The Universe in a Computer

Nishikanta Khandai, NISER

Physics of Galaxy Formation, IUCAA

June 8-24, 2026

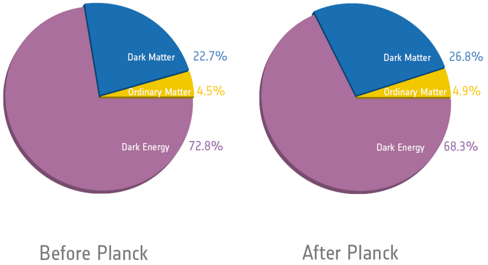
With Ranit Behera (IUCAA)

Plan of the Talk

- Background: The N-Body Problem of Self-Gravitating Systems
- Exact Methods $\mathcal{O}(N^2)$: Direct Summation approach
- Approximate Methods in Fourier Space $\mathcal{O}(N \log N)$: Grid Based Methods
- Approximate Methods in Real Space $\mathcal{O}(N \log N)$: Multipole Methods
- Hybrid Methods: P^3M, TPM, TreePM
- Other Issues: Individual Timesteps, Parallelisation

The N-body Problem of Self-Gravitating Systems

- Gravity is the dominant force on large scales.
- Matter in the Universe is dominated by Cold Dark Matter $\sim 80\%$



Credit: ESA and the Planck Collaboration.

- For this lecture we will look at Classical Newtonian gravity.

The N-body Problem of Self-Gravitating Systems

- Determines the dynamical evolution of small-N systems (solar system, star clusters) to large-N systems (globular clusters, galaxies, clusters of galaxies and the large scale structure of the Universe.

Figure: Messier 80

The N-body Problem of Self-Gravitating Systems

Figure: Colliding Galaxies: NGC2207 (left) and IC 2163 (Right)

The N-body Problem of Self-Gravitating Systems

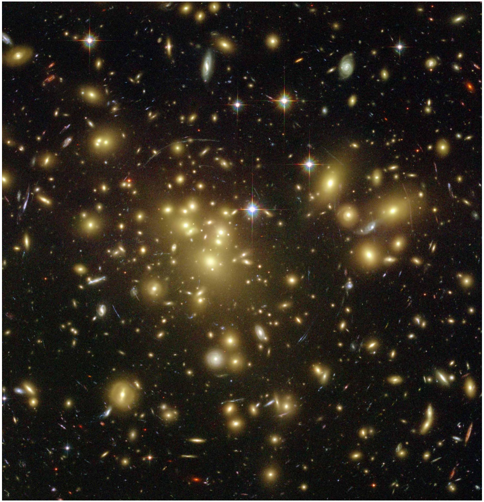
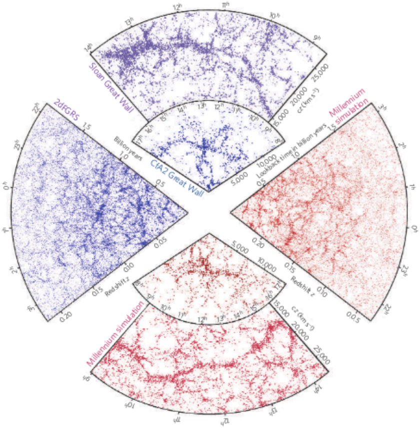


Figure: Abell 1689

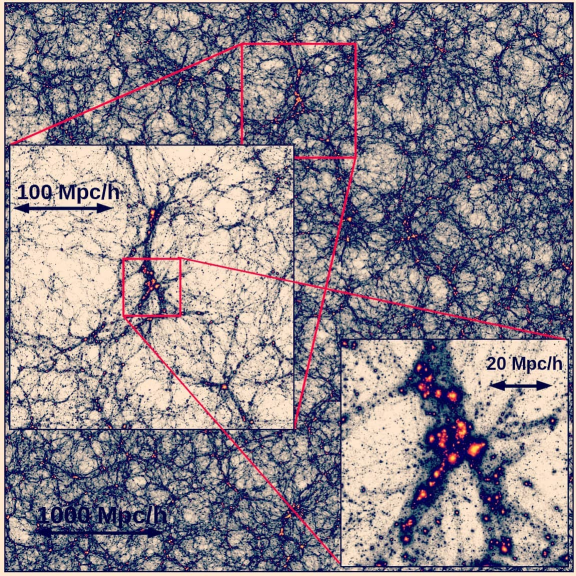
Structure Formation

Structure Formation: How can we explain the distribution and properties of objects that we see in the Universe?



Springel et al, 2004 2dF and SDSS survey

Dark Matter Simulations: SMDPL2



Rana et al 2021

Beyond Structure Formation: Galaxy Formation and Cosmology

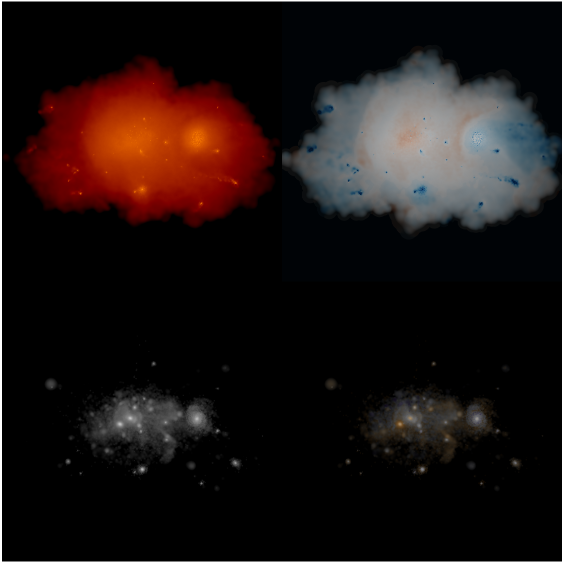
- Can we predict the number of spirals and elliptical galaxies?
- Can we predict the SEDs of galaxies and how they evolve across cosmic time?
- What are the important mechanisms affecting the formation and and evolution of galaxies?
- How do galaxies obtain gas and convert it to stars?
- What is the relation between galaxies and their host dark matter halos?

- Observational Cosmology is currently putting constraints at the percent level.
- Limitations are due to poor understanding of galaxy formation.
- . . .

Beyond Structure Formation: Galaxy Formation and Evolution

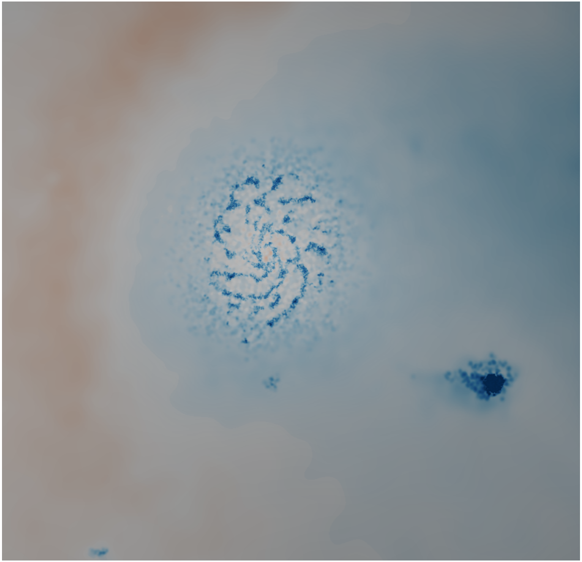
- Quasar/Galaxy LF and Clustering
- Galaxy Morphology, Photometry
- Galaxy Formation with Patchy Reionisation
- IGM, CGM, metals in gas.
- High Redshift DLAs (contribution of IGM)
- Low redshift CGM, IGM
- Molecular and atomic gas in galaxies.
- GW from merging SMBHs (SMBH flyby). Host galaxy properties.
- Atomic and molecular hydrogen in galaxies.
- Recycling of gas.
- SMBH accretion histories and constraints on the SMBH mass function.
- ...

Low Redshift Results: Galaxy Images (mass conserving)



Roy et al 2026, SPHERE

Low Redshift Results: Galaxy Images (mass conserving)



Roy et al 2026, SPHERE

The N-body Problem of Self-Gravitating Systems

- Statement of the Problem:
  Given N objects interacting gravitationally with masses m_i, initial positions and velocities determine their trajectories at any later time δt?

$$
\begin{aligned}
[\mathbf{r}_i(t), \mathbf{v}_i(t)] &\Rightarrow [\mathbf{r}_i(t+\delta t), \mathbf{v}_i(t+\delta t)] && (1)\\
\left[\mathbf{r}_i^0, \mathbf{v}_i^0\right] &\Rightarrow \left[\mathbf{r}_i^1, \mathbf{v}_i^1\right] && (2)
\end{aligned}
$$

- For $N = 2$: Analytical Solutions
- For $N = 3$ there are interesting problems e.g. Earth-Satellite-Moon, Pythagorean problem.
- For $N > 2$: no analytical solutions, need to find numerical solutions.

Collisional and Collisionless Systems: (See Binney and Tremaine)

- Relaxation time

$$t_{relax} = n_{relax} t_{cross} \sim \frac{0.1N}{\ln N} t_{cross} \tag{3}$$

  $n_{relax} \approx \frac{v^2}{\Delta v^2} \rightarrow$ the number of crossing for velocity to change by order itself.

- Typical Numbers
  - Globular Cluster
  - $N \approx 10^5 \quad t_{cross} = 10^6 yr \Rightarrow t_{relax} \approx 10^{10} yr \sim t_{Age}$
  - 2-body relaxation is important. Collisional system.

  - Galaxy
  - $N \approx 10^{11} \quad t_{cross} = 10^8 yr \Rightarrow t_{relax} \approx 10^{16} yr > t_{Age}$
  - 2-body relaxation is not important. Collisionless system.

The N-body Problem of Self-Gravitating Systems

- Simple Numerical Solution:

$$\mathbf{r}_i^1 = \mathbf{r}_i^0 + \mathbf{v}_i^0 \delta t + \mathbf{a}_i^0 \frac{\delta t^2}{2} + \mathcal{O}(\delta t^3) \tag{4}$$

$$\mathbf{v}_i^1 = \mathbf{v}_i^0 + \mathbf{a}_i^0 \delta t + \mathbf{j}_i^0 \frac{\delta t^2}{2} + \mathcal{O}(\delta t^3) \quad \mathcal{O}(N) \tag{5}$$

$$\mathbf{a}_i^0 = \ddot{\mathbf{r}}_i^0 = -G \sum_{j=1, j \neq i}^{N} \frac{m_j (\mathbf{r}_i^0 - \mathbf{r}_j^0)}{\left| \mathbf{r}_i^0 - \mathbf{r}_j^0 \right|^3} \quad \mathcal{O}(N^2) \tag{6}$$

With $\mathbf{j}_i^0 = \dot{\mathbf{a}}_i^0 = \frac{\dot{\mathbf{a}}_i^1 - \dot{\mathbf{a}}_i^0}{\delta t} \Rightarrow$ Jerk term

The N-body Problem of Self-Gravitating Systems

- The Leap-Frog Integrator

$$\mathbf{r}_i^1 = \mathbf{r}_i^0 + \mathbf{v}_i^{1/2}\delta t + \mathcal{O}(\delta t^3) \tag{7}$$

$$\mathbf{v}_i^{1/2} = \mathbf{v}_i^{-1/2} + \mathbf{a}_i^0\delta t + \mathcal{O}(\delta t^3) \tag{8}$$

- Time Symmetric Integrator: Energy Conservation

Direct Summation

- Direct Summation: (Pros)
  - Easy to implement.
  - The intergrator is time symmetric $\Rightarrow$ Energy conserving.
  - Can tune accuracy by reducing δt
  - Accuracy a must for Collisional System
- Direct Summation: (Cons)
  - Expensive for $N > 10^6$ particles
  - Cannot study galaxies and LSS in general with this approach
  - Need for expensive convergence of orbits tests since system is collisional.
- Direct Summation: (Way out)
  - Two and three body regularisation: NBODY code by Aarseth
  - Buy specialised hardware like GRAPE to speed up calculation
  - Buy GPUs to speed up calculation

Particle Dynamics in an Expanding Universe

EOM Physical Coordinates

$$
\begin{aligned}
\ddot{\mathbf{r}}(t) &= \ddot{a}(t)\mathbf{x}(t) + 2\dot{a}(t)\dot{\mathbf{x}}(t) + a\ddot{\mathbf{x}}(t) = -\nabla\Phi_{tot} = -\nabla(\Phi_{av} + \phi) \\
\nabla_r^2 \Phi_{tot} &= \nabla_r^2(\Phi_{av} + \phi) = 4\pi G\rho - \Lambda = (4\pi G\bar{\rho} - \Lambda) + 4\pi G\bar{\rho}\delta \\
\rho(\mathbf{r}) &= \sum_i m_i \delta_D^3(\mathbf{r} - \mathbf{r}_i) \qquad \delta(\mathbf{r}, t) = \frac{\rho(\mathbf{r}, t) - \bar{\rho}(t)}{\bar{\rho}(t)} \qquad (9)
\end{aligned}
$$

EOM Comoving Coordinates: $\mathbf{r} = a(t)\mathbf{x}$

$$
\begin{aligned}
\ddot{\mathbf{x}} + 2\frac{\dot{a}}{a}\dot{\mathbf{x}} &= -\frac{\nabla_r \phi(\mathbf{x}, t)}{a} = -\frac{\nabla_x \phi(\mathbf{x}, t)}{a^2} \\
\nabla_x^2 \phi &= 4\pi G a^2 \bar{\rho}\delta = \frac{3}{2} H_0^2 \Omega_{nr} \frac{\delta}{a} \qquad \rho = \bar{\rho}(1 + \delta) \\
\rho(\mathbf{x}) &= \sum_i m_i \delta_D^3(\mathbf{x} - \mathbf{x}_i) \approx \sum_i m_i W(\mathbf{x} - \mathbf{x}_i, \epsilon) \qquad (10)
\end{aligned}
$$

Friedmann Equation:

$$
\frac{\dot{a}^2}{a^2} + \frac{k}{a^2} = H_0^2 \left[\Omega_{rad} \left(\frac{a_0}{a} \right)^4 + \Omega_{nr} \left(\frac{a_0}{a} \right)^3 + \Omega_\Lambda \right] \qquad (11)
$$

Basic Equations

The Linear Limit

$$\ddot{\delta} + 2\frac{\dot{a}}{a}\dot{\delta} = 4\pi G\bar{\rho}_{nr}\delta \tag{12}$$

$$\delta_{\mathbf{x}}(t) = \int \frac{d^3\mathbf{k}}{(2\pi)^3}\delta_{\mathbf{k}}(t)\exp(i\mathbf{k}.\mathbf{x}) \tag{13}$$

$$\delta_{\mathbf{k}}(t) = \int \frac{d^3\mathbf{x}}{V}\delta_{\mathbf{x}}(t)\exp(-i\mathbf{k}.\mathbf{x}) \tag{14}$$

$$\ddot{\delta}_{\mathbf{k}} + 2\frac{\dot{a}}{a}\dot{\delta}_{\mathbf{k}} = 4\pi G\bar{\rho}_{nr}\delta_{\mathbf{k}} \tag{15}$$

Hubble's parameter $H(a) = \dot{a}/a$ is a solution (decaying).
Second solution can be constructed using the Wronskian.

$$\begin{aligned} X &= 1 + \Omega_{nr}\left(\frac{1}{a} - 1\right) + \Omega_{\Lambda}(a^2 - 1) \\ b(t) &\propto \frac{X^{1/2}}{a}\int^{a}\frac{da}{X^{3/2}} \end{aligned} \tag{16}$$

Exercise

$$
\begin{aligned}
\frac{d\mathbf{u}}{db} &= -\frac{3}{2}\frac{Q}{b}(\mathbf{u}-\mathbf{g}) \\
\nabla^2\psi &= \frac{\delta}{b}
\end{aligned}
\tag{17}
$$

where

$$
\begin{aligned}
\mathbf{g} &\equiv -\nabla\psi = -\frac{2}{3H_0^2\Omega_{nr}}\left(\frac{a}{b}\right)\nabla\phi \\
Q &= \left(\frac{\bar{\rho}_{nr}}{\rho_c}\right)\left(\frac{\dot{a}b}{a\dot{b}}\right)^2 = \left(\frac{\bar{\rho}_{nr}}{\rho_c}\right)\frac{1}{f^2}
\end{aligned}
\tag{18}
$$

- $\mathbf{u} \equiv d\mathbf{x}/db \Rightarrow$ generalised velocity,
- $\psi \Rightarrow$ generalised potential
- $\mathbf{g} \Rightarrow$ the generalised force,
- $f = d\ln(b)/d\ln(a) \Rightarrow$ Growth Factor

Initial Conditions

- Linear Regime: $d\mathbf{u}/db = 0$
- Early Times: $\mathbf{u} = \mathbf{g}$
- Einstein deSitter Universe, $a = b$ and $Q = 1$. The EOM reduces to:

$$\frac{d\mathbf{u}}{da} = -\frac{3}{2a}(\mathbf{u} - \mathbf{g}) \tag{19}$$

$$\mathbf{u}(\mathbf{x}, b) = \mathbf{u}(\mathbf{x}) = \mathbf{g}(\mathbf{x}, b) = \mathbf{g}(\mathbf{x}) \tag{20}$$

In the Zeldovich approximation velocities are constant in Lagrangian space $\Rightarrow$ initial velocity.

$$\mathbf{u}(\mathbf{x}, b) = \mathbf{u}(\mathbf{q}) = \mathbf{g}(\mathbf{q}) \tag{21}$$

$$\mathbf{x} = \mathbf{q} + b\mathbf{g}(\mathbf{q}) \tag{22}$$

The negative time variable θ defined as:

$$d\theta = \frac{H_0 dt}{a^2} \tag{23}$$

transforms the EOM

$$\begin{aligned} \ddot{\tilde{x}} &= -\frac{3}{2}\Omega_{nr} a \nabla_x \chi \\ \nabla_x^2 \chi &= \delta \end{aligned} \tag{24}$$

$\ddot{\tilde{x}} \equiv \frac{d^2x}{d\theta^2}$
$\phi = \frac{3}{2}\frac{\Omega_{nr} H_0^2}{a}\chi$.

Approximate N-Body techniques: Grid Based Methods

- EOM is simple to integrate $\Rightarrow \mathcal{O}(N)$.
- Time consuming part is the force calculation due to gravitational interaction of bodies $\Rightarrow \mathcal{O}(N^2)$.
- $\mathcal{O}(N^2) \Rightarrow$ due to pairwise interaction of gravity and its long-range nature.
- N-Body methods: resort to quicker and approximate techniques to compute the gravitational interactions between particles.
- Also need to soften forces since systems are collisionless.

- Broadly 2 methods to compute interactions (third hybrid one)

- Important: need to also assign size to particles, since we are dealing with a collisionless system.

Eulerian or Grid Based methods (PM)

- Method One: solve the Poisson Equation to get the potential (hence force):

$$\nabla^2\phi(\mathbf{r}) = 4\pi G\rho(\mathbf{r}) \tag{25}$$

$$-k^2\phi_k \propto \rho_k \qquad \text{In Fourier space} \tag{26}$$

Use the $\mathcal{O}(N \log N)$ Fast Fourier Transform to do the forward and inverse transform.

PM Method Continued

- Pros
  - $\mathcal{O}(N \log N) \Rightarrow$ extremely fast.
  - Simple to implement, algebraic equations.
  - Easy to Parallelise with MPI (not OpenMP)
  - Periodic Boundary Conditions from FFT.

- Cons
  - Resolution limit $\Rightarrow$ grid size.
  - Parallisation limited to slab parallisation in FFTW.

- Remedy
  - Use PencilFFT for further parallisation (See Blue Waters Simulation by Feng et al)
  - Adaptive meshes which adjust to local density, AMR
  - ART (Kravtsov, Klypin, Kokhlov), MLAPM (Knebe et al), Nyx (Almgren et al), Enzo (Bryan, Norman et al)

Lagrangian or Particle Based methods (BH-Tree)

- Compute direct forces in real space.

- Approximation: Structure the entire mass distribution into groups nested into larger groups $\Rightarrow$ Tree.
- Can keep the monopole term or higher order multipoles (expensive)
- Pros:
  - Scales as $\mathcal{O}(N \log N)$
  - Extremely Accurate
- Cons:
  - Slower than PM by ~ 100
  - Consumes more memory
  - Need to add terms due to periodic boundary conditions.
- Remedy:
  - Combine with PM to get both speed and accuracy, TPM: Bode & Ostriker TreePM:Bagla 2002

BH-Tree

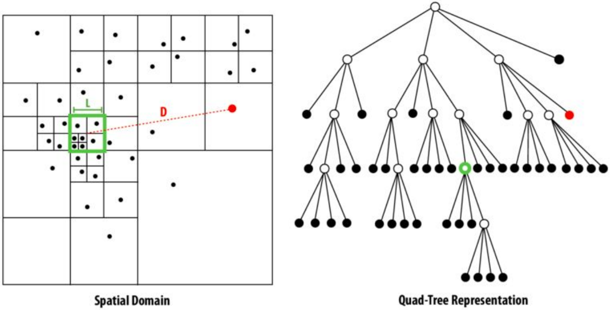


Figure: BH-Quad Tree: Credits cs.cmu.edu

- Cell Opening Criterion $\theta < L/D$, Choice of $\theta < 1.0$
- Other Multipole Methods (not covered) Fast Multipole Method (memory intensive)

Hybrid Methods: P^3M Hockney and Eastwood

- PM method has a limiting resolution L_{grid}
- In LSS clustering increases with time.
- $\sim$ 50% of the mass is in collapsed regions (subgrid level)
- Remedy: Use direct summation for particles in cells with $\delta_{cell} > \delta_{crit} \sim \mathcal{O}(10)$
- Add the tidal forces from particles outside the cell (PM method)
- For particles $\delta_{cell} < \delta_{crit}$ use PM method.
- Problem: In strong clustering regime will still be time consuming (P-P part)
- Interpolation errors from PM.

Hybrid Methods: *TPM* Bode & Ostriker

- In principle similar to P^3M
- Use BHTree for forces for particles in cells with $\delta_{cell} > \delta_{crit} \sim \mathcal{O}(10)$
- Add the tidal forces from particles outside the cell (PM method)
- For particles $\delta_{cell} < \delta_{crit}$ use PM method.
- PM is insensitive to clustering.
- Interpolation errors from PM still persist.
- All particles do not have the same errors in force.

TreePM Method: Bagla

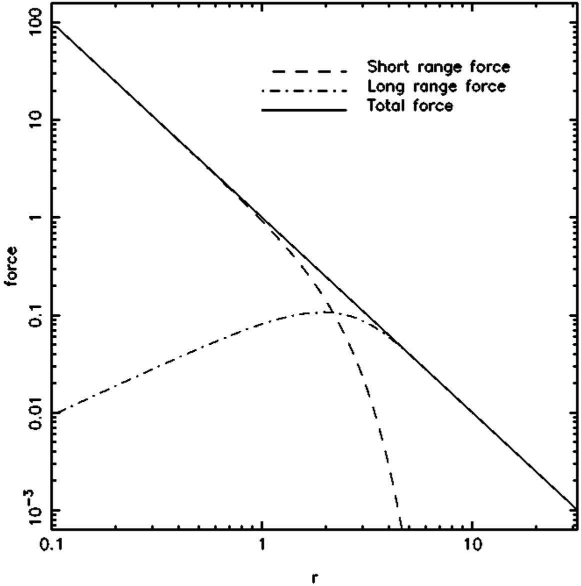


Figure: TreePM: Bagla 1999

TreePM Method: Bagla

- The inverse square force is divided into a long range and a short range force

$$\phi_{\mathbf{k}} = -4\pi G \frac{\rho_{\mathbf{k}}}{k^2} = \phi^l_{\mathbf{k}} + \phi^s_{\mathbf{k}} \tag{29}$$

$$\phi^l_{\mathbf{k}} = -4\pi G \frac{\rho_{\mathbf{k}}}{k^2} \exp(-k^2 r_s^2) \tag{30}$$

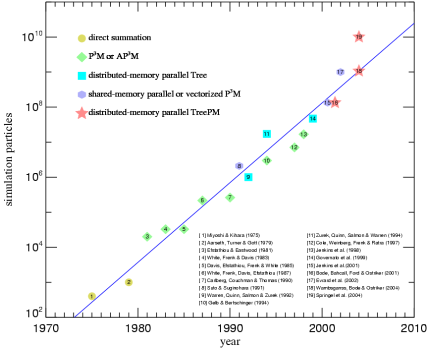
$$\phi^s_{\mathbf{k}} = -4\pi G \frac{\rho_{\mathbf{k}}}{k^2}(1 - \exp(-k^2 r_s^2)) \tag{31}$$

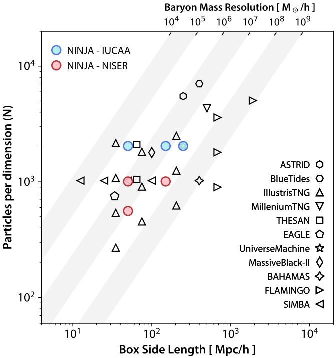
$$\mathbf{f}^s_{\mathbf{k}}(\mathbf{r}) = -\frac{Gm\mathbf{r}}{r^3}\left(erfc\left(\frac{r}{2r_s}\right) + \frac{r}{r_s\sqrt{\pi}} \exp\left(-\frac{r^2}{4r_s^2}\right)\right) \tag{32}$$

TreePM Method: Bagla

- Interpolation errors in the PMForce is exponentially suppressed.
- Scales as $\mathcal{O}(N \ln N)$
- Is currently adopted in Gadget (Springel).
- Some of the largest cosmological simulations employ TreePM.

Simulations over the Years





Credit: *Left:* Volker Springel et al. *Right* Ranit Behera et al

Hierarchical Timesteps

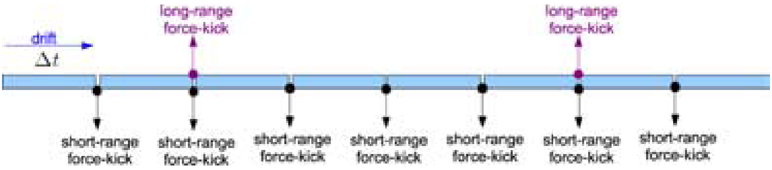


Figure: Credit: Volker Springel

$$U(\Delta t) = K_{lr}\left(\frac{\Delta t}{2}\right)\left[K_{sr}\left(\frac{\Delta t}{2m}\right) D\left(\frac{\Delta t}{m}\right) K_{sr}\left(\frac{\Delta t}{2m}\right)\right]^{m} K_{lr}\left(\frac{\Delta t}{2}\right) \quad (33)$$

Domain Decompostion

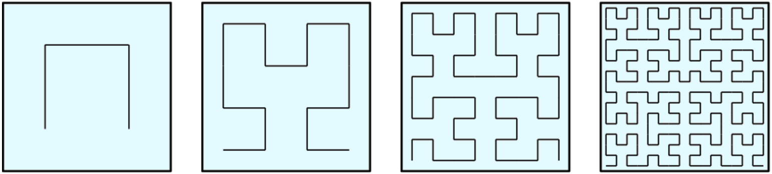


Figure: 1D Peano Hilbert Curve. Credit: Volker Springel

Domain Decompostion

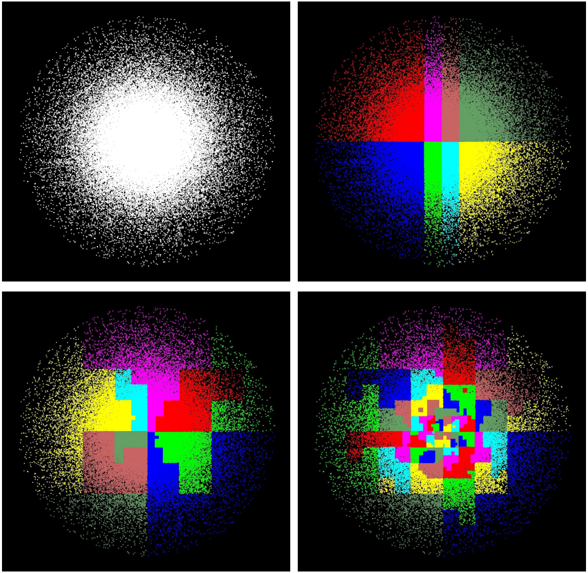


Figure: Credit: Volker Springel

Steps to Run a Cosmological Simulation (gravity)

- Choose your favourite Cosmology: Ω_x, $P(k)$
- Set up Initial Conditions
  - Gaussian Random Density Fields with $P(k) \Rightarrow \delta_{\mathbf{k}}(a_i)$
  - Zeldovich Approximation: $\Rightarrow \mathbf{v}_{\mathbf{k}}(a_i)$
- Run the Simulation (e.g. Gadget)
- Analyse