The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

Borde, Arnaud and Palanque-Delabrouille, Nathalie and Rossi, Graziano and Viel, Matteo and Bolton, James S. and Yeche, Christophe and Rich, James and Le Goff, Jean-Marc (2014) New approach for precise computation of Lyman-alpha forest power spectrum with hydrodynamical simulations. Journal of Cosmology and Astroparticle Physics, 2014 (July). 005/1-005/31. ISSN 1475-7516

Access from the University of Nottingham repository:
http://eprints.nottingham.ac.uk/42503/1/1401.6472.pdf

Copyright and reuse:

The Nottingham ePrints service makes this work by researchers of the University of Nottingham available open access under the following conditions.

This article is made available under the University of Nottingham End User licence and may be reused according to the conditions of the licence. For more details see:
http://eprints.nottingham.ac.uk/end_user_agreement.pdf

A note on versions:

The version presented here may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the repository url above for details on accessing the published version and note that access may require a subscription.

For more information, please contact eprints@nottingham.ac.uk

Prepared for submission to JCAP

New approach for precise computation of Lyman-alpha forest power spectrum with hydrodynamical simulations

Arnaud Borde,[a] Nathalie Palanque-Delabrouille,[a,b] Graziano Rossi,[a,f] Matteo Viel,[c,d] James S. Bolton,[e] Christophe Yèche,[a] Jean-Marc LeGoff,[a] and Jim Rich[a]

[a]CEA, Centre de Saclay, IRFU/SPP, F-91191 Gif-sur-Yvette, France
[b]Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[c]INAF, Osservatorio Astronomico di Trieste, Via G. B. Tiepolo 11, 34131 Trieste, Italy
[d]INFN/National Institute for Nuclear Physics, Via Valerio 2, I-34127 Trieste, Italy
[e]School of Physics and Astronomy, University of Nottingham, University Park, Nottingham, NG7 2RD, UK
[f]Department of Astronomy and Space Science, Sejong University, Seoul, 143-747, Korea
E-mail: nathalie.palanque-delbrouille@cea.fr, arnaud.borde@cea.fr

Abstract. Current experiments are providing measurements of the flux power spectrum from the Lyman-α forests observed in quasar spectra with unprecedented accuracy. Their interpretation in terms of cosmological constraints requires specific simulations of at least equivalent precision. In this paper, we present a suite of cosmological N-body simulations with cold dark matter and baryons, specifically aiming at modeling the low-density regions of the inter-galactic medium as probed by the Lyman-α forests at high redshift. The simulations were run using the `GADGET-3` code and were designed to match the requirements imposed by the quality of the current SDSS-III/BOSS or forthcoming SDSS-IV/eBOSS data. They are made using either $2 \times 768^3 \simeq 1$ billion or $2 \times 192^3 \simeq 14$ million particles, spanning volumes ranging from $(25\,\mathrm{Mpc.h}^{-1})^3$ for high-resolution simulations to $(100\,\mathrm{Mpc.h}^{-1})^3$ for large-volume ones. Using a splicing technique, the resolution is further enhanced to reach the equivalent of simulations with $2 \times 3072^3 \simeq 58$ billion particles in a $(100\,\mathrm{Mpc.h}^{-1})^3$ box size, i.e. a mean mass per gas particle of $1.2 \times 10^5 M_\odot.h^{-1}$. We show that the resulting power spectrum is accurate at the 2% level over the full range from a few Mpc to several tens of Mpc. We explore the effect on the one-dimensional transmitted-flux power spectrum of four cosmological parameters (n_s, σ_8, Ω_m and H_0) and two astrophysical parameters (T_0 and γ) that are related to the heating rate of the intergalactic medium. By varying the input parameters around a central model chosen to be in agreement with the latest Planck results, we built a grid of simulations that allows the study of the impact on the flux power spectrum of these six relevant parameters. We improve upon previous studies by not only measuring the effect of each parameter individually, but also probing the impact of the simultaneous variation of each pair of parameters. We thus provide a full second-order expansion, including cross-terms, around our central model. We check the validity of the second-order expansion with independent simulations obtained either with different cosmological parameters or different seeds. Finally, a comparison to the one-dimensional Lyman-α forest power spectrum obtained with BOSS by Palanque-Delabrouille et al. [1] shows an excellent agreement.

arXiv:1401.6472v3 [astro-ph.CO] 11 Jun 2014

Contents

1 Introduction **1**

2 Simulation grid **3**
- 2.1 Methodology 3
- 2.2 Simulation variable parameters 4
  - 2.2.1 Cosmological parameters 4
  - 2.2.2 Astrophysical parameters 5
  - 2.2.3 Grid values 6

3 Pipeline **6**
- 3.1 `CAMB` 6
- 3.2 `2LPT` 7
- 3.3 `Gadget-3` 8
- 3.4 `extract` 10
- 3.5 Post-processing 10

4 Convergence tests **11**
- 4.1 Mass resolution 12
- 4.2 Box size 14
- 4.3 Summary of convergence requirements 14

5 Splicing **14**

6 Results and discussions **16**
- 6.1 Assessment checks 17
- 6.2 Comparison to SDSS-III/BOSS DR9 data 19
- 6.3 Power spectrum 20
- 6.4 Density-temperature relation 21

7 Conclusions **22**

A List of all simulations **24**

1 Introduction

In the intergalactic medium, light is absorbed at the Lyman-α absorption wavelength $\lambda_{\mathrm{Ly}\alpha} \sim 1216\,\text{Å}$ by neutral hydrogen. Combined with cosmological redshifting, it produces an absorption spectrum which is observed on any background source as a map of transmission fraction as a function of redshift [2]. For light sources at sufficiently high redshift for the absorption of the intergalactic matter to be sufficiently strong, the continuous nature of the absorption spectrum is easily observable as the Lyman-α forest. Although this spectrum can be seen as a series of merged absorption lines, simulations have shown that it is in reality a map of density fluctuations in the intervening intergalactic medium seen in redshift space, with peaks of absorption at the density peaks of the absorbing gas [3, 4]. Moreover, the fluctuations in the Lyman-α forest absorption can be used as a tracer of the varying density of intergalactic gas expected from the growth of structure from primordial fluctuations

in the Universe [5]. An intergalactic medium heated exclusively by photo-ionization can be modeled with hydrodynamic simulations [6–11] and the physics at play in this model is well understood.

The amplitude and shape of the power spectrum of mass fluctuations can be measured through the information embedded in the Lyman-α forest observable in quasar spectra [5, 12–18]. These can later be used to constrain cosmology [13, 19–21], the baryonic acoustic oscillation peak position [22] or the sum of the masses of neutrinos [23, 24]. Small numbers of high-resolution spectra were first used to measure the Lyman-α forest power spectrum: 1 Keck HIRES spectrum [5], 19 spectra from the Hershel telescope on La Palma or the AAT [25], 8 Keck HIRES spectra [17], a set of 30 Keck HIRES and 23 Keck LRIS spectra [21], or a set of 27 high resolution UVES/VLT QSO spectra at redshifts ~ 2 to 3 [11, 26, 27]. The Sloan Digital Sky Survey [28] lead to a major breakthrough providing a much larger sample of 3035 medium-resolution ($R = \lambda/\Delta\lambda_{\rm FWHM} \approx 2000$) quasar spectra for the measurement of the Lyman-α forest power spectrum by McDonald et al. [29]. The large number of observed quasars allowed detailed measurements with well characterized errors of the power spectrum up to larger scales, probing the linear regime and providing cosmological constraints [23, 30].

The next step is carried out by the Sloan Digital Sky Survey III [31] through the Baryon Oscillation Spectroscopic Survey (BOSS, Dawson et al. [32]). Quasars at redshift $z > 2$, which are useful for Lyman-α forest analyses, are specifically targeted, leading to a much higher number of such quasar spectra than in previous surveys (Dawson et al. [32] and references therein). It thus allows a measurement of the Lyman-α power spectrum in both three-dimensional and one-dimensional redshift space. The 60,000 quasars spectra with Lyman-α forest absorption [33, 34] of the Data Release 9 [35] have already permitted the measurement of the BAO peak position and new constraints on the history of the expansion of the universe [36–38] using the three-dimensional power spectrum. A measurement of the one-dimensional power spectrum P_{1D} with a significant improvement over previous studies in the achieved precision has also been conducted [1]. Other background sources, such as Lyman-break galaxies, are also being investigated for a dense mapping of the Lyman-α forest [39].

Whereas the measurement of the three-dimensional power spectrum uses only information from the flux correlation of pixel pairs in different quasar spectra and thus provides information on rather large scales, the one-dimensional power spectrumP_{1D}, defined by

$$P_{1D}(k_\parallel) = \int_0^\infty \frac{dk_\perp k_\perp}{2\pi} P_{3D}(k_\parallel, k_\perp) \,, \tag{1.1}$$

uses the correlation of pixel pairs on the same quasar spectrum and thus provides a complementary, useful information on smaller scales that are fundamental to constrain the physical parameters of the Lyman-α forest. The one-dimensional P_{1D} is probing scales at the transition from linear to non-linear regime. Therefore, cosmological simulations are required to provide insight on the non-linear physics of the intergalactic medium on the small scales probed by P_{1D}. Such simulations are then used to constrain various cosmological and astrophysical parameters that have an effect on the power spectrum [11, 24, 30, 40–43].

Here, we present a set of 28 cosmological smoothed particles hydrodynamics (SPH) and N-body simulations that reproduce the impact on the one-dimensional matter power spectrum of the values taken by the most relevant cosmological and astrophysical parameters. Only the baryonic particles undergoes a SPH treatment, i.e. they receive an additional hydrodynamic acceleration, and their internal entropy per unit mass is evolved as an independent thermodynamic variable. All our simulations are run using `GADGET-3`, last described by [44]. The requirements in terms of box size, resolution and redshift coverage of our simulations were derived from the Data Release 9 quasar catalogue [33, 35] of the Baryon Oscillation Spectroscopic Survey [32]. We extrapolate these requirements so that this

suite of simulations may also be used for future spectroscopic surveys such as eBOSS[1] (planned for 2014-2018) or DESI[2] [45] (2018-2023). The full suite of simulations will be made available upon request to the authors.

The outline of the paper is as follows. In section 2 we describe our grid and the values chosen for the different parameters we varied. In section 3, we present the simulations pipeline, along with our solutions to issues such as the generation of the initial conditions or the radiative cooling and heating processes that occur in the intergalactic medium (IGM). In section 4 we present tests that were made to determine the required characteristics of our simulations in the light of our goals. We describe, in section 5, the splicing technique we apply in order to obtain simulations with the desired resolution and box size. We demonstrate the validity of our grid approach and present final discussions on this suite of simulations in section 6. Conclusions and perspectives are given in section 7. A recapitulation of all the simulations performed for this study is given in appendix A.

2 Simulation grid

Ideally, in order to derive confidence intervals on each parameter of a cosmological model with eight to ten free parameters, one would like to compute theoretical predictions for thousands of models, exploring most of the parameter space. Statistical frameworks have been studied to optimize the precision of the model for a reduced number of simulations, such as Latin hypercube sampling [46]. While this method is superior to a random sampling of the parameters for instance as regards the attained precision [47], it still requires a large number of simulations. Latin hypercube sampling has only been tested so far to predict the power spectrum on large-scales, using low resolution simulations, of order 128^3 particles for a 450 Mpc.h^{-1} box [48, 49].

When dealing with Lyman-α data, running large numbers of simulations is not possible due to the high execution time of each hydrodynamical simulation. Hence, various approximate methods have been developed in which a restricted number of simulations is used either to calibrate a flux-to-matter power-spectrum bias function or to Taylor expand the flux power-spectrum with respect to cosmological parameters in the vicinity of a best-fit model. For cosmological predictions of the power spectrum in the Lyman-alpha regime where hydrodynamical simulations are required, the grid approach as presented in [50] is generally adopted (cf. [51] for instance for a recent application). This is the method we have selected for this work.

2.1 Methodology

We model the variations of the one-dimensional power spectrum with relevant input parameters by a second-order Taylor expansion about our best-guess model:

$$\begin{aligned} f(\mathbf{x}+\mathbf{\Delta x}) = & \; f(\mathbf{x}) \\ & + \sum_i \frac{\partial f}{\partial x_i}(\mathbf{x})\Delta x_i \\ & + \frac{1}{2}\sum_i \sum_j \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x})\Delta x_i \Delta x_j \,. \end{aligned} \tag{2.1}$$

With n parameters, the total number of simulations required to get the Taylor expansion coefficients is $1+2n+n(n-1)/2$, where the terms account for, respectively, the central (or best-guess) model, two

[1] http://www.sdss3.org/future/eboss.php and http://www.sdss3.org/future/sdss4.pdf
[2] http://desi.lbl.gov

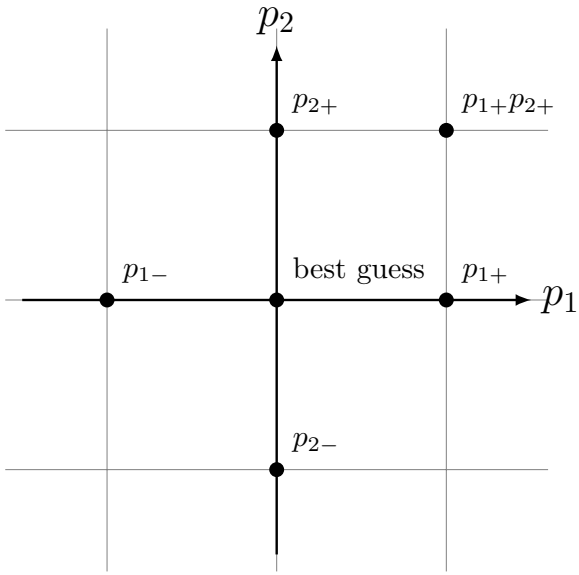


Figure 1. Illustration of the required grid for a second-order Taylor expansion in a two-dimensional parameter space.

other values of each parameter to derive the first and second-order derivatives, and the simultaneous variation of each pair of parameters to compute the cross derivatives (cf. figure 1). With this lattice, all derivatives are approximated to second order except the cross derivatives which are approximated to first order. This approximation is justified by the fact that the parameters are reasonably decoupled, and it allowed us to reduce the CPU time consumption since second-order cross derivatives would require additional $n(n-1)/2$ simulations.

2.2 Simulation variable parameters

To model of the physics of the Universe, we introduced two categories of parameters that are varied in the simulations: cosmological parameters that describe the cosmological model in the simplest case of ΛCDM assuming a flat Universe with mass-less neutrinos, and astrophysical parameters that model the astrophysics within the IGM and the relation between temperature and density of the gas. A summary of all simulations performed to compute the coefficients of the Taylor expansion is given in appendix A.

2.2.1 Cosmological parameters

This first category contains four parameters: the amplitude of the matter power spectrum σ_8, the spectral index of primordial density fluctuations n_s, the matter density Ω_m and the Hubble constant H_0. The values for our central model are in agreement with the latest best-fit values from Planck [52], which we recall in table 1. We chose the range of variation for these parameters so as to include other recent constraints from the Wilkinson Microwave Anisotropy Probe seven years data [53], the South Pole Telescope data [54] and the SuperNova Legacy Survey three year data [55, 56], thus taking into account the fact that results from Planck for H_0 (respectively Ω_m) are low (respectively high) compared to other measurements. Central values at redshift $z = 0$ and range for each of the cosmological parameters are given in table 2. We also give in the appendix the values of σ_8 at redshift $z = 3.0$, (pivot redshift of Lyα power spectrum measurements of [1] and [30]), and of the shape parameter $\Gamma = \Omega_m h$, often used in the first Lyα studies.

Parameter	Best fit	68% limits
n_s	0.9624	0.9616 ± 0.0094
σ_8	0.8344	0.834 ± 0.027
Ω_m	0.3175	0.314 ± 0.020
H_0	67.11	67.4 ± 1.4

Table 1. Cosmological parameters values from Planck temperature power spectrum alone. We give best fit and 68% confidence limits.

2.2.2 Astrophysical parameters

This second category includes two redshift-dependent parameters that describe the temperature-density relation of the IGM for $\rho/\langle\rho\rangle \leq 10$:

$$T(\rho, z) = T_0(z) \times \left(\frac{\rho}{\langle\rho\rangle}\right)^{\gamma(z)-1}, \tag{2.2}$$

where ρ is the baryonic density, $T_0(z)$ is a normalization temperature and $\gamma(z)$ a logarithmic slope. At the post processing step (cf. 3.5), we scaled the effective optical depth $\tau_{\rm eff} = -\ln(\langle F\rangle) = -\ln(\langle e^{-\tau}\rangle)$, where F is the flux and τ is the optical depth, so that it followed a power law $\tau_{\rm eff}(z) = \tau_A \times (1+z)^{\tau_S}$. We allow for different mean flux normalizations and evolutions with redshift by varying the parameters τ_A and τ_S.

In the absence of a clear consensus on the heating history of the IGM, we took the $T(\rho)$ measurements from Becker et al. [57] assuming $\gamma = 1.3$ as our central model, and we chose a wide variation around these values so that other recent measurements [58–60] fall into the explored range. The evolution with redshift of $\gamma(z)$ and $T_0(z)$ in our simulations is therefore designed to reproduce the $T(\rho)$ measurements presented by Becker et al. [57] through an adaptation of the cooling routines in the simulation code. Thus we only need to fix those two parameters at a given redshift, in our case $z = 3.0$, which corresponds to the central redshift of our study. In practice, we do not set $T_0(z = 3)$ and $\gamma(z = 3)$ but instead use two internal code parameters, `AMPL` and `GRAD`, that alter the amplitude and density dependence of the photo-ionization heating rates, such that $\epsilon_f = \texttt{AMPL} \times \delta^{\texttt{GRAD}} \times \epsilon_i$ where ϵ's are the heating rates and δ is the over-density. T_0 and γ are evaluated after the simulations have run, as explained at the end of section 3. Given the one-to-one correspondence between (T_0, γ) and (`AMPL`, `GRAD`), we prefer to keep on quoting T_0 and γ since these parameters have a physical meaning and can be compared to other studies.

There are also a number of additional astrophysical effects which will impact on the Lyman-alpha flux power spectrum which we have not considered in this work. For example, the hydrogen reionisation history will alter the pressure smoothing scale of gas in the IGM, particularly at redshifts approaching the tail-end of the reionisation at $z \sim 6$ [61]. Galactic winds will impact on the distribution of HI around haloes, although the filling factor of the winds will be small at $z > 2$ [62]. Finally helium reionisation may induce fluctuations in the ionisation and thermal state of the IGM at $z \sim 3$, impacting on the power spectrum at large scales [63]. Note, however, these will generally influence the power spectrum to a lesser extent than the variations in the effective optical depth and the thermal state of the IGM we consider here, see e.g. [64]. As a result, we do not include them within our current analysis. These second order effects will nevertheless be important to consider for precision measurements with the Lyman-alpha forest power spectrum, although note that modelling these self-consistently will still ultimately require radiation hydrodynamics calculations which are currently computationally prohibitive.

2.2.3 Grid values

The central values and variation ranges of the parameters of our study are summarized In table 2. With six varying parameters, this represents a total of 28 cosmological simulations in our grid. We also varied the mean flux as explained in the previous section, but this was done a posteriori and did not require any additional simulation.

Parameter	Central value	Range
n_s	0.96	± 0.05
σ_8	0.83	± 0.05
Ω_m	0.31	± 0.05
H_0	67.5	± 5
$T_0(z = 3)$	14000	± 7000
$\gamma(z = 3)$.	1.3	± 0.3
τ_A	0.0025	± 0.0020
τ_S	3.7	± 0.4

Table 2. Central values and variation ranges of the cosmological parameters for our simulation grid.

3 Pipeline

All the components of our simulation work flow are represented on figure 2. The first part of the pipeline is the production of the initial condition snapshot. This is done in the linear approximation with perturbations treated up to second order. The simulations are then performed using both N-body and hydrodynamic (SPH) treatments. The post-processing stage takes the result of the simulations and computes the power spectra that will be compared to data through the Taylor expansion described earlier.

The products of our suite of simulations are obtained at 13 predefined redshifts, equally spaced every $\Delta z = 0.2$ from $z = 2.2$ to 4.6. Our selection of redshifts reflects the possibilities of current and forthcoming large-scale spectroscopic surveys. In SDSS, the lower bound results from the UV cut-off of CCDs at $\lambda \sim 350$ nm that prevents the observation of Lyman-α below $z \sim 2.2$. The upper bound results from the quasar luminosity function that peaks near $z \sim 2$ and drops significantly at $z > 3$. The density of QSOs at $z > 4$ is of order 0.3 per square degree to a limiting magnitude $g < 22$ as is the case for SDSS-III/BOSS, and even to $g < 23$ as expected for the future DESI survey, only reaches a density of 2 per square degree. This is less than an order of magnitude smaller than at $z \sim 2$ [65].

3.1 CAMB

The Code for Anisotropies in the Microwave Background (CAMB)[3] [66] is a numerical Boltzmann code written in Fortran 90. It is a parallelized line-of-sight integration code developed from CMBFAST [67] and COSMICS [68], which is widely used (and thus tested) to calculate not only the lensed cosmic microwave background temperature and polarization spectra but also linear matter power spectra for different species of particles (in our case baryons and dark matter).

CAMB is here used to compute the transfer functions and linear power spectra that will be used in the next step to compute the initial displacement of particles.

[3] http://camb.info

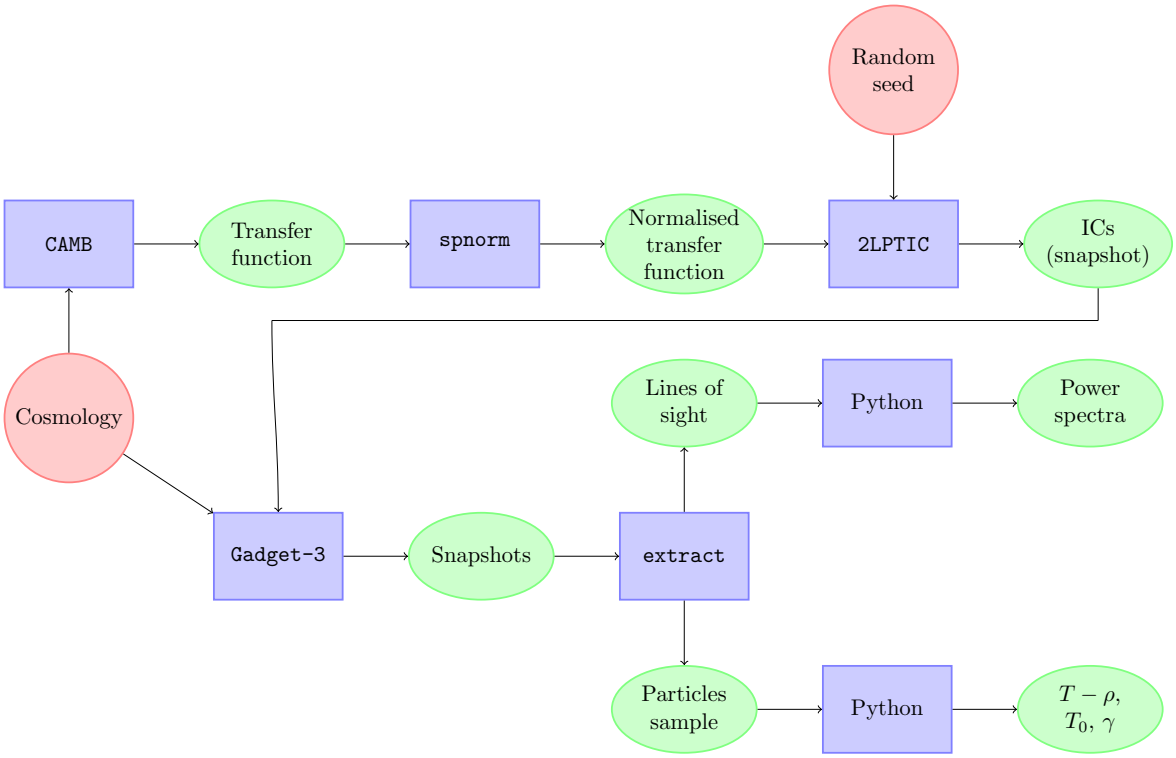


Figure 2. Our simulation pipeline: red circles represent input from the user, blue rectangles are software packages and scripts and yellow ellipses correspond to outputs from the software.

3.2 2LPT

All our simulations are tuned to obtained a given σ_8 at $z = 0$. This is done with the `spnorm` Python script that rescales the total matter power spectrum $P_S^{\rm CAMB}$ issued from CAMB before generating the initial conditions such that

$$P_S = P_S^{\rm CAMB} \times \left[\frac{\sigma_8(z_i)}{\sigma_8^{\rm CAMB}(z_i)} \right]^2 , \tag{3.1}$$

where z_i is the redshift at which the initial conditions are run, $\sigma_8^{\rm CAMB}$ is the value of σ_8 obtained with CAMB for a chosen cosmological model, and

$$\sigma_8(z_i) = \frac{\sigma_8(z=0)}{\sigma_8^{\rm CAMB}(z=0)} \times \sigma_8^{\rm CAMB}(z_i) \, . \tag{3.2}$$

Thus, in the 2LPTIC code, the power spectra are taken from CAMB at $z = 0$ and scaled back to the initial redshift $z_{\rm i} = 30$ by explicitly forcing the simulation to achieve the desired value of σ_8 at $z = 0$. While `CAMB` includes radiation, this is not the case for `GADGET`. However, the impact of the radiation component is very small in terms of the matter power spectrum at the scales relevant for the present work. Consequently, the low-redshift evolution of the simulation reproduces the matter and transmitted flux power spectra in a regime in which the radiation contribution can be safely ignored. The rescaled power spectra are then used as input to the `2LPTIC`[4] code that provides initial conditions based on second-order Lagrangian Perturbation Theory (2LPT), rather than first-order (Zel'dovich

[4] http://cosmo.nyu.edu/roman/2LPT/

approximation). The choice of second-order precision initial conditions is motived by the discussion in Crocce et al. [69] and the fact that we also run cosmological simulations including neutrinos as a new particle type [70]. Indeed, because of their high velocity, neutrinos require initial conditions taken at rather low redshift in order to reduce Poisson noise [43, 71]. Initial conditions for all the grid simulations are run with the same seed.

3.3 Gadget-3

GADGET-3 (GAlaxies with Dark matter and Gas intEracT) is a massively parallel tree-SPH code for cosmological simulations, originally developed by Volker Springel and collaborators [44, 72]. It is written in ANSI C, and uses the standardized message passing interface (MPI) along with several open-source libraries (GSL[5], FFTW[6]). Gravitational interactions are computed via a hierarchical multipole expansion using the standard N-body method, and gas-dynamics are followed with smoothed particle hydrodynamics (SPH); collisionless dark matter and gas are both represented by particles.

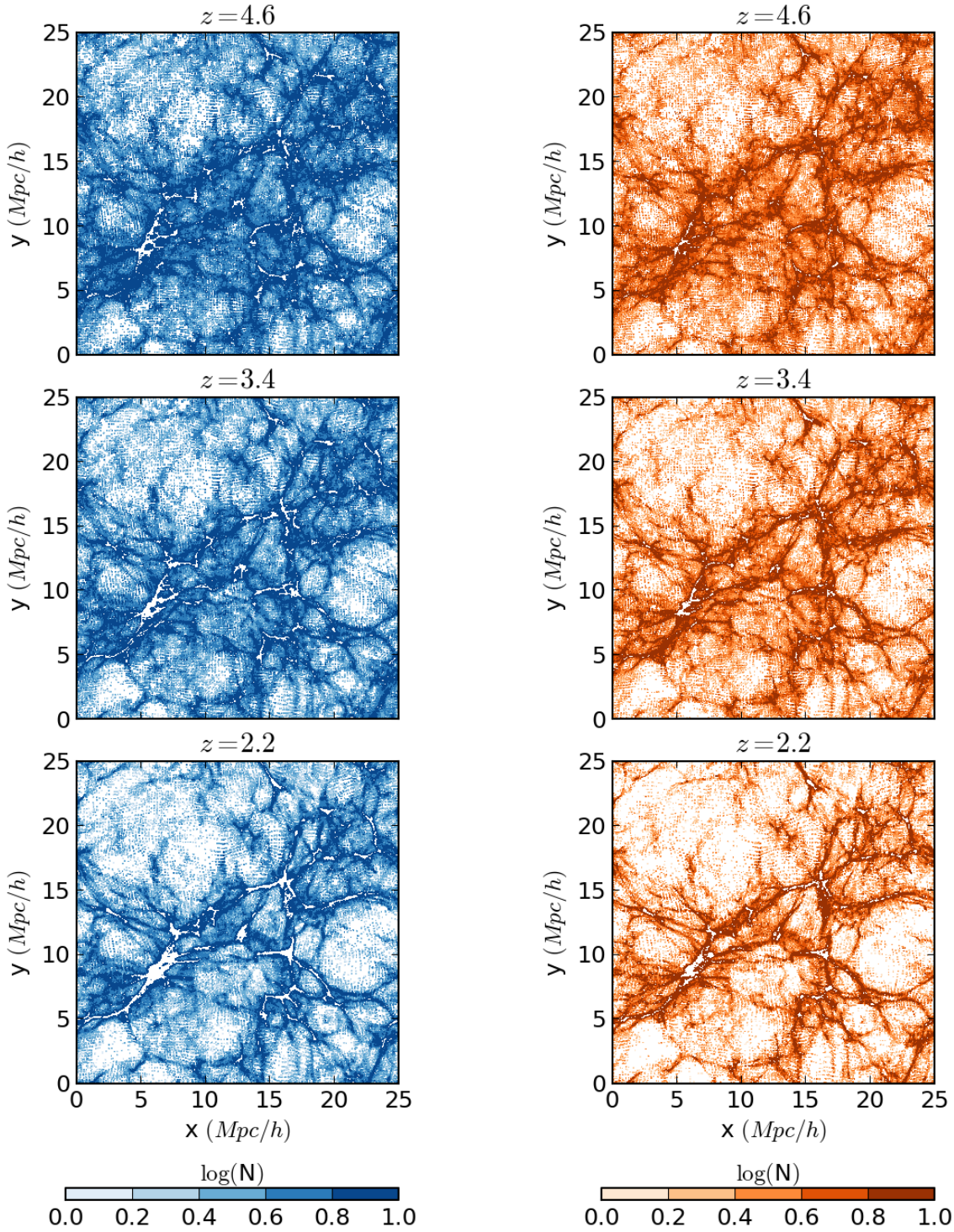
Since its original version (GADGET-1), the code underwent a series of improvements and optimizations over several years (GADGET-2 and 3), to maximize the work-load balance and the efficiency in memory consumption and communication bandwidth. In what follows, we briefly describe the key features of the code.

GADGET-3 follows a collisionless fluid with the standard N-body method, and an ideal gas with smoothed particle hydrodynamics (SPH). The code solves simultaneously for the dynamics of the collisionless component and of the ideal gas, both subject to and coupled by gravity in an expanding background space. The N-body implementation only differs from other cosmological codes by the accuracy of the gravitational field computation. A number of further physical processes have also been implemented in GADGET-3, from radiative cooling/heating physics to non-standard dark matter dynamics, star formation and feedback. In figure 3, we present the evolution of a filament with redshift and in figure 7 we show the image of a snapshot made with splotch [7]. Such realizations can be used for visual confirmation before quantitative analysis as well as for public outreach and education.

Several optimization strategies have been added in GADGET-3. These include a Peano-Hilbert space decomposition, a massively parallel version of the Fast Fourier Transform library, the possibility of splitting the simulation data across several files (to facilitate and speed-up the input/output process), and the fact that the code can be run on an arbitrary number of processors. In its current version, GADGET-3 is highly efficient in memory consumption (it allocates up to 80 bytes per particle) and communication bandwidth, is versatile and flexible, accurate and fast. Another important aspect is the scalability of the code, i.e. its performance when the number of processors is increased, which has currently been tested up to 16,000 cores.

We started all our simulations at $z = 30$ with initial conditions based on second-order Lagrangian perturbation theory[69], and adopted the same gravitational softening for the different species considered (i.e. gas, dark matter, stars), which however varies with the length of the box and the size of the mesh chosen. Specifically, we set the gravitational softening length to 0.8 kpc/h for the simulation having 25 Mpc/h boxsize and resolution 2×768^3, while the softening is 3.25 kpc/h for the other two runs, i.e., the 25 Mpc/h boxsize and 2×192^3 resolution, and the 100 Mpc/h boxsize and 2×768^3 resolution. We used the 'QUICKLYA' routine in GADGET-3 to simulate the Lyman-α forest, assuming the gas of primordial composition with a helium mass fraction of $Y = 0.24$. We neglect metals and the evolution of elementary abundances, as well as feedback processes and galactic winds.

[5] http://www.gnu.org/software/gsl/
[6] http://www.fftw.org/
[7] http://www.mpa-garching.mpg.de/~kdolag/Splotch



(a) Baryonic gas

(b) Dark matter

Figure 3. Slice of baryon and dark matter snapshots (2.5 Mpc h^{-1} depth), at three different redshifts, extracted from a simulation with 192^3 particles per type in a $(25\,\mathrm{Mpc}\,h^{-1})^3$ box. As expected, there are very few differences between the distributions for the two types of particles. Color represents particle number density. Out-of-scale densities (whether underflow or overflow) are white.

Along the lines of Viel et al. [24], we adopted a simplified criterion for star formation: all gas particles whose overdensity with respect to the mean is above 1000 and whose temperature is less than 10^5 K are turned into star particles immediately.

3.4 `extract`

The `GADGET-3` snapshots contain various fields among which the position **p** and velocity **v** for both dark matter and gas particles. It also contains fields that are specific to the SPH treatment of gas particles: internal energy U, density ρ, electron fraction N_e, hydrogen fraction $N_{\rm H}$ and smoothing length h. We use these fields to extract two samples:

- **a particle sample:** we extract a subsample of particles to study the temperature-density relation. For each particle the temperature is derived with the formula

$$k_B T = U \times (\gamma - 1) \times \mu M_{\rm H}\,, \tag{3.3}$$

  where $\mu = 1/(X_{\rm H}(0.75 + N_e) + 0.25)$. γ is the adiabatic index (5/3 for monoatomic gas), $M_{\rm H}$ is the mass of an hydrogen atom, k_B is the Boltzmann constant and $X_{\rm H}$ is the hydrogen fraction by mass. Figure 10 illustrates typical temperature-density diagrams obtained from this particle sample.

- **a line of sight sample: following the traditional procedure in one-dimensional flux power studies [21, 73]**, we extract lines of sight (LOS) from the simulation cube choosing random origin and axis. For each pixel of each LOS, we derive density ρ, temperature T, peculiar velocity v and optical depth τ, all for H I only using the SPH equation:

$$A(\mathbf{r}) = \sum_j m_j \frac{A_j}{\rho_j} W\left(\left|\mathbf{r} - \mathbf{r}_j\right|, h_j\right) \tag{3.4}$$

  where A is a scalar quantity, **r** a position in the cube, h the smoothing length, and W a kernel function. The index j runs on all particles. We use the 3D cubic spline kernel:

$$W(q_j) = \begin{cases} [1 + q_j^2(-1.5 + 0.75 q_j)].\frac{1}{\pi} & \left|q_j\right| \leq 1 \\ [0.25(2 - q_j)^3].\frac{1}{\pi} & 1 < \left|q_j\right| \leq 2 \\ 0 & \left|q_j\right| \geq 2 \end{cases} \tag{3.5}$$

  where $q_j = \left|\mathbf{r} - \mathbf{r_j}\right| / h_j$. These LOS are not mock spectra, in the sense that they do not match any properties (such as noise, resolution, metals absorption, ...) of observational data. The quantity of particular interest for our study is the optical depth for H I, from which we compute the transmitted flux for each pixel.

3.5 Post-processing

The post-processing stage allows us to extract two categories of outputs. The first one is the large-box high-resolution power spectrum that is derived by an appropriate combination of the power spectra from 3 lower-resolution or smaller-box simulations, using the splicing technique described in section 5. At this stage, we fix the photo-ionization rate (or equivalently the UV flux) by requiring the effective optical depth at each redshift to follow the empirical power law $\tau_{\rm eff}(z) = \tau_A \times (1+z)^{\tau_S}$, where $\tau_A = 0.0025$ and $\tau_S = 3.7$ in agreement with observations [74]. The rescaling coefficients, determined

independently for each redshift bin using all the line-of-sight pixels, are typically between −20% and +20%. We perform this normalization a posteriori since it is computationally much cheaper than finding and fixing the appropriate photo-ionization rate a priori for each of the simulations. As explained in [75], however, this is justified by the fact that when the gas is highly ionized and in photo-ionization equilibrium, as is the case for the Lyman-α forest, the total heating rate per unit volume is independent of the amplitude of the UV flux. Gas dynamics can thus be considered not to be affected by the UV flux. The power spectrum is then computed from the scaled flux, and averaged over all lines of sight.

The second category results from the particle sample. It is used to derive the parameters $T_0(z)$ and $\gamma(z)$ in the IGM. This is performed by estimating the location of the most populated region of the diagram using the mode of the 2D distribution, for particles lying in the region defined by $\log(\delta) \in [-0.5, 0.0]$ and $\log(T/1\,\mathrm{K}) < 5.0$, with $\delta = \rho/\langle\rho\rangle$. Given the large tail of particles toward the high temperature regions where clusters of galaxies reside, in particular at low redshift, the mode was preferred to the mean since it is not affected by the precise choice of the (δ, T) bounds used to define the IGM. We estimate the mode by taking bins of 1000 K and computing the barycenter of the five highest bins. A linear fit is then performed using these points.

4 Convergence tests

We base our minimal requirements for the resolution and box size of our simulations on the largest currently-available spectroscopic survey: SDSS-III/BOSS [32]. Those requirements are driven by the extension of the Lyman-α forest that can be probed experimentally and by the measurement errors on the power spectrum , which set the convergence levels to be achieved in the simulations.

The quasar coadded spectra provided by the SDSS pipeline [76] are computed with a constant pixel width of $\Delta v = 69\,\mathrm{km\,s^{-1}}$. The largest mode is bounded by the Nyquist-Shannon limit at $k_{\mathrm{Nyquist}} = \pi/\Delta v = 4.5 \times 10^{-2}\,(\mathrm{km/s})^{-1}$. Instrumental constraints, however, make this theoretical limit very difficult to obtain with reasonable precision from data, and the largest mode measured in BOSS data by Palanque-Delabrouille et al. [1] is $k_{\mathrm{max}} = 2.0 \times 10^{-2}\,(\mathrm{km/s})^{-1}$. The smallest mode is driven by the extension of the Lyman-α forest which lies between the Lyman-α and Lyman-β emissions respectively at 1216 Å and 1026 Å. The exploitable Lyman-α forest, however, is smaller than the separation of the two emission peaks due to their respective widths. Palanque-Delabrouille et al. [1] computed the 1D power spectrum from forest lengths corresponding to a third of the total available range in order to restrain the redshift span to $\Delta z = 0.2$ at most. This led to $k_{\mathrm{min}} \sim 1.0 \times 10^{-3}\,(\mathrm{km/s})^{-1}$. We therefore consider simulations that should cover the minimal range $1 \times 10^{-3}\,(\mathrm{km/s})^{-1} < k < 2 \times 10^{-2}\,(\mathrm{km/s})^{-1}$, which corresponds approximately to $0.1\,(\mathrm{Mpc}/h)^{-1} < k < 2\,(\mathrm{Mpc}/h)^{-1}$ at $z \sim 3$.

In numerical simulations, the two relevant parameters are the size of the box L that determines the smallest k-mode ($k_{\mathrm{min}} = 2\pi/L$), and the ratio $N^{1/3}/L$, where N is number of particles, that drives the largest k-mode. One may note that due to the computational algorithms used nowadays in simulations, such as smooth-particles hydrodynamics (SPH) or adaptive mesh refinement (AMR) in which "resolution follows density", particle spacing in high-density regions will be significantly smaller than $L/N^{1/3}$. Because the 1D power spectrum results from an integral over the 3D power spectrum up to $k = \infty$ (cf. Eq. 1.1), the resolution of the simulations has to be of the size of the smallest structures in the transverse direction. For structures in local hydrostatic equilibrium, this would be the Jeans scale, of order a few 100 kpc at $z = 3$. In an SPH approach, over-dense regions are sampled with much higher spatial resolution than average. Under-dense regions, on the other hand, might not necessarily be in local hydrostatic equilibrium. The decisive solution to ensure that the simulations do resolve the relevant structures is therefore to perform convergence tests.

The simulations used for the convergence test are all run with the same random seed and with the following cosmological parameters: $(\Omega_m, \Omega_b h^2, h, \sigma_8, n_s) = (0.31, 0.021, 0.675, 0.83, 0.96)$. We ran two sets of simulations: the first set with simulations having the same box size L of $\sim 20\,\mathrm{Mpc}/h$ but changing the particle loading N^3 and therefore the mass resolution, the second with simulations having the same mass resolution but varying volumes, keeping L/N fixed at ~ 0.12. These two sets are listed in table 3 and the results at three different redshifts are presented in figures 4(a) and 4(b). Hereafter, we will use the notation (L,N) to represent a simulation with N^3 particles for each species (dark matter and baryons) in a box of size $L\,\mathrm{Mpc}\,h^{-1}$ on a side.

Mass-resolution test	Box-size test
(20,1024)	**(120,1024)**
(20,768)	(90,768)
(20,512)	(80,683)
(20,384)	(60,512)
(20,192)	(20,171)

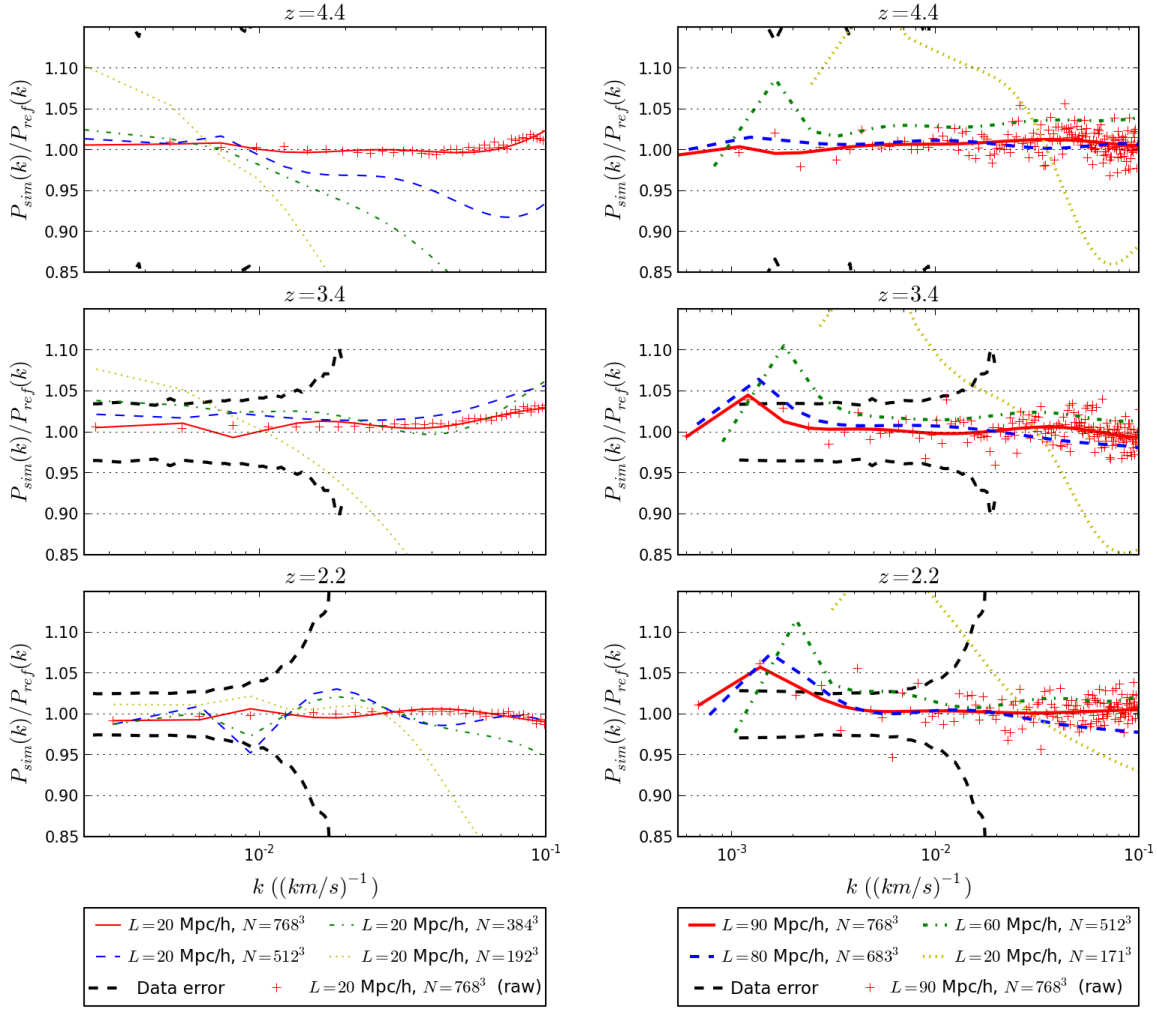
Table 3. The two sets of simulations used for convergence tests, with the reference simulation indicated in bold. (L, N) refers to a simulation with N^3 particles per species in a box of size $L\,\mathrm{Mpc}\,h^{-1}$ on a side.

These convergence tests are more stringent than what has been done before, justified by our aim to use our simulation suite for comparison to data of higher quality. For instance, to probe the effect of the box size, [11] compared to a reference simulation with $(L, N) = (120, 200)$ and thus $L/N \sim 0.60$, and [77] to a reference simulation $(80, 400)$ i.e., $L/N \sim 0.20$. This is to be compared to our L/N of 0.12. As regards the convergence on the mass-resolution, we explored a similar range of mass-resolutions as [77] (in contrast, [11] restricted to a minimum mass per particle 3 times larger), but using a $20\,\mathrm{Mpc}\,h^{-1}$ box instead of $10\,\mathrm{Mpc}\,h^{-1}$.

The most difficult redshifts at which to achieve convergence are those at $z > 3$, since the mean flux level becomes very small at such epochs and under-dense regions, which are less well sampled than average in an SPH framework, are producing absorption. High redshift bins, however, are very important since gravitational collapse tends to suppress the differences in the linear-theory power spectra. These bins therefore highlight primordial differences between the matter power spectra resulting from different contributions of the various cosmological constituents. Some convergence problems can also arise at $z \sim 2$ due to the fact that strong systems, which are very non-linear, might be simulated inadequately due to cosmic variance or lack of resolution. Low-redshift bins are also those where the measurements from QSO spectra have the smallest statistical error bars, making the convergence criteria tighter.

4.1 Mass resolution

We computed the ratio of the power spectra of each of the simulations listed in the first column of table 3 to the power spectrum of the (120,1024) simulation. The results presented in figure 4(a) show that an excess of power on large scales (small k) and a lack of power on small scales (large k) appear with decreasing resolution. As expected, this effect is stronger at higher redshift where the Lyman-α forest probes low density regions, which are less well resolved in the SPH treatment since it is the mass (and not the spatial) resolution that is kept fixed. Further details about this effect can be found in Bolton and Becker [77].



(a) Mass-resolution tests, the reference simulation has $L = 20\,\mathrm{Mpc}\,h^{-1}$ and 2×1024^3 particles.

(b) Box-size tests, the reference simulation has $L = 120\,\mathrm{Mpc}\,h^{-1}$ and 2×1024^3 particles.

Figure 4. Convergence tests for mass resolution and box size at three different redshifts. The curves are 5th-order polynomial functions fitted to the data for better visibility. All values of the power spectrum ratio are shown for the (20,768) (left) and the (90,768) (right) cases. The dashed black curves illustrate $1\sigma_{\mathrm{stat}}$ uncertainties measured in BOSS data. Data uncertainties exceed the plot boundaries at $z = 4.2$.

The dashed curves in figure 4(a) illustrate the level of the $1\sigma_{\rm stat}$ statistical uncertainties observed in the BOSS analysis [1] at each redshift. At $z = 4.2$, the experimental uncertainties are larger than the maximum ±15% departure allowed on the plot and no longer appear.

Simulations with a mass resolution at least as good as for the (20,512) simulation all deviate by less than 2.5% from the highest mass-resolution power spectrum over our minimal k-range. This corresponds to a mean mass per particle of $M = 2.2 \times 10^5 M_\odot.h^{-1}$. Extending to $k_{\rm max} = 0.1\,({\rm km/s})^{-1}$, the (20,512) simulation deviates by $\sim 10\%$ at the largest redshift,.

4.2 Box size

The results of figure 4(b) show that the box size has an effect on all scales, and not only on the large scales that approach the Nyquist limit. This is due to the non-linear coupling of modes during gravitational evolution, and to the fact that even on scales close to the box size, mass-fluctuations are not fully linear.

As before, the dashed curves in figure 4(b) illustrate the level of the $1\sigma_{\rm stat}$ data uncertainties at each redshift. To reach $k_{\rm min} = 1.0 \times 10^{-3}\,({\rm km/s})^{-1}$, we see that we need a box size of at least $90\,{\rm Mpc}\,h^{-1}$. The most significant constraint comes from the largest scales that cannot be probed (or not with adequate precision) otherwise.

4.3 Summary of convergence requirements

In conclusion, the ideal simulation for our study should use a $\sim\ 100\,{\rm Mpc}\,h^{-1}$ box and a mass resolution roughly equivalent to a (20,614) simulation, which translates into 3072^3 particles of each species. The mean mass of a gas particle is then $M = 1.2 \times 10^5 M_\odot.h^{-1}$.

Although convergence tests are specific to each problem and each statistical property for which convergence is sought, we can briefly compare to the results obtained by other studies. To infer the dark matter power spectrum from the Lyman-α forest in high-resolution QSO absorption spectra covering $0.003 < k < 0.03\,({\rm km/s})^{-1}$, Viel et al. [11] chose a $(60, 400)$ simulation, i.e. a mass per gas particle of $\sim 4 \times 10^7 M_\odot.h^{-1}$. To resolve the high redshift Lyman-α forest in smoothed particle hydrodynamics simulations, a problem similar to our own, Bolton and Becker [77] found that a box size of at least $40\,{\rm Mpc}.h^{-1}$ is preferable at all redshifts. They also found that while a mean gas particle mass $M_{\rm gas} \leq 1.6 \times 10^6 M_\odot.h^{-1}$ is required at $z = 2$, a mass resolution at least 8 times better is needed at $z = 5$, i.e. $M_{\rm gas} \leq 2 \times 10^5 M_\odot.h^{-1}$. Our requirements are thus more stringent than selected in past, both in terms of box size and mass resolution.

Several tens of such simulations, as needed to compute our grid of cosmological simulations, would require several tens of millions of hours to be run, which is not an acceptable computational time. We address and solve this issue with the splicing technique presented in the next section.

5 Splicing

In the previous section we have estimated that simulating a flux power spectrum covering the range $k = 1 \times 10^{-3}\,({\rm km/s})^{-1}$ to $k = 2 \times 10^{-2}\,({\rm km/s})^{-1}$) with a unique simulation at sufficient precision for every redshift in the range $2.2 < z < 4.6$ requires $N = 3072^3$ particles of each species in a box of size $L = 100\,{\rm Mpc}\,h^{-1}$. To obtain power spectra of equivalent resolution and box size in a reasonable computational time, we use the technique described in [78]. In this method, competing demands of large box size and high resolution are solved by splicing together the power spectra from pairs of large and small box simulations, using $L = 100\,{\rm Mpc}\,h^{-1}$ for the large-scale power, and $L = 25\,{\rm Mpc}\,h^{-1}$ for the small-scale power, both with $N = 768^3$. One must then correct the large box size simulation for the lack of resolution, and the small box size for the lack of non-linear coupling between the highest

and the lowest k-modes. The corrections are computed using a transition $(25, 192)$ simulation that has same resolution as a $(100, 768)$ and same box size as a $(25, 768)$.

One needs to distinguish three regimes when computing the full power spectra:

- $\mathbf{k} < \mathbf{k_{min,25}}$, where $k_{min,25} = 2\pi/25\,\mathrm{Mpc}\,h^{-1}$ is the minimum k present in a $L = 25\,\mathrm{Mpc}\,h^{-1}$ box. The spliced flux power P_F is the flux power $P_{F,100,768}$ of the $(100, 768)$ simulation here taken as our reference, corrected for its low-resolution by a k-independent factor evaluated at $k_{min,25}$:

$$P_F(k) = P_{F,100,768}(k) \times \frac{P_{F,25,768}(k_{min,25})}{P_{F,25,192}(k_{min,25})}\,.$$

  The possibility of using a constant factor for the largest k-modes has been tested in [78].

- $\mathbf{k_{min,25}} < \mathbf{k} < \mathbf{k_{Nyq,100}/4}$, where $k_{Nyq,100} = 768\pi/100\,\mathrm{Mpc}\,h^{-1}$ is the Nyquist wave number of the large box. In this regime we use a similar correcting ratio, but taken at the wave number k at which the flux power is calculated:

$$P_F(k) = P_{F,100,768}(k) \times \frac{P_{F,25,768}(k)}{P_{F,25,192}(k)}\,.$$

  This is mathematically equivalent to considering the high-resolution simulation (25,768) as our reference, and correcting it for its small box size.

- $\mathbf{k} > \mathbf{k_{Nyq,100}/4}$. At these large k-modes, the resolution correction is no longer a small factor. We thus take the $(25, 768)$ simulation as our reference, and correct for its limited box size by a k-independent factor evaluated at the fixed splicing point $k = k_{Nyq,100}/4$:

$$P_F(k) = P_{F,25,768}(k) \times \frac{P_{F,100,768}(k_{Nyq,100}/4)}{P_{F,25,192}(k_{Nyq,100}/4)}\,.$$

The splicing technique is applied for each redshift at which we compute the power spectrum. We illustrate the method and its accuracy on figure 5, using a set of smaller-resolution simulations to enhance the contrast between the different power spectra, as well as to allow the comparison to a full resolution run (labelled "exact" on the figure) with 1024 particles in a $(100\,\mathrm{Mpc}/h)^3$ box. For this illustration, the large box-size, the large resolution and the transition simulations are (100,256), (25,256) and (25,64) simulations respectively. The spliced power spectrum obtained at $z = 3.0$ is presented on the left of figure 5, along with the exact power spectrum and the individual runs entering the splicing estimate. The correction coefficient with respect to the reference power spectrum in each regime of k-modes is shown on the right plot of figure 5. In the intermediate regime, the resolution correction increases towards smaller scales, reaching 40% for the set of simulations illustrated here. It shows less scale-dependence when taken as a box-size correction to the large-resolution power spectrum, but it is noisier since it requires taking the ratio of two simulations with different box sizes and thus different natural k-modes. Although the correction factors show discontinuities at the boundary where the simulation chosen as reference changes, the spliced power spectrum is continuous by construction. In the large-mode splicing regime, at $k > k_{Nyq,100}/4$, it is unclear whether a constant box-size correction or even any correction at all is indeed the optimal combination, since both the correction factor and the residuals are at the same level of about 0.95. This regime, however, is only probed by the medium resolution SDSS-III/BOSS data in the highest redshift bins where measurement uncertainties significantly exceed the splicing errors. Its optimization is thus beyond the scope of this paper.

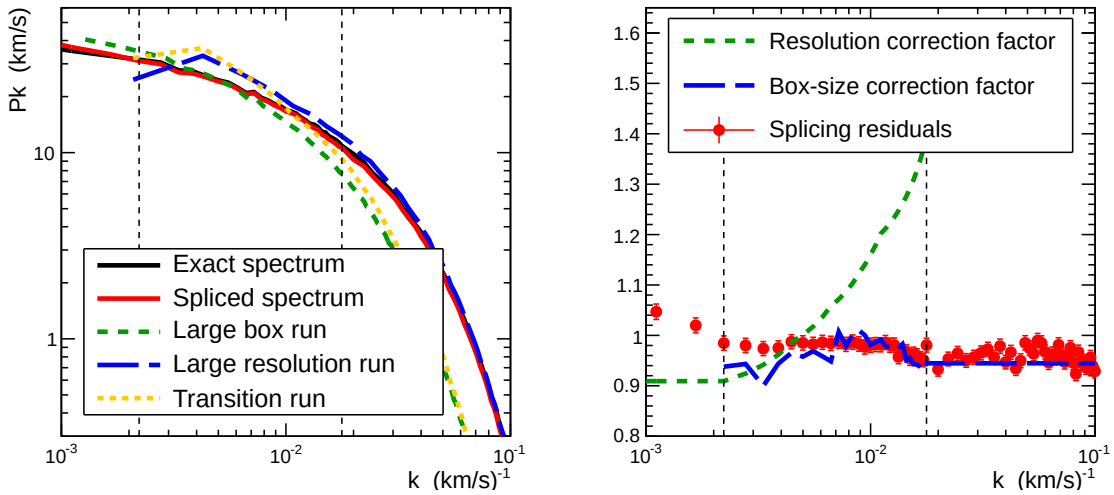


Figure 5. Illustration of the splicing technique with simulations of either 64^3 or 256^3 particles in a volume of $(25\,\mathrm{Mpc}/h)^3$ or $(100\,\mathrm{Mpc}/h)^3$. Dashed vertical lines illustrate the regime boundaries. Left: power spectrum at $z = 3.0$ for the exact (100,1024) run, the spliced technique and the individual components entering the splicing. Right: correction factors computed with respect to the power spectrum chosen as reference in each regime. Although the correction presents a discontinuity, the spliced power spectrum is continuous by construction. Splicing residuals are overlaid.

For the box size and resolution chosen for our simulation suite, the last regime begins at $k = 5.3\times10^{-2}\,(\mathrm{km/s})^{-1}$ for $z = 3.0$, which is beyond the maximum mode that can be reached with BOSS or eBOSS data. The maximum correction factor, obtained for $k = 2.0\times10^{-2}\,(\mathrm{km/s})^{-1}$, is thus smaller than in the previous illustration. It ranges from 22% at $z = 4.6$ to 5% at $z = 2.2$.

We estimate the accuracy of the technique from the splicing residuals, defined as the ratio of the spliced to the exact power spectrum. The splicing residuals show no dependence with redshift. The residuals at $z = 3.0$ are overlaid on the right plot of figure 5. In figure 6, they are plotted for $z = 2.2$ and $z = 4.2$, along with the statistical uncertainty at the same redshifts obtained in the most recent BOSS analysis [1]. Over the k-range of interest for BOSS data, the residuals have an average of -0.98 with an *rms* of 0.01. The largest excess is seen near $k = 10^{-3}\,(\mathrm{Mpc}/h)^3$. A simulation with a larger box size would be needed to reduce the splicing residuals further. For the purpose of this study, the splicing technique is accurate at the 2% level over the entire k-range of interest.

6 Results and discussions

Several checks were performed to validate our simulations. We first verified that the power spectrum of independent simulations obtained either with different cosmological parameters or different seeds is consistent with the power spectrum derived from the Taylor expansion of Sec. 2. We then present a comparison of our central model with the one-dimensional Lyman-α forest power spectrum obtained with BOSS by [1]. This allows us to quantify the agreement between our simulation and the measured power spectrum. Finally, we discuss some characteristics of our simulations. In particular, we describe the effect on the flux power spectrum of some of the parameters we have varied, and we show the $T-\rho$ diagrams from which we derive the two parameters $T_0(z)$ and $\gamma(z)$ that describe the IGM.

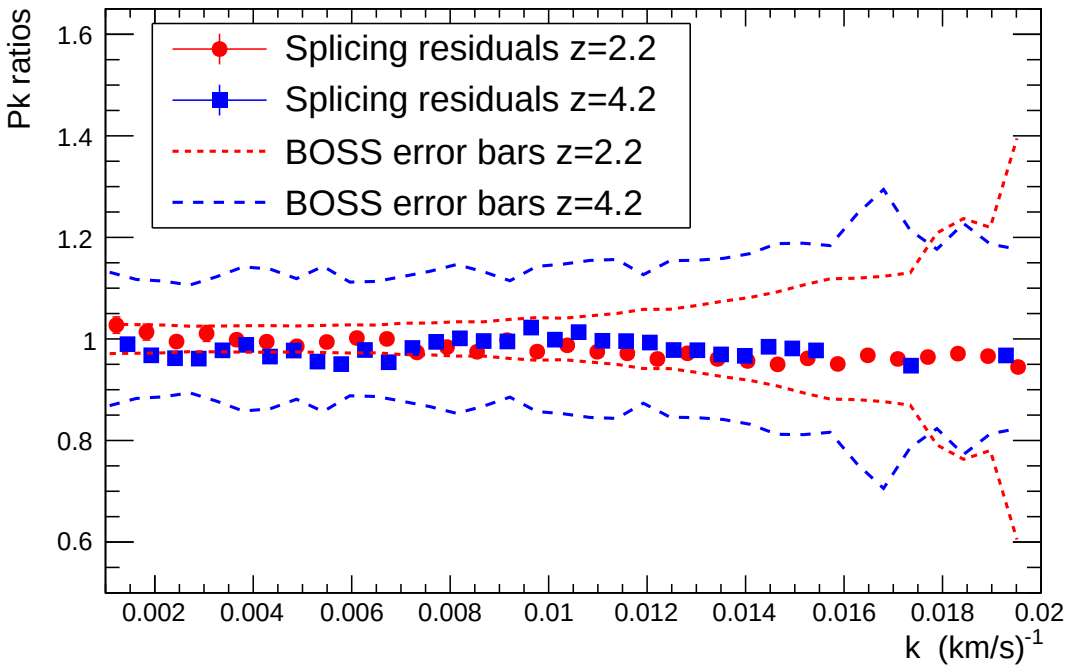


Figure 6. Residuals of the spliced to the exact power spectrum for $z = 2.2$ and $z = 4.2$. The dashed curves illustrates the level of statistical uncertainties in current data (1σ).

6.1 Assessment checks

We performed three categories of assessment checks. The first one verifies the statistical errors in the simulation relative to the number of lines of sight used to compute the one-dimensional power spectrum. The second one assesses the accuracy of our second-order Taylor expansion to model the power spectrum by comparing its prediction to the simulated power spectrum for simulations other than those used in the grid. The last category tests the impact of cosmic variance from the use of a given random seed.

For each simulation, the one-dimensional power spectrum was computed from 100.000 lines of sight. This large number is necessary to ensure that the simulation uncertainties remain about an order of magnitude smaller than current most precise data measurements (taken from [65]). We ensured that the simulation uncertainties were not artificially limited by an oversampling of the simulated volume: this was done by considering different numbers n of lines of sight (from 5.000 to 100.000). For each redshift and mode, the power spectrum value is taken as the mean over the n lines of sight and the uncertainty on the mean as the *rms* of the distribution divided by $\sqrt{n}$. We checked that the uncertainty on each point of the power spectrum scaled as the square-root of the number of lines of sight used to compute it, at better than the percent level.

To test the accuracy with which our Taylor expansion reproduces the power spectrum for different cosmologies, we performed simulations with input cosmological and astrophysical parameters different from those that were used to compute the derivatives. We tested the most relevant parameters for our study. One simulation was run with different n_s and σ_8, two others with different T_0 and γ, and a last set with all input parameters different from their values in the grid simulations. In each case, we computed the power spectra corresponding to the twelve redshift bins in the range

$z = [2.1 - 4.5]$. We then performed a simple fit of the six parameters (n_s, σ_8, Ω_m, H_0, AMPL, GRAD) using our second-order Taylor expansion as model.

The results are summarized in Tab. 4. The last column shows the fitted values over 100.000 lines of sight. The uncertainty is estimated as the *rms* of the distribution for each parameter fitted over 10 subsamples of 10.000 lines of sight each, divided by $\sqrt{10}$. The configurations were chosen so as to probe different relevant regions of the parameter phase space. These tests give results in excellent agreement with the input parameters. The level of accuracy achieved with these validation tests is 3 to 5 times better than the errors we expect on these parameters from a fit to data given the uncertainties of Palanque-Delabrouille et al. [1]. We note that the accuracy drops rapidly, however, as we test values outside the range that was used to compute the derivatives (cf. table 2), as indicated by the almost 3σ discrepancy on the fitted value of T_0 in the last test. Since our variation range was purposely chosen to be wide enough to include all recent results, this is not expected to cause any problem in the future. These checks thus demonstrate that our Taylor expansion adequately models the power spectrum for any set of input parameters within the range of table 2.

Finally, we produced a new simulation with the same parameters as our central simulation but using a different random seed to compute the initial conditions. Snapshots of the resulting gas distribution in the two cases are shown in figure 7. The derived power spectra for the two seeds are in excellent agreement on low scales. On the largest scales, the two power spectra can differ by up to 2 to 3σ at all redshifts, indicating a sample variance contribution to the uncertainty on the simulated power spectrum due to the fact that the simulation box has a size close to the largest modes measured. We again performed a simple fit on the power spectrum measured with the new seed using our Taylor expansion as model. The results are given in table 5. They show that cosmic variance has an impact on the power spectrum that exceeds the simulation statistical uncertainty and will therefore need to be included as a systematic uncertainty when comparing our model to data.

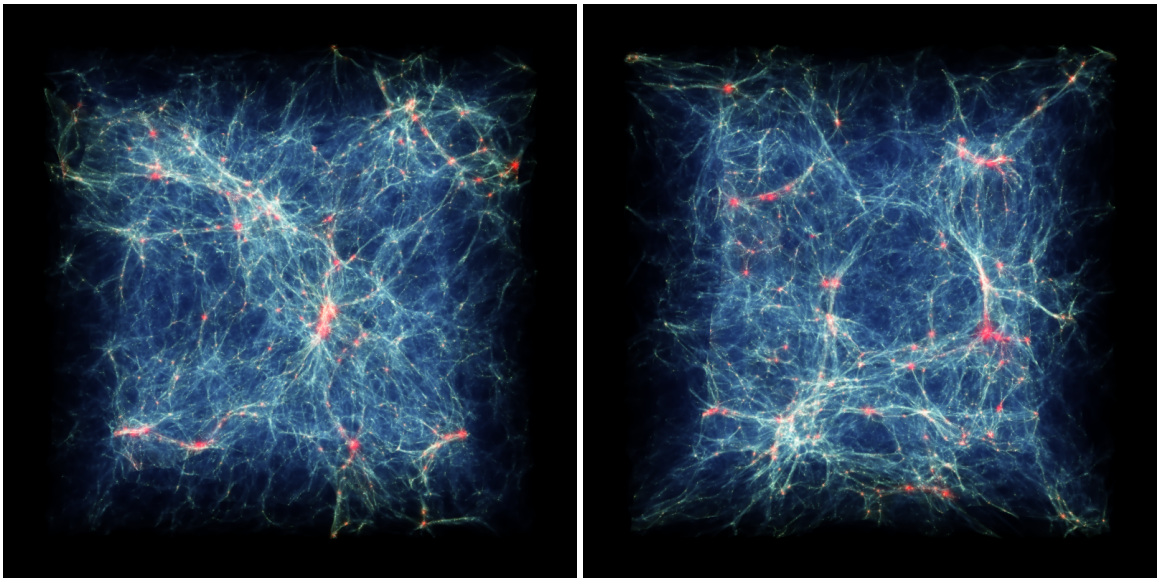


Figure 7. Visualisation using `splotch` of the baryonic gas from a `GADGET-3` snapshot taken at $z = 2.2$ for two simulations run with identical parameters but different random seeds to compute the initial conditions. Both simulations are using $2x768^3$ particles in a $(25\,\mathrm{Mpc}\,h^{-1})^3$ box. Color represents gas temperature (from blue to red) and density is mapped to intensity. Left for the random seed used for the grid, right for a different random seed.

Test configuration	Parameter	Input value	Fitted value
$n_s - \sigma_8$	n_s	0.93	0.931 ± 0.002
	σ_8	0.85	0.846 ± 0.008
	Ω_m	0.31	0.310 ± 0.003
	H_0	67.5	67.2 ± 1.1
	$T_0(z = 3)$	14000	14230 ± 600
	$\gamma(z = 3)$.	1.32	1.33 ± 0.03
$T_0 - \gamma$	n_s	0.96	0.961 ± 0.002
	σ_8	0.83	0.830 ± 0.009
	Ω_m	0.31	0.310 ± 0.003
	H_0	67.5	67.2 ± 1.1
	$T_0(z = 3)$	10000	10130 ± 200
	$\gamma(z = 3)$.	1.47	1.47 ± 0.02
$T_0 - \gamma$	n_s	0.96	0.961 ± 0.001
	σ_8	0.83	0.830 ± 0.008
	Ω_m	0.31	0.310 ± 0.003
	H_0	67.5	67.3 ± 1.1
	$T_0(z = 3)$	10000	10420 ± 300
	$\gamma(z = 3)$.	1.16	1.15 ± 0.02
All parameters	n_s	0.93	0.927 ± 0.002
	σ_8	0.86	0.848 ± 0.004
	Ω_m	0.30	0.300 ± 0.003
	H_0	66	67.7 ± 1.1
	$T_0(z = 3)$	10000	10470 ± 400
	$\gamma(z = 3)$.	1.16	1.19 ± 0.03
All parameters	n_s	0.935	0.935 ± 0.002
	σ_8	0.846	0.833 ± 0.005
	Ω_m	0.285	0.282 ± 0.004
	H_0	68	69.2 ± 1.1
	$T_0(z = 3)$	5840	6720 ± 320
	$\gamma(z = 3)$.	1.32	1.30 ± 0.02

Table 4. Comparison of the simulated parameters and the fitted parameters for different sets of input parameters.

6.2 Comparison to SDSS-III/BOSS DR9 data

In Palanque-Delabrouille et al. [1], the one-dimensional Lyman-α forest power spectrum is measured with 13,821 quasar spectra from SDSS-III/BOSS DR9 selected on the basis of their high quality, large signal-to-noise ratio, and good spectral resolution. The power spectra are measured over twelve redshift bins from $\langle z \rangle = 2.2$ to $\langle z \rangle = 4.4$, and scales from 0.001 $(\mathrm{km/s})^{-1}$ to 0.02 $(\mathrm{km/s})^{-1}$ (see figure 8).

In order to compare the measurements to the power spectrum obtained for our central model model, we normalized the simulation power spectrum at each redshift by constraining the effective optical depth to follow the power law evolution $\tau_{\mathrm{eff}}(z) = \tau_A \times (1 + z)^{\tau_S}$, where $\tau_A = 0.0025$ and $\tau_S = 3.7$. To account for the effect of the correlated Si III absorption, we correct the simulated power spectrum by a multiplicative term, $1 + a^2 + 2a\cos(vk)$ with $a = f_{\mathrm{Si\,III}}/(1 - \langle F \rangle(z))$ following the

Parameter	Input value	Fitted value
n_s	0.96	0.969 ± 0.004
σ_8	0.83	0.839 ± 0.005
Ω_m	0.31	0.28 ± 0.01
H_0	67.5	65± 1
$T_0(z = 3)$	14000	13750 ± 1000
$\gamma(z = 3)$.	1.32	1.38 ± 0.03

Table 5. Comparison of the simulated parameters and the fitted parameters for a different seed in the simulation.

suggestion of McDonald et al. [29]. The parameter $f_{\mathrm{Si\,III}}$ is adjusted and v is fixed at 2271 km/s. We model the imperfection of the resolution of BOSS spectra though a multiplicative term. Finally, we allow for imperfection in the noise estimate of the BOSS spectra with eight additive terms (one for each redshift bin).

Figure 8 illustrates the good agreement between the data and the simulations. Without any adjustment of the cosmological and astrophysical parameters, the χ^2 per number of degrees of freedom is already better than 1.2. The good agreement between data and simulation covers the whole redshift range, $z = [2.1 - 4.5]$, in contrast with the cosmological analysis described in Palanque-Delabrouille et al. [1] which was performed over the reduced redshift range $z = [2.1 - 3.7]$. This simple comparison demonstrates the improvement obtained with these simulations over the previous generation of simulations [24].

6.3 Power spectrum

Figure 9 illustrates the impact on the power spectrum of our four cosmological parameters. We compare the power spectrum computed from our best-guess model to the one obtained when varying each parameter, one at a time. We note that the dependence on the value of the four parameters is as expected according to their physical meaning. We briefly explain the different behaviors below.

The spectral index n_s represents the evolution of the primordial density fluctuations with respect to k through $\mathcal{P}(k) \propto k^{n_s-1}$. A larger n_s therefore increases the power at large k, as seen in the top left panel of figure 9.

The parameter σ_8 measures the *rms* amplitude of the linear matter density fluctuations today in spheres of size $8\,h^{-1}$ Mpc, and thus determines the normalization of the matter power spectrum. To first order, increasing the value of σ_8 therefore increases the power spectrum on all scales, as shown in the top right panel of figure 9. A slightly larger effect, however, is seen on large scales, since an excess in the amplitude of the fluctuations will favor the merging of small scale fluctuations, thus enhancing the power on larger scales. This tiny trend is purely non-linear and not expected in the evolution with σ_8 of the linear power-spectrum.

The present-day Hubble constant H_0 (in units of velocity/distance) allows the conversion from distance-space to k-space (units of inverse velocity). Therefore, if H_0 is increased, a given distance will correspond to a higher k, thus leading to an increase of power since the power spectrum, which is a decreasing function of k, is shifted to the right. This is indeed what is observed in the lower left panel of figure 9.

Finally, the parameter Ω_m quantifies the fraction of matter density in a flat Universe. Because Ω_m and the dark energy density Ω_Λ vary in opposite directions, a higher Ω_m delays the onset of dark energy domination, thus increasing the time available for structure formation. In addition, in a larger Ω_m universe, more structures (in particular small ones that would not collapse otherwise) will be

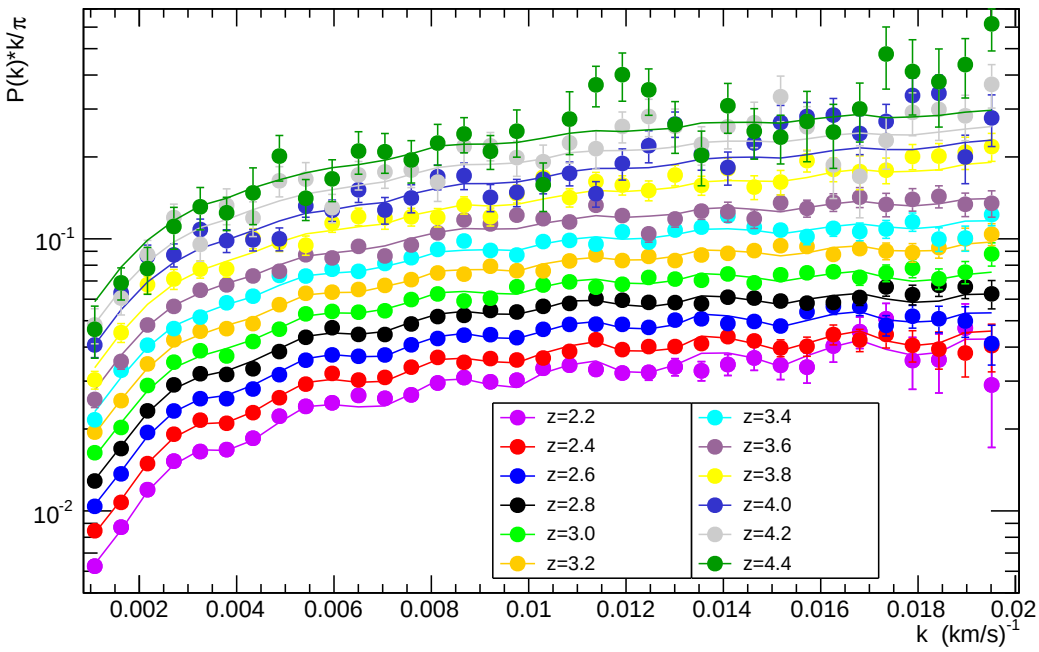


Figure 8. One-dimensional Lyman-α forest power spectrum obtained with BOSS spectra. The dots are the measured power spectrum by Palanque-Delabrouille et al. [1]). The solid line represents the power spectrum for our central model after adjustment of nuisance parameters to account for imperfect modeling of the instrumental parameters in the 1D power spectrum measurement.

formed, leading to an increase of the power spectrum, especially at high k. This is in agreement with the plots in the lower right panel of figure 9.

6.4 Density-temperature relation

In figure 10, we present the $T - \rho$ diagrams obtained from our central simulation at each of the snapshot redshifts. We can distinguish three different populations – the IGM, the stars, and the clusters – with a clear evolution with redshift for each of them.

The IGM is described by the low density and low temperature particles. This is the regime that dominates at high redshift. At later times, however, fewer and fewer particles reside in this part of the $T - \rho$ diagram, since they are captured by collapsing over-densities. We use this region to extract the $T_0(z)$ and $\gamma(z)$ parameters, displayed in figure 11, where they are compared to the measurements of Becker et al. [57].

The particles with higher temperature correspond to clusters and galactic gas. As expected, their density increases as structures are formed in the simulation box. They therefore become more prominent at lower redshifts.

In our simulations dedicated to the study of the IGM though the Lyman-α forest measurements, star formation undergoes a simplified treatment, which reflects as the sharp cut-off at $\log(\delta) \simeq 3$. Any particle sufficiently dense and cool is transformed into a star particle. The latter is used for gravity force calculation, but does not undergo SPH treatment like baryonic gas does.

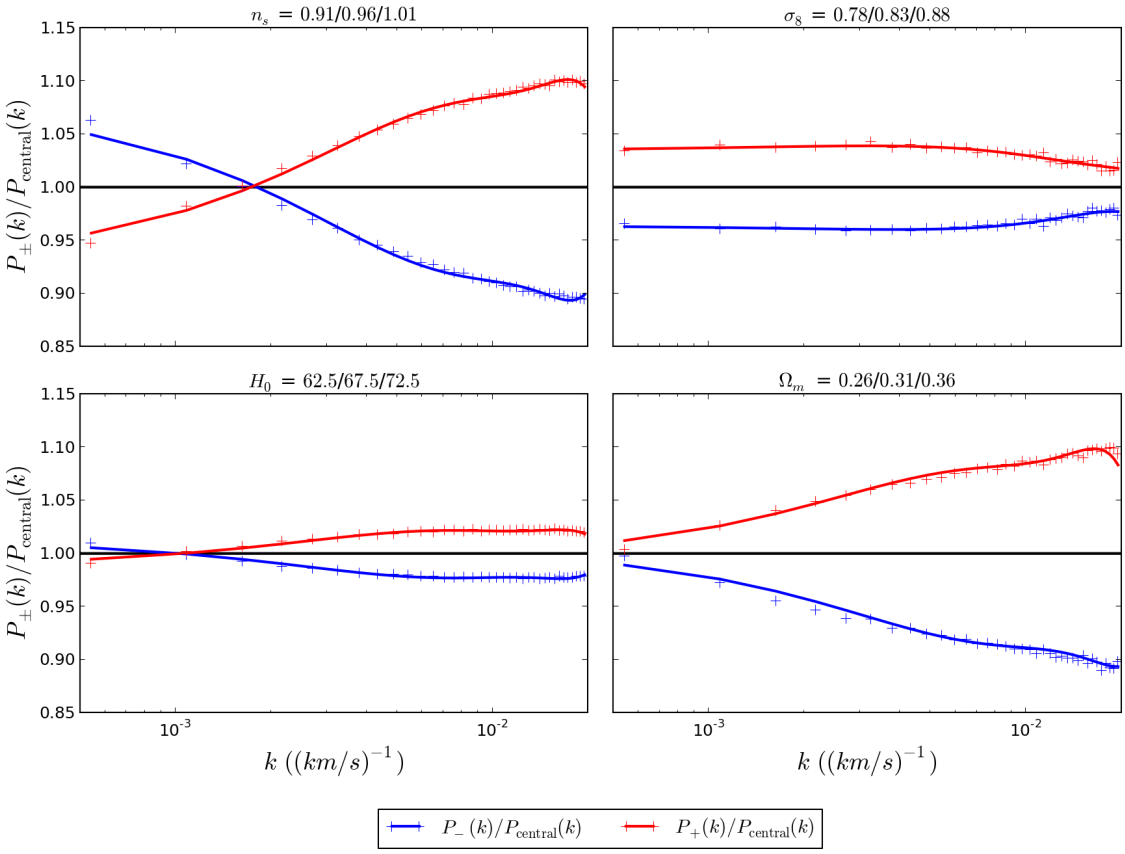


Figure 9. Effect of the parameters n_s, σ_8, H_0 and Ω_m on the power spectrum (limited to the k-range of our study) at $z = 3.2$. $P_+(k)$ and $P_-(k)$ refer to the power spectra extracted from the simulations using the upper and lower limit on each considered parameter respectively. The fit to the points is a 5$^{\rm th}$ order polynomial function.

7 Conclusions

We have designed and produced a grid of cosmological simulations, which may be used to extract constraints on cosmological parameters from Lyman-α surveys, whether current like SDSS-III/BOSS or future like SDSS-IV/eBOSS. These simulations cover the redshift range 2.2 – 4.6. They explore the cosmological parameters n_s, σ_8, H_0, and Ω_m over a large range centered on Planck measurements, as well as the astrophysical parameters T_0 and γ in a range covering most recent results.

Using the splicing technique of McDonald [78], we computed 1D power spectra from simulations equivalent to a 100 Mpc h^{-1} box filled with 3072^2 particles of each species (here dark matter and baryonic gas), abbreviated to (100, 3072) using our standard notation, from lower-resolution (100,768) and smaller box-size (25,768) simulations, combined using a transition (25,192) simulation. We show that the splicing technique allows us to approximate the exact full-resolution large-box simulation with an accuracy at the 2% level.

While one full-size high-resolution (100,3072) simulation would have required of order one million hours, one equivalent set of 3 simulations consumes an average of 70,000 hours of CPU time, with most of the time in the simulation pipeline (see figure 2) being spent on performing the hydrodynamical simulations. The data volume produced by each set is 1.6 terabytes. Therefore, the whole grid represents about 2 millions hours of CPU time and a volume of 45 terabytes of data.

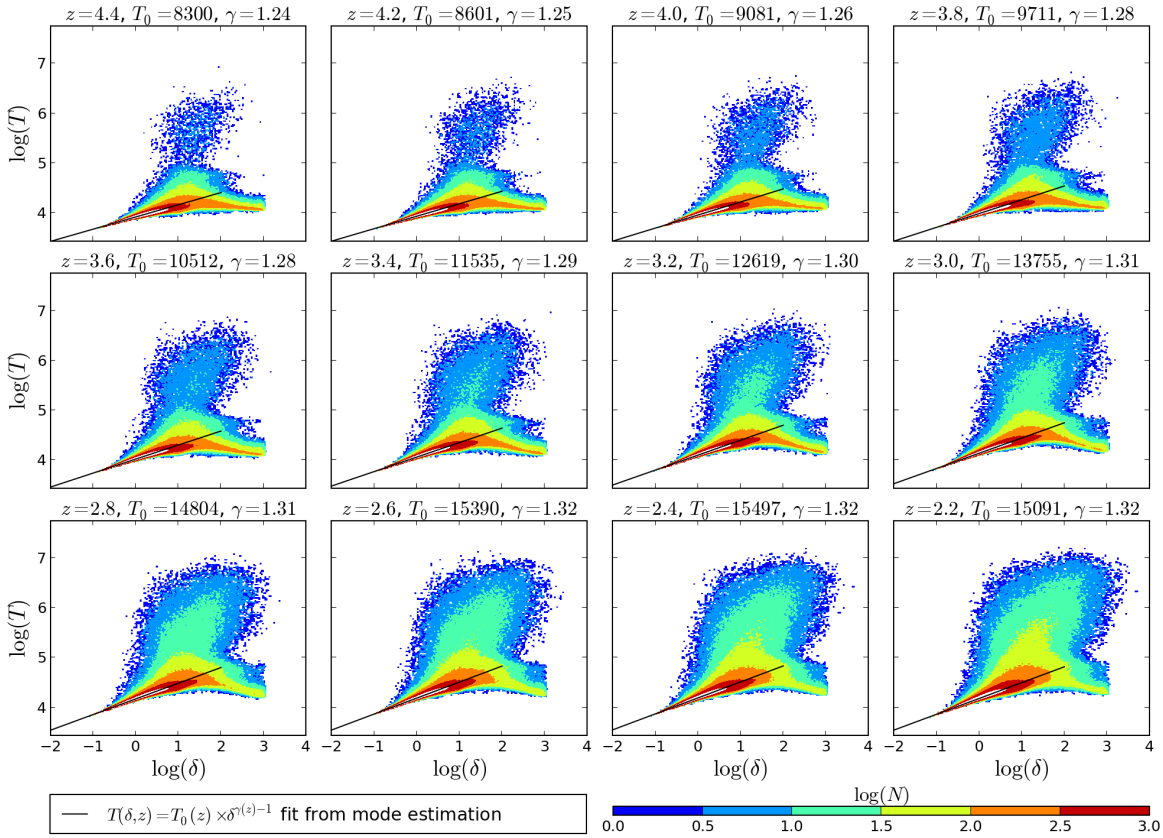


Figure 10. Temperature-density diagrams at various redshift. Color represents the particle density in logarithmic scale. The black line represents the fitted $T-\rho$ relation from several mode-estimated points. δ is the normalized density $\rho/\langle\rho\rangle$.

From the 1D power-spectra that we computed at each point of the grid, we derived a second-order Taylor expansion around our best-guess model. It describes the evolution of the 1D power spectrum with changes in either the cosmological or the astrophysical parameters that we studied. We have performed several check runs to ensure the quality and validity of our simulation grid, using either different seeds, or off-the-grid values of the cosmological and astrophysical parameters. These checks were all consistent with the power spectrum predicted using our second-order Taylor expansion, thus validating it. We compared our central simulation to published data from BOSS and showed that they were already in good agreement without any adjustment of any of the simulation parameters. In forthcoming work, we will use this Taylor expansion for a quantitative comparison to data in order to extract best-fit cosmological parameters.

These simulations are accompanied by a set of simulations where massive neutrinos are included. These required additional developments for an efficient treatment and a proper account of the additional particles (at all levels of the pipeline: in `CAMB`, in the setup of the initial conditions for thermal velocities, in `Gadget-3`, etc.), but are otherwise produced with a pipeline similar to the one presented in this study. The details about the simulations with massive neutrinos can be found in the companion paper [70]. Additional parameters can yet be included in the same context. However, due to the presence of the cross terms that are necessary for an accurate modeling of the likelihood function that illustrates the variation of the power spectrum in all directions of this growing parameter-space, adding new parameters will become more and more expensive in terms of calculation time.

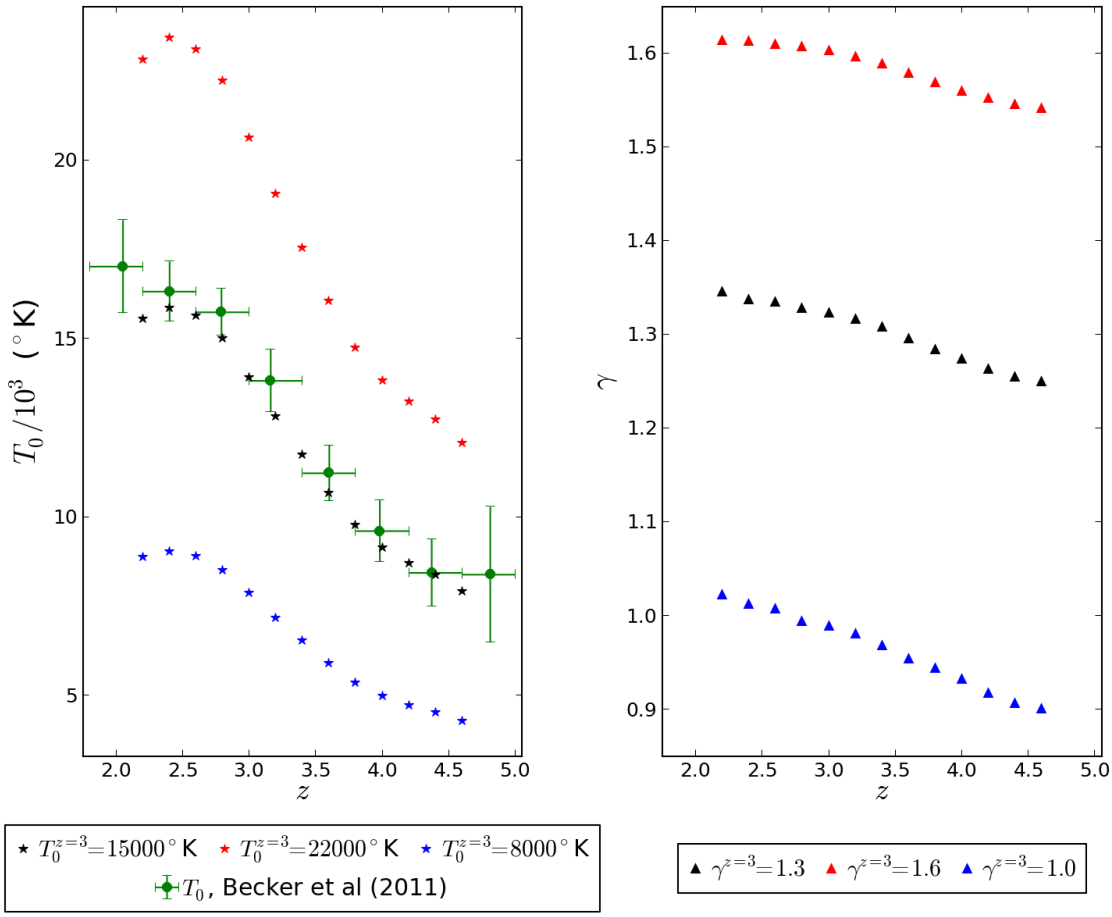


Figure 11. Evolution with redshift of T_0 and γ for the different values of these two parameters used in the grid. The plotted values are extracted from a sample of particles using a mode estimation as explained in the text. We overlaid the measurements of Becker et al. [57] (for $\gamma = 1.3$) for comparison.

A List of all simulations

We summarize in table 6 and 7 all the simulations mentioned in the paper. For box size and number of particles, (L, N) refers to a simulation with N^3 particles per species (gas or dark matter, thus $2\times N^3$ particles total) in a box of size $L\,\mathrm{Mpc}\,h^{-1}$ on a side. Standard parameters are $(n_s, \sigma_8, \Omega_m, H_0, T_0, \gamma) = (0.83, 0.96, 0.31, 67.5, 14.000, 1.3)$. Unless parameter names are explicitly listed, values are given for all parameters in the order just mentioned. All parameters are assumed to have their standard value unless specified otherwise. Except for the simulations performed for the convergence tests or to compute the exact power spectrum in the splicing test, all simulations are using the splicing technique to combine each set of three simulations into a single one of equivalent size to the largest box and equivalent mass-resolution to the best mass resolution. In the first column of the table 6, ∂_i and ∂_{ii} indicate simulations needed to compute the first and second order derivatives with respect to parameter i, and ∂_{ij} indicate additional simulations required to compute cross-derivatives with respect to parameters i and j, where i and j are explicitly given in the same column. In the last column of the same table, we also give the values of the shape parameter $\Gamma = \Omega_m h$ and of $\sigma_8(z = 3.0)$ where $z = 3.0$ corresponds to the pivot redshift of Lyα power spectrum measurements in [1] and [78]. These

parameters are not used in this work but we make them available since they are closely related to observations and were often used in the pioneering work on Lyα forest measurements.

Grid			
Type	Box size, particles (L, N)	Simulation parameters	$\sigma_8(z = 3)$; Γ
Central	(25,768)+(100,768)+(25,192)	Standard	0.26; 0.21
$\partial_i, \partial_{ii} : n_s$	(25,768)+(100,768)+(25,192)	$n_s = 0.91$	0.26; 0.21
$\partial_i, \partial_{ii} : n_s$	(25,768)+(100,768)+(25,192)	$n_s = 1.01$	0.26; 0.21
$\partial_i, \partial_{ii} : \sigma_8$	(25,768)+(100,768)+(25,192)	$\sigma_8 = 0.83$	0.25; 0.21
$\partial_i, \partial_{ii} : \sigma_8$	(25,768)+(100,768)+(25,192)	$\sigma_8 = 0.93$	0.28; 0.21
$\partial_i, \partial_{ii} : \Omega_m$	(25,768)+(100,768)+(25,192)	$\Omega_m = 0.26$	0.27; 0.18
$\partial_i, \partial_{ii} : \Omega_m$	(25,768)+(100,768)+(25,192)	$\Omega_m = 0.36$	0.25; 0.24
$\partial_i, \partial_{ii} : H_0$	(25,768)+(100,768)+(25,192)	$H_0 = 62.5$	0.26; 0.19
$\partial_i, \partial_{ii} : H_0$	(25,768)+(100,768)+(25,192)	$H_0 = 72.5$	0.26; 0.22
$\partial_i, \partial_{ii} : T_0$	(25,768)+(100,768)+(25,192)	$T_0 = 7.000$	0.26; 0.21
$\partial_i, \partial_{ii} : T_0$	(25,768)+(100,768)+(25,192)	$T_0 = 21.000$	0.26; 0.21
$\partial_i, \partial_{ii} : \gamma$	(25,768)+(100,768)+(25,192)	$\gamma = 1.0$	0.26; 0.21
$\partial_i, \partial_{ii} : \gamma$	(25,768)+(100,768)+(25,192)	$\gamma = 1.6$	0.26; 0.21
$\partial_{ij} : n_s - \sigma_8$	(25,768)+(100,768)+(25,192)	$n_s = 1.01, \sigma_8 = 0.93$	0.28; 0.21
$\partial_{ij} : n_s - \Omega_m$	(25,768)+(100,768)+(25,192)	$n_s = 1.01, \Omega_m = 0.36$	0.25; 0.24
$\partial_{ij} : n_s - H_0$	(25,768)+(100,768)+(25,192)	$n_s = 1.01, H_0 = 72.5$	0.26; 0.22
$\partial_{ij} : n_s - T_0$	(25,768)+(100,768)+(25,192)	$n_s = 1.01, T_0 = 21.000$	0.26; 0.21
$\partial_{ij} : n_s - \gamma$	(25,768)+(100,768)+(25,192)	$n_s = 1.01, \gamma = 1.6$	0.26; 0.21
$\partial_{ij} : \sigma_8 - \Omega_m$	(25,768)+(100,768)+(25,192)	$\sigma_8 = 0.93, \Omega_m = 0.36$	0.27; 0.24
$\partial_{ij} : \sigma_8 - H_0$	(25,768)+(100,768)+(25,192)	$\sigma_8 = 0.93, H_0 = 72.5$	0.28; 0.22
$\partial_{ij} : \sigma_8 - T_0$	(25,768)+(100,768)+(25,192)	$\sigma_8 = 0.93, T_0 = 21.000$	0.28; 0.21
$\partial_{ij} : \sigma_8 - \gamma$	(25,768)+(100,768)+(25,192)	$\sigma_8 = 0.93, \gamma = 1.6$	0.28; 0.21
$\partial_{ij} : \Omega_m - H_0$	(25,768)+(100,768)+(25,192)	$\Omega_m = 0.36, H_0 = 72.5$	0.25; 0.26
$\partial_{ij} : \Omega_m - T_0$	(25,768)+(100,768)+(25,192)	$\Omega_m = 0.36, T_0 = 21.000$	0.25; 0.24
$\partial_{ij} : \Omega_m - \gamma$	(25,768)+(100,768)+(25,192)	$\Omega_m = 0.36, \gamma = 1.6$	0.25; 0.24
$\partial_{ij} : H_0 - T_0$	(25,768)+(100,768)+(25,192)	$H_0 = 72.5, T_0 = 21.000$	0.26; 0.22
$\partial_{ij} : H_0 - \gamma$	(25,768)+(100,768)+(25,192)	$H_0 = 72.5, \gamma = 1.6$	0.26; 0.22
$\partial_{ij} : T_0 - \gamma$	(25,768)+(100,768)+(25,192)	$T_0 = 21.000, \gamma = 1.6$	0.26; 0.21

Table 6: Simulations used in this work for the grid.

Convergence tests		
Type	Box size, particles (L, N)	Simulation parameters
Resolution	(20,1024)	Standard
Resolution	(20,768)	Standard
Resolution	(20,512)	Standard
Resolution	(20,384)	Standard
Resolution	(20,192)	Standard
Box size	(120,1024)	Standard
Box size	(90,768)	Standard
Box size	(80,683)	Standard
Box size	(60,512)	Standard

Box size	(20,171)	Standard
Splicing tests		
Type	Box size, particles (L, N)	Simulation parameters
Grid-like	(25,256)+(100,256)+(25,64)	Standard
Exact	(100,1024)	Standard
Validity checks		
Type	Box size, particles (L, N)	Simulation parameters
Random seed	(25,768)+(100,768)+(25,192)	Standard
Off-grid 1	(25,768)+(100,768)+(25,192)	$(0.93, 0.85, 0.31, 67.5, 14.000, 1.32)$
Off-grid 2	(25,768)+(100,768)+(25,192)	$(0.96, 0.83, 0.31, 67.5, 10.000, 1.47)$
Off-grid 3	(25,768)+(100,768)+(25,192)	$(0.96, 0.83, 0.31, 67.5, 10.000, 1.16)$
Off-grid 4	(25,768)+(100,768)+(25,192)	$(0.93, 0.86, 0.30, 66, 10.000, 1.16)$
Off-grid 5	(25,768)+(100,768)+(25,192)	$(0.935, 0.846, 0.285, 68, 5.840, 1.32)$

Table 7: Simulations used in this work for the tests and final validity checks.

Acknowledgments

We acknowledge PRACE (Partnership for Advanced Computing in Europe) for awarding us access to resource curie-thin nodes based in France at TGCC, under allocation number 2012071264.
This work was also granted access to the resources of CCRT under the allocation 2013-t2013047004 made by GENCI (Grand Equipement National de Calcul Intensif).
A.B., N.P.-D., G.R. and Ch.Y. acknowledge support from grant ANR-11-JS04-011-01 of Agence Nationale de la Recherche.
M.V. is supported by ERC-StG "CosmoIGM".
JSB acknowledges the support of a Royal Society University Research Fellowship.
We thank Volker Springel for making `GADGET-3` available to our team.

References

[1] Palanque-Delabrouille, N., C. Yèche, A. Borde, et al. The one-dimensional Ly-alpha forest power spectrum from BOSS. *Astronomy & Astrophysics*, 559:A85, 2013. arXiv:1306.5896.

[2] Lynds, R. The Absorption-Line Spectrum of 4c 05.34. *The Astrophysical Journal*, 164:L73, 1971.

[3] Bi, H. G., G. Boerner, and Y. Chu. An alternative model for the Ly-alpha absorption forest. *Astronomy & Astrophysics*, 266:1–5, 1992.

[4] Miralda-Escude, J. and M. J. Rees. Tests for the minihalo model of the Lyman-alpha forest. *Monthly Notices of the Royal Astronomical Society*, 260:617–624, 1993.

[5] Croft, R. A. C., D. H. Weinberg, N. Katz, et al. Recovery of the Power Spectrum of Mass Fluctuations from Observations of the Ly-alpha Forest. *The Astrophysical Journal*, 495:44–62, 1998. arXiv:astro-ph/9708018.

[6] Cen, R. A hydrodynamic approach to cosmology - Methodology. *The Astrophysical Journal Supplement Series*, 78:341, 1992.

[7] Zhang, Y., P. Anninos, and M. L. Norman. A Multispecies Model for Hydrogen and Helium Absorbers in Lyman-Alpha Forest Clouds. *The Astrophysical Journal*, 453, 1995. arXiv:astro-ph/9508133.

[8] Hernquist, L., N. Katz, D. H. Weinberg, et al. The Lyman-Alpha Forest in the Cold Dark Matter Model. *The Astrophysical Journal*, 457:1–11, 1996. arXiv:astro-ph/9509105.

[9] Hui, L. and N. Y. Gnedin. Equation of state of the photoionized intergalactic medium. *Monthly Notices of the Royal Astronomical Society*, 292:27–42, 1997. arXiv:astro-ph/9612232.

[10] Hui, L., N. Y. Gnedin, and Y. Zhang. The Statistics of Density Peaks and the Column Density Distribution of the Ly-alpha Forest. *The Astrophysical Journal*, 486:599–622, 1997. arXiv:astro-ph/9608157.

[11] Viel, M., M. G. Haehnelt, and V. Springel. Inferring the dark matter power spectrum from the Lyman alpha forest in high-resolution QSO absorption spectra. *Monthly Notices of the Royal Astronomical Society*, 354:684–694, 2004. arXiv:astro-ph/0404600.

[12] Gnedin, N. Y. Probing the Universe with the Ly-alpha forest - II. The column density distribution. *Monthly Notices of the Royal Astronomical Society*, 299:392–402, 1998. arXiv:astro-ph/9706286.

[13] Hui, L., A. Stebbins, and S. Burles. A Geometrical Test of the Cosmological Energy Contents Using the Ly-alpha Forest. *The Astrophysical Journal*, 511:L5–L8, 1999. arXiv:astro-ph/9807190.

[14] Gaztanaga, E. and R. A. C. Croft. Predictions for the clustering properties of the Lyman-alpha forest – I. One-point statistics. *Monthly Notices of the Royal Astronomical Society*, 309:885–904, 1999. arXiv:astro-ph/9811480.

[15] Nusser, A. and M. Haehnelt. A first step towards a direct inversion of the Lyman forest in QSO spectra. *Monthly Notices of the Royal Astronomical Society*, 303:179–187, 1999. arXiv:astro-ph/9806109.

[16] Feng, L.-l. and L.-z. Fang. Non-Gaussianity and the Recovery of the Mass Power Spectrum from the Ly-alpha Forest. *The Astrophysical Journal*, 535:519–529, 2000. arXiv:astro-ph/0001348.

[17] McDonald, P., J. Miralda-Escude, M. Rauch, et al. The Observed Probability Distribution Function, Power Spectrum, and Correlation Function of the Transmitted Flux in the Ly-alpha Forest. *The Astrophysical Journal*, 543:1–23, 2000. arXiv:astro-ph/9911196.

[18] Hui, L., S. Burles, U. Seljak, et al. On Estimating the QSO Transmission Power Spectrum. *The Astrophysical Journal*, 552:15–35, 2001. arXiv:astro-ph/0005049.

[19] Alcock, C. and B. Paczyński. An evolution free test for non-zero cosmological constant. *Nature*, 281: 358–359, 1979.

[20] McDonald, P. and J. Miralda-Escude. Measuring the Cosmological Geometry from the Ly-alpha Forest along Parallel Lines of Sight. *The Astrophysical Journal*, 518:24–31, 1999. arXiv:astro-ph/9807137.

[21] Croft, R. A. C., D. H. Weinberg, M. Bolte, et al. Toward a Precise Measurement of Matter Clustering: Ly-alpha Forest Data at Redshifts 2–4. *The Astrophysical Journal*, 581:20–52, 2002. arXiv:astro-ph/0012324.

[22] McDonald, P. and D. Eisenstein. Dark energy and curvature from a future baryonic acoustic oscillation survey using the Lyman-alpha forest. *Physical Review D*, 76:063009, 2007. arXiv:astro-ph/0607122.

[23] Seljak, U., A. Makarov, P. McDonald, et al. Cosmological parameter analysis including SDSS Ly-alpha forest and galaxy bias: Constraints on the primordial spectrum of fluctuations, neutrino mass, and dark energy. *Physical Review D*, 71:103515, 2005. arXiv:astro-ph/0407372.

[24] Viel, M., M. G. Haehnelt, and V. Springel. The effect of neutrinos on the matter distribution as probed by the intergalactic medium. *Journal of Cosmology and Astroparticle Physics*, 2010:015–015, 2010. arXiv:1003.2422.

[25] Croft, R. A. C., D. H. Weinberg, M. Pettini, et al. The Power Spectrum of Mass Fluctuations Measured from the Ly-alpha Forest at Redshift z = 2.5. *The Astrophysical Journal*, 520:1–23, 1999. arXiv:astro-ph/9809401.

[26] Kim, T.-S., M. Viel, M. G. Haehnelt, et al. Erratum: The power spectrum of the flux distribution in the Lyman-alpha forest of a large sample of UVES QSO Absorption Spectra (LUQAS). *Monthly Notices of the Royal Astronomical Society*, 351:1471–1472, 2004.

[27] Kim, T. S., M. Viel, M. G. Haehnelt, et al. The power spectrum of the flux distribution in the Lyman-alpha forest of a large sample of UVES QSO absorption spectra (LUQAS). *Monthly Notices of the Royal Astronomical Society*, 347:355–366, 2004. arXiv:astro-ph/0308103.

[28] York, D. G., J. Adelman, J. E. Anderson, Jr., et al. The Sloan Digital Sky Survey: Technical Summary. *The Astronomical Journal*, 120:1579–1587, 2000. arXiv:astro-ph/0006396.

[29] McDonald, P., U. Seljak, S. Burles, et al. The Ly-alpha Forest Power Spectrum from the Sloan Digital Sky Survey. *The Astrophysical Journal Supplement Series*, 163:80–109, 2006. arXiv:astro-ph/0405013.

[30] McDonald, P., U. Seljak, R. Cen, et al. The Linear Theory Power Spectrum from the Ly-alpha Forest in the Sloan Digital Sky Survey. *The Astrophysical Journal*, 635:761–783, 2005. arXiv:astro-ph/0407377.

[31] Eisenstein, D. J., D. H. Weinberg, E. Agol, et al. SDSS-III: Massive Spectroscopic Surveys of the Distant Universe, the Milky Way and Extra-Solar Planetary Systems. *The Astronomical Journal*, 142: 72, 2011. arXiv:1101.1529v1.

[32] Dawson, K. S., D. J. Schlegel, C. P. Ahn, et al. The Baryon Oscillation Spectroscopic Survey of SDSS-III. *The Astronomical Journal*, 145:10, 2013. arXiv:1208.0022.

[33] Pâris, I., P. Petitjean, E. Aubourg, et al. The Sloan Digital Sky Survey quasar catalog: ninth data release. *Astronomy & Astrophysics*, 548:A66, 2012. arXiv:1210.5166.

[34] Lee, K.-g., S. Bailey, L. E. Bartsch, et al. The BOSS Ly-alpha forest sample from SDSS data release 9. *The Astronomical Journal*, 145:69, 2013. arXiv:1211.5146.

[35] Ahn, C. P., R. Alexandroff, C. Allende Prieto, et al. The Ninth Data Release of the Sloan Digital Sky Survey: First Spectroscopic Data from the SDSS-III Baryon Oscillation Spectroscopic Survey. *The Astrophysical Journal Supplement Series*, 203:21, 2012. arXiv:1207.7137.

[36] Busca, N. G., T. Delubac, J. Rich, et al. Baryon acoustic oscillations in the Ly-alpha forest of BOSS quasars. *Astronomy & Astrophysics*, 552:A96, 2013. arXiv:1211.2616.

[37] Slosar, A., V. Iršič, D. Kirkby, et al. Measurement of baryon acoustic oscillations in the Lyman-alpha forest fluctuations in BOSS data release 9. *Journal of Cosmology and Astroparticle Physics*, 2013: 026–026, 2013. arXiv:1301.3459.

[38] Kirkby, D., D. Margala, A. Slosar, et al. Fitting methods for baryon acoustic oscillations in the Lyman-alpha forest fluctuations in BOSS data release 9. *Journal of Cosmology and Astroparticle Physics*, 2013:024–024, 2013. arXiv:1301.3456.

[39] Lee, K.-g., J. F. Hennawi, M. White, et al. Observational Requirements for Lyman-alpha Forest Tomographic Mapping of Large-Scale Structure at z ∼ 2. 2013. arXiv:1309.1477.

[40] Viel, M. and M. G. Haehnelt. Cosmological and astrophysical parameters from the Sloan Digital Sky Survey flux power spectrum and hydrodynamical simulations of the Lyman alpha forest. *Monthly Notices of the Royal Astronomical Society*, 365:231–244, 2006. arXiv:astro-ph/0508177.

[41] Bolton, J. S., M. Viel, T.-S. Kim, et al. Possible evidence for an inverted temperature–density relation in the intergalactic medium from the flux distribution of the Ly-alpha forest. *Monthly Notices of the Royal Astronomical Society*, 386:1131–1144, 2008. arXiv:0711.2064.

[42] Bird, S., H. V. Peiris, M. Viel, et al. Minimally parametric power spectrum reconstruction from the Lyman-alpha forest. *Monthly Notices of the Royal Astronomical Society*, 413:1717–1728, 2011. arXiv:1010.1519.

[43] Bird, S., M. Viel, and M. G. Haehnelt. Massive neutrinos and the non-linear matter power spectrum. *Monthly Notices of the Royal Astronomical Society*, 420:2551–2561, 2012. arXiv:1109.4416.

[44] Springel, V. The cosmological simulation code GADGET-2. *Monthly Notices of the Royal Astronomical Society*, 364:1105–1134, 2005. arXiv:astro-ph/0505010.

[45] Schlegel, D., F. Abdalla, T. Abraham, et al. The BigBOSS Experiment. 2011. arXiv:1106.1706.

[46] Tang, B. Orthogonal Array-Based Latin Hypercubes. *Journal of the American Statistical Association*, 88:1392–1397, 1993.

[47] McKay, M. D., R. J. Beckman, and W. J. Conover. A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output From a Computer Code. *Technometrics*, 42:55–61, 2000.

[48] Heitmann, K., D. Higdon, M. White, et al. The Coyote Universe. II. Cosmological Models and Precision Emulation of the Nonlinear Matter Power Spectrum. *The Astrophysical Journal*, 705: 156–174, 2009. arXiv:0902.0429.

[49] Schneider, M. D., S. Cole, C. S. Frenk, et al. Fast Generation of Ensembles of Cosmological N-Body Simulations via Mode Resampling. *The Astrophysical Journal*, 737:11, 2011. arXiv:1103.2767.

[50] Viel, M., M. G. Haehnelt, and A. Lewis. The Lyman alpha forest and WMAP year three. *Monthly Notices of the Royal Astronomical Society: Letters*, 370:L51–L55, 2006. arXiv:astro-ph/0604310.

[51] Wang, M.-y., R. A. C. Croft, A. H. G. Peter, et al. Lyman-alpha Forest Constraints on Decaying Dark Matter. 2013. arXiv:1309.7354.

[52] Planck Collaboration, P. A. R. Ade, N. Aghanim, et al. Planck 2013 results. XVI. Cosmological parameters. 2013. arXiv:1303.5076.

[53] Komatsu, E., K. M. Smith, J. Dunkley, et al. Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation. *The Astrophysical Journal Supplement Series*, 192:18, 2011. arXiv:1001.4538.

[54] Hou, Z., C. L. Reichardt, K. T. Story, et al. Constraints on Cosmology from the Cosmic Microwave Background Power Spectrum of the 2500 deg2 SPT-SZ Survey. *The Astrophysical Journal*, 782:74, 2014. arXiv:1212.6267.

[55] Conley, A., J. Guy, M. Sullivan, et al. Supernova Constraints and Systematic Uncertainties from the First Three Years of the Supernova Legacy Survey. *The Astrophysical Journal Supplement Series*, 192: 1, 2011. arXiv:1104.1443.

[56] Sullivan, M., J. Guy, A. Conley, et al. SNLS3: Constraints on Dark Energy Combining the Supernova Legacy Survey Three-Year Data with Other Probes. *The Astrophysical Journal*, 737:102, 2011. arXiv:arXiv:1104.1444v2.

[57] Becker, G. D., J. S. Bolton, M. G. Haehnelt, et al. Detection of extended HeII reionization in the temperature evolution of the intergalactic medium. *Monthly Notices of the Royal Astronomical Society*, 410:1096–1112, 2011. arXiv:1008.2622.

[58] Garzilli, A., J. S. Bolton, T.-S. Kim, et al. The intergalactic medium thermal history at redshift z = 1.7-3.2 from the Lyα forest: a comparison of measurements using wavelets and the flux distribution. *Monthly Notices of the Royal Astronomical Society*, 424:1723–1736, 2012. arXiv:1202.3577.

[59] Lidz, A., C.-A. Faucher-Giguère, A. Dall'Aglio, et al. A Measurement of Small-Scale Structure in the 2.2 < z < 4.2 Ly-alpha Forest. *The Astrophysical Journal*, 718:199–230, 2010. arXiv:0909.5210.

[60] Schaye, J., T. Theuns, M. Rauch, et al. The thermal history of the intergalactic medium. *Monthly Notices of the Royal Astronomical Society*, 318:817–826, 2000. arXiv:astro-ph/9912432.

[61] Gnedin, N. Y. and L. Hui. Probing the Universe with the Lyalpha forest - I. Hydrodynamics of the low-density intergalactic medium. *Monthly Notices of the Royal Astronomical Society*, 296:44–55, 1998. astro-ph/9706219.

[62] Theuns, T., M. Viel, S. Kay, et al. Galactic Winds in the Intergalactic Medium. *The Astrophysical Journal*, 578:L5–L8, 2002. astro-ph/0208418.

[63] McQuinn, M., L. Hernquist, A. Lidz, et al. The signatures of large-scale temperature and intensity fluctuations in the Lyman α forest. *Monthly Notices of the Royal Astronomical Society*, 415:977–992, 2011. arXiv:astro-ph.CO/1010.5250.

[64] McDonald, P., U. Seljak, R. Cen, et al. Physical effects on the Lyα forest flux power spectrum: damping wings, ionizing radiation fluctuations and galactic winds. *Monthly Notices of the Royal Astronomical Society*, 360:1471–1482, 2005. astro-ph/0407378.

[65] Palanque-Delabrouille, N., C. Magneville, C. Yèche, et al. Luminosity function from dedicated SDSS-III and MMT data of quasars in $0.7 < z < 4.0$ selected with a new approach. *Astronomy & Astrophysics*, 551:A29, 2013. arXiv:1209.3968.

[66] Lewis, A., A. Challinor, and A. Lasenby. Efficient Computation of Cosmic Microwave Background Anisotropies in Closed Friedmann-Robertson-Walker Models. *The Astrophysical Journal*, 538: 473–476, 2000. arXiv:astro-ph/9911177.

[67] Seljak, U. and M. Zaldarriaga. A Line-of-Sight Integration Approach to Cosmic Microwave Background Anisotropies. *The Astrophysical Journal*, 469:437, 1996. arXiv:astro-ph/9603033.

[68] Bertschinger, E. COSMICS: Cosmological Initial Conditions and Microwave Anisotropy Codes. 1995. arXiv:astro-ph/9506070.

[69] Crocce, M., S. Pueblas, and R. Scoccimarro. Transients from initial conditions in cosmological simulations. *Monthly Notices of the Royal Astronomical Society*, 373:369–381, 2006. arXiv:astro-ph/0606505.

[70] Rossi, G., N. Palanque-Delabrouille, A. Borde, et al. A Suite of Hydrodynamical Simulations for the Lyman-Alpha Forest with Massive Neutrinos. 2014. arXiv:1401.6464.

[71] Ali-Haimoud, Y. and S. Bird. An efficient implementation of massive neutrinos in non-linear structure formation simulations. *Monthly Notices of the Royal Astronomical Society*, 428:3375–3389, 2012. arXiv:1209.0461.

[72] Springel, V., N. Yoshida, and S. D. White. GADGET: a code for collisionless and gasdynamical cosmological simulations. *New Astronomy*, 6:79–117, 2001. arXiv:astro-ph/0003162.

[73] Gnedin, N. Y. and A. J. S. Hamilton. Matter power spectrum from the Lyman-alpha forest: myth or reality? *Monthly Notices of the Royal Astronomical Society*, 334:107–116, 2002. astro-ph/0111194.

[74] Meiksin, A. A. The physics of the intergalactic medium. *Reviews of Modern Physics*, 81:1405–1469, 2009. arXiv:0711.3358.

[75] Theuns, T. Numerical simulations of quasar absorbers. In Williams, P., C.-G. Shu, and B. Menard, editors, *IAU Colloq. 199: Probing Galaxies through Quasar Absorption Lines*, pages 185–204, 2005, astro-ph/0507570.

[76] Bolton, A. S., D. J. Schlegel, E. Aubourg, et al. Spectral Classification and Redshift Measurement for the SDSS-III Baryon Oscillation Spectroscopic Survey. *The Astronomical Journal*, 144:144, 2012. arXiv:1207.7326.

[77] Bolton, J. S. and G. D. Becker. Resolving the high redshift Ly-alpha forest in smoothed particle hydrodynamics simulations. *Monthly Notices of the Royal Astronomical Society: Letters*, 398:L26–L30, 2009. arXiv:0906.2861.

[78] McDonald, P. Toward a Measurement of the Cosmological Geometry at $z \sim 2$: Predicting Ly-alpha Forest Correlation in Three Dimensions and the Potential of Future Data Sets. *The Astrophysical Journal*, 585:34–51, 2003. arXiv:astro-ph/0108064.