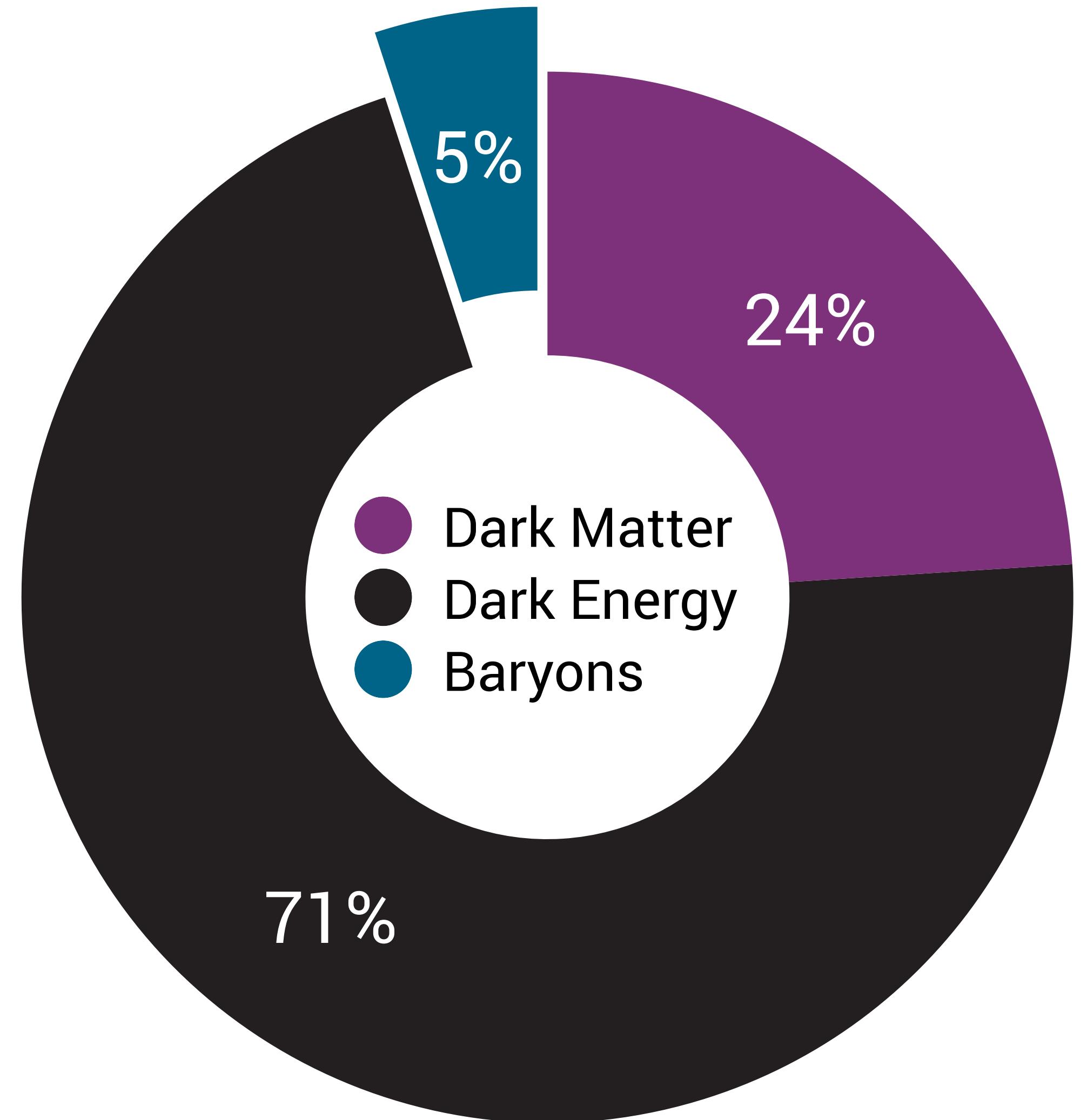


# Simulating the Universe

Josh Borrow

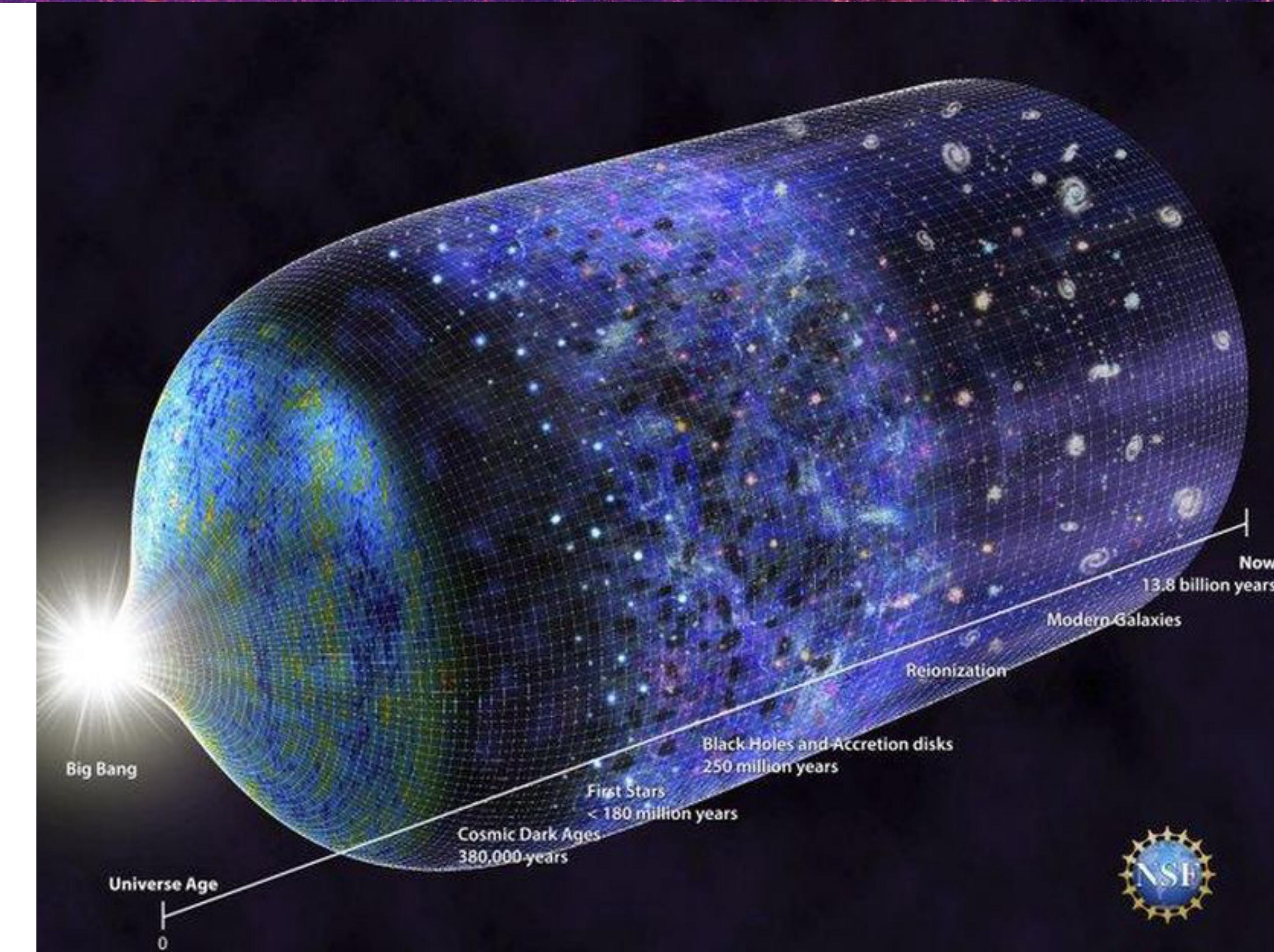
# What is the Universe Made of?

- Can't see the majority of the Universe
- Complete nightmare



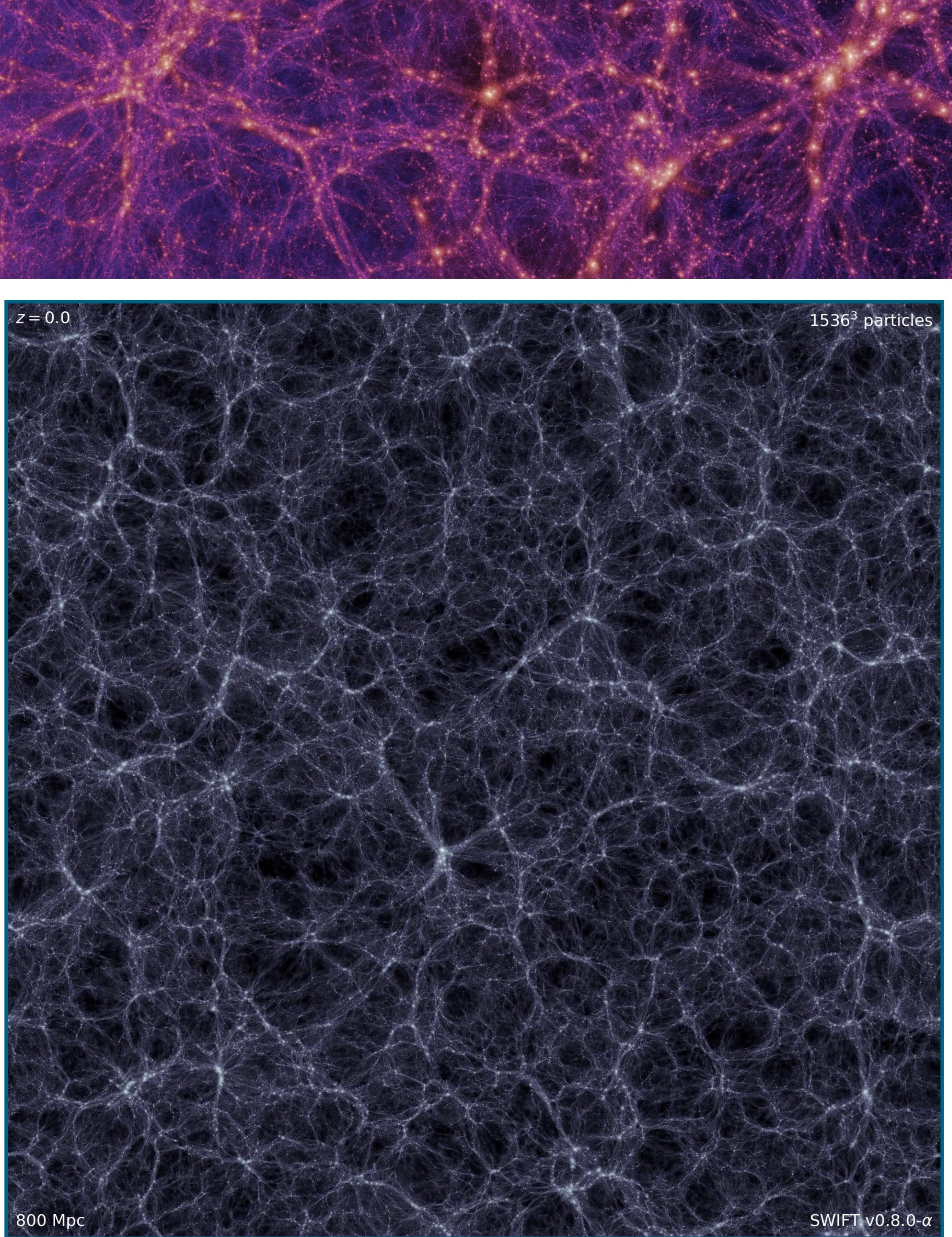
# Dark Energy

- Honestly, nobody has a clue about this one
- Affects growth of large-scale structure
- Particle physics estimates differ by a factor of  $10^{120}$
- Thankfully this doesn't actually change galaxy formation much



# Dark Matter

- Dominates gravitational attraction in the Universe
- Probably a WIMP
- Can be modelled as a cold, collision-less fluid
- Important: cannot 'cool'



# Baryons!

- Everything else
- Gas, stars, black holes, planets, people...
- We can see these, so they are quite important.



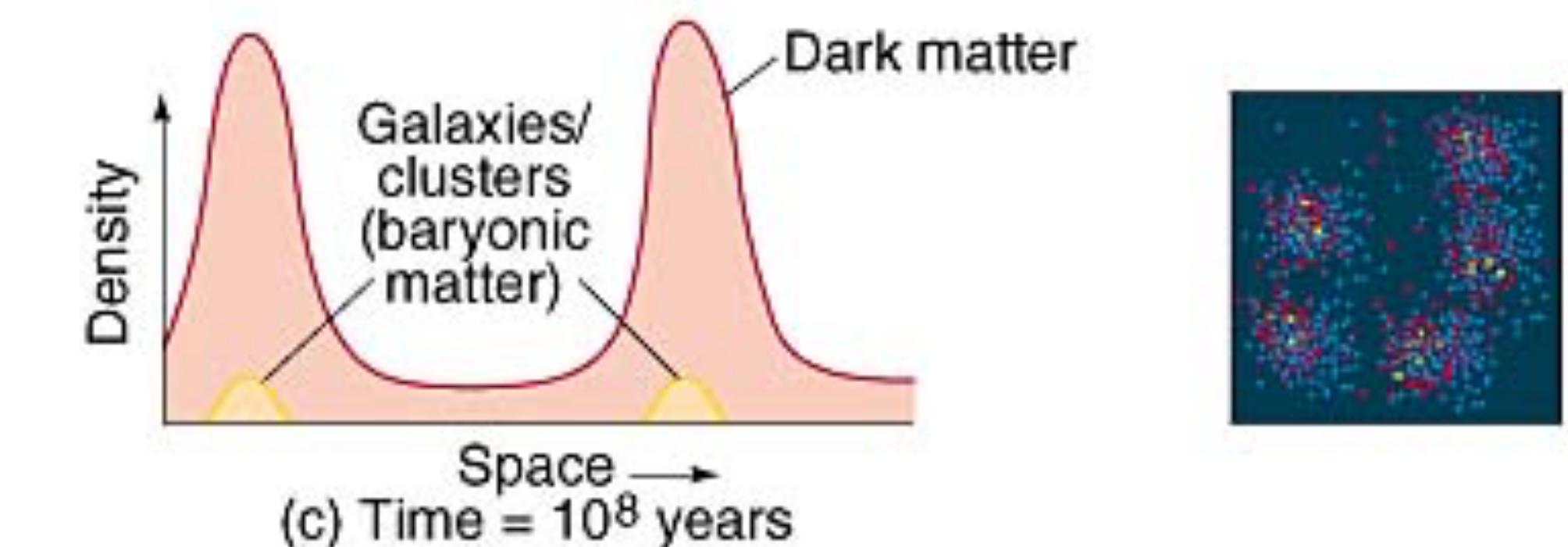
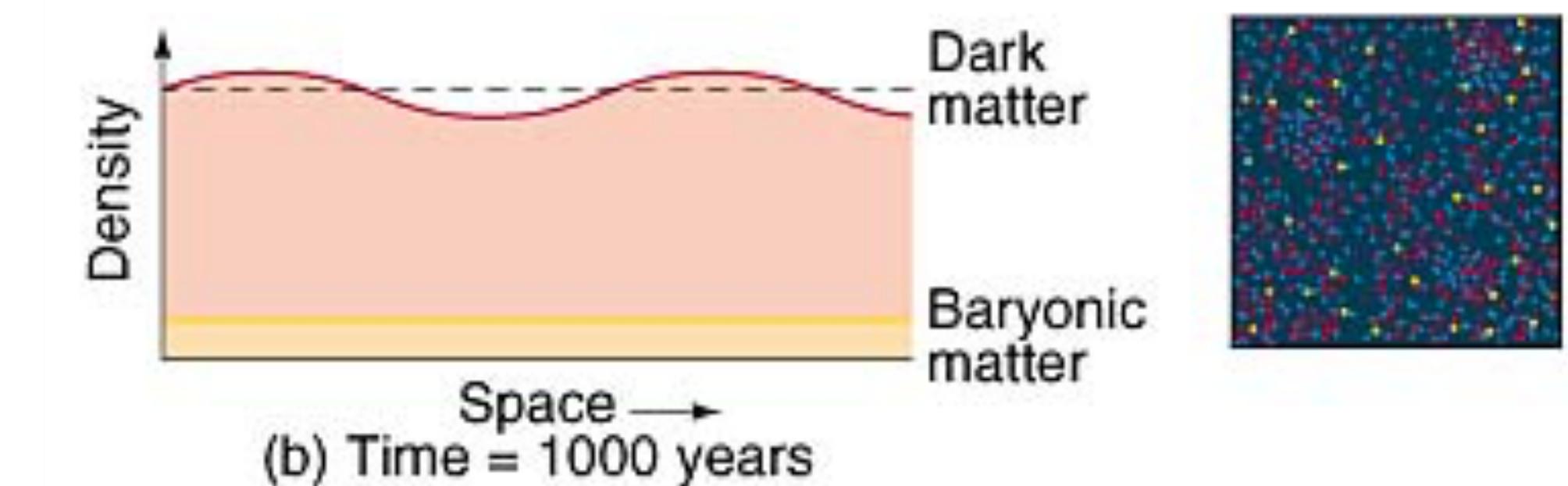
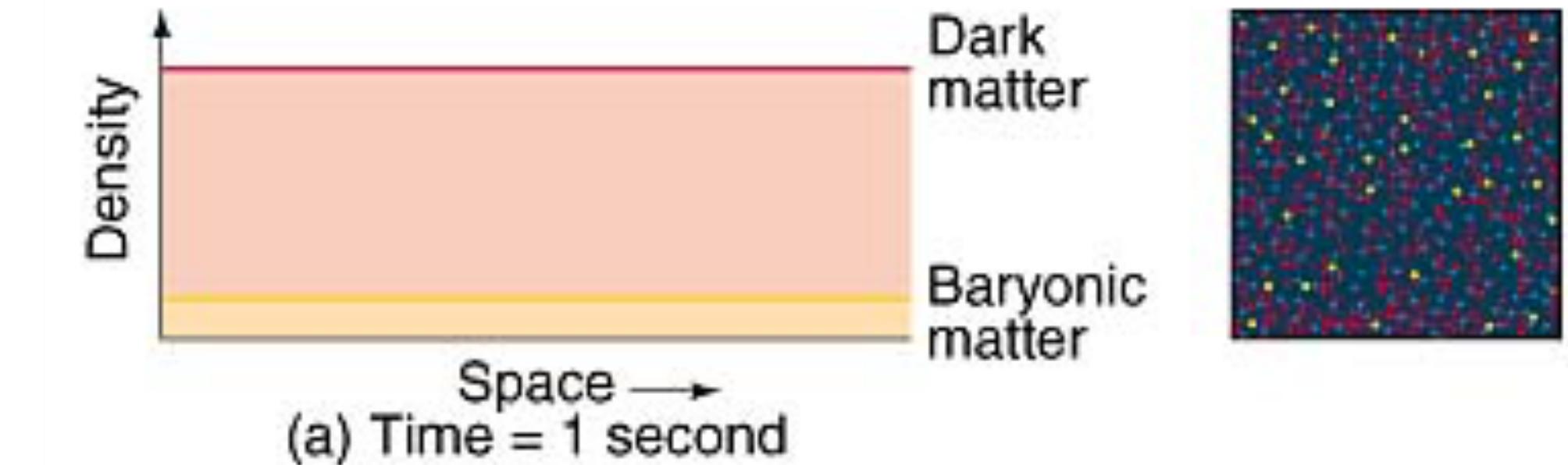
# Galaxy Formation

- Where do galaxies come from?
- Dark matter and dark energy are actually quite easy, the difficult bit is the baryons.
- How can we connect cosmology (amount of stuff) to observables (e.g. distribution of masses of galaxies)?



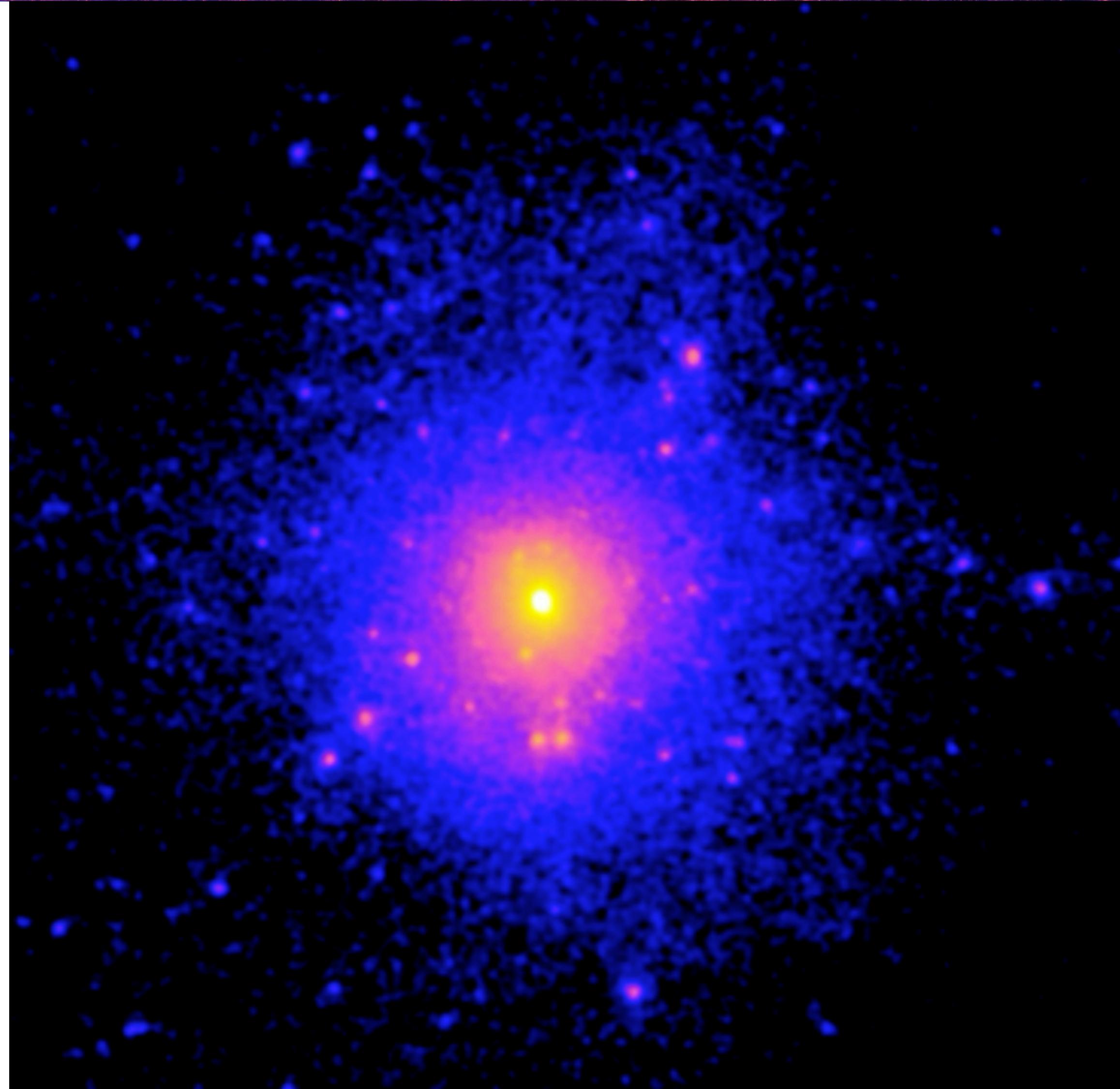
# Dark Matter

- Dark matter drives galaxy formation
- Dense dark matter halos form
- These merge to form larger halos
- Gas falls into these halos and makes galaxies



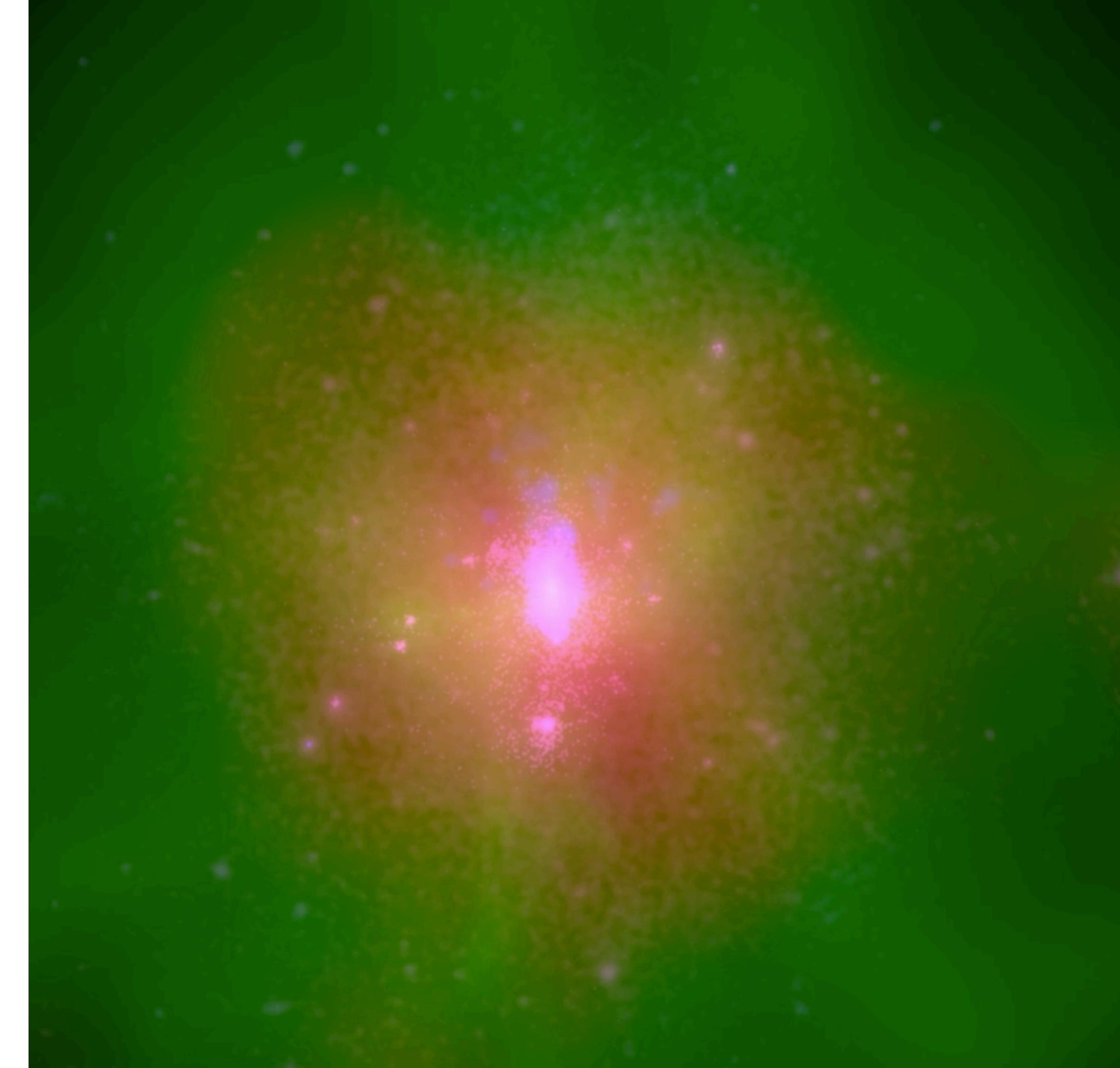
# Dark Matter Halos

- Dark matter is collisionless
- Cannot 'cool'
- Hence halos are spherical (conservation of angular momentum)
- They have a very well characterised density profile (NFW)!



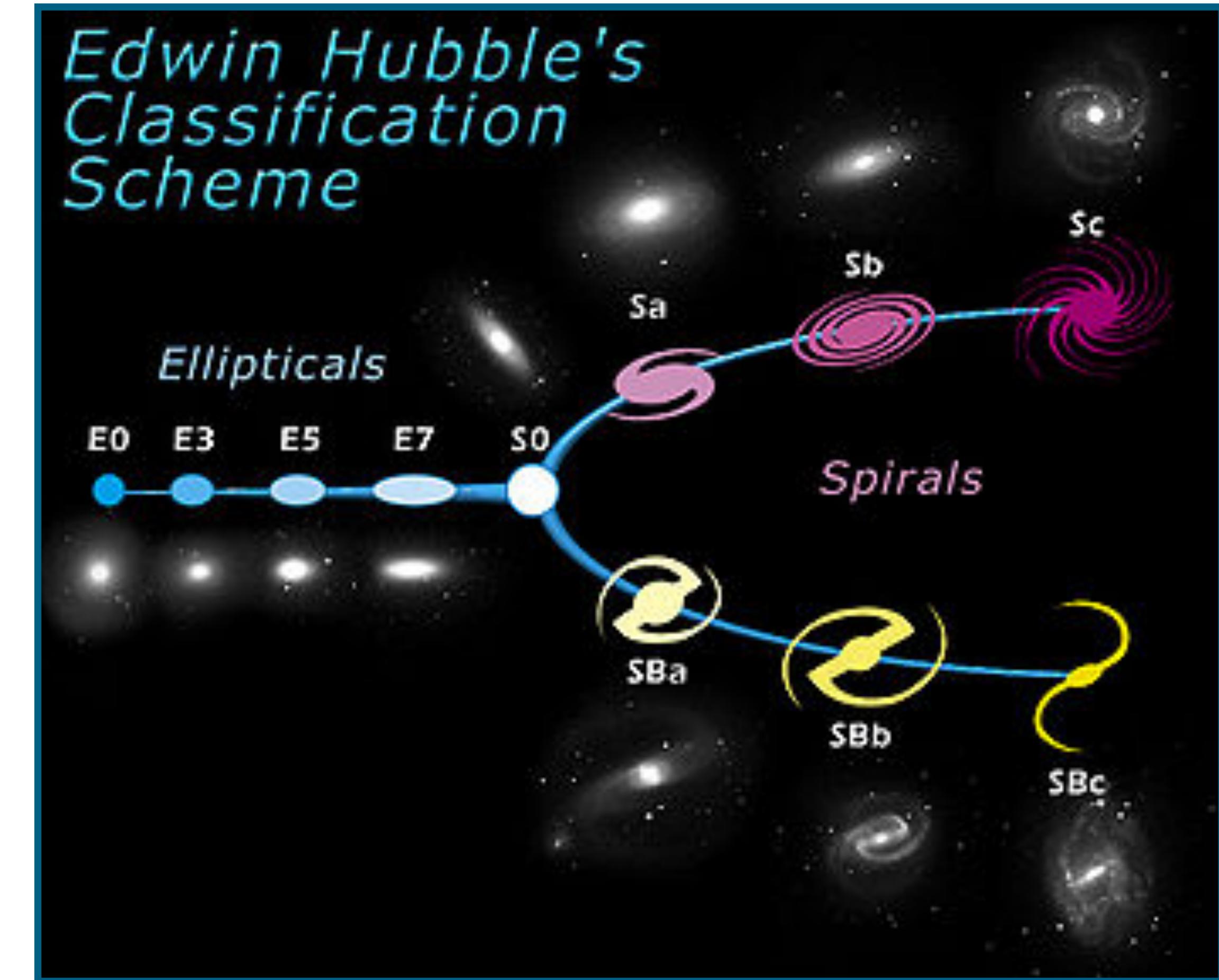
# Gas Cooling

- Gas, unlike dark matter, *can* loose angular momentum
- This allows it to fall into the deepest part of the halo's potential well
- This makes the gas even more dense
- The energy loss from cooling is nonlinear!



# Galaxies?

- So we've got a dense ball of dark matter
- A dense disk of gas
- That's a galaxy, right?



I think  
you'll  
find  
it's a  
bit  
more  
complicated  
than  
that

# Ben Goldacre

Bestselling author of *Bad Science* and *Bad Pharma*



# Galaxies in the Real Universe

# Timescales: The Biggest Killer

- Typical timescales in galaxy formation are ~100s of millions of years
- Your grand<sup>400000</sup>children will be able to see significant changes in the galaxy population

TODAY

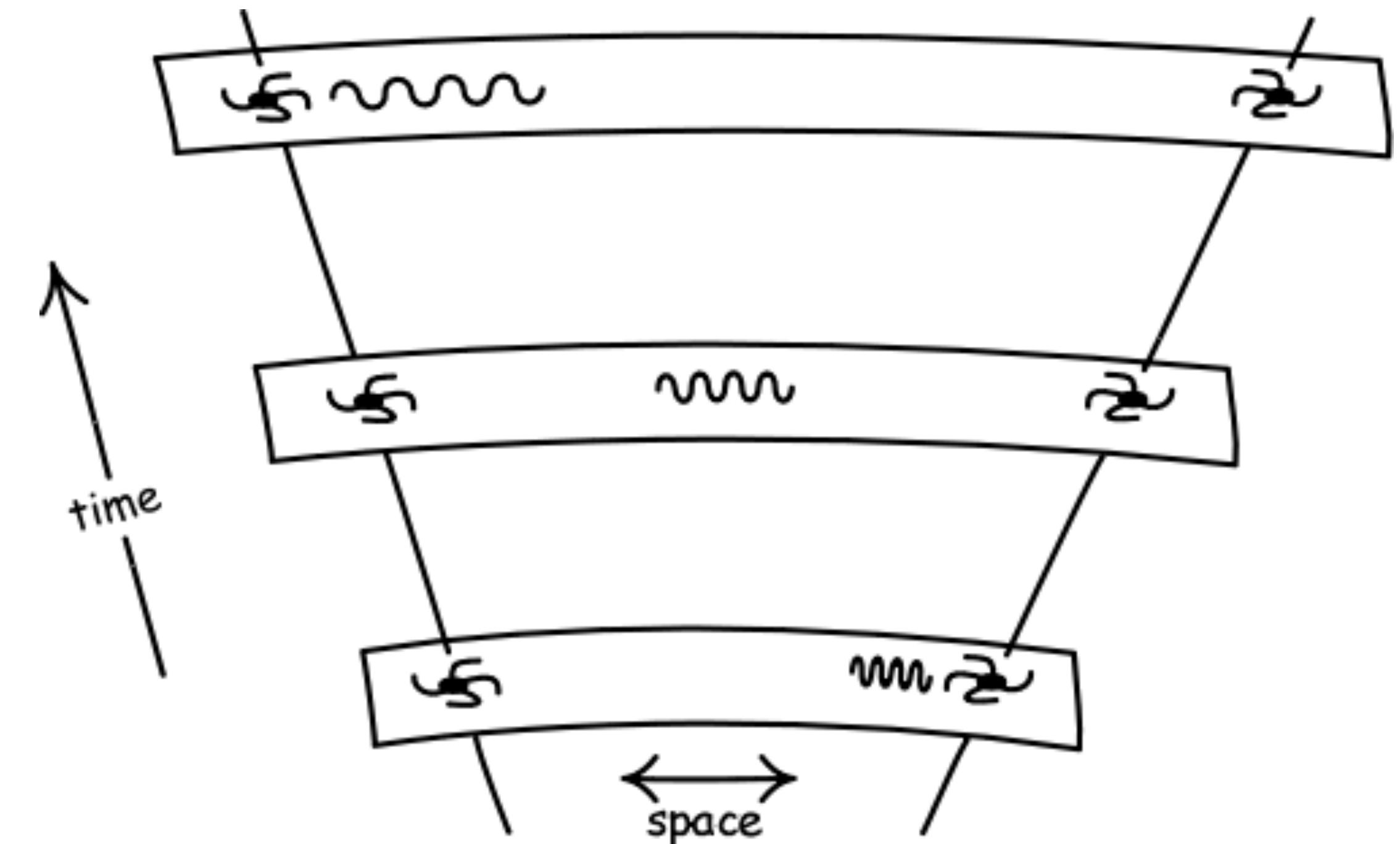


WHEN I DIE (HOPEFULLY IN QUITE A WHILE)



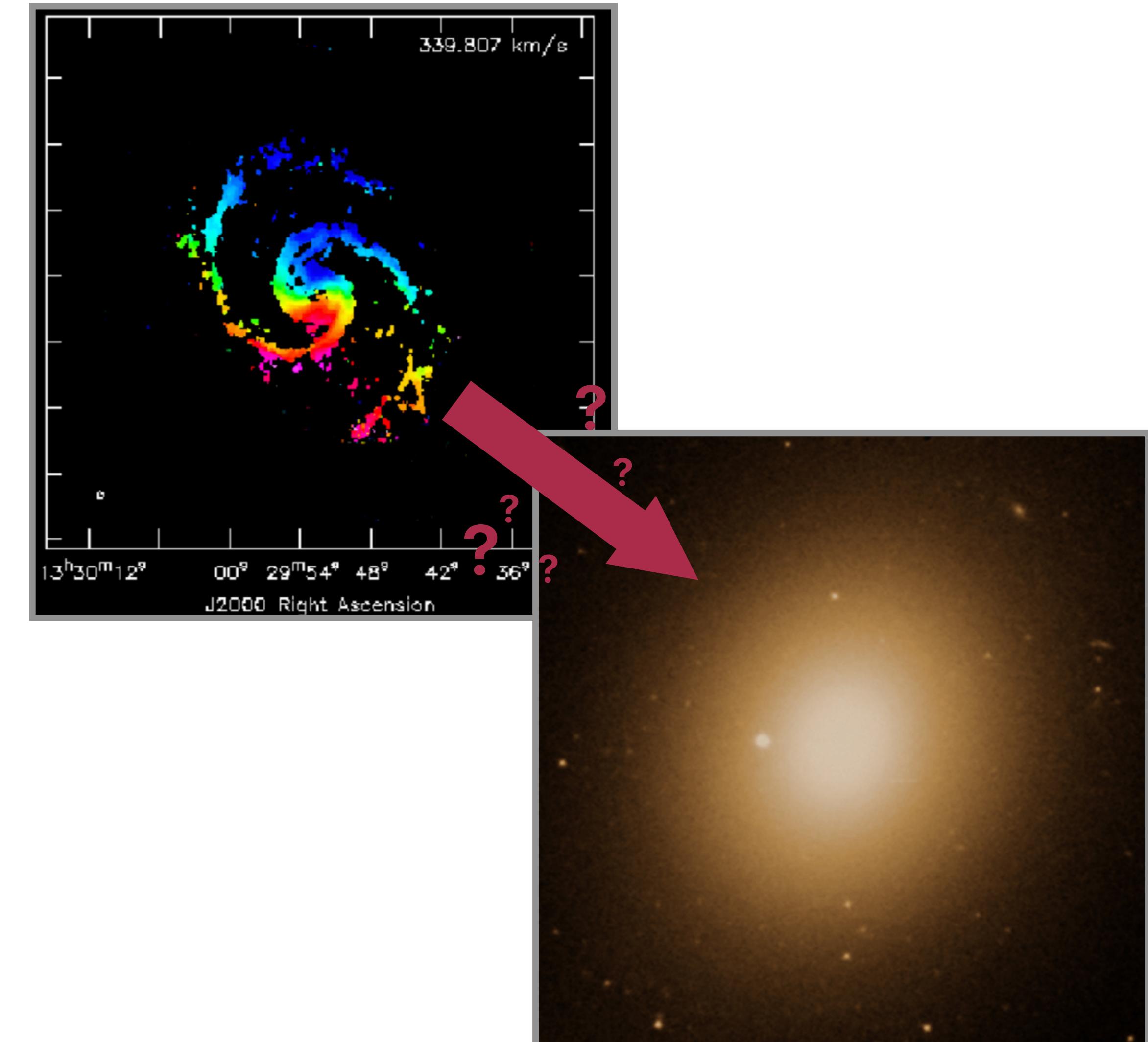
# Addendum: Cosmological Redshift

- Astronomers use redshift as a proxy for distance
- Hubble flow,  $v = H_0 d$  gives recession velocity, which then gives redshift
- Objects at higher redshifts are also seen further back in time because of light's finite speed of travel.



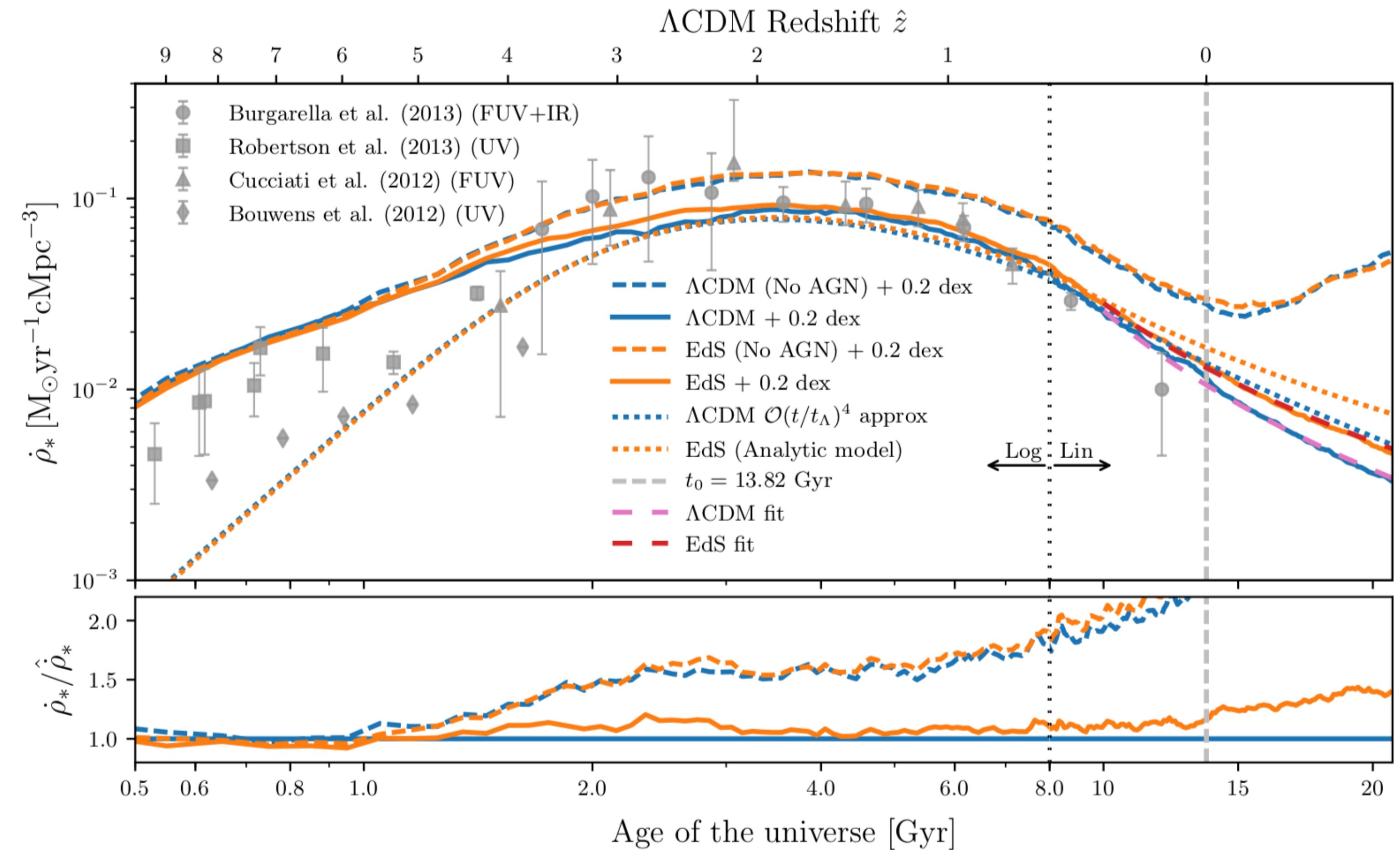
# Progenitor Matching

- We must match *different* galaxies to their nearby counterparts (by mass or other properties)
- This is rubbish as we can't see the direct correspondence between objects



# Statistics

- What we can do is collect statistical information about the galaxy population
- This is a very famous plot that shows the star formation history of the Universe

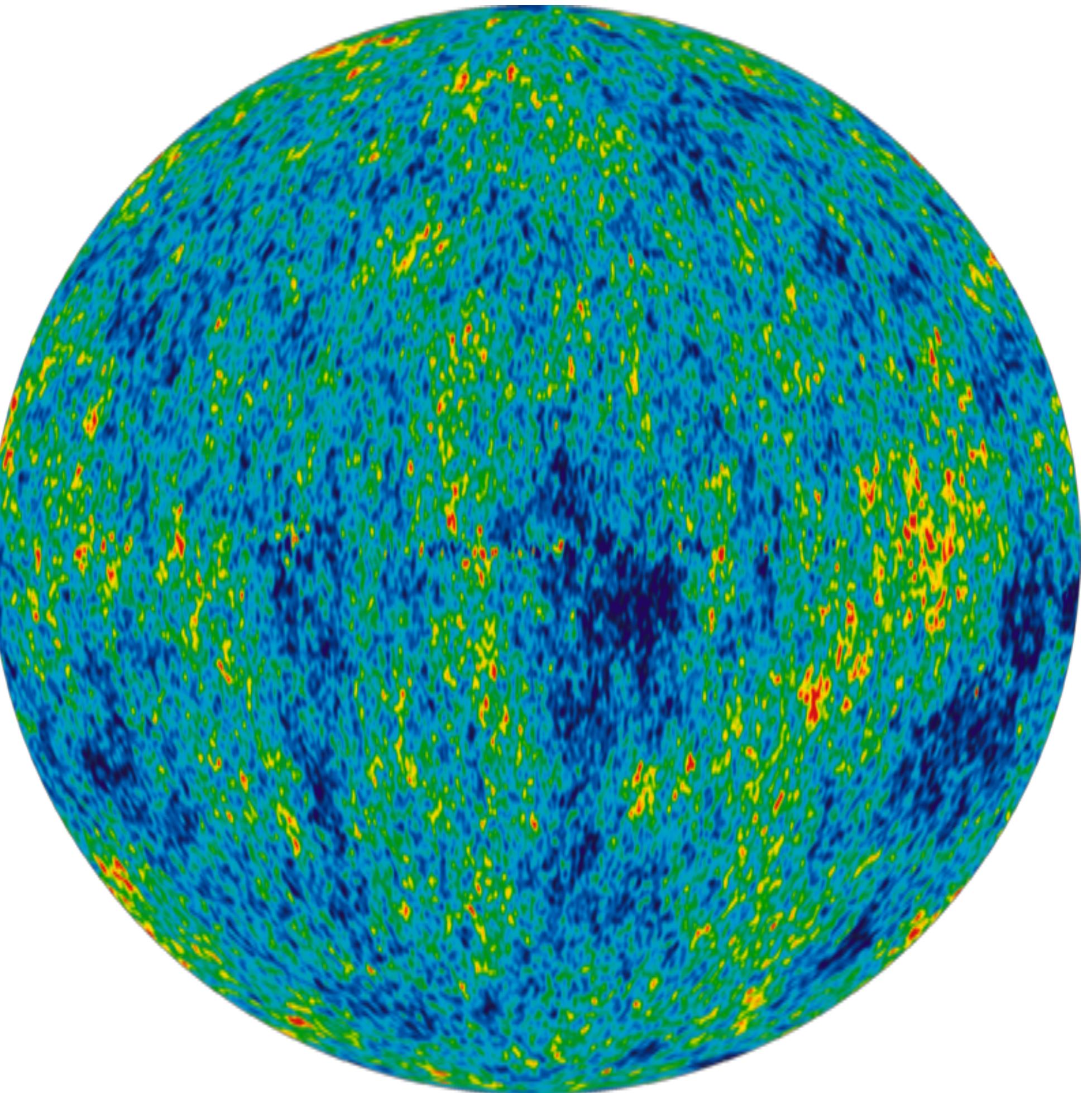


A vibrant, multi-colored nebula with swirling patterns of red, blue, and purple, set against a dark background of stars.

# Understanding Galaxy Formation

# Where can we go from here?

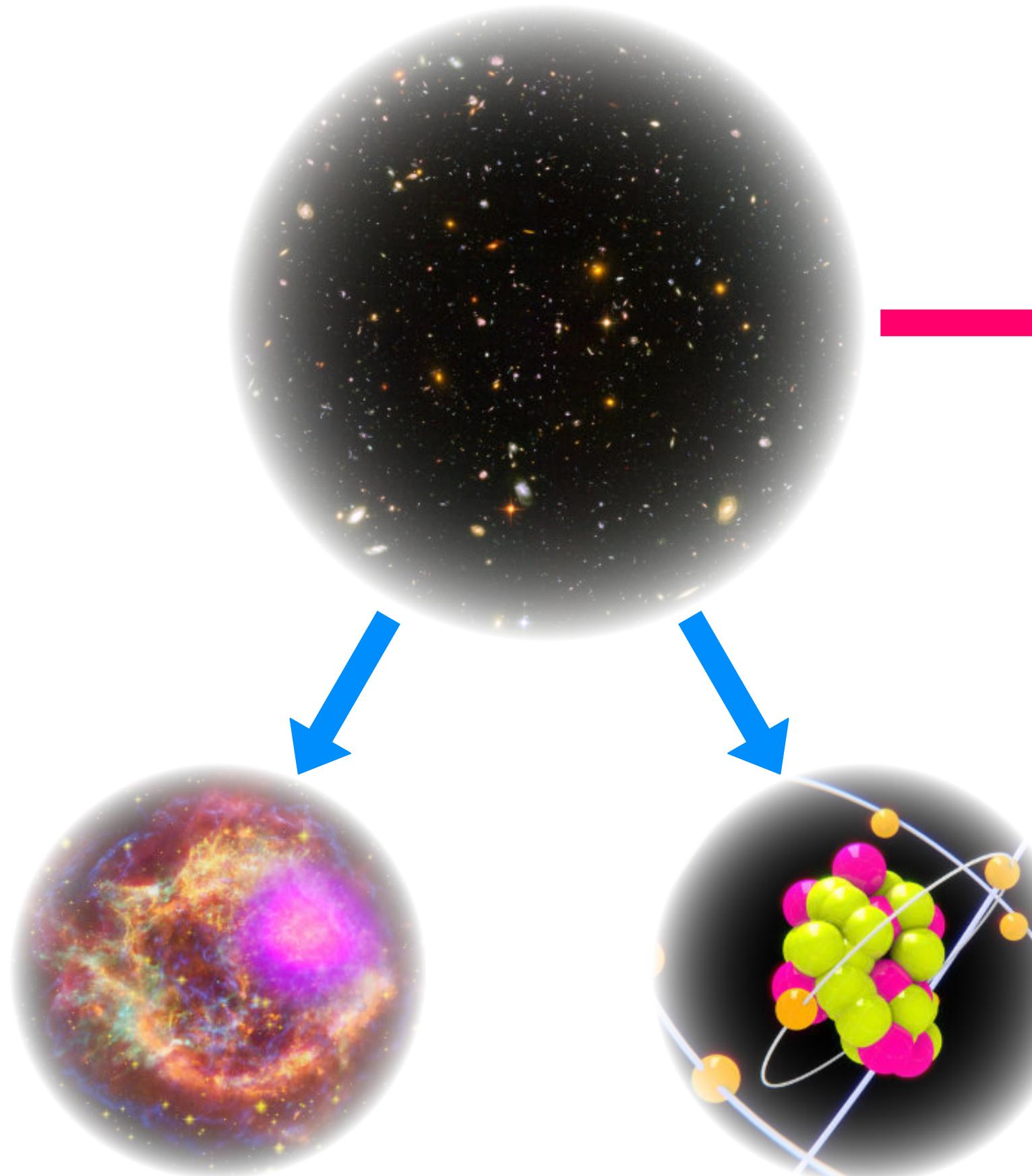
- We have a conceptual idea of galaxy formation
- Also have some statistical properties of the Universe
- No idea how a specific galaxy formed
- Difficult to pin down what processes are important!



The cosmic microwave background (CMB) as seen by WMAP

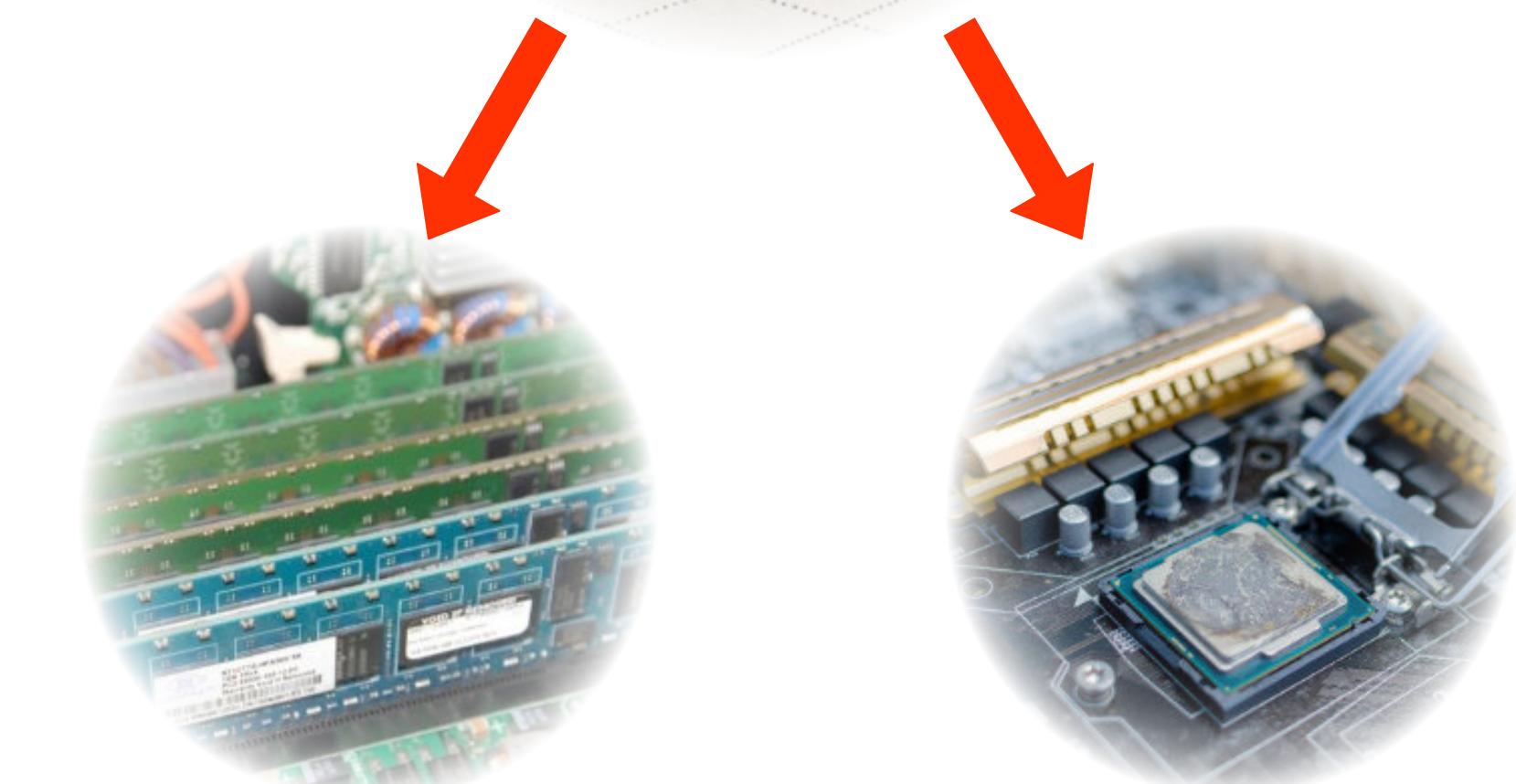
# Running a Simulation of the Universe

**The Universe**



Infinite Volume

**HPC System**



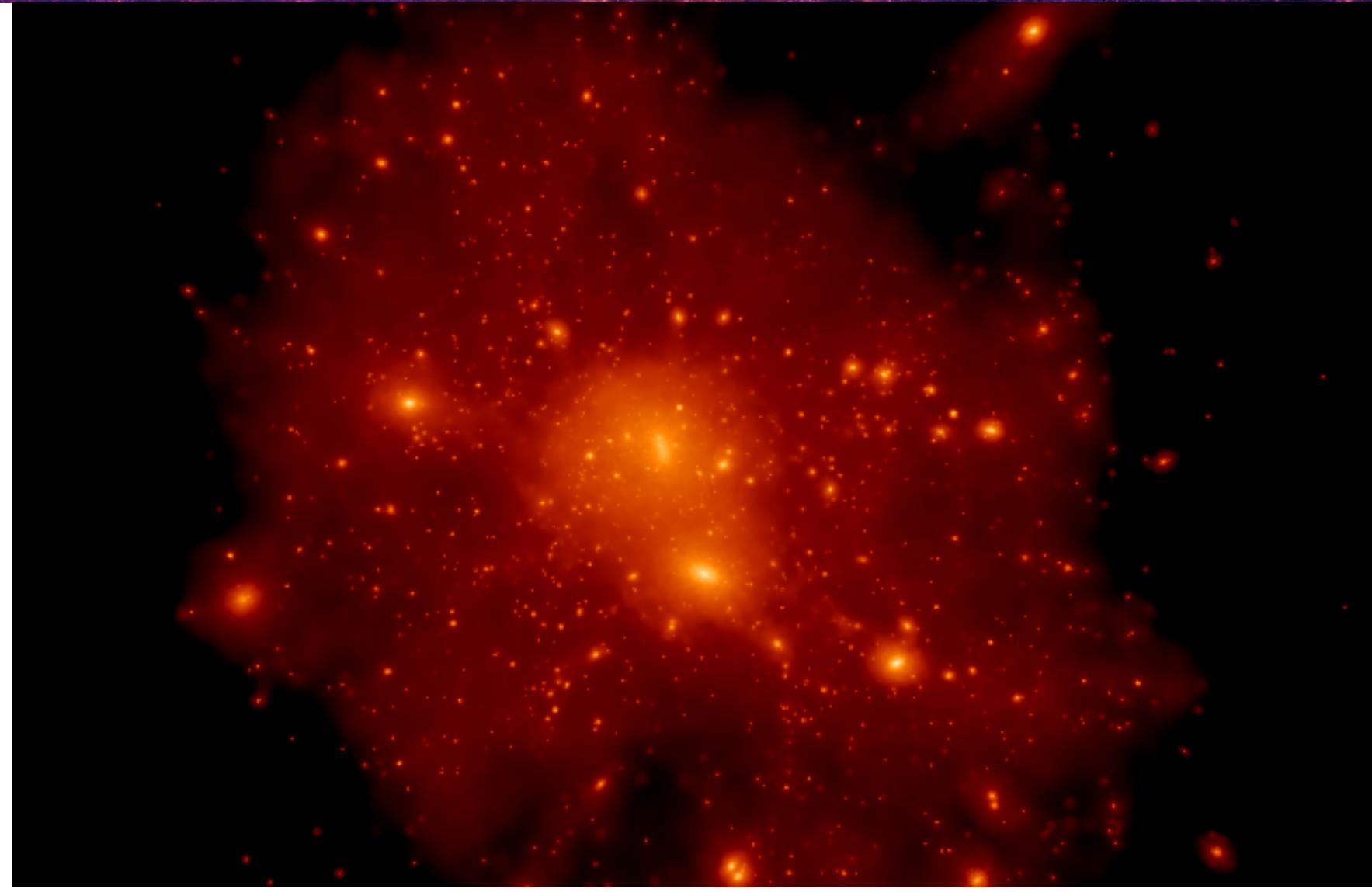
Limited Memory

Finite Scale

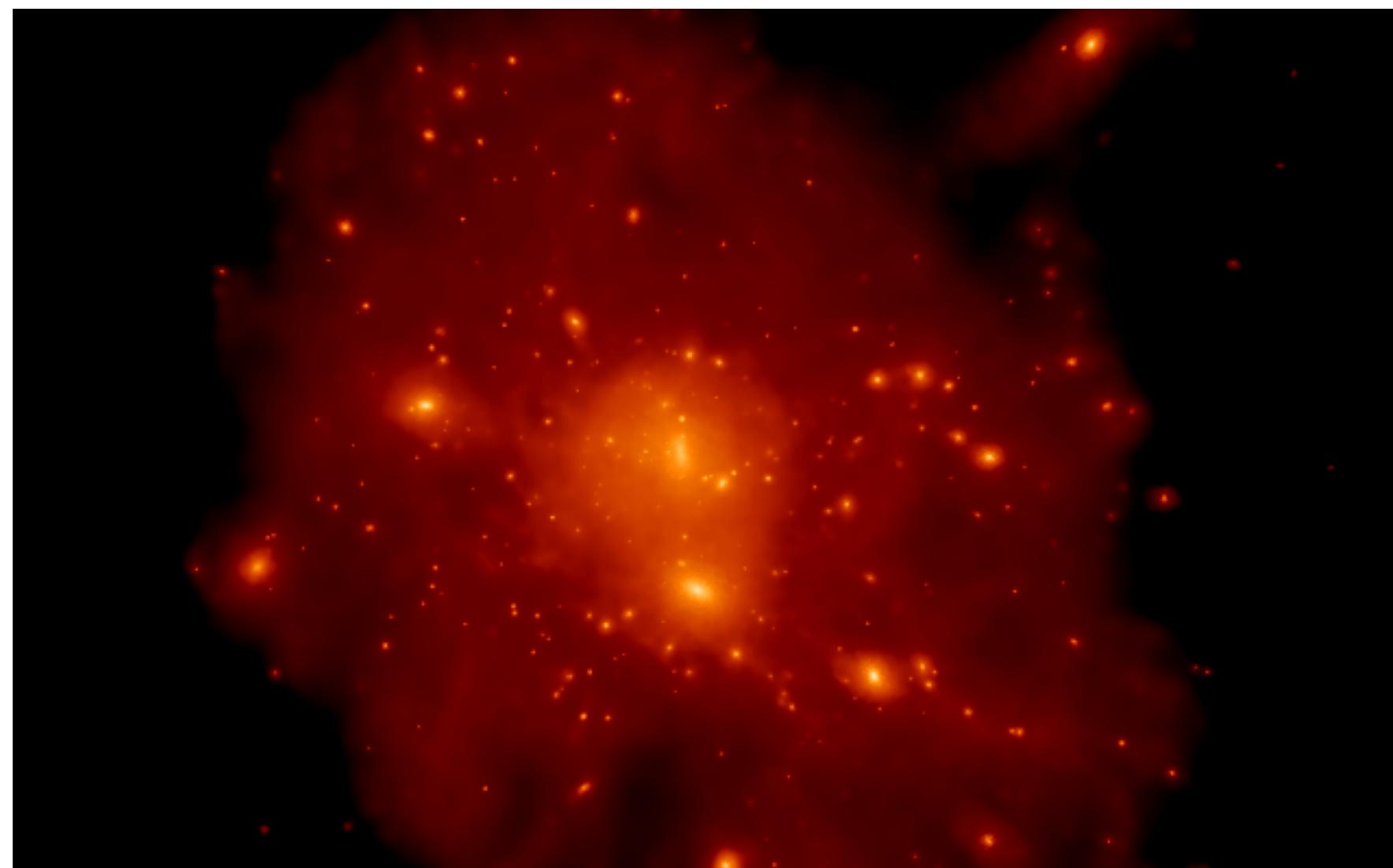
# Step 1: Get the Dark Matter Right

- Cosmology tells us *how much* dark matter there is, but not of what *type*.
- We can run gravity-only simulations of dark matter to determine how it affects structure formation in the Universe.
- These allow us to predict e.g. how many satellite galaxies we'd expect around the Milky Way

CDM: WIMP

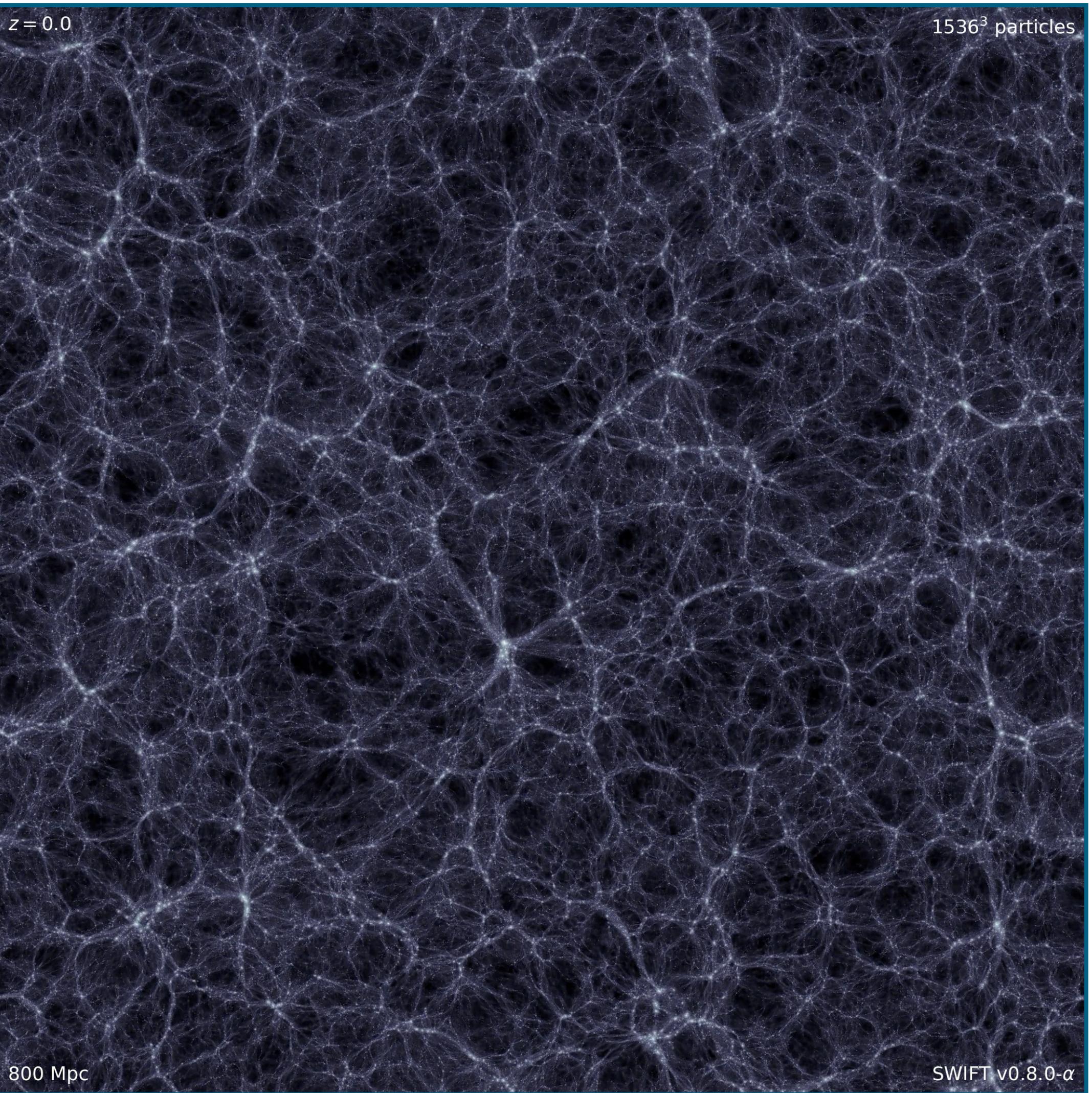


WDM: Neutrinos



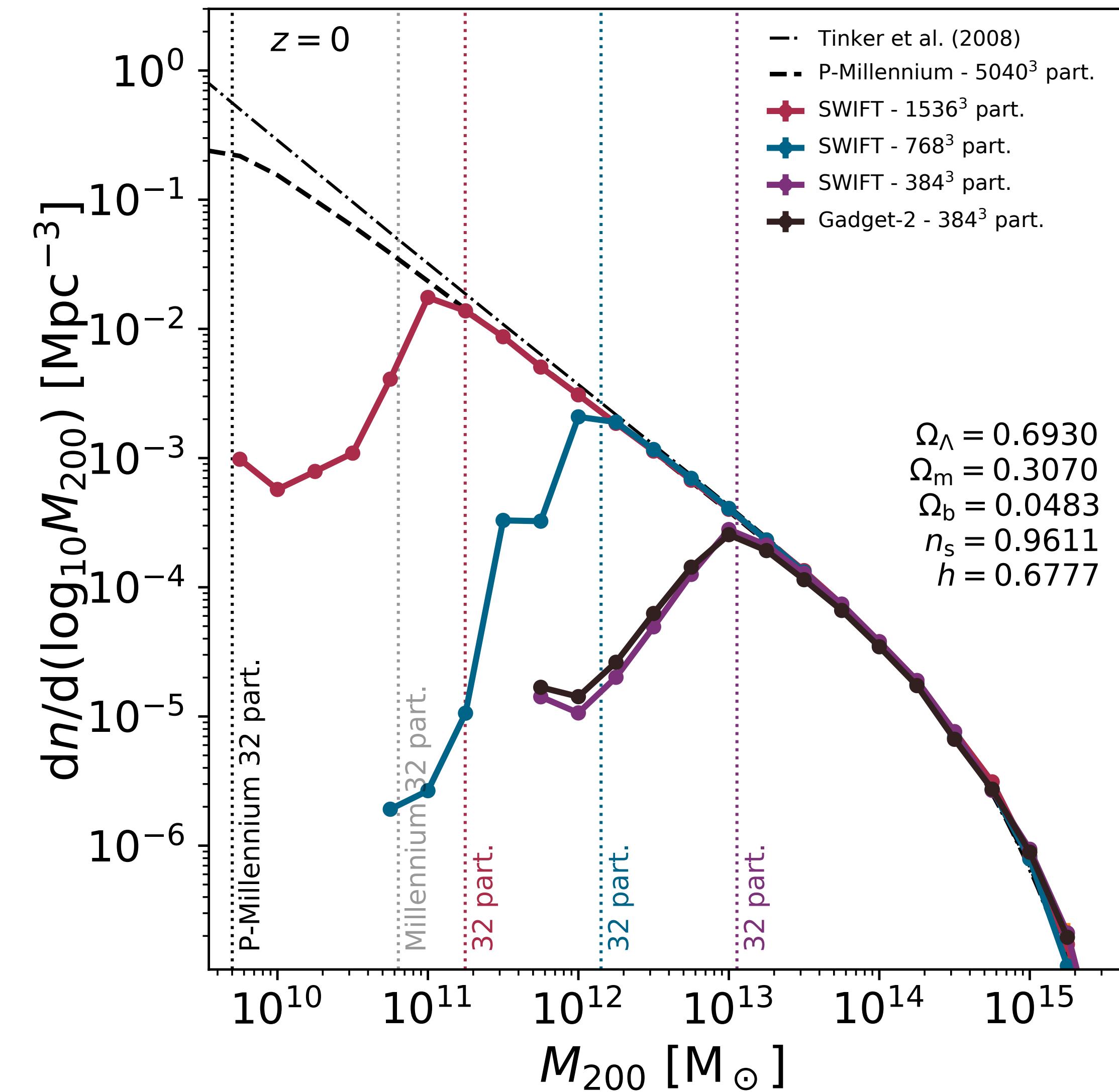
# Step 1: Get the Dark Matter Right

- We can run these 'N-body' simulations super fast these days
- $5000^3$  or higher (125 billion) particle simulations are not uncommon



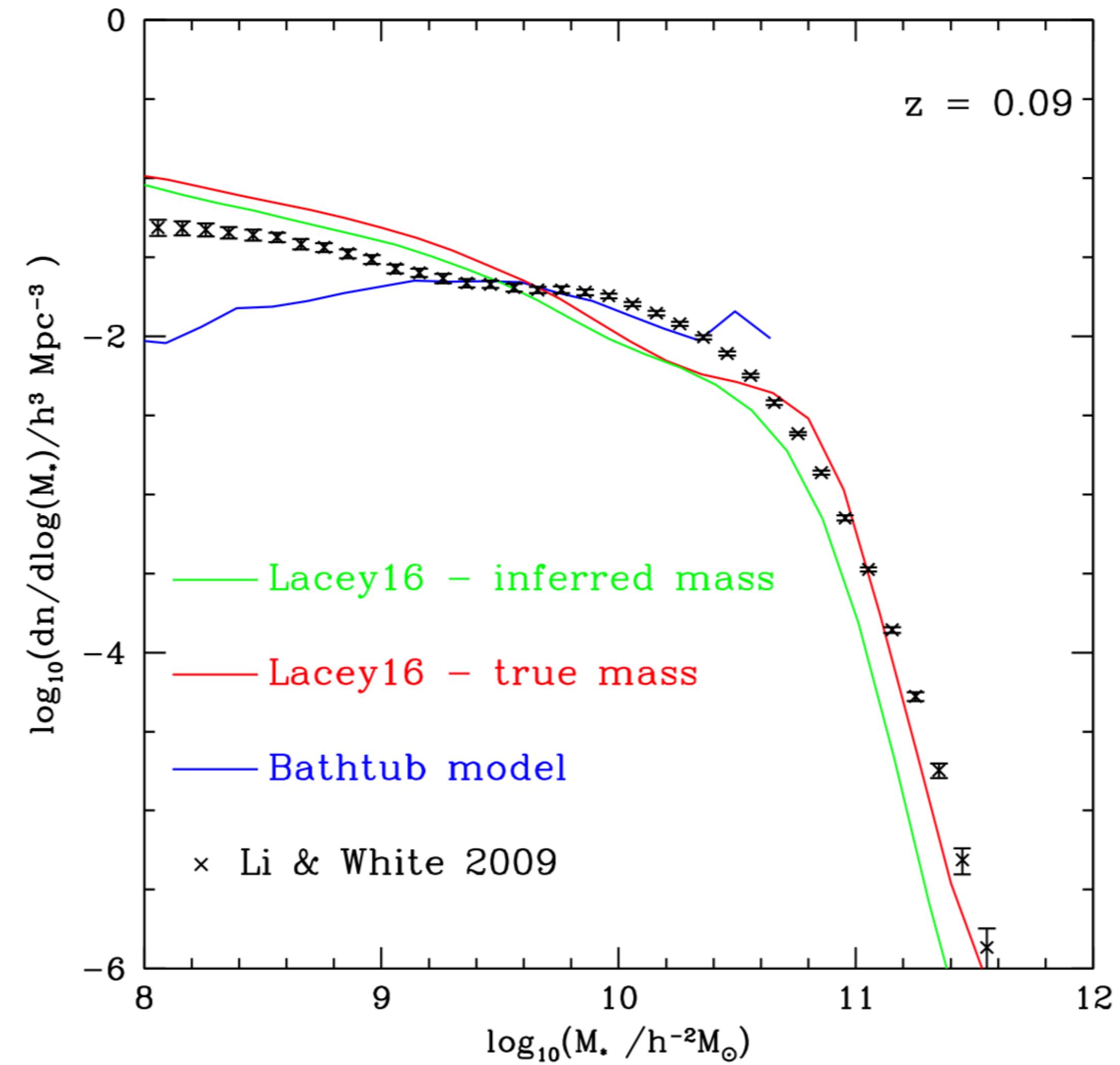
# The Halo Mass Function

- Shows the abundance of dark matter halos as a function of mass
- This is self-similar all the way down
- Clearly that's not the same for baryons...



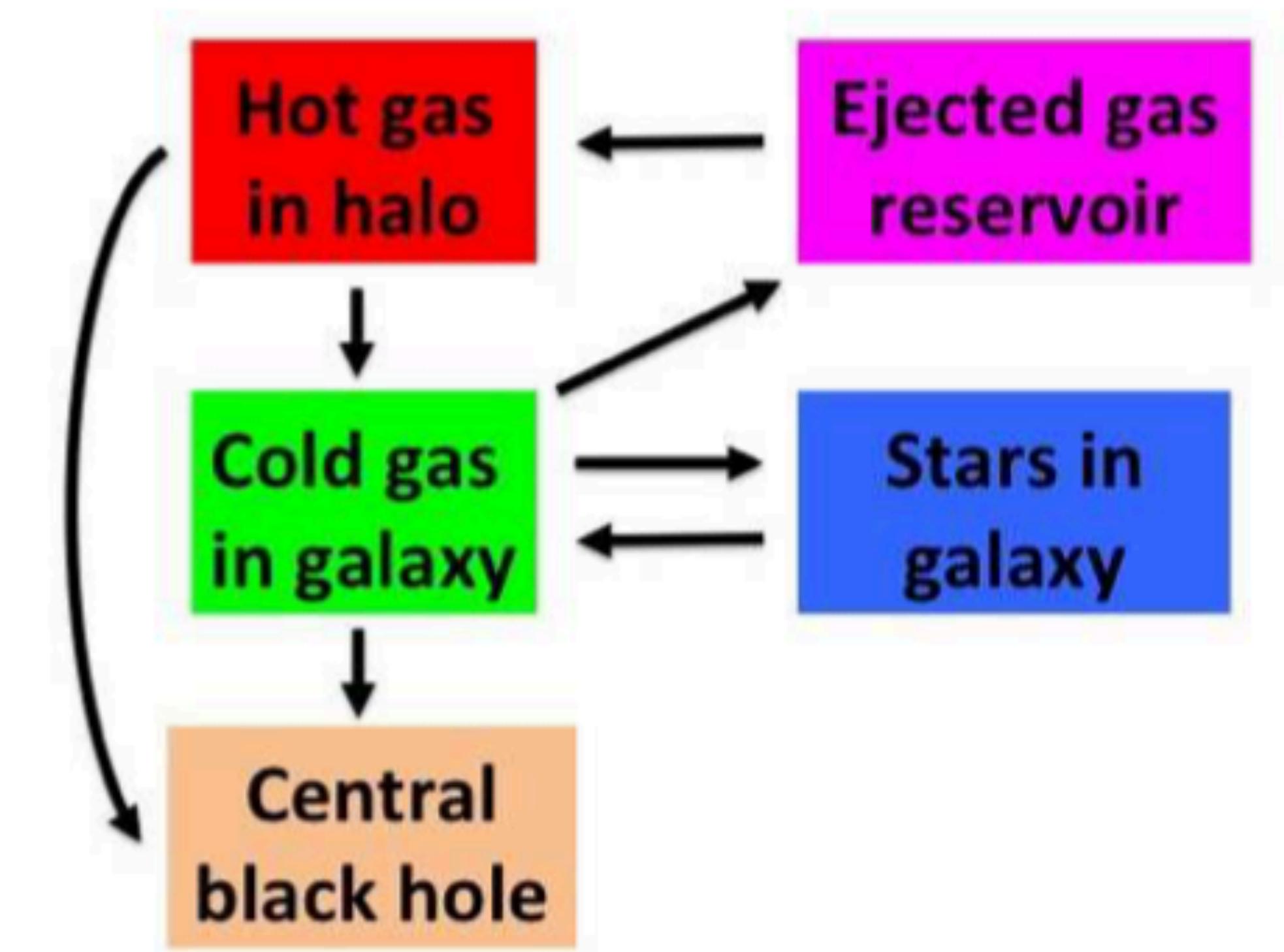
# Semi-Analytic Models

- Use the dark matter-only simulations as a base
- Stick on top a 'semi-analytic' model of galaxy formation that has equations to model galaxy properties
- Caveat: these are boring AF and have no pretty pictures



# What do SAMs say about Galaxy Formation?

- A lot.
- Pretty much everything we know comes from these models.
- Important processes:
  - Star formation and stellar “feedback”
  - Black holes and their associated “AGN feedback”
  - Gas cooling

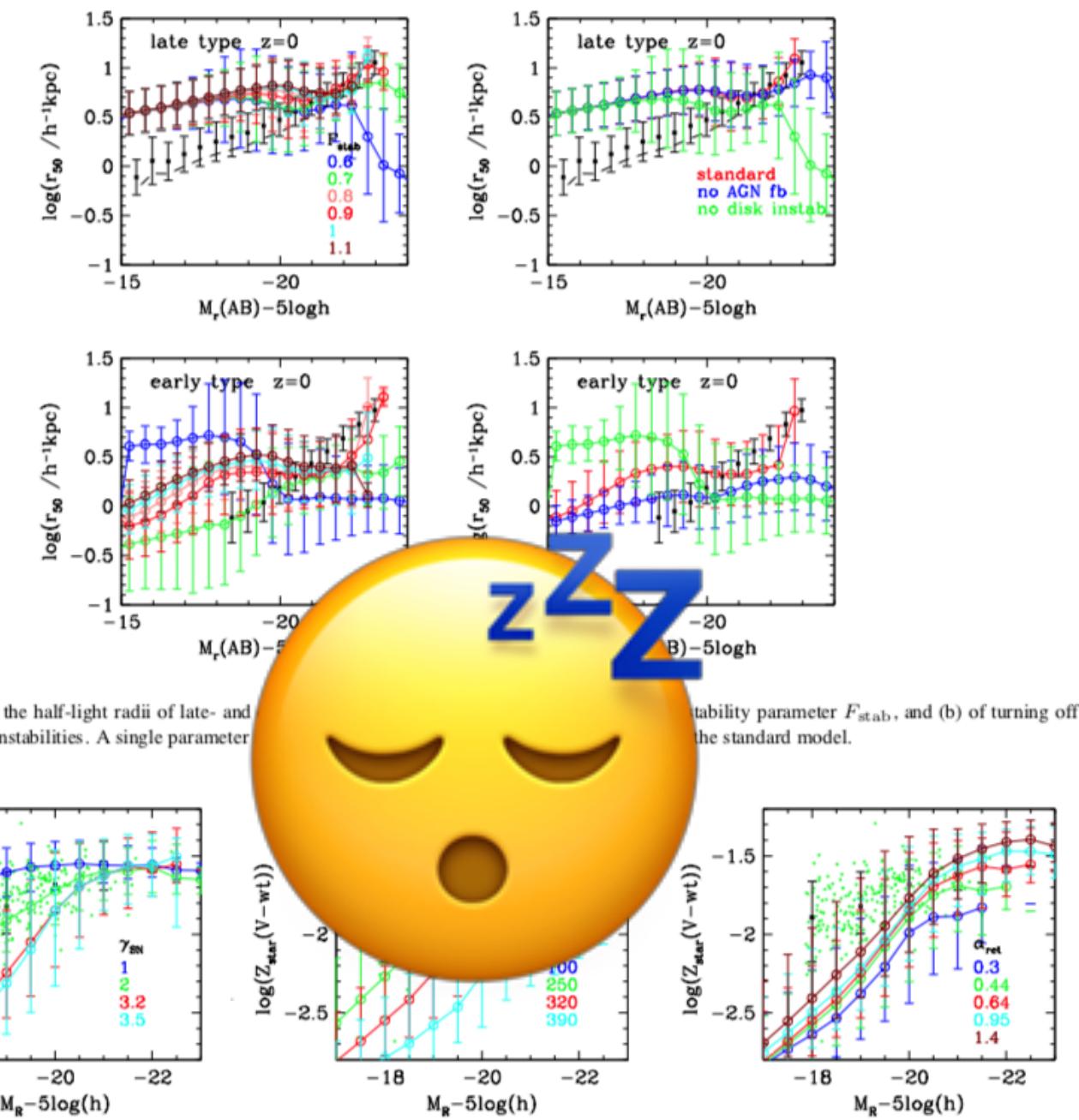


**Figure 1.** Flow chart showing the different baryonic components in a halo, and transfers between them.

# The Problem with SAMs

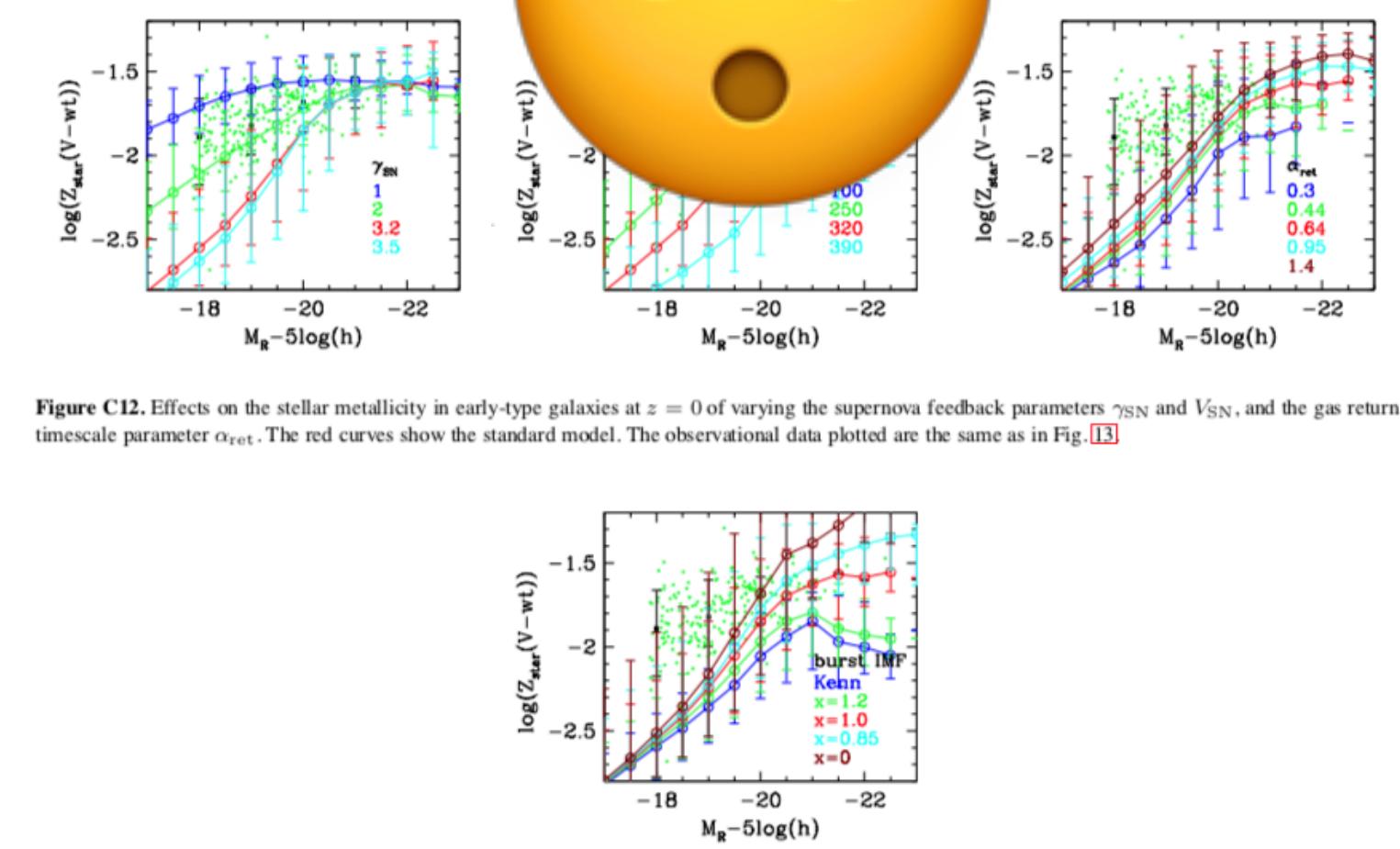
- Boring
- Ugly
- No resolved properties!
- Cannot study the interaction of galaxies with their environment
- Difficult to study galaxy interactions

52 Lacey et al.



**Figure C11.** Effects on the half-light radii of late- and AGN feedback or disk instabilities. A single parameter

stability parameter  $F_{\text{stab}}$ , and (b) of turning off the standard model.



**Figure C12.** Effects on the stellar metallicity in early-type galaxies at  $z = 0$  of varying the supernova feedback parameters  $\gamma_{\text{SN}}$  and  $V_{\text{SN}}$ , and the gas return timescale parameter  $\alpha_{\text{ret}}$ . The red curves show the standard model. The observational data plotted are the same as in Fig. 13

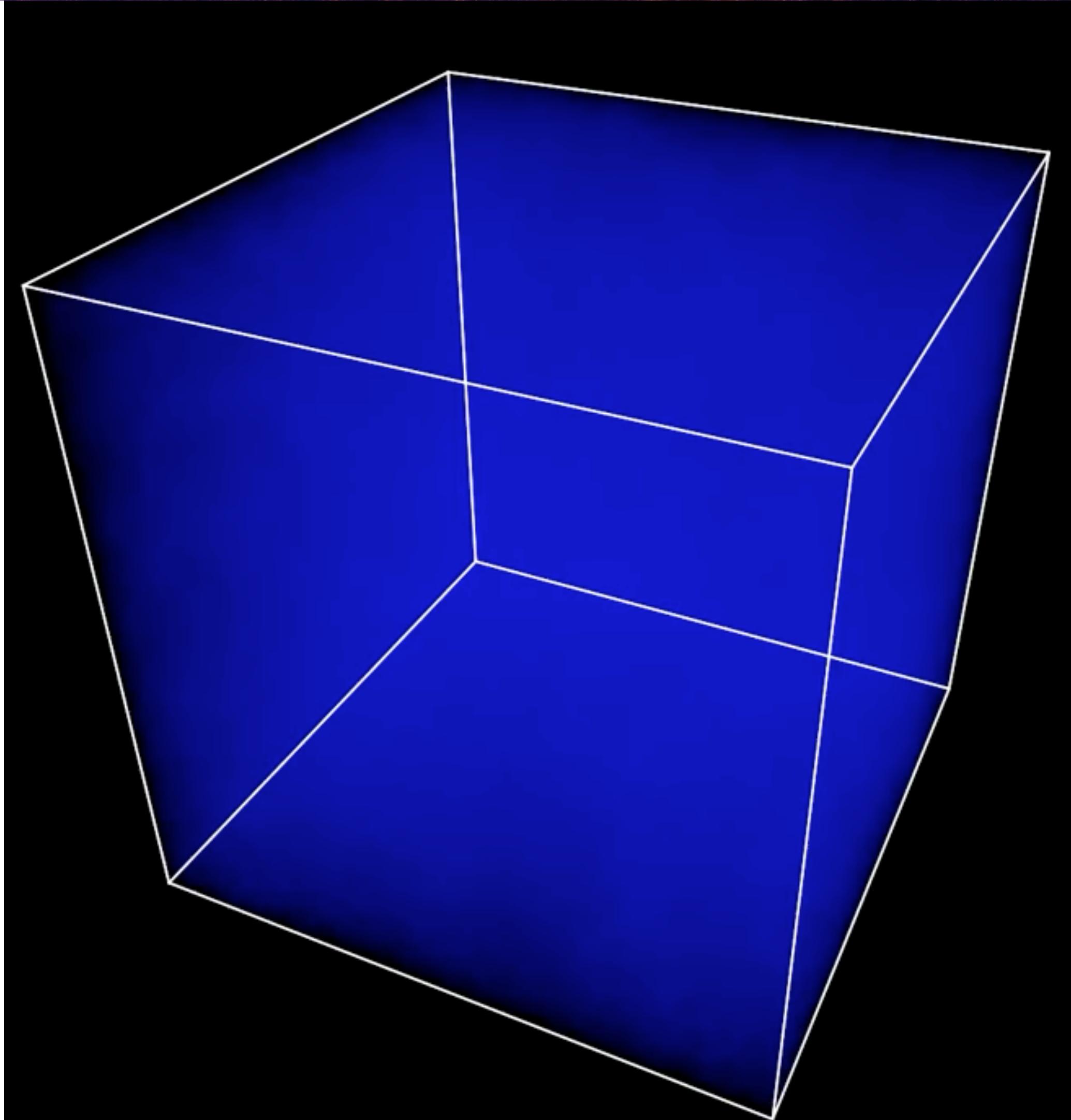
**Figure C13.** Effects on the stellar metallicity in early-type galaxies at  $z = 0$  of changing the slope  $x$  of the starburst IMF.



# Cosmological Hydrodynamics Simulations

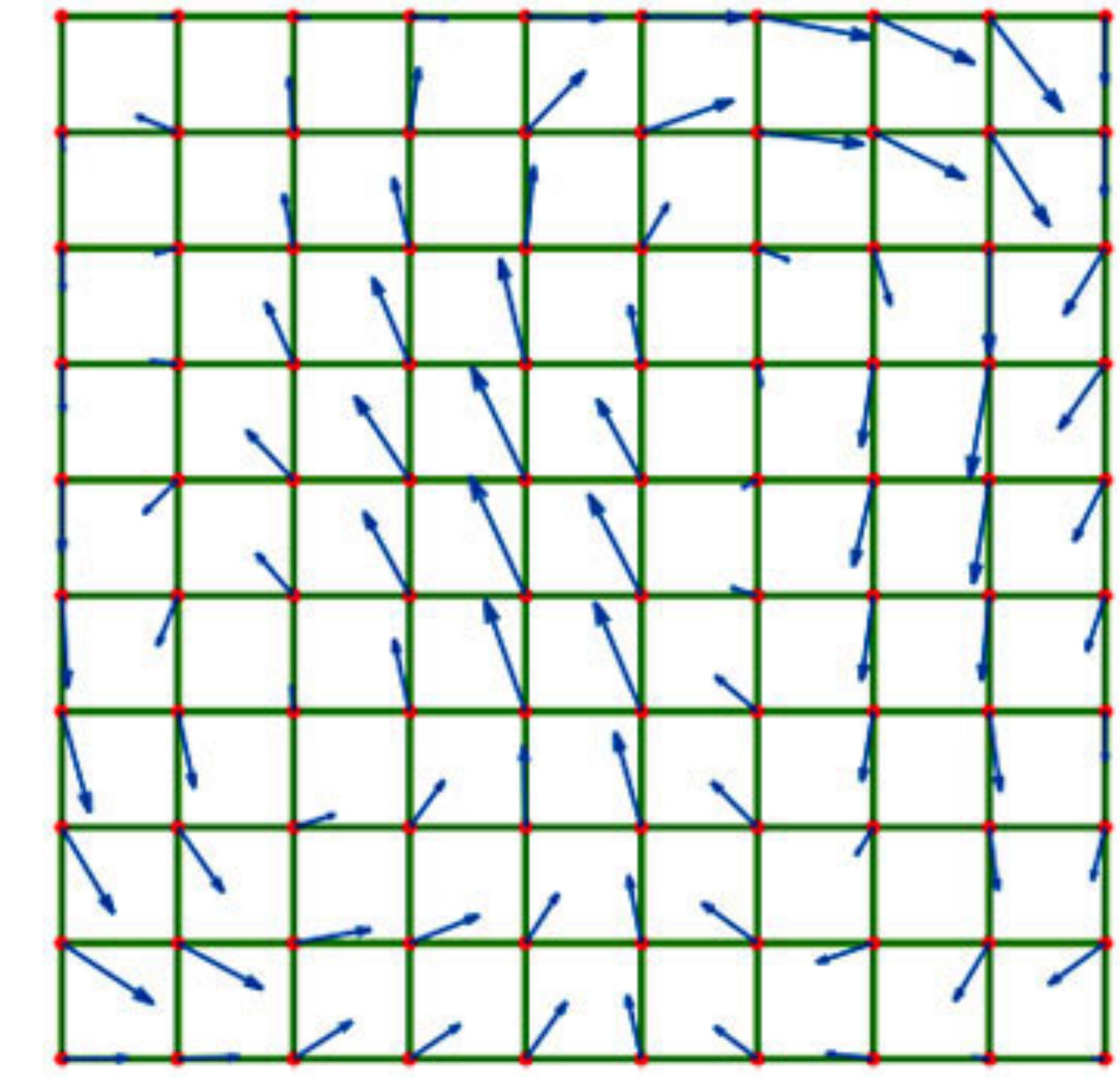
# Step 2: Do it ‘Properly’

- Need to couple (at least) three main key pieces of physics
  - Cosmology
  - Hydrodynamics
  - Gravity
  - (+ Radiative transfer..???)
  - (+ Magnetic fields..???)

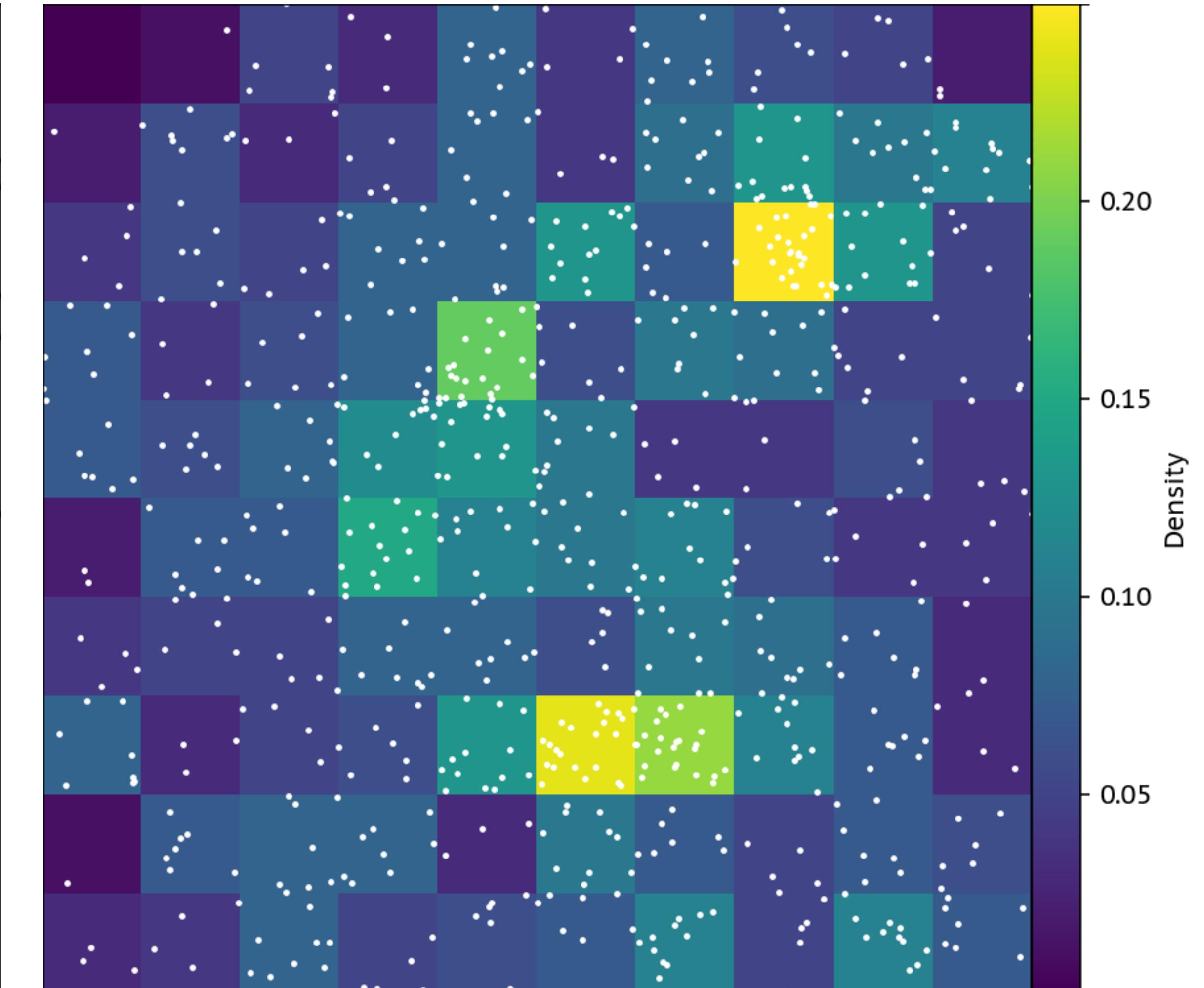
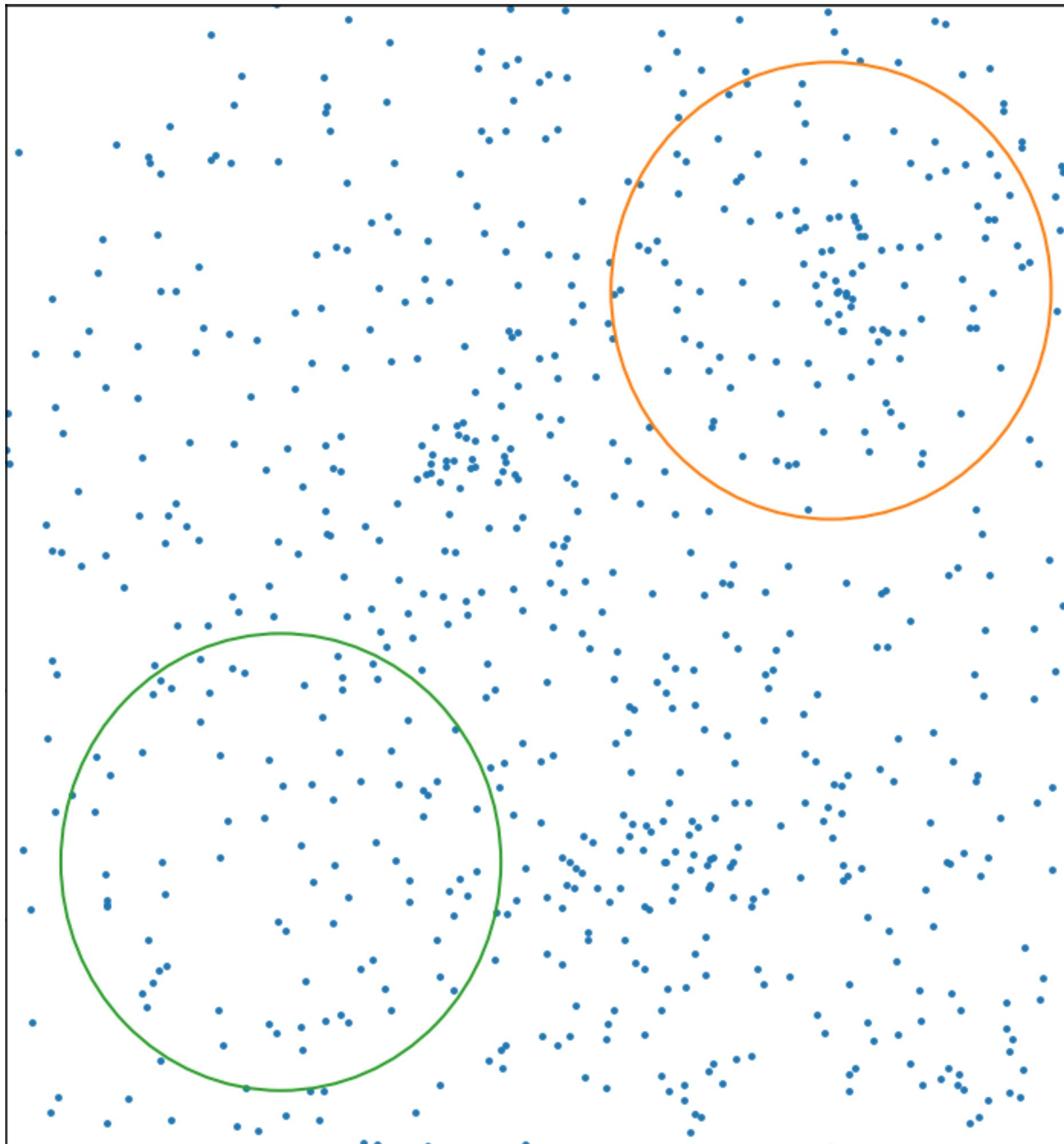


# Hydrodynamics

- Typical way to do this: lay down a grid and compute fluid flows between cells
- Several problems:
  - Lack of adaptivity
  - **Cannot track fluid flow over time**
  - Makes gravity difficult



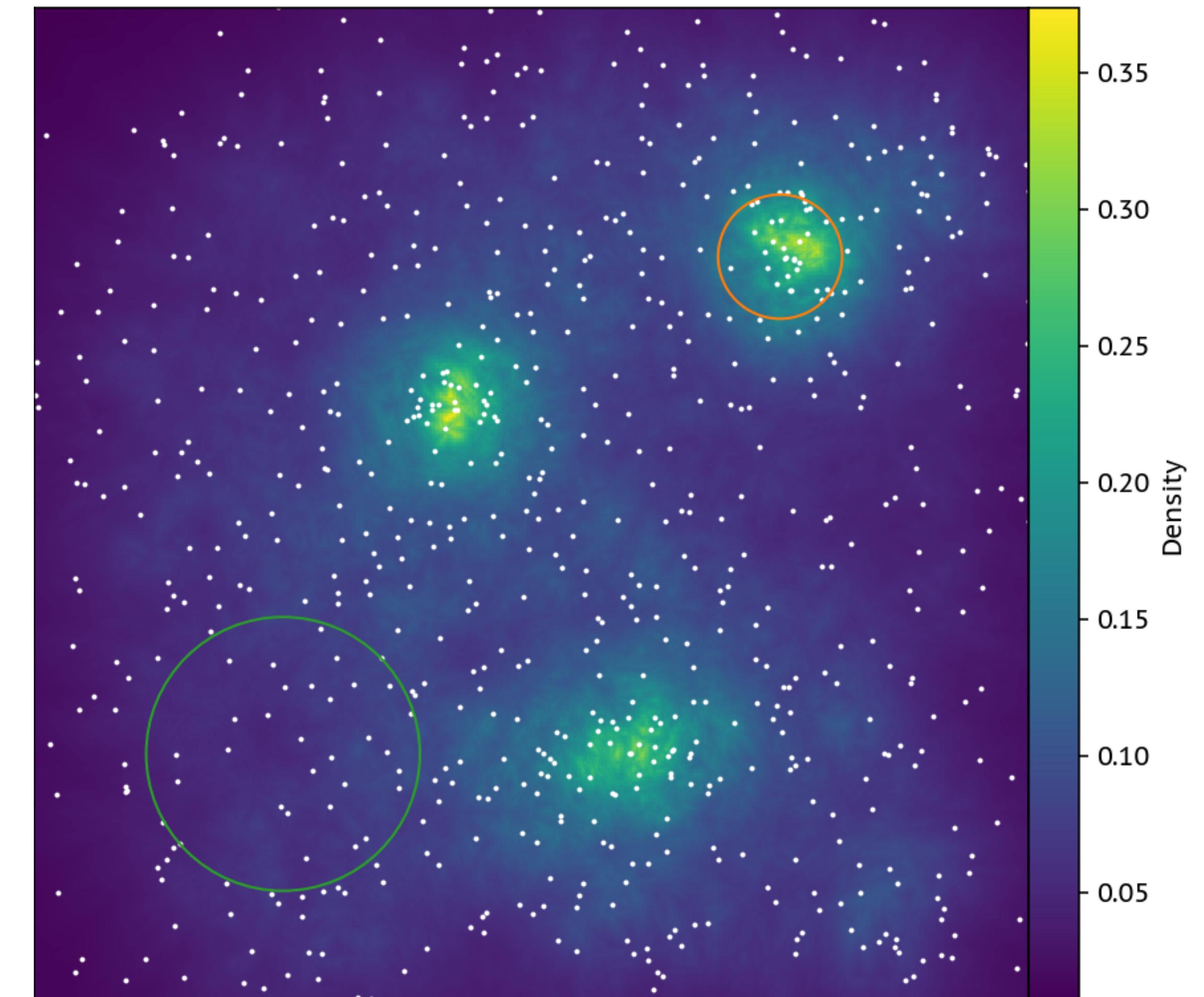
# Hydrodynamics with Particles



# Improving our Density Estimate

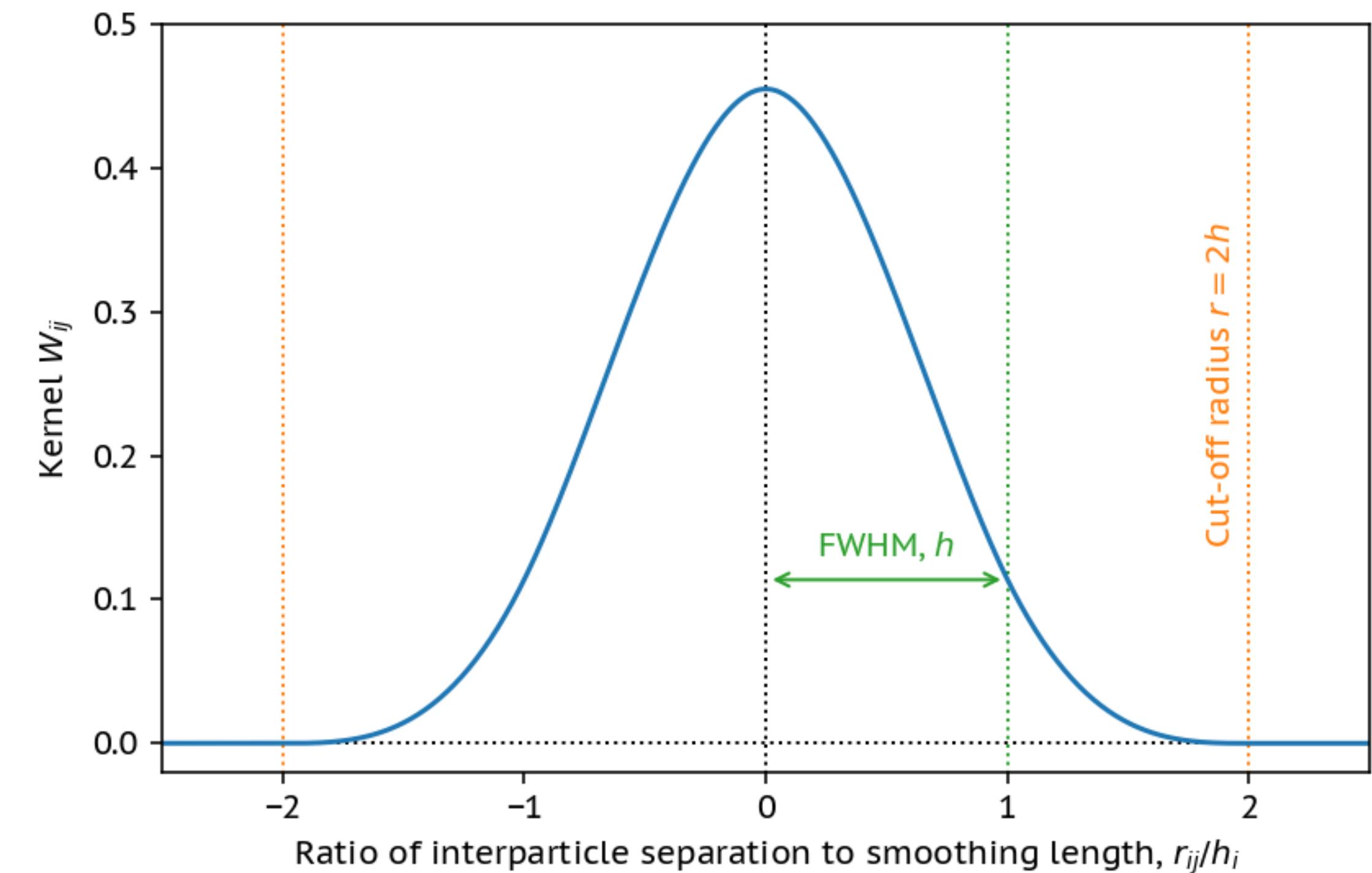
- What if we change the volume we consider for each particle?
- Set R such that our volume encloses 30 particles.
- Density for each particle is then

$$\rho_i = \frac{30m_{\text{part}}}{\frac{4}{3}\pi R_i^3}$$



# Turning Down the Noise

- We can do better! We care less about particles that are further away.
- Weight the contribution from each particle with a “kernel”.
- Size kernel to enclose 30 particles within cut-off radius.

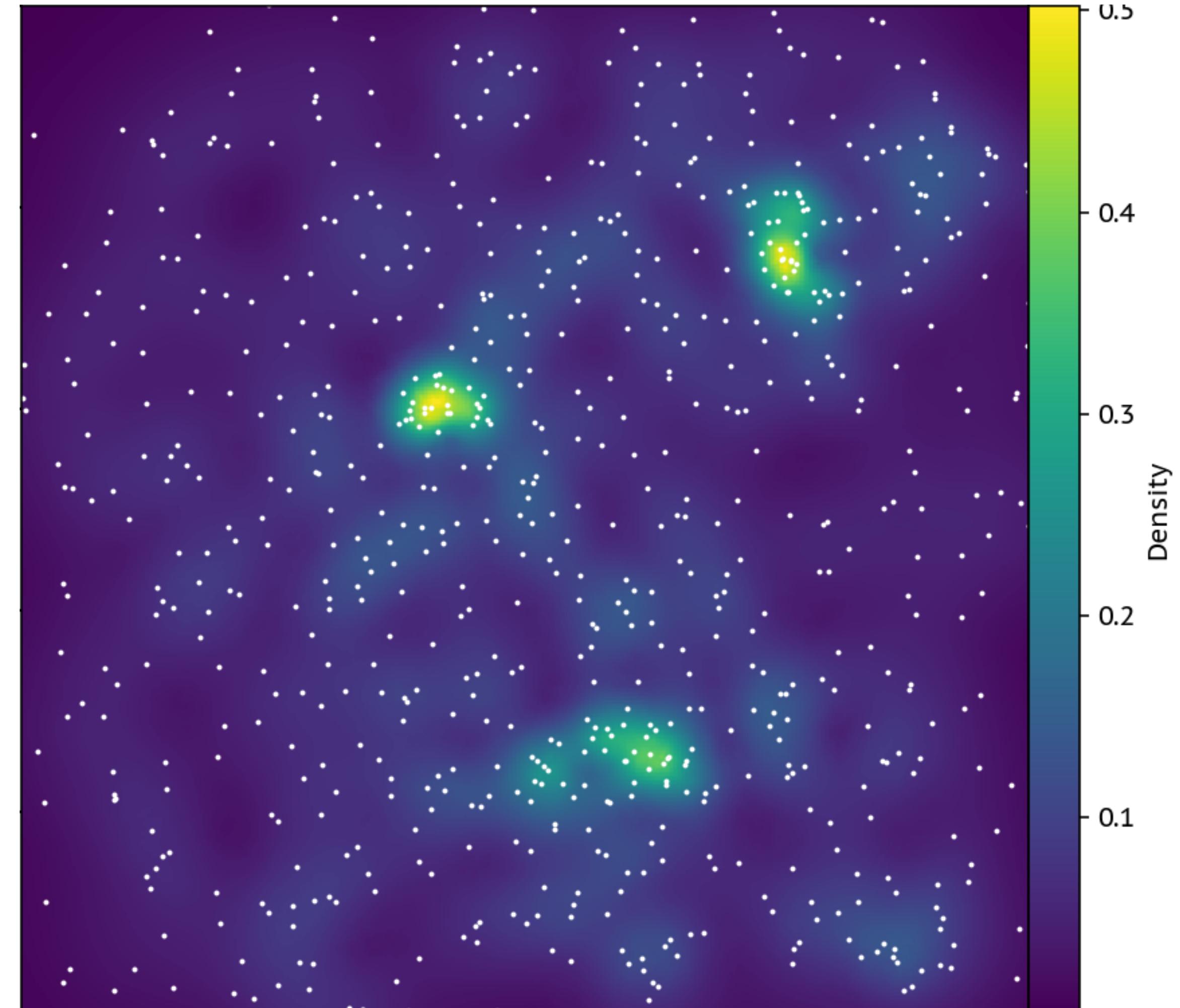


# A Smooth Density Estimate

- Density for each particle now given by:

$$\rho_i = \sum_{j=1}^{30} m_{\text{part}} W(r_{ij}, h_i)$$

- Density (+ temperature, tracked by particles) gives pressure, which gives dynamics



# Adding in the ‘Dynamics’

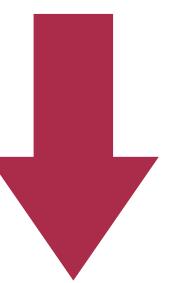
$$L(q, \dot{q}) = \frac{1}{2} \sum_{i=1}^N m_i \dot{r}_i^2 - \sum_{i=1}^N m_i u_i,$$

$$\left. \frac{\partial u_i}{\partial q_i} \right|_A = - \frac{P_i}{m_i} \frac{\partial \Delta V_i}{\partial q_i}, \quad \phi_i(\mathbf{q}) = \kappa h_i^{n_d} \frac{1}{\Delta \tilde{V}} - N_{ngb} = 0,$$

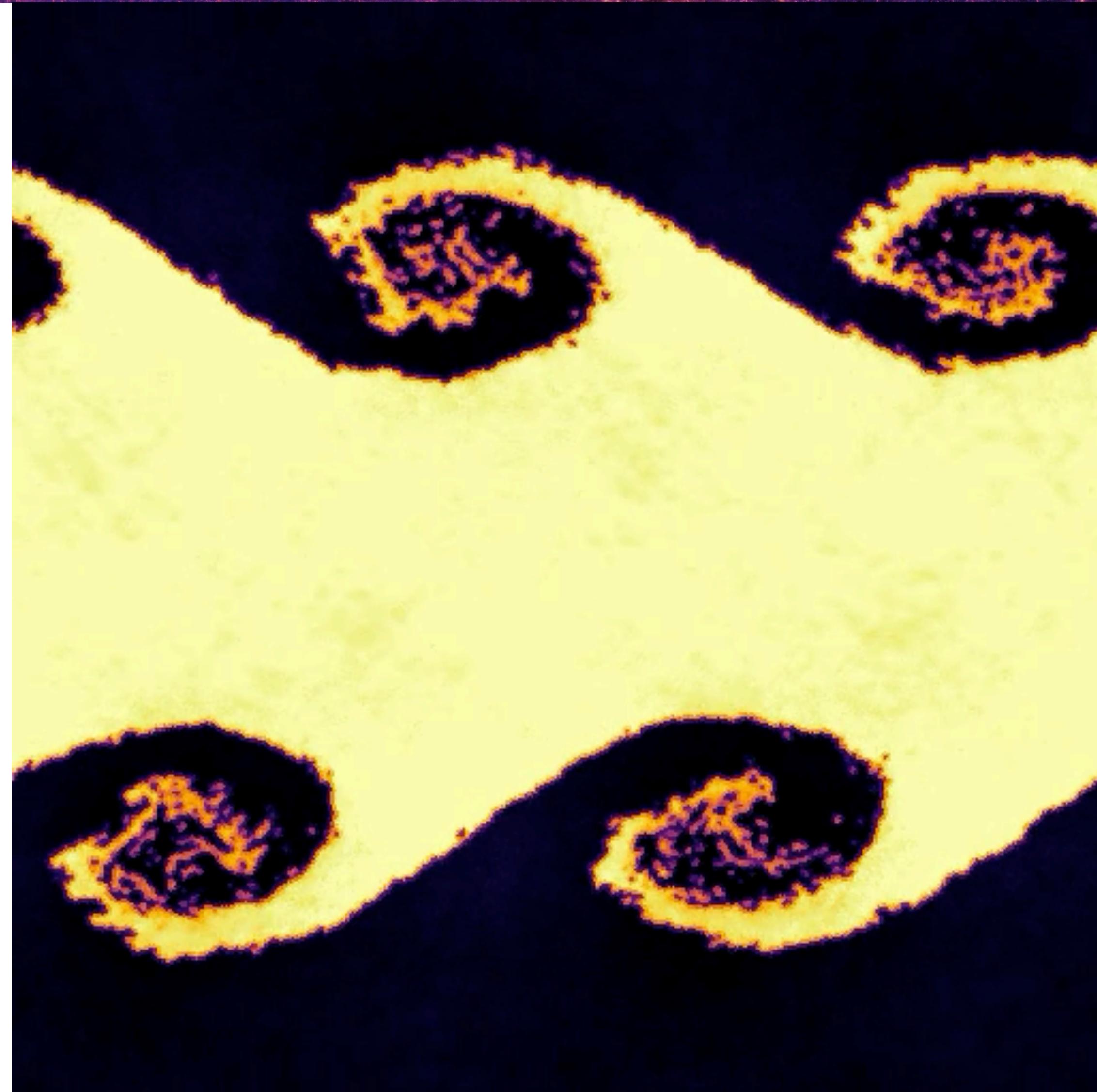
+



$$m_i \frac{d\mathbf{v}_i}{dt} = \sum_{j=1}^N P_j \nabla_i \Delta \tilde{V}_j \psi_j + P_j \nabla_i \Delta V_j.$$

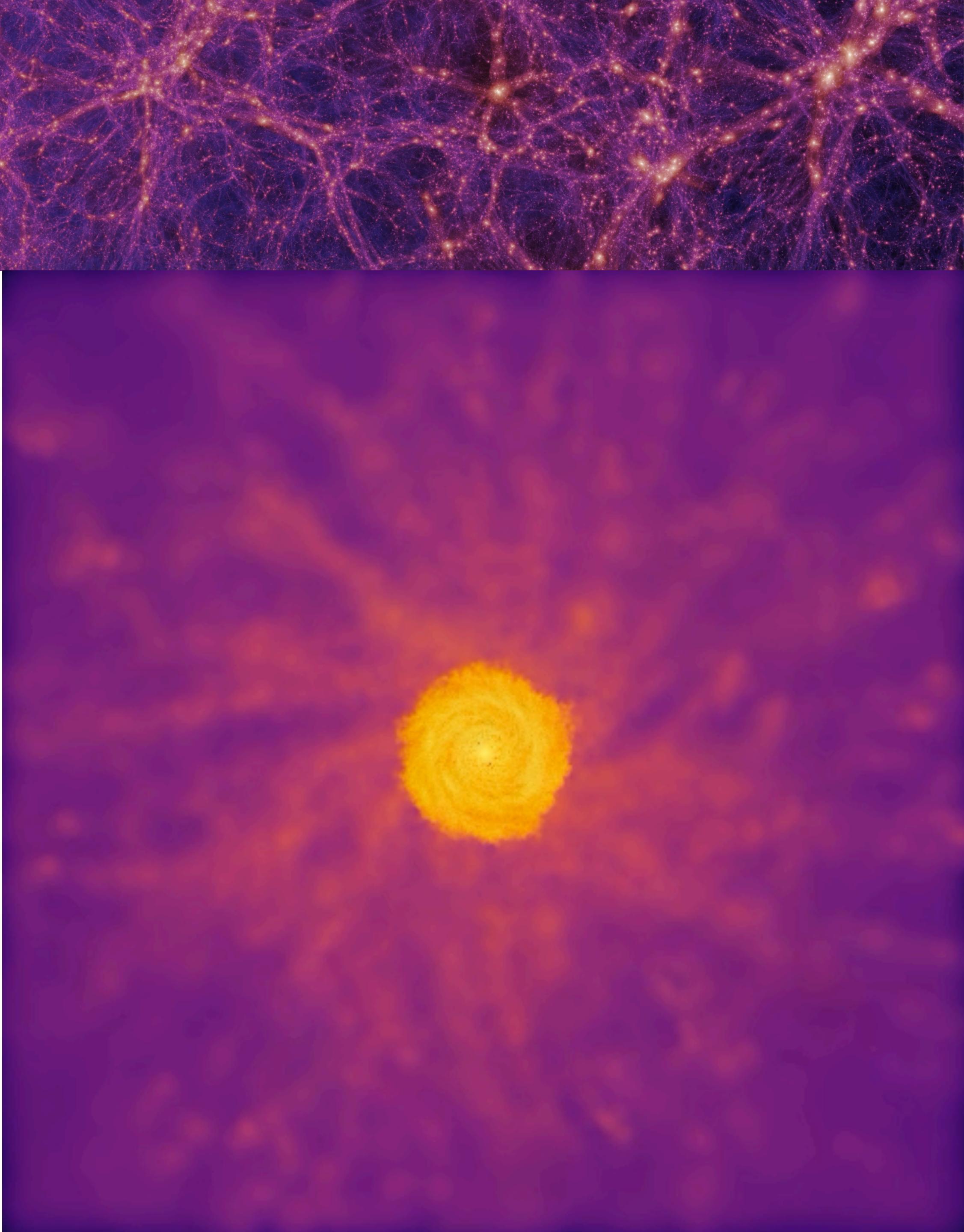


$$\frac{d\mathbf{v}_i}{dt} = - \sum_{j=1}^N x_i x_j \left[ \frac{f_{ij} P_i}{y_i^2} \nabla_i W_{ij}(h_i) + \frac{f_{ji} P_j}{y_j^2} \nabla_i W_{ji}(h_j) \right]$$



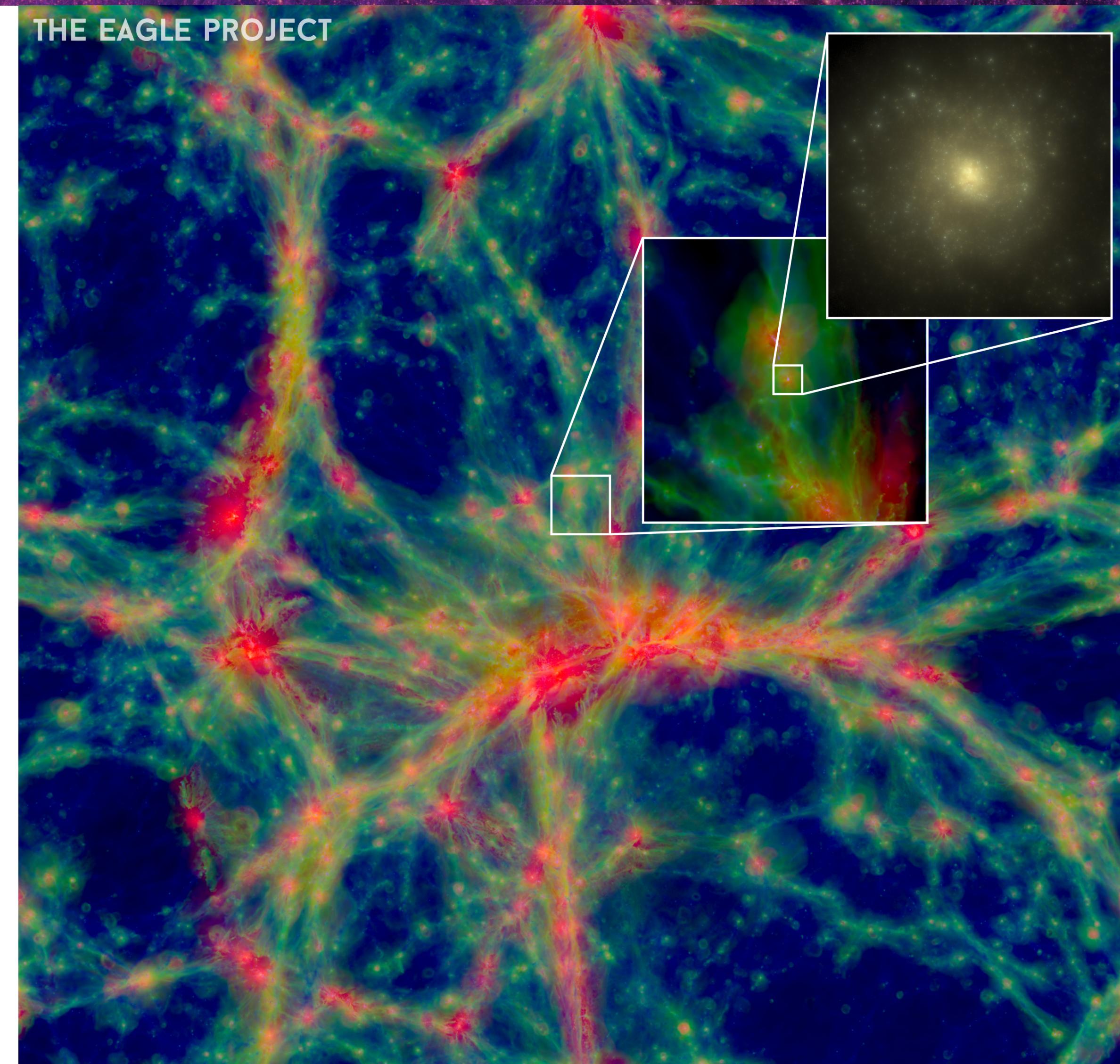
# Sub-Grid Physics

- We cannot resolve everything
- Typical resolution is about  $10^6$  solar masses
- Lots of processes occur at smaller scales:
  - Gas cooling (atomic transitions...)
  - Star formation
  - Black hole accretion



# Galaxy Formation Models

- This 'sub-grid' physics is directly related to the SAMs we discussed before
- They contain similar information, but the full models can be much simpler



# Implementation

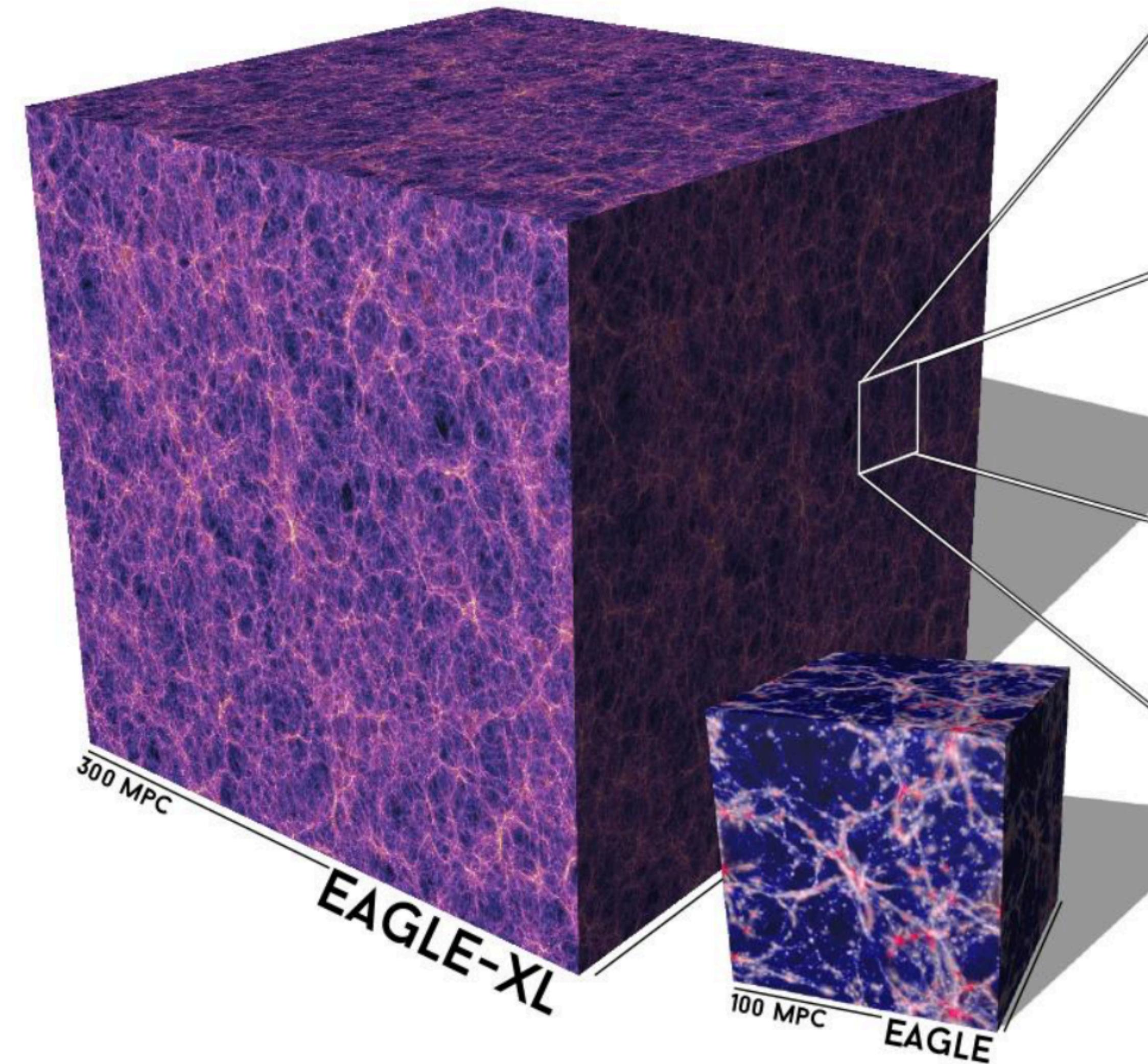
- 100'000 lines of open-source C99 code
- Hybrid (MPI + threads)
- Collaboration between astronomy, computer science, and industry
- Supported by DiRAC

The screenshot shows a GitLab project page for 'SWIFTsim'. The header includes the GitLab logo, navigation links for Projects, Groups, More, and This project, along with a search bar and various status indicators. The main content area displays the project details, including its logo (a stylized bird), name ('SWIFTsim'), description ('SPH With Inter-dependent Fine-grained Tasking'), and statistics: Files (63.4 MB), Commits (8,743), Branches (76), Tags (10), Readme, Changelog, and GNU GPLv3. Below this is a 'Contribution guide' section. A commit history is shown, with one recent commit from Matthieu Schaller: 'Removed the top-level FFT task from the space structure. This is not used any more.' (commit d6fa0c3a). The bottom part of the screenshot shows a table of repository branches: master, swiftsim, and several others. The table includes columns for Name, Last commit, and Last update.

Name	Last commit	Last update
argparse	Free the string allocated by strdup whe...	1 month ago
doc	Documented some of the basic meta-d...	1 day ago
examples	Merge branch 'vрascalupdates' into 'm...	8 hours ago
m4	Added more CPUID strings to the ax_g...	3 months ago

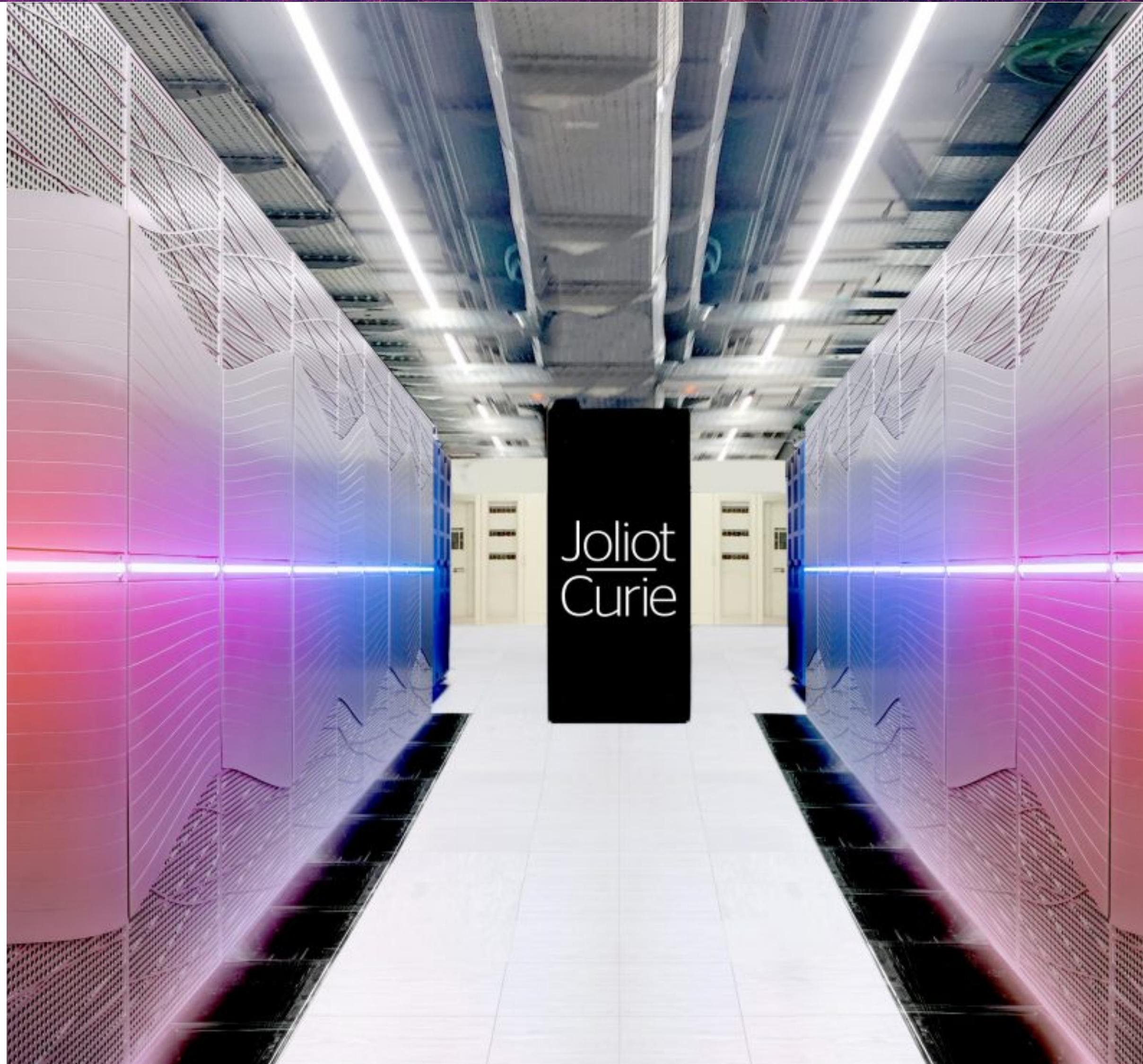
# How big?

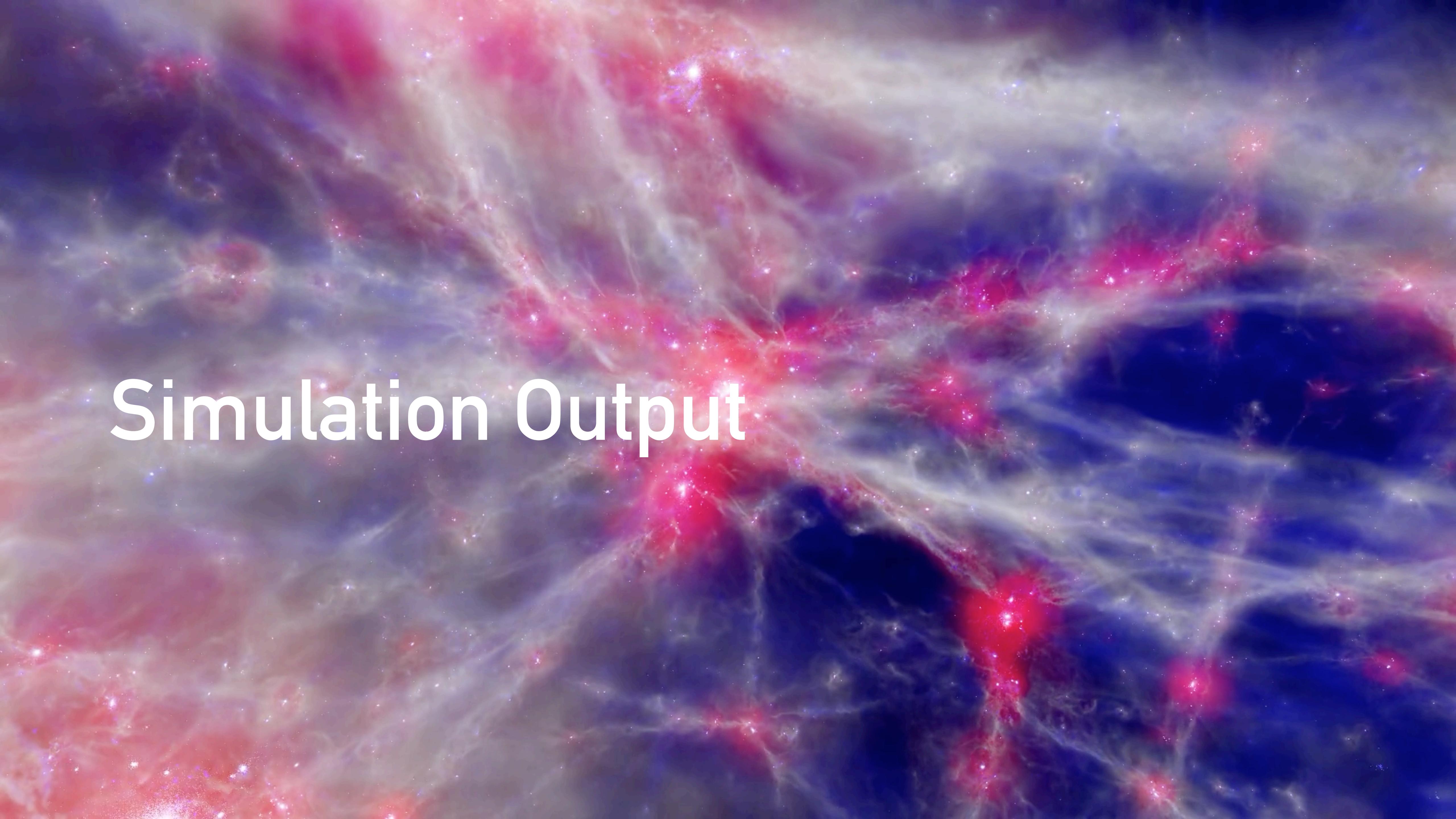
- Our next big run:
- Nearly **100 billion particles**
- 30'000 galaxies with a mass similar to our own Milky Way
- Each galaxy resolved by 10'000 - 100'000 particles



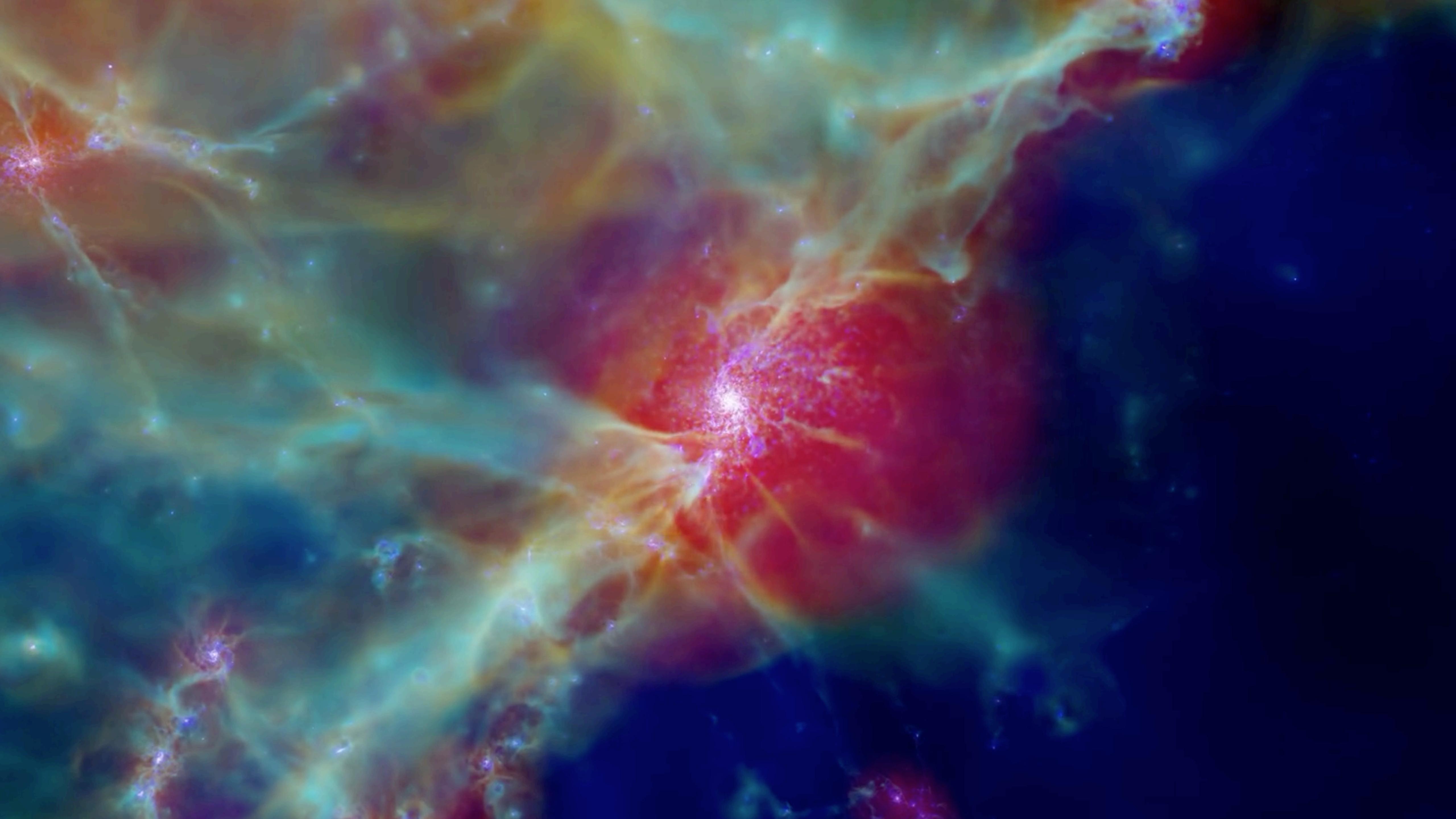
# Computing Resources

- Need Tier-0 resources
- 10s of millions of core hours
- 10'000s of CPUs
- ~200-300 TB of RAM
- Snapshots ~4-5 TB each



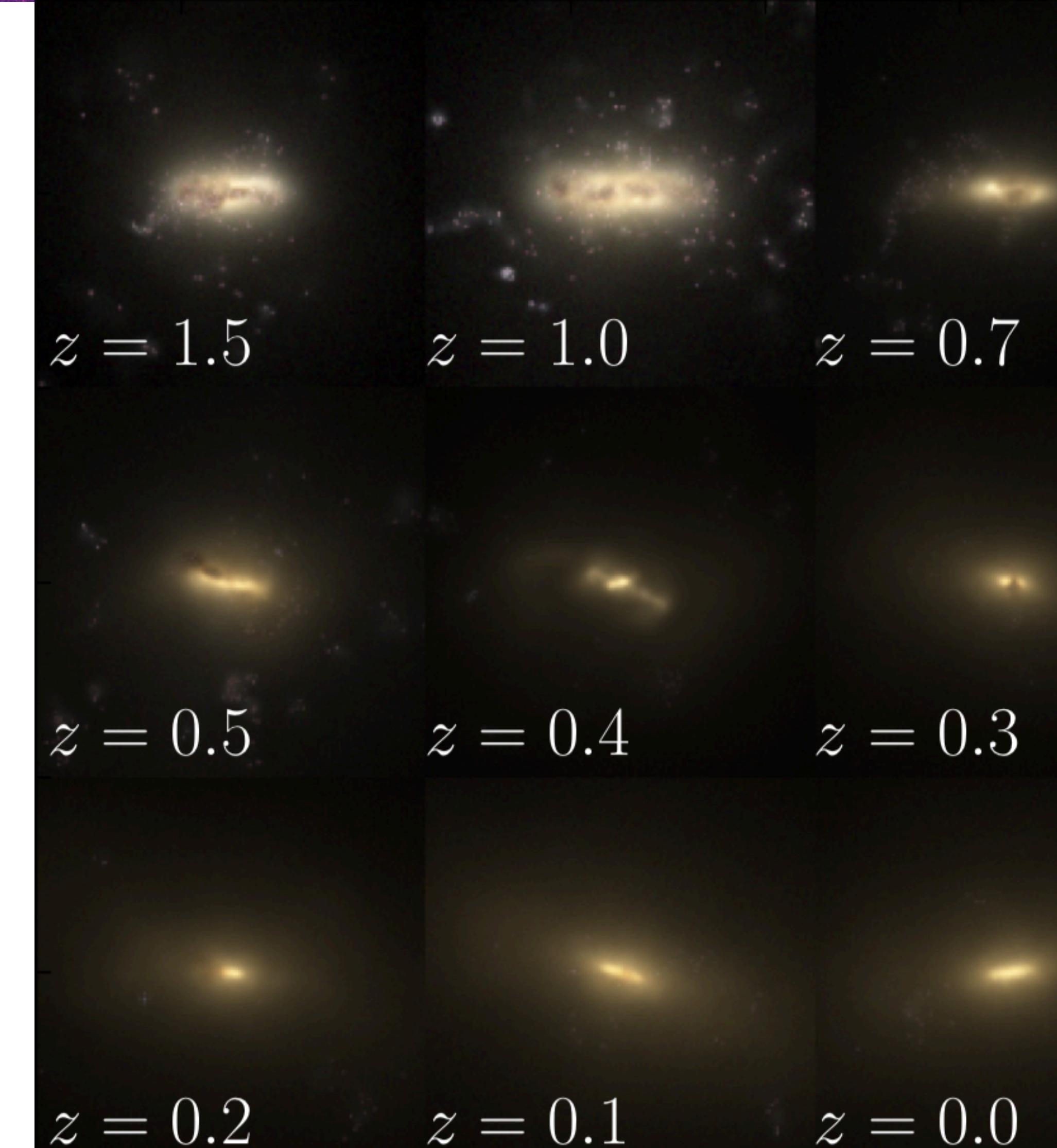
The background of the image is a complex, multi-colored simulation output. It features intricate, wispy filaments of light in shades of red, orange, yellow, green, blue, and purple, set against a dark, star-filled background. The colors are most intense in the central and lower-right regions, creating a sense of depth and motion.

Simulation Output



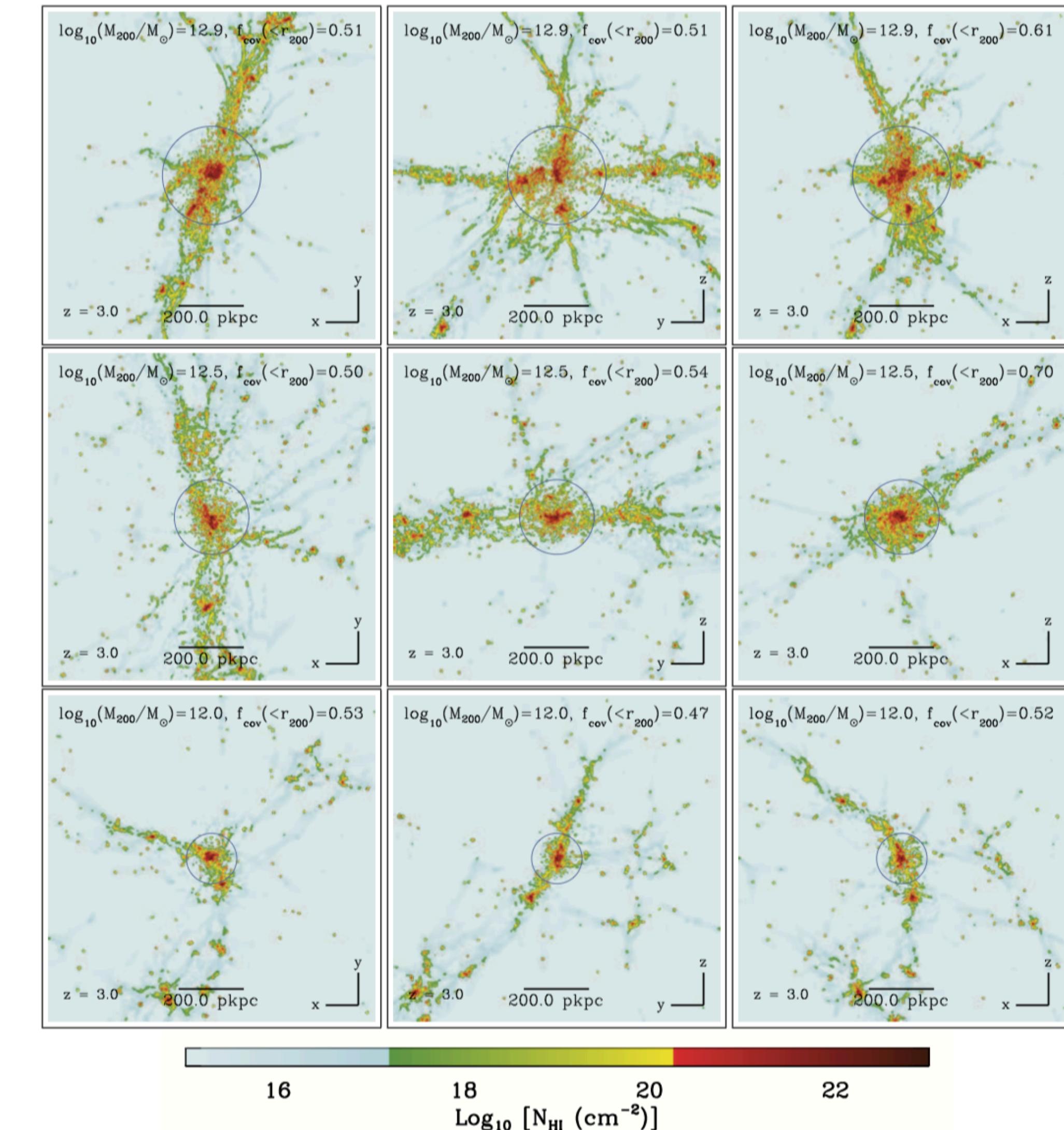
# Seeing Galaxies Grow and Change

- We can see the emergence of the “Tuning Fork” by following *individual* galaxies through their evolution
- Turns out that it goes backwards, with Ir -> Spiral -> Elliptical.



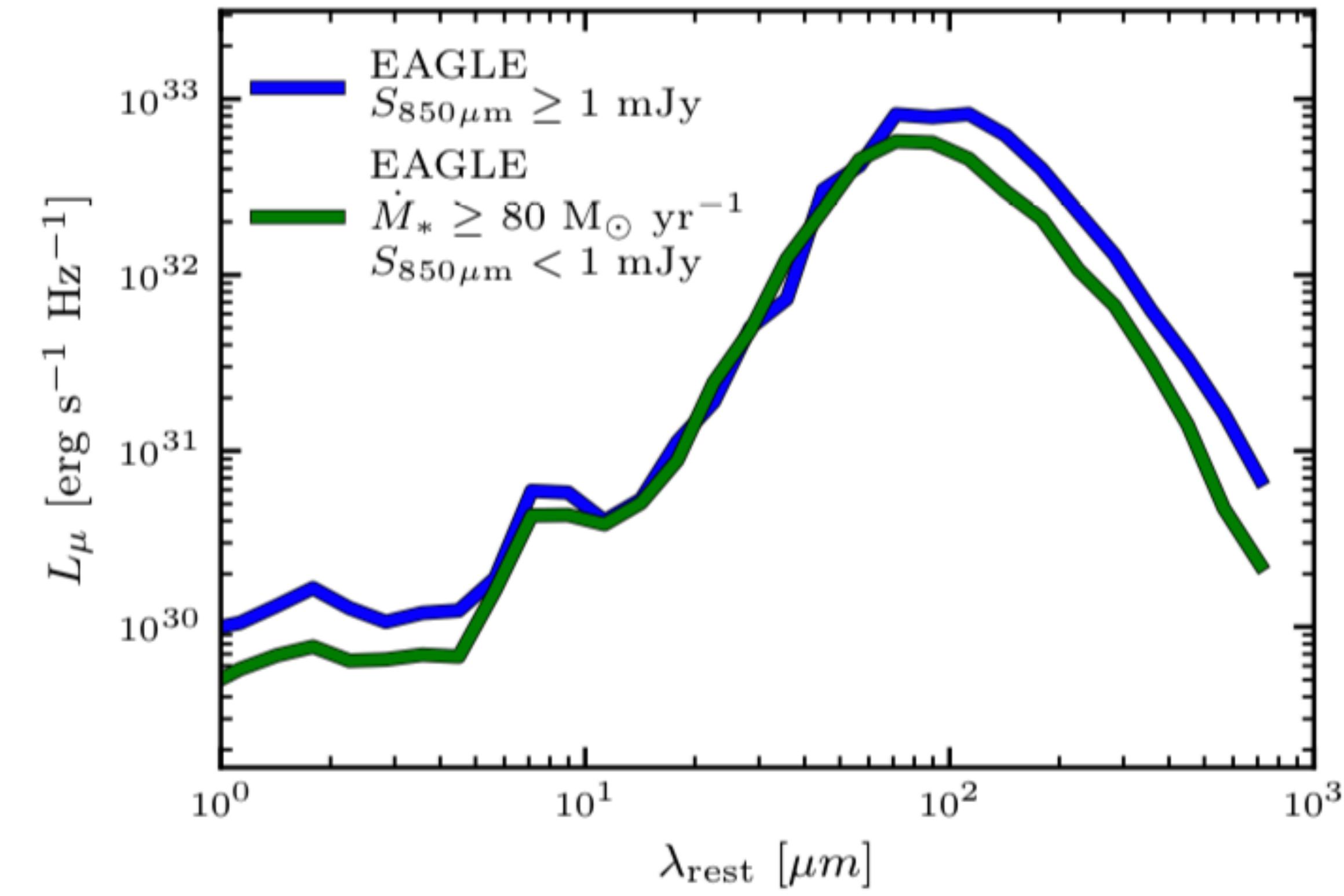
# Effects on Large-Scale Structure

- Can see how the gas physics changes the morphology of the large-scale structure
- This feeds back into the calibrations (e.g. dust absorption) that we make when observing distant objects.



# Huge Predictive Power

- Cosmological simulations have huge predictive power
- They can be used to tell observers where to point their telescopes for the best chance of finding rare objects
- We can also compare abundances of weird objects to see if some are missing, and figure out why.



**Figure 9.** Composite broadband spectral energy distributions (SEDs) for the  $S_{850\mu\text{m}} \geq 1 \text{ mJy}$  and highly star-forming Submm-Faint galaxies. The leftward shifting of the peak in the dust emission for the highly star-forming Submm-Faint sample signifies hotter dust temperatures than the  $S_{850\mu\text{m}} \geq 1 \text{ mJy}$  galaxies.

# Summary

- Large timescales make cosmology and astrophysics really challenging
- We can address some of this by using cosmological simulations
- To perform a useful calculation, you need some of the largest supercomputers in the world
- We can make pretty movies...