Cosmological N-body simulation: Techniques, scope and status

J. S. Bagla

Harish-Chandra Research Institute, Chhatnag Road, Jhunsi, Allahabad 211 019, India

Cosmological N-body simulations have become an essential tool for studying formation of large scale structure. These simulations are computationally challenging even though the available computing power gets better every year. A number of efficient algorithms have been developed to run large simulations with better dynamic range and resolution. We discuss key algorithms in this review, focusing on techniques used and their efficacy. N-body simulations solve a model that is an approximation of the physical model to be simulated, we discuss limitations arising from this approximation and techniques employed for solving equations. Apart from simulating models of structure formation, N-body simulations have also been used to study aspects of gravitational clustering. Simulating formation of galaxies requires us to take many physical process into account; we review numerical implementations of key processes.

I. Introduction

LARGE scale structures like galaxies and clusters of galaxies are believed to have formed by amplification of small perturbations[1–9]. Galaxies are highly overdense systems, matter density ρ in galaxies is thousands of times larger than the average density $\bar{\rho}$ in the universe. Typical density contrast ($\delta \equiv \rho/\bar{\rho} - 1$) in matter at these scales in the early universe, e.g. at the time of decoupling of matter and radiation was much smaller than unity. Thus the problem of galaxy formation and the large scale distribution of galaxies is essentially one of evolving density perturbations from small initial values to the large values we encounter today.

The universe is assumed to be homogeneous and isotropic at large scales and this assumption is consistent with observations and there are strong limits on departures from homogeneity and isotropy[10]. A homogeneous and isotropic universe is described by Friedman equations in the general theory of relativity. As long as perturbations in the gravitational potential are small, we can treat density fluctuations as perturbations about a Friedman universe. If perturbations are in nonrelativistic matter, as appears to be the case in our universe, we can work in the Newtonian limit for studying their evolution. The back reaction of perturbations on the universe is not taken into account, i.e. we do not worry about the effect of perturbations on the average global properties of the universe. Studies have shown that this back-reaction is ignorable in most cases[11,12]. Local variations in density, etc. can lead to dispersion in values of cosmological parameters determined through local observations[13] but this problem can be controlled by making measurements at larger scales.

Gravity is the dominant force at large scales and is believed to drive growth of perturbations. Magnetic field is the only other interaction that can lead to formation of large scale structures, this leads to very distinct signatures and does not appear to be the dominant factor in our universe[14–16]. We will assume that gravity is the only interaction responsible for growth of perturbations at large scales. Equations that describe the evolution of density perturbations in nonrelativistic matter due to gravitational interaction in an expanding universe have been known for a long time[17]. These equations can be solved analytically for small density contrasts, and for highly symmetric situations. But apart from such special cases, few solutions are known. Many approximate solutions are known[9,18–25] and are useful in understanding the evolution of perturbations in the quasilinear regime, these fail when density contrast becomes large ($\delta \gg 1$). Cosmological N-body simulations are an essential tool for evolving density perturbations in the nonlinear regime. Fluctuations in the gravitational potential do not grow by a large amount even as density contrast increases by several orders of magnitude[22,23], therefore the Newtonian approximation continues to be a valid framework. At galactic scales, gas dynamical and other effects play an important role and need to be taken into account for a detailed solution of the problem.

N-body simulations are used for a variety of applications. Simulations of specific models of dark matter allow us to make predictions for these models and compare with observations. Simulations allow us to carry out numerical experiments with initial conditions that have little to do with the real universe. The purpose of such experiments is to understand the physics of gravitational collapse in an expanding universe. Simulations are also used for testing approximate solutions for growth of density perturbations, comparisons with N-body simulations allow us to validate these approximations and understand when these approximations are useful. Lastly, we can calibrate methods for analysing observations on mock catalogues made from N-body simulations. We can test whether a particular method works or not because in N-body all the details are known whereas the same is not true of the real universe.

e-mail: jasjeet@mri.ernet.in

The physical parameters of the problem make cosmological simulations a challenging task. Unlike simulations of systems like globular clusters that can be treated as (relatively) isolated systems, the universe does not have a boundary. Therefore cosmological simulations need to be run with periodic boundary conditions. An exception are simulations of large spherical volumes that do not suffer significant deformation during the course of evolution.

Observations suggest that density perturbations are present at all scales that have been probed by observations. Amplitude of fluctuations at small scales is large and it drops at large scales ($l \gg 10$ Mpc, where 1 Mpc = 3.08×10^{24} cm). Figure 1 shows the root mean square fluctuations in mass as a function of length scale for one of the popular models. Fluctuations at scales much larger than 100 Mpc are generally not relevant for structure formation at the scale of galaxies. Thus the physical size corresponding to the periodic simulation box should be at least 100 Mpc, unless these are meant for studying large scale structure at early times when the amplitude of fluctuations is small and a smaller simulation box is acceptable. We will revisit this issue and discuss it more quantitatively but this approximate figure will suffice at present.

If we wish to study the distribution of galaxies in detail then the mass of each particle in the simulation should be much smaller than the mass of a typical galaxy. This, with the requirement that the simulation box should be more than 100 Mpc across implies that the N-body simulation must be done with at least 10^8 particles. The large number of particles required is one of the things that makes cosmological simulations challenging.

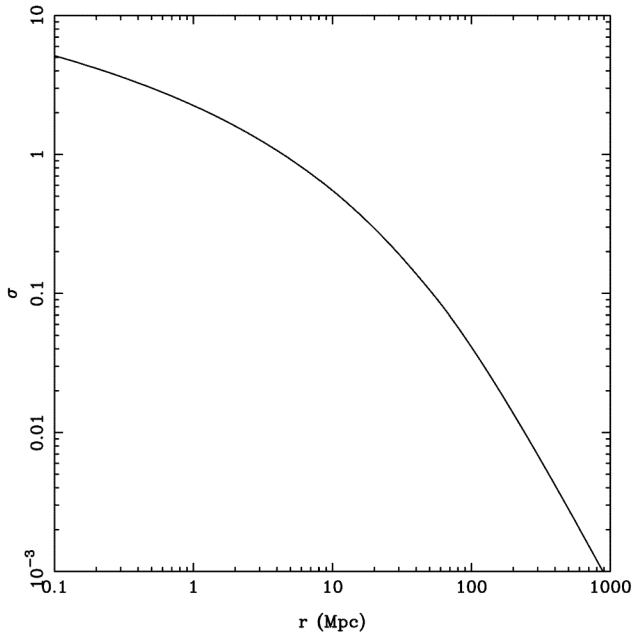

Figure 1. Root mean square fluctuations in mass σ are shown as a function of scale. The amplitude of σ is for the ΛCDM model with $\Omega_\Lambda = 0.7$, $\Omega_B = 0.05$, $h = 0.7$ and $n = 1.0$. The linearly extrapolated amplitude is plotted here and no nonlinear corrections have been included in generating this plot.

Observations show that the dominant component of matter in the universe does not radiate light and thus cannot be seen, except through its gravitational effect on visible matter[26]. Observational evidence points towards non-relativistic dark matter[27–29]. This makes our task simple as motions of non-relativistic matter can be studied in the fluid limit at large scales. (Relativistic dark matter will have significant pressure/velocity dispersion and one cannot take the fluid limit. In such a case one needs to solve the Boltzmann equation in order to correctly model the effects of free streaming[30].) Little is known about what constitutes this dark matter though it is often assumed to be made of Weakly Interacting Massive Particles (WIMP). There are strong limits on the interaction cross section of WIMPs from astrophysical considerations[31]. Non-relativistic, non-interacting (collisionless) dark matter is known as Cold Dark Matter (CDM)[32–34]. Dark matter interacts only through gravity and that makes simulating growth of perturbations somewhat simpler.

Dark matter dominates over normal matter in terms of density by a large factor[28], therefore N-body simulations with the entire matter density in dark matter provide a good first approximation for the distribution of matter. In any case, gravity is the dominant interaction and normal matter is expected to follow dark matter at large scales.

Cosmological simulations differ from other types of N-body simulations as the background is not static and expansion of the universe has to be taken into account. It is convenient to work in comoving coordinates that expand with the universe. Equations in comoving coordinates deal with the evolution of perturbations and the average quantities (density, velocity, etc.) are scaled out. We will discuss this in greater detail in the following section.

In N-body simulations, each particle represents a very large number of dark matter particles and interaction of two particles in N-body simulations should mimic the interaction of two 'fluid elements'. The fluid elements being simulated have a physical size and at scales comparable to this, the fluid elements should feel much less force than two point particles. This is done by assuming that the particles have a finite size and density profile, this leads to an effective softening of force at small scales. Clearly, the force should be softened at scales comparable to the (local) average interparticle separation in the N-body simulation. A much smaller softening length can lead to unwanted two-body relaxation[35,36]. The form of force softening also plays an important role, force softened in such a way that it matches inverse square force beyond the softening length is better[37,38]. If the softened force approaches inverse square force only asymptotically then the difference is equivalent to error in force at large separations and can lead to spurious effects.

This ends the overview of the physical requirements and associated approximations for cosmological simulations. The next section contains a detailed discussion of

cosmological N-body simulations that take only gravitational interaction into account. It is very important to understand limitations of N-body simulations as using simulations outside the domain of validity can easily lead to incorrect results. Limitations of N-body simulations are discussed in the following section. N-body simulations have been used to further our understanding of gravitational clustering in the nonlinear regime, we discuss some relevant issues. This is followed by an overview of other processes that must be taken into account for a complete study of galaxy formation, we also discuss numerical implementations of these physical processes in N-body simulations.

II. Gravity only simulations

In this section we discuss N-body simulations where only gravitational interaction is taken into account. In an N-body simulation, a given density–velocity field is represented by a set of particles[39]. Density as a function of position is obtained by averaging over this distribution of particles.

Evaluation of force and solving the equation of motion are two key components of N-body simulations. Setting up relevant initial conditions for cosmological simulations is another important aspect. We discuss the equation of motion and its integration in the following subsection. This is followed by a discussion of algorithms for force calculations. We end this section with a discussion on setting up the initial conditions.

Equation of motion

The equations that govern the evolution of a given distribution of particles are obtained by starting with the standard (Newtonian) equation of motion for a set of gravitationally interacting particles in the physical/proper coordinates. Transforming to comoving coordinates allows us to focus on perturbations in density and velocity. Information about expansion of the universe appears in these equations through the scale factor $a(t)$, obtained as a solution to the Friedman equations[1–8]. The scale factor and other cosmological parameters in these equations carry information about the average quantities that are scaled out in the coordinate transformation.

$$\ddot{\mathbf{x}} + 2\frac{\dot{a}}{a}\dot{\mathbf{x}} = -\frac{1}{a^2}\nabla\phi\,, \tag{1}$$

$$\nabla^2\phi = 4\pi G\bar{\rho}(t)a^2\delta = \frac{3}{2}H_0^2\Omega_0\frac{\delta}{a}. \tag{2}$$

Here $\mathbf{x}$ is the comoving position of a particle and is related to the physical/proper position $\mathbf{r} = a(t)\mathbf{x}$, with $a(t)$ being the scale factor. ϕ is the gravitational potential due to density perturbations, H_0 is the Hubble's constant and Ω_0 is the density parameter of non-relativistic matter at the present epoch[1,2,4–6,9]. It is assumed that the relativistic components do not cluster, or are negligible in our universe. These equations are valid for non-relativistic matter ($v \ll c$, $\phi \ll c^2$) at scales that are much smaller than the Hubble radius ($l \ll c/H_0$)[1,2].

The expansion of the universe acts as a viscous force in comoving coordinates. This drag opposes gravitational infall and as a result the growth of density perturbations is slower in an expanding universe. The time scale over which gravitational infall occurs (in absence of expansion) is comparable with the expansion time scale ($a/\dot{a}$), therefore velocities of particles do not become very large during infall. As a result, integrating equation of motion is simpler in cosmological N-body simulations.

The basic idea for numerical integration is as follows. The equation of motion expresses the second derivative of position in terms of position, velocity and time. Position and velocity at later times are expressed in terms of position and velocity at earlier times using a truncated Taylor series. The simplest truncation is not sufficiently accurate and the resulting error is of order h^2 in one time step, where h is the time step[39–41]. The key constraint in cosmological simulations is that force evaluation is very time consuming and one wishes to minimize the number of force evaluations per time step. Mainly due to this reason, cosmological N-body simulations use the Leap-Frog method[39–41] for integrating the equation of motion as it requires only one evaluation of force and the error is of order h^3. Performance of the Leap-Frog integrator can be improved considerably by making small modifications[42], but such modifications are often more useful in non-cosmological N-body simulations.

Time step h is typically chosen so that momentum is conserved and energy evolves according to the Irvine–Layzer equation[43–45]. Monitoring consistency with the Irvine–Layzer equation requires care and adds significantly to the number of operations to be carried out in an N-body simulation[46], hence it is usual to carry out test runs and fix the value of h. Additional tests can be devised, e.g. we can require that the clustering of power law models evolves in a scale invariant manner even in the strongly nonlinear regime.

Optimum value of time step h depends on the distribution of particles and it changes as this distribution evolves. It is common to use a time step that varies with time so that the N-body code does not use too small a time step when a larger value will do, or use too large an h when a smaller value is required for conserving integrals of motion. It is possible to generalize even further and choose a different h for each particle as well, motivation for this being that a few particles in a very dense regions require a small h whereas most particles are not in such regions. There are several methods of implementing this in N-body simulations, e.g. see ref. 47. Main consideration is to ensure that the positions and velocities of all particles are synchronized at frequent intervals. Using individual time steps can speed up N-body simulations by a significant amount.

Calculating force

Gravitational force in the Newtonian limit falls as $1/r^2$, hence it is a long range force and we cannot ignore force due to distant particles. This makes calculation of force the most time-consuming task in N-body simulations. As a result, a lot of attention has been focused on this aspect and many algorithms and optimizing schemes have been developed. We will discuss the major algorithms in some detail and briefly summarize other developments. We refer the reader to ref. 48 for a detailed review. In the following discussion, we also review some algorithms that are not used in cosmological N-body simulations as these can be used in hybrid algorithms.

Direct summation or particle–particle method. The most obvious approach to the problem of force calculation is to carry out a direct pairwise summation over all particles. This is also called the particle–particle (PP) method and this works very well for a small number of particles. Most early simulations used this method, see ref. 49 for an early approach, ref. 50 for a discussion of characteristics of this method, and ref. 51 for an overview of the direct summation method. Earliest cosmological simulations also used this method[52]. The number of terms in the pairwise summation increases in proportion with N^2, where N is the number of particles. This rapid variation limited the early cosmological simulations to about 10^3 particles. Limitations of the PP approach were realized as focus shifted to other methods.

The last decade has seen a revival of sorts for this method with the advent of the GRAPE chip[53]. With the latest version of GRAPE[54] it is possible to simulate systems with more than 10^6 particles. Development of efficient parallel algorithms[55] has also made this method competitive.

It is difficult to implement periodic boundary conditions in the PP method. The only optimized method available for this is the Ewald summation[56]. One of the first concrete proposals to use this method in cosmological N-body simulations is in ref 57, though Ewald summation had been used for testing other methods[46]. Adding periodic boundary conditions to a PP code remains a difficult proposition and this method is not used very often for cosmological N-body simulations.

Tree method. The key limitation of the PP method is the rapid increase in computational load with the number of particles in the N-body simulation. This in turn arises from adding individual contribution to the force due to each particle. The force of a distant group of particles can be approximated by the force due to a single pseudo particle located at the centre of mass of the group, with mass equal to the total mass of the group of particles. This approximation changes the scaling of the number of calculations from N^2 to $N \log N$.

Efficient division of particles into groups can be done by arranging particles in a tree structure[58,59]. The simulation volume is taken to be a cube and is divided into smaller cubes with 1/8 the volume each at every stage till the smallest cells have only one particle in them. Larger cells serve as groups of particles for a rapid calculation of force. An essential ingredient is the criterion for deciding whether a group of particles can be considered distant or not. This is called the cell acceptance criterion and the error in approximation is controlled by the choice of this criterion[47,58,60]. See refs 47, 60–62 for a detailed study of characteristics of the tree code, in particular of errors and timing as a function of the distribution of particles and the cell acceptance criteria.

Accuracy of the tree approximation can be improved by retaining information about moments of the particle distribution in the group, e.g. the quadruple moment.

The tree code can be optimized by vectorizing the calculation of force[63,64]. The set of acceptable distant groups of particles is almost the same for neighbouring particles and sharing of interaction lists amongst neighbouring particles can further improve the performance of the tree code[65]. The tree code can be parallelized efficiently[66], and a parallel tree code has also been implemented using the GRAPE chip[67]. The parallel algorithm divides the simulation box into domains with equal number of particles and calculations for each domain are done by a different processor. This scheme can be improved by dividing into domains with equal computational load[66].

Periodic boundary conditions are difficult to implement with tree codes, the level of difficulty being similar to that with the PP codes. Some innovative schemes have been tried[68]. Implementations using the Ewald summation[56] add a large computational overhead[47,69]. But in spite of these difficulties, tree codes have been used very effectively for cosmological N-body simulations.

Fast multipole method. The performance of tree codes can be improved upon by using including higher moments of mass distribution in cells. These and use of some other optimization schemes lead to the fast multipole method. The number of computational operations in the fast multipole method[70] scales as N, the number of particles.

Inclusion of higher moments can be also modelled in terms of pseudo-particles for easy implementation on the GRAPE chip[71,72].

An explicitly momentum conserving extension of the tree code has also been proposed and implemented[73,74], here the number of computations required scales linearly with the number of particles.

These codes also suffer from the problem of open boundary conditions and cannot be adapted very easily to cosmological problems.

Particle–mesh method. Particle–mesh (PM) method[39,46,75–79] has been used extensively for cosmological simulations and was the first method to be used for ($N \sim 10^5$) simulations[75]. In PM codes, the fact that the Poisson's equation (eq. (2)) is

a simple algebraic equation in Fourier space is combined with Fast Fourier Transforms (FFT)[40,41]. FFT requires sampling of functions at uniformly spaced points, and a grid/mesh is used for this. Usually the simulation volume is taken to be a cube with equal number of grid/mesh points along each axis.

Particles are used for representing the density and velocity field and we compute density at grid points by using weight functions[39]. Density contrast δ and the potential ϕ is defined on the grid for solving the Poisson equation. Use of particles and a mesh gives this method the very appropriate name.

By using Fourier methods we get periodic boundary conditions for free. The use of a mesh also softens the force naturally at small scales, though this softening leads to underestimation of force to fairly large scales[77,78]. Thus it would seem that PM codes are ideal for cosmological N-body simulations if we wish to study the large-scale structure of the universe.

The PM algorithm has been parallelized[80], typically this involves using parallel FFT[81]. The computational load is divided amongst processors by dividing the simulation box into slabs of equal size.

Softening at the mesh scale ensures collisionless evolution, but this also means that PM codes cannot resolve structure at length scales smaller than a mesh. Even in dense regions force softening is at the grid scale. This seriously limits the effective dynamic range of simulations run with PM codes. A related problem is that the use of mesh makes force of a particle anisotropic at small scales.

Improving force resolution in high density regions can improve the effectiveness of PM codes. Several techniques have been proposed for achieving this as poor resolution is the main shortcoming of PM codes for cosmological N-body simulations.

Adaptive mesh refinement. In Adaptive Mesh Refinement (AMR) the grid is refined in high density regions. A new mesh with smaller spacing is introduced and the low resolution force calculated using the coarse global grid is improved upon using the refined mesh[82–85]. Several levels of refinement can be introduced; indeed they are required in order to resolve substructure in dense haloes[84,85]. Care is required to ensure conservation of momentum and angular momentum in AMR codes.

P^3M: particle–particle + particle–mesh. The basic idea here is to add a 'correction' to the force computed using the PM method. This correction is computed by summing the contribution of close neighbours using the particle–particle method, hence the name PP + PM = P^3M. It is assumed that this correction depends only on the distance, i.e. it is assumed to be isotropic, and is generally added out to a distance of about 1.5–2 times the distance between neighbouring mesh points[46,86]. P^3M was the first method to be used for high resolution cosmological N-body simulations.

The PP part of the calculation can also be done with the GRAPE chip[87], though it requires some innovation as the chip is designed to return the force and not the short range correction.

The P^3M has been parallelized, though load balancing for such a code is not very simple as the overdense regions that require more CPU time are not distributed uniformly[88–90] (see below).

The P^3M has some undesirable features.

- The correction for the force is assumed to be isotropic, whereas the standard PM force has anisotropies at grid scale due to the anisotropic mesh structure. Thus the resulting force (long range + short range correction) must be anisotropic at the grid scale.
- The short range correction in force is added only up to 1.5–2 grid lengths, whereas the PM method underestimates the force out to a much larger distance[78].
- The refined inter-particle force is softened at scales much smaller than the average inter-particle separation, this can lead to two-body scattering and relaxation[35,36]. Results of astrophysical interest like the correlation function may also get modified in the process[91].
- P^3M simulations slow down at late times when the distribution of particles becomes highly clustered. At this stage computation of the short-range correction of force dominates the total number of compute operations required.

While the last item here relates to the efficiency of the code, other items in the list are far more serious as these raise doubts about the accuracy with which force is calculated. In order to retain the good features of P^3M codes and address some of the issues listed above, several variations on the theme have been suggested.

Tree + PM = TPM, GOTPM, TreePM. In this section we discuss a series of hybrid codes that combine the PM and the tree method in the same spirit as the PP and PM methods are combined in the P^3M code. We will discuss these in order of the significance of departure from the P^3M method.

- The Grid of Trees PM (GOTPM) code[92] replaces the PP part of P^3M codes with a local tree in each region. This speeds up calculation of the short-range force correction and the time taken for this calculation is less sensitive to the degree of clustering. This code takes care of the last undesirable feature listed for P^3M codes but it does not address any of the other issues. GOTPM is a parallel code where multiple levels of decomposition is used to achieve load balance[92].
- The TPM code[93–95] also addresses the problem of two body scattering in P^3M codes. In this code the short-range correction to force is added only if the particle is in a high density region. Density is computed at the position of each particle and a local tree is constructed in high density regions for computing the short-range correction to the long-range PM force. If the high density regions have a

large number of particles then the region is fragmented into sub-regions. TPM is a parallel code, ideally suited for distributed parallel computing.

- The Tree + PM (TreePM) code[96] makes significant changes to the P^3M approach. The force is partitioned between a short-range and a long-range component instead of using the force computed with a PM code as the long-range force. In the TreePM code the long-range force is truncated below a certain scale r_s and the long-range force becomes very small below this scale. In particular, the long-range force is small at grid scale if r_s is chosen appropriately and therefore anisotropies in the long-range force are also less important compared to the P^3M, GOTPM and TPM codes. The short-range force is computed using a global tree. The short-range force has to be taken into account out to larger distances than in a typical P^3M code. By tuning the model parameter r_s and a few other parameters in the code, it is possible to keep error in force below 1% for most of the particles[97]. The TreePM code is fairly simple to implement as the mathematical model of this method is well defined[97–99]. This code has been parallelized in a manner similar to that used for tree codes with some provision for the long-range force calculation[100]. The TreePM code solves all the other problems of a P^3M code but it does not address the issue of two-body scattering and relaxation.
- The Adaptive TreePM (ATreePM) code[101] addresses the problem of two-body scattering in TreePM code by using an adaptive softening length instead of a constant softening length. This can be thought of as an equivalent of AMR without using a refined grid. Local number density of particles is used to determine the softening length for particles. In order to ensure momentum conservation, force is symmetrized for particles that are closer than the softening length of either one of the two particles. Determination of local density and explicit symmetrization add a computational overhead. This is offset to some extent by the speedup at early times when the softening length is large for all the particles. Using a hierarchy of time steps optimizes this method further.

The variety of techniques used for computing gravitational force discussed above is evidence of the work being done in developing better algorithms.

It is also evident that full use has been made of parallel computing in order to achieve good performance. It is important to construct ways of comparing performance of different N-body codes. We list some suggestions here.

1. Dynamic range: The range of scales over which interaction force is computed reliably. There is rarely any problem at larger scales so the dynamic range is typically determined by the smallest separations over which the force errors are small. A useful unit for this scale is the average inter-particle separation. We can fix this scale by requiring that at all larger scales error in force due to one particle be less than 1%. For good codes, this scale should coincide with the softening length.
2. Trajectories of particles: The code should integrate the equation of motion in a reproducible manner and momentum should be conserved. The N-body code should reproduce well-known results about the correlation function and other statistical measures in the quasi-linear regime. Time step should be much smaller than the crossing time of particles in dense haloes, the ratio of time step to the smallest crossing time is a good estimate and this number should be smaller than 10^{-1}.
3. Efficiency: The N-body code should be efficient and we should be able to run large simulations in as little time as possible. This requirement is likely to conflict with the first two requirements, and one should compare both the dynamic range and efficiency. It has been proposed that error in force should be plotted against time taken for computing force for comparing codes[47], we believe this to be the correct approach.
4. Resource requirement: The N-body code must store positions and velocities of all the particles, i.e. at least 6 numbers per particle. Most advanced codes discussed here store many more numbers per particle in order to speed up force calculation. This requirement can restrict our ability to run large simulations as the memory on computers is limited. In case of parallel codes, the relevant quantity is the maximum memory utilization per particle for one processor.

This ends the discussion of different algorithms for computing force in cosmological N-body simulations.

Setting up initial conditions

N-body simulations are generally started from fairly homogeneous initial conditions, i.e. the density contrast is much smaller than unity at all scales of interest in the simulation. In this regime we can use linear perturbation theory to compute all quantities of interest. In linear theory, the evolution of density contrast can be described as a combination of a growing and a decaying mode. At late times only the growing mode survives and hence we must choose the initial density and velocity field to put the system in the growing mode. Density contrast δ is related to ϕ through eq. (2), and in linear theory the velocity field can also be expressed in terms of the gravitational potential ϕ in the growing mode. Therefore our problem reduces to generating the gravitational potential and using it to set up the density contrast and velocity field. The relation between the velocity and the potential in Zel'dovich approximation[18] is the same as that in the linear theory, therefore it is often said that the Zel'dovich approximation is used to set up the initial conditions for N-body simulations[102].

Given an initial gravitational potential ϕ_{in}, there are two schemes for generating the initial density field.

- The particles are distributed uniformly and their masses are chosen so that $M = \bar{\rho}(t_{in})(1 + \delta_{in})dV$, where δ_{in} is evaluated at the position of the particle. Here dV is the volume in the simulation box per particle and $\bar{\rho}(t_{in})$ is the average density at time t_{in}. We can either start with zero velocities, in which case we have to increase the amplitude of ϕ_{in} by a factor 5/3 to account for the presence of decaying mode. Alternatively we can choose to put the system in the growing mode and assign the appropriate velocity to each particle.
- Starting with a uniform distribution the particles are displaced using the velocity in linear theory for the growing mode. It is important to ensure that the maximum displacement is smaller than the average inter-particle separation in the simulation box. It can be shown that the resulting distribution of particles will represent the required density field[79] if the initial distribution did not have any inhomogeneities. We can retain the initial velocity field used to displace particles. If the amplitude of displacements used is larger than the average inter-particle separation, it becomes necessary to recompute the potential from displaced positions and assign initial velocities with this potential[46]. Such large displacements can lead to an incorrect realization of the power spectrum.

Schemes outlined above require an initial uniform distribution of particles. This is important as any inhomogeneities present in the initial distribution will combine with the density perturbations that are generated by displacing particles and will modify the initial conditions.

- The commonly used solution is to place particles on a cubic grid, this is a uniform but not a random distribution.
- An intuitive solution is to put particles at random inside the simulation box. This distribution has $\sqrt{n}$ fluctuations which result in spurious clustering that tend to dominate over the fluctuations we wish to simulate.
- Particles are placed in lattice cells but at a random displacement from the centre of the cell[79]. This removes the regularity of grid without sacrificing uniformity. The amplitude of fluctuations can be controlled by reducing the amplitude of displacement about the centre of the cell.
- The Glass initial conditions are obtained by evolving an arbitrary distribution of particles in an N-body simulation with a repulsive force. It can be shown that the amplitude of perturbations oscillates and decreases as $a^{-1/4}$ (ref. 79).

The initial gravitational potential. The initial density field is taken to be a Gaussian random field in most models. Linear evolution does not modify the statistics of density fields except for evolving the amplitude of perturbations. As the potential and density contrast are related through a linear equation, it follows that the gravitational potential is also a Gaussian random field. A Gaussian random field is completely described in terms of its power spectrum[103]. The Fourier components of a Gaussian random field (both the real and the imaginary part) are random numbers with a normal distribution with variance proportional to the power spectrum of the random field. This property is used to generate the Gaussian random field in Fourier space and an inverse transform gives us the initial potential in real space.

If some special features are required in the initial conditions, e.g. if we want a large planar perturbation, then we need to impose constraints on the Gaussian random field to be generated[104–107]. Gaussian random fields with a variable resolution[108] are needed for adaptive mesh refinement codes.

III. Limitations of N-body simulations

Here we discuss the scope and limitations of N-body simulations. We will consider several issues concerning the domain of validity for N-body simulations.

N-body simulations take initial density fluctuations over a finite range of scales into account. Do the fluctuations that are not taken into account make a difference to results of N-body simulations?

Several N-body studies have shown that fluctuations at small scales do not affect structure that forms at large scales[109,111–113] in a significant manner. Effects are of course there if larger scales are in linear regime but by the time larger scales reach $\delta \sim 1$, influence of smaller scales is ignorable. These studies used power spectrum, correlation function and visual appearance to reach this conclusion. Therefore in any N-body simulations that are used, the scale where root mean square fluctuations are unity should be clearly resolved for results to be reliable and independent of the small scale cutoff. On the other hand we expect some effect of small scale fluctuations on how larger density perturbations relax[114–118]. We can conclude that small scale fluctuations do not influence large scales in a significant manner but it is an important issue and further studies are required to make the effect or the lack of it more quantitative.

N-body simulations assume that the density in the simulation box is same as the average density in the universe. Therefore we must choose a simulation box such that the amplitude of fluctuations in the universe (or the model being simulated) at that scale is ignorable. Studies have shown that violating this requirement leads to an underestimate of correlation function though the mass function of small mass haloes does not change by much[119,120]. Effects at large scales can be significant[121,122]. In other words, the formation of small haloes is not disturbed but their distribution is affected by non-inclusion of long wave modes. One way of quantifying fluctuations at large scales, and hence their effect on structure formation at small scales is the amplitude of fluctuations at the scale of the simulation box. Figure 2 shows lines of constant root mean square fluctuations in mass σ. These lines are plotted as a function of scale and redshift for the ΛCDM model ($\Omega_\Lambda = 0.7$, $\Omega_B = 0.05$, $h = 0.7$ and $n = 1.0$). Curves are for $\sigma = 0.1$, 0.05, 0.025 and 0.01 in

ascending order. We can find out the lowest redshift up to which a given simulation box may be used once we fix σ that is considered acceptable at the box size.

However, a threshold in σ does not carry any information about the shape of the power spectrum and that is extremely relevant here. We find that the best way of quantifying the effect of long wave modes is to check whether including them in the simulation will change the number of massive haloes or not[123]. This can be estimated using the Press–Schechter mass function[124]. If there is a measurable effect on the number of haloes and collapsed mass, the long wave modes are relevant and must be considered. The large-scale structure in N-body simulations does not converge until all such modes are taken into account[123]. Our results for the ΛCDM model are:

- A box size of 150 h^{-1} Mpc is needed for simulations that are evolved to the present epoch ($z = 0$).
- Simulations that are evolved up to $z \sim 3$ should be 50 h^{-1} Mpc across for generic applications.
- Simulations for studying early reionization ($z \simeq 10$) should be at least 20 h^{-1} Mpc across if these are to have a representative density field.

Comparing these with the curves in Figure 2, we find that $\sigma = 0.025$ is a reasonable choice except at very high redshifts. $\sigma = 0.01$ is a better choice at higher redshifts, and this variation in the threshold value is indicative of the dependence on the shape of the power spectrum[123].

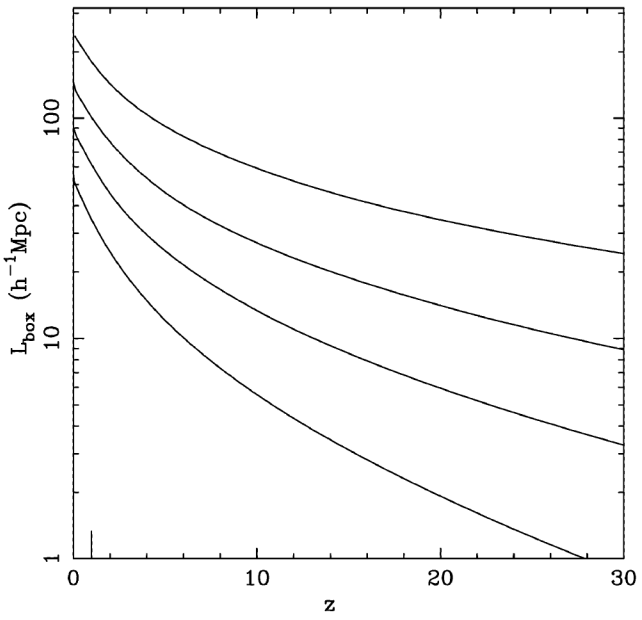

Figure 2. Lines of constant root mean square fluctuations in mass σ are plotted as a function of scale and redshift for the ΛCDM model (see text for details). Curves are for $\sigma = 0.1$, 0.05, 0.025 and 0.01 in ascending order. We can find out the lowest redshift up to which a given simulation box may be used once we fix σ that is considered acceptable at the box size. Such a threshold does not carry any information about the shape of the power spectrum. Comparing with an approach based on convergence of collapsed mass, we find that $\sigma = 0.025$ is a reasonable choice at low redshifts[123].

The effect of a finite simulation box on modes comparable with the box size can be estimated analytically[125] and this can also be used to check whether the chosen physical size of the simulation box is sufficiently large or not.

Methods have been developed to take the missing long wave modes into account[126,127], these are wave modes larger than the simulation box. These methods evolve a realization of long wave modes independently and their effect is combined with the evolution of small scales in the N-body simulation. This allows one to work with smaller simulation boxes, but the methods have limitations and cannot be used for simulating very small regions. Indeed, the criterion based on convergence of collapsed mass can be used to fix the smallest box size for which these methods can be employed[123].

Other limitations relate to two-body relaxation, poor accuracy in force calculation and discreteness. It has been asserted that in collisionless simulations forces are not required to be very accurate as the discreteness noise dominates over errors in force[128,129]. This is more relevant for non-cosmological simulations, where many of these tests were carried out. More recent studies of force softening have pointed out that the form of softening can be very important[37] in controlling the discreteness noise.

Tests that are specific for cosmological simulations[91] show that unless mass and force resolution are comparable, two-body relaxation can leave detectable signatures. Thus it is important to avoid discreteness effects and maintain parity between mass and force resolution in N-body codes.

IV. Nonlinear gravitational clustering

N-body simulations have been used extensively to understand aspects of gravitational clustering. Of the four fundamental interactions, gravity is the weakest and it is impossible to carry out laboratory experiments in gravitational dynamics of a many-body system as other interactions overwhelm gravity at these scales. N-body simulations can be used as a test-bed for doing numerical experiments in order to understand various aspects of nonlinear gravitational clustering. Indeed, we cannot claim to have understood the process of structure formation till we develop sufficient insight so as to make N-body simulations redundant except for the purpose of computing detailed predictions for models of structures formation.

Considerable progress has been made using quasilinear approximations, nonlinear scaling relations and N-body simulations, though many important questions are yet to be settled. Basic premise here is that if we are interested in only a limited amount of information about the final state then it should be possible to simplify the calculation by making

an ansatz that captures the essential physical process at work. One can also invert this and find out which physical process dominates in a given situation. We review some of the issues that have been studied in detail.

The distribution of matter is very close to homogeneous at early times, how is it confined to central region of potential wells at late times?

- Infall is described very well by first order Lagrangian perturbation theory. This, extrapolated to mildly nonlinear density contrasts is called the Zel'dovich approximation[18]. This approximation suggests that the first structures to form as a result of collapse are typically surfaces of high density, the so-called pancakes. Comparisons with N-body simulations show this to be a very good approximation up to this stage[130]. However, these surfaces of high density thicken and disappear if we extrapolate the approximation scheme to late times, whereas these pancakes do not thicken in N-body simulations. Clearly, the Zel'dovich approximation and even higher order Lagrangian perturbation theory[131–133] lacks the key ingredient that helps confine matter to potential wells.
- An artificial viscosity term can be added to the equation of motion (eq. (1)), and it can be transformed into Burger's equation[19]. This prevents interpenetration of infalling streams and pancakes remain thin, this is called the Adhesion approximation[19,20]. The approximation puts matter in the right place but fails to find a physical reason for the viscosity term.
- The equation of motion (eq. (1)) contains a velocity-dependent term where the expansion of the universe acts as a viscous drag. Further, it can be shown that the gravitational potential ϕ varies very slowly. If we assume that the potential changes at the rate expected in the linear theory while retaining its initial spatial dependence then the drag force is sufficient to confine matter to the central regions of potential wells[22,23]. This approximation (Linear Evolution of Potential (LEP), aka Frozen Potential Approximation (FPA)) compares well with N-body simulations. We can conclude that the drag term in the equation of motion plays an important role in confining matter to potential wells and the interaction of infalling particles plays a less significant role.

How strongly do density fluctuations at small scales influence density perturbations at large scales? This is relevant as N-body simulations ignore perturbations at scales smaller than those probed in the simulation.

- If there are no perturbations at large scales then small scale fluctuations generate a $P(k) \propto k^4$ spectrum at small k, this effect can be derived in the second order perturbation theory as well as from the full set of equations[1]. Amplitude of this limiting spectrum increases rapidly at early times[111]. This effect of course cannot be seen if there are fluctuations at large scales with $P(k) \propto k_n$ with $n < 4$ at small k.
- If there are small density perturbations present at large scales then there is a discernable visual effect of small scales on perturbations at large scales[109–113] for generic initial conditions. Small scales can influence the power spectrum and higher moments of particle distribution at these scales, the detailed effect depends on the form of fluctuations at small scales[134].
- If there are nonlinear density perturbations present at large scales then there is no discernable visual effect of small scales on perturbations at large scales[109–113] for generic initial conditions. There is no significant effect in the power spectrum or the two-point correlation function[109,111] in such a case. If the small scales and the scale of nonlinearity are well separated then there is no effect in higher moments either[134].
- Clearly, nonlinear gravitational clustering works to remove the effect of perturbations at large scales.
- For special initial conditions, it has been found that collisions between clumps formed due to small scale perturbations can lead to a faster relaxation of perturbations at large scales. E.g. if larger scales are modelled as a single plane wave then the resulting pancake is thinner in presence of significant small-scale fluctuations[118]. It remains to be seen whether this is an important effect for generic initial conditions.

Does gravitational clustering erase memory of initial conditions? Is there a one-to-one correspondence between some characterization of initial perturbations and the final state? Note that this is different from removing effects of perturbations at scales much smaller than the scale of nonlinearity. Clearly, if the answer to the first question is yes then we cannot recover any information about the initial density perturbations from determination of density perturbations in the nonlinear regime.

- N-body simulations show that gravitational clustering does not erase memory of initial conditions[135–140]. The final power spectrum is a function of the initial power spectrum and this relationship can be written as a one-step mapping.
- The functional form of this mapping depends on the initial power spectrum[140].
- An analytical understanding of some aspects of this mapping can be developed using simple models and approximations[141,142].
- It is found that density profiles of massive haloes have a form independent of initial conditions[143–145]. It is important to note that there is considerable scatter in density profiles obtained from N-body simulations and it is difficult to state whether a given functional form is always the best fit or not. There are a large number of recent studies (e.g. see refs 146, 147) but most of these focus on the ΛCDM model and do not explore other initial conditions.

Is it possible to predict the masses and distribution of haloes that form as a result of gravitational clustering?

- The initial density field is taken to be a Gaussian random field, and for hierarchical models[1] the simple assumption that each peak undergoes collapse independent of the surrounding density distribution can be used to estimate the mass function[124,148] and several related quantities.
- N-body simulations show that this simple set of approximations is incorrect, however the resulting mass function is fairly accurate over a wide range of masses.
- Merger rates can be computed using the extended Press–Schechter formalism[148], these are useful for many applications[149–151]. It is also possible to estimate clustering properties of haloes using this formalism[152,153].
- Modifying some of these assumptions can lead to improved predictions[154–156].

We have highlighted some of the key issues in nonlinear gravitational clustering in this section and reviewed how N-body simulations and various approximations have been used to develop insight. N-body simulations have been and are being used for a variety of other applications, the most notable being computing detailed predictions for models of structure formation.

V. Gastrophysical effects

N-body simulations that take only gravitational interactions into account can be used to obtain the large-scale distribution of galaxies and clusters. Further details can only be obtained using simulations with gas physics. There are two very different approaches to including gas physics in N-body simulations.

The classical approach is to use a grid to solve fluid equations[157–168]. Fluid interactions are short range and information from nearby grid points is sufficient to evolve fluid properties at any grid point. Of course, fluids must respond to gravitational field of the matter distribution. This type of a code is well adapted for capturing shocks and discontinuities. Resolution of the code can be enhanced in high density regions by using mesh refinement. In cosmological simulations it is important to improve the mass resolution of dark matter particles along with an enhancement in resolution for hydrodynamics. Grid codes can be easily generalized to include other effects such as magnetohydrodynamics[169,170]. Such codes have been used effectively for cosmological applications[157,162,164–168].

The Smooth Particle Hydrodynamics (SPH)[171,172] where particles are assigned fluid properties is a very different approach to the same problem. The fluid properties like pressure, density, temperature, etc. at any point can be found by averaging over particles in the region using a weight function. Use of interpolation for determining fluid properties makes it impossible to resolve shocks and discontinuities in SPH codes. There are several known limitations[173–175] that one should be aware of. SPH codes are relatively easy to implement and several implementations are in use. Several implementations for cosmological simulations are in use[176–178]. Variations of SPH with focus on entropy equation[179,180] or on simulation of multi-phase media[181,182] have been developed.

Both types of codes have been compared and give similar results[183,184] for astrophysical applications.

Effects other than hydrodynamics can also be incorporated in a similar fashion in both types of codes. For example, elementary chemical reactions like formation of hydrogen molecules, cooling, heating, etc. are important local effects. Key effects like star formation and feedback from stellar and other sources are difficult to include as vastly different scales are of relevance. As a result, much of the treatment of these effects has remained phenomenological.

Radiation transport is a non-local effect and is difficult to take into account[185–191]. The radiation field is either assumed to be isotropic and homogeneous, or it is assumed to originate at a few point sources. The time scales over which the radiation field evolves are much shorter than most other time scales in the problem. Thus we can assume that the density and velocity fields do not change while studying evolution of the radiation field.

State-of-the-art simulations can be used for studying a variety of problems. We summarize three systems here:

- The inter-galactic medium[192–197] is believed to contain mostly pristine material with elements other than hydrogen and helium present in only trace amounts. The radiation field can be assumed to be isotropic and homogeneous at $z \leq 3$, though at higher redshifts it is patchy. Densities in the inter-galactic medium do not become very large.
- The intra-cluster medium[99,184,198–207] is believed to be close to hydro-static equilibrium. Gas in the intra-cluster medium is so hot that many radiative processes can be ignored safely. There are several sources of energy in clusters of galaxies, most of these are distributed uniformly with some contribution from a central AGN. Magnetic fields may be important enough to make the physics of the problem more complicated.
- Formation of first stars[208–215] requires simulating cooling of hydrogen (mainly) through various processes till molecules form and lead to temperatures low enough for formation of stars.

Focus in future will be on incorporating further effects so that more complex problems like galaxy formation can be studied in full detail. N-body simulations are useful not only as tools for evolving complex systems, these can also be used to understand which effects play a more important role in different phases of this evolution. Next decade holds the promise of understanding physics of galaxy formation with help of N-body simulations.

1. Peebles, P. J. E., *Large Scale Structure of the Universe*, Princeton University Press, Princeton, 1980.
2. Padmanabhan, T., *Structure Formation in the Universe*, Cambridge University Press, Cambridge, 1993.
3. Peebles, P. J. E., *An Introduction to Physical Cosmology*, Princeton University Press, Princeton, 1993.
4. Peacock, J. A., *Cosmological Physics*, Cambridge University Press, Cambridge, 1998.
5. Coles, P. and Lucchin, F., *Cosmology: The Origin and Evolution of Cosmic Structure*, John Wiley, Chichester, 2002, 2nd edn.
6. Padmanabhan, T., *Theoretical Astrophysics: Galaxies and Cosmology*, Cambridge University Press, Cambridge, 2002, vol. 3.
7. Dodelson, S., *Modern Cosmology*, Academic Press, Amsterdam, 2003, 2nd edn.
8. Liddle, A. R., *An Introduction to Modern Cosmology*, John Wiley, Chichester, 2003, 2nd edn.
9. Bernardeau, F., Colombi, S., Gaztanaga, E. and Scoccimarro, R., *Phys. Rep.*, 2002, **367**, 1
10. Ellis, G. F. R. *et al.*, *The Evolution of the Universe: Report of the Dahlem Workshop on the Evolution of the Universe, 51* (ed. Bonner, G. and Gottlober, S.), John Wiley, New York.
11. Buchert, T. and Carfora, M., *Class. Quan. Grav.*, 2002, **19**, 6109.
12. Buchert, T. and Carfora, M., *Phys. Rev. Lett.*, 2003, **90**, 031101.
13. Turner, E., Cen, R. and Ostriker, J. P., *Astronom. J.*, 1992, **103**, 1427.
14. Sethi, S. K., *MNRAS*, 2003, **342**, 962.
15. Gopal, R. and Sethi, S. K., *JPaA*, 2003, **24**, 51.
16. Sethi, S. K. and Subramanian, K., *MNRAS*, 2005, **356**, 778.
17. Peebles, P. J. E., *A&A*, 1974, **32**, 391.
18. Zel'dovich Ya, B., *A&A*, 1970, **5**, 84.
19. Gurbatov, S. N., Saichev, A. I. and Shandarin, S. F., *MNRAS*, 1989, **236**, 385.
20. Weinberg, D. H. and Gunn, J. E., *MNRAS*, 1990, **247**, 260.
21. Matarrese, S., Lucchin, F., Moscardini, L. and Saez, D., *MNRAS*, 1992, **259**, 437.
22. Brainerd, T. G., Scherrer, R. J. and Villumsen, J. V., *ApJ*, 1993, **418**, 570.
23. Bagla, J. S. and Padmanabhan, T., *MNRAS*, 1994, **266**, 227.
24. Sahni, V. and Coles, P., *Phys. Rep.*, 1995, **262**, 1.
25. Hui, L. and Bertschinger, E., *ApJ*, 1996, **471**, 1.
26. Trimble, V., *ARA&A*, 1987, **25**, 425.
27. Ma, C.-P., Bertschinger, E., Hernquist, L., Weinberg, D. H. and Katz, N., *ApJ*, 1997, **484**, 1.
28. Spergel, D. N. *et al.*, *ApJS*, 2003, **148**, 175.
29. Tegmark, M. *et al.*, astro-ph/0310723.
30. Ma, C.-P. and Bertschinger, E., *ApJ*, 1994, **429**, 22.
31. Hennawi, J. F. and Ostriker J. P., *ApJ*, 2002, **572**, 41.
32. Bond, J. R. and Efstathiou, G., *ApJL*, 1984, **285**, 45.
33. Davis, M., Efstathiou, G., Frenk, C. S. and White, S. D. M., *ApJ*, 1985, **292**, 371.
34. Liddle, A. R. and Lyth, D. H., *Phys. Rep.*, 1993, **231**, 1.
35. Binney, J. and Knebe, A., *MNRAS*, 2002, **333**, 378.
36. Melott, A. L., Shandarin, S. F., Splinter, R. J. and Suto, Y., *ApJL*, 1997, **479**, 79.
37. Athanassoula, E., Fady, E., Lambert, J. C. and Bosma, A., *MNRAS*, 2000, **314**, 475.
38. Dehnen, W., *MNRAS*, 2001, **324**, 273.
39. Hockney, R. W. and Eastwood, J. W., *Computer Simulation using Particles*, McGraw Hill, New York, 1988.
40. Antia, H. M., *Numerical Methods for Scientists and Engineers*, Tata McGraw Hill, New York, 1991.
41. Press, W. H., Flannery, B. P., Teukolsky, S. A. and Vetterling, W. T., *Numerical Recipes*, Cambridge University Press, Cambridge, 1986.
42. Hut, P., Makino, J. and McMillan, S., *ApJ*, 1995, **443**, 93.
43. Irvine, W. M., *Doctoral Dissertation*, Harvard University, 1961.
44. Layzer, D., *ApJ*, 1963, **138**, 174.
45. Dmitriev, N. A. and Zel'dovich, Ya, B., *Sov. Phys. JETP*, 1964, **18**, 793.
46. Efstathiou, G., Davis, M., Frenk, C. S. and White, S. D. M., *ApJS*, 1985, **57**, 241.
47. Springel, V., Yoshida, N. and White, S. D. M., *New Astron.*, 2001, **6**, 79.
48. Bertschinger, E., *ARA&A*, 1998, **36**, 599.
49. Aarseth, S. J. and Hoyle, F., *Astrophis. Norv.*, 1964, **9**, 313.
50. Makino, J. and Hut, P., *ApJS*, 1988, **68**, 833.
51. Aarseth, S. J., *PASP*, 1999, **111**, 1333.
52. Aarseth, S. J., Turner, E. L. and Gott, J. R. III, *ApJ*, 1979, **228**, 664.
53. Toshikazu, E., Makino, J., Fukushige, T., Taiji, M., Sugimoto, D., Ito, T. and Okumora, S. K., *PASJ*, 1993, **45**, 269.
54. Makino, J., Fukushige, T., Koga, M. and Namura, K., *PASJ*, 2003, **55**, 1163.
55. Makino, J., *New Astron.*, 2002, **7**, 373.
56. Ewald, P. P., *Ann. Phys.*, 1921, **64**, 253.
57. Rybicki, G. B., In *The Use of Supercomputers in Stellar Dynamics* (eds Hut, P. and McMillan, S.), Springer, Berlin, 1986, p. 181.
58. Barnes, J. and Hut, P., *Nature*, 1986, **324**, 446.
59. Jernigan, J. G. and Porter, D. H., *ApJS*, 1989, **71**, 871.
60. Salmon, J. K. and Warren, M. S., *J. Comp. Phys.*, 1994, **111**, 136.
61. Hernquist, L., *ApJS*, 1987, **64**, 715.
62. Barnes, J. E. and Hut, P., *ApJS*, 1989, **70**, 389.
63. Hernquist, L., *J. Comput. Phys.*, 1990, **87**, 137.
64. Makino, J., *J. Comput. Phys.*, 1990, **87**, 148.
65. Barnes, J. E., *J. Comput. Phys.*, 1990, **87**, 161.
66. Dubinski, J., *New Astron.*, 1996, **1**, 133.
67. Makino, J., *PASJ*, 2004, **56**, 521.
68. Bouchet, F. R. and Hernquist, L., *ApJS*, 1988, **68**, 521.
69. Hernquist, L., Bouchet, F. R. and Suto, Y., *ApJS*, 1991, **75**, 231.
70. Greengard, L. and Rokhlin, V., *J. Comput. Phys.*, 1987, **73**, 325.
71. Kawai, A. and Makino, J., *ApJL*, 2001, **550**, 143.
72. Kawai, A., Makino, J. and Ebisuzaki, T., *ApJS*, 2004, **151**, 13.
73. Denhen, W., *ApJL*, 2000, **536**, 39.
74. Dehnen, W., *J. Comput. Phys.*, 2002, **179**, 27.
75. Klypin, A. A. and Shandarin, S. F., *MNRAS*, 1983, **204**, 891.
76. Miller, R. H., *ApJ*, 1983, **270**, 390.
77. Bouchet, F. R., Adam, J. C. and Pellat, R., *A&A*, 1985, **144**, 413.
78. Bouchet, F. R. and Kandrup, H. E., *ApJ*, 1985, **299**, 1.
79. Bagla, J. S. and Padmanabhan, T., *Pramana*, 1997, **49**, 161.
80. Merz, H., Pen, U.-L. and Trac, H., astro-ph/0402443, 2004.
81. http://www._tw.org/
82. Villumsen, J. V., *ApJS*, 1989, **71**, 407.
83. Suisalu, I. and Saar, E., *MNRAS*, 1995, **274**, 287.
84. Kravtsov, A. V., Klypin, A. A. and Khokhlov, A. M., *ApJS*, 1997, **111**, 73.
85. Knebe, A., Green, A. and Binney, J., *MNRAS*, 2001, **325**, 845.
86. Couchman, H. M. P., *ApJL*, 1991, **368**, L23.
87. Brieu, P. P., Summers, F. J. and Ostriker, J. P., *ApJ*, 1995, **453**, 566.
88. Theuns, T., *Comput. Phys. Commun.*, 1994, **78**, 238.
89. Macfarland, T., Couchman, H., Pearce, F. R. and Pichlmeier, J., *New Astron.*, 1998, **3**, 687.
90. Brieu, P. P. and Evrard, A. E., *New Astron.*, 2000, **5**, 163.
91. Splinter, R. J., Melott, A. L., Shandarin, S. F. and Suto, Y., *ApJ*, 1998, **497**, 38.
92. Dubinski, J., *New Astron.* 2004, **9**, 111.
93. Xu, G., *ApJS*, 1995, **98**, 355.

94. Bode, P., Ostriker, J. P. and Xu, G., *ApJS*, 2000, **128**, 561.
95. Bode, P. and Ostriker, J. P., *ApJS*, 2003, **145**, 1.
96. Bagla, J. S., *J. Astrophys. Astron.*, 2002, **23**, 185; astro-ph/9911025.
97. Bagla, J. S. and Ray, S., *New Astron.*, 2003, **8**, 665.
98. White, M., *ApJS*, 2002, **143**, 241.
99. Borgani, S. *et al.*, *MNRAS*, 2004, **348**, 1078.
100. Ray, S. and Bagla, J. S., astro-ph/0405220, 2004.
101. Bagla, J. S. and Ray, S., *Work in Progress*.
102. Joyce, M. and Marcos, B., astro-ph/0410451, 2004.
103. Bardeen, J. M., Bond, J. R., Kaiser, N. and Szalay, A. S., *ApJ*, 1986, **304**, 15.
104. Bertschinger, E., *ApJL*, 1987, **323**, 103.
105. Hoffman, Y. and Ribak, E., *ApJL*, 1991, **380**, 5.
106. Hoffman, Y. and Ribak, E., *ApJ*, 1992, **384**, 448.
107. van de Weygaert, R. and Bertschinger, E., *MNRAS*, 1996, **281**, 84.
108. Bertschinger, E., *ApJS*, 2001, **137**, 1.
109. Little, B., Weinberg, D. H. and Park, C., *MNRAS*, 1991, **253**, 295.
110. Evrard, A. E. and Crone, M. M., *ApJL*, 1992, **394**, 1.
111. Bagla, J. S. and Padmanabhan, T., *MNRAS*, 1997, **286**, 1023.
112. Peebles, P. J. E., *ApJ*, 1985, **297**, 350.
113. Couchman, H. M. P. and Peebles, P. J. E., *ApJ*, 1998, **497**, 499.
114. Peebles, P. J. E., *ApJ*, 1990, **365**, 27.
115. Weinberg, M. D., *MNRAS*, 2001, **328**, 311.
116. Ma, C.-P. and Boylan-Kolchin, M., *Phys. Rev. Lett.*, 2004, **93**, 012301.
117. Ma, C.-P. and Bertschinger, E., *ApJ*, 2004, **612**, 28.
118. Bagla, J. S., Prasad, J. and Ray, S., astro- ph/0408229, 2004.
119. Gelb, J. M. and Bertschinger, E., *ApJ*, 1994, **436**, 467.
120. Gelb, J. M. and Bertschinger, E., *ApJ*, 1994, **436**, 491.
121. Kauffmann, G. A. M. and Melott, A. L., *ApJ*, 1992, **393**, 415.
122. Barkana, R. and Loeb, A., *ApJ*, 2004, **609**, 474.
123. Bagla, J. S. and Ray, S., astro-ph/0410373, 2004.
124. Press, W. H. and Schechter, P., *ApJ*, 1975, **187**, 452.
125. Seto, N., *ApJ*, 1999, **523**, 24.
126. Tormen, G. and Bertschinger, E., *ApJ*, 1996, **472**, 14.
127. Cole, S., *MNRAS*, 1997, **286**, 38.
128. Hernquist, L., Hut, P. and Makino, J., *ApJL*, 1993, **402**, 85.
129. Hernquist, L., Hut, P. and Makino, J., *ApJL*, 1993, **411**, 53.
130. Coles, P., Melott, A. L. and Shandarin, S. F., *MNRAS*, **260**, 765.
131. Buchert, T., *A&A*, 1993, **267**, L51.
132. Bouchet, F. R., Colombi, S., Hivon, E. and Juszkiewicz, R., *A&A*, 1995, **296**, 575.
133. Catelan, P., *MNRAS*, 1995, **276**, 115.
134. Bagla, J. S. and Prasad, J., *Work in Progress*, 2005.
135. Hamilton, A. J. S., Kumar, P., Lu, E. and Matthews, A., *ApJ*, 1991, **374**, L1.
136. Jain, B., Mo, H. J. and White, S. D. M., *MNRAS*, 1995, **276**, L25.
137. Peacock, J. A. and Dodds, S. J., *MNRAS*, 1994, **267**, 1020.
138. Peacock, J. A. and Dodds, S. J., *MNRAS*, 1996, **280**, L19.
139. Padmanabhan, T., Cen, R., Ostriker, J. P. and Summers, F. J., *ApJ*, 1996, **466**, 604.
140. Smith, R. E. *et al.*, *MNRAS*, 2003, **341**, 1311.
141. Nityananda, R. and Padmanabhan, T., *MNRAS*, 1994, **271**, 976.
142. Padmanabhan, T., *MNRAS*, 1996, **278**, L29.
143. Navarro, J. F., Frenk, C. S. and White, S. D. M., *ApJ*, 1996, **462**, 563.
144. Tormen, G., Bouchet, F. R. and White, S. D. M., *MNRAS*, 1997, **286**, 865.
145. Subramanian, K., Cen, R. and Ostriker, J. P., *ApJ*, 2000, **538**, 528.
146. Power, C. *et al.*, *MNRAS*, 2003, **338**, 14.
147. Fukushige, T., Kawai, A. and Makino J., *ApJ*, 2004, **606**, 625.
148. Bond, J. R., Cole, S., Efstathiou, G. and Kaiser, N., *ApJ*, 1991, **379**, 440.
149. Bower, R. G., *MNRAS*, 1991, **248**, 332.
150. Lacey, C. and Cole, S., *MNRAS*, 1992, **262**, L627.
151. Cohn, J. D., Bagla, J. S. and White, M., *MNRAS*, 2001, **325**, 1053.
152. Mo, H. J. and White, S. D. M., *MNRAS*, 1996, **282**, 1096.
153. Mo, H. J., Jing, Y. P. and White, S. D. M., *MNRAS*, 1997, **284**, 189.
154. Monaco, P., *MNRAS*, 1997, **290**, 439.
155. Sheth, R. K. and Tormen, G., *MNRAS*, 1999, **308**, 119.
156. Sheth, R. K., Mo, H. J. and Tormen, G., *MNRAS*, 2001, **323**, 1.
157. Evrard, A. E., *MNRAS*, 1988, **235**, 911.
158. Stone, J. M. and Norman, M. L., *ApJS*, 1992, **80**, 753.
159. Bryan, G. L., Norman, M. L., Stone, J. M., Cen, R. and Ostriker, J. P., *Comput. Phys. Commun.*, 1995, **89**, 149.
160. Norman, M. L., *RmxAC*, 2004, **9**, 66.
161. Xu, G., *MNRAS*, 1997, **288**, 903.
162. Cen, R., *ApJS*, 1992, **78**, 341.
163. Ryu, D., Ostriker, J. P., Kang, H. and Cen, R., *ApJ*, 1993, **414**, 1.
164. Cen, R. and Ostriker, J. P., *ApJ*, 1999, **517**, 31.
165. Quilis, V., Ibanez, J. M. A. and Saez, D., *ApJ*, 1996, **469**, 11.
166. Shapiro, P. R., Martel, H., Villumsen, J. V. and Owen, J. M., *ApJS*, 1996, **103**, 269.
167. Couchman, H. M. P., Thomas, P. A. and Pearce, F. R., *ApJ*, 1995, **452**, 797.
168. Pearce, F. R. and Couchman, H. M. P., *New Astron.*, 1997, **2**, 411.
169. Stone, J. M. and Norman, M. L., *ApJS*, 1992, **80**, 791.
170. Stone, J. M., Mihalas, D. and Norman, M. L., *ApJS*, 1992, **80**, 819.
171. Gingold, R. A. and Monaghan, J. J., *MNRAS*, 1977, **181**, 375.
172. Monaghan, J. J., *ARA&A*, 1992, **30**, 543.
173. Hernquist, L., *ApJ*, 1993, **404**, 717.
174. Steinmetz, M. and Mueller, E., *A&A*, 1993, **268**, 391.
175. Bate, M. and Burkert, A., *MNRAS*, 1997, **288**, 1060.
176. Steinmetz, M., *MNRAS*, 1996, **278**, 1005.
177. Katz, N., Weinberg, D. H. and Hernquist, L., *ApJS*, 1996, **105**, 19.
178. Davé, R., Dubinski, J. and Hernquist, L., *New Astron.*, 1997, **2**, 277.
179. Serna, A., Alimi, J.-M. and Chieze, J.-P., *ApJ*, 1996, **461**, 884.
180. Springel, V. and Hernquist, H., *MNRAS*, 2002, **333**, 649.
181. Ritchie, B. W. and Thomas, P. A., *MNRAS*, 2001, **323**, 743.
182. Springel, V. and Hernquist, H., *MNRAS*, 2003, **339**, 312.
183. Kang *et al.*, *ApJ*, 1994, **430**, 83.
184. Frenk *et al.*, *ApJ*, 1999, **525**, 554.
185. Abel, T., Norman, M. L. and Madau, P., *ApJ*, 1999, **523**, 66.
186. Razoumov, A. O. and Scott, D., *MNRAS*, 1999, **309**, 287.
187. Kessel-Deynet, O. and Burkert, A., *MNRAS*, 2000, **315**, 713.
188. Sokasian, A., Abel, T. and Hernquist, L. E., *New Astron.*, 2001, **6**, 88.
189. Gnedin, N. Y. and Abel, T., *New Astron.*, 2001, **6**, 437.
190. Cen, R., *ApJS*, 2002, **141**, 211.
191. Maselli, A., Ferrara, A. and Ciardi, B., *MNRAS*, 2003, **345**, 379.
192. Hernquist, L., Katz, N., Weinberg, D. H. and Jordi, M.-E., *ApJL*, 1996, **457**, 51.
193. Wadsley, J. W. and Bond, J. R., astro-ph/9612148, 1996.
194. Theuns, T., Leonard, A., Efstathiou, G., Pearce, F. R. and Thomas, P. A., *MNRAS*, 1998, **301**, 478.
195. Bryan, G. L., Machacek, M., Anninos, P. and Norman, M. L., *ApJ*, 1999, **517**, 13.
196. Meiksin, A. and White, M., *MNRAS*, 2001, **324**, 141.
197. Croft, R. A. C., Hernquist, L., Springel, V., Westover, M. and White, M., *ApJ*, 2002, **580**, 634.
198. Evrard, A. E., *ApJ*, 1990, **363**, 349.
199. Thomas, P. A. and Couchman, H. M. P., *MNRAS*, 1992, **257**, 11.

200. Schindler, S. and Mueller, E., *A&A*, 1993, **272**, 137.
201. Navarro, J. F., Frenk, C. S. and White, S. D. M., *MNRAS*, 1995, **275**, 720.
202. Anninos, P. and Norman, M. L., *ApJ*, 1996, **459**, 12.
203. Roettiger, K., Loken, C. and Burns, J. O., *ApJS*, 1997, **109**, 307.
204. Lewis, G. F. *et al.*, *ApJ*, 2000, **536**, 623.
205. Metzler, C. A. and Evrard, A. E., *ApJ*, 1994, **437**, 564.
206. Lewis, G. F. *et al.*, *ApJ*, 2000, **536**, 623.
207. Jubelgas, M., Springel, V. and Dolag, K., *MNRAS*, 2004, **351**, 423.
208. Anninos, P., Zhang, Y., Abel, T. and Norman, M. L., *New Astron.* 1997, **2**, 209.
209. Abel, T., Anninos, P., Zhang, Y. and Norman, M. L., *New Astron.*, 1997, **2**, 181.
210. Abel, T., Anninos, P., Norman, M. L. and Zhang, Y., *ApJ*, 1998, **508**, 518.
211. Bromm, V., Coppi, P. S. and Larson, R. B., *ApJL*, 1999, **527**, 5.
212. Abel, T., Bryan, G. L. and Norman, M. L., *ApJ*, 2000, **540**, 39.
213. Abel, T., Bryan, G. L. and Norman, M. L., *Science*, 2002, **295**, 93.
214. Bromm, V., Coppi, P. S. and Larson, R. B., *ApJ*, 2002, **564**, 23.
215. Jang-Condell, H. and Hernquist, L., *ApJ*, 2001, **548**, 68.

ACKNOWLEDGEMENTS. Developing insight into developments in cosmological N-body simulations would not have been possible without access to a high performance computing facility. Cluster computing facilities at the Harish-Chandra Research Institute (http://cluster.mri.ernet.in) were used heavily for trying, testing and also developing some of the algorithms. I thank Suryadeep Ray, with whom I have tried out many experiments with N-body algorithms. I also thank H. K. Jassal and Suryadeep Ray for a careful reading of the manuscript. This research has made use of NASA's Astrophysics Data System.