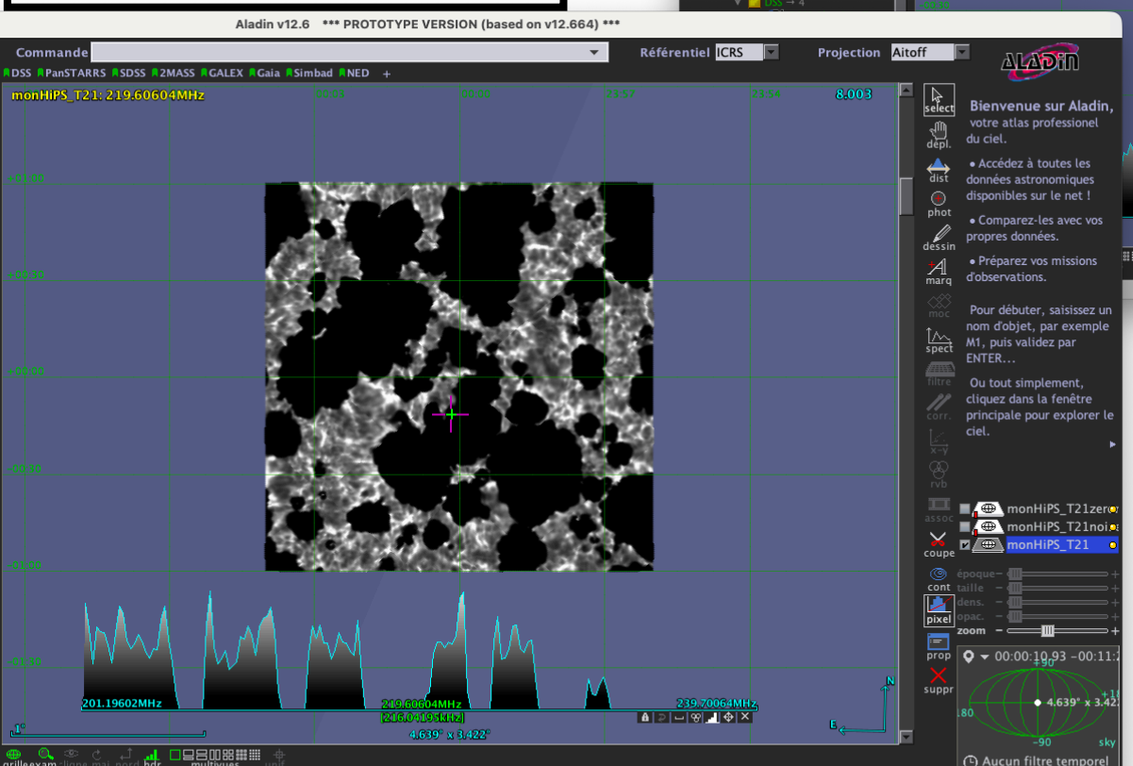


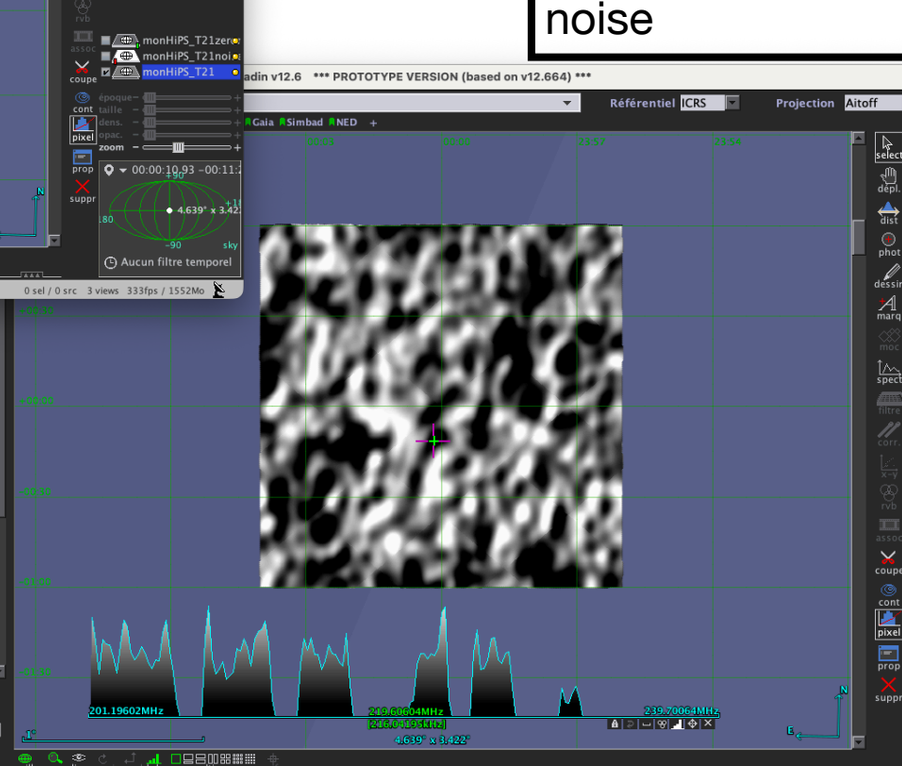


SKA response without instrumental noise

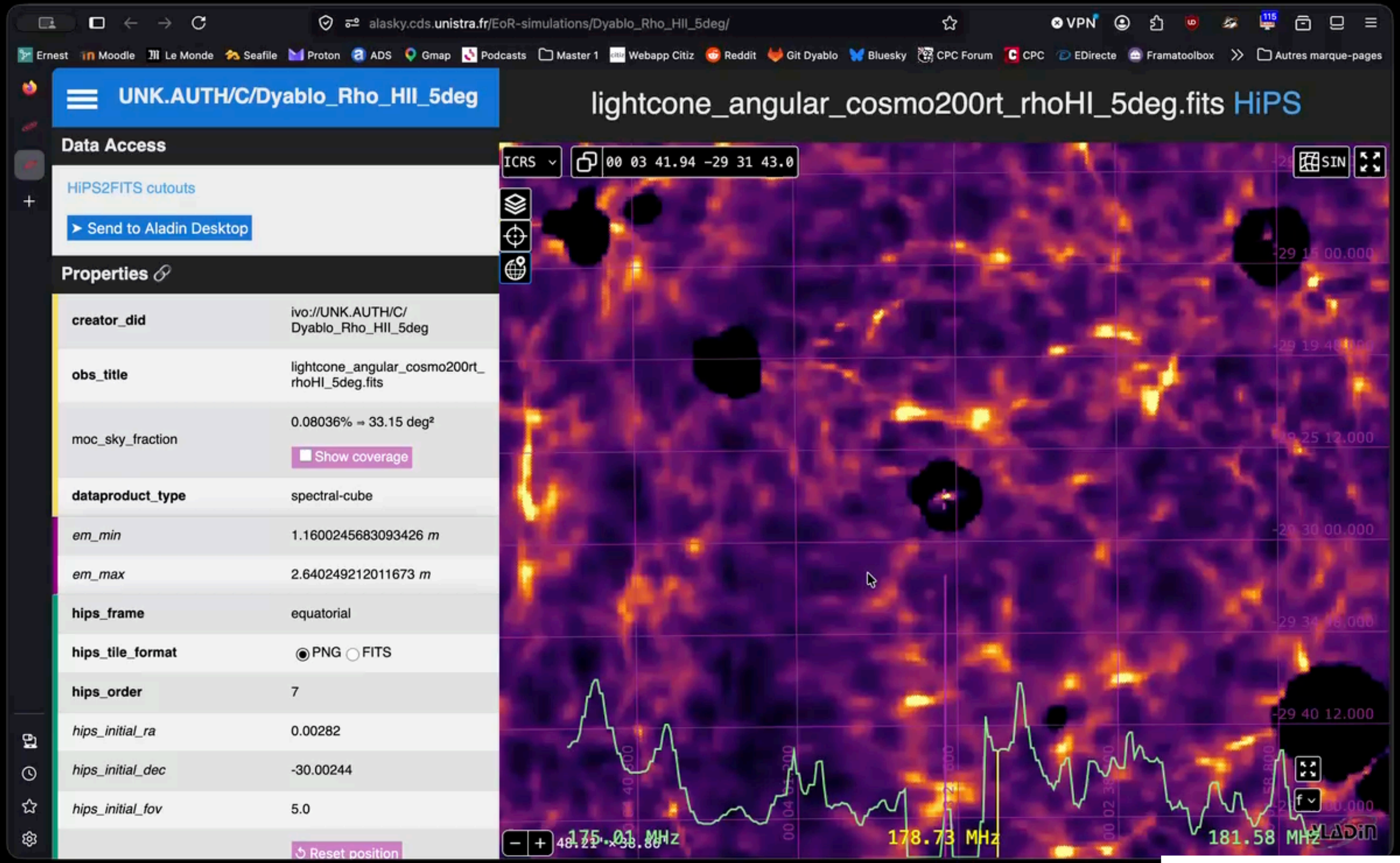


Pure 21cm radio emission





SKA response with instrumental noise



Field 5deg x 5deg, 2048 x 2048 x 6000 frequency channels
Final HiPS3D dataset ~ 250 GB
Remote browsing via Alasky (and Aladin Desktop)

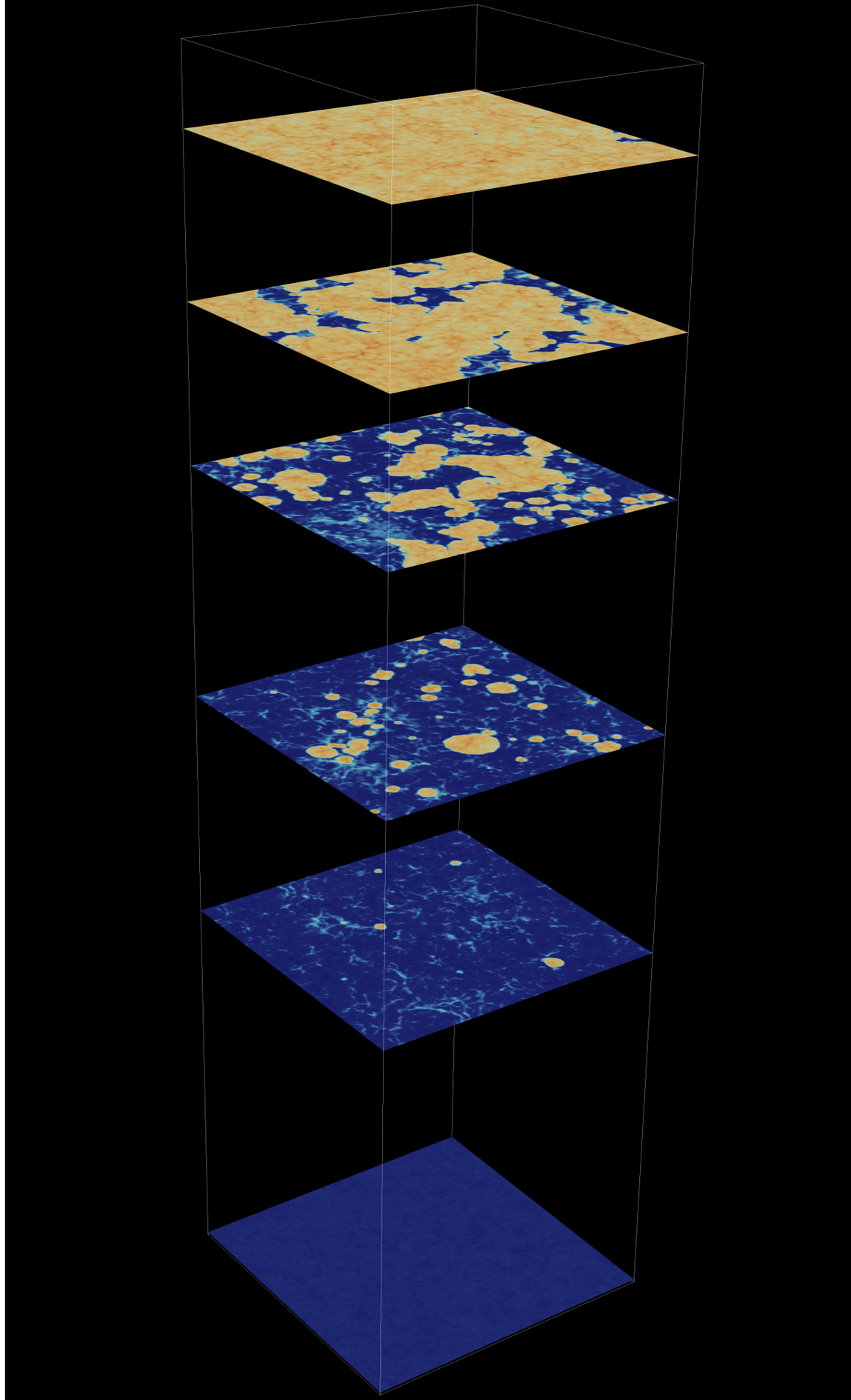
Take Away Message : ‘Dyablo-EoR’ lives

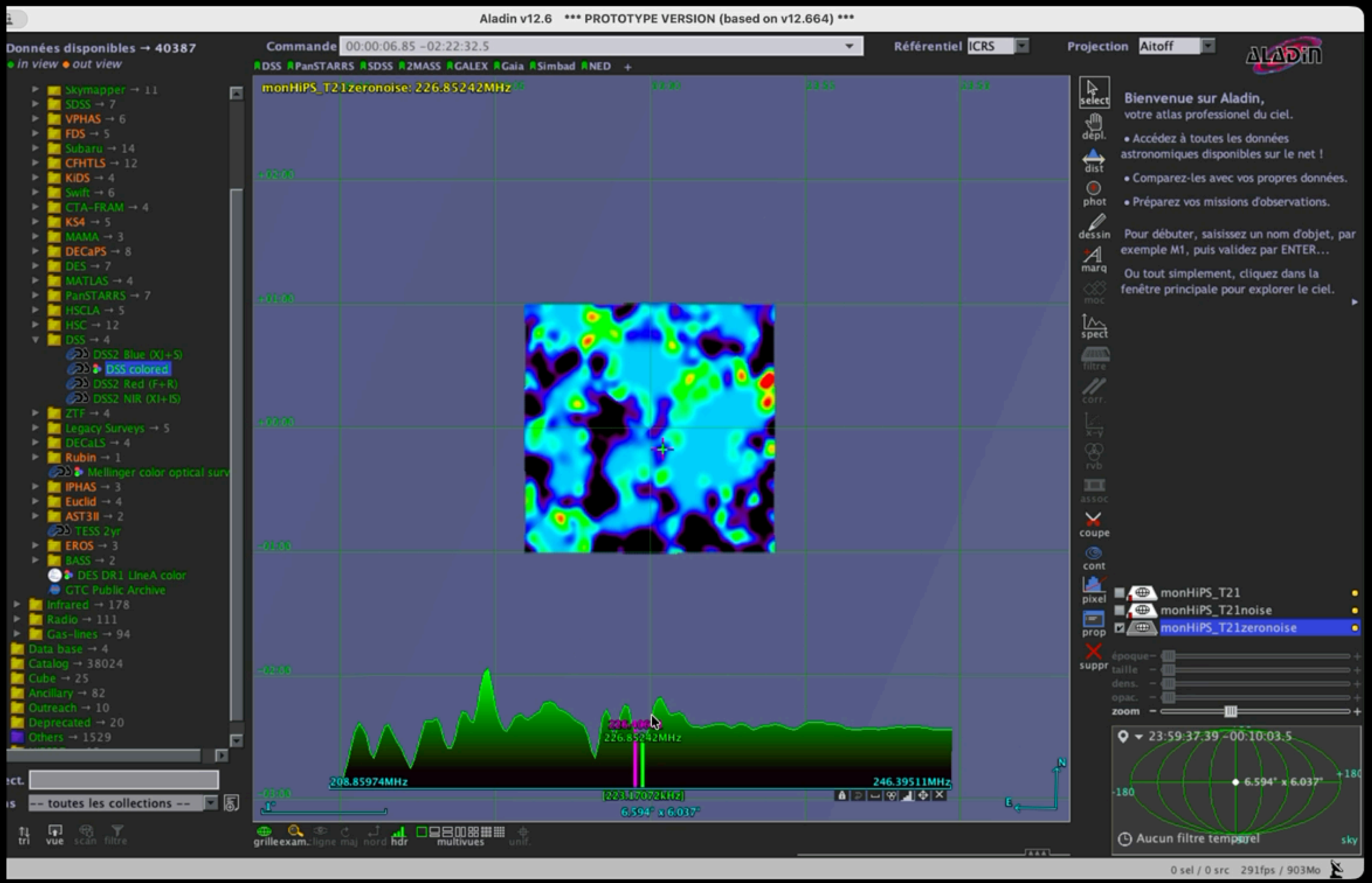
We are almost there for ‘Dyablo-SKA’ i.e. scaling up to models and datasets at the level of SKA requirements.

Complement other existing simulation (e.g. SPH Licorice by B. Semelin) or semi-analytical models (21cmfast or Zeus21).

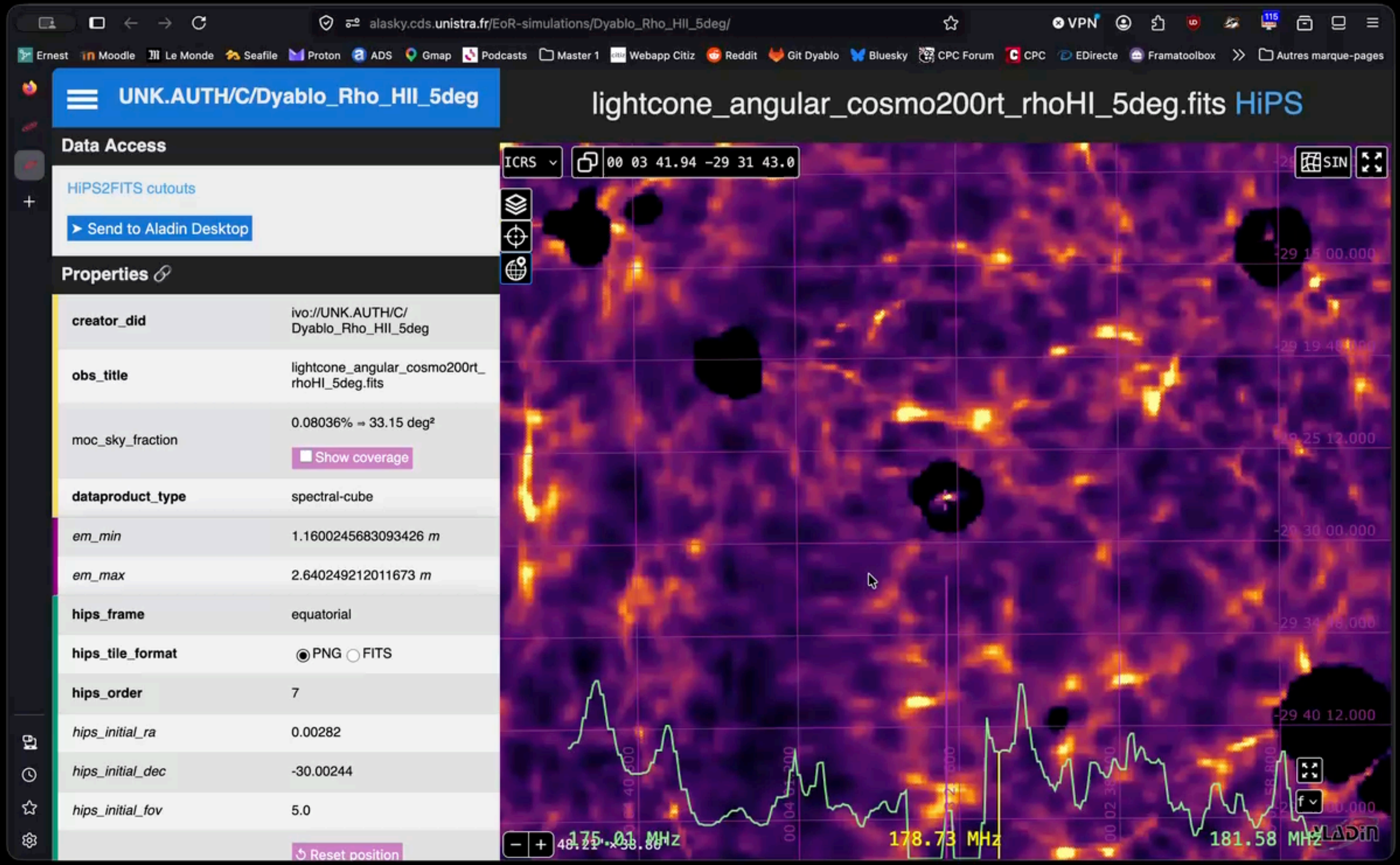
Features underway :

- X-Rays (Almost Done) / Multi-species thermochemistry
- Zoom simulations (Almost Done)
- subgrid model for sources at low resolution (Almost Done)
- advanced radiative transfer methods beyond M1
- writing a paper…





SKA-Low AA4 - 1000 hours - 2 km Baseline , made with tools21cm (Giri+ 2020)
2 deg x 2 deg, 4000 frequency channels (110 MHz - 250 MHz)
2048 x 2048 x 4000 (15 Gb)



Field 5deg x 5deg, 2048 x 2048 x 6000 frequency channels
Final HiPS3D dataset ~ 250 GB
Remote browsing via Alasky (and Aladin Desktop)