The history of cosmic baryons: discoveries using advanced computing

Michael L Norman
Physics Department and CASS, UC San Diego, La Jolla, CA 92093
mlnorman@ucsd.edu

Abstract. We live in the era of the cosmological concordance model. This refers to the precise set of cosmological parameters which describe the average composition, geometry, and expansion rate of the universe we inhabit. Due to recent observational, theoretical, and computational advances, these parameters are now known to approximately 10% accuracy, and new efforts are underway to increase precision tenfold. It is found that we live in a spatially flat, dark matter-dominated universe whose rate of expansion is accelerating due to an unseen, unknown dark energy field. Baryons--the stuff of stars, galaxies, and us--account for only 4% of the total mass-energy inventory. And yet, it is through the astronomical study of baryons that we infer the rest. In this talk I will highlight the important role advanced scientific computing has played in getting us to the concordance model, and also the computational discoveries that have been made about the history of cosmic baryons using hydrodynamical cosmological simulations. I will conclude by discussing the central role that very large scale simulations of cosmological structure formation will play in deciphering the results of upcoming dark energy surveys.

1. Cosmology’s Standard Model

A principal goal of modern cosmology is the precise measurement of the half dozen cosmological parameters that describe the shape, matter-energy contents, and expansion history of our universe. Recent observations of the cosmic microwave background (CMB), high redshift supernovae, and galaxy clusters have yielded a concordant set of parameters with a precision of about 10% [1]. In addition, the spectrum of matter fluctuations which give rise to cosmic structure have been measured at the epoch of recombination over a range of scales $10^4 > \lambda > \sim 10$ Mpc, and found to have a shape consistent with the popular cosmological constant + cold dark matter ($\Lambda$CDM) model, and its amplitude has been measured to about 10% accuracy. The new cosmological frontier is pushing the accuracy of these measurements to 1% accuracy, or better, and extending the range of scales over which the matter power spectrum is measured. At this level of accuracy, important inferences about the nature of dark energy and dark matter can be made [2].

It is found that we live in a spatially flat, dark-matter dominated universe whose rate of expansion is accelerating due to an unseen, unknown dark energy field [3]. In units of the critical density, dark matter and dark energy contribute about 24% and 73%, respectively. Baryons, the stuff of stars, galaxies, and us, account for only 4% of the total mass-energy inventory. And yet, it is through the astronomical study of baryons that we infer the rest. Through observational, theoretical, and computational studies, the history of baryons from the big bang to the present time is coming into focus. In this paper I will highlight the important role advanced scientific computing has played in getting us to the concordance model, and also the computational discoveries that have been made about the history of cosmic baryons using hydrodynamical cosmological simulations.
2. A Brief History of Baryons

Figure 1 summarizes our current understanding of the history of the universe. Baryons—cosmology-speak for particles containing quarks such as protons and neutrons—emerged from the intense radiation field of the big bang through a process known as baryogenesis. The physics of this is not well understood since it occurred at an energy scale we have not yet probed in the laboratory [4]. The second major event in the history of baryons was the formation of the lightest atomic nuclei—H, D, He3, He4 and Li7—when the typical energy scale was a few MeV. This occurred about 1-100 sec after the big bang (ABB). The physics of big bang nucleosynthesis (BBNS) is very well understood by now, and this topic is considered to be in the precision era [5]. A key result is the determination of the mean baryon density in the universe determined by the observed D/H ratio in primordial intergalactic clouds. The next major event occurred about 380,000 yr ABB when the energy scale was a few eV and electrons and light nuclei combined to form neutral atoms. The sudden disappearance of electrons at the epoch of recombination left the universe transparent to radiation, which we see as the cosmic microwave background (CMB) today. The study of the CMB has made spectacular advances in the last decade, and is now also in the precision era [6]. The recent measurement of the spectrum of temperature fluctuations in the CMB by the Wilkinson Microwave Anisotropy Probe have not only confirmed the cosmic baryon density using BBNS, but also measured the ratio of baryons to non-baryonic dark matter (cold dark matter)[1]. After recombination, matter density fluctuations are amplified by gravitational instability. Those fluctuations that enter the nonlinear regime form the observed structures in the universe (galaxies, clusters, etc.), and require large scale numerical simulation for their accurate dynamical evolution.
Figure 2. Gridding the universe. The computational volume expands with cosmic scale factor \( a(t) \), as given by a particular solution of the Friedmann equation. The dynamical equations for the matter contents are transformed to comoving coordinates, and solved assuming periodic boundary conditions. Initial conditions are set down at \( t_1 \) when density fluctuations are still in the linear regime.

3. Universe in a Box

How does one go about simulating the universe? The evolution of structure in the universe is an initial value problem suitable for computation. Globally, the universe obeys the Friedmann equations, which are derived from the Einstein equation in the limit of a homogeneous and isotropic space-time metric. The Friedmann equations specifies the first and second time derivatives of the metric scale factor \( a(t) \) in terms of the mean mass-energy densities \( \Omega_i \), where \( i=(b,v,\text{dm},\text{de}) \) for baryons, neutrinos, dark matter, and dark energy, respectively. As mentioned above, these parameters are now known to about 10% precision. Locally, that is on scales much smaller than the horizon scale \( d_H \sim ct \), the matter constituents obey Newtonian dynamical equations for dark matter and baryons, which interact only via their self-consistent gravitational field. In practice what is done is to transform the dynamical equations to a frame which is comoving with the expanding universe (figure 2). With this change of variables, the dynamical equations become Newton’s laws of N-body dynamics for the dark matter, which is assumed to behave as a collisionless phase fluid, Euler’s equations for the primordial gas, and the Poisson equation for the gravitational field. Each of these equations is slightly altered in form to handle the coordinate transformation, but otherwise familiar. The important point is that numerical techniques that have been developed over the past 50 years of computational physics research for simulating fluids and plasmas are directly applicable to numerical cosmology. The computational challenges are nonetheless substantial, as we now discuss.

Simulations are carried out in a 3D Cartesian cube in comoving coordinates with assumed periodic boundary conditions. The later is motivated not by the assumption that the universe is a crystal, but rather the cosmological principle, which states that on sufficiently large scales, the universe is homogeneous and isotropic. Obviously, cosmological structure is not homogeneous and isotropic on the scale of galaxies, etc., but on much larger scales, its statistics is. Operationally, then, one selects a volume which is large enough to contain a representative statistical sample of the universe as a whole for the objects of interest. Initial conditions must also be specified. We impose linear amplitude Gaussian perturbations with a power spectrum given by observations (figure 8) on an otherwise uniform distribution of baryons and dark matter at some early time in the universe, typically shortly after recombination. These fluctuations are then evolved into the nonlinear regime as they are amplified by gravitational instability.
3.1 Computational Challenges

Cosmological simulations of structure formation present both a multiscale challenge, and a multiphysics challenge. For very large scale simulations, they also present a data management challenge which we briefly discuss at the end. Regarding the first, roughly 4 orders of magnitude separate the size of a typical galaxy (~10 kpc) and galaxy large scale structure (LSS), which exhibits a cellular “filament and void” pattern on ~30-100 Mpc scales (figure 3). If one wishes to both resolve the internal structure of galaxies with 10 resolution elements, say, and at the same time simulate a volume comparable to the modern surveys (~1 Gpc^3), one arrives at a spatial dynamic range requirement of 6 orders of magnitude in 3D! The highest dynamic range that has been achieved to date in a large statistical volume is the Millenium Simulation \cite{7} carried out in Europe by the Virgo Consortium. They achieved a dynamic range of 10^5 in every collapsing structure with the mass of a small galaxy or larger everywhere within a simulation volume 700 Mpc on a side. This dark matter-only N-body simulation used 10^{10} particles, and consumed 350,000 cpu-hrs on 512 processors of an IBM p690 cluster.

Impressive as this is, baryons were not included in the Millenium Simulation, only the clustering of cold dark matter. In order to understand the physical processes that form the baryonic objects we observe, a host of physical processes must be modelled which make cosmological structure formation a multiphysics challenge as well. A schematic diagram of these processes is given in figure 4. The clustering of dark matter creates potential wells into which the primordial gas sinks. The gas becomes dense and hot enough for it to radiate energy away, cool, condense further, and form stars. Stars emit radiation which escapes their parent galaxies and builds up the metagalactic radiation background. Radiation above 13.6 eV ionizes the intergalactic medium, altering its ability to cool and condense to form galaxies. Thus, there are tight two way couplings between the three principal “actors” in baryonic cosmology: intergalactic gas, galaxies (and quasars), and the radiation background. In order to simulate this complicated multiphysics problem, one needs robust, accurate algorithms for multispecies gas dynamics, stiff reaction kinetics, good microphysical models for photo-heating and cooling, and multidimensional radiative transfer. Star formation is treated as a sub-grid phenomenological model which is calibrated against observations. In addition to radiation, stellar ensembles feed back energy and heavy elements into the intergalactic gas, which heats and modifies the cooling properties of the gas. All this must be simulated self-consistently, ideally over the vast range of scales previously discussed.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{multiscale.png}
\caption{The multiscale challenge for all cosmological structure formation simulations.}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{multiphysics.png}
\caption{The multiphysics challenge for hydrodynamic cosmological simulations in particular.}
\end{figure}
3.2 The Enzo Code

The Enzo code [8], developed at the Laboratory for Computational Astrophysics (LCA) at UC San Diego over the past 5 years, and before that at the National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign, confronts both of these technical challenges. Enzo is a grid-based hybrid code (hydro + N-body) which uses the adaptive mesh refinement (AMR) algorithm of Berger & Collela [9] to improve spatial resolution in regions of large gradients, such as in gravitationally collapsing objects. The method is attractive for cosmological applications because it: (1) is spatially- and time-adaptive, (2) uses accurate and well-tested grid-based methods for solving the hydrodynamics equations, and (3) can be well optimized and parallelized. The central idea behind AMR is to solve the evolution equations on a grid, adding finer meshes in regions that require enhanced resolution. Mesh refinement can be continued to an arbitrary level, based on criteria involving any combination of overdensity (dark matter and/or baryon), Jeans length, cooling time, etc., enabling us to tailor the adaptivity to the problem of interest. The code solves the following physics models: collisionless dark matter and star particles, using the particle-mesh N-body technique [10]; gravity, using FFTs on the root grid and multigrid relaxation on the subgrids; cosmic expansion; gas dynamics, using the piecewise parabolic method (PPM)[11]; multispecies nonequilibrium ionization and H\textsubscript{2} chemistry, using backward Euler time differencing [12]; radiative heating and cooling, using subcycled forward Euler time differencing [12]; and a parameterized star formation/feedback recipe [13]. At the present time, magnetic fields and radiation transport are being installed. Enzo is publicly available at http://cosmos.ucsd.edu/enzo.

Enzo is implemented using a combination of C/C++ and Fortran, and is parallelized using the MPI message-passing library. I/O is handled using the HDF5 library from NCSA. The code has been optimized for the IBM Power3/4 architectures by Robert Harkness at the San Diego Supercomputer Center (SDSC). The code runs on any distributed memory parallel supercomputer or PC cluster, although 64-bit architectures are required for applications using the implicit kinetic solver. Enzo is used in two basic modes. In unigrid mode, AMR is switched off, and parallelization is used for block-decomposing the root grid. This is the mode employed for intergalactic medium simulations, where, similar to turbulence simulations, we want high uniform spatial resolution. Unigrid simulations as large as 1024\textsuperscript{3} using as many as 1024 processors have been successfully carried out on SDSC’s IBM Data Star system. In AMR mode, used for galaxy formation simulations, the root grid is typically smaller (e.g., 128\textsuperscript{3} or 256\textsuperscript{3}), but still block decomposed for parallel execution. AMR typically generates thousands of subgrids, which are dynamically load-balanced across the processors [14]. In general, the computational load increases with level of refinement up to some maximum which can be 10-100 times that required to advance the root grid one timestep.

Figure 5. Enzo simulation of galaxy formation and large-scale structure. **Left:** survey volume simulated with 1 billion cells and particles on SDSC Blue Horizon. **Right:** Zoom in on young galaxies using adaptive mesh refinement (AMR).
3.2 Example: Galaxy Formation within the Cosmic Web

An example of Enzo’s capabilities is shown in figure 5. Here we solve the N-body equations for 1 billion dark matter particles and the equations of multispecies Eulerian hydrodynamics, photoionization by an ultraviolet radiation background, and self-gravity on a uniform grid of 1 billion \((1024^3)\) cells. Shown is a volumetric rendering of the baryon density, revealing a network of filaments referred to as the cosmic web [15]. The calculation was done in 2004 on 512 processors of SDSC’s IBM Power3 Blue Horizon computer, and produced 10TB of raw data and 6 TB of derived data. Galaxies form along the filaments and at the intersection of filaments. Despite the large grid, a spatial dynamic range of only 1000 is insufficient to even begin resolving the galaxies. This calculation serves as a survey volume for follow-on adaptive mesh refinement (AMR) simulations which resolve the galaxies’ internal structure. At right is shown an old AMR simulation of galaxy formation done at NCSA in 1998. Using AMR, a dynamic range of \(1.5 \times 10^4\) is achieved where galaxies condense. We supplement the physics used for the unigrid simulation with a recipe for converting cold, dense gas within the galaxies into star particles, which are evolved dynamically with a particle-in-cell technique. Due to computer power and data handling limitations at the time, only 1/64 of the survey volume \((256^3\) base grid) could be simulated at high resolution. Now, with more powerful parallel supercomputers and data management technologies, we can in principle simulate the entire volume at high spatial resolution. Making that a practical reality is the overarching goal of a cosmology simulation data grid project, recently initiated in collaboration with LLNL. This is briefly described in the last section.

4. History of Cosmic Baryons: Discoveries using Advanced Computing

Here we highlight a few discoveries made using hydrodynamic cosmological simulations by us and others. They are presented in order of cosmic time, not chronologically by publication date.

4.1 First Structures and First Stars

The concordance model is a model where cold dark matter dominates the mass density of the universe. Therefore its gravitational clustering drives the baryonic dynamics. Such models are called hierarchical models because they form structures “bottom-up”; i.e., through the hierarchical mergers of smaller collapsed objects which formed previously. For example, a cluster of galaxies forms by the merger of galaxy groups, which forms from the merger of individual galaxies. This recursion leads to an obvious question: what were the very first objects to form, and when did they form? Were they galaxies or something much smaller? My former PhD students Tom Abel, now an assistant professor at Stanford, and Greg Bryan, now an assistant professor at Columbia, and I set out to answer this question in 1994. By 2001, thanks to Enzo and careful attention to the microphysics of primordial gas, we arrived at a robust conclusion [16]. We found that dark matter fluctuations with a characteristic mass scale of \(5 \times 10^5\) solar masses gather enough baryons within their potential wells at \(t \approx 50-100\) Myr ABB to form massive primordial stars (figure 6). These so-called Population III stars are of primordial composition, and have calculable properties [17].

The robustness of this prediction stems from the fact that both the initial conditions and the physics of the dark matter and gas are well known, making this a well-posed computational problem. The use of adaptive mesh refinement allows us to resolve all the important length and timescales in the problem over a range of \(10^{12}\) in isolated objects [18]. So powerful is AMR at resolving the collapsing gas that we terminate the calculation when our physics model breaks down at particle densities of \(10^{12}\) particles/cc. At these densities, radiative transfer effects not included in the simulation become important.

The significance of our result is that massive Pop III stars are known to feed back ionizing radiation and heavy elements to their environs when they explode as supernovae. They also leave behind compact remnants such as neutron stars and black holes. If they form with a sufficiently high space density in the early universe, they could begin the process of reionizing the intergalactic gas, as well as enriching it with heavy elements. Understanding the reionization and chemical enrichment of the intergalactic medium is an active field of current research.
Figure 6. Massive primordial stars are the first objects to form in the cosmological concordance model, roughly 50-100 Myr after the big bang. Top row: a time sequence showing the formation of collapsed baryons at z~24, which form a massive collapsing protostar by z~18. Middle row: projections of the baryon density at z=18.2 on various linear scales resolved by our AMR technique, from cosmological scales (6 kpc) to the protostellar core scale (0.06 pc). Bottom row: same as middle row, but showing projections of gas temperature. The massive protostar forms in the cold gas at the center of the first object.

Figure 7. Schematic, showing the origin of the Lyman α forest absorption in the spectra of high redshift quasars. (Courtesy M. Murphy)

4.2 Nature of the Lyman alpha Forest
The Lyman α forest refers to a picket fence of absorption features seen in the spectra of high redshift quasars blueward of the quasars’ intrinsic hydrogen Ly α emission feature. Shortly after its discovery by Lynds in 1972, it was proposed that the absorption lines were caused by intervening intergalactic
gas clouds along the LOS to the quasar, each absorbing at the hydrogen Ly \( \alpha \) resonant wavelength, cosmologically redshifted. Until the mid 1990s, no satisfactory physical model existed for the origin and distribution of the clouds. Hydrodynamic cosmological simulations [19] and high spectral resolution observations [20] have revolutionized all that. This author took part in this “revolution”. Based on the impressive agreement between simulations and observations on a variety of HI absorption line statistics [20] it is now widely accepted that the H I Ly \( \alpha \) absorption is caused by mildly overdense, highly photoionized intergalactic gas that closely traces the dark matter distribution in cold dark matter (CDM) models of structure formation. According to these simulations, on scales of a Mpc or more, the dark matter and baryons trace out a network of sheets and filaments (figure 5) referred to as the cosmic web arising from the growth of primordial matter fluctuations [15]. Because of the cosmological redshift, a filament at redshift \( z \) along the LOS to the quasar absorbs at a wavelength \( \lambda = 1216 \) \((1+z)\) Angstrom. Thus, the absorption spectrum is a 1D sample of the matter distribution along that LOS. Many high quality spectra of this sort have been obtained over the past decade, and have been used to measure the matter power spectrum on scales an order of magnitude smaller that currently accessible using CMB anisotropies [21].

### 4.3 Missing Baryons at \( z=0 \)

We now have three independent techniques to measure the cosmic baryon density \( \Omega_b \): BBNS which probes the universe at 1 sec, CMB anisotropies which probe the universe at 380,000 yr, and the mean level of absorption in the Lyman \( \alpha \) forest [22], which probes the universe at 3-7 Gyr. All three methods agree that \( \Omega_b \approx 0.04 \). However, when astronomers tally up the baryons they can see in the local universe (\( z=0 \)) as stars and gas in galaxies, or hot gas bound to clusters of galaxies, they find only about 10% of the 4%, or \( \Omega_b \) (observed today)\( \approx 0.004 \) [23]. This is known as the missing baryon problem. Given the measurement of the amount of baryons in the universe at three very different epochs, all which agree, it has been assumed that the baryons are in some form which is difficult to detect at \( z=0 \). This mystery was resolved by Renyue Cen and Jeremiah Ostriker in 2000 [24]. They performed a simulation very similar to the one shown in figure 5, but continued the simulation all the way to the present day. They found that the cosmic web filaments become shock heated to temperatures in the range \( 10^5 < T < 10^7 \) K where they emit very little radiation in the EUV and X-ray. They found that 30-40% of the intergalactic medium resides in this phase, which they dubbed the warm-hot intergalactic medium (WHIM). Since their discovery, X-ray astronomers have searched in vain to see this WHIM, however it is known to be below the detection threshold of our current X-ray satellites [25]. Astronomers instead have begun trying to detect the WHIM in absorption in the spectra of high redshift quasars; something that is observationally feasible. The so-called X-ray forest is directly analogous to the Ly \( \alpha \) forest seen at higher redshifts, only now one looks for resonance lines of highly ionized oxygen in the IGM. A Search and Discovery article in the April 2005 issue of Physics Today reports on a potential observational confirmation of Cen & Ostriker’s computational discovery.

### 4.4 Baryons and Dark Matter in X-ray Clusters

It has been known since the early days of X-ray astronomy that clusters of galaxies are copious emitters of X-rays. These X-rays are produced by thermal bremsstrahlung in the hot, diffuse intracluster plasma which is gravitationally bound to the cluster [26]. From the X-ray luminosity and temperature of the gas, one can determine the mass of the emitting gas since \( L_x=3\pi\epsilon_x(T) \), where \( \epsilon_x(T) \) is the known X-ray emissivity per gram. If one makes the assumption that the X-ray emitting gas is isothermal, spherically symmetric, and in hydrostatic equilibrium in the cluster potential, then from the X-ray intensity distribution on can build analytic models to derive the density distribution of the dark matter which confines the hot gas. It is found that the dynamical mass \( M_d \) of the cluster exceeds the mass of the hot gas by 8-10. If one further makes the argument that a galaxy cluster is a fair sample of the universe as a whole, then the ratio \( M_d/M_\star = \Omega_d/\Omega_\star \), where \( \Omega_\star \) is the sum of the baryon and dark matter density—one of the fundamental cosmological parameters. Assuming \( \Omega_d=0.04 \), measured using BBNS or otherwise, one can thus infer \( \Omega_\star \) from clusters. This was made rigorous by Simon White and collaborators [27], who in 1993 carried out hydrodynamic cosmological simulations...
of X-ray clusters in dark matter-dominated models. They found that $\Omega_m \approx 0.3$, far less than the $\Omega_m = 1$ result sought by inflation theorists. This was one of the earliest and most convincing demonstrations that we live in a low matter density universe. The cluster determination of $\Omega_m$ is one of the three key inputs to determining the cosmological parameters of the concordance model [28].

5. **Exciting Opportunities in Cosmology using Ultrascale Computing**

The recent precision measurements of the cosmological parameters [1] put the field of cosmology in a place not unlike particle physics, where the goal going forward is to refine and test the standard model with yet higher precision measurements. Fundamental science questions driving the field include the nature of dark energy and dark matter, the formation and evolution of galaxies and quasars, and how and when the intergalactic medium was reionized. In this last section, I touch on how advanced simulation can contribute to research into the nature of dark matter and dark energy, as these are closest to DOE Office of Science interests.

5.1 **Precision Cosmology using the Lyman $\alpha$ Forest**

The intimate physical connection between absorption and dark matter density has stimulated many researchers to explore the possibility of using observations of the Ly $\alpha$ forest as a cosmological probe of the $z=2$-4 universe in much the same way as observations of CMB anisotropies have been used to probe the $z=1280$ universe. The forest probes the primordial power spectrum on scales an order of magnitude smaller than the highest resolution CMB experiments. The slope of the matter power spectrum on small scales is a probe of inflation [29]; a sharp cutoff in the power spectrum at high wave numbers is a probe of the dark matter particle’s mass (figure 8)[30]. The observational function that assumes the role of the CMB angular power spectrum is the flux power spectrum. This is essentially the Fourier transform of transmitted flux spectrum averaged over many lines of sight.

The key to extracting cosmological information from the Ly $\alpha$ forest is to understand how the absorption relates to the underlying matter power spectrum. Mathematically, it is assumed that if $P_i(k)$...
is the flux power spectrum, and \( P(k) \) is the matter power spectrum, then the two are related by \( P(k) = b^2(k) P_s(k) \), where \( b(k) \) is the scale-dependent bias factor \[21\]. If one knew \( b(k) \) by some means, then one could measure the matter power spectrum by measuring \( P_s(k) \). The approach to evaluating \( b(k) \) is to carry out a large suite of hydrodynamical cosmological simulations of the Ly \( \alpha \) forest varying \( P(k) \), the cosmological, and astrophysical parameters. From these outputs synthetic absorption spectra are generated by passing lines of sight through the volume. An ensemble of such spectra are used to derive \( P_s(k) \) and hence \( b(k) \). Rupert Croft and collaborators used this approach to measure \( P(k) \) to \(-20\%\) accuracy over a range of scales 1.5 orders of magnitude smaller than current CMB measurements \[21\]. We applied this approach in \[31\] and uncovered a parameter degeneracy between the matter power amplitude and the mean temperature of the photoionized gas, confirming an earlier result.

Extending this approach to precision measurements (~1%) encounters two obstacles, both technical in nature. The first is that real observed spectra are imperfect; systematic uncertainties due to continuum fitting, noise, finite spectral resolution, and non-Ly \( \alpha \) absorption features must be understood and removed. The Tytler group at UCSD, with whom I collaborate, specializes in this type of work. The second is that the bias parameter is an unknown function of the cosmological and astrophysical parameters. In principle, an exhaustive search of parameter space can determine this function. In practice this is infeasible because of the high dimensionality of the parameter space (~10), and the high cost of the individual simulations. Fortunately, an exhaustive search is not required. McDonald \[32\] showed that \( b(k) \) becomes a weak function of all but a few parameters at small \( k \). Based on simulations varying domain size and spatial resolution, he concluded that a simulation which faithfully evolves fluctuations over the range of scales \( 40 \text{ Mpc}/h > \lambda(\text{comoving}) > 40 \text{ kpc}/h \) would contain essentially all of the useful information contained in the Ly \( \alpha \) forest. Here \( h \) is the Hubble constant in units of 100 km/s/Mpc. The lower end of this range is set by the physical thickness of the absorbing structures. The upper end of this range is set by the turnover in the matter power spectrum. This is a range of scales of 1000. The best that has been achieved to date is our \( 1024^3 \) simulation in a \( L=80 \text{ Mpc} \) box shown in figure 5. The range of scales present in this simulation is only 256, since \( \lambda_{\text{min}} = 2\Delta x \) is the minimum wavelength that can be represented on a grid of spacing \( \Delta x \), and \( \lambda_{\text{max}} = L/2 \) is the wavelength of the largest Fourier mode that can fit in a box of sidelen\( ght L \).

We are planning to use NERSC’s Seaborg system to carry out a simulation of the Ly \( \alpha \) forest that will provide the most accurate estimate for \( b(k) \) over this range of scales currently feasible. This will be a \( 2048^3 \) simulation of the concordance model in a box 200 Mpc on a side. This, combined with the extensive parameter survey that we have performed in smaller boxes will permit us to resolve the parameter degeneracy between matter power and IGM temperature. It is estimated this calculation will require ~4 TB of RAM and 1.4 M cpu-hrs on 4096 Seaborg processors.

5.2 Using Clusters to Measure the Dark Energy Equation of State

High redshift supernovae have been used to trace the expansion history of the Universe over roughly half its present age \[3\]. The discovery that the expansion is currently accelerating implies that space is pervaded by dark energy whose pressure is negative, like the quantum mechanical vacuum energy. The dark energy equation of state is written \( p = w \rho \), with current observations consistent with \( w = -1 \) as Einstein’s cosmological constant term would have. New observations of high redshift supernovae are planned via a joint NASA/DOE space mission in order to place tighter constraints on \( w \). Complimentary to supernovae as a method for measuring the expansion history of the universe is the number density of rich galaxy clusters as a function of redshift. Galaxy clusters can be observed to high redshift using X-ray, the Sunyaev-Zeldovich effect, and weak lensing surveys \[33\]. Extraordinary precision will be needed in both approaches to see if the dark energy density is indeed constant, or slowly evolving as in some proposals. In order to assess the systematic biases which might arise in the cluster surveys, mock catalogs of clusters from hydrodynamic simulations carried out in large survey volumes are needed. The large limiting redshifts for these surveys, plus the large area of the sky to be surveyed, translate into many cubic Gpc of space that needs to be simulated at quite high dynamic range.
Figure 9. Tiling the lightcone with high resolution AMR simulations for analysis of weak lensing shear, SZ clusters, etc. The benefit of this strategy is high spatial resolution and constant angular resolution within a large survey volume.

Our approach to this will be to optimally tile the light cone between us and the limiting redshift of z=1.5-2 with a large number of statistically independent, large AMR simulations (figure 9). The concept of simulating clusters on the lightcone was introduced by Evard et al. [34], and refined by White and Hu [35]. Evard used a single large dark matter simulation in a 3 Gpc box to extract the light cone at a constant comoving spatial resolution. White and Hu used multiple simulations in box sizes chosen to tile the light cone of a given opening angle, which is important when comparing with observations of a fixed angular resolution. We will adopt the approach of White and Hu, but instead of the small uniform grid dark matter simulations they used, we will use large AMR hydrodynamic simulations with a dynamic range of $>10^4$. Each simulation will be done on a $512^3$ root grid and 6 levels of refinement. We estimate that each of a dozen simulations covering a different redshift interval will require 0.5-1 TB of RAM, and require of order 0.2-0.5 M cpu-hrs on machine like Seaborg. Data management will be a major challenge working at this scale. We are beginning to explore the use of distributed data grids and workflow software to manage this challenge, in particular the Storage Resource Broker (SRB) technology [36] developed by Reagan Moore’s group at SDSC. We are pursuing this as a part of a collaborative project between UCSD and LLNL. We plan to provide simulated sky maps and mock catalogs for the Large Synoptic Survey Telescope (LSST) project [37], which has as a principal science goal the mapping of dark matter in the universe via weak lensing measurements.

Acknowledgements The author would like to acknowledge the contributions of his students and collaborators to the work presented here, in particular Tom Abel, Greg Bryan, and Brian O’Shea. Enzo has been developed with the support of the National Science Foundation via grants ASC-9318185, AST-9803137, and the National Computational Science Alliance under NSF Cooperative Agreement ASC-9740300, PACI Subaward ACI-9619019. Simulations have been carried out the NCSA and SDSC with LRAC allocation MCA98020.

References
[1] Spergel, D N et al. 2003. Astrophys. J. Supp., 148, 175
[2] Seljak, U et al. 2005. Phys. Rev. D, vol. 71, Issue 10, id. 103515
[3] Perlmutter, S 2003. Physics Today, April 2003.
