UC Berkeley, November 16, 2016

Solving cosmic structure formation with computational methods

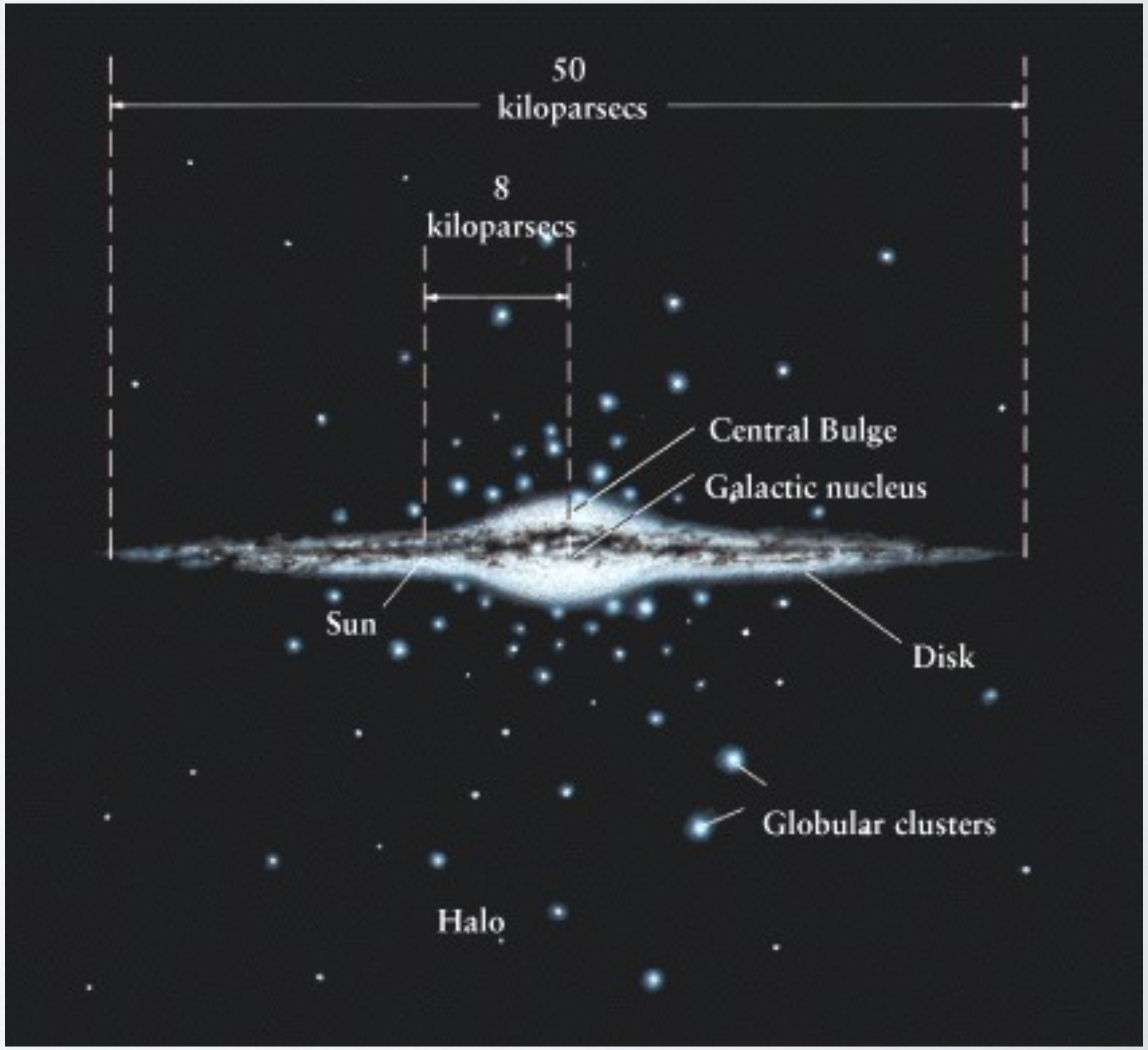
Davide Martizzi
Astronomy Department UC Berkeley

Overview of this talk

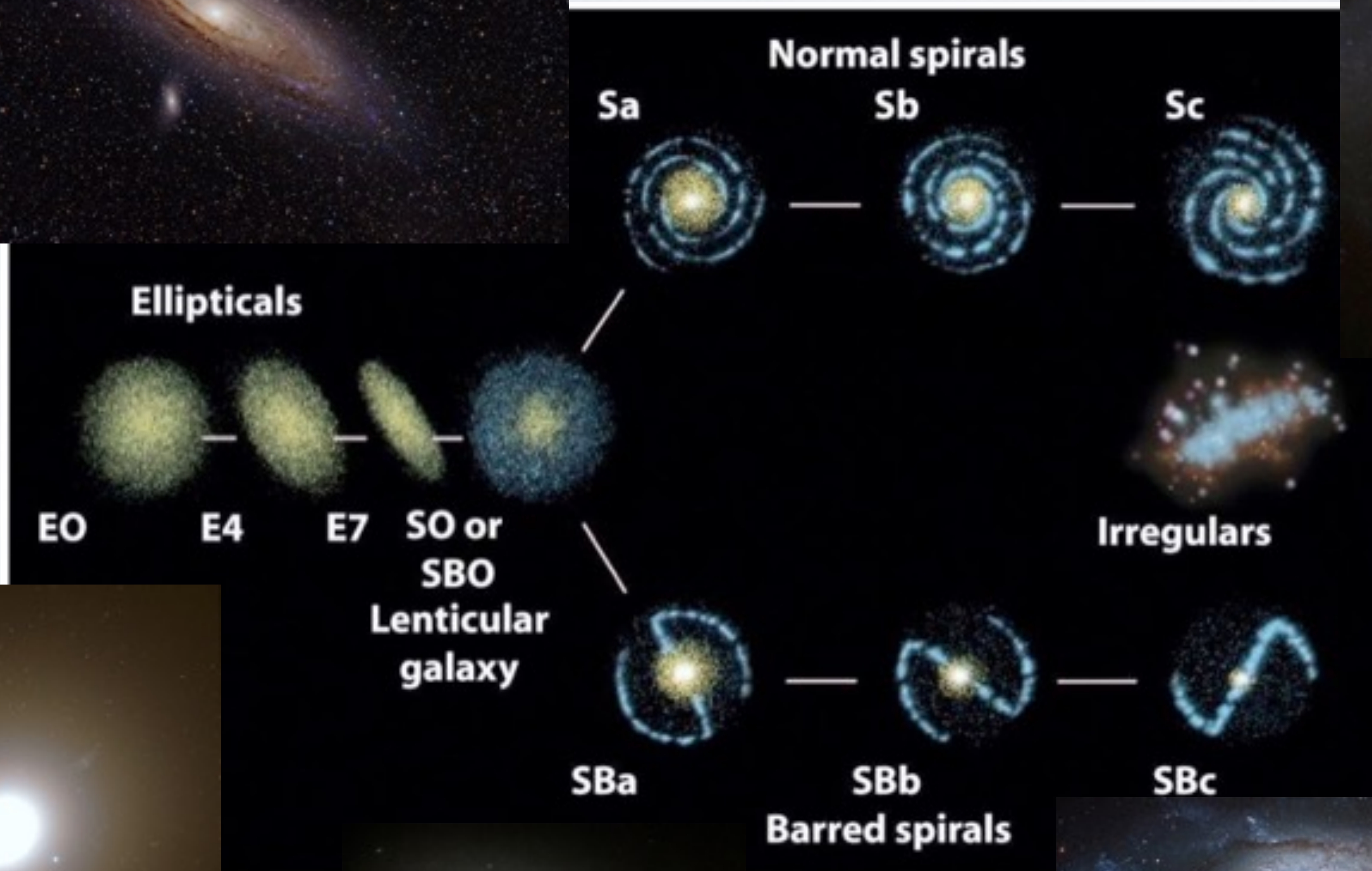
- Motivation for this research.
- Basic concepts of the theory of cosmic structure formation.
- Simulation techniques and challenges in the field.

The Milky Way

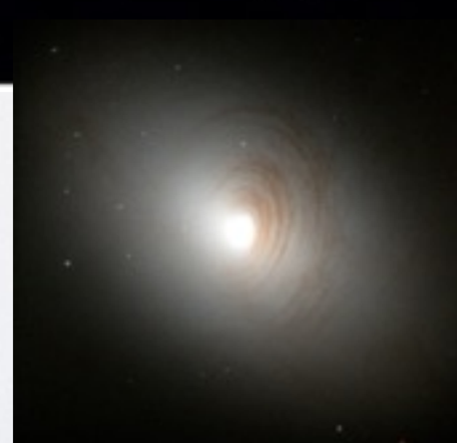
$1\ pc = 3.0857 \times 10^{16}\ m$
$8\ kpc = 2.5 \times 10^{20}\ m$



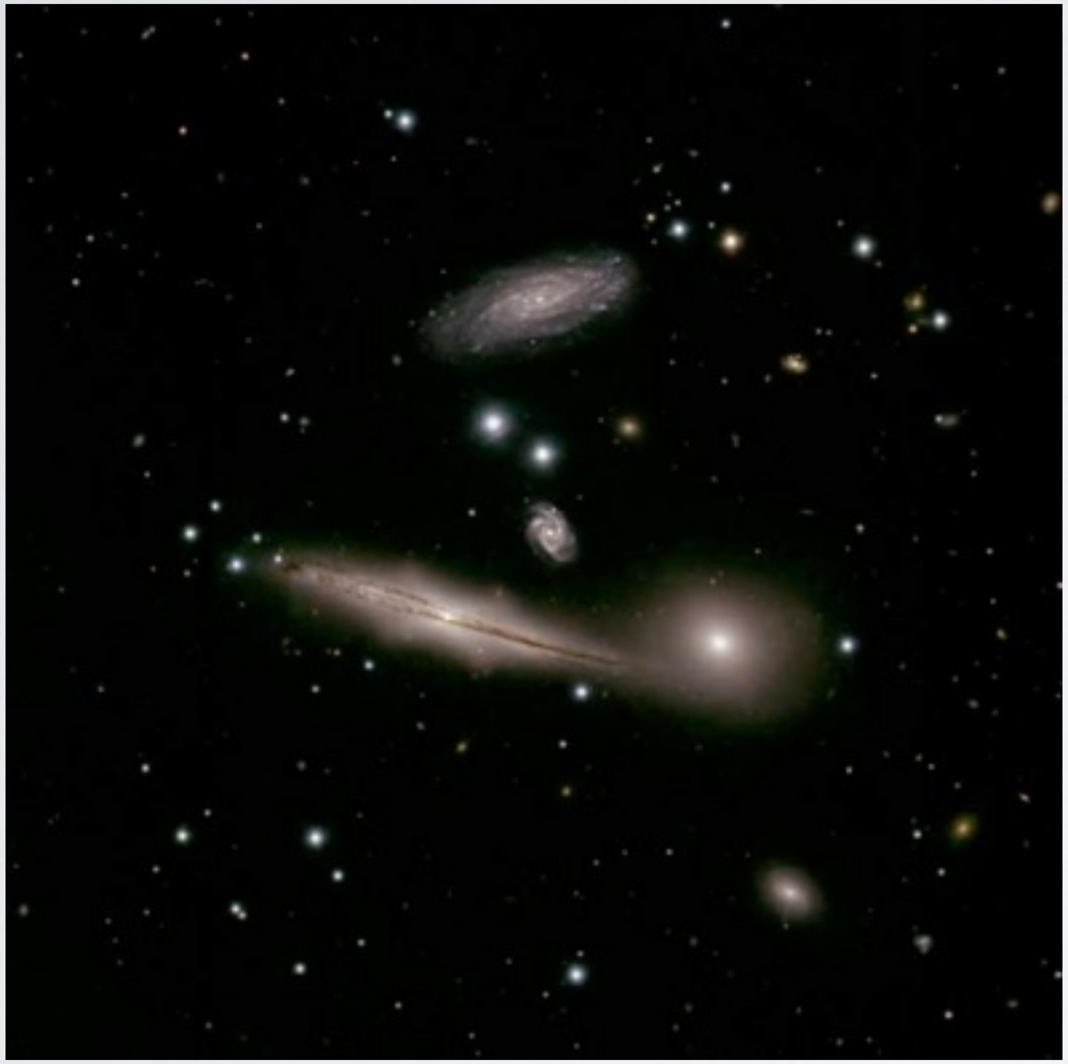
Galaxies



Up to 100 billion stars!



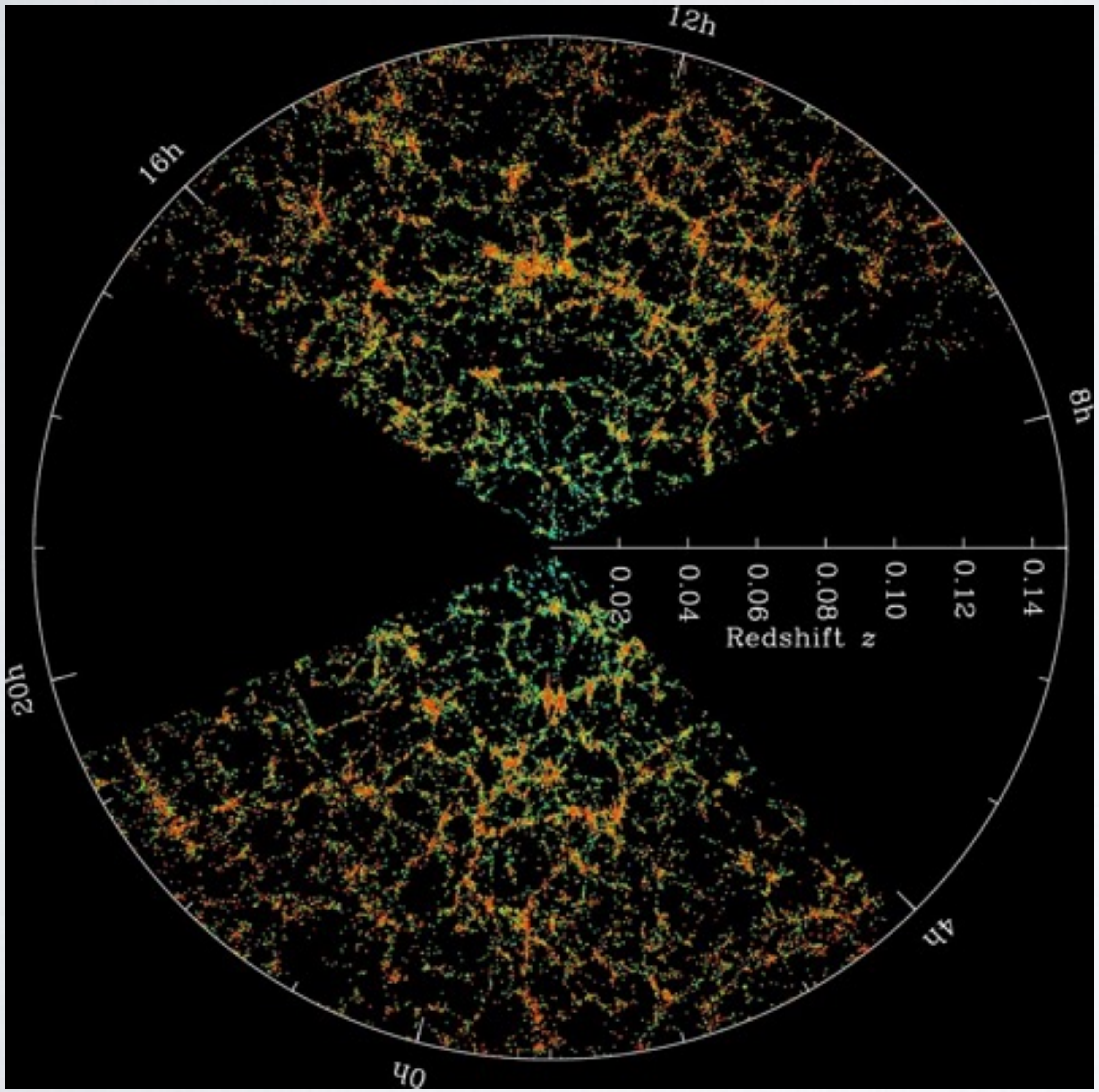
Galaxy Groups and Clusters



Hickson Group
Galaxy Group
<50 galaxies.

Virgo Cluster
Galaxy Cluster
Up to ~100 galaxies.
Sizes up to a few Mpc.

Large Scale Structure



Sloan Digital Sky Survey (SDSS)

Cosmic Web

This is how the Universe looks like on scales ~100 Mpc.

There is a hierarchy of structures of different sizes that constitute our Universe.

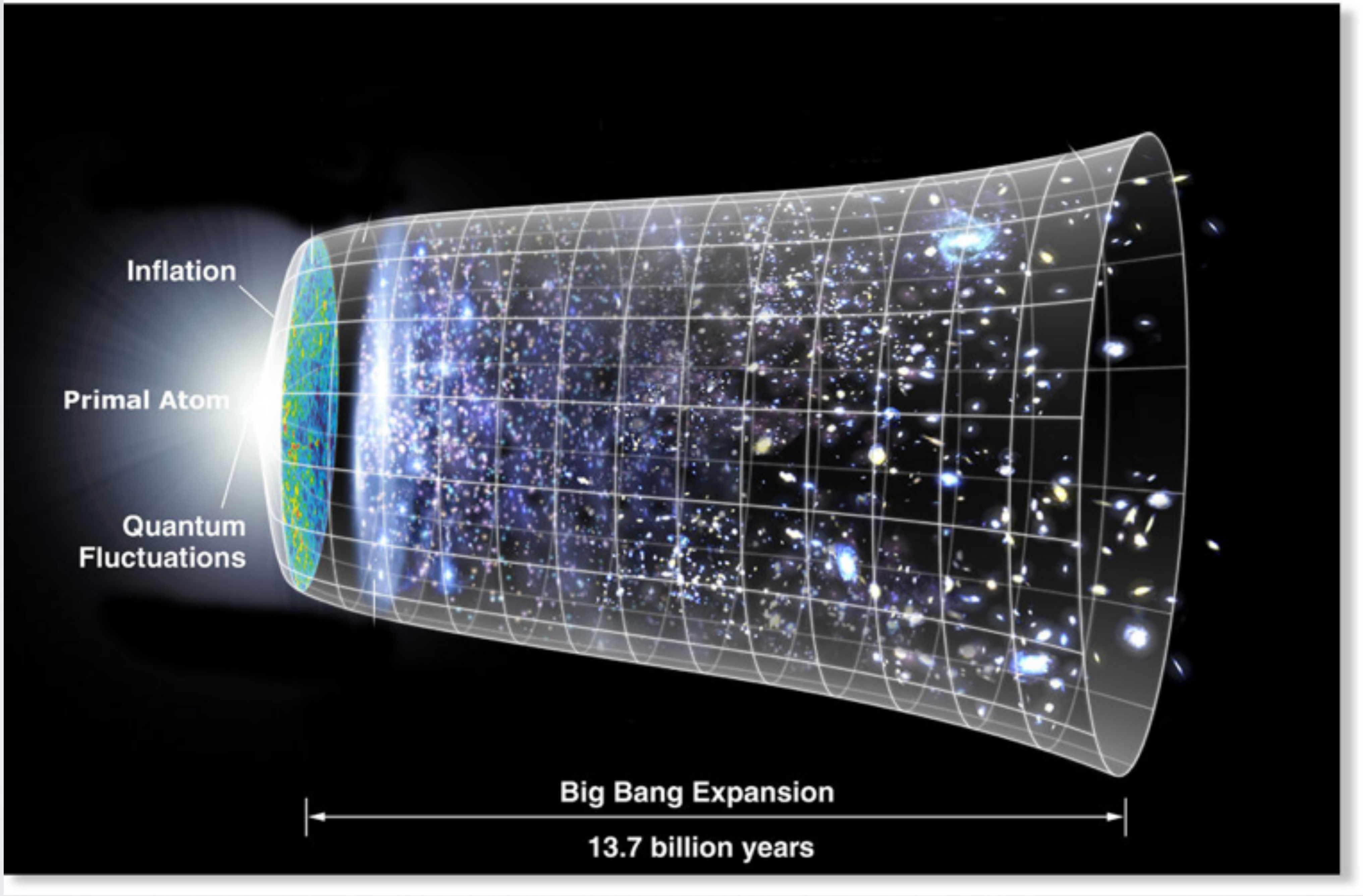
Cosmological Structure Formation

Understanding structure formation in the context of an evolving universe.

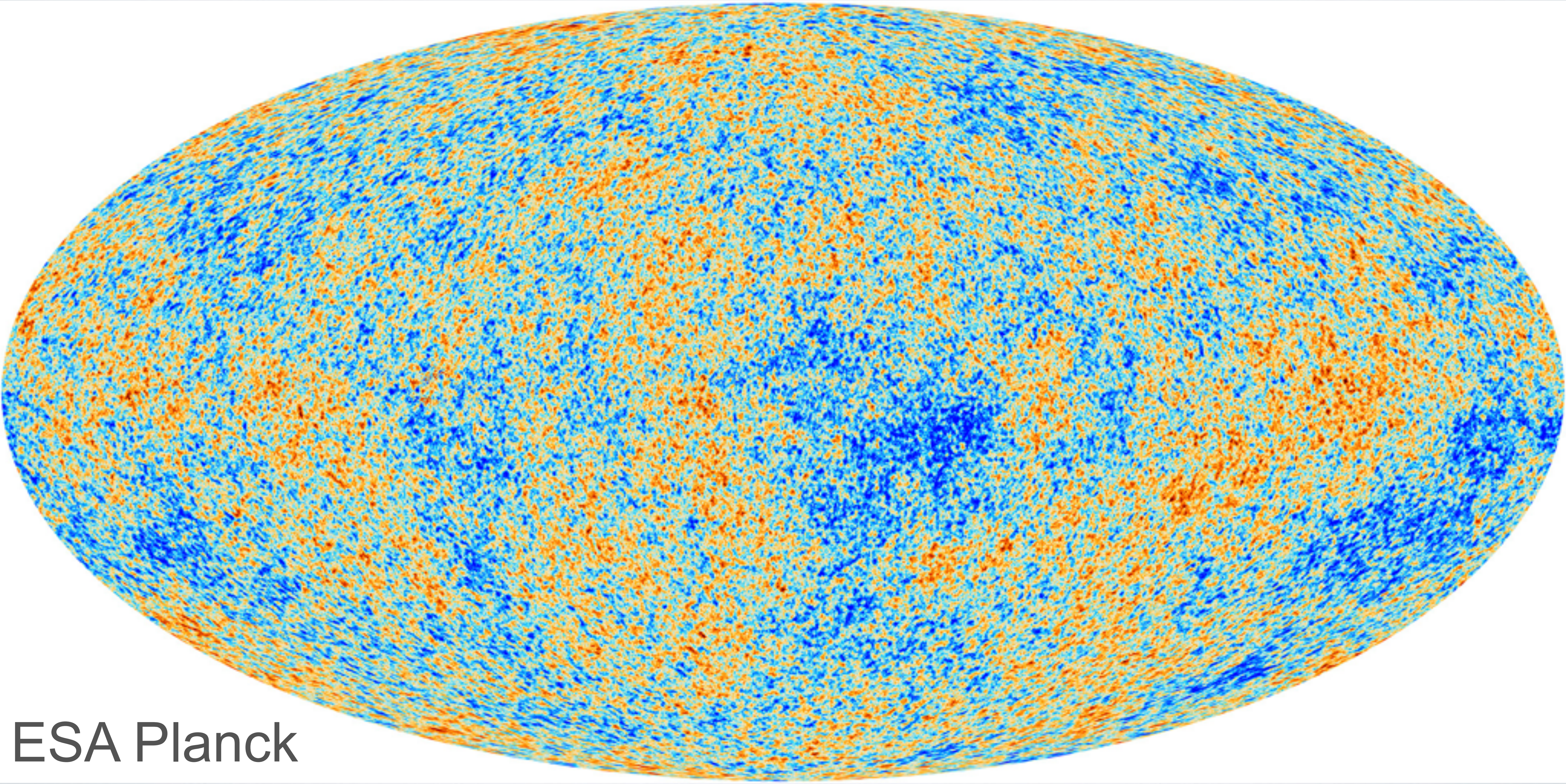
It's a work in progress that requires:

- Observations to characterize the "ingredients" of the Universe.
- A theory for structure formation.
- Calculations to make testable predictions from the theory.

Big Bang Cosmology

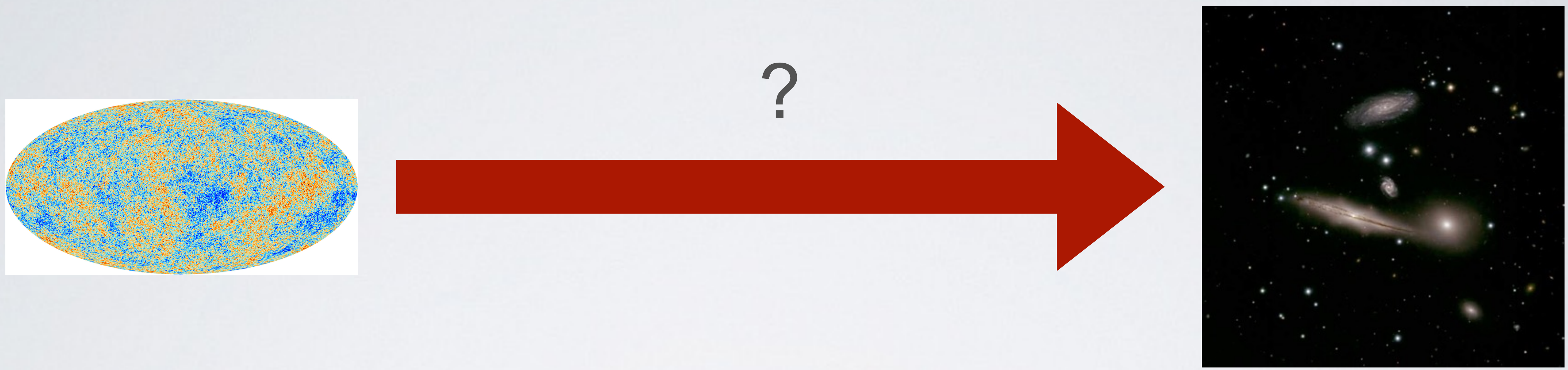


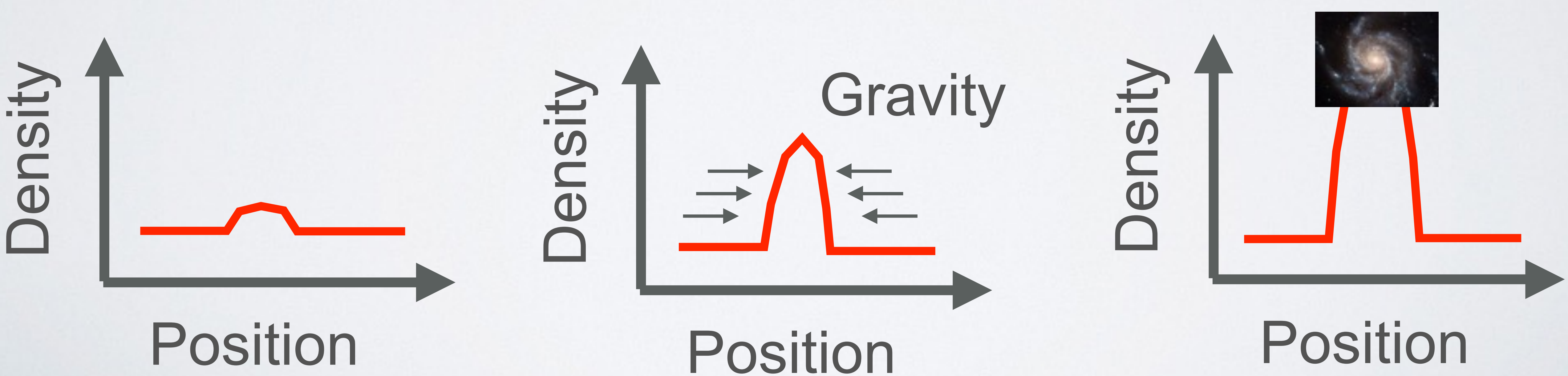
Cosmic Microwave Background Anisotropies



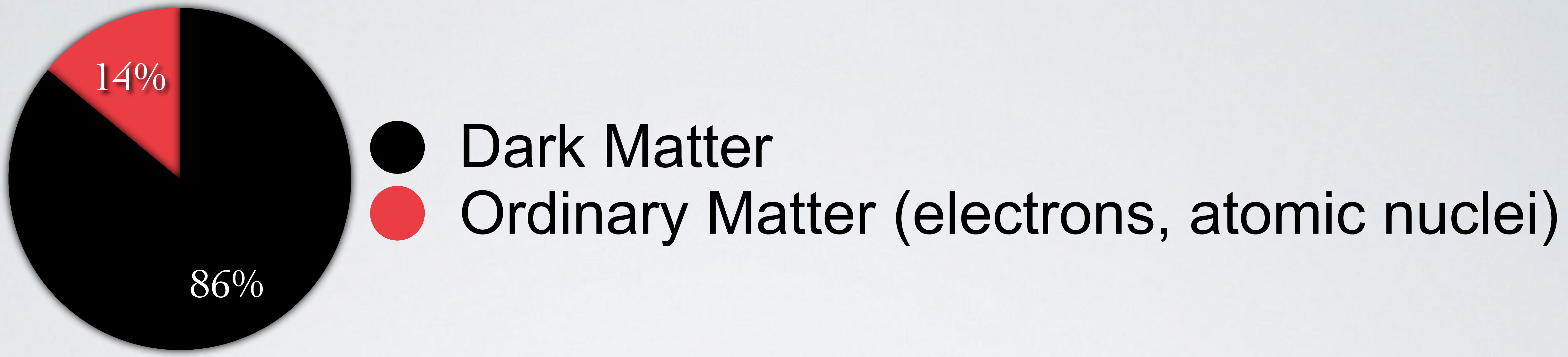
ESA Planck

Growth of primordial perturbations





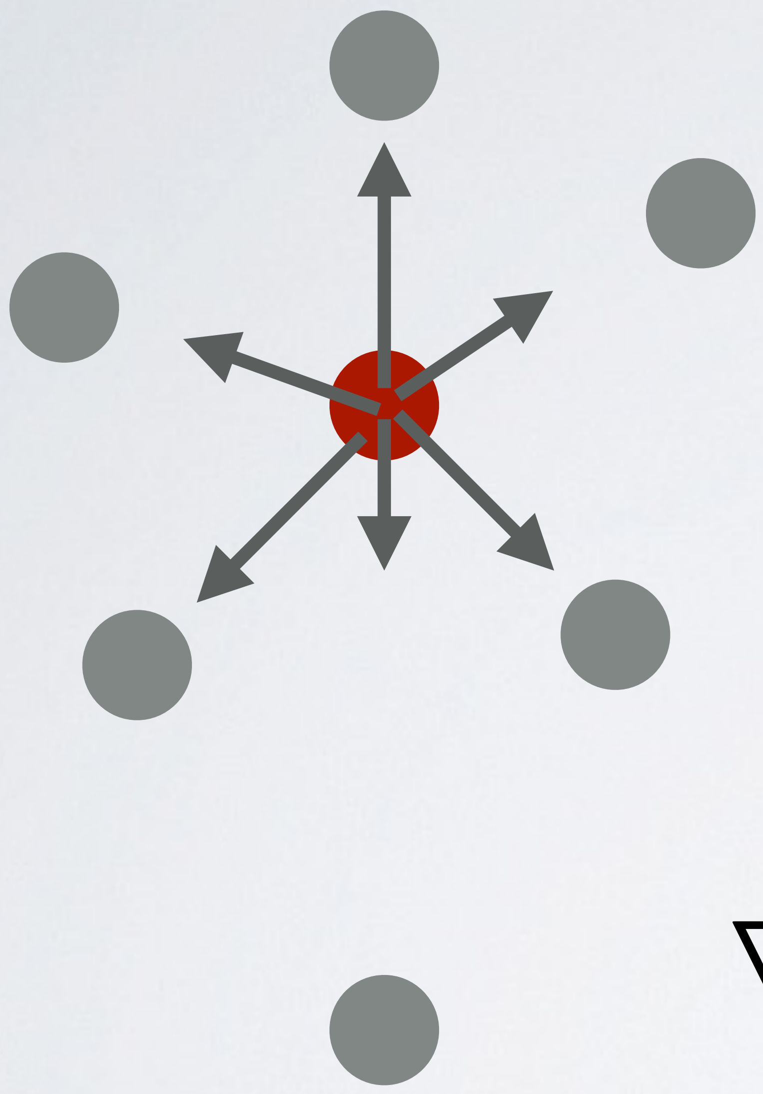
Simple model of the Universe



A series of hypotheses based on observational evidence:

- The Universe expands.
- Dark Matter only interacts gravitationally and dominates the growth of cosmic structure.
- Ordinary Matter can be neglected to zero-th order.

Dark Matter Dynamics



Gravity is a long-range force.

Dark Matter particles respond to the gravitational potential generated by the global distribution of Dark Matter.

$$\nabla^2 \Phi = 4\pi G \rho$$ Poisson's Equation

$$\mathbf{F}(\mathbf{x}) = -\nabla \Phi(\mathbf{x})$$ Gravitational force

Dark Matter Dynamics

Dark Matter particles only see the global gravitational potential.

f(**x**,**v**,t) = Probability for a dark matter particle of having velocity **v** while being at position **x**.

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \dot{\mathbf{x}}\frac{\partial f}{\partial \mathbf{x}} - \nabla\Phi\frac{\partial f}{\partial \mathbf{v}} = 0$$

Boltzmann Equation

$$\rho = \int f d^3\mathbf{v}$$

Dark Matter 3D - Density

$$\nabla^2\Phi = 4\pi G\rho$$

Poisson’s Equation for the gravitational potential

Problems, problems, problems…

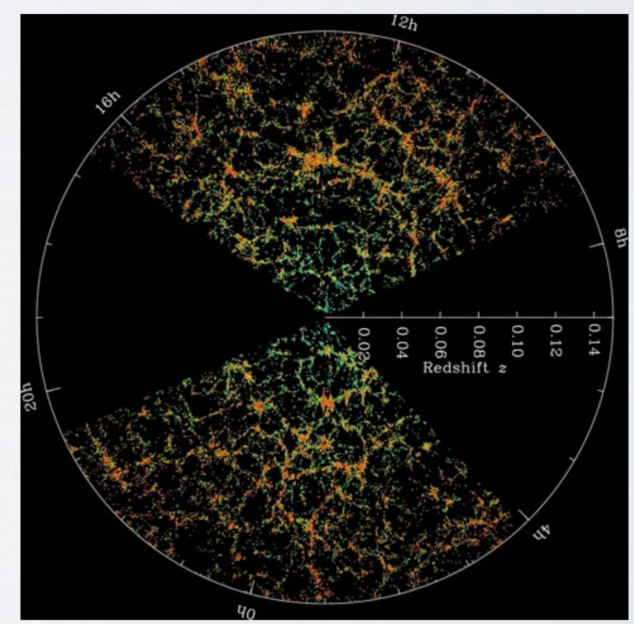
$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \dot{\mathbf{x}}\frac{\partial f}{\partial \mathbf{x}} - \nabla\Phi\frac{\partial f}{\partial \mathbf{v}} = 0$$ Boltzmann Equation

1. BE is a non-linear integro-differential equation in 6d. General non-trivial solutions are hard to find!

2. Solutions for cosmic structure formation have to be accurate over a huge dynamic range.

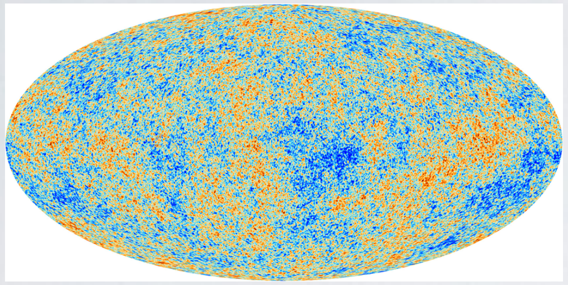
From galaxies…

To clusters…



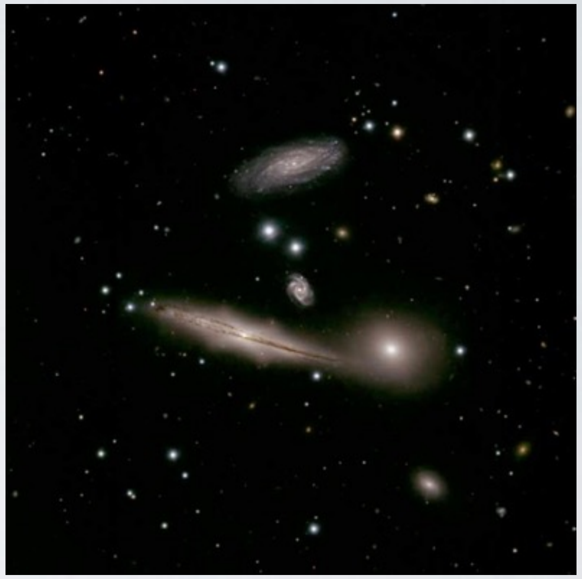
To the LSS.

Cosmological Simulations



Initial conditions motivated by Big Bang theory.

Theory
+
Simulation Algorithm
+
Supercomputer



A snapshot of a simulated universe at a given time.

Can be compared to observations to test theory!

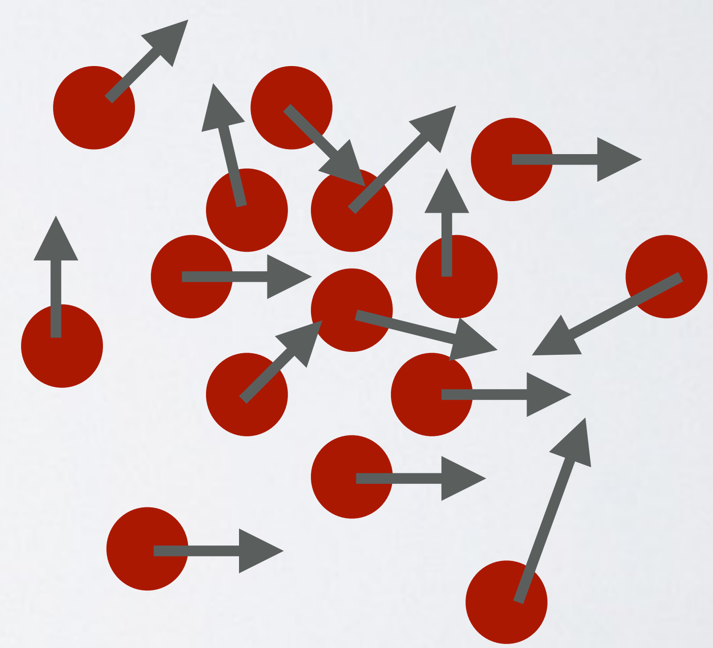
N-body approach

Discretization of the 6-D distribution of DM.

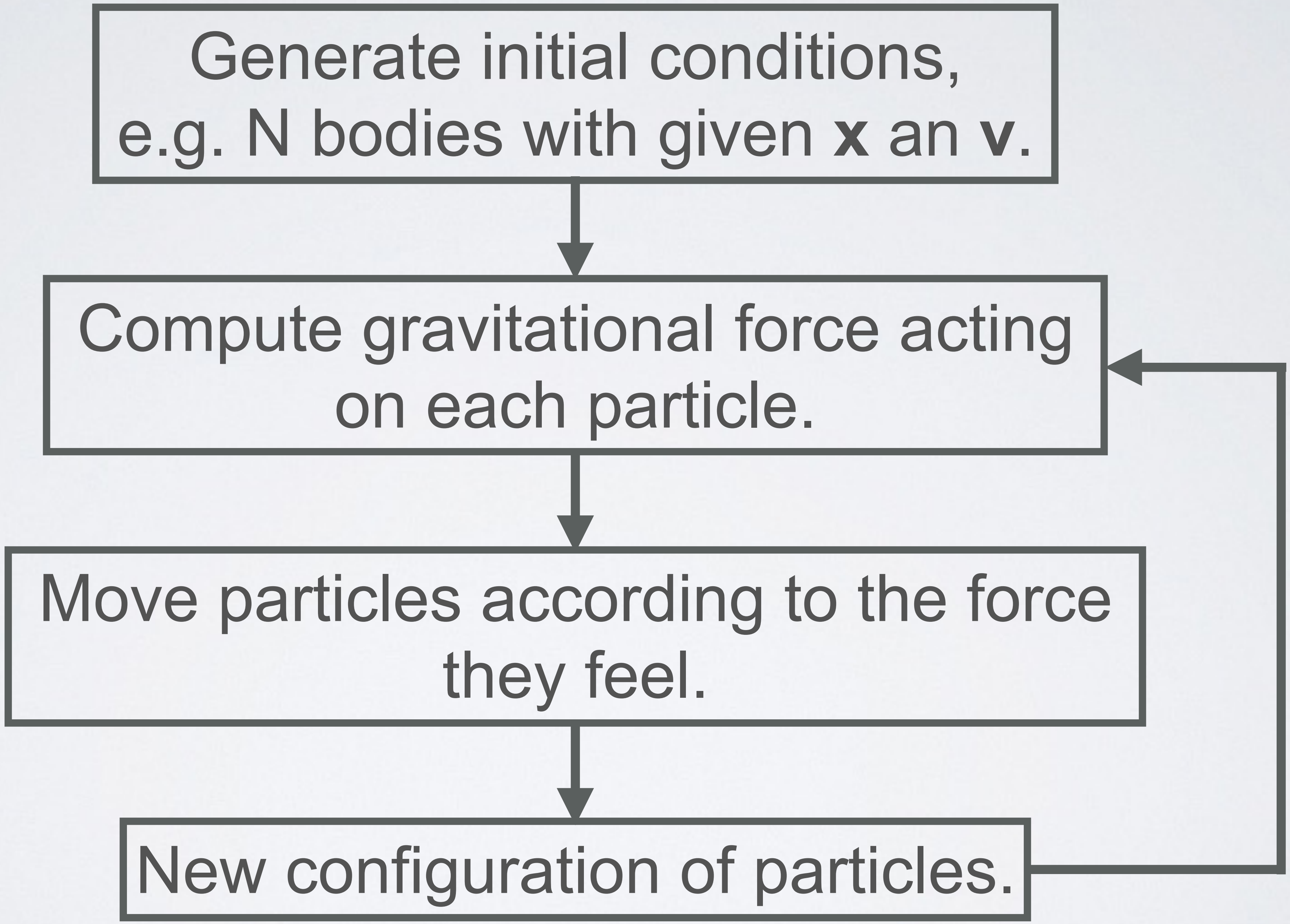
Continuous, smooth $f(\mathbf{x},\mathbf{v},t)$

Monte Carlo sampling

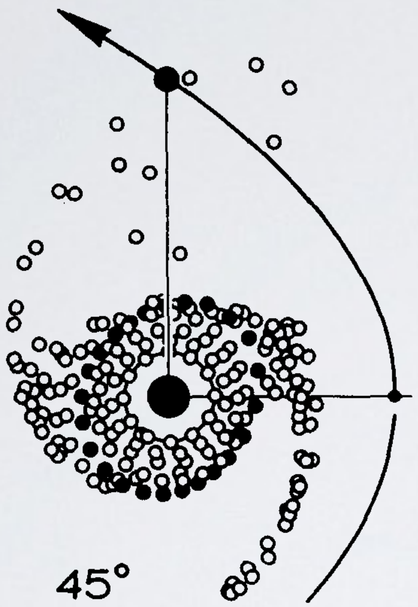
N particles of mass ***m***



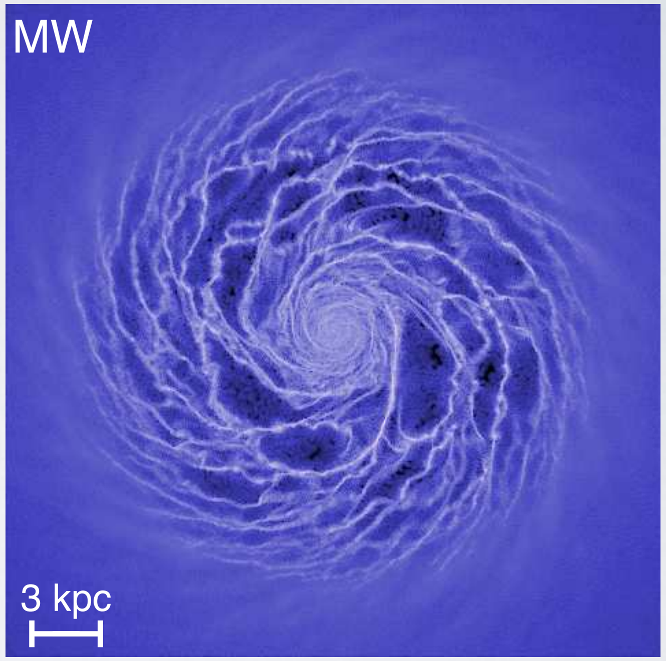
Typical algorithm for DM dynamics



Number of bodies and resolution



N~100 particles
(Toomre & Toomre 1972)



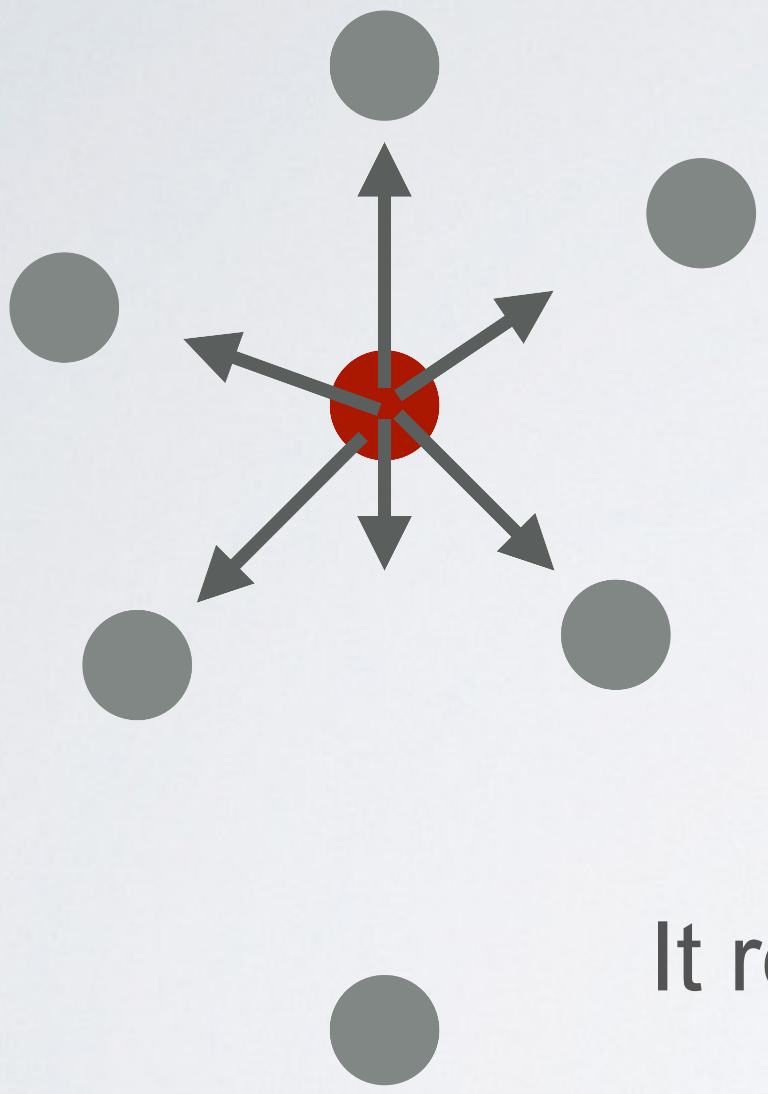
N~10^8 particles
(Hopkins et al. 2011)

Better resolution -> More precise predictions.

Large N -> High computational cost.

Fast algorithms required!

Direct summation



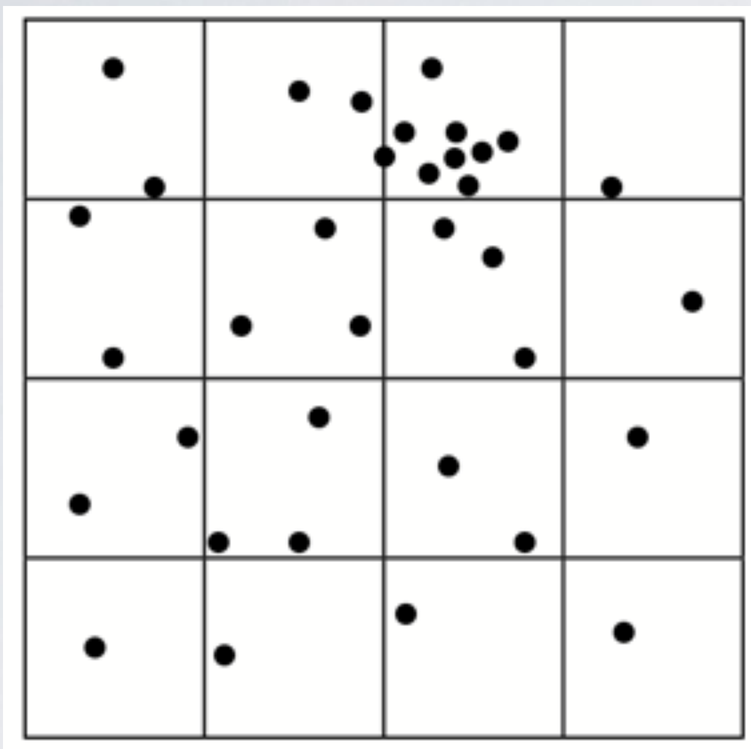
What's the gravitational force acting on a particle?

$$\mathbf{F}_i = \sum_j^N \frac{G m_i m_j (\mathbf{x}_j - \mathbf{x}_i)}{|\mathbf{x}_i - \mathbf{x}_j|^3}$$

It requires $O(N^2)$ operations per time step.

Bottleneck for large ***N***!

Particle Mesh Methods



$$\nabla^2\Phi = 4\pi G\rho$$

Scaling
O($\boldsymbol{N}_{\mathbf{mesh}}$ ln($\boldsymbol{N}_{\mathbf{mesh}}$))

Build a Cartesian mesh.

↓

Interpolate particles to obtain density $\boldsymbol{\rho}(\mathbf{x})$ at the mesh nodes.

↓

Compute FFT of density $\rho(\mathbf{k})$

↓

Solve Poisson equation, get $\Phi(\mathbf{k})$ FT of gravitational potential.

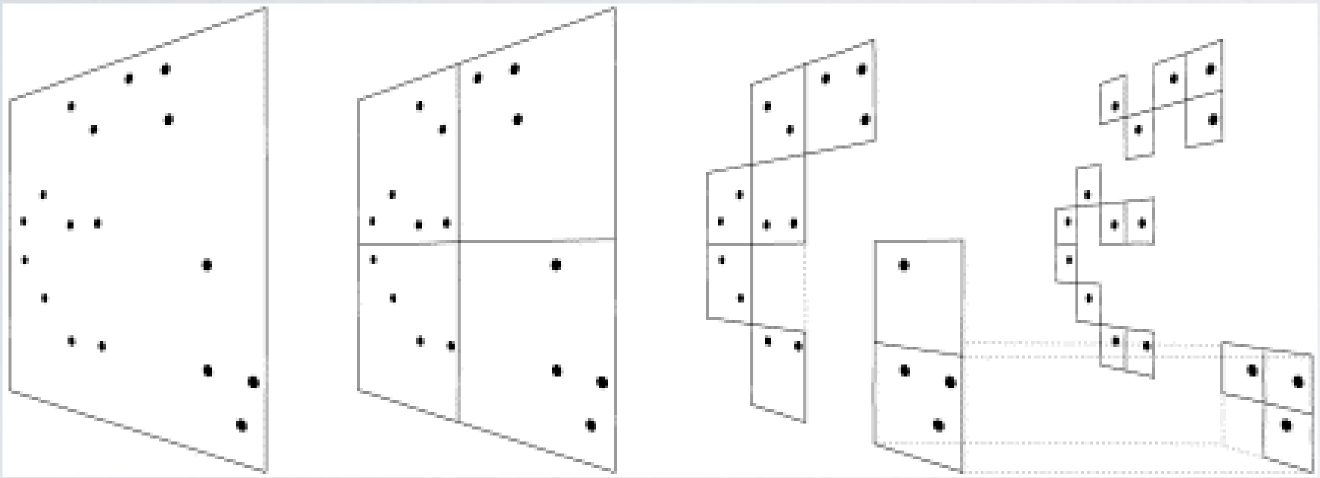
↓

Inverse FFT to get grav. pot. in at mesh nodes $\Phi(\mathbf{x})$.

↓

Compute forces at nodes and interpolate back to particles.

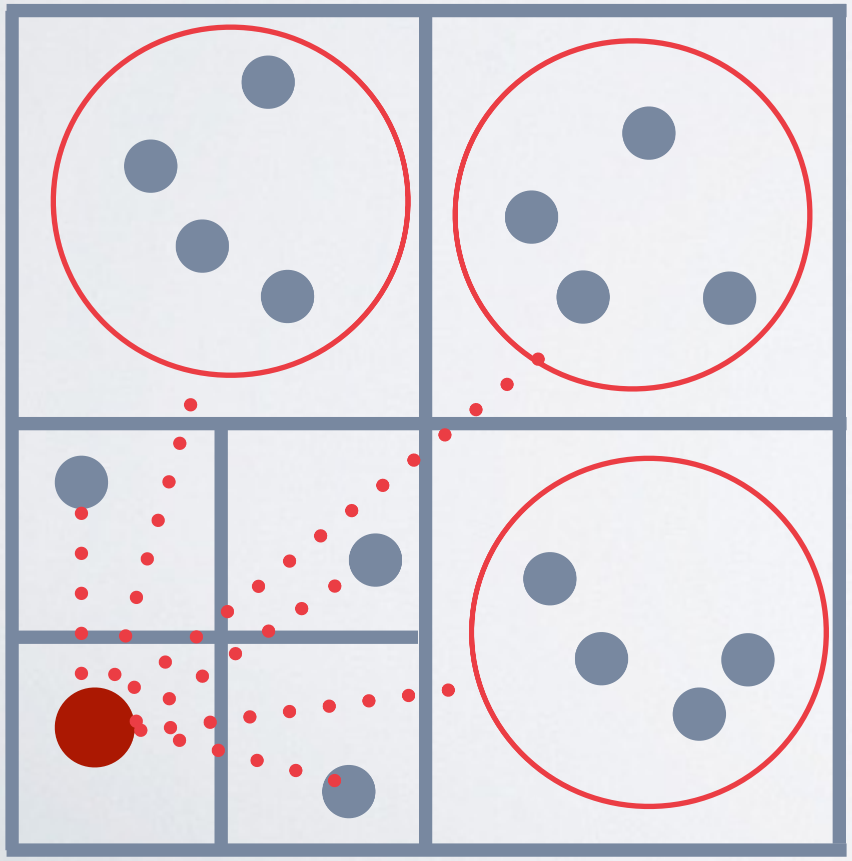
Tree Methods



Oct tree, Barnes & Hut, 1986

Step 1: Organize particles in a tree structure.

Step 2: walk the tree and sum force contribution from tree branches.



Standard tree code scaling:
~O($\boldsymbol{N}$ ln($\boldsymbol{N}$))

Modern techniques as fast as O($\boldsymbol{N}$).

Modern simulations

- Include both Dark and Ordinary Matter. Require combination of N-body and Computational Fluid Dynamics methods. Better physical modeling.
- 10^6-10^{12} bodies. Better scaling of the algorithms.
- 10^5-10^7 CPU hours on computing clusters. Code parallelization.
- 10-100 TB of data to analyze. Big Data.
- Performed with optimized “community” codes. Open source vs. “limited access” codes.

Where do we run the simulations?

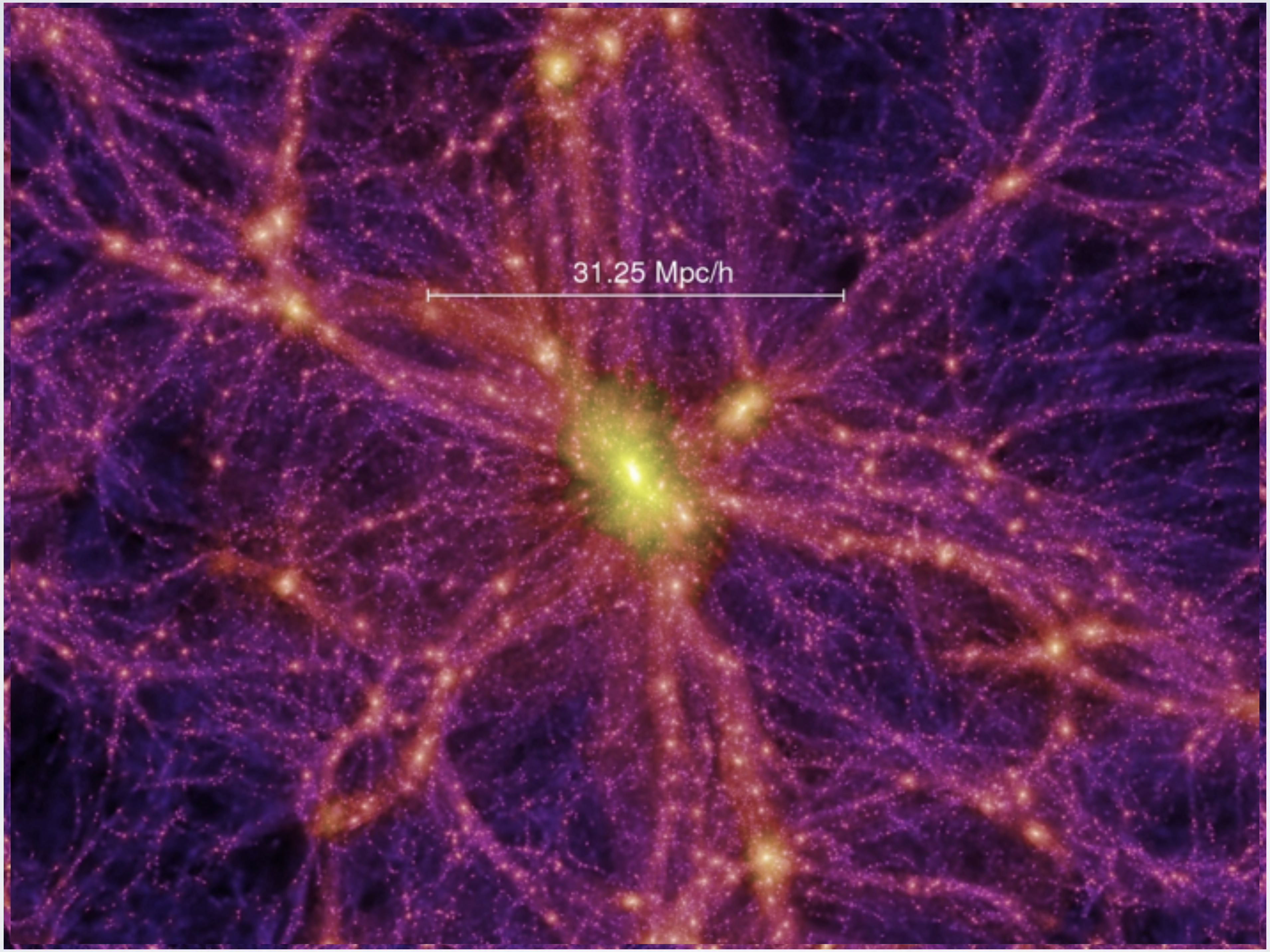
Pleiades supercomputer
at NASA’s Ames Research Center

Example:

- ~200000 CPUs.
- GPU nodes.
- ~800 TB of memory.
- 6.28 Pflop/s.

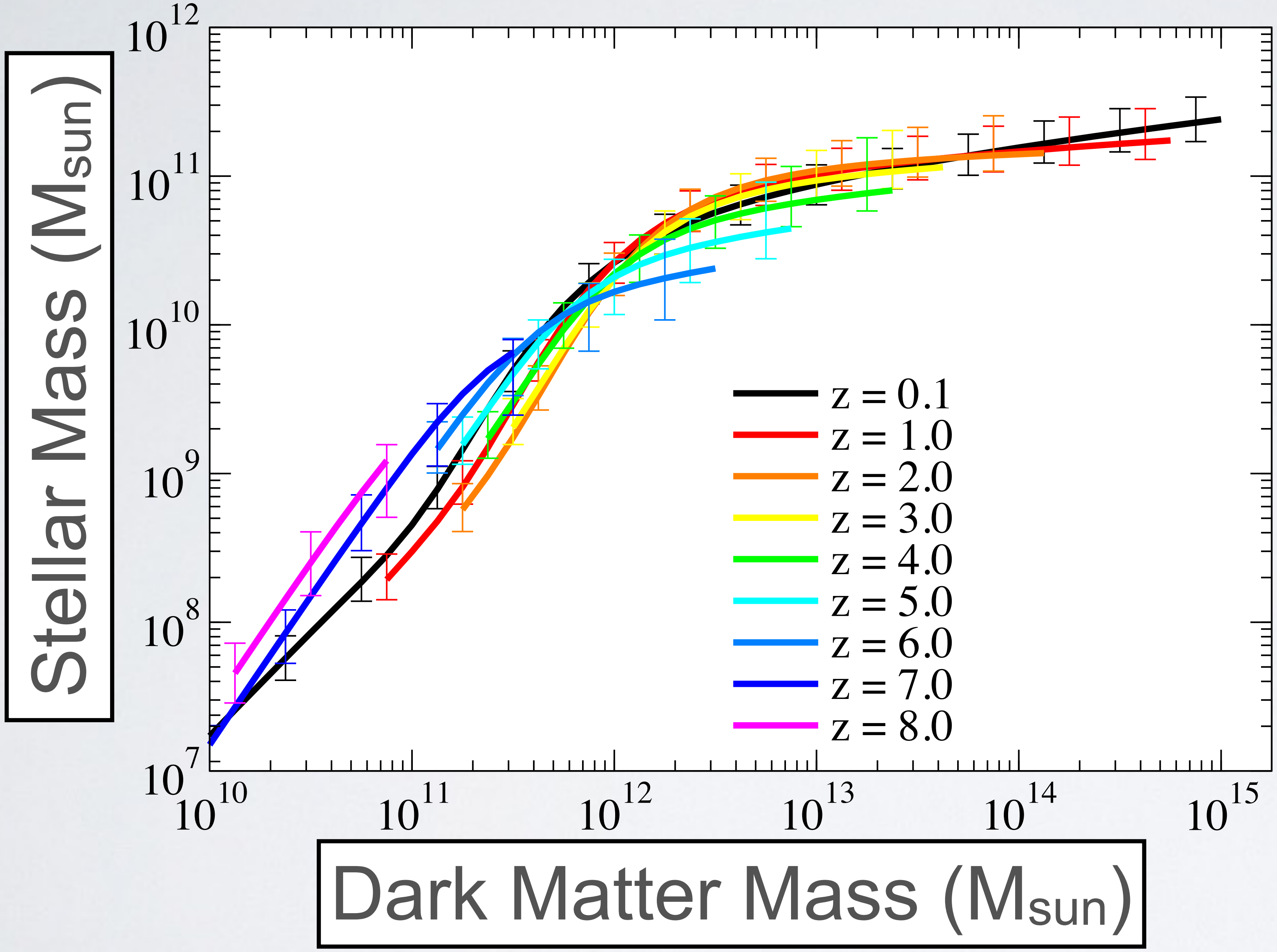
See also XSEDE (NSF).

Millenium Run



Springel et al. (2005).

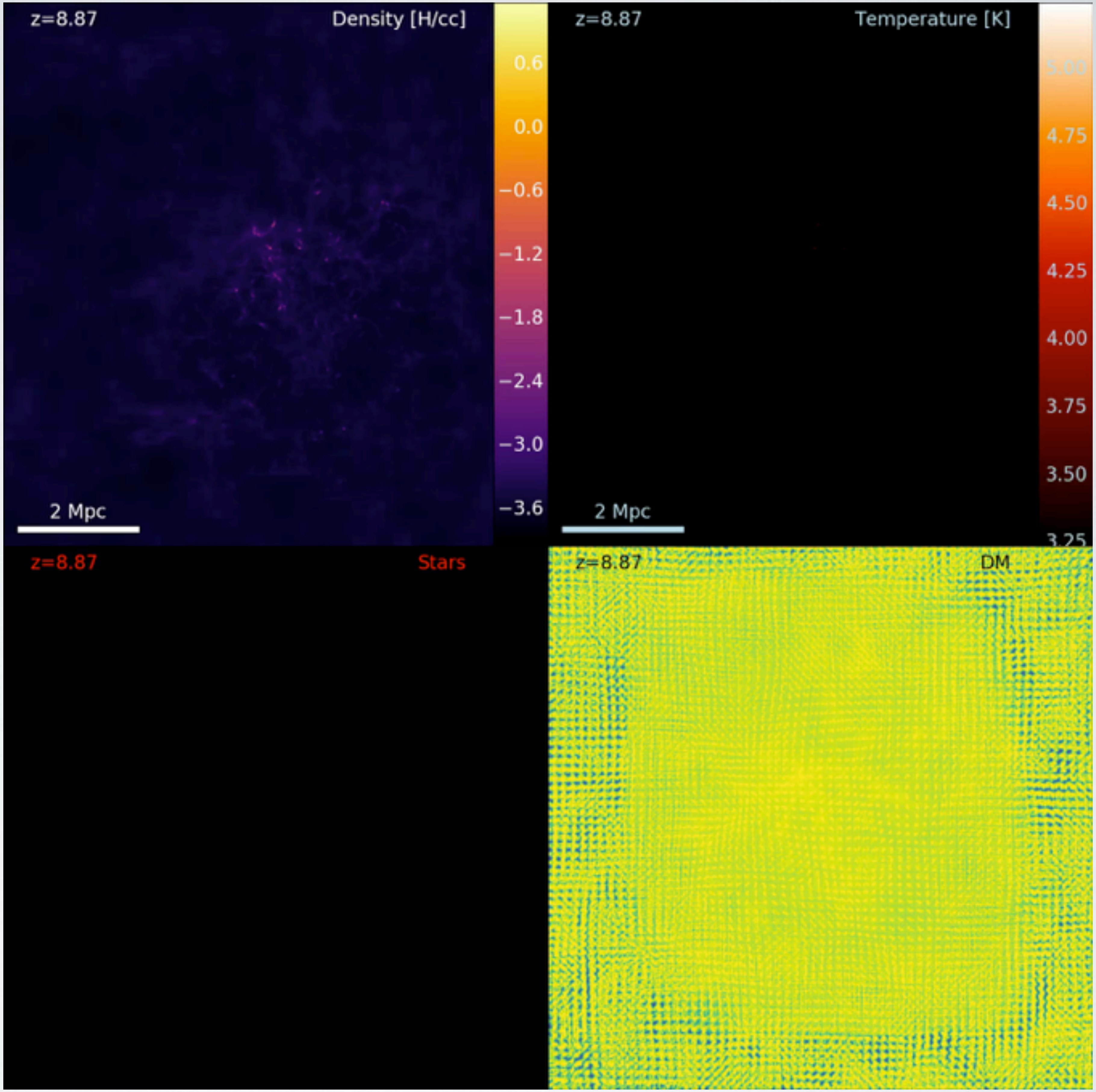
What can we learn?



Behroozi+ (2013)

Ordinary Matter does not exactly trace Dark Matter.

RH
AP
SO
DY
G

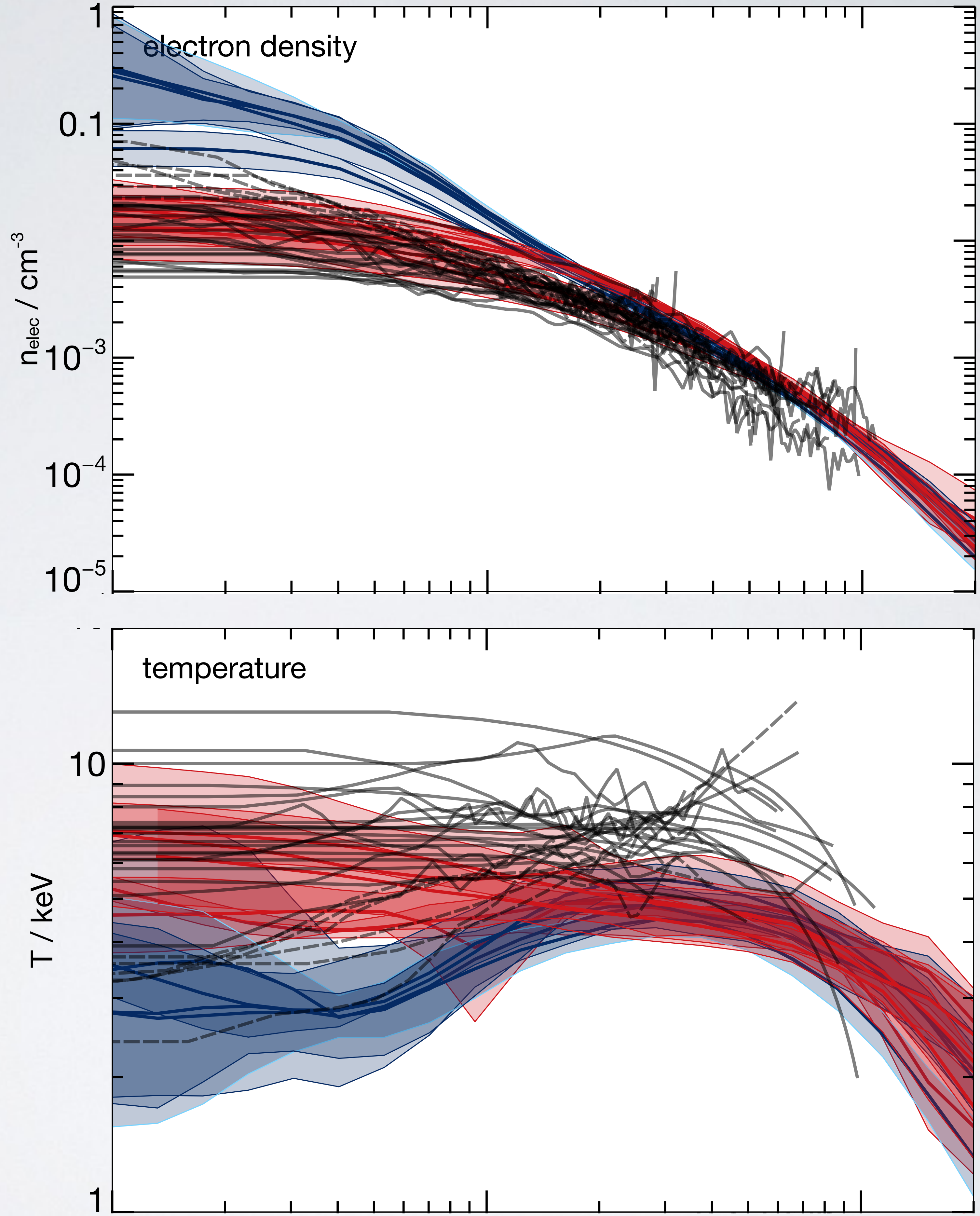


Martizzi+ 2016

Videos by: Pawel Biernacki

RH
AP
SO
DY
G

Hahn+ 2016

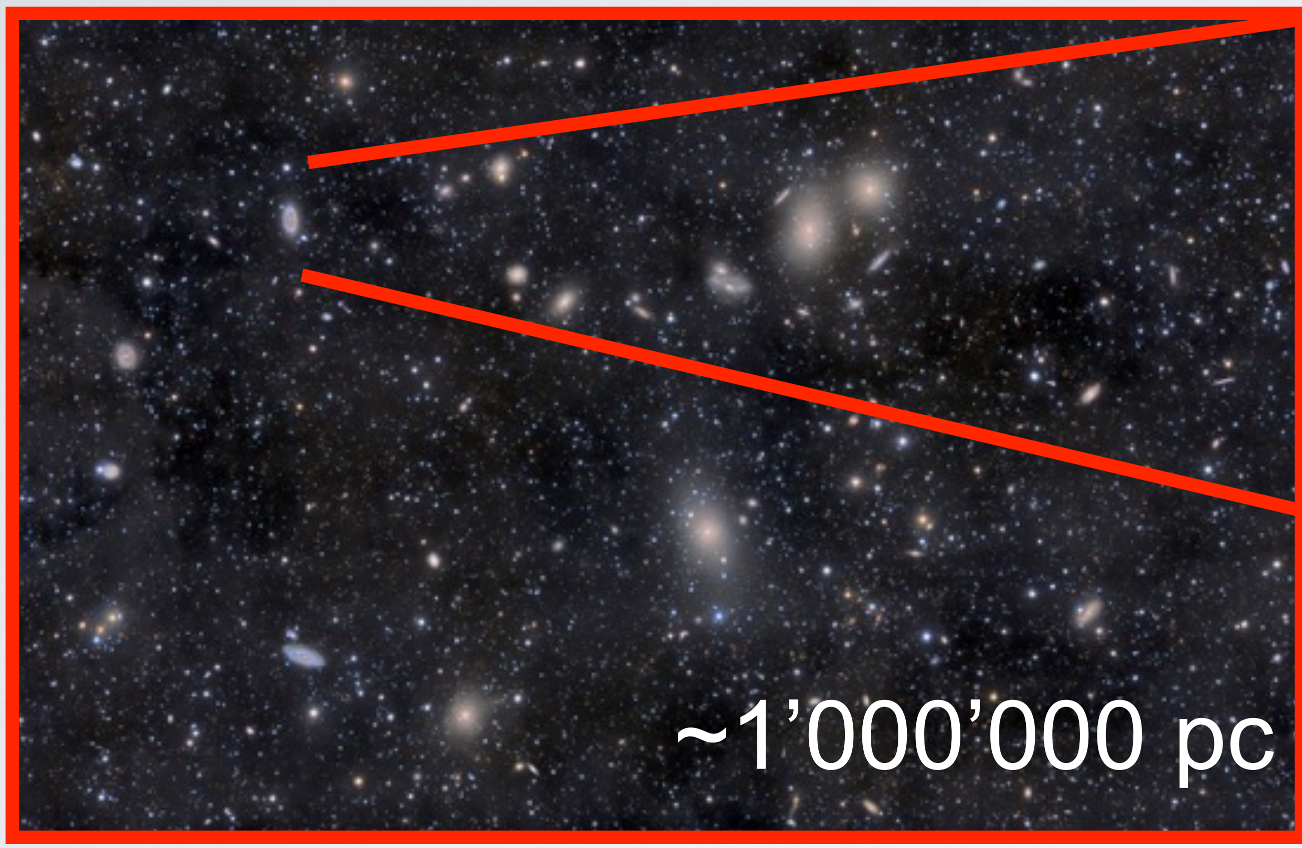


What’s next?

Open questions

- Do we fully understand galaxy formation in a cosmological context?
- Do we reproduce realistic populations of galaxies?
- Are simulated galaxies realistic at all times?
- How do primordial galaxies look like?
- How do internal processes modify galaxies and their surroundings?
- A lot more…

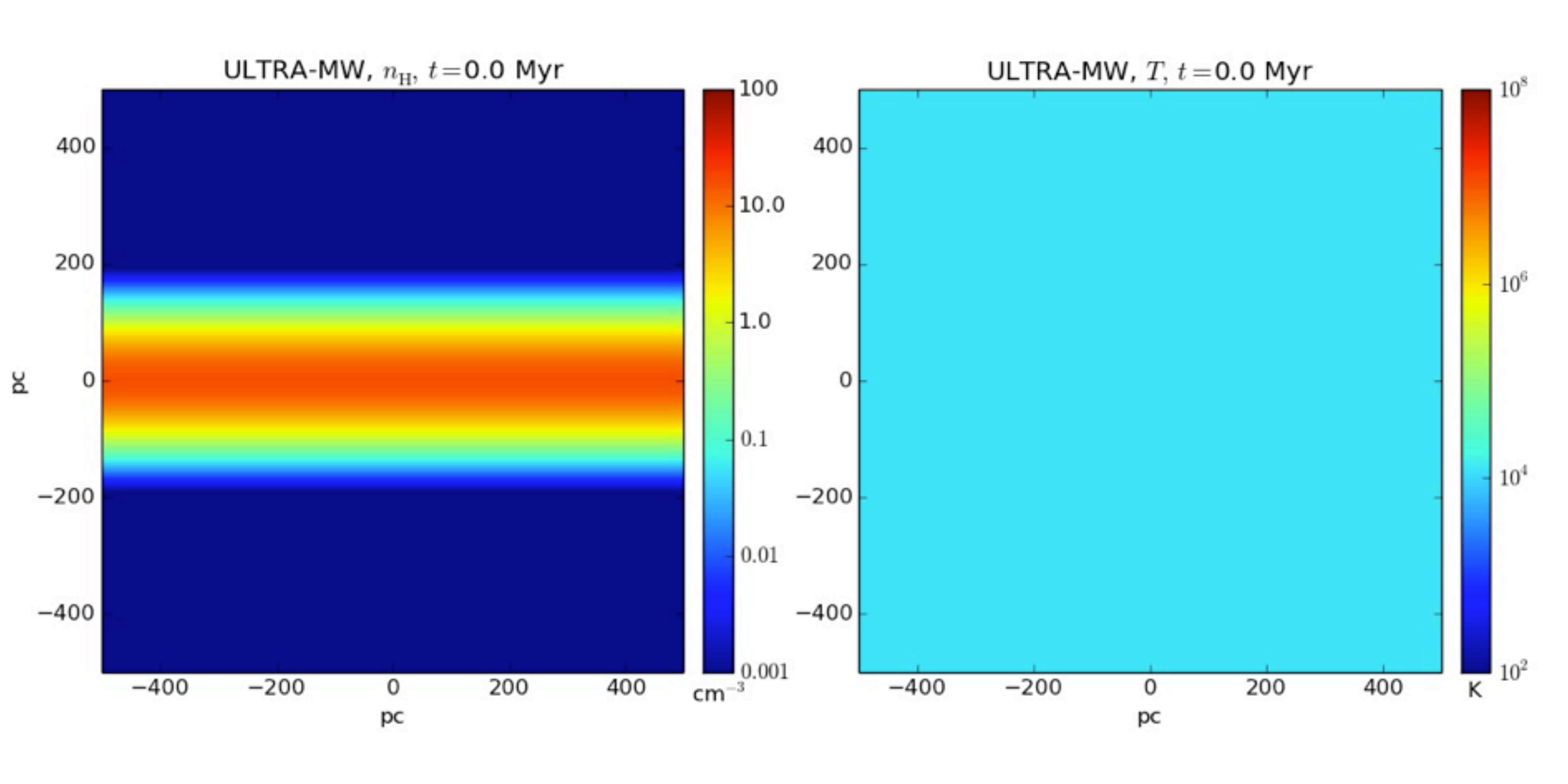
Galaxy Formation Studies



“Small scale” processes influence the big picture.

We can’t simply throw more resolution at the problem.

Example: stellar explosions in a galaxy



Martizzi+ 2016

Higher Resolution

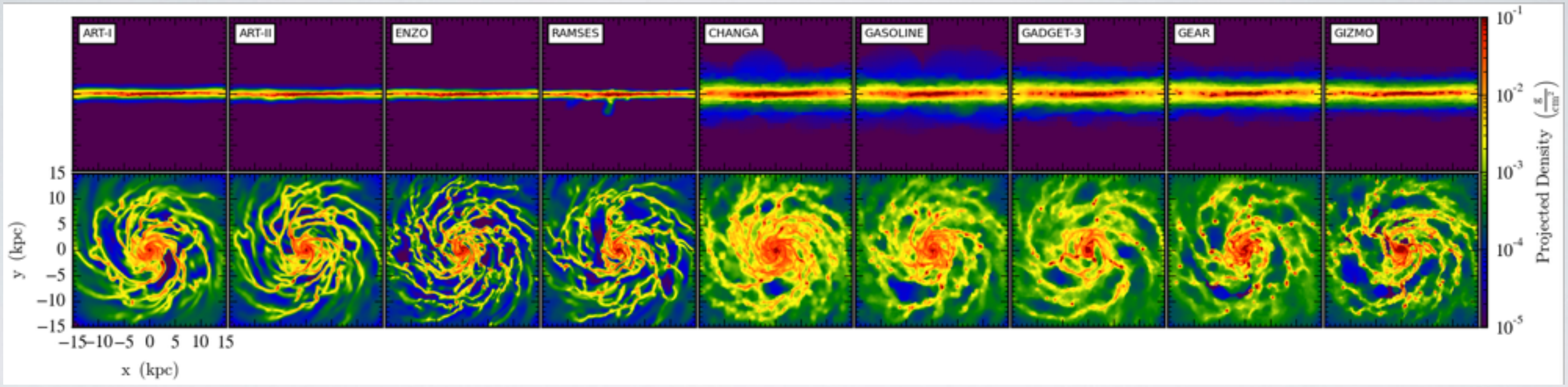
Physical modeling:

- Interstellar/intergalactic gas.
- Star formation.
- Stellar explosions.
- Radiation transfer.
- High energy processes.
- Black hole physics.

Scientific Computing:

- More accurate N-body +CFD methods.
- New architectures (CPUs, GPUs, Coprocessors),
- Code optimization.
- New data management strategies.

More accurate methods



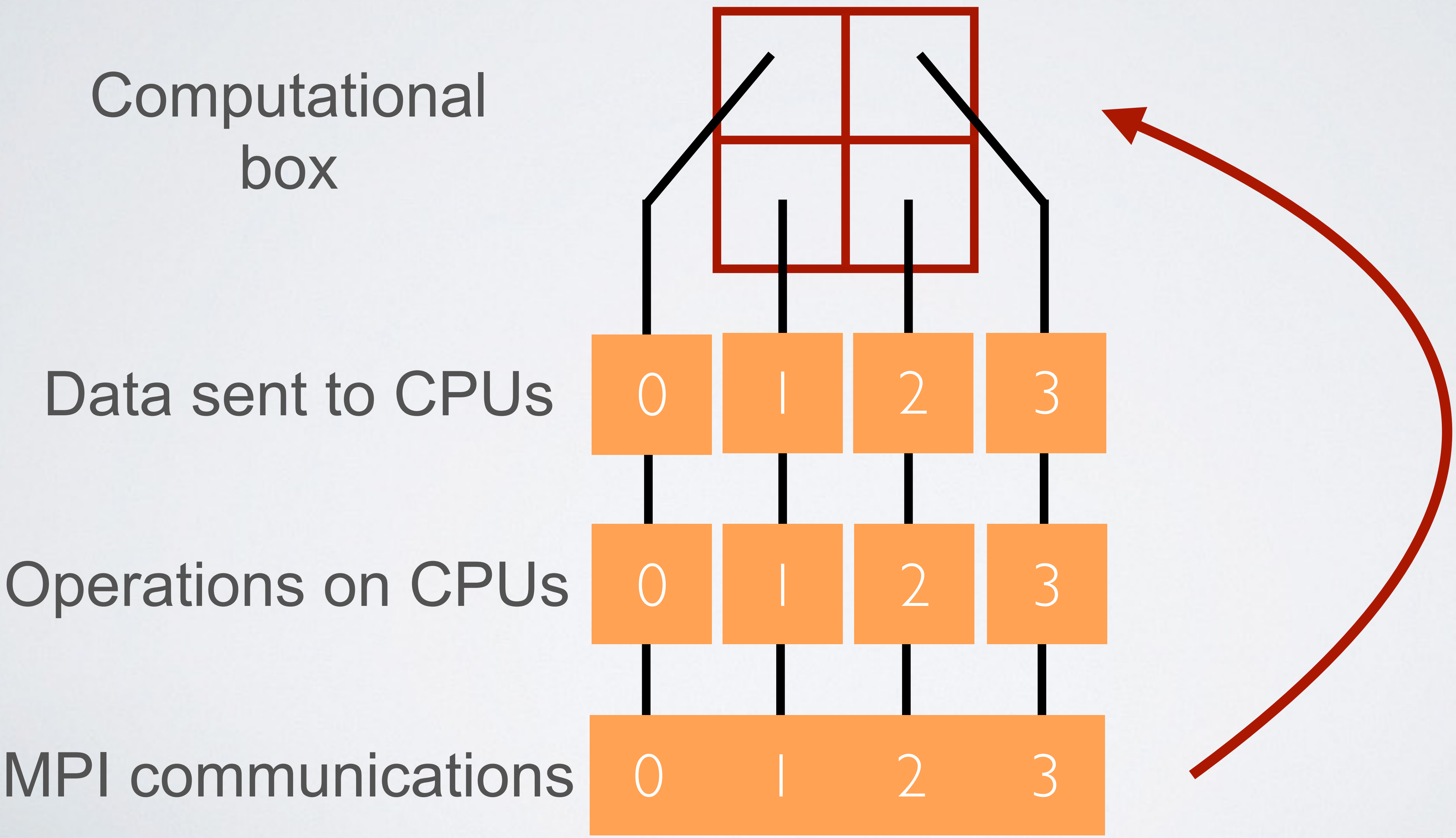
AGORA comparison project (Kim+ 2016)

Different methods/codes can give different qualitative/quantitative results.

We need to assess the limits of current methods and design better ones!

A note on parallelization

Most available codes use
domain decomposition + MPI communication.



Better parallelization

MPI-based parallelization can have problems with:
- Load balancing.
- Synchronization.
- Speed of the CPUs.

Proposed solutions, work in progress:
- Better load balancing/domain decomposition schemes.
- Hybrid MPI+OpenMP parallelization. With coprocessors?
- GPGPU computing on GPU clusters.
- ???

A lot of work!
Some algorithms (e.g. tree methods) are not easily portable to new architectures!

Sharing and data management

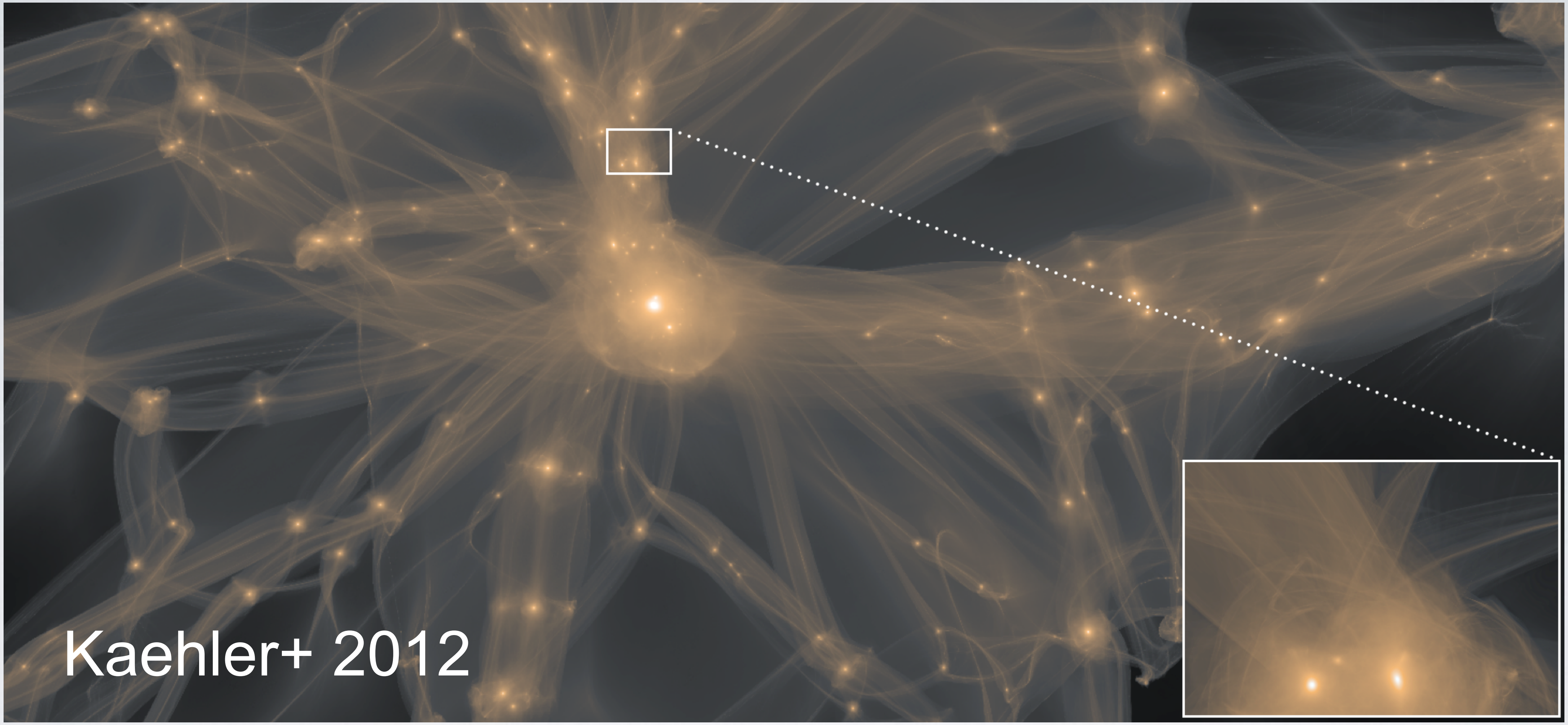
Computation can be a bottleneck but data management challenges are extremely important:

- Huge simulations -> ~100 TB of data.
- Data cannot be easily transferred. Public databases?
- Is it cheaper to re-run a simulation or to keep the data?

This is science -> Results from simulations have to be reproducible by anyone. Possible solutions:

- Public simulation codes + initial conditions.
- Private codes + public simulation databases.

Visualization



Useful for:

- Intuitive communication of scientific results.
- Intuition-driven discoveries.
- Public outreach and education.

Summary

- We have a well establish framework for studying cosmic structure formation with numerical simulations.
- Many solutions in scientific computing have been used to develop this field.
- Observation + theory + simulation = deeper understanding of cosmic structure formation.
- Many open problems: need for better physical modeling + new technologies.