

# Lawrence Berkeley National Laboratory

## LBL Publications

### Title

Modelling the Lyman- $\alpha$  forest with Eulerian and SPH hydrodynamical methods

### Permalink

<https://escholarship.org/uc/item/6bd9r3q0>

### Journal

Monthly Notices of the Royal Astronomical Society, 518(3)

### ISSN

0035-8711

### Authors

Chabanier, Solène

Emberson, JD

Lukić, Zarija

et al.

### Publication Date

2022-12-02

### DOI

10.1093/mnras/stac3294

Peer reviewed

# Modeling the Lyman- $\alpha$ forest with Eulerian and SPH hydrodynamical methods

Solène Chabanier<sup>1\*</sup>, J.D. Emberson<sup>2</sup>, Zarija Lukić<sup>1</sup>, Jesus Pulido<sup>3</sup>, Salman Habib<sup>2,4</sup>, Esteban Rangel<sup>2</sup>, Jean Sexton<sup>1</sup>, Nicholas Frontiere<sup>2</sup>, and Michael Buehlmann<sup>4</sup>

<sup>1</sup>Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

<sup>2</sup>CPS Division, Argonne National Laboratory, Lemont, IL 60439, USA

<sup>3</sup>Los Alamos National Laboratory, Los Alamos, NM 87545, USA

<sup>4</sup>HEP Division, Argonne National Laboratory, Lemont, IL 60439, USA

## ABSTRACT

We compare two state-of-the-art numerical codes to study the overall accuracy in modeling the intergalactic medium and reproducing Lyman- $\alpha$  forest observables for DESI and high-resolution data sets. The codes employ different approaches to solving both gravity and modeling the gas hydrodynamics. The first code, Nyx, solves the Poisson equation using the Particle-Mesh (PM) method and the Euler equations using a finite volume method. The second code, CRK-HACC, uses a Tree-PM method to solve for gravity, and an improved Lagrangian smoothed particle hydrodynamics (SPH) technique, where fluid elements are modeled with particles, to treat the intergalactic gas. We compare the convergence behavior of the codes in flux statistics as well as the degree to which the codes agree in the converged limit. We find good agreement overall with differences being less than observational uncertainties, and a particularly notable  $\lesssim 1\%$  agreement in the 1D flux power spectrum. This agreement was achieved by applying a tessellation methodology for reconstructing the density in CRK-HACC instead of using an SPH kernel as is standard practice. We show that use of the SPH kernel can lead to significant and unnecessary biases in flux statistics; this is especially prominent at high redshifts,  $z \sim 5$ , as the Lyman- $\alpha$  forest mostly comes from lower-density regions which are intrinsically poorly sampled by SPH particles.

**Key words:** methods: numerical – galaxies: intergalactic medium – quasars: absorption lines

## 1 INTRODUCTION

At high redshift, the vast majority of the baryon content of the universe is present in the intergalactic medium (IGM). Neutral hydrogen in the IGM scatters light at 1216 Å, producing characteristic absorption features in the spectra of distant quasars (QSO), dubbed the Lyman- $\alpha$  ( $Ly\alpha$ ) forest. Observationally, the  $Ly\alpha$  forest has been investigated for over half a century, but its recognition as a crucial source of cosmological information is more recent (Croft et al. 1998; Rauch 1998). Effectively, each quasar spectrum yields a nonlinear one-dimensional map of the IGM density at high redshifts. On large scales, the gas follows the matter density and is pressure-smoothed on small-scales; the forest traces density fluctuations in the linear to the mildly nonlinear regime at small scales, down to  $0.5 h^{-1}\text{Mpc}$ , and at high redshifts, in the range  $2 < z < 5$ .

In its ability to probe smaller scales, the  $Ly\alpha$  forest is sensitive to the small-scale matter clustering signal – potentially filtered by the effects of relativistic particles (e.g., massive neutrinos or exotic dark matter (DM) candidates) – without being strongly affected by the highly nonlinear gravitational collapse and galaxy formation details that plague low-redshift probes. As such, it has been used to put strong constraints on the sum of neutrino masses (Palanque-Delabrouille

et al. 2015a,b; Yèche et al. 2017b; Palanque-Delabrouille et al. 2019), on warm dark matter (WDM) mass (Viel et al. 2005, 2008, 2013a; Baur et al. 2016; Yeche et al. 2017a; Baur et al. 2017; Palanque-Delabrouille et al. 2019) and on fuzzy dark matter models (Iršič et al. 2017; Armengaud et al. 2017). The  $Ly\alpha$  forest is also sensitive to the thermal state of the IGM and can therefore set constraints on reionization and on its thermal history (Hui & Gnedin 1997; Zaldarriaga et al. 2001; Bolton & Becker 2009; Meiksin 2009; Becker & Bolton 2013; Lee et al. 2015; McQuinn 2016; Boera et al. 2019; Walther et al. 2019; Gaikwad et al. 2021).

The  $Ly\alpha$  forest probes matter clustering as traced by neutral hydrogen on both large and small spatial scales. Flux transmission correlation measurements on large scales,  $\sim 100 h^{-1}\text{Mpc}$ , are usually performed using pixels across different lines of sight (LOS) and have been used to measure the Baryon Acoustic Oscillation (BAO) peak with the Baryon Oscillation Spectroscopic Survey (BOSS) and eBOSS data at  $z \sim 2.4$  (Slosar et al. 2011, 2013; Bautista et al. 2017; du Mas des Bourboux et al. 2017; de Sainte Agathe et al. 2019) providing additional constraints on dark energy. On the other hand, information on small scales is accessible through correlations along the LOS, or equivalently through the 1D power spectrum,  $P_{1D}$ , for which measurements can be divided into two categories. The first category consists of a large sample of medium resolution quasar spectra, in the tens of thousands, from the Sloan Digital Sky

\* E-mail: schabanier@lbl.gov

Survey (SDSS-III, Eisenstein et al. 2011) BOSS program (Dawson et al. 2013) and the SDSS-IV (Blanton et al. 2017) eBOSS (Dawson et al. 2016) extension. These measurements were performed for scales  $0.001 < k < 0.02 \text{ km}^{-1}\text{s}$  and redshifts  $2.1 < z < 4.7$  (McDonald et al. 2006; Palanque-Delabrouille et al. 2013; Chabanier et al. 2019), with the latest measurement being based on  $> 45,000$  quasar spectra and reaching  $\sim 1.5\%$  statistical precision already at low redshifts. The second category of measurements originates from 8m-class telescope data sets, e.g., Keck/HIRES, VLT/UVES, or X-Shooter (Vogt et al. 1994; Dekker et al. 2000; Vernet et al. 2011), which are significantly smaller (tens to hundreds of spectra) but with much higher resolution and higher S/N (Viel et al. 2013a; Iršič et al. 2016; Walther et al. 2018; Boera et al. 2019; Gaikwad et al. 2021; Karaçaylı et al. 2022), probing both smaller scales and extending the redshift range to  $1.8 < z < 5.4$ , albeit at lower statistical precision.

The improvement in instrumentation and data analyses of future and ongoing Ly $\alpha$  surveys, such as the Dark Energy Spectroscopic Instrument (DESI, Levi et al. 2013; DESI Collaboration et al. 2016) or WEAVE-QSO (Pieri et al. 2016), will strongly increase the precision of  $P_{1D}$  observational measurements. In particular, DESI will observe a total of  $\sim 800,000$  Ly $\alpha$  quasars with  $z \geq 2.1$ , i.e. about 4 times the largest data set to date. Also, thanks to the strong improvement in resolution – factor of two better than eBOSS – and the observation of fainter quasars, made possible with better target selection and the larger mirror, DESI will be able to measure  $P_{1D}$  down to scales of  $5 \text{ h}^{-1}\text{Mpc}$  at redshifts as high as  $z \sim 4.9$  at an unprecedented accuracy, 1% at low redshifts and a few percent at the highest redshifts.

However, robust cosmological constraints using small-scale Ly $\alpha$  data can only be achieved if theoretical predictions from hydrodynamical simulations can reach percent-level accuracy in order to match the quality of the observations. At high redshifts and small scales, gravitational clustering is in the mildly nonlinear regime, and because the Ly $\alpha$  forest is sensitive to the intrinsic properties of the baryonic gas and details of reionization, no analytical solution for the Ly $\alpha$  flux fluctuations over time is available. One must therefore resort to cosmological hydrodynamical simulations for obtaining accurate theoretical predictions.

Structure formation in the universe is often studied by gravity-only N-body methods. (For a recent review, see Angulo et al. 2021.) While these techniques do not include gas dynamics and a host of astrophysical effects, they are by now relatively well-understood. Internal tests and code comparison campaigns have shown that different N-body codes (often using different algorithms) can agree at the per-cent level on small scales, depending on the chosen force and mass resolution. The situation is quite different, however, for the significantly more complex case of cosmological hydrodynamical codes, where even without including subgrid physics modeling, different numerical methods have historically disagreed at the level of tens of percent (Frenk et al. 1999; Agertz et al. 2007; Sembolini et al. 2016). As methods have been sharpened, and new ones developed, the results from different gas dynamics schemes have come closer. In particular, comparisons of (non-radiative) Nyx and CRK-HACC simulations for individual galaxy clusters have shown excellent agreement, without the need to add any unphysical conductivity or switches of various types to the CRKSPH solver (Frontiere et al. 2022). Nevertheless, potential problems with over- and under-mixing of gas, implementation of different types of numerical viscosity schemes, differing evolution algorithms, and choices of various limiters, all contribute to making code comparison a nontrivial task.

It is common practice to test code accuracy on a number of idealized problems, where either analytical or highly accurate numerical

solutions are known. These problems often possess high degrees of symmetry or reduced dimensionality, which do not make them very realistic. Codes are often tuned to address specific types of problems so that they may do well on certain of these tests and not so well on others. It is not easy, therefore, to predict how a particular method will perform in a given realistic scenario, especially since there is no rigorous a priori theory of error convergence for any method. Consequently, comparing different methods and codes (with default settings) on multiple realistic test problems is a reasonable way to assess the robustness of the numerical methods. It avoids problems with false convergence of a given algorithm; if a variety of methods, based on different assumptions and techniques, all agree well, confidence in the final result is certainly enhanced.

Although there are multiple classes of methods, the differences between smoothed particle hydrodynamics (SPH) and Eulerian techniques have been highlighted previously (Ageretz et al. 2007; Hubber et al. 2013). The main differences were related to SPH’s inability to resolve fluid instabilities such as Kelvin-Helmholtz or Rayleigh-Taylor and the consequential effects on mixing of multi-phase media. Modern SPH methods (e.g., Read & Hayfield 2012; Saitoh & Makino 2013; Hopkins 2013; Frontiere et al. 2017) have been developed to overcome these difficulties and have been incorporated in cosmological hydrodynamics codes. In any case, modeling the Ly $\alpha$  forest is much simpler (smooth single-phase gas, with shocks limited to high-density regions, which have little effect on the measured signals), and therefore one may expect different methods to agree better than in other, more difficult, astrophysical circumstances, although previous comparisons still reported differences at the  $\sim 5\text{-}10\%$  level (Regan et al. 2007; Bird et al. 2013; Walther et al. 2021) between different codes. Due to advances in computational methods and computing power, as well as increased accuracy of observational data, it is of considerable interest to examine agreements between different codes at higher levels of accuracy.

In this paper, we therefore examine the current state-of-the-art in Ly $\alpha$  forest statistics as reproduced by optically-thin hydrodynamical simulations. We study the convergence of Ly $\alpha$  forest properties using two hydrodynamical codes following completely different numerical approaches. One is the Lagrangian, particle-based, Conservative Reproducing Kernel SPH (CRKSPH) method (Frontiere et al. 2017) implemented in the CRK-HACC code (Habib et al. 2016; Frontiere et al. 2022) and the other is the Eulerian, grid-based, finite-volume hydrodynamical method used in the Nyx code (Almgren et al. 2013; Sexton et al. 2021). Good levels of agreement between these codes in realistic cosmological simulation studies will significantly bolster our confidence that both are converging to the correct result. Indeed, by improving the sampling of the density field in lower-density regions with the (volume-weighted) Delaunay Tesselation Field Estimator (DTFE) (Schaap 2007; Rangel et al. 2016) for CRK-HACC, we find that the Ly $\alpha$  1D power spectra of the two codes agree at better than 1% at the redshifts of interest,  $2.0 \leq z \leq 5.0$ , without any fine tuning.

The plan of this paper is as follows. In Section 2 we describe the different ways in which Nyx and CRK-HACC solve the gravitational and hydrodynamics equations along with a description of the simulation sets we perform and analyze for the purpose of the comparison. In Section 3 we outline the numerical methods used to derive the Ly $\alpha$  properties from the simulation snapshots. We discuss morphological differences in Section 4 and differences in the temperature and density fields between the two codes in Section 5. We present results on the Ly $\alpha$  flux statistics, the mean flux, flux probability distribution function and (1D and 3D) flux power spectrum in Section 6 for small boxes with varying spatial resolution, and study the impact of box

size in Section 7. We finally conclude in Section 8 with a discussion of the main results.

## 2 SIMULATIONS

At first sight, it appears relatively straightforward to directly simulate the Ly $\alpha$  forest. The gas that creates the absorption features – the diffuse IGM – is only at moderate overdensity,  $0 \leq \delta \leq 10$ , making the simulation task relatively modest in difficulty. The evolution of the gas depends on gravity and pressure forces, and some simple radiative processes. Since the gas is very close to the primordial composition, the background ionizing radiation is relatively spatially uniform, and since almost all of the gas (by volume) is optically thin to the radiation, the required simulation machinery is not complicated. At first order, any cosmological hydrodynamics code modeling dark matter and gas with uniform radiative heating and cooling should be able to adequately capture the diffuse IGM. (See [Meiksin 2009](#) and [McQuinn 2016](#) for thorough reviews of the relevant physics.)

It is important to note, however, that there are secondary effects linked to galaxy formation and evolution, such as sources of galactic feedback, inhomogeneity of the ionizing background, or heating from helium reionization, which can strongly impact the thermal state and distribution of gas in the IGM, affecting  $P_{\text{ID}}$  at the several percent level ([Viel et al. 2013b](#); [Chabanier et al. 2020](#); [Montero-Camacho & Mao 2021](#)). Before incorporating these effects, it is important to confirm that the primary-level modeling of the Ly $\alpha$  forest is in fact robust at the  $\sim 1\%$  level, the main purpose of this paper.

We will employ simple – but still fairly realistic – assumptions regarding the physics of the Ly $\alpha$  forest, which is based on a time-evolving, spatially uniform UV background. Simulations of this kind neglect the effects of inhomogeneous reionization, which produces temperature and UV background fluctuations on very large scales ([Oñorbe et al. 2019](#)), and neglect galaxy formation models. We emphasize that this type of simulation is used as a forward model in virtually any recent inference approach applied to the Ly $\alpha$  power spectrum (for example, see: [Boera et al. 2019](#); [Walther et al. 2019](#); [Palanque-Delabrouille et al. 2019](#); [Pedersen et al. 2021](#); [Walther et al. 2021](#)). In the following, we first describe the codes we use, and then present the suite of simulations performed for this work.

### 2.1 Nyx

Nyx is a publicly available<sup>1</sup>, parallel, adaptive mesh, cosmological simulation code that solves the equations of compressible hydrodynamics describing the evolution of baryonic gas coupled with an N-body treatment of the dark matter in an expanding universe ([Almgren et al. 2013](#); [Sexton et al. 2021](#)). Nyx’s hydrodynamics is based on an Eulerian formulation, which is a very efficient approach for the low-density regions of the intergalactic medium. The code uses a second order (dimensionally-unsplitted) piecewise linear (PLM) or piecewise parabolic method (PPM, [Colella & Woodward 1984](#)) to construct the fluxes through the interfaces of each cell. The Poisson equation for self-gravity of the gas and dark matter is solved using a geometric multigrid method.

Nyx is built on the AMReX ([Zhang et al. 2019](#)) adaptive mesh refinement (AMR) library and is written in C++. The approach to AMR uses a nested hierarchy of logically-rectangular grids with simultaneous refinement in both space and time. AMR levels are

advanced at their own timestep (sub-cycling) and jumps by factors of 2 and 4 are supported between levels. The integration algorithm on the grid hierarchy is a recursive procedure in which coarse grids are advanced in time, fine grids are advanced multiple steps to reach the same time as the coarse grids and the data at different levels are then synchronized. A separate synchronization is done for gravity. We use MPI to distribute AMR grids across nodes and use logical tiling with OpenMP to divide a grid across threads for multi-core CPU machines (exposing coarse-grained parallelism) and/or CUDA/HIP/DPC++ to spread the work across GPU threads on GPU-based machines (fine-grained parallelism).

Details of Nyx’s Ly $\alpha$  forest modeling are given in [Lukić et al. 2015](#), but we quickly summarize it here as well. To model the Ly $\alpha$  forest, Nyx follows the abundance of six species: neutral and ionized hydrogen, neutral, once and twice ionized helium, and free electrons. For these species, all relevant atomic processes – ionization, recombination, and free-free transitions are modelled. Heating and cooling source terms are calculated using a sub-cycled approach in order to avoid running the whole code on a short, cooling timescale. It is assumed that all gas elements are optically thin to ionizing photons, such that their ionization state can be fully described by a uniform and isotropic UV+X-ray background radiation field ([Oñorbe et al. 2017](#)). Nyx also has the capability to simulate inhomogeneous reionization models ([Oñorbe et al. 2019](#)), which affect Ly $\alpha$  observables at large scales.

### 2.2 CRK-HACC

CRK-HACC is the newly developed hydrodynamic extension of the high-performance cosmological N-body code HACC (Hardware/Hybrid Accelerated Cosmology Code; [Habib et al. 2016](#)) which couples the gravity solver of HACC with a Lagrangian hydro solver based on the CRKSPH method ([Frontiere et al. 2017](#)). As a modern development of SPH ([Lucy 1977](#); [Gingold & Monaghan 1977](#)), CRKSPH leverages its key benefits (precise enforcement of conservation laws, Galilean invariance, ease of parallelization, and inherently adaptive refinement), while at the same time addressing many of the original difficulties with SPH (e.g., zeroth-order field reproduction, overly aggressive artificial viscosity models). The gravity solver in CRK-HACC uses a high-order spectral particle-mesh method handing off at small scales to short-distance algorithms (FMM, P<sup>3</sup>M, tree) depending on the choice of computational architecture. The implementation of the hydro solver in CRK-HACC follows the initial design philosophy of HACC in order to achieve easy portability and full scalability on all modern supercomputing platforms including heterogeneous CPU+GPU systems. Full details of the CRK-HACC framework are provided in [Frontiere et al. \(2022\)](#). Although originally designed for large-volume, high dynamic-range cosmology simulations, CRK-HACC is easily adapted to the lower dynamic range milieu characteristic of the Ly $\alpha$  forest, as described further below.

The CRK-HACC simulations performed here allow for radiative cooling and heating with the assumption that the gas remains optically thin while exposed to a spatially uniform and time-varying ultraviolet background. We assume a primordial abundance of hydrogen and helium with recombination, collisional ionization, and cooling rates taken from [Lukić et al. \(2015\)](#). The energy evolution of each individual particle is evaluated from the resulting cooling function using the exact integration scheme of [Townsend \(2009\)](#). In these runs, the central densities of collapsed objects tend to grow quite large since we have not enabled any astrophysical feedback mechanisms. To avoid computational inefficiencies associated with resolving these over-cooled regions (due to its Lagrangian nature, CRK-HACC is au-

<sup>1</sup> <https://amrex-astro.github.io/Nyx/>

tomatically adaptive) we create collisionless star particles out of any gas particle whose density exceeds 1000 times the mean density and has a temperature below  $10^5$  K (this is analogous to the QUICKLYA option in the Gadget code). We have checked that this does not impact the Ly $\alpha$  measurements, which are instead sensitive to lower density regions.

There are two parameters controlling the force resolution within CRK-HACC. In the first place, the gravitational force resolution has a fixed scale that is set by the Plummer softening length,  $r_{\text{soft}}$ . Our fiducial choice for each run here is to set the softening length to be 1/8 the mean inter-particle separation (i.e.,  $r_{\text{soft}} = L/8N$  where  $L$  is the box size and  $N$  is the one-dimensional particle count of either dark matter or baryon particles). This was used for each run except for H\_L10R78 where we set  $r_{\text{soft}} = L/16N$  to ensure that the softening length was no larger than  $5 \text{ h}^{-1} \text{ kpc}$  in any of the runs. The hydrodynamic force resolution, on the other hand, is spatially adaptive to the local gas density, as specified by the SPH smoothing lengths,  $h$ , of individual baryon particles. In general, the number of neighbors encompassed within a sphere of radius  $h$  centered on each baryon particle is  $n_{\text{nb}} = (4\pi/3)n_h^3$ , where  $n_h$  is an adjustable parameter. The standard choice in CRK-HACC takes  $n_h = 4$  meaning that each smoothing sphere contains roughly 268 neighbors.

### 2.3 The Simulation Set

We performed a set of 12 simulations, comprising of 6 CRK-HACC runs and 6 Nyx runs. Simulation box sizes of 10 and 40  $\text{h}^{-1} \text{ Mpc}$  were used, with the number of cells/gas particles varying from  $128^3$  to  $2048^3$ . A summary of the simulation characteristics is presented in Table 2.3. We also use a set of seven additional CRK-HACC simulations with  $256^3$  gas particles to test the convergence of Ly $\alpha$  statistics with respect to the gravitational softening length,  $r_{\text{soft}}$ , and the number of neighbors to define the SPH smoothing length,  $n_{\text{nb}}$ . For the number of neighbors, we use  $n_{\text{nb}} = 268$  for most of the simulations. Lowering  $n_{\text{nb}}$  was found to provide increased contrast in the density field, but it also increases noise due to the high-order nature of the CRKSPH formalism used in CRK-HACC. We also use  $n_{\text{nb}} = 113$  to study the impact of the number of neighbors on the convergence of Ly $\alpha$  flux statistics in low-resolution simulations. Both CRK-HACC and Nyx include radiative heating from an ultraviolet background as well as radiative cooling from a primordial composition of hydrogen and helium. We use the ‘‘middle reionization’’ ultraviolet background model from Oñorbe et al. (2017). We emphasize that, aside from the discussion presented above, no parametric tuning was attempted to bring the codes into better agreement for the comparison runs.

We generated initial conditions for a  $\Lambda$ CDM cosmology consistent with the WMAP-7 (Komatsu et al. 2011) measurements. In particular, we have  $\Omega_b = 0.045$ ,  $\Omega_m = 0.27$ ,  $\Omega_\Lambda = 0.73$ ,  $h = 0.70$ ,  $\sigma_8 = 0.8$  and  $n_s = 0.95$ . Initial conditions were generated at a starting redshift of  $z = 99$  using the Zel’dovich approximation (Zel’dovich 1970) to perturb  $2 \times N^3$  dark matter plus baryon particles each arranged on a uniform mesh that are offset from each other by half the mesh spacing. The Zel’dovich displacement field was computed on a  $2048^3$  ( $4096^3$ ) mesh for the 10 (40)  $\text{h}^{-1} \text{ Mpc}$  box allowing particle counts up to  $N = 1024$  (2048) without requiring any interpolation from the displacement field. In this way, all runs at fixed box size pertain to the same random realization of the universe. The simulations were run to a final redshift of  $z = 2$  (except for the H\_L10R10 and H\_L40R20 runs which were stopped at  $z = 3$  due to the computational cost of simulating further) and we focus on Ly $\alpha$  measurements in the range  $2 \leq z \leq 5$  relevant for both high and medium resolution surveys.

## 3 NUMERICAL METHODS

In this section we describe the simulation output data and how it is processed to produce synthetic Ly $\alpha$  forest sightlines. For each simulation, we have snapshots at  $z = 2, 3, 4$  and  $5$ , except for H\_L10R10 and H\_L40R20 for which we do not include  $z = 2.0$ . For most of the comparisons, we include only redshifts  $z = 3$  and  $z = 5$ , which are relevant for medium and high resolution observational data sets, respectively.

The first step to compute Ly $\alpha$  forest quantities is to extract sightlines for each simulation box. For the Nyx outputs, we take skewers along the three simulation axes, keeping periodic boundary conditions with the rays passing through all cell centers. For CRK-HACC outputs, we deposit particle quantities on a grid and follow the same procedure. In general, there are a number of ways in which particle data can be interpolated to a grid. We choose here to use an SPH scattering method that interpolates particle quantities onto cell centers via an SPH kernel. This is the most natural choice to create a volume-filling interpolation that makes use of the smoothing lengths internally computed within CRK-HACC. In contrast, methods such as nearest grid point (NGP) or cloud-in-cell (CIC) interpolation will suffer from empty cells particularly in low-density regions that are important for the Ly $\alpha$  calculation.

The SPH interpolation works in the following manner: for each particle quantity  $\psi$ , we compute the field  $\psi(x_i)$  at each mesh point  $\mathbf{x}_i$  via

$$\psi(\mathbf{x}_i) = \frac{\sum_j \psi_j W_{ij} V_j}{\sum_j W_{ij} V_j}, \quad (1)$$

where  $V_j = m_j/\rho_j$  is the volume of a particle with mass  $m_j$  and density  $\rho_j$  while  $W_{ij} = W(\mathbf{x}_i - \mathbf{x}_j, h_j)$  is the SPH weight of a particle at location  $\mathbf{x}_j$  with smoothing length  $h_j$ . The sums in Eqn. (1) are performed over all  $j$  particles whose smoothing spheres intersect with the mesh point  $\mathbf{x}_i$  while the functional form of  $W_{ij}$  corresponds to a 4th-order ( $C^4$ ) Wendland function (Dehnen & Aly 2012). The particle quantities that are interpolated in this manner include density, temperature, and velocity. Each of these quantities are calculated internally within the CRKSPH formalism of CRK-HACC along with the smoothing lengths of each individual particle. We find that for good convergence of Ly $\alpha$  quantities we need a grid with  $N_{\text{grid}} = 2 \times N_{\text{part}}$ , where  $N_{\text{grid}}$  is the number of grid elements on each side and  $N_{\text{part}}$  is the number of particles on each side.

After initial comparisons of the CRK-HACC and Nyx results, we found that the main differences in their Ly $\alpha$  measurements could be attributed to differences in the density field. Visual inspection of sightlines showed that the CRK-HACC density field was noticeably smoother and slower to converge with increasing resolution compared to Nyx. Interestingly, this was only observed in the density field whereas temperature and velocity displayed stronger agreement with Nyx. The contrast in the density field could be enhanced by lowering the SPH neighbor count,  $n_{\text{nb}}$ , used in CRK-HACC. However, making this number too small ( $n_h \leq 2$  or  $n_{\text{nb}} \lesssim 34$ ) resulted in numerical issues during the evolution of the simulation as the high-order CRKSPH solver is sensitive to low particle sampling noise (Frontiere et al. 2017).

To investigate the discrepancy in density fields further, we carried out separate CRK-HACC Ly $\alpha$  measurements with a density field computed using a DTFE method, which is known to have good properties for density estimation at low particle densities, limiting the effects of shot noise (e.g., see Kaehler et al. 2012, for a visual comparison of density fields obtained from SPH and tessellation methods). In this case, a parallel DTFE code (Rangel et al. 2016) developed specifi-

| Name    | Box size [Mpc/h] | Gas cells/particles | Average resolution [kpc/h] |
|---------|------------------|---------------------|----------------------------|
| L10R78  | 10               | 128 <sup>3</sup>    | 78                         |
| L10R39  | 10               | 256 <sup>3</sup>    | 39                         |
| L10NR20 | 10               | 512 <sup>3</sup>    | 20                         |
| L10R10  | 10               | 1024 <sup>3</sup>   | 10                         |
| L40R39  | 40               | 1024 <sup>3</sup>   | 39                         |
| L40R20  | 40               | 2048 <sup>3</sup>   | 20                         |

**Table 1.** The set of simulations with the name of the simulation, the box size in  $h^{-1}$ Mpc, the number of gas cells/particles and the equivalent average resolution in  $h^{-1}$ kpc, which is the exact resolution in the Nyx case and the mean inter-particle separation in the CRK-HACC case. When referring to Nyx and CRK-HACC outputs we write N\_ and H\_ respectively, followed by the simulation label.

cally for cosmological applications was used to compute a density<sup>2</sup> for each particle that is subsequently interpolated to the grid via Eqn. (1), and led to much improved results, as discussed further below. In this way, we are replacing only the SPH density estimate with a DTFE density estimate while preserving the use of an SPH interpolation kernel to map the particle density to the grid. We also tried replacing the SPH grid interpolation with a DTFE grid interpolation for which the DTFE particle densities are linearly interpolated to grid points, but found that this only served to pick up small-scale noise in the DTFE density field along the gridded sightlines. In other words, the advantage of the DTFE is to provide a better particle density estimate in low-density regions while not necessarily providing a better particle to grid interpolation scheme. In what follows, we use the notation CRK-HACC.sph and CRK-HACC.dtf to distinguish between Ly $\alpha$  measurements derived from density fields obtained using the SPH density taken from the simulation output and from the DTFE, respectively. Note that both cases use the same temperature and velocity fields.

The second step is to compute the (normalized) transmitted flux  $F$  at every pixel, with  $F = e^{-\tau}$  where  $\tau$  is the optical depth for Ly $\alpha$  photon scattering. The latter is defined as

$$\tau_\nu = \frac{\pi e^2}{m_e c} f_{12} \int \frac{n_{\text{HI}}}{\Delta \nu_D} \phi_\nu dr, \quad (2)$$

where  $\nu$  is the observed frequency,  $e$  is the electron charge,  $c$  is the speed of light,  $f_{12}$  is the oscillator strength for the Ly $\alpha$  resonance transition,  $n_{\text{HI}}$  is the neutral hydrogen density,  $\Delta \nu_D = (b_T/c)\nu_0 = (\sqrt{2k_B T/m_H}/c)\nu_0$  is the Doppler width with  $b_T$  the Doppler parameter,  $m_H$  the mass of hydrogen,  $k_B$  the Boltzmann constant and  $\phi_\nu$  is the line profile.

In general, the line profile is a Voigt profile, but we use the Doppler profile instead for several reasons. For line center optical depths of less than 1000, the Doppler profile gives identical fluxes. We are only interested in Ly $\alpha$  forest systems, which have line center optical depths less than 10. For Lyman Limit Systems (LLS) and Damped Ly $\alpha$  systems (DLAs), our simulations are not designed to produce the correct density and temperature for HI density mapping in any case. The HI density in these systems should have self-shielding corrections, which cannot be evaluated properly without coupled

radiative transfer-hydrodynamics in the simulations. If we were to use Voigt profiles with these high column density systems, the damping wings would not only be inaccurate, but those errors would then contaminate nearby regions. The Doppler profile is

$$\phi_\nu = \frac{1}{\sqrt{\pi}} \exp \left[ - \left( \frac{\nu - (1 - \frac{v_{\parallel}}{c})\nu_0}{\Delta \nu_D} \right)^2 \right]. \quad (3)$$

In velocity space, peculiar velocities modify the optical depth by shifting the absorption positions and broadening the lines. Thus, in redshift space we have

$$\tau_\nu = \frac{\pi e^2 f_{12} \lambda_0}{m_e c H} \int \frac{n_{\text{HI}}}{v_{\text{th}}} \exp \left[ - \left( \frac{\nu - \nu' - v_{\parallel}}{v_{\text{th}}} \right)^2 \right] dv' \quad (4)$$

where  $\lambda_0$  is the rest-frame wavelength and  $H$  is the Hubble expansion rate at the given redshift. All the Ly $\alpha$  fields computation, along with the different flux statistics presented in the following section are performed using the gimlet postprocessing software (see for instance Friesen et al. 2016). For each output, we use  $N_{\text{gas}}^2 \times 3$  lines of sight, where  $N_{\text{gas}}$  is the number of resolution elements.

#### 4 MORPHOLOGICAL COMPARISON

In this section, we compare morphological differences between the most resolved outputs from the L10R10 runs. This comparison is only qualitative in order to highlight the differences in the properties of the simulation data between the codes. It aims at building our understanding of the impact of these two different numerical methods on observable Ly $\alpha$  flux statistics, which are studied in the following sections.

In Fig. 1, we show one example skewers of CRK-HACC.dtf, CRK-HACC.sph and Nyx at  $z = 3$  and  $z = 5$ , with the baryon density,  $\rho_b$ , the temperature,  $T$ , and the redshift space flux  $F = e^{-\tau}$ . The skewer on the left pierces low to moderate densities with  $-1.0 \leq \log(\rho_b/\bar{\rho}_b) \leq 0.5$ , i.e. the typical density regime from which originates the Ly $\alpha$  signal. For CRK-HACC skewers, we used a grid with 2 times more cells than particles in each dimension. The skewers display good agreement between Nyx and CRK-HACC. Although the initial conditions are the same, there will be differences in how structure will develop in the codes, as a function of position at low redshift. We observe two systematic discrepancies between Nyx and CRK-HACC.sph. First, the CRK-HACC.sph results tend to overestimate the density in under-dense regions, e.g., at  $x \sim 3.5 h^{-1}$ Mpc or  $x \sim 7.3 h^{-1}$ Mpc (indicated by dashed grey vertical lines at  $z = 3$ ), the Nyx voids are systematically deeper than those in CRK-HACC.sph. Second, while the CRK-HACC.sph and Nyx skewers capture the same overall structures, the former misses some small-scale fluctuations

<sup>2</sup> The DTFE density is obtained by first applying a Voronoi tessellation to the full set of gas particles. This stage is computed in parallel using a uniform volume decomposition across MPI ranks with ghost zones used to pad the boundaries. The end product of the Voronoi tessellation is the assignment of a Voronoi cell to each gas particle defining the contiguous region of space for which that gas particle would be the nearest neighbor of any point in that volume. We define the DTFE density for each particle as the inverse volume of its Voronoi cell multiplied by its mass.

because of the inherent smoothing of the density estimate with the chosen parameter choices. Using the DTFE to estimate the density field of CRK-HACC significantly improves the agreement between the two codes in the very low density regions (with voids being deeper) but still tends to over-smooth small-scale fluctuations. The density fields from the same skewer in lower resolution runs at  $z = 3$  can be found in App. A1, where typical differences between the two codes are easier to catch by eye than on the high-resolution outputs.

Fig. 2 shows an example slice through our simulations. We do not show slices for the DTFE CRK-HACC density field as the main differences have already been discussed. We show the baryon density and temperature and the transmitted flux from the L10R10 simulations at  $z = 3$ . We also include differences of the simulated fields. The LOS axis is plotted vertically, so in the flux images, one can imagine each column of pixels as a separate spectrum. Again, the qualitative agreement is very good and we see the same structures in all panels. The major differences in density arise near the high-density regions, near the bottom and top clusters and along the filaments. The highest density regions are more extended in CRK-HACC due to the formation of star particles out of high-density gas particles. Nyx, on the other hand, allows baryons to cool and collapse much more, producing more concentrated halo density profiles and thinner filaments, as seen by the blue stripes along filaments. We observe similar patterns in the temperature maps, with the largest differences mostly in accretion shocks enveloping the filaments and halos, where the shocks in CRK-HACC are slightly puffier. Structures appear very similar in the other panels, but in flux, they are distorted in position along the LOS by peculiar velocities and broadened along the LOS by thermal broadening. Where the other panels also show smooth transitions between extreme values, the flux often has sharp features along the LOS, quickly changing from a saturated region (red) to a transmissive region (blue). The largest differences in flux once again come from the high density regions, where different structures are created by the formation of star particles in CRK-HACC and over-cooling in Nyx. These largest differences arise in saturated or almost saturated regions so that they do not strongly impact the Ly $\alpha$  forest. However, subtler differences arise in the low-density regions, which more significantly impact Ly $\alpha$  flux statistics, as we will discuss in Section 6.

## 5 DENSITY AND TEMPERATURE STATISTICS

We now consider the density and temperature statistics in the simulations. Fig. 3 shows the temperature density diagrams of the baryonic component in the most resolved simulations, N\_L10R10 and H\_L10R10, at  $z = 3$  and  $z = 5$ , which illustrate the phases of the IGM as captured in both codes. We do not use results for the DTFE snapshots in order to ensure that we show a self-consistent relationship between temperature and density values computed internally within the simulation. The PDF is calculated as the volume-weighted histogram of  $\log \rho_b$  and  $\log T$ . We note that the CRK-HACC particle data is not used directly, as this would result in different spatial sampling. Instead, we use the SPH values deposited on the uniform grid. We also fit the  $\rho_b$ - $T$  relation in the diffuse IGM, which is a key component of the robustness of Ly $\alpha$  forest predictions. It is approximated by the power law

$$T = T_0 \left( \frac{\rho_b}{\bar{\rho}_b} \right)^{\gamma-1}, \quad (5)$$

where  $T_0$  is the temperature at the mean density and  $\gamma$  is the slope of the power law, both of which are set by the ionizing background.

The fit is performed using cells in the low-density regime, with  $10^{-1} \bar{\rho}_b \leq \rho_b \leq 10 \bar{\rho}_b$  and  $T \leq 10^5$  K. We show the evolution of  $T_0$  and  $\gamma$  with redshift for all the simulations in  $10 h^{-1}$  Mpc boxes in Fig. 4.

Returning to Fig. 3, we note that the diffuse phase is similar between the two codes, as expected. The  $\rho_b$ - $T$  relation overlaps at all redshifts for converged simulations, although the temperature in CRK-HACC might be slightly higher at low redshift, and slightly lower at high redshift. The activation of star formation at high density in CRK-HACC is clearly apparent in Fig. 3, since points past  $\rho_b \sim 10^4 \bar{\rho}_b$  on the galaxy phase tail are systematically lost. Finally, CRK-HACC is much faster to converge on  $T_0$  and  $\gamma$ , in particular at high redshift where there is very little improvement with spatial resolution.

## 6 FLUX STATISTICS

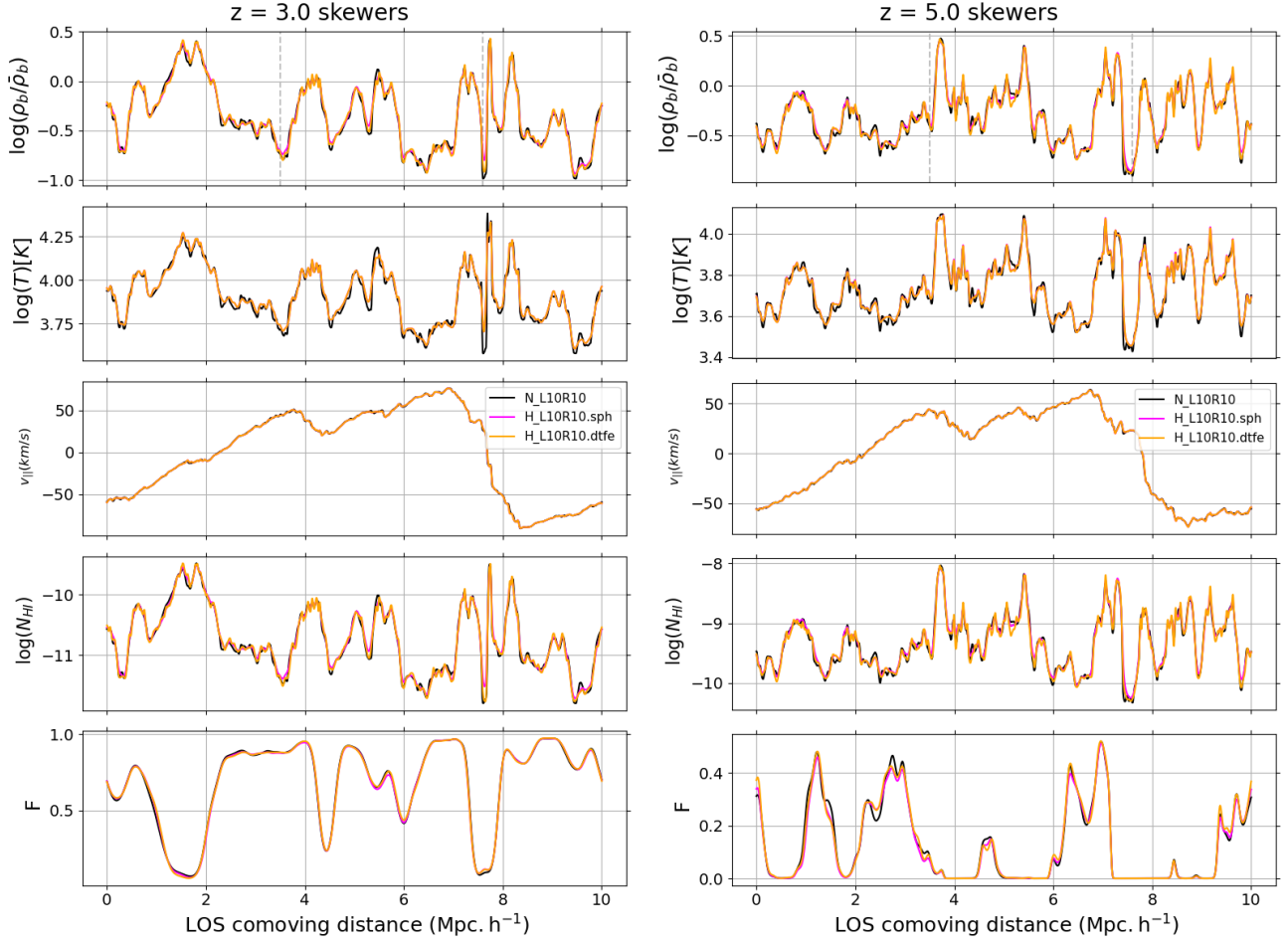
In this section, we compare the observationally relevant Ly $\alpha$  flux statistics that are commonly used to constrain astrophysics and cosmology, the 0-, 1-, and 2-point flux statistics, i.e., the mean flux, the flux probability distribution function (PDF) and the flux power spectrum. We show the comparison in the small box at  $z = 3$  and  $z = 5$ , which are the most relevant for observations. We will discuss box-size effects in the following section.

### 6.1 Mean flux

The simplest Ly $\alpha$  flux statistic is the mean transmitted flux,  $\langle F \rangle$ , or equivalently the effective optical depth  $\tau_{\text{eff}} = -\log(\langle F \rangle)$ , which is a measure of the opacity in the IGM. Observations shows that it smoothly evolves from low to high values with decreasing redshift as the expansion gradually lowers the proper hydrogen density and the UVB intensity decreases the neutral hydrogen fraction. Current mean flux measurements vary between different groups, e.g. at  $z = 2.8$   $\bar{F} \sim 0.69$  at the 2-3% level (Palanque-Delabrouille et al. 2015b; Walther et al. 2019) and  $\bar{F} \sim 0.72$  at the 1% level (Becker et al. 2007; Becker & Bolton 2013). We compute the mean flux using all pixels in the box. Fig. 5 shows the mean flux as a function of redshift for all  $10 h^{-1}$  Mpc boxes at all snapshots, varying spatial resolution only.

The two codes have a similar convergence trend, with higher discrepancies at higher redshifts. Low-resolution runs have more absorption than the high resolution ones for Nyx and CRK-HACC.sph as they overestimate the density in low-density regions. But we observe a non-monotonic convergence for CRK-HACC.dtfе at high redshift and an opposed sense of convergence at low redshifts. For low redshifts,  $2 \leq z \leq 3$ , both Nyx and CRK-HACC converge at the percent level, with differences between N\_L10R10 - N\_L10R20 and H\_L10R10 - H\_L10R20 below 1%, using either SPH or DTFE for the density field. At high redshifts,  $4 \leq z \leq 5$ , the mean flux is less converged with respect to spatial resolution, but still within observational uncertainties with 8%, 1% and 2% differences for CRK-HACC.sph, CRK-HACC.dtfе and Nyx respectively. The increase in the spatial resolution requirement with redshift to model the Ly $\alpha$  forest is fully expected because of three effects. First, the comoving filtering scale increases with redshift. Then, the majority of the Ly $\alpha$  signal comes from mildly-dense regions at  $z = 2.0$  and from under-dense regions at  $z = 5$ , which are more difficult to capture (Bolton & Becker 2009). And finally, thermal broadening is weaker at higher redshifts.

An important finding is that the Nyx results converge faster than CRK-HACC.sph, as the former requires a lower average resolution than the latter to reach convergence. In Fig. 6 we show the mean effective



**Figure 1.** Skewers from the snapshots in  $10 h^{-1}\text{Mpc}$  with  $1024^3$  cells/particles (L10R10) at  $z = 3$  (left) and  $z = 5$  (right) for Nyx in black and CRK-HACC using the SPH and DTFE methods to estimate the density fields in magenta and orange respectively. The baryon density  $\rho_b$  (first row), the temperature  $T$  (second row), the velocity parallel to the LOS (third row), the neutral Hydrogen density  $n_{\text{HI}}$  (fourth row) and the redshift-space flux  $F$  (fifth row) are displayed. We show a skewer that goes through a low density region (for a more obvious visual difference between the SPH and DTFE results, see the lower-resolution comparison data shown in Fig. A1 in the Appendix), typical of the density regime from which originates the Ly $\alpha$  forest signal.

resolution as a function of density for the two codes when varying the number of cells/particles. For Nyx, given that all cells have the same size, we simply take the resolution  $r$  as

$$r = \frac{L}{N_{\text{cells}}}, \quad (6)$$

where  $L$  is the box size in  $h^{-1}\text{kpc}$  and  $N_{\text{cells}}$  the number of cells in one dimension. For CRK-HACC we define the effective resolution as the local inter-particle separation which we derive from the normalized smoothing length,  $h/n_h$ . We then bin particles in density and plot the mean normalized smoothing length for each bin in Fig. 6.

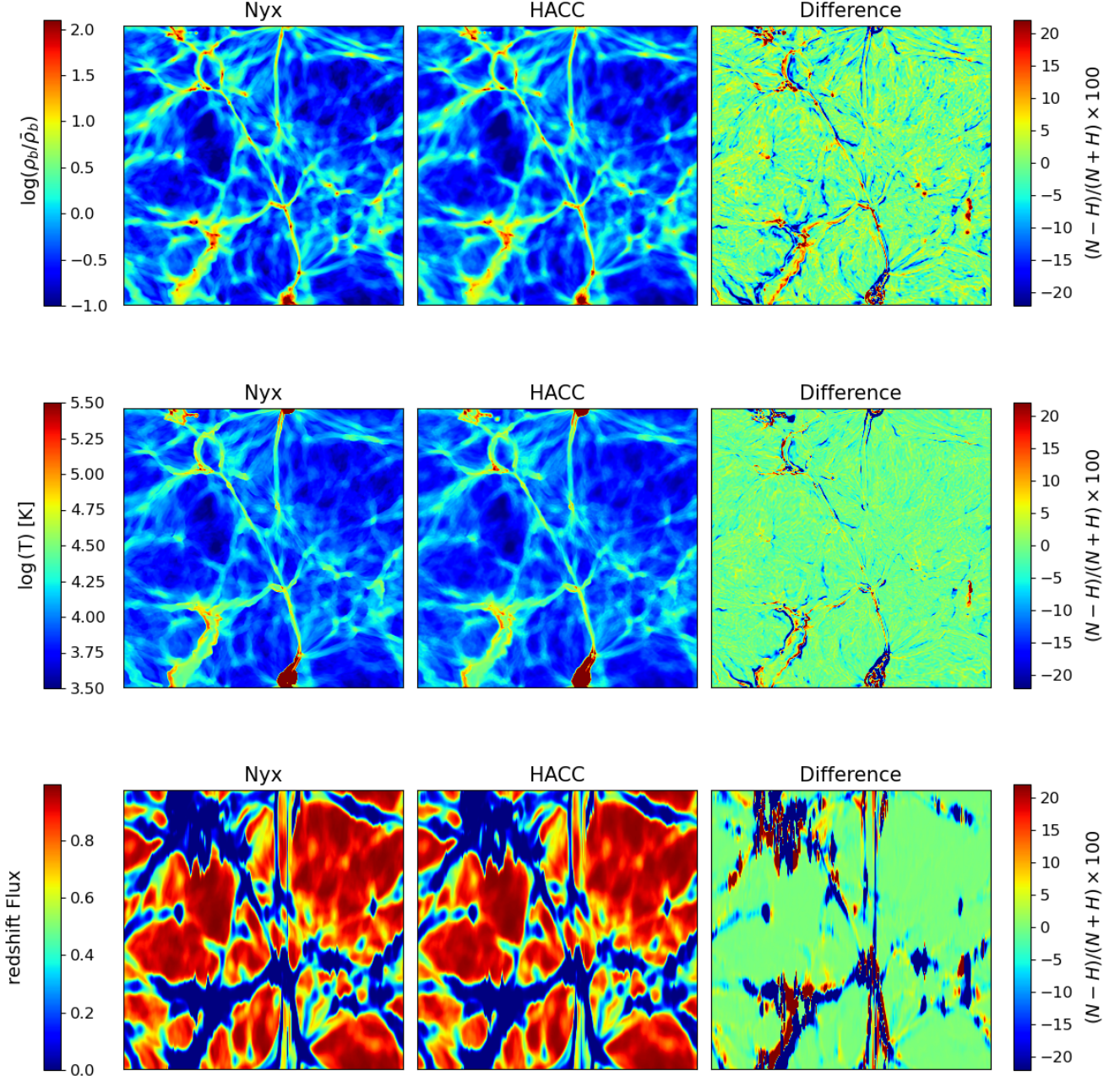
It is clear that for the same number of particles/cells, if the average resolution is the same, the effective resolution in low-density regions is better for Nyx, while it is better in high-density regions for CRK-HACC. This does not come as a surprise, as particles are attracted to high-density regions at the expense of low-density regions. Therefore, a Lagrangian code will always need more particles than cells for an Eulerian code to reach the same effective resolution in the Ly $\alpha$  forest. We also highlight the density regions from which the Ly $\alpha$  signal originates at  $z = 2$  and  $z = 5$  (the lighter and darker gray bars in Fig. 6). To define the minimum and maximum density of these regions, we consider cells with flux  $0.85 \leq F \leq 0.95$  and

$0.1 \leq F \leq 0.2$  respectively, and take the median density. This comparison of effective resolution with density nicely shows why Nyx and CRK-HACC are closer at low redshift compared to high redshift. The use of DTFE significantly improves the estimation of the density field in the lower density regions, making the internal convergence of the mean flux of CRK-HACC as fast as the convergence of Nyx. We also find that the improvement when using DTFE increases with redshift because of the increased importance of lower density regions to the Ly $\alpha$  signal at  $z = 5$  than at  $z = 2.0$ .

Equally importantly, when comparing converged results (in the sense of being “internally converged”, i.e., when increasing the spatial resolution does not induce any further changes), the Nyx and two CRK-HACC methods each converge to the same answer with differences below 1% for  $2 \leq z \leq 3$  and 2.5% for  $4 \leq z \leq 5$ . The behavior of the mean flux evolution, however, is a little different between the two codes when using the DTFE density field. At the best resolution, Nyx starts with a higher mean flux at high redshift and ends up with a lower mean flux at low redshift than CRK-HACC. We note that the number of neighbors parameter,  $n_{\text{nb}}$ , for the CRK-HACC runs has no impact on the mean flux.

It is common practice to rescale the effective optical depth so that

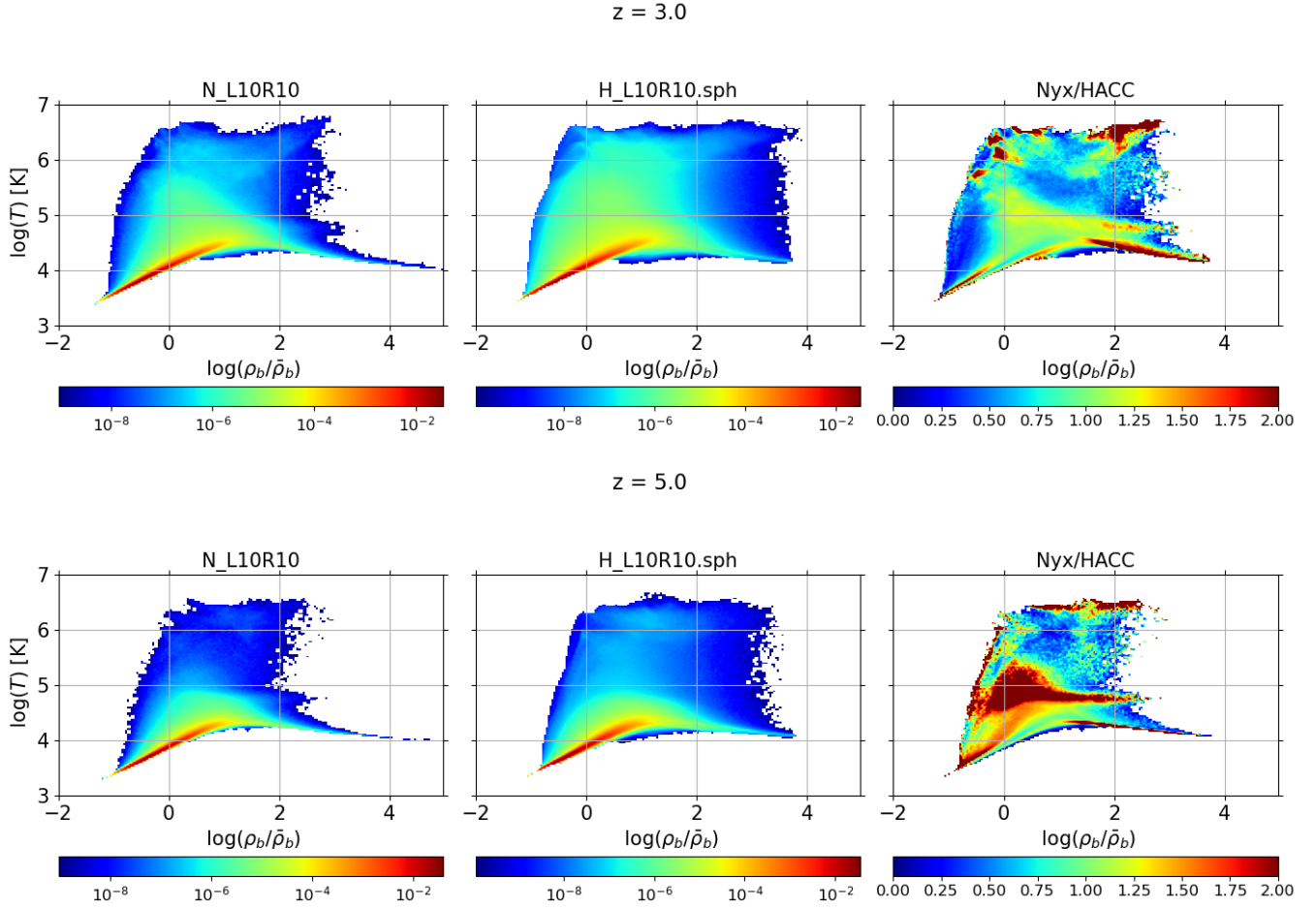




**Figure 2.** Slices through snapshots of the  $10 h^{-1}\text{Mpc}$  with  $1024^3$  cells/particles simulations (L10R10) at  $z = 3$  of Nyx (left) and CRK-HACC (middle) and the difference (right) at  $z = 3$ , showing the baryon density  $\rho_b$  (top), the temperature  $T$  (middle) and the transmitted flux  $F$  (bottom). The CRK-HACC density was estimated with the SPH method.

the simulated mean flux matches the observed mean flux, which is roughly equivalent to adjusting the intensity of the specific UVB used in the simulation. Indeed simulations of the IGM must assume a particular ionizing background and common prescriptions of the ionizing background provide a table of the photoionization rate and heating per atom for the primordial species HI, HeI, and HeII at many redshifts. The reality is that our constraints on the ionizing background are much weaker than our knowledge of cosmology but this uncertainty can be parametrized by rescaling the simulated

optical depth (Lukić et al. 2015), which is more or less equivalent to changing the photoionization rate  $\Gamma$  in the simulation since  $\Gamma \propto n_{\text{HI}} \propto 1/\Gamma_{\text{HI}}$ . For relative flux fluctuations  $\delta_F = F/\langle F \rangle - 1$ , a small difference in mean flux directly propagates as an amplitude shift in power spectra. For all other flux statistics, we rescale the optical depths by a small amount to match N\_L10R10, i.e.  $\bar{F} = 0.66051489$  at  $z = 3$  and  $\bar{F} = 0.11177497$  at  $z = 5$ .



**Figure 3.** Temperature-density diagrams for Nyx (left) and CRK-HACC (middle) outputs in a  $10 h^{-1} \text{Mpc}$  box with  $1024^3$  cells/particles (L10R10) and their ratio (right). We show results at  $z = 3$  (top) and  $z = 5$  (bottom). Here we use the SPH method to estimate the CRK-HACC density field. The high gas density tail is not seen in the CRK-HACC simulation since it has been converted into stars.

## 6.2 Flux PDF

The flux PDF, the probability density function of the pixel fluxes, has been a probe of interest for a long time, estimated with both medium-resolution (Desjacques et al. 2007; Lee et al. 2015) and high-resolution quasar spectra (McDonald et al. 2000; Lidz et al. 2006; Kim et al. 2007; Calura et al. 2012; Rollinde et al. 2013; Rorai et al. 2017; Gaikwad et al. 2021) with precision between 5-10%, to measure the amplitude of matter fluctuations or the thermal history of the IGM through  $T_0$  and  $\gamma$  (Bolton et al. 2008; Viel et al. 2009; Calura et al. 2012; Lee et al. 2015; Rorai et al. 2017; Gaikwad et al. 2021). We compute the flux PDF,  $P(F)$ , in 50 equally spaced flux bins with the integral constraint  $\int P(F) dF = 1$ . The results are shown in Fig. 7 at  $z = 3$  and  $z = 5$ .

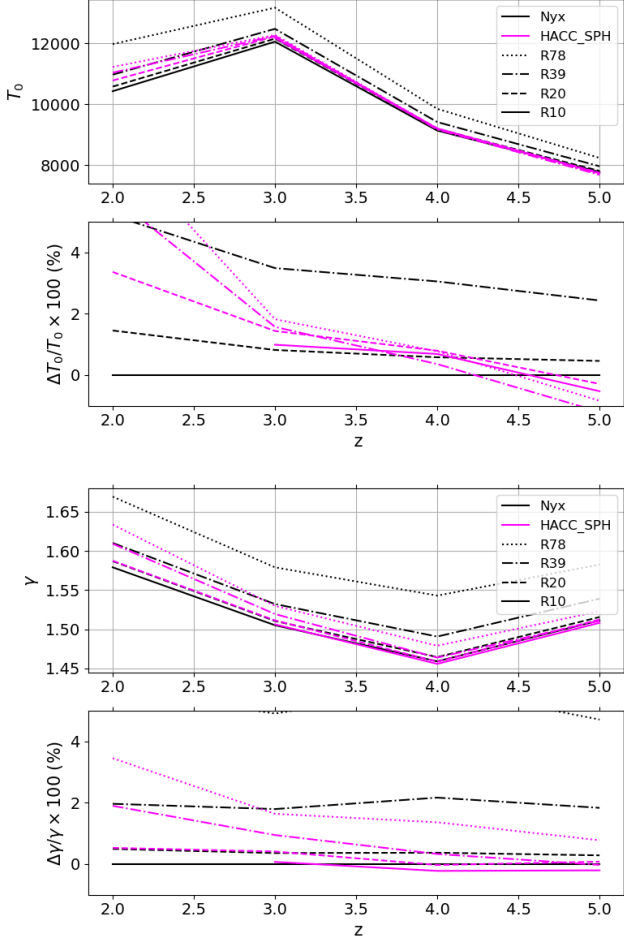
We observe that in this case the resolution requirement also increases with redshift. At  $z = 3$ , the convergence is relatively rapid for both codes and they agree with each other at the 1% level except for  $F \sim 0$  and  $F \sim 1$ . Nyx has a larger probability at the extremas and the integral constraint pushes the intermediate probabilities down relative to CRK-HACC values. At  $z = 5$ , we observe a threshold in the flux above which the simulation cannot produce the flux value, since the low-density gas is missed in the low-resolution runs. This threshold is systematically higher for Nyx for the same number of

cells/particles, for the resolution-related reasons detailed in Sec. 6.1. For low-medium flux values, the results internally converge at the 2% level for Nyx and CRK-HACC.dtf, while only 5-10% for CRK-HACC.sph; all three point to the same answer, with differences  $\leq 2\%$  between H\_L10R10 and N\_L10R10. We note that the improvement between CRK-HACC.sph and CRK-HACC.dtf is larger at  $z = 5$  since the DTFE algorithm better represents low density regions, the origin of the Ly $\alpha$  signal at  $z = 5$ , as opposed to the lower redshift signal that originates from low- to mid-density regions.

## 6.3 1D power spectrum

Spatial correlations of the Ly $\alpha$  flux constitute an extremely powerful cosmological probe of matter density fluctuations at small scales (as small as 1 Mpc) and high redshifts. This increasingly used 2-point statistic has given the strongest constraints to date on neutrino masses and dark matter models (Yèche et al. 2017b; Iršič et al. 2017; Armengaud et al. 2017; Palanque-Delabrouille et al. 2019) and has also produced competitive measurements on the IGM thermal history (Walther et al. 2018; Boera et al. 2019; Gaikwad et al. 2021).

The first category of  $P_{1D}$  measurements have been performed using tens of thousands of medium-resolution SDSS quasar spec-

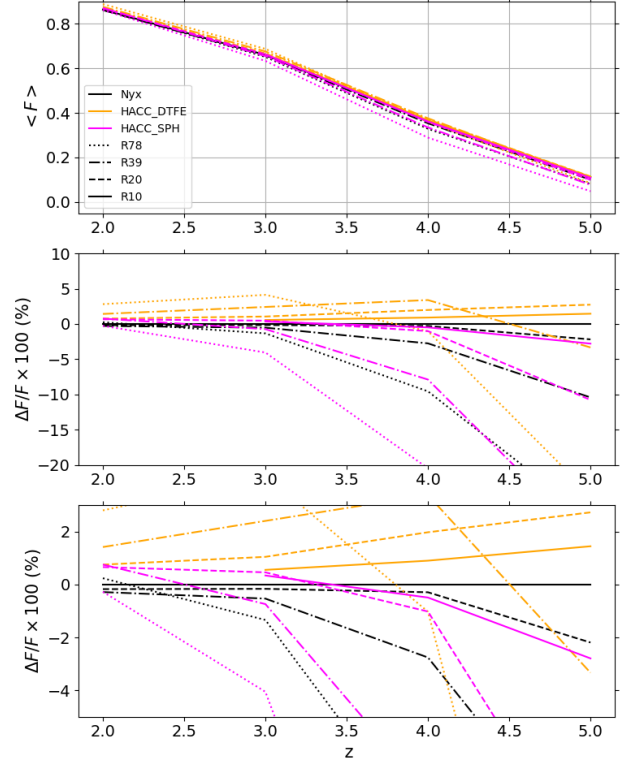


**Figure 4.**  $T_0$  and  $\gamma$  evolution with redshift for all  $10 h^{-1} \text{Mpc}$  simulations. Bottom panels show relative differences with respect to N\_L10R10. For CRK-HACC we use the SPH method to estimate the density field.

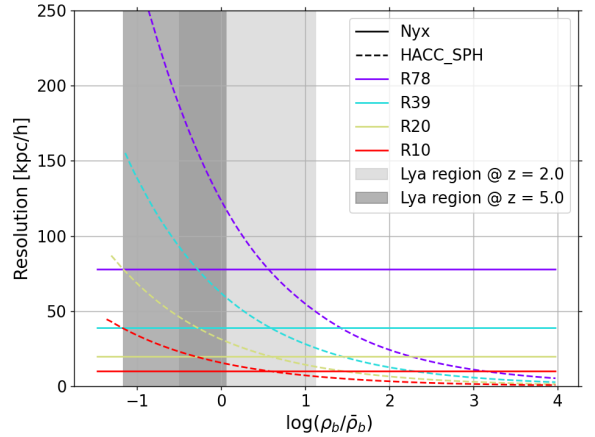
tra (Palanque-Delabrouille et al. 2013; Chabanier et al. 2019). They measure flux fluctuations on scales  $0.0015 \text{ km}^{-1} \text{ s} \leq k \leq 0.02 \text{ km}^{-1} \text{ s}$  with great precision, down to 1.5% at low redshifts and large scales, thanks to the extremely favorable statistics of the quasar samples. However they miss the small-scale flux fluctuations, important to study alternative DM models or the IGM thermal history, due to lack of resolution and low SNR. By increasing the number of Ly $\alpha$  quasar spectra by a factor of 4 and improving the resolution and noise, the on-going DESI survey will probe scales smaller than its predecessor, in the range  $0.0015 \text{ km}^{-1} \text{ s} \leq k \leq 0.04 \text{ km}^{-1} \text{ s}$ , with percent level precision at low redshifts.

The second category used tens to hundreds of quasar spectra (Viel et al. 2013a; Iršič et al. 2016; Yèche et al. 2017b; Walther et al. 2018; Boera et al. 2019; Gaikwad et al. 2021; Karaçaylı et al. 2022) with high resolution and high S/N allowing them to probe smaller scales than the former category but without the large scales required to constrain cosmology, with scales in the complementary range  $0.02 \text{ km}^{-1} \text{ s} \leq k \leq 0.1 \text{ km}^{-1} \text{ s}$ . The higher resolution comes at the expense of a higher exposure time resulting in smaller data sets, and therefore a lower precision, i.e., 5-15%.

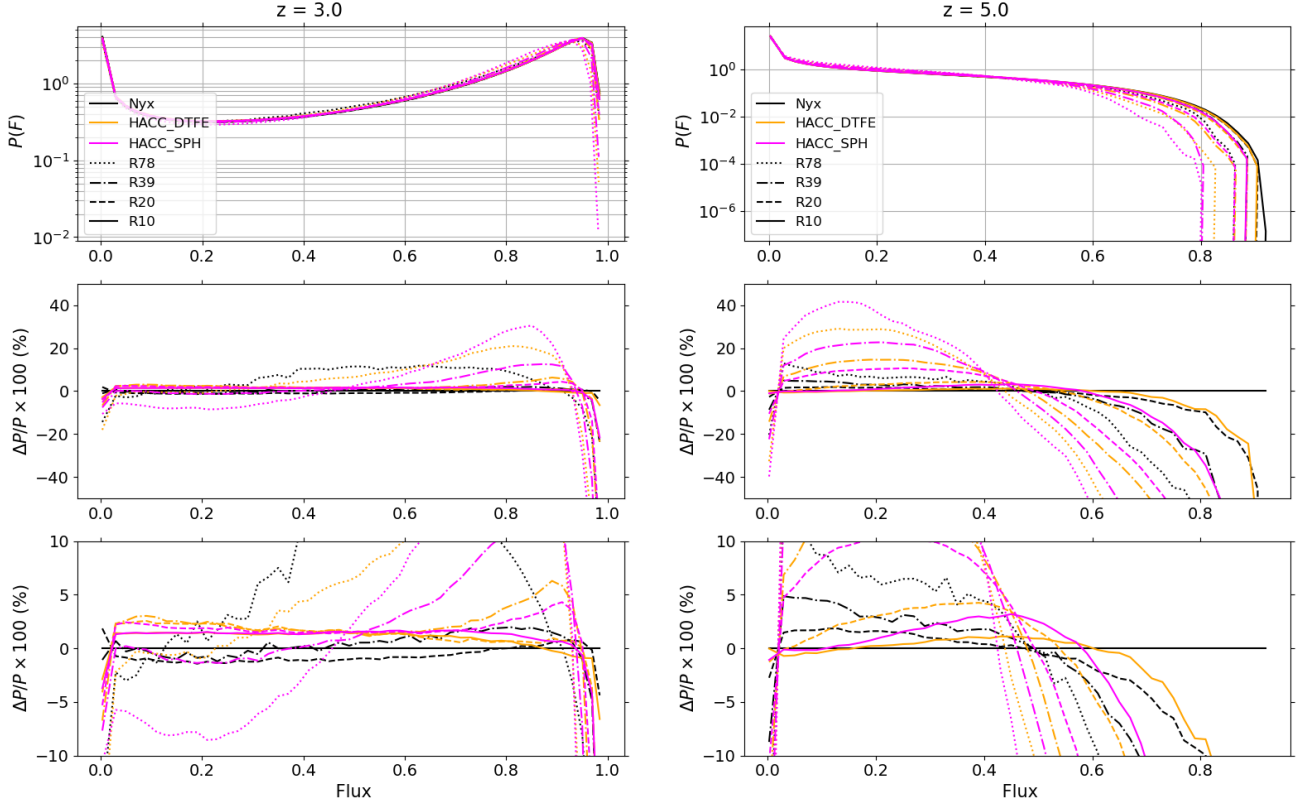
Analogously to density fluctuations, we can define the flux perturbations as  $\delta_F = F / \langle F \rangle - 1$ , where  $\langle F \rangle$  is averaged over all



**Figure 5.** Mean flux as a function of redshift for all Nyx and CRK-HACC simulations in  $10 h^{-1} \text{Mpc}$  boxes for different spatial resolution (top). Middle and bottom panels show relative differences with respect to N\_L10R10, the most-resolved Nyx simulation with  $10 h^{-1} \text{kpc}$  resolution.



**Figure 6.** Effective resolution as a function of baryon density for Nyx (solid) and CRK-HACC.sph (dashed) runs in  $10 h^{-1} \text{Mpc}$  boxes varying spatial resolution. The light gray and dark gray shaded regions show the density regime from which the Ly $\alpha$  forest originates at  $z = 5$  and  $z = 2$  respectively.



**Figure 7.** Flux PDF for Nyx and CRK-HACC at  $z = 3$  (left) and  $z = 5$  (right). Middle and bottom panels show relative differences with respect to N\_L10R10, the most resolved Nyx simulation with  $10 \text{ h}^{-1} \text{ kpc}$  resolution.

skewers. Then, the 1D Ly $\alpha$  flux power spectrum,  $P_{1D}$  is obtained by Fourier transforming  $\delta_F$  along each line of sight and averaging in  $k_{\parallel}$ . We display the dimensionless  $P_{1D}$  in Fig. 8 at  $z = 3$  and at  $z = 5$  in the  $10 \text{ h}^{-1} \text{ Mpc}$  boxes. The shaded areas indicate observational uncertainties averaged over the wavelength range for DESI-like surveys and high-resolution surveys in darker and light gray respectively. DESI-like uncertainties cover the large-scale range  $0.005 \text{ km}^{-1} \text{ s} \leq k \leq 0.04 \text{ km}^{-1} \text{ s}$  and are from the forecast of Valuri et al. (2022), which is based on eBOSS uncertainties (Chabanier et al. 2019), with a factor 4 increase in the number of quasars and improved resolution. We use the  $z = 4.6$  forecast for redshift 5.0, given that 4.6 is the highest redshift bin available in the eBOSS measurement. For the high-resolution uncertainties, we use those from Karaçaylı et al. (2022) and (Boera et al. 2019) for  $z = 3$  and  $z = 5$  respectively, in the  $0.02 \text{ km}^{-1} \text{ s} \leq k \leq 0.1 \text{ km}^{-1} \text{ s}$  wavelength range. All the aforementioned error estimates represent statistical and systematic uncertainties added in quadrature. In the following paragraphs, we first compare the convergence of Nyx with that of CRK-HACC.sph, then we introduce the results of CRK-HACC.dtf. Finally, we investigate how results evolve when artificially increasing the spatial resolution using Richardson extrapolation.

### 6.3.1 Comparison of Nyx and CRK-HACC.sph outputs

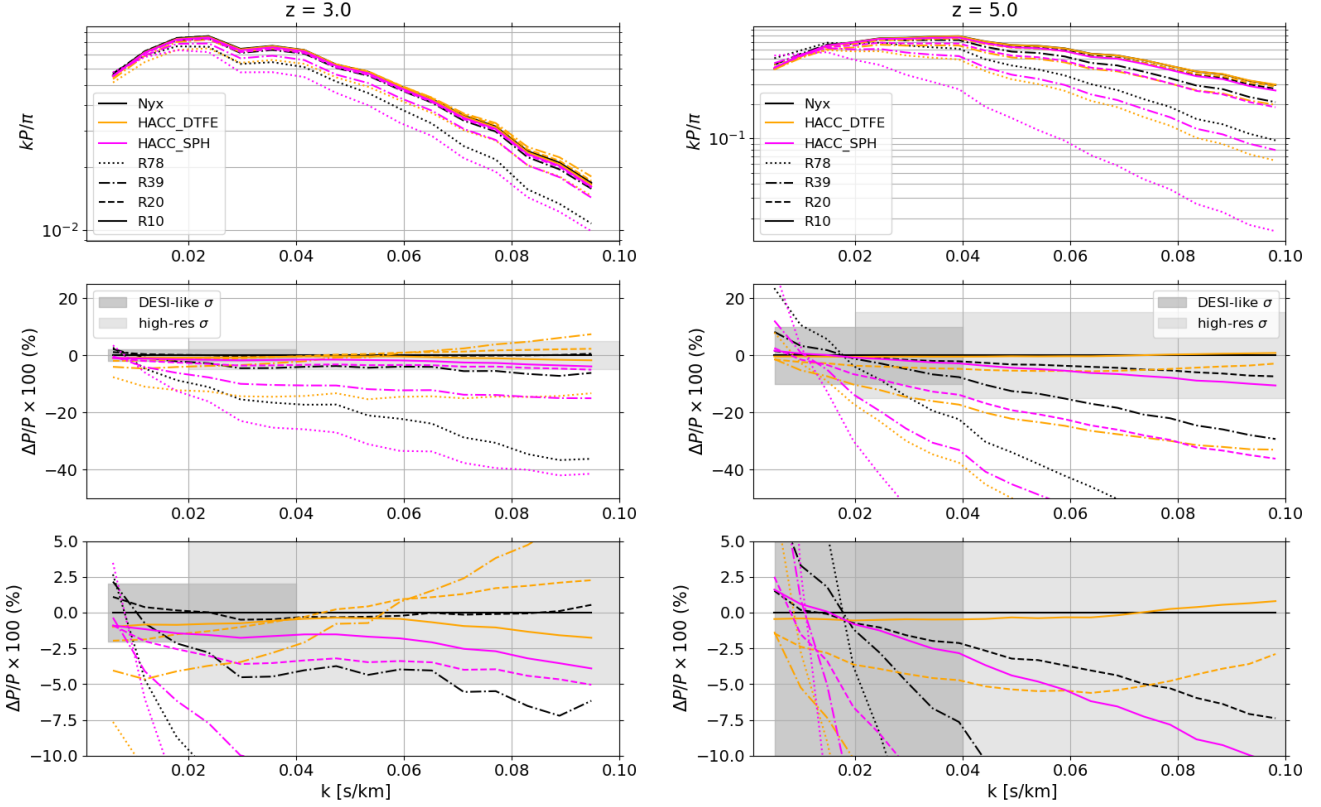
For both codes and as a function of redshift, lower-resolution runs tend to have more large-scale power, with a larger tilt, so that they fall off at smaller scales than higher-resolution runs. These larger discrepancies at small scales are expected because flux fluctuations are more sparsely sampled so there is less power for the low-resolution runs.

Similarly for the previous flux statistics, the increase in spatial resolution requirement with redshift to model the Ly $\alpha$   $P_{1D}$  is because of the larger filtering scale at high redshift, the signal originating from lower-density regions, and of weaker thermal broadening.

At  $z = 3$ , the convergence is again faster for Nyx, indeed N\_L10R20 and N\_L10R10 are almost exactly the same at all scales, even as small as  $0.1 \text{ km}^{-1} \text{ s}$  proving that N\_L10R20 is already fully converged, in agreement with previous findings (Tytler et al. 2009; Lukić et al. 2015). CRK-HACC.sph is slower to converge as there are still differences between H\_L10R20 and H\_L10R10, at the  $\sim 2\%$  level. When comparing the most resolved power spectra, the agreement between Nyx and CRK-HACC.sph is at the limit of observational uncertainties. For large scales,  $k \leq 0.04 \text{ km}^{-1} \text{ s}$ , probed by medium resolution surveys, Nyx and CRK-HACC runs tend to the same answer at the 1.5% level, and disagree at only  $< 5\%$  at small scales, i.e.  $0.04 \leq k/\text{km}^{-1} \text{ s} \leq 0.1$ , probed by high-resolution surveys.

At  $z = 5$ , the convergence is significantly slower for both codes, with better convergence for Nyx at all scales. Agreement between N\_L10R20 and N\_L10R10 is better than 2.5% on large scales and better than 7.5% on small scales. In contrast, agreement between H\_L10R20 and H\_L10R10 is only at 10% on large scales and 30% on small scales. Finally, the agreement between Nyx and CRK-HACC.sph at the best spatial resolution is better than current observational uncertainties, but still quite large, about 2.5% on large scales and 10% on small scales.

Finally, as discussed in the previous sections, CRK-HACC.sph requires more particles to be converged, i.e., H\_L10R10 is converged but H\_L10R20 is not, as opposed to Nyx for which  $20 \text{ h}^{-1} \text{ kpc}$  resolution is enough to be converged with respect to current observational



**Figure 8.** 1D Ly $\alpha$  power spectrum for Nyx (black lines) and CRK-HACC (orange and magenta lines for the DTFE and SPH density estimation, respectively) at  $z = 3$  (left) and  $z = 5$  (right) in the  $10 h^{-1}\text{Mpc}$  boxes for different spatial resolutions. Middle and bottom panels show relative differences with respect to N\_L10R10, the most-resolved Nyx simulation with  $10 h^{-1}\text{kpc}$  resolution. The shaded regions show  $k$ -averaged observational uncertainties of DESI-like (darker gray) and high-resolution (lighter gray) surveys. Note the excellent agreement between the high-resolution Nyx and CRK-HACC results, when using DTFE with the latter.

uncertainties. Also, at equivalent average resolution, we find that CRK-HACC.sph has a systematically lower Ly $\alpha$   $P_{1D}$  compared to Nyx. This is in accordance with the previous results shown for the individual skewers (Fig. 1) or the effective resolution as a function of density (Fig. 6) from which it is clear that CRK-HACC.sph computes systematically higher densities in low-density regions and is thus more opaque to Ly $\alpha$  transmission. Nevertheless, we find that the differences between CRK-HACC.sph and Nyx at equivalent average resolution decrease with increasing resolution – they are at the limit of the observational uncertainties when comparing the best-resolution runs. The requirement of  $10 h^{-1}\text{kpc}$  average resolution for CRK-HACC.sph at  $z = 3$  looks reasonable from Fig. 6 given that the effective resolution is about  $20 h^{-1}\text{kpc}$  on the range of density producing the Ly $\alpha$  signal, as shown with the black dashed line (CRK-HACC.sph resolution) lying below the magenta solid line ( $20 h^{-1}\text{kpc}$  constant resolution of Nyx). For  $z = 5$  though, precision modeling would require an average resolution of  $5 h^{-1}\text{kpc}$  at least.

If the difference between the two codes at the best resolution is at the limit of observational uncertainties, there are still 1.5% scale-independent and scale-dependent biases at large and small scales respectively. We perform a series of tests to investigate the sources driving these discrepancies.

- We lower the numbers of neighbors to  $n_{\text{nb}} = 113$ , instead of  $n_{\text{nb}} = 268$  for the low-resolution run H\_L10R39. This improves the results by 2% and 10% at most at small scales at  $z = 3$  and  $z = 5$  (see

App. A2). Therefore, we do not expect this parameter to significantly alter the results at higher resolution.

- We take the Nyx velocity instead of the CRK-HACC.sph velocity to compute the flux power spectra and the results remain unchanged.

- If the  $\gamma$  thermal parameters are very close for H\_L10R10 and N\_L10R10, the temperature at mean density,  $T_0$ , shows a few percent difference (see Fig. 4). We compute the flux power spectra for H\_L10R10 by rescaling the temperature and density so that the new  $T_0$  and  $\gamma$  match that of N\_L10R10, and recomputing the resulting neutral hydrogen fraction, but it shows very little change (0.5% at the smallest scales at  $z = 3$ ).

- We remove the effect of high density regions, since we see in Figs. 1 and 2 that these regions are quite different, notably because of the activation of collisionless star particles in CRK-HACC.sph and the difference in the velocity field. We remove these regions by thresholding to densities less than 50 times the mean density. If the density of a cell exceeds  $50\bar{\rho}_b$ , we set  $\rho_b = 50\bar{\rho}_b$  and set the temperature of the gas using the fluctuating Gunn-Peterson approximation-like equation of state following Eqn. 5, with  $T_0$  and  $\gamma$  matching the simulation fit. We then recompute the optical depth and Ly $\alpha$  flux the regular way, and the resulting flux power spectra show no significant deviation.

- From Figs. 1 and 6, showing skewers and the resolution as function of density in the simulations, we observe biased sampling in void regions for CRK-HACC.sph. To assess the impact of this discrepancy on the flux power spectrum, we use the Nyx density instead of the

CRK-HACC.sph density, when the latter is below a certain threshold,  $\rho_{\text{thresh}}$ , and we recompute the optical depth the regular way. We gradually increase  $\rho_{\text{thresh}}$  from  $-0.75$  to  $0.5$  in logarithmic units of  $\bar{\rho}_b$ . We find that values of  $\log(\rho_{\text{thresh}}/\bar{\rho}_b) = -0.5$  and  $0.5$  for  $z = 5$  and  $z = 3$  respectively, are enough to make the CRK-HACC.sph and Nyx results agree at better than 1% at all scales, confirming that the low density regions are responsible for these remaining discrepancies. These results and further discussion are available in App. A3.

### 6.3.2 Comparison with CRK-HACC.dtfc

The realization that the low-density regions are responsible for the main discrepancies between Nyx and CRK-HACC.sph is the driving motivation for using DTFE in order to improve the sampling of low-density regions. The corresponding Ly $\alpha$   $P_{1D}$  results for CRK-HACC.dtfc outputs are shown in Fig. 8 in orange.

We first observe that, as opposed to Nyx and CRK-HACC.sph, the convergence of CRK-HACC.dtfc is non-monotonic at  $z = 3$  as the different power spectra oscillate around the final most converged power spectrum, reflecting the DTFE sensitivity (as a local method) to particle sampling noise. The CRK-HACC.dtfc power spectra have some times more power at small scales than the most resolved Nyx power spectrum. Also, the convergence is faster to reach for CRK-HACC.dtfc than for CRK-HACC.sph, especially at  $z = 5$ .

More importantly, the most converged power spectra of Nyx and CRK-HACC.dtfc agree on the whole range of scales at the percent level at  $z = 3$  and at the sub-percent level at  $z = 5$ . Power spectra of L10R20 of Nyx and CRK-HACC.dtfc already agree at 2% at  $z = 5$ , whereas they are not fully converged with respect to spatial resolution. Unlike CRK-HACC.sph, an average resolution of  $20 \text{ h}^{-1} \text{ kpc}$  appears sufficient to consider the Ly $\alpha$  power spectra as being converged at the percent level at  $z = 3$  and at a few percent at  $z = 5$ , considerably decreasing the computational time needed to reach the required precision given by observational surveys. We reach a better agreement between Nyx and CRK-HACC.dtfc at  $z = 5$  since the Ly $\alpha$  signal comes from lower density regions at  $z = 5$  than at  $z = 3$ , due to the improvement from using the DTFE algorithm. Note that this did not involve any fine-tuning of parameters, neither for the simulation run nor for the analyses, to get the results in such remarkable agreement. This is to our knowledge the first time that Ly $\alpha$  statistics results from grid-based and SPH hydrodynamical simulations have been brought to such an extreme level of agreement, thanks to the better sampling of low density regions with the DTFE algorithm, which strongly increases our confidence that both codes are converging to the correct answer.

### 6.3.3 Richardson extrapolation

In this section we study the agreement between the codes and their convergence properties when extrapolated to the ‘‘continuum’’ solution using Richardson extrapolation. This technique has been used to study the convergence properties of cosmological simulations where a priori predictions are difficult to make or not available (e.g., Lidz et al. 2006; Heitmann et al. 2010; Lukić et al. 2015), but for completeness we repeat the main reasoning here. A numerical method that is  $n$ -th order accurate in space produces a numerical approximation  $P(h)$ , where  $h$  is the discretization element, such that

$$P(h) = P + Ah^n + O(h^{n+1}), \quad (7)$$

with  $P$  being the exact value,  $Ah^n$  the leading error, and the last term the higher-order error. In general, the order of accuracy, especially for multi-dimensional, multi-physics codes like those considered here,

is best derived from the numerical solutions themselves, even given a priori expectations for the convergence. Using three numerical solutions for  $P$  with three spatial resolutions  $h$ ,  $rh$  and  $r^2h$ , where  $r$  is the constant refinement ratio, and assuming that convergence is uniform, the order of accuracy is:

$$n = \frac{\ln\left(\frac{P(r^2h) - P(rh)}{P(rh) - P(h)}\right)}{\ln(r)}. \quad (8)$$

Knowing  $n$ , Richardson extrapolation allows to extrapolate  $P$  to the  $n$ -th order of accuracy ‘‘continuum’’ solution,  $h \rightarrow 0$ , using two numerical results for  $P$ :

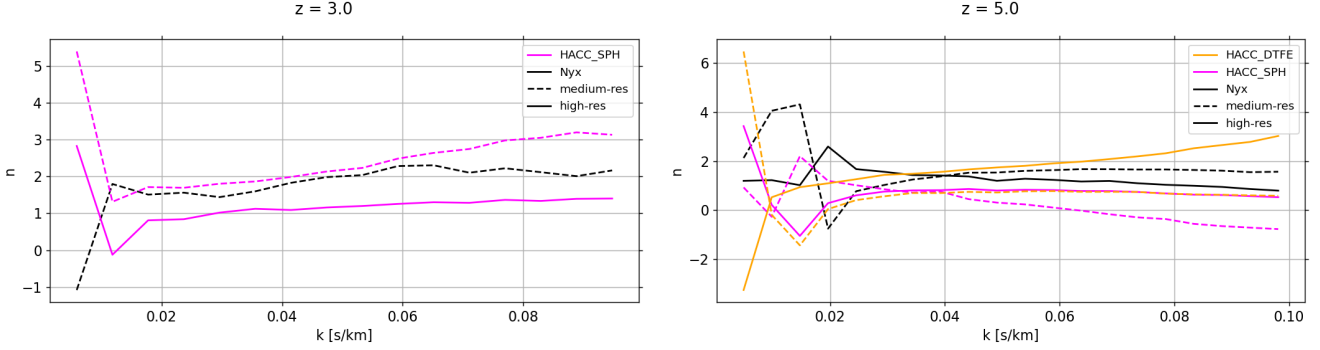
$$P_R = \frac{r^n P(h) - P(rh)}{r^n - 1}. \quad (9)$$

This ‘‘extrapolation to the limit’’ is usually checked at three resolution values to verify the value of  $n$ , before the final extrapolation is implemented.

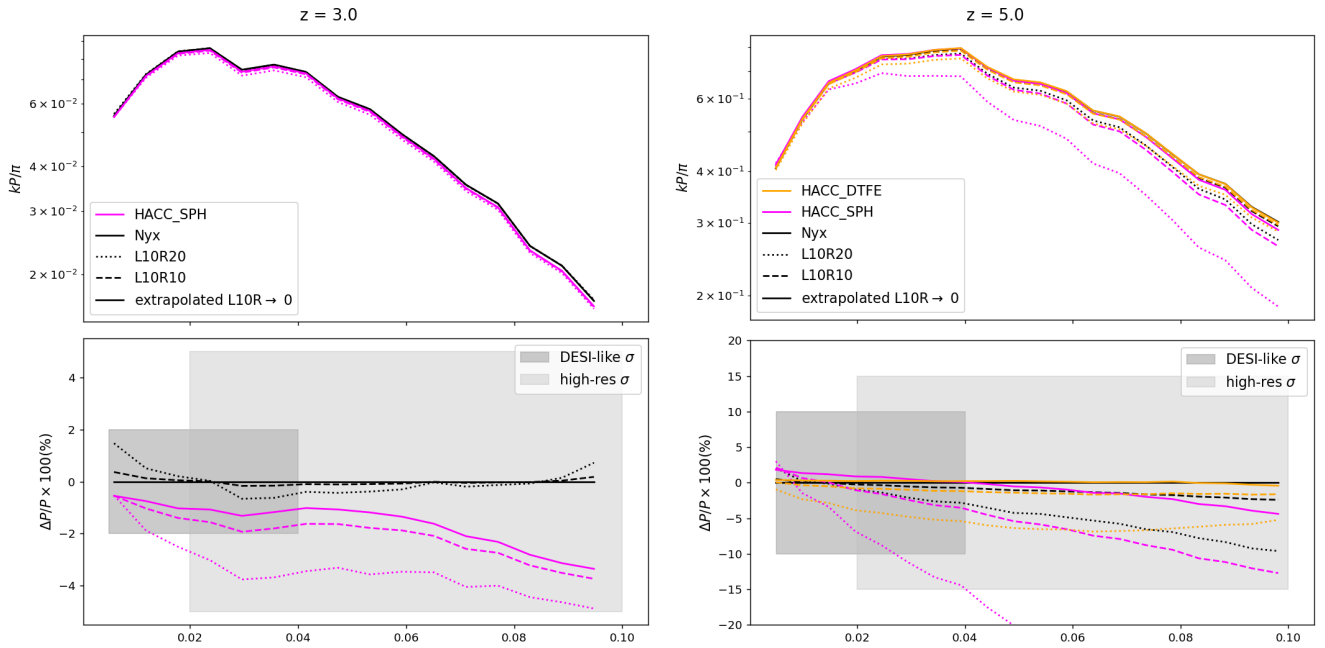
To quantitatively assess the convergence of  $P_{1D}$ , we computed the order of accuracy,  $n$ , along with the differences between the 1D Ly $\alpha$  power spectra derived from the simulation,  $P_N$  and  $P_H$  for Nyx and CRK-HACC respectively, and the power spectra derived from the Richardson extrapolation,  $P_R$ , at  $z = 3$  and  $z = 5$ , as shown in Fig. 9. Note that we do not include results for CRK-HACC with DTFE outputs at  $z = 3$  since Richardson extrapolation requires monotonic convergence which is not the case with DTFE as shown in Fig. 8.

We use the medium-resolution set, composed of the L10R78, L10R39 and L10R20 runs, and the high-resolution set, composed of the L10R39, L10R20 and L10R10 runs, to compute  $n_{\text{medium-res}}$  and  $n_{\text{high-res}}$  and assess how it evolves with increasing spatial resolution. At  $z = 3$ , we do not display  $n_{\text{high-res}}$  for Nyx because  $P_{N\_L10R10}$  and  $P_{N\_L10R20}$  are almost exactly matching (see Fig. 8) and thus the calculation for  $n$  becomes dominated by sub-percent numerical noise differences between the two runs. For Nyx, at both  $z = 3$  and  $z = 5$ ,  $n$  is in remarkable agreement with the theoretical value  $n = 2$ . For CRK-HACC, at  $z = 3$   $n$  is also nicely converging to  $n = 2$ , with a slightly lower value than Nyx. At  $z = 5$ , for CRK-HACC.sph,  $n$  is negative when using the medium resolution set and tends toward 1 when using the high-resolution set, which is not surprising because the power spectrum is quite far from being properly convergent at this redshift. For CRK-HACC.dtfc,  $n_{\text{medium-res}}$  almost exactly matches  $n_{\text{high-res}}$ , and  $n_{\text{high-res}}$  is in very good agreement with the  $n = 2$  theoretical value, except at very small scales where it diverges toward high values.

Finally, to further assess the convergence and agreement with respect to the spatial resolution for both Nyx and CRK-HACC.sph, we use Richardson extrapolation to extrapolate power spectra to the limiting case of an infinitely small resolution element. We show the results at  $z = 3$  and  $z = 5$  in Fig. 10. At  $z = 3$ , as expected for Nyx, extrapolation does not improve the result since  $P_{N\_L10R20}$  and  $P_{N\_L10R10}$  are already almost identical. For CRK-HACC.sph, results are hardly improved with differences below 1% at all scales between  $P_{H\_L10R10}$  and  $P_R$ , while  $P_{H\_L10R10}$  and  $P_{H\_L10R20}$  are a few percent different confirming that we need an average resolution of  $10 \text{ h}^{-1} \text{ kpc}$  for CRK-HACC.sph to reach the 1% convergence at  $z = 3$  at scales as small as  $0.1 \text{ km}^{-1} \text{ s}$ . It is important to note however that the  $P_{1D}$  of the extrapolated Nyx and CRK-HACC.sph runs,  $N\_L10R0$  and  $H\_L10R0$ , still display a 1% systematic offset at large scales and scale-dependent bias at small scales up to 2%. But these are at the limit of current observational uncertainties, and also appear to be degenerate with the mean flux on large scales, hence unlikely to impact cosmological parameter inference. At  $z = 5$ , Richardson extrapolation improves the results by 2% at small scales, compared to 10% discrepancies



**Figure 9.** Order of accuracy  $n$  at  $z = 3$  (left) and  $z = 5$  (right) for Nyx (black) and CRK-HACC runs using the SPH (magenta) and DTFE (orange) density field estimators. The power-law index was estimated using the sets of medium resolution (L10R78, L10R39 and L10R20) and high resolution (L10R39, L10R20 and L10R10) simulations (dashed and solid lines, respectively). For more details, see text (Section 6.3.3).



**Figure 10.** 1D Ly $\alpha$  power spectrum for Nyx and CRK-HACC.sph at  $z = 3$  (left) and  $z = 5$  (right) in the  $10 h^{-1} \text{Mpc}$  boxes for different spatial resolution. Middle and bottom panels show relative differences with respect to N\_L10R0, the Richardson extrapolated power spectra with resolution  $R \rightarrow 0$ .

between  $P_{N\_L10R20}$  and  $P_{N\_L10R10}$ . This shows that while there is room for improvement, the errors are still significantly below the observational uncertainties, which are about 10-15% at high redshifts. This demonstrates that  $10 h^{-1} \text{kpc}$  resolution is enough for Nyx, or presumably any grid-based code, at  $z = 5$ .  $P_{R,H\_L10R0}$  displays up to 8% more power compared to  $P_{H\_L10R10}$ , and it is even closer to the Nyx results, with agreement at the 2% level even at small scales.

The results confirm that Richardson extrapolation works well to estimate simulated power spectra to a higher resolution than the native resolution of the simulations. In practice, we could use  $80 h^{-1} \text{kpc}$  and  $40 h^{-1} \text{kpc}$  resolution simulations to extrapolate a converged power spectra for Nyx instead of running a  $20 h^{-1} \text{kpc}$  simulation at redshifts  $z \sim 3$ . For HACC, if using the DTFE density field estimator,  $80 h^{-1} \text{kpc}$  and  $40 h^{-1} \text{kpc}$  resolution simulations are also sufficient to extrapolate a converged power spectrum, whereas if using the SPH density estimator,  $40 h^{-1} \text{kpc}$  and  $20 h^{-1} \text{kpc}$  averaged resolution simulations are required to extrapolate to a converged result.

## 6.4 3D power spectrum

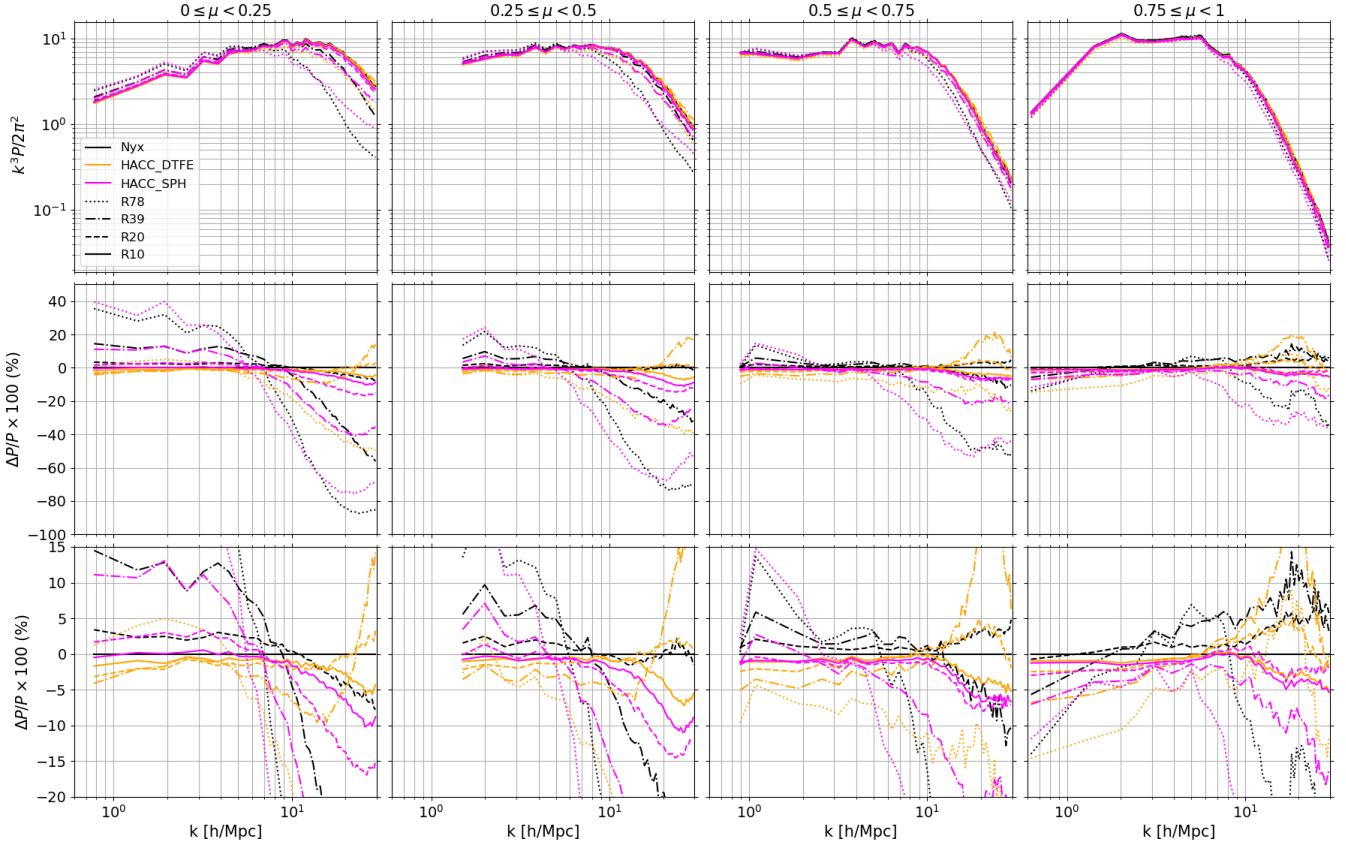
In contrast to the 1D Lyman- $\alpha$  power spectra that yield Ly $\alpha$  correlations only along the LOS, the 3D Lyman- $\alpha$  power spectra,  $P_{3D}$ , gives Ly $\alpha$  correlations in all directions. The two are related by

$$P_{1D}(k_{\parallel}) = \int_0^{\infty} P_{3D}(k_{\perp}, k_{\parallel}) \frac{dk_{\perp} k_{\perp}}{2\pi}. \quad (10)$$

The  $P_{3D}$  has not been measured with observational data yet because of lack of surveys with high enough LOS density. Since the quasar density for DESI will allow such measurements, algorithms have already been developed to estimate  $P_{3D}$  with real data (Font-Ribera et al. 2018). As such, it is timely to prepare the theoretical modeling of these statistics. Here, we only show results at  $z = 3$  given that the first measurement will be performed at low redshifts, where the quasar density is the highest. We show the equivalent results at  $z = 5$  in App. A4.

To visualize the anisotropy of the power spectrum field we follow

$z = 3.0$



**Figure 11.** 3D Ly $\alpha$  power spectra at  $z = 3$  for Nyx (black line) and CRK-HACC (orange and magenta lines for the DTFE and SPH density estimation, respectively) in the  $10 \text{ h}^{-1} \text{ Mpc}$  boxes for different spatial resolutions. Middle and bottom panels show relative differences with respect to N\_L10R10, the most-resolved Nyx simulation with  $10 \text{ h}^{-1} \text{ kpc}$  resolution. The bottom panel shows a more restricted range on the y-axis to make it easier to distinguish the different results.

the convention of using  $P_{3D}(k, \mu)$ , where  $\mu$  is the cosine of the angle between the mode and the LOS, i.e.,  $\mu = k_{\parallel}/k$ . We average the power spectrum in  $4 \mu$  bins. The resulting 3D flux power spectra are shown in Fig. 11. From left to right, the power  $\mu$  wedges shown are increasingly parallel to the LOS. That is, the panel on the left can be thought of as transverse power and the panel on the right as LOS power, i.e., close to  $P_{1D}$ .

The observations on resolution and redshift evolution trends of the  $P_{1D}$  section still hold and can be generalized for every  $\mu$  bin. For Nyx and CRK-HACC.sph, lower-resolution runs have more power at large scales and less power at small scales compared to high-resolution runs because flux fluctuations are more sparsely sampled (the discrepancies are larger at high redshifts, see Fig. A4). In contrast, the lower-resolution power spectra from CRK-HACC.dtfе display a different sense of convergence as a function of  $k$ , with less power at large scales and higher power at small scales. We also observe that the LOS power (right panel) tends to agree much better across resolutions and codes. This is due to the thermal broadening along the LOS, which acts to smooth out fluctuations in this direction. The perpendicular perturbations are not filtered in this way, and show much more small-scale power, which enhances discrimination between simulation results run at different resolutions.

The agreement between the most resolved results is remarkably good for all  $\mu$  bins on the range of scales probed by DESI-like surveys, i.e. up to  $10 \text{ h}^{-1} \text{ Mpc}$ . Surprisingly, the systematic  $\sim 1\%$

offset observed on the largest scales for the  $P_{1D}$  with CRK-HACC.sph (see Fig. 8) disappears as we move to a direction perpendicular to the LOS.

## 7 BOX SIZE EFFECTS

The discussion so far has been limited to the study of flux statistics from the small-box simulations. Although the L10R20 and L10R10 simulations appear to be converged with respect to resolution, they are definitely not converged in box size. By redshift  $z = 2.0$ , the box-scale mode is significantly nonlinear for a  $10 \text{ h}^{-1} \text{ Mpc}$  box size, which accounts for a certain amount of missing power in the simulations. Additionally, the missing large-scale modes also significantly affect the properties of bulk flows in the simulations. This is generally the least converged quantity in small-box simulations, and results in a lower temperature distribution (affecting the broadening of the lines) and smaller peculiar velocities (affecting the redshift-space distortions). In order to compare more representative predictions from both codes, we also consider statistics from larger box simulations, in a box of  $40 \text{ h}^{-1} \text{ Mpc}$  of side, with resolution up to  $20 \text{ h}^{-1} \text{ kpc}$ , since previous studies show that this box-size is enough to avoid finite-size effects (McDonald 2003; Lidz et al. 2006; Tytler et al. 2009; Lukić et al. 2015).

We find results similar to Tytler et al. (2009) and Lukić et al. (2015)



when comparing results between the L10R20 and L40R20 runs for Nyx and CRK-HACC (see App. A5). In particular, the mean flux is larger for the bigger box, with more enhanced discrepancies at high redshift, at the 3% level at most. For the 1D  $Ly\alpha$  power spectrum, in addition to not probing scales as large as the L40 runs, the L10 runs display up to 10% discrepancies with the L40  $Ly\alpha$  power spectra at both redshifts.

More importantly, in the context of this paper, when comparing differences between CRK-HACC and Nyx results for simulations in the  $40 \text{ h}^{-1}\text{Mpc}$  boxes with varying spatial resolution, we reach similar conclusions as in the case of the  $10 \text{ h}^{-1}\text{Mpc}$  boxes (Section 6). Using the SPH density field for CRK-HACC, the Nyx results converge faster for all the flux statistics and the differences between the two codes when comparing the most converged results are the same. In particular, the offsets between the two power spectra at  $z = 3.0$  and  $z = 5.0$  are identical; thus the level of agreement between Nyx and CRK-HACC in terms of  $Ly\alpha$  statistics does not depend on the box size.

## 8 CONCLUSIONS

In this paper, we compare  $Ly\alpha$  forest statistics in numerical simulations using two very different hydrodynamical solvers, the Lagrangian SPH CRK-HACC code (Habib et al. 2016; Frontiere et al. 2017) and the Eulerian grid-based Nyx code (Almgren et al. 2013; Sexton et al. 2021). The primary aim is to establish a level of confidence in theoretical predictions extracted from hydrodynamical simulations to interpret the data from future or ongoing  $Ly\alpha$  surveys, such as DESI (DESI Collaboration et al. 2016) or WEAVER-QSO (Pieri et al. 2016).

We used a set of 12 main simulations, comprising of 6 CRK-HACC and 6 Nyx runs, and 7 supporting CRK-HACC simulations, in box sizes ranging from  $10$  to  $40 \text{ h}^{-1}\text{Mpc}$  with number of cells/particles ranging from  $128^3$  to  $2048^3$ , translating into average resolutions ranging from  $78$  to  $10 \text{ h}^{-1}\text{kpc}$ . Comparing results between the two codes was performed on the same Cartesian grid, which required a choice of density estimation for the CRK-HACC results. We used two approaches – the standard SPH method where density is averaged over all neighbors in a smoothing sphere (CRK-HACC.sph results) and the Delaunay Tessellation Field Estimator (CRK-HACC.dtf results) for which the volume is covered by contiguous and non-overlapping tetrahedra.

Morphological comparisons of line-of-sight skewers and 2D slices show good agreement between the two codes. We find two main discrepancies between the CRK-HACC.sph and Nyx outputs, with the former overestimating the density in low-density regions and missing small-scale fluctuations because of the inherent smoothing nature of the SPH density estimator, as seen in Fig. 1. The use of the DTFE estimator significantly improves the comparisons, as we discuss later below.

We compute the  $Ly\alpha$  flux field for all outputs and compare the 0-, 1-, and 2-point flux statistics, i.e., the mean flux, the flux probability distribution function and the flux power spectrum, the most commonly used statistics to extract astrophysical and cosmological information from observations. For all statistics, both codes have similar convergence trends with higher resolution requirement and larger discrepancies between the codes at high redshifts and/or at small scales, the former being due to the fact that at high redshifts the Jeans length is larger, thermal broadening weaker, and the  $Ly\alpha$  signal originates from lower density regions.

We find that the Nyx results converge significantly faster than CRK-HACC.sph, requiring less cells than the latter to predict converged

$Ly\alpha$  flux statistics. This is partly due to the intrinsically density-adaptive nature of Lagrangian codes, since particles are attracted to high-density regions at the expense of low-density regions making it harder to control resolution in the low-density regime, without increasing the particle sampling density. This behavior is shown in Fig. 6, which presents the effective resolution as a function of density in simulations varying the number of cells/particles. Agreements between the converged predictions of Nyx and CRK-HACC.sph are summarized below:

- Mean flux: better than 1% for  $2 \leq z \leq 3$  and better than 2.5% for  $z \leq 5$
- Flux PDF: better than 5% at  $z = 3$  and better than 30% for  $F \leq 0.8$  at  $z = 5$  since low-flux values are not produced in the CRK-HACC.sph output
- Flux 1D power spectrum: better than 2.5% and 5% at large and small scales respectively at  $z = 3$ , which is at the limit of observational uncertainties, and better than 5% and 10% at large and small scales respectively at  $z = 5$

To understand the sources of the remaining discrepancies between Nyx and CRK-HACC.sph outputs, we test how the flux power spectrum is modified when varying the number of neighbors,  $n_{\text{nb}}$ , taking the Nyx velocity instead of the CRK-HACC velocity to compute the CRK-HACC flux field, rescaling  $T_0$  and  $\gamma$  to match between the two codes, removing high-density regions and using the Nyx density field instead of the CRK-HACC density field in low-density regions to compute the CRK-HACC flux field. We find that differences in the low-density regions between Nyx and CRK-HACC outputs are responsible for the systematic discrepancies between the codes while tests for other sources for possible differences showed negligible impact.

In order to improve the sampling of the lower-density regions in CRK-HACC outputs, we use the DTFE algorithm of Rangel et al. (2016), yielding density fields in much better agreement with Nyx, with more small-scale fluctuations and deeper voids (see Fig. 1). The convergence of  $Ly\alpha$  flux statistics of CRK-HACC.dtf outputs is as fast as those from Nyx; comparisons between converged predictions of Nyx and CRK-HACC.dtf show agreement at:

- Mean flux: better than 1% at all redshifts
- Flux PDF: better than 5% at  $z = 3$  and better than 30% for  $F \leq 0.8$  at  $z = 5$  (i.e. with an upper bound for the flux 30% higher than for CRK-HACC.sph) for the flux PDF
- Flux 1D power spectrum: better than 1.5% at all scales at  $z = 3$  and better than 1% at all scales at  $z = 5$  for the flux power spectrum.

We stress that this is the first time that two completely different hydrodynamical methods, using Eulerian and Lagrangian approaches, agree on  $Ly\alpha$  predictions (or any cosmological observable which requires hydrodynamical modeling in the nonlinear regime!) at the percent level. Moreover, this was achieved without any fine tuning of code parameters, raising our confidence that both methods are converging towards the correct answer. Observational measurements of the  $Ly\alpha$  forest are now also reaching percent level precision, thanks to the strong increase in the number of observed quasars. Those measurements will enable much better constraints of cosmological parameters, including neutrino mass and potential non-CDM models, but only if we are able to produce theoretical models at a similar or better level of accuracy. The comparison of computational methods we present here is, therefore, particularly timely.

In the interest of completeness, we note that at the few percent level  $Ly\alpha$  forest observables are also sensitive to secondary effects related to galaxy formation and evolution, such as AGN and supernova feedback (Viel et al. 2013b; Chabanier et al. 2020; Montero-

Camacho & Mao 2021; Burkhart et al. 2022). Incorporating these effects into hydrodynamical simulations requires implementation of (relatively ad hoc) subgrid models, which are commonly calibrated on astrophysical observations and more strongly differ between different implementations. Such physical models were not a part of the analysis presented here. Nevertheless, we wish to emphasize that the simulations we produced, incorporating atomic cooling and UV background heating, are the simulations commonly used by the broader Ly $\alpha$  community to constrain cosmological and astrophysical parameters with Ly $\alpha$  observables. We foresee that this situation will be stable in the near future, and do not expect simulation campaigns to simultaneously vary cosmological and galaxy formation parameters constraining both from the Ly $\alpha$  forest. Instead, the impact of these effects (or uncertainties associated with them) can be parameterized by running an ensemble of simulations and varying subgrid parameters (as in Chabanier et al. 2020 for the impact of AGN feedback) and including those as corrections in either post-processing of simulation forward models or in the likelihood calculation itself.

## ACKNOWLEDGEMENTS

We thank Eric Armengaud for the forecast of DESI observational uncertainties. This research was supported by the Exascale Computing Project (17-SC-20-SC), a collaborative effort of the U.S. Department of Energy Office of Science and the National Nuclear Security Administration. Lawrence Berkeley Laboratory's work was supported under the U.S. Department of Energy contract DE-AC02-05CH11231. Argonne National Laboratory's work was supported under the U.S. Department of Energy contract DE-AC02-06CH11357. This work used resources of the Oak Ridge Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC05-00OR22725. Additionally, this study utilized resources of the Argonne Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC02-06CH11357. This research also used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility located at Lawrence Berkeley National Laboratory, operated under Contract No. DE-AC02-05CH11231 using NERSC award m3921 in 2021/22.

## DATA AVAILABILITY

The data used for testing the simulation results in this paper will be made available on request.

## REFERENCES

- Agertz O., et al., 2007, *MNRAS*, **380**, 963
- Almgren A. S., Bell J. B., Lijewski M. J., Lukic Z., Van Andel E., 2013, *ApJ*, **765**, 39
- Angulo R. E., Zennaro M., Contreras S., Aricò G., Pellejero-Ibañez M., Stücker J., 2021, *MNRAS*, **507**, 5869
- Armengaud E., Palanque-Delabrouille N., Yèche C., Marsh D. J. E., Baur J., 2017, *Monthly Notices of the Royal Astronomical Society*, **471**, 4606
- Baur J., Palanque-Delabrouille N., Yèche C., Magneville C., Viel M., 2016, *J. Cosmology Astropart. Phys.*, **2016**, 012
- Baur J., Palanque-Delabrouille N., Yèche C., Boyarsky A., Ruchayskiy O., Armengaud É., Lesgourgues J., 2017, *J. Cosmology Astropart. Phys.*, **2017**, 013
- Bautista J. E., et al., 2017, *Astronomy & Astrophysics*, **603**, A12
- Becker G. D., Bolton J. S., 2013, *MNRAS*, **436**, 1023
- Becker G. D., Rauch M., Sargent W. L. W., 2007, *ApJ*, **662**, 72
- Bird S., Vogelsberger M., Sijacki D., Zaldarriaga M., Springel V., Hernquist L., 2013, *MNRAS*, **429**, 3341
- Blanton M. R., et al., 2017, *AJ*, **154**, 28
- Boera E., Becker G. D., Bolton J. S., Nasir F., 2019, *ApJ*, **872**, 101
- Bolton J. S., Becker G. D., 2009, *MNRAS*, **398**, L26
- Bolton J. S., Viel M., Kim T. S., Haehnelt M. G., Carswell R. F., 2008, *MNRAS*, **386**, 1131
- Burkhart B., Tillman M., Gurvich A. B., Bird S., Tonnesen S., Bryan G. L., Hernquist L. E., Somerville R. S., 2022, arXiv e-prints, p. arXiv:2204.09712
- Calura F., Tesconi E., D'Odorico V., Viel M., Cristiani S., Kim T. S., Bolton J. S., 2012, *MNRAS*, **422**, 3019
- Chabanier S., et al., 2019, *J. Cosmology Astropart. Phys.*, **2019**, 017
- Chabanier S., Bournaud F., Dubois Y., Palanque-Delabrouille N., Yèche C., Armengaud E., Peirani S., Beckmann R., 2020, *MNRAS*, **495**, 1825
- Colella P., Woodward P., 1984, *Journal of Computational Physics*, **54**, 174
- Croft R. A. C., Weinberg D. H., Katz N., Hernquist L., 1998, *ApJ*, **495**, 44
- DESI Collaboration et al., 2016, arXiv e-prints, p. arXiv:1611.00036
- Dawson K. S., et al., 2013, *AJ*, **145**, 10
- Dawson K. S., et al., 2016, *AJ*, **151**, 44
- Dehnen W., Aly H., 2012, *MNRAS*, **425**, 1068
- Dekker H., D'Odorico S., Kaufer A., Delabre B., Kotzlowski H., 2000, in Iye M., Moorwood A. F., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 4008, Optical and IR Telescope Instrumentation and Detectors. pp 534–545, doi:10.1117/12.395512
- Desjacques V., Nusser A., Sheth R. K., 2007, *MNRAS*, **374**, 206
- Eisenstein D. J., et al., 2011, *AJ*, **142**, 72
- Font-Ribera A., McDonald P., Slosar A., 2018, *J. Cosmology Astropart. Phys.*, **2018**, 003
- Frenk C. S., et al., 1999, *ApJ*, **525**, 554
- Friesen B., Almgren A., Lukic Z., Weber G., Morozov D., Beckner V., Day M., 2016, *Computational Astrophysics and Cosmology*, **3**, 4
- Frontiere N., Raskin C. D., Owen J. M., 2017, *Journal of Computational Physics*, **332**, 160
- Frontiere N., Emberson J. D., Buehlmann M., Adamo J., Habib S., Heitmann K., Faucher-Giguère C.-A., 2022, arXiv e-prints, p. arXiv:2202.02840
- Gaikwad P., Srianand R., Haehnelt M. G., Choudhury T. R., 2021, *MNRAS*, **506**, 4389
- Gingold R. A., Monaghan J. J., 1977, *MNRAS*, **181**, 375
- Habib S., et al., 2016, *New Astron.*, **42**, 49
- Heitmann K., White M., Wagner C., Habib S., Higdon D., 2010, *ApJ*, **715**, 104
- Hopkins P. F., 2013, *MNRAS*, **428**, 2840
- Hubber D. A., Falle S. A. E. G., Goodwin S. P., 2013, *MNRAS*, **432**, 711
- Hui L., Gnedin N. Y., 1997, *MNRAS*, **292**, 27
- Iršič V., Viel M., Haehnelt M. G., Bolton J. S., Becker G. D., 2017, *Physical Review Letters*, **119**, 031302
- Iršič V., et al., 2016, *Monthly Notices of the Royal Astronomical Society*, **466**, 4332
- Kaehler R., Hahn O., Abel T., 2012, arXiv e-prints, p. arXiv:1208.3206
- Karaçaylı N. G., et al., 2022, *MNRAS*, **509**, 2842
- Kim T. S., Bolton J. S., Viel M., Haehnelt M. G., Carswell R. F., 2007, *MNRAS*, **382**, 1657
- Komatsu E., et al., 2011, *ApJS*, **192**, 18
- Lee K.-G., et al., 2015, *The Astrophysical Journal*, **799**, 196
- Levi M., et al., 2013, arXiv e-prints, p. arXiv:1308.0847
- Lidz A., Heitmann K., Hui L., Habib S., Rauch M., Sargent W. L. W., 2006, *ApJ*, **638**, 27
- Lucy L. B., 1977, *AJ*, **82**, 1013
- Lukic Z., Stark C. W., Nugent P., White M., Meiksin A. A., Almgren A., 2015, *MNRAS*, **446**, 3697
- McDonald P., 2003, *ApJ*, **585**, 34
- McDonald P., Miralda-Escudé J., Rauch M., Sargent W. L. W., Barlow T. A., Cen R., Ostriker J. P., 2000, *ApJ*, **543**, 1
- McDonald P., et al., 2006, *ApJS*, **163**, 80
- McQuinn M., 2016, *Annual Review of Astronomy and Astrophysics*, **54**, 313–362

Meiksin A. A., 2009, *Reviews of Modern Physics*, 81, 1405–1469

Montero-Camacho P., Mao Y., 2021, *MNRAS*, 508, 1262

Oñorbe J., Hennawi J. F., Lukić Z., 2017, *ApJ*, 837, 106

Oñorbe J., Davies F. B., Lukić Z., Hennawi J. F., Sorini D., 2019, *MNRAS*, 486, 4075

Palanque-Delabrouille N., et al., 2013, *Astronomy & Astrophysics*, 559, A85

Palanque-Delabrouille N., et al., 2015a, *Journal of Cosmology and Astroparticle Physics*, 2015, 045–045

Palanque-Delabrouille N., et al., 2015b, *Journal of Cosmology and Astroparticle Physics*, 2015, 011–011

Palanque-Delabrouille N., Yèche C., Schöneberg N., Lesgourgues J., Walther M., Chabanier S., Armengaud E., 2019, arXiv e-prints, p. arXiv:1911.09073

Pedersen C., Font-Ribera A., Rogers K. K., McDonald P., Peiris H. V., Pontzen A., Slosar A., 2021, *J. Cosmology Astropart. Phys.*, 2021, 033

Pieri M. M., et al., 2016, in Reylé C., Richard J., Cambrésy L., Deleuil M., Pécontal E., Tresse L., Vauglin I., eds, SF2A-2016: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics. pp 259–266 (arXiv:1611.09388)

Rangel E., Li N., Habib S., Peterka T., Agrawal A., Liao W.-K., Choudhary A., 2016, in 2016 IEEE International Conference on Cluster Computing (CLUSTER). pp 30–39, doi:10.1109/CLUSTER.2016.40

Rauch M., 1998, *Annual Review of Astronomy and Astrophysics*, 36, 267

Read J. I., Hayfield T., 2012, *MNRAS*, 422, 3037

Regan J. A., Haehnelt M. G., Viel M., 2007, *MNRAS*, 374, 196

Rollinde E., Theuns T., Schaye J., Pâris I., Petitjean P., 2013, *MNRAS*, 428, 540

Rorai A., et al., 2017, *MNRAS*, 466, 2690

Saitoh T. R., Makino J., 2013, *ApJ*, 768, 44

Schaap W. E., 2007, Ph. D. Thesis

Sembolini F., et al., 2016, *MNRAS*, 457, 4063

Sexton J., Lukić Z., Almgren A., Daley C., Friesen B., Myers A., Zhang W., 2021, *Journal of Open Source Software*, 6, 3068

Slosar A., et al., 2011, *Journal of Cosmology and Astroparticle Physics*, 2011, 001–001

Slosar A., et al., 2013, *Journal of Cosmology and Astroparticle Physics*, 2013, 026–026

Townsend R. H. D., 2009, *ApJS*, 181, 391

Tytler D., Paschos P., Kirkman D., Norman M. L., Jena T., 2009, *MNRAS*, 393, 723

Valluri M., et al., 2022, arXiv e-prints, p. arXiv:2203.07491

Vernet J., et al., 2011, *A&A*, 536, A105

Viel M., Lesgourgues J., Haehnelt M. G., Matarrese S., Riotto A., 2005, *Physical Review D*, 71, 063534

Viel M., Becker G. D., Bolton J. S., Haehnelt M. G., Rauch M., Sargent W. L. W., 2008, *Phys. Rev. Lett.*, 100, 041304

Viel M., Bolton J. S., Haehnelt M. G., 2009, *MNRAS*, 399, L39

Viel M., Becker G. D., Bolton J. S., Haehnelt M. G., 2013a, *Phys. Rev. D*, 88, 043502

Viel M., Schaye J., Booth C. M., 2013b, *MNRAS*, 429, 1734

Vogt S. S., et al., 1994, in Crawford D. L., Craine E. R., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 2198, Instrumentation in Astronomy VIII. p. 362, doi:10.1117/12.176725

Walther M., Hennawi J. F., Hiss H., Oñorbe J., Lee K.-G., Rorai A., O’Meara J., 2018, *The Astrophysical Journal*, 852, 22

Walther M., Oñorbe J., Hennawi J. F., Lukić Z., 2019, *ApJ*, 872, 13

Walther M., Armengaud E., Ravoux C., Palanque-Delabrouille N., Yèche C., Lukić Z., 2021, *J. Cosmology Astropart. Phys.*, 2021, 059

Yèche C., Palanque-Delabrouille N., Baur J., Bourboux H. d. M. d., 2017a, *Journal of Cosmology and Astroparticle Physics*, 2017, 047

Yèche C., Palanque-Delabrouille N., Baur J., Bourboux H. d. M. d., 2017b, *Journal of Cosmology and Astroparticle Physics*, 2017, 047–047

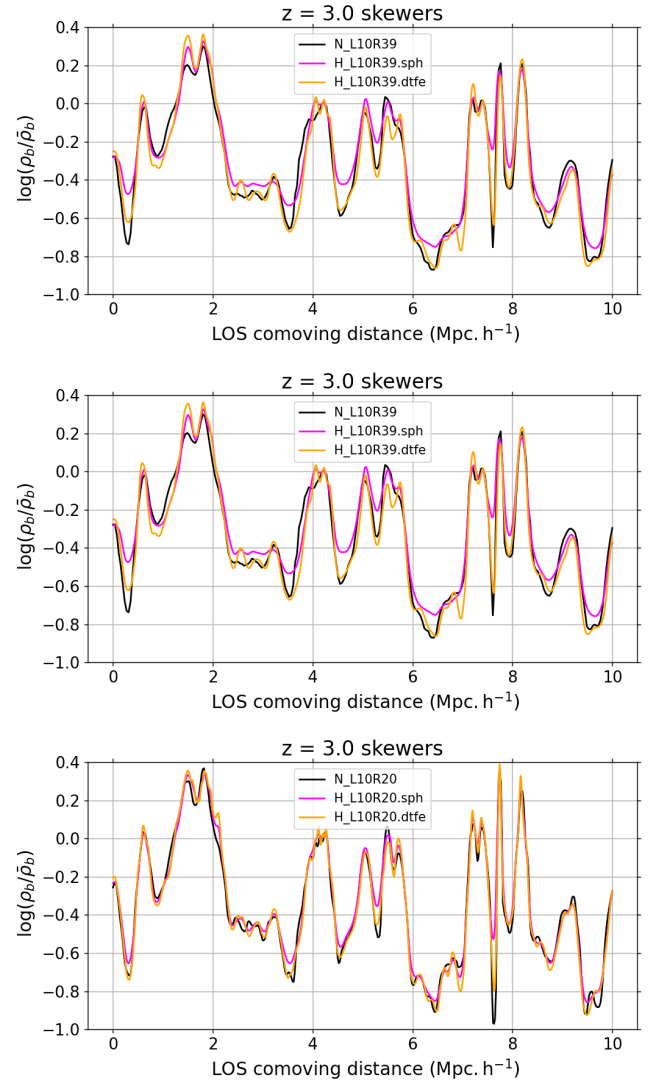
Zaldarriaga M., Hui L., Tegmark M., 2001, *ApJ*, 557, 519

Zel’dovich Y. B., 1970, *A&A*, 500, 13

Zhang W., et al., 2019, *Journal of Open Source Software*, 4, 1370

de Sainte Agathe V., et al., 2019, *Astronomy & Astrophysics*, 629, A85

du Mas des Bourboux H., et al., 2017, *Astronomy & Astrophysics*, 608, A130



**Figure A1.** Density field skewers from the snapshots in the  $10 h^{-1} \text{Mpc}$  box with  $128^3$  (top),  $256^3$  (middle) and  $512^3$  (bottom) cells/particles at  $z = 3$  for Nyx (black) and CRK-HACC using the SPH and DTFE methods to estimate the density fields (magenta and orange respectively). Results from the high-resolution skewer with  $1024^3$  cells/particles are depicted in Fig. 1, left panel.

## APPENDIX A: ADDITIONAL EXAMPLES AND CODE TESTS

### A1 Density fields in low-resolution simulations

At high resolution it is not easy to visually discern the differences between code outputs. To make these differences more obvious, Fig. A1 shows the same low-density skewer as in Fig. 1 to highlight the differences, i.e., the smoothing of small-scale fluctuations and the over-estimation of the density field in low-density regions using the SPH algorithm and the improvement due to the DTFE algorithm. Note that the DTFE density estimation-based results are much closer to the Nyx outputs at all resolutions (the SPH over-smoothing is very clear in the top two panels of the figure), but there remain differences in the details.

## A2 Impact of the number of neighbors on the 1D Ly $\alpha$ power spectra for CRK-HACC.sph

We study the impact of SPH density estimation on the Ly $\alpha$  1D power spectrum by varying the the number of neighbors,  $n_{\text{nb}}$ , for each SPH smoothing sphere. The results are shown in Fig. A2. (See Sec. 2.2 for a complete description of the parameter,  $n_{\text{nb}}$ .) Lowering the number of neighbors, which increases the spatial resolution, makes the CRK-HACC power spectra prediction come in to slightly better agreement with Nyx, 2% at  $z = 3$  and 10% at  $z = 5$ . We were unable to test whether or not reducing  $n_{\text{nb}}$  further improves the results since the increased sampling noise associated with fewer neighbors led to computational convergence issues in the high-order CRKSPH solver precluding the simulation from finishing. The results in Fig. A2 can be contrasted with those in Fig. 8, which show the much-improved agreement when using the DTFE algorithm with CRK-HACC.

We stress that the SPH neighbor count is an internal code parameter impacting both the simulation evolution as well as the density estimate used in the Ly $\alpha$  calculations. We have shown that lowering the default value of  $n_{\text{nb}}$  in CRK-HACC from 268 to 113 improves Ly $\alpha$  agreement with Nyx by reducing the amount of smoothing in the SPH density estimate, particularly in the sparsely sampled low-density regions where the two codes most strongly disagree. The value of  $n_{\text{nb}} = 113$  is still high compared to other common SPH approaches where  $n_{\text{nb}} \sim 32 - 64$  and it is possible that using such a smaller neighbor count for the density estimate would further improve agreement between CRK-HACC and Nyx. Of course, lowering the neighbor count also increases sampling noise meaning at some point the agreement may reverse. Hence, there is likely an optimal value of  $n_{\text{nb}}$  for the density estimate, and this value will depend on the SPH scheme. The important point, however, is not to identify an optimal value for  $n_{\text{nb}}$ , but rather to determine the numerically converged result when the total particle count is continually increased (at fixed neighbor count). Using a different neighbor count may change the rate of convergence, but ultimately the final converged results should agree. The main utility of the DTFE density estimate is that it was found to dramatically increase the rate of convergence of the CRK-HACC Ly $\alpha$  measurements by providing a better density estimate while also not requiring any notion of neighbor counts. For this reason, it is likely that other SPH codes would also benefit from the usage of the DTFE density estimate in the context of Ly $\alpha$  measurements.

## A3 Impact of low-density regions on the 1D Ly $\alpha$ power spectra

An investigation of the low-density behavior of the Ly $\alpha$  1D power spectrum from CRK-HACC.sph is shown in Fig. A3. As described in the main text (Section 6.3.1), we carried out a number of tests and confirmed that the difference between the Nyx and CRK-HACC.sph arose from the low density regions. To see what regions were responsible for the discrepancy, we simply replaced the CRK-HACC.sph density field with the Nyx density field on the evaluation grid, when the density fell below a certain threshold value, taken to be a tunable parameter. For this test, we use a mesh with  $N_{\text{grid}} = N_{\text{part}}$  for the CRK-HACC outputs to match the Nyx mesh. We find that CRK-HACC.sph comes into increasing agreement as the parameter  $\rho_{\text{thresh}}$  is gradually increased; thus highlighting that the main differences between the two codes originates from low-density regions. We find that replacing CRK-HACC.sph densities with Nyx densities for  $\log(\rho_{\text{thresh}}/\bar{\rho}_b) < -0.5$  (0.5) is sufficient to reach sub-percent agreement between the two codes at  $z = 5$  ( $z = 3$ ) over all scales. This sets the density scale to investigate the source of the possible discrepancy.

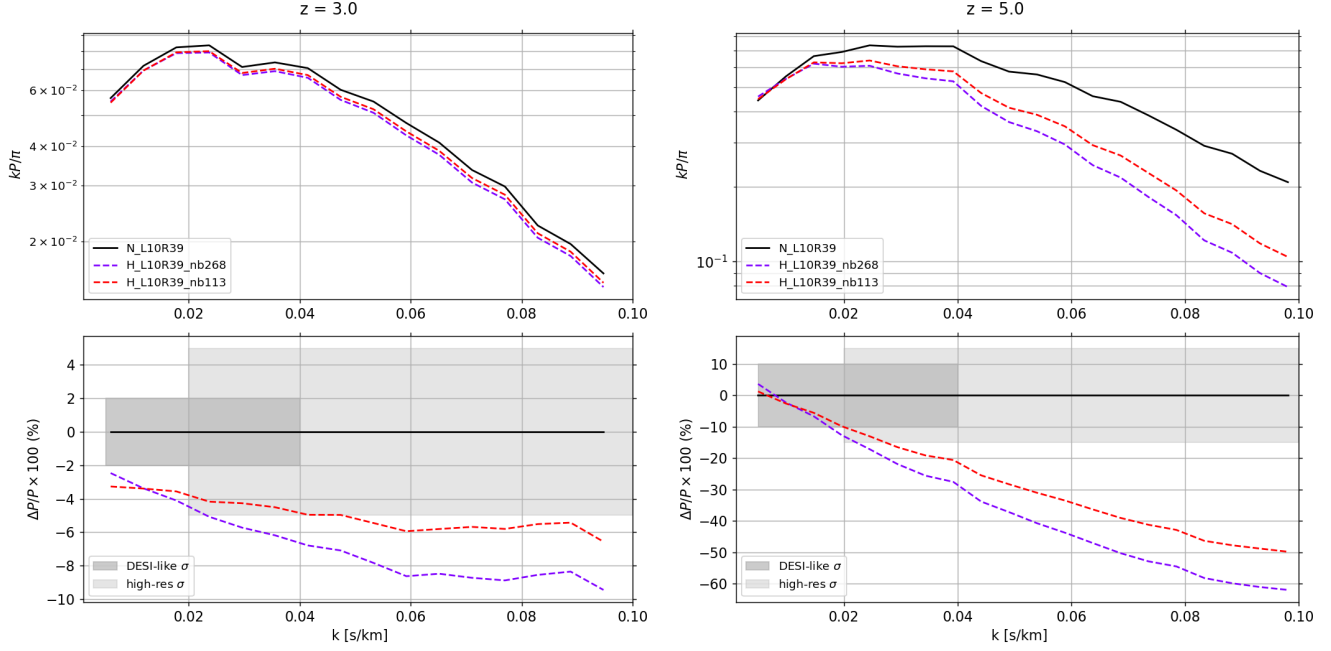
## A4 3D power spectrum at high redshifts

Fig. A4 shows the 3D Ly $\alpha$  power spectra at  $z = 5$  which can be contrasted with the  $z = 3$  results shown in Fig. 11 (Sec. 6.4). The tendency of reduced error as one moves to the right (towards the LOS result) is less marked in this case compared to Fig. 11, due to the reduced thermal broadening at high redshifts. As at  $z = 3$ , the power spectra from CRK-HACC.dtfc display a different sense of convergence as a function of  $k$  compared to the Nyx and CRK-HACC.sph results, with less power at large scales and higher power at small scales compared to the reference.

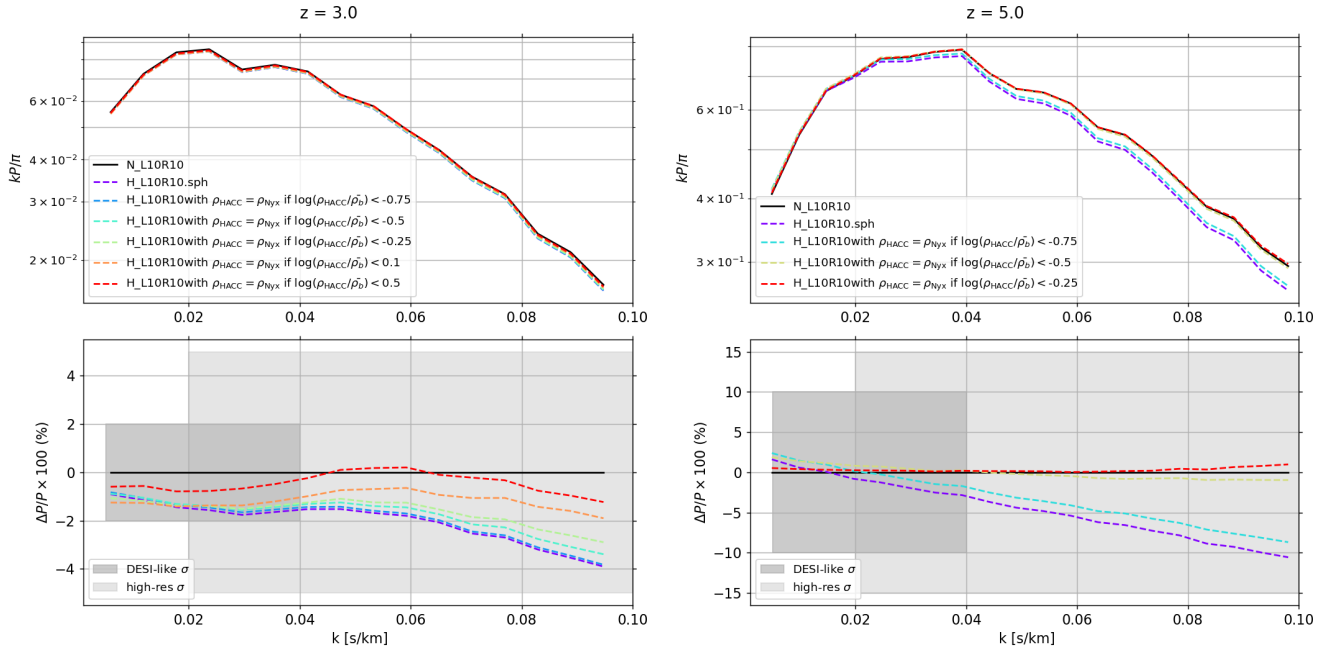
## A5 Large box results

Results from our investigation in box size effects are discussed here, adding to the presentation already provided in Sec. 7. Fig. A5 and Fig. A6 show the impact of box size on the 1D and 3D Ly $\alpha$  power spectra respectively at  $z = 3$ . The main purpose of showing these results is to demonstrate the consistent offset between the power spectra for both the Nyx and CRK-HACC.sph runs, which is due to the missing power in the smaller box, as mentioned already in Sec. 7. Consequently, there is no additional relative correction needed for box size compensation. To complete the discussion, we also include data on the affect of spatial resolution in the larger  $40 \text{ h}^{-1} \text{ Mpc}$  box: Fig. A7 shows the impact of varying resolution on the 1D Ly $\alpha$  power spectra at  $z = 3$ . The smooth behavior is consistent with the similar investigation carried out in Sec. 6.3.3 for the  $10 \text{ h}^{-1} \text{ Mpc}$  boxes.

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.

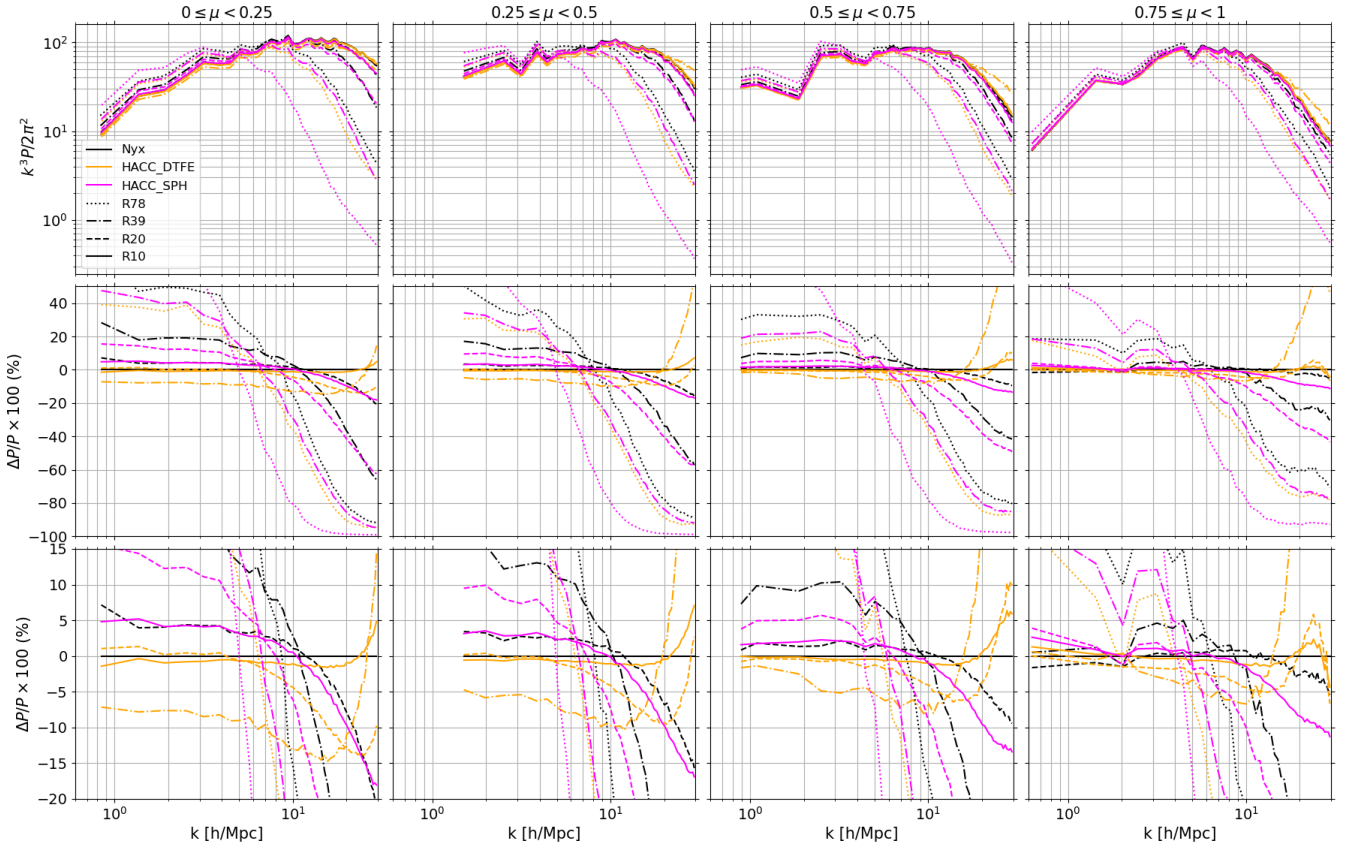


**Figure A2.** 1D Ly $\alpha$  power spectrum for Nyx (solid) and CRK-HACC.sph (dashed) at  $z = 3$  (left) and  $z = 5$  (right) in the  $10 h^{-1}$ Mpc boxes. For CRK-HACC.sph, results are shown for different number of neighbors,  $n_{\text{nb}}$ , contained in each smoothing sphere, with  $n_{\text{nb}} = 268$  and  $n_{\text{nb}} = 113$  in purple and red respectively. Middle and bottom panels show relative differences with respect to N\_L10R10, the Nyx most-resolved simulation with  $10 h^{-1}$ kpc resolution. The shaded regions show  $k$ -averaged observational uncertainties of DESI-like (darker gray) and high-resolution surveys respectively. These results can be contrasted with those in Fig. 8, which show the much-improved agreement when using the DTFE algorithm with CRK-HACC.

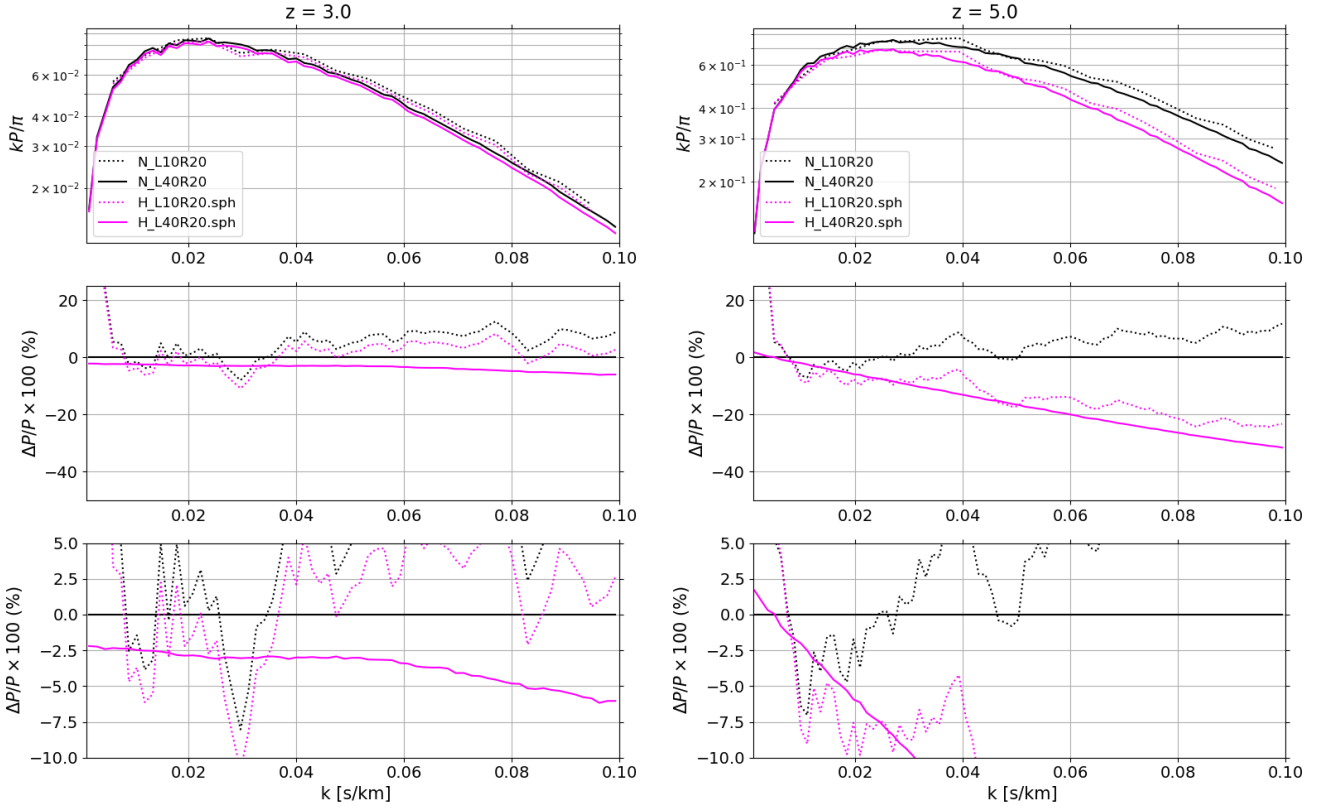


**Figure A3.** 1D Ly $\alpha$  power spectrum for Nyx (solid) and CRK-HACC.sph (dashed) at  $z = 3$  (left) and  $z = 5$  (right) in the  $10 h^{-1}$ Mpc boxes. For CRK-HACC.sph, we use the Nyx density,  $\rho_{\text{Nyx}}$  instead of the CRK-HACC density,  $\rho_{\text{HACC}}$ , when the latter is above  $\rho_{\text{thresh}}$  (in mean density units). Middle and bottom panels show relative differences with respect to N\_L10R10, the most-resolved Nyx simulation with  $10 h^{-1}$ kpc resolution. The shaded regions show  $k$ -averaged observational uncertainties of DESI-like (darker gray) and high-resolution (lighter gray) surveys respectively.

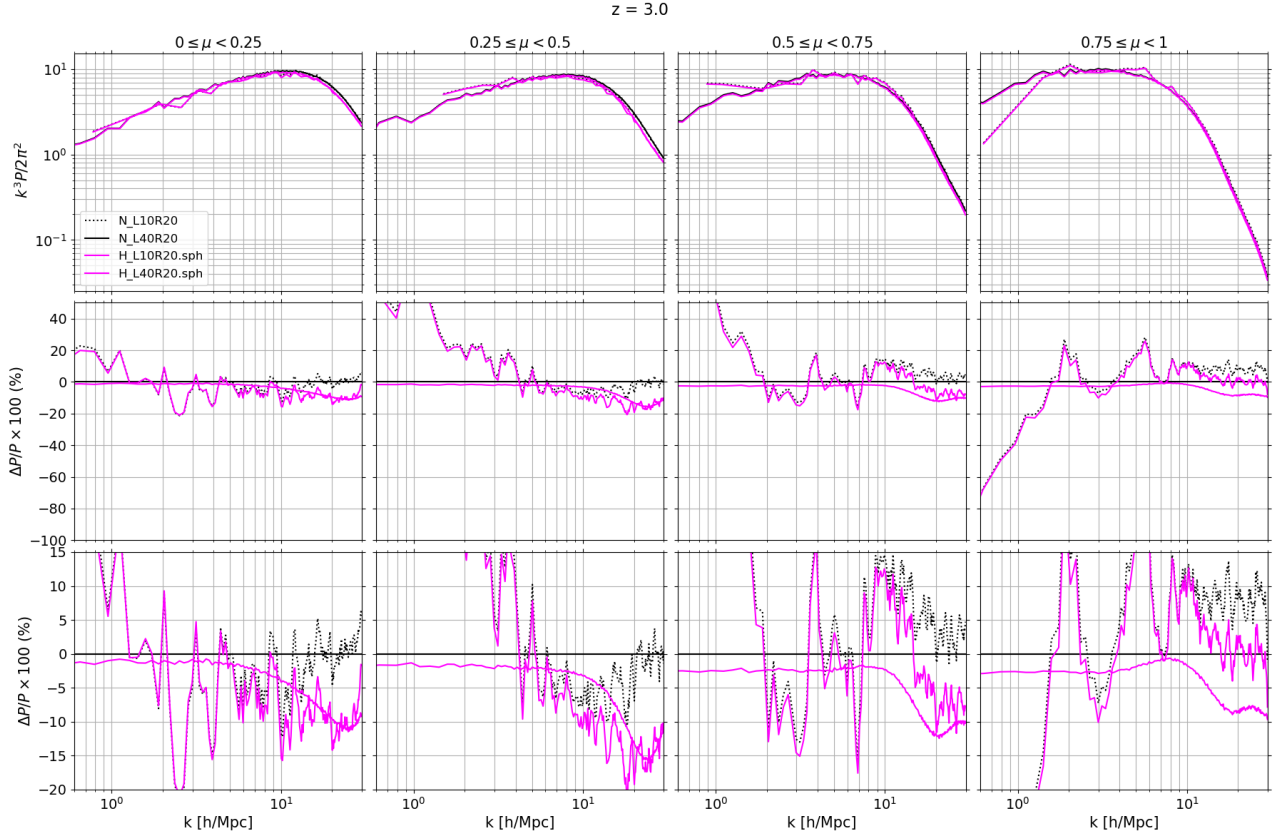
$z = 5.0$



**Figure A4.** 3D Ly $\alpha$  power spectra at  $z = 5$  for Nyx (black), CRK-HACC.sph (magenta) and CRK-HACC.dtfе (orange) in  $10 h^{-1}$ Mpc boxes at different spatial resolutions. Middle and bottom panels show relative differences with respect to N\_L10R10. The bottom panel shows a more restricted range on the y-axis to make it easier to distinguish the different results.

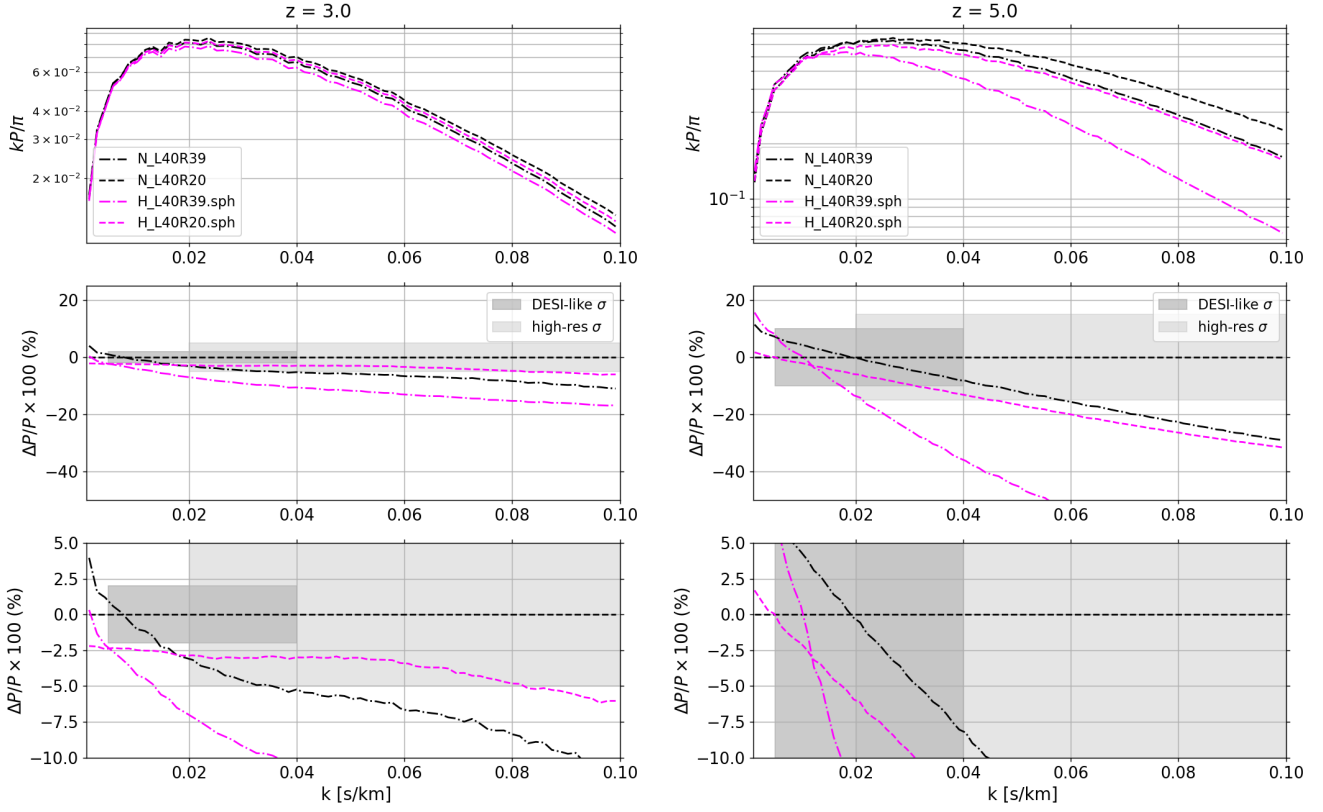


**Figure A5.** 1D Ly $\alpha$  power spectra for Nyx (black) and CRK-HACC.sph (magenta) at  $z = 3$  (left) and  $z = 5$  (right) in the  $40 h^{-1}\text{Mpc}$  (solid) and  $10 h^{-1}\text{Mpc}$  (dotted) boxes with an average resolution of  $20 h^{-1}\text{kpc}$ . Middle and bottom panels show relative differences with respect to N\_L40R20. The bottom panel shows a more restricted range on the y-axis.



**Figure A6.** 3D Ly $\alpha$  power spectrum for Nyx (black) and CRK-HACC.sph (magenta) at  $z = 3$  in the  $40 \text{ h}^{-1}\text{Mpc}$  (solid) and  $10 \text{ h}^{-1}\text{Mpc}$  (dotted) boxes with an averaged resolution of  $20 \text{ h}^{-1}\text{kpc}$ . Middle and bottom panels show relative differences with respect to N\_L40R20. The bottom panel shows a more restricted range on the y-axis.





**Figure A7.** 1D Ly $\alpha$  power spectra for Nyx and CRK-HACC.sph at  $z = 3$  (left) and  $z = 5$  (right) in the  $40 h^{-1} \text{Mpc}$  boxes at different spatial resolutions. Middle and bottom panels show relative differences with respect to N\_L40R10. Shaded regions show  $k$ -averaged observational uncertainties of DESI-like (darker gray) and high-resolution surveys (silver).