01 Aug 2019

Radiative Transfer Distortions of Lyman **A** Emitters: A New Fingers-Of-God Damping in the Clustering in Redshift Space

Chris Byrohl

Shun Saito
Missouri University of Science and Technology, saitos@mst.edu

Christoph Behrens

Recommended Citation

C. Byrohl et al., "Radiative Transfer Distortions of Lyman A Emitters: A New Fingers-Of-God Damping in the Clustering in Redshift Space," *Monthly Notices of the Royal Astronomical Society*, vol. 489, no. 3, pp. 3472-3491, Oxford University Press, Aug 2019.
The definitive version is available at https://doi.org/10.1093/mnras/stz2260

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **489,** 3472–3491 (2019)
Advance Access publication 2019 August 16

doi:10.1093/mnras/stz2260

Radiative transfer distortions of Lyman α emitters: a new Fingers-of-God damping in the clustering in redshift space

Chris Byrohl [1]★ Shun Saito [1,2,3] and Christoph Behrens [4]

[1] *Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Straße 1, D-85740 Garching bei München, Germany*
[2] *Department of Physics, Missouri University of Science and Technology, 1315 N. Pine Street, Rolla, MO 65409, USA*
[3] *Kavli Institute for the Physics and Mathematics of the Universe (WPI), Todai Institutes for Advanced Study, The University of Tokyo, 5-1-5 Kashiwa-no-Ha, Kashiwa, Chiba 277-8582, Japan*
[4] *Institut für Astrophysik, Georg-August Universität Göttingen, Friedrich-Hund-Platz 1, D-37075 Göttingen, Germany*

Accepted 2019 July 28. Received 2019 July 26; in original form 2019 June 17

ABSTRACT

Complex radiative transfer (RT) of the Lyman α photons poses a theoretical challenge to galaxy surveys that infer the large-scale structure with Lyman α emitters (LAEs). Guided by RT simulations, prior studies investigated the impact of RT on the large-scale LAE clustering, and claimed that RT induces a selection effect which results in an anisotropic distortion even in real space but in an otherwise negligible effect in redshift space. However, our previous study, which relies on a full RT code run on the Illustris simulations, shows that the anisotropic selection effect was drastically reduced with higher spatial resolution. Adopting the same simulation framework, we further study the impact of RT on the LAE clustering in redshift space. Since we measure LAE's radial position through a spectral peak of Lyman α emission, the frequency shift due to RT contaminates the redshift measurement and hence the inferred radial position in redshift space. We demonstrate that this additional RT offset suppresses the LAE clustering along the line of sight, which can be interpreted as a novel Fingers-of-God (FoG) effect. To assess the FoG effect, we develop a theoretical framework modelling the impact of the RT similar to that of the small-scale peculiar velocity which is commonly studied in the context of the redshift space distortion (RSD). Although our findings strongly encourage a more careful RSD modelling in LAE surveys, we also seek a method to mitigate the additional FoG effect due to RT by making use of other information in a Lyman α spectrum.

Key words: radiative transfer – galaxies: high-redshift – large-scale structure of Universe.

1 INTRODUCTION

Current and future redshift surveys detecting galaxies with prominent Lyman α emission, so-called Lyman α emitters (LAEs), can give competitive astrophysical and cosmological constraints. For example, the Hobby-Eberly Telescope Dark Energy Experiment (hereafter HETDEX; Hill et al. 2008; Adams et al. 2011) is currently operating and will eventually detect close to a million of LAEs in a redshift range of $1.9 \leq z \leq 3.5$ over a sky patch of 450 deg^2 (Leung et al. 2017). HETDEX will measure cosmic expansion history and growth history of matter fluctuations through baryon acoustic oscillations (BAOs) and redshift space distortion (RSD), and also shed more light on the properties of star-forming galaxies by offering a range of statistical measures.

However, HETDEX and other future Lyman α redshift surveys might have to deal with severe modifications of the detected spatial clustering signal due to the complex radiative transfer (RT) of Lyman α given its resonant nature and high optical depth in astrophysical environments. Often Lyman α photons, particularly those produced within the star-forming regions of galaxies, scatter many times in the interstellar medium (ISM) and the circumgalactic medium (CGM) before reaching the observer. Additionally, column depths of neutral hydrogen outside the galaxy's host halo might be sufficient to substantially attenuate the remaining flux in the intergalactic medium (IGM) by scattering photons out of the line of sight (LOS).

Scatterings of neutral hydrogen change both frequency and position of photons before escaping towards the observer. This can introduce new distortion effects in the cosmological signal both in real space and in redshift space. In real space this corresponds to a selection effect favouring certain LAEs to be detected over others based on their environment, introducing both isotropic and

★ E-mail: cbyrohl@mpa-garching.mpg.de

anisotropic modifications to the two-point statistics. An isotropic distortion effectively corresponds to changing the bias due to emitters of different mass range being favourably detected. Similarly the detection of emitters might be affected by their large-scale environment, which can also give rise to anisotropic distortions as demonstrated by Zheng et al. (2011). While these real space distortions automatically propagate into the redshift space signal, there can be additional distortions purely arising in redshift space when the RT modifies the spectral features from which the line-of-sight position is inferred. Similarly to the selection effects, this can cause both isotropic and anisotropic distortions in the two-point statistics.

The complexity of the RT limits analytic solutions to symmetric toy models, while in physical environments explicit RT simulations are needed, which are usually run as a post-processing step given the expensive numerical cost. Prior simulations of LAEs in their large-scale structure environment examining possible distortion effects include Zheng et al. (2010)/Zheng et al. (2011), Behrens & Niemeyer (2013), and Behrens et al. (2018) running an RT code on top of hydrodynamic simulations tracing neutral hydrogen distribution down to scales of ∼kpc (for latter publication), but no further subgrid model is assumed below those scales. More studies of LAEs in large cosmological volumes such as Inoue et al. (2018) and Gurung-López et al. (2019b) exist that however use a different approach: A dark matter only *N*-body simulation is used to map the large-scale structure and identify haloes that possibly contain LAEs. Different sophisticated semi-analytical models are used to approximate the spectra arising on the ISM/CGM scales that remain unresolved in the simulations themselves. After calibration of such semi-analytical model a good match with observations can be obtained.

Zheng et al. (2010)/Zheng et al. (2011) find a strong anisotropic selection effect occurring in the real space clustering signal caused by a correlation of the observed flux with the large-scale velocity gradient. This selection effect has been challenged by Behrens & Niemeyer (2013) and Behrens et al. (2018), who could not reproduce such effect. Behrens et al. (2018) show that prior findings might have been strongly overestimated due to a lack of spatial resolution and a simplified emitter model.

All these studies are primarily concerned with real space distortions, either not evaluating (Behrens & Niemeyer 2013; Behrens et al. 2017) or not finding (Zheng et al. 2010, 2011) additional effects in redshift space. In this work, we revisit the idea of possible, additional RSDs of LAEs in redshift space. For the analysis, we reuse the RT simulations run by Behrens et al. (2018), covering a redshift range from $z = 2.0$ to $z = 5.85$ as described later.

The structure of this paper is as follows. In Section 2, we describe the theoretical background needed to model the newly found distortion described later in the paper. Afterwards, in Section 3, we introduce the RT simulations performed by Behrens et al. (2018) used here and how to reduce them to mock catalogues of LAEs. In Section 4, we present results of detected spectra and inferred positions before showing how these affect the two-point statistics. In Section 5, we summarize our findings along with a discussion of possible shortcomings in our findings and how future surveys might be affected by/corrected for RT RSDs, before concluding in Section 6.

2 THEORETICAL BACKGROUND OF RSD

In this section we provide a brief review on the galaxy clustering in redshift space. In real space which we considered in our previous work (Behrens et al. 2017), the fluctuation in number density of LAEs for a given sample is given by

$$1 + \delta_{\rm g}(\boldsymbol{x}) = \frac{n_{\rm g}(\boldsymbol{x})}{\bar{n_{\rm g}}}. \tag{1}$$

Then we consider the two-point statistics to characterize strength of the clustering signal, i.e. the correlation function, $\xi_{\rm g}$, or the power spectrum, $P_{\rm g}$, as its Fourier-counterpart mapped by the Fourier transform (FT):[1]

$$1 + \xi_{\rm g}(\boldsymbol{r}) = \left\langle \left\{1 + \delta_{\rm g}(\boldsymbol{x})\right\} \left\{1 + \delta_{\rm g}(\boldsymbol{x} + \boldsymbol{r})\right\} \right\rangle, \tag{4}$$

$$(2\pi)^3 P_{\rm g}(\boldsymbol{k}) \delta_{\rm D}(\boldsymbol{k} + \boldsymbol{k}') = \langle \delta_{\rm g}(\boldsymbol{k}) \delta_{\rm g}(\boldsymbol{k}') \rangle, \tag{5}$$

where statistical homogeneity is implicitly assumed. In real space where statistical isotropy is given if no selection effect is present, the arguments of the two-point statistics depend only on the scale, i.e. $\xi_{\rm g}(r)$ and $P_{\rm g}(k)$.

Now let us consider a contamination of *any form* of velocity, $\vec{v}$, along an LOS of a galaxy in measuring its redshift and hence its radial position. Mathematically this is equivalent to a mapping from real to redshift space:

$$\boldsymbol{s} = \boldsymbol{r} + \frac{\boldsymbol{v} \cdot \hat{\boldsymbol{r}}}{aH(a)} \hat{\boldsymbol{r}}, \tag{6}$$

where $\hat{\boldsymbol{r}}$ is a unit vector along the LOS direction, a is the scale factor of the Universe, and $H(a)$ is the Hubble expansion rate. It has been extensively discussed in the literature that the peculiar velocity of a galaxy contaminates its redshift space position, making the clustering pattern anisotropic known as RSD (see e.g. Hamilton 1998, for a review). This occurs simply because this is the effect only along the LOS, which breaks statistical isotropy.

The complexity to accurately model the two-point statistics in redshift space originates from the fact that the Jacobian of the mapping in equation (6) is *non-linear* in terms of the velocity, v. In the following, let us discuss the impact of the non-linear mapping on the non-linear power spectrum and the correlation function as generally as possible. For this purpose, we begin with the following expression of the redshift-space density field (Taruya, Nishimichi & Saito 2010) which is *exact* under the global plain-parallel approximation (i.e. the LOS is fixed with one global direction as $\hat{\boldsymbol{r}} \approx \hat{\boldsymbol{z}}$, see Beutler et al. 2014):

$$\delta_{\rm g}^s(\boldsymbol{k}) = \int {\rm d}^3x \left\{\delta_{\rm g}(\boldsymbol{x}) - f\partial_z u_z(\boldsymbol{x})\right\} {\rm e}^{{\rm i}\boldsymbol{k}\cdot\boldsymbol{x} + {\rm i}fk_z u_z(\boldsymbol{x})}, \tag{7}$$

where we introduce a scaled velocity, $\vec{u} \equiv \vec{v}/(faH)$, and $f \equiv {\rm d}\ln D/{\rm d}\ln a$ is the linear growth function. We specifically denote a quantity in redshift space with a superscript 's' throughout this paper. We then find an expression for the redshift-space power spectrum as

$$P_{\rm g}^s(\boldsymbol{k}) = \int {\rm d}^3r\, {\rm e}^{{\rm i}\boldsymbol{k}\cdot\boldsymbol{r}} \left\langle {\rm e}^{-{\rm i}fk_z\Delta u_z} \times \left\{\delta_{\rm g}(\boldsymbol{x}) + f\partial_z u_z(\boldsymbol{x})\right\} \left\{\delta_{\rm g}(\boldsymbol{x}') + f\partial_z u_z(\boldsymbol{x}')\right\} \right\rangle, \tag{8}$$

where $\boldsymbol{r} \equiv \boldsymbol{x} - \boldsymbol{x}'$ and $\Delta\boldsymbol{u} \equiv \boldsymbol{u}(\boldsymbol{x}) - \boldsymbol{u}(\boldsymbol{x}')$. Equation (8) apparently involves higher order correlations between the density $\delta_{\rm g}$ and the

[1] We adopt the following convention for the FT:

$$A(\boldsymbol{k}) = \int {\rm d}^3x\, {\rm e}^{{\rm i}\boldsymbol{k}\cdot\boldsymbol{x}} A(\boldsymbol{x}), \tag{2}$$

$$A(\boldsymbol{x}) = \int \frac{{\rm d}^3k}{(2\pi)^3}\, {\rm e}^{-{\rm i}\boldsymbol{k}\cdot\boldsymbol{x}} A(\boldsymbol{k}). \tag{3}$$

velocity field u_z. Linearizing equation (8) in terms of δ_g and u_z yields

$$P_g^{s,L}(\boldsymbol{k}) = (1 + f\mu^2)^2 P_g(k), \tag{9}$$

where μ is the cosine of an angle between $\vec{k}$ and the LOS, defined as $\mu \equiv k_z/k$. This equation, well known as the Kaiser formula (Kaiser 1987), shows that the clustering in redshift space is more enhanced closer to the LOS direction, which is a valid picture on large scales and nothing but the main target of RSD measurements (see e.g. White, Song & Percival 2009). On the other hand, equations (7) and (8) imply that even random velocity suppresses the redshift-space clustering on small scales along the LOS, often quoted as the Finger-of-God (FoG) effect (Jackson 1972). To see this more explicitly, let us rewrite equation (8) in terms of the cumulants as (Scoccimarro 2004; Taruya et al. 2010)

$$\begin{aligned} P_g^s(\boldsymbol{k}) = \int \mathrm{d}^3 r\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \exp\left\{\left\langle \mathrm{e}^{-\mathrm{i}fk_z\Delta u_z}\right\rangle_c\right\} \\ \times \left\{\left\langle \mathrm{e}^{-\mathrm{i}fk_z\Delta u_z}\mathcal{A}(\boldsymbol{x})\mathcal{A}(\boldsymbol{x}')\right\rangle_c \right. \\ \left. + \left\langle \mathrm{e}^{-\mathrm{i}fk_z\Delta u_z}\mathcal{A}(\boldsymbol{x})\right\rangle_c \left\langle \mathrm{e}^{-\mathrm{i}fk_z\Delta u_z}\mathcal{A}(\boldsymbol{x}')\right\rangle_c\right\}, \end{aligned} \tag{10}$$

where $\mathcal{A}(\vec{x}) \equiv \delta_g(\vec{x}) + f\partial_z u_z(\vec{x})$ is used just to simplify the notation. As Zheng & Song (2016) pointed out, the overall exponential factor, $\exp\left\{\left\langle \mathrm{e}^{-\mathrm{i}fk_z\Delta u_z}\right\rangle_c\right\}$, contains terms which depend only on the one-point cumulants. These terms survive even when two-point correlations such as $\langle \Delta u_z \mathcal{A}\rangle$ are zero, and can be integrated out because it no longer depends on the scale. That is to say,

$$\begin{aligned} \exp\left\{\left\langle \mathrm{e}^{-\mathrm{i}fk_z\Delta u_z}\right\rangle_c\right\} &= \exp\left\{\sum_{n=1}^{\infty}(-\mathrm{i}fk_z)^n \frac{\left\langle \Delta u_z^n\right\rangle_c}{n!}\right\} \\ &= \exp\left\{\sum_{m=1}^{\infty}(-\mathrm{i}fk_z)^{2m} \frac{2\left\langle u_z(\boldsymbol{x})^{2m}\right\rangle_c}{(2m)!}\right\} \\ &\quad \times \exp\left\{\sum_{m=1}^{\infty}(-\mathrm{i}fk_z)^{2m} \frac{\left\langle \Delta u_z^{2m}\right\rangle_c - \left\langle u_z(\boldsymbol{x})^{2m}\right\rangle_c - \left\langle u_z(\boldsymbol{x}')^{2m}\right\rangle_c}{(2m)!}\right\}. \end{aligned} \tag{11}$$

We have used the fact that the terms with odd power in the second and third lines vanish because of symmetry in a galaxy pair. The exponential factor in the second line of equation (11) does not depend on the scale. For example, if u_z follows a Gaussian distribution with zero mean and variance of σ_u^2, only $m = 1$ term survives in the exponential factor, corresponding to the FoG damping factor commonly assumed:

$$D_{\mathrm{FoG}}^{\mathrm{Gaussian}}(k,\mu) = \mathrm{e}^{-f^2k^2\mu^2\sigma_u^2}. \tag{12}$$

Notice that, since the FoG damping factor depends only on the one-point cumulants, it can be derived at the level of the density field, equation (7). Namely, if the velocity field follows a probability distribution function (PDF), $P(u_z)$, we have

$$D_{\mathrm{FoG}}(k,\mu) = \left|\left\langle \mathrm{e}^{\mathrm{i}fk\mu u_z}\right\rangle\right|^2 = \left|\int \mathrm{d}u_z\, P(u_z)\mathrm{e}^{\mathrm{i}fk\mu u_z}\right|^2 \tag{13}$$

that is the FT of the one-point PDF, $P(u_z)$ (e.g. Hikage & Yamamoto 2016). Another common velocity PDF is an exponential distribution (e.g. Scoccimarro 2004). The FT of such distribution of the $\exp(-\sqrt{2}\,|u_z|\,/\sigma_u)/(\sqrt{2}\sigma_u)$, is a Lorentzian damping function[2]

$$D_{\mathrm{FoG}}^{\mathrm{Lorentzian}}(k,\mu) = \left\{\frac{1}{1 + f^2k^2\mu^2\sigma_u^2}\right\}^2. \tag{14}$$

We note again that the PDFs $P(u_z)$ are assumed for the one-point (rather than pairwise) velocity distributions. This results in a different interpretation of the σ for the Gaussian and the exponential distribution in the literature (e.g. Scoccimarro 2004).

Here we stress that the FoG damping inevitably arises as long as the velocity field has a non-zero dispersion and higher order moments. For instance, Agrawal et al. (2017) confirmed the damping due to the non-linear mapping even assuming a linear velocity field. In addition, the two-point correlations between the density and velocity fields (i.e. the second and third lines in equation 10) are essential to accurately model the non-linear redshift-space power spectrum as several authors have shown (see e.g. Taruya et al. 2010; Okumura, Seljak & Desjacques 2012; Matsubara 2014; Zheng, Song & Oh 2019; Vlah & White 2019).

So far we have shown that, in Fourier space, the FoG damping factor depending only on one-point PDF can be expressed as an *overall multiplicative* factor. In the following, let us instead discuss the configuration space as a complementary approach. The two-point correlation function (TPCF) in redshift space is generally written as (e.g. Scoccimarro 2004)

$$1 + \xi_g^s(\boldsymbol{s}) = \int \mathrm{d}\pi\ \left\{1 + \xi_g(r)\right\}\mathcal{P}(u_z;\boldsymbol{r}), \tag{15}$$

where $\pi \equiv s_z - u_z$ is the vector along the LOS direction, $\hat{z}$ in configuration space. $\mathcal{P}(u_z;\vec{r})$ is the *pairwise* velocity PDF given by the FT of the pairwise velocity generating function $\mathcal{M}(if\gamma,\vec{r})$:

$$\mathcal{P}(u_z;\boldsymbol{r}) = \int \frac{\mathrm{d}\gamma}{2\pi}\mathrm{e}^{-\mathrm{i}\gamma u_z}\mathcal{M}\left(if\gamma;\boldsymbol{r}\right), \tag{16}$$

$$\mathcal{M}(if\gamma;\boldsymbol{r}) = \frac{\left\langle \exp\left(if\gamma\Delta u_z\right)\left[1 + \delta_g(\boldsymbol{x})\right]\left[1 + \delta_g(\boldsymbol{x}')\right]\right\rangle}{1 + \xi(r)}. \tag{17}$$

An advantage of equation (15) is that the redshift-space correlation function can be expressed only in terms of quantities in real space. However, the complexity arises because of a convolution with the pairwise velocity PDF which is weighted by density fields at two points and hence scale-dependent. At linear level, the mean of the pairwise velocity PDF is related to coherent infall motion and hence the Kaiser factor, while its dispersion is related to the velocity power spectrum (Fisher 1995). Similarly to Fourier space at non-linear level, however, one has to take into consideration the correlation between the density and velocity fields as well as the contribution from the one-point PDF, as several authors have recently studied (see e.g. Scoccimarro 2004; Reid & White 2011; Uhlemann, Kopp & Haugg 2015; Bianchi, Percival & Bel 2016).

As we will explain in detail in the next section, we investigate another velocity offset due to the RT effect in addition to the peculiar velocity of a galaxy. We will study the impact of the RT velocity component on the redshift-space clustering, mainly focusing on the FoG damping factor in Fourier space, $D_{\mathrm{FoG}}(k,\mu)$, and the pairwise velocity PDF.

[2]We note that the exponential PDF is often adopted for the *pairwise* velocity PDF (see equation 15) rather than for the velocity PDF (Davis & Peebles 1983; Ballinger, Peacock & Heavens 1996). In this case, there is no square factor in the damping factor in equation (14). We avoid this choice, because the pairwise velocity PDF is generally scale-dependent and hence its mean and dispersion are not necessarily constants.

Table 1. In this table, we summarize key quantities of the redshift snapshots considered relevant to our analysis: Spatial resolution Δ, number of LAEs considered $N_{\rm LAE}$, and median radius $r_{\rm crit,200}$ encompassing 200 times the critical density of the Universe for those emitters for the post-processed snapshots. Beside the radius' physical size, we also state the angular size as seen for an observer. For each redshift, we also state the average neutral fraction, $f_{\rm IGM}$, at a characteristic hydrogen number density of $10^{-4}\,{\rm cm}^{-3}$. We also quote the conversion factor $(aH)^{-1}$ from peculiar velocity to comoving distance at each redshift.

Redshift	2.00	3.01	4.01	5.85
$f_{\rm IGM}$ (10^{-5})	2	3.7	6.8	35
Δ (pkpc)	1.2	0.8	0.7	0.5
$N_{\rm LAE}$	45 594	45 434	39 782	23 114
$r_{\rm crit,200}$ (pkpc)	40.6	30.4	24.3	17.8
$(aH)^{-1}\left(\frac{\rm Mpc\,h^{-1}}{\rm km\,s^{-1}}\right)$	0.0105	0.0094	0.0085	0.0073

3 METHODS

3.1 Radiative transfer simulations

We utilize our previous RT simulations in Behrens et al. (2018) with slight modifications as explained below. We run a Monte Carlo RT code of Ly α photons on top of the Illustris simulations (Vogelsberger et al. 2014; Nelson et al. 2015) as a post process at redshift outputs of 2.00, 3.01, 4.01, and 5.85 with its box size of $L_{\rm box} = 75\,{\rm cMpc\,h^{-1}}$. The Illustris simulations provide a distribution of galaxies and their neutral hydrogen content in a context of the large-scale structure, necessary for studying the impact of Lyman α RT on to statistics used in cosmology. Before applying the RT, we convert the Voronoi tessellation in Illustris on to an octet-tree data structure with a maximal resolution of $\Delta = 3.3\,{\rm ckpc}$. The refinement criterion is triggered for cells containing 32 or more Voronoi cells' defining positions. More details on the processing of the Illustris datasets can be found in Behrens et al. (2018).

We then explicitly place Ly α photons in a centre of the dark matter haloes and weight them by the haloes' respective Ly α luminosity that is based on star formation rate (SFR) of each halo:

$$L_{\rm int} = \frac{\rm SFR}{\rm M_\odot\,yr^{-1}} \times 10^{42}\,{\rm erg\,s^{-1}}. \quad (18)$$

With above relation we only consider Lyman α radiation after recombination locally sourced by ionizing radiation in the star-forming regions. The proportionality constant can vary by a factor of a few and depends on made assumptions such as the initial mass function, metallicity and fraction of recombinations radiating Lyman α photons (e.g. Furlanetto et al. 2005). For this work, the value of the proportionality constant is irrelevant as we detect emitters at a fixed number density threshold. We use the SFR given by the Illustris simulation, which is given as its instantaneous value at given redshift and computed close to the multiphase description in Springel & Hernquist (2003).

Note that we assume properties of the LAEs such as positions, velocities, and SFR only from host halo catalogues (i.e. 'group' catalogues) and hence ignore satellite galaxies. We impose a minimum threshold of 0.1 $\rm M_\odot\,yr^{-1}$ on the SFR and $10^{10}\,\rm M_\odot$ on the halo mass to limit ourselves to well resolved haloes and to limit the required computational resources. We summarize characteristics of our simulated LAEs in Table 1.

Additionally, we set the initial frequency profile emerging from the unresolved ISM to be a Gaussian whose width σ_i is set by the virial temperature of the halo, see Behrens et al. (2018). While the initial profile should have a significant impact on the observed properties, we lack a profound ISM sub-grid model for the scope of this paper. For the fiducial sample of LAEs at a number density of $n_{\rm LAE} = 10^{-2}\,{\rm Mpc^{-3}\,h^3}$ in our survey, we find a mean input width $\sigma_i = 137\,{\rm km^{-1}}$ at $z = 3.01$. This is roughly consistent with recent findings in shell models by Gronke (2017) finding $\sigma_i = 172^{+75}_{-60}$ km $\rm s^{-1}$ fitted to an LAE sample at median redshift of $z = 3.83$.

Varying the Gaussian widths for $T < T_{\rm vir}$, we found that this only has an insignificant impact on the spectra emerging after reprocessing on CGM scales. As cosmological simulations such as Illustris are unable to resolve the ISM regions, we explicitly cut out the unresolved ISM as defined by a hydrogen number density threshold of $0.13\,{\rm cm^{-3}}$ for the gas. Also, we ignore the impact of dust attenuation on the RT for simplicity. Given these simplifications, we do not expect our simulations to agree well with the observed luminosity function as already discussed in Behrens et al. (2018).

After the luminosity-weighted photons are spawned with an isotropic angular distribution in the LAE's rest frame, the photons are propagated in a straight line until a scattering with a neutral hydrogen atom occurs. Then at each scattering point, the attenuated contribution along the line of sight towards the observer is computed ('peeling-off' photon) while the original photon is re-emitted and propagated/scattered subsequently.

In comparison to the simulations presented in Behrens et al. (2018), the only modification in our RT simulation stems from an increased initial Monte Carlo photon count: we increase this count from 100 to 1000 in order to properly sample the spectra as a function of wavelength. The requirement for the photon count was less important before as only the total flux was relevant for the analysis in Behrens et al. (2018).

As a result of those RT simulations, catalogues of attenuated photon contributions reaching the observer are created, including information such as the observed intensity as a function of wavelength and position perpendicular to the LOS and the positions of photons' originating LAE. In Fig. 1, we visualize the reprocessed Lyman α photons escaping the simulation box and project the surface brightness on to the cube's faces. The top and the right face of the cube contain the line-of-sight direction and are shown in redshift space. One can easily notice a strong anisotropy in redshift space, which will be the focus of this study. The position of the individually observed LAE along the LOS however depends upon a detection algorithm whose methodology we introduce in the next subsection.

3.2 Analysis of simulated LAEs in redshift space

3.2.1 LAE spectra and redshift space positions

To determine the position of LAEs in redshift space, we calculate the flux and spectral information by applying a spherical aperture of 3 arcsec radius (our default case) around a known LAE's position (from the halo catalogues), which also already fixes the angular position of the selected LAE. The aperture size is chosen to correspond roughly to the size of the host haloes in our sample (4.8–3.0 arcsec for redshifts 2.0–5.85; see Table 1). We consider contributions only from the targeted source, allowing us to separate out any issues due to source confusion (unlike a real observation). We stress that this detection algorithm is different from the one in our previous work in real space (Behrens et al. 2017) where we adopted an adaptive Friends-of-Friends (FoF) grouping only along the directions perpendicular to the LOS. The previous FoF algorithm

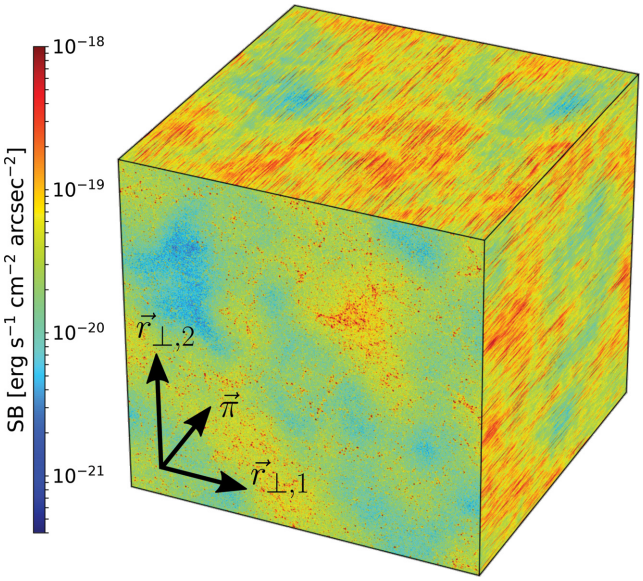

Figure 1. Lyman α intensity map of the simulation box with a length of 75 Mpc h^{-1} at a redshift of 3.01 projected on to the cube's faces. $\vec{\pi}$ denotes the line-of-sight direction, while $\vec{r}_{\perp,1}$ and $\vec{r}_{\perp,2}$ indicate the perpendicular direction to former. Thus, the top and right cube faces show a directional alignment stemming from RSDs. Most of the visible distortion is due to Lyman α RT and the subject of this paper. Individual spectra of LAEs are later reconstructed from such cube.

naturally leads to the source confusion, making an interpretation of the redshift-space clustering more complicated.

The spectrum of an LAE is computed with respect to the comoving frame at the line-of-sight distance of a given emitter. We impose a fixed spectral resolution in terms of the velocity shift as 24.7 km s^{-1} ($R \sim 12\,000$) for which we found the Monte Carlo sampling shot noise to be negligible but still have a sufficient spatial resolution mapping the velocity differences into comoving distances in redshift space. As we show later, the resulting spectra I_λ will have one or more peaks. In this paper, we consider two localization methods for the line-of-sight position, $s_{\rm app}$, of the LAEs: Either by using the global spectral maximum at $\lambda_{\rm max}$ or by the spectral maximum at $\lambda_{\rm max,red}$ only considering the red wing with respect to the LAE's rest frame, i.e. $\Delta\lambda > 0$.

Once a peak has been identified, this allows to define a corresponding apparent line-of-sight velocity of the emitter as

$$v_{\rm app} = c \times \frac{\lambda_{\rm max} - \lambda_{{\rm Ly}\,\alpha}}{\lambda_{{\rm Ly}\,\alpha}}, \tag{19}$$

where c is the speed of light, $\lambda_{{\rm Ly}\,\alpha}$ the rest-frame Lyman α line centre wavelength ($\sim$1216 Å). The comoving position in redshift space is given by equation (6), i.e.

$$\boldsymbol{s}_{\rm app} = \boldsymbol{r} + \frac{v_{\rm app}}{aH}\hat{\boldsymbol{r}}. \tag{20}$$

The apparent line-of-sight position of the emitter is set by adding the apparent velocity to the position from the halo catalogues. This completes the localization of the emitters and we construct mock LAE catalogues containing the position of detected emitters in redshift space as well as their velocity and spectra.

While we focus on the distortions to the line-of-sight component of the position, we also tested the distortion that RT induces for the angular component. We found that, except for source confusion, deviations from the LAEs' actual positions to be negligible in comparison to those arising in the line-of-sight direction.

Wavelengths and frequencies in this paper are always evaluated at the emitters' cosmological redshift. It is important to note that there are two distinct velocities contributing to $v_{\rm app}$: the peculiar velocity of a halo, $v_{\rm pec}$, and the velocity offset induced by the RT, $v_{\rm RT}$. While $v_{\rm RT}$ is not a physical velocity as it only represents the shift of Lyman α photons from the line centre, we sometimes refer to it short-handedly as radiative velocity for the sake of brevity. We can estimate $v_{\rm RT}$ from measured $v_{\rm app}$ and $v_{\rm pec}$ from the halo catalogues as

$$v_{\rm app} = \boldsymbol{v}_{\rm pec} \cdot \hat{\boldsymbol{r}} + v_{\rm RT}. \tag{21}$$

Velocities and positions without vector notation are the magnitude along the line of sight if not stated otherwise. The same decomposition is also done for $\lambda_{\rm max,red}$, so that the RT velocity offset is inferred from the peak in the red part of the line centre with respect to the halo's rest frame.

While we choose the host haloes' velocity for the peculiar velocity, we checked that the qualitative reasoning remains the same for other choices such as the linear velocity field or the velocity of the star-forming regions. We will show the convenience and significance of this velocity decomposition to determine $v_{\rm RT}$ in Sections 4.2 and 4.3.

3.2.2 *Measuring the two-point clustering statistics*

Next, we use the LAE catalogues created in Section 3.2.1 to compute the two-point statistics, $\xi_{\rm g}(\boldsymbol{r})$ and $P_{\rm g}(\boldsymbol{k})$.

We use `halotools` (Hearin et al. 2016)[3] to compute the TPCF of our LAE samples with the Landy–Szalay estimator (Landy & Szalay 1993)

$$\xi_g(\boldsymbol{r}) = \frac{{\rm DD}(\boldsymbol{r}) - 2{\rm DR}(\boldsymbol{r}) + {\rm RR}(\boldsymbol{r})}{{\rm RR}(\boldsymbol{r})}, \tag{22}$$

where $\boldsymbol{r} = \boldsymbol{x} - \boldsymbol{x}'$ denotes the separation between a pair of emitters. DD, DR, and RR represent LAE–LAE, LAE–random, and random–random pair counts found at the given separation for a given spatial binning width $\Delta r = \frac{5}{6}$ Mpc h^{-1}. The pair separation can either be evaluated in real space or redshift space. In real space, we expect an isotropic clustering when ignoring RT, so that for such case we can characterize $\xi(\vec{r})$ as $\xi(r)$ with $r \equiv |\boldsymbol{r}|$. As the line-of-sight positions change in redshift space, we express the signal as a function of parallel (π) and perpendicular ($r_{\perp}$) separation to the line-of-sight component.

To estimate $P_{\rm g}(\boldsymbol{k})$, we make use of the Fast Fourier Transformation (FFT). For this purpose, we first assign LAEs to a three-dimensional grid with 512 cells in each direction (i.e. the Nyquist frequency is $k_{\rm Nyq} \sim 21.4$ h Mpc^{-1}.) with the Triangle Shape Cloud interpolation. Next we perform the FFT to obtain the LAE number density field on the grid in Fourier space, and then measure the power spectrum as

$$P_{\rm g}(\boldsymbol{k}) = \frac{1}{N_{\rm modes}} \sum_{\boldsymbol{k}\,\text{in bin}} \delta_{\rm g}(\boldsymbol{k})\delta_{\rm g}(\boldsymbol{k})^*, \tag{23}$$

where $N_{\rm modes}$ is the number of Fourier modes on the grid which fit within a range of a given bin, e.g. $[k - \Delta k/2,\ k + \Delta k/2]$ and $[\mu - \Delta\mu/2,\ \mu + \Delta\mu/2]$. We suppress the aliasing effect by applying the interlacing technique as presented in Sefusatti et al. (2016) and

[3] http://halotools.readthedocs.io/en/latest/

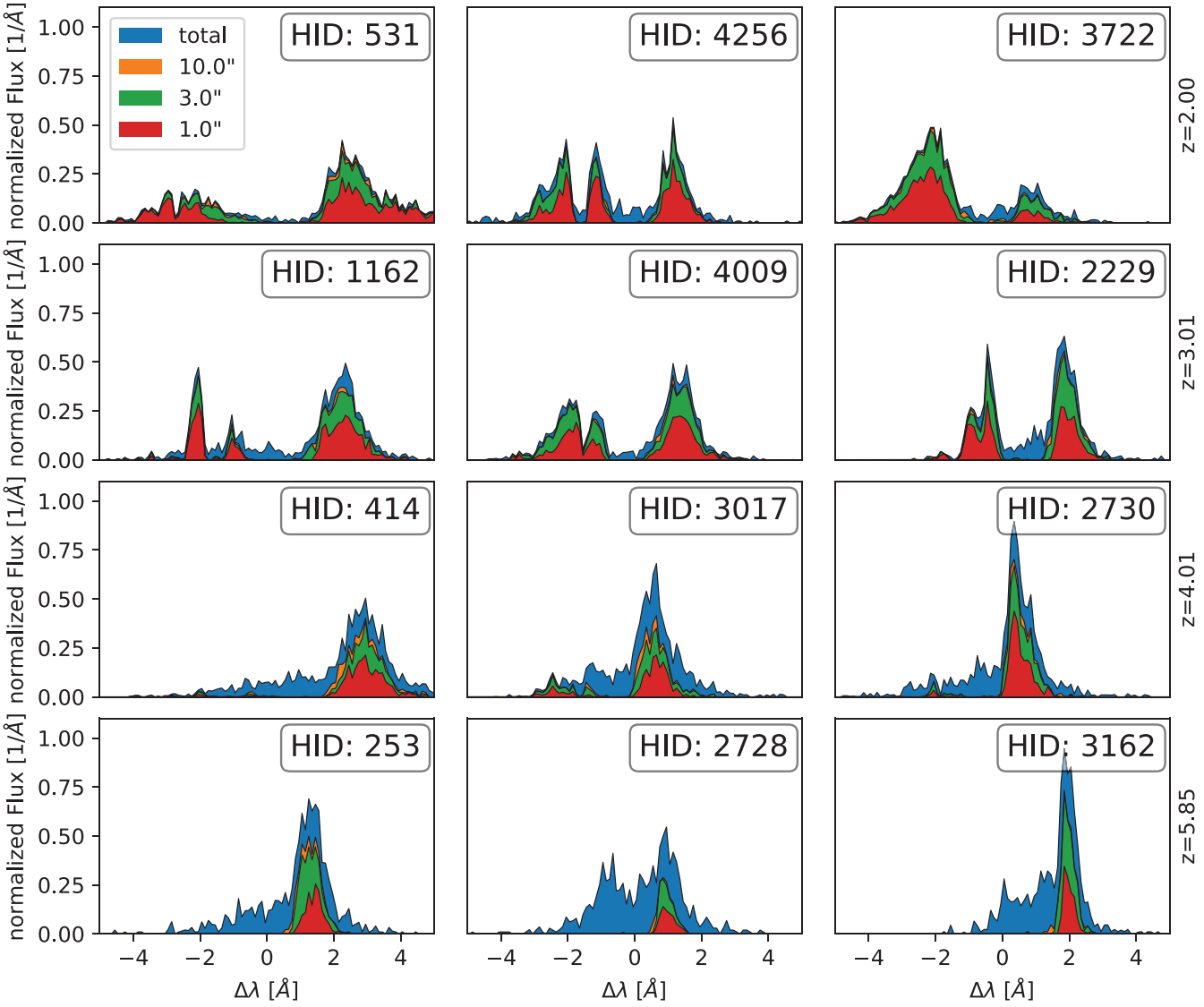

Figure 2. Randomly selected spectra at different redshifts in the comoving rest frame at the emitters' position. The halo IDs (HIDs) are given with respect to position in Illustris' halo catalogues for given redshift. Different colours show the flux for different aperture sizes. As larger apertures always enclose the flux of smaller aperture, only the excess flux over flux from smaller apertures is shown. The spectra show plenty of different characteristics with double peaks being common for lower redshifts and rare at high redshifts. Emission becomes more diffuse at higher redshift as apparent from larger flux contributions from larger apertures.

also subtract out the Poisson shot noise which properly takes the interlacing correction into account.

Another two-point statistics is the pairwise velocity distribution $\mathcal{P}(u_z, \vec{r})$, which encodes all the information in real space required to describe induced RSD (see equation 15) on the clustering signal. We compute the distribution by

$$\mathcal{P}(u_z, \boldsymbol{r}) = \frac{\mathrm{DD}(u_z, \boldsymbol{r})}{\mathrm{DD}(\boldsymbol{r})}, \tag{24}$$

i.e. counting the direct LAE-pairs within at a given separation $r \in [r - \Delta r/2,\, r + \Delta r/2]$ and line-of-sight velocity $u_z = \hat{\boldsymbol{r}} \cdot (\boldsymbol{v}_2 - \boldsymbol{v}_1) \in [u_z - \Delta u_z/2,\, u_z + \Delta u_z/2]$. The sign convention is chosen such that an infalling motion corresponds to a positive pairwise velocity (also across periodic boundaries). For the binning, we chose $\Delta r = 1.0\,\mathrm{Mpc\,h^{-1}}$ and $\Delta u_z = 60\,\mathrm{km\,s^{-1}}$, because the number of pairs for central galaxies of interest here quickly goes down below these scales.

4 RESULTS

4.1 Spectra

In Fig. 2 we show a random selection of LAE spectra for the simulated redshift range between $z = 2.0$ and $z = 5.85$. The spectra are evaluated with respect to the comoving rest frame at the emitters' position. Most emitters at redshifts $z = 2.00$ and $z = 3.01$ show a characteristic double-peaked spectrum, which is expected for optically thick environments as shown in fully homogeneous and isotropic analytic toy models (Adams 1972; Harrington 1973; Neufeld 1990) or more recent simulated isotropic shell models for

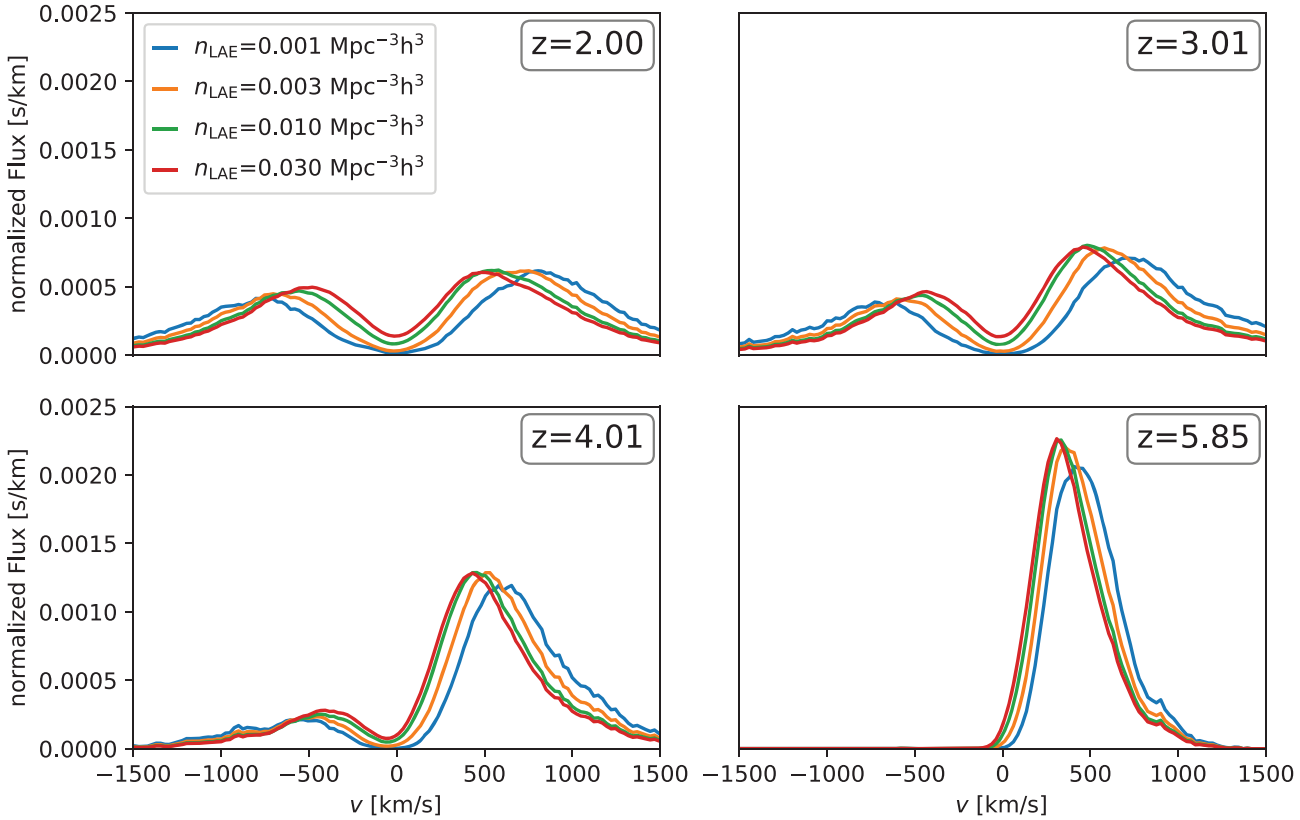

Figure 3. Stacked spectra in the haloes' frame for different number densities and redshifts for an aperture radius of 3 arcsec. The overall flux is normalized with respect to the sample of chosen number density $n_{\rm LAE}$.

small offset velocities (Ahn, Lee & Lee 2003; Verhamme et al. 2008).

At $z = 4.01$ double-peaked spectra become significantly sparser and are only exceptional cases at $z = 5.85$. This redshift evolution is mostly related to the decreasing transmissivity in the IGM at higher redshifts (Laursen, Sommer-Larsen & Razoumov 2011) due to the increasing neutral hydrogen density, explaining the disappearing of blue peaks at high redshift and a stronger trough around the line centre. At the same time the spatial luminosity profiles become more extended as scattering in the CGM increases with redshift. This trend shows up in Fig. 2, as relative contributions to the flux at fixed finite aperture becomes lower at higher redshift.

Our simulations seem to overestimate the abundance of double-peaked emitters when compared to observations at high redshifts $z = 2.0-3.0$, where the IGM interaction is limited. Typically observations find fractions of ~ 30 per cent at these redshifts in star-forming galaxies (Kulas et al. 2012; Herenz et al. 2017), while most of our emitters show double-peaked emitters. We shortly discuss the shortcomings in the modelled spectral shape and thus also the overprediction in the double-peaked profiles in Section 5.2.

More complex spectra are present depending on the underlying H I distribution and velocity structure, such as emitters with $n \geq 3$ peaks particularly in the blue part of the spectrum. See HID 4256 at $z = 2.0$ or HID 1162 at $z = 3.01$ for such examples.

The redshift evolution becomes particularly apparent when computing the stacked profiles $I_{\rm stacked}$ of the emitters as shown in Fig. 3. Here we vary redshifts and LAE number densities $n_{\rm LAE}$ for which we consider only the $n_{\rm LAE} L_{\rm box}^3$ emitters with the highest apparent luminosity. Imposing a number density threshold has several advantages when compared to a surface brightness threshold, including independence from the proportionality constant in the luminosity model in equation (18) and controlled shot noise behaviour for measuring LAE power spectra. The spectra are stacked in the haloes' rest frame. Stacking in the halo's rest frame reveals a trough at $v = 0\,{\rm km\,s^{-1}}$ in the spectra caused by a combination of IGM attenuation and strong frequency diffusion into the wings due to high-optical depths. As the number density $n_{\rm LAE}$ is reduced, the peaks of the stacked profiles move to higher offsets. At the same time, the dispersion $\sigma_{\rm stacked}$, the stacked profiles' second central moment for the respective red or blue spectral peak, also slightly increases but only on the per cent level. We motivate the significance of this dispersion in Section 5.2. The change in $\sigma_{\rm stacked}$ mostly stems from a significant correlation between brighter sources and more massive sources which in turn have a larger trough/peak separation.

4.2 Distinguishing distortion contributions

Fig. 4 shows the one-point PDF, $P(v)$, decomposed into $v_{\rm RT}$ and $v_{\rm pec}$ for $z = 3.01$ as suggested in equation (21). The projected PDFs on to the $v_{\rm pec}$ and $v_{\rm RT}$ axes give the one-point PDF, $P(v_{\rm pec})$, and $P(v_{\rm RT})$, respectively. There is no noticeable correlation between the two velocity components in Fig. 4. We checked that the same holds true for other redshift outputs as well. This result is expected to some extent, since two physically distinct processes are responsible for each velocity component. As we discuss later, this independence allows us to model additional RSD due to the RT effect independently of the peculiar velocity.

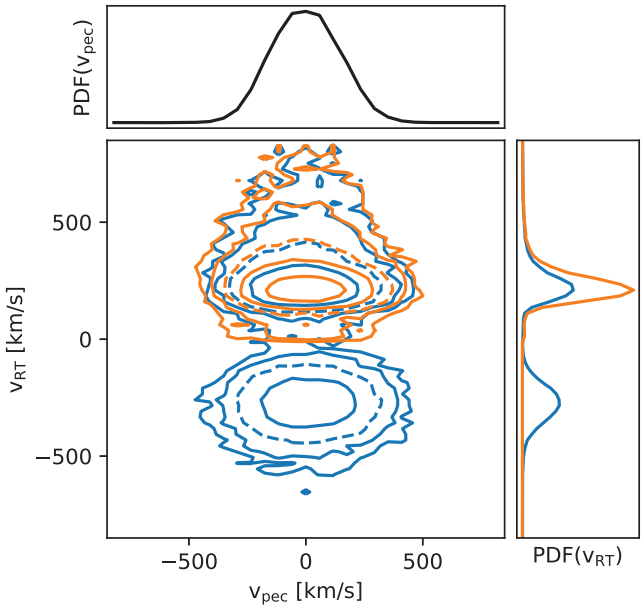

Figure 4. Contour plot of emitters' peculiar and radiative velocities. Detection of the global peak (*blue*) and the peak in the red wing of the emitters' rest frame (*orange*).

4.3 RT velocity PDF

As we discussed in Section 2, the velocity PDF, i.e. a probability distribution of radial position of LAEs with respect to real-space position (see e.g. equation 13), determines the redshift space clustering. We show the PDF of the peaks for the brightest emitters detected by the observer for a given LAE number density $n_{\rm LAE}$ in Fig. 5. The peak distributions show a strong redshift evolution: As we move to higher redshifts, the blue peaks are strongly suppressed and practically non-existent at $z = 5.85$. When decreasing the number density, the selected emitters are restricted to the more luminous ones, which in turn have a higher average peak offset due to their higher optical depth.

Although we have already seen similar trends in the stacked spectra as expected, it is important to notice that the velocity PDFs and the stacked spectra are *not* the same. We discuss a possible correlation between the distribution of stacked spectra and the velocity PDF in Section 5.2. Since only the stacked spectra are directly observable, such relation will prove itself crucial in estimating the additional RT effect in observational surveys stemming from the velocity PDF.

In Fig. 5, we also provide a comparison of the $v_{\rm RT}$–PDF with a Gaussian of same mean and variance for the red peaks. We find that the distributions quickly fall off towards the line centre while they extend more towards larger velocity offsets. This leads to a higher kurtosis than for the Gaussian and also a positive skewness, which will affect the quality of using a Gaussian approximation for the damping in the next sections.

4.4 Configuration space: TPCF and the pairwise velocity PDF

In Fig. 6 we show the correlation functions $\xi(\pi, r_{\rm ll})$ measured from the mock observations with an LAE number density of $n_{\rm LAE} = 0.01\,{\rm h^3\,Mpc^{-3}}$. Different columns show the clustering at different redshifts in an increasing order, while different rows show different velocity contributions (as defined in Section 3.2.1) added on to the real space configuration to obtain the redshift space clustering. In the first row we set the velocity contribution $v_{\rm app}$ to zero, so that we plot the real space clustering. The second row shows the redshift space result using the peculiar velocity $v_{\rm pec}$ from the halo catalogues and thus explicitly omitting the contribution from RT. We stress that the second row is often presented as the redshift-space clustering of LAEs in the literature (Zheng et al. 2011; Gurung-Lopez et al. 2019a) but does not yet directly correspond to an observable radial position from the redshift measurement. Instead, the apparent $v_{\rm app}$ is the observable containing both contributions from the complex RT and the peculiar velocity. The third/fourth ($v_{\rm app}/v_{\rm app,red}$) row shows inferred overall position from the peaks in the spectra, which includes both peculiar velocity and RT effects. The fifth/sixth ($v_{\rm RT}/v_{\rm RT,red}$) rows show the RT component of the velocity only as the residual of the apparent and peculiar velocity.

Given the same simulation set-up as Behrens et al. (2018), we expect a very similar result to those shown for the real-space clustering in the first row. Only minor differences arise from an increased Monte Carlo photon count and the simplified detection algorithm. Note that there is a slight anisotropy in the clustering signal in real space. As we stressed in Behrens et al. (2018), the slight anisotropy of this dataset does not originate from an RT effect and was statistically consistent with zero.

As introduced in Section 2, we confirm two competing RSD effects in Fig. 6. We see that on the shown scales ($1-10\,{\rm h^{-1}\,Mpc}$) the Kaiser effect dominates the RSDs from the peculiar velocity field $v_{\rm pec}$ over the usual FoG effect due to random motion of the LAEs in the second row. Once $v_{\rm RT}$ is added, however, the small-scale damping from $v_{\rm RT}$ is significant on these scales and thus the overall redshift space clustering with the apparent velocities $v_{\rm app}$ is elongated along the line of sight despite the squashing from the Kaiser effect (third and fourth rows).

In order to investigate the distortions of the TPCF in redshift space more quantitatively, we discuss the pairwise velocity PDF which encodes the full information of RSD (see equation 15). Note that we do not report the measurement of the multipole moment as we did in Behrens et al. (2018). Before showing the pairwise velocity PDF, let us first extend the discussion in 2 in the presence of two physically distinct velocity contaminations $v_{\rm pec}$ and $v_{\rm RT}$ (see equation 21). In Section 4.2, we show that there is no apparent correlation between $v_{\rm pec}$ and $v_{\rm RT}$ at the level of the one-point PDF, i.e. $\langle v_{\rm pec}(\boldsymbol{x}) v_{\rm RT}(\boldsymbol{x})\rangle = 0$. We further assume that $v_{\rm RT}$ has no correlation with the density field or the peculiar velocity at scales of interest, i.e. $\langle u_{\rm RT}(\boldsymbol{x})\delta_{\rm g}(\boldsymbol{x}')\rangle = \langle u_{\rm RT}(\boldsymbol{x}) u_{\rm pec}(\boldsymbol{x}')\rangle = 0$. Under this simple setting, equation (17) follows that

$$\mathcal{M}(if\gamma;\boldsymbol{r}) = \frac{\left\langle {\rm e}^{if\gamma\Delta u_{{\rm pec},z}}\left[1+\delta_{\rm g}(\boldsymbol{x})\right]\left[1+\delta_{\rm g}(\boldsymbol{x}')\right]\right\rangle}{1+\xi(r)} \langle {\rm e}^{if\gamma\Delta u_{{\rm RT},z}}\rangle. \tag{25}$$

Its FT, the pairwise velocity PDF, is written as

$$\mathcal{P}(u_z;\boldsymbol{r}) = (\mathcal{P}_{\rm pec} * \mathcal{P}_{\rm RT})(u_z;\boldsymbol{r}), \tag{26}$$

where $*$ denotes the convolution for simplicity with $\mathcal{P}_{\rm RT}$ given by

$$\begin{aligned}\mathcal{P}_{\rm RT}(u_z) &= \int \frac{{\rm d}\gamma}{2\pi} {\rm e}^{-{\rm i}\gamma u_z} \langle {\rm e}^{if\gamma\Delta u_{{\rm RT},z}}\rangle \\ &= \int \frac{{\rm d}\gamma}{2\pi} {\rm e}^{-{\rm i}\gamma u_z} \left| \int {\rm d}u\, P_{\rm RT}(u) {\rm e}^{if\gamma u}\right|^2. \end{aligned} \tag{27}$$

Notice that the $u_{\rm RT}$ contribution in equation (25) is not weighted by the density field at different scales and, as a result, the ensemble average becomes an integration over one-point PDF, $P_{\rm RT}(u_z)$. In

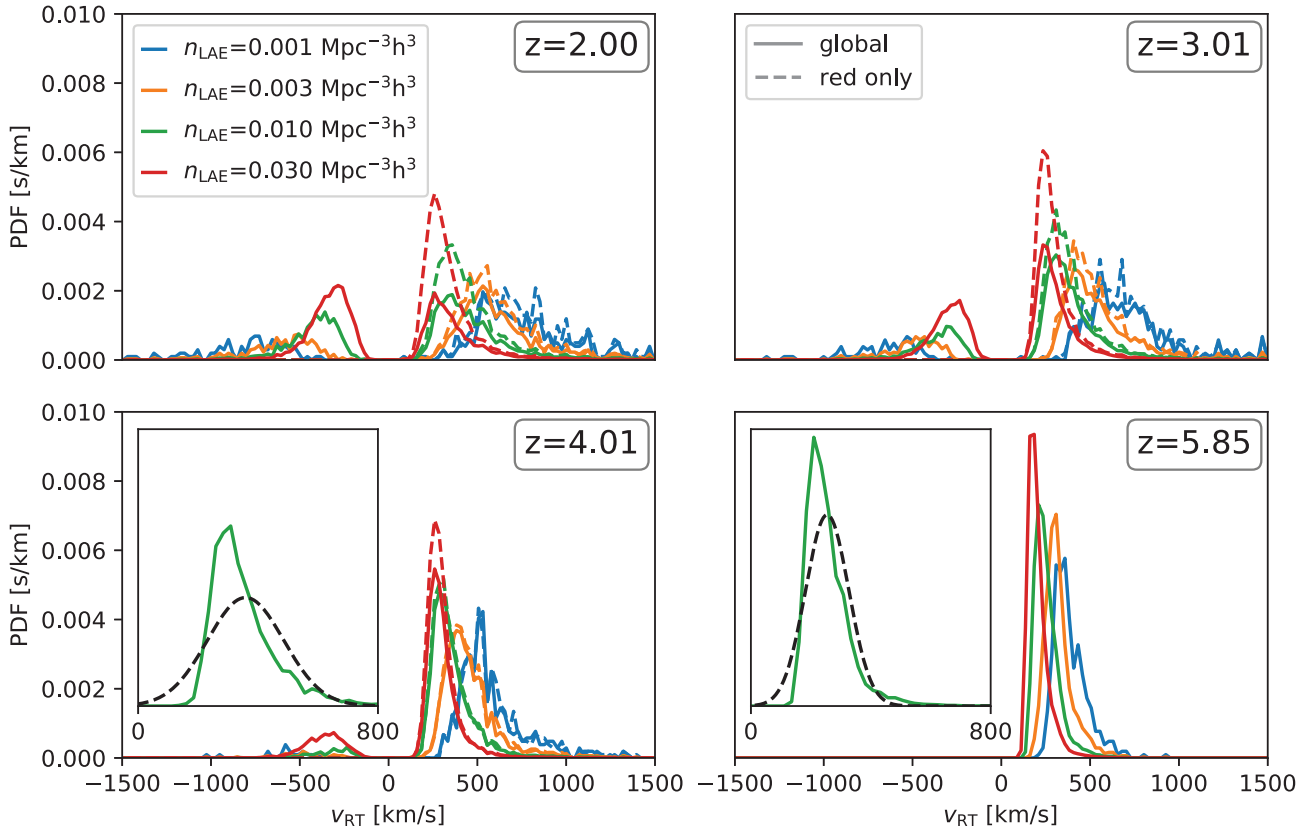

Figure 5. RT velocity offset $v_{RT,global}/v_{RT,red}$ distributions for different number densities n_{LAE} and redshifts. Solid lines show the velocity offset being deduced from the global peak of each LAE's spectra and dashed lines with the velocity offset deduced from the red peak only. For redshift $z = 3.01$ and 4.01 we show the distribution of $v_{RT,global}$ for the red peak at $n_{LAE} = 0.01\text{Mpc}^{-3}\,\text{h}^3$ along with a Gaussian of matching mean and standard deviation. The scaling of the y-axis remains the same as for the other subplots. The velocity distributions for the respective wings are positively skewed with respect to the velocity offset to the line centre.

other words, the scale dependence in the pairwise velocity PDF in equation (26) comes only from the peculiar velocity part, $\mathcal{P}_{pec}(u_z;\vec{r})$.

In Fig. 7, we show the measured pairwise velocity PDFs at two different scales (solid and dashed lines for 1 and 10 cMpc h^{-1}, respectively) in cases of both the red-peak only (upper panel) and the global peak (lower panel) at $z = 3.01$. First of all, when we adopt v_{pec} (purple lines) as a velocity, we confirm a trend well known in the literature (see e.g. Scoccimarro 2004). We see a positive peak which corresponds to a coherent infall motion on large scale ($r_\parallel = 10$ cMpc h^{-1}), while the PDF follows a distribution with an exponential tail at small separation, $r_\parallel = 1$ Mpc h^{-1}. On the other hand, we do not confirm such a trend in the case of v_{RT}. In general, a large tail of the pairwise v_{RT}–PDF contributes to that of the pairwise v_{app}–PDF even at a large separation of 10 cMpc h^{-1}. This impact is clearly more significant for the global peak only case than for the red-only peak case.

To see the scale dependence of the pairwise velocity PDFs more explicitly, we show three low-order moments of the PDFs (mean, dispersion, and kurtosis) as a function of a separation scale in Fig. 8, where we further confirm the aforementioned trends. The mean infall velocity (upper panel) is dominated by v_{pec} particularly on large scales, and the impact of v_{RT} on the mean is marginal even in the global peak case. This is the reason why the Kaiser squashing effect is seen on large scales in Fig. 6. Meanwhile, the dispersion (middle panel) is mainly dominated by v_{RT} for both detection algorithms, and notably has little scale dependence. Interestingly, the dominant contribution to the kurtosis (lower panel) depends on the peak finding algorithm: As expected, the kurtosis of v_{pec} deviates from the Gaussian value of 3 at small scales due to its extended tail (e.g. Scoccimarro 2004). In the case of the red peak detection method, a much wider distribution due to v_{RT} dominantly contributes to the kurtosis than for v_{pec}. On the other hand, both v_{pec} and v_{RT} roughly equally contribute to the kurtosis for the global peak detection method. We do not find a strong scale dependence of the kurtosis for both cases. Since the dispersion and higher order moments of the pairwise velocity PDF contribute to the FoG elongation (Scoccimarro 2004), this evidence suggests that random velocity following the one-point PDF plays a main role in the FoG elongation in Fig. 6.

In Fig. 8, we also show the three moments of the v_{RT} estimated from v_{tot} and v_{pec}, assuming that $\mathcal{P}_{pec}$ and $\mathcal{P}_{RT}$ are independent of each other (dashed lines compared with solid lines with the same colours). Although they are roughly consistent, we see small discrepancies between solid and dashed lines, implying that either $\langle v_{RT}(\boldsymbol{x})\delta_g(\boldsymbol{x}')\rangle$ or $\langle v_{RT}(\boldsymbol{x})v_{pec}(\boldsymbol{x}')\rangle$ is not exactly equal to zero.

4.5 Fourier space: the FoG damping factor

In the previous subsection, we show that the impact of v_{RT} on the pairwise velocity PDF can be roughly understood in terms of the one-point PDF, $P_{RT}(u_z)$. However, it is not straightforward to quantify its impact on the observable TPCF because of the involved convolution (see equation 26). It is simpler to work in Fourier space, since the convolution becomes a multiplication after FT.

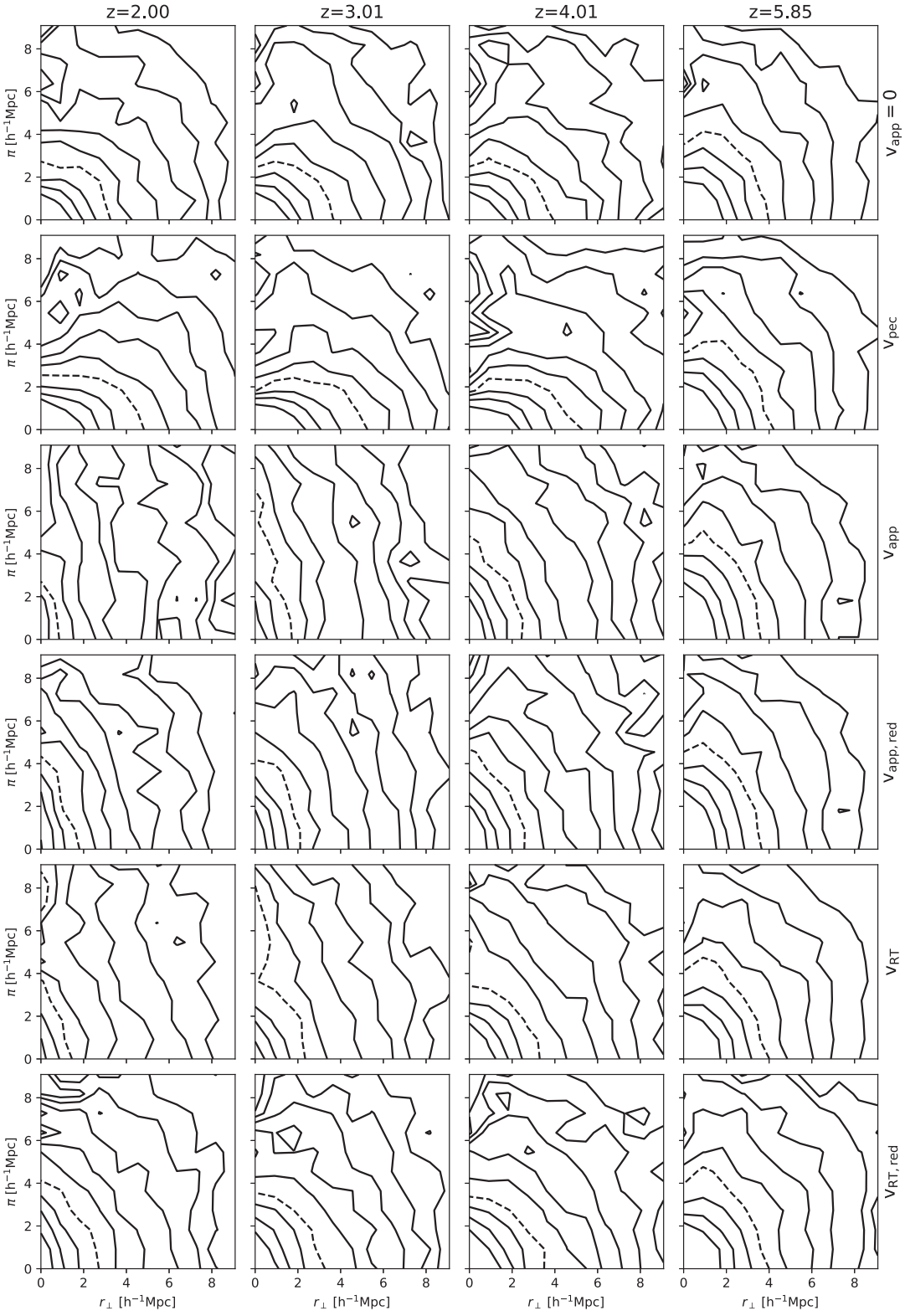

Figure 6. TPCF for disentangled RSD of visible LAEs for $n_{\rm LAE} = 10^{-2}h^3\,{\rm Mpc}^{-3}$ in redshift space. The dashed contour corresponds to $\xi = 1$ with contours decreasing by a factor of 1.4 further away from the origin. In the first row, we set the apparent velocity $v_{\rm app}$ to zero, i.e. we show the real space clustering. In the second row, we only consider the peculiar velocity $v_{\rm pec}$ of the corresponding host halo. The third and fourth rows show the apparent overall velocity shifts $v_{\rm app}$ detected from the global peak and the red peak, respectively. The last two rows show the clustering when only the RT contribution is considered.

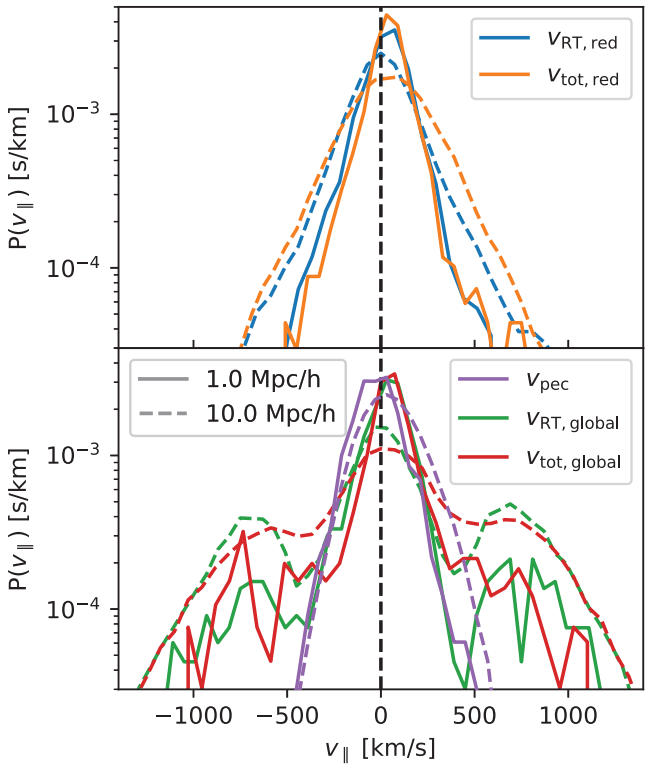

Figure 7. The pairwise velocity distribution $\mathcal{P}$ for the peculiar and RT velocity component evaluated at different length scales for $z = 3.01$ with $n_{LAE} = 0.01 h^3\,\mathrm{Mpc}^{-3}$ along the line of sight. The upper panel shows the RT and total apparent velocities ($v_{RT,red}/v_{tot,red}$) based on the red peak only detection method. The lower panel shows the same for the global peak detection method plus the peculiar velocity distribution v_{pec}. Different line styles indicate different spatial separations $\vec{r}$.

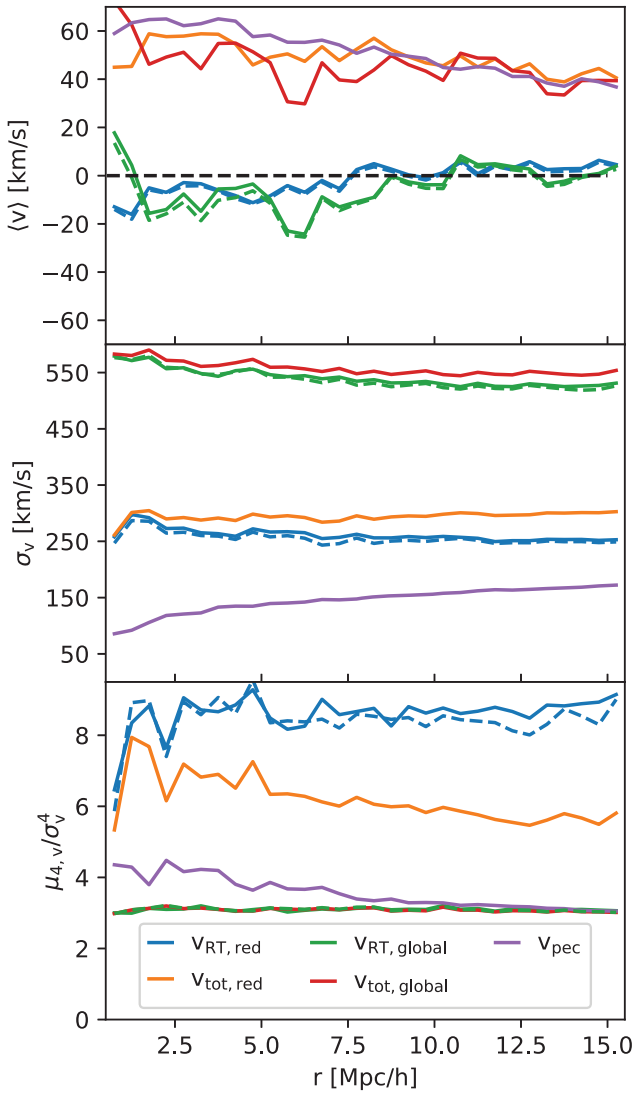

Figure 8. The mean, standard deviation, and kurtosis of the pairwise velocity distribution $\mathcal{P}$ for the peculiar and RT velocity component as a function of the total separation r of emitter pairs at $z = 3.01$ with $n_{LAE} = 0.01\,h^3\,\mathrm{Mpc}^{-3}$ along the line of sight. The solid lines show the respective velocity contributions as listed in the legend, i.e. the apparent and RT velocity offset for each the global and red-only detection method along with the peculiar velocity. In the dashed lines we show the expected RT velocity contribution v_{RT} at a given scale under the assumption that v_{RT} and v_{pec} are independent distributions. Thus, the difference between the solid and dashed line indicates the strength of any correlation between these two contributions present on a given scale. The colour coding is chosen to be consistent with Fig. 7.

Since equation (26) follows

$$1 + \xi^s_g(r_\perp, \pi) = \left(\left[1 + \xi^s_{g,pec}(r_\perp, \pi) \right] * \mathcal{P}_{RT} \right)(u_z), \tag{28}$$

we obtain the FoG damping in Fourier space with a help from equation (27),

$$D^{RT}_{FoG}(k, \mu) \equiv \frac{P^s_{g,tot}(\boldsymbol{k})}{P^s_{g,pec}(\boldsymbol{k})} = \left| \int \mathrm{d}u_z \, P_{RT}(u_z) \mathrm{e}^{\mathrm{i} f k \mu u_z} \right|^2, \tag{29}$$

where $P^s_{g,tot}(\vec{k})$ and $P^s_{g,pec}(\vec{k})$ denote the redshift-space power spectra when we adopt the total apparent velocity and the peculiar velocity only as a velocity contamination, respectively. This relation allows us to directly compare the line-of-sight damping from RT as found in the measured redshift-space power spectra with an expected damping from the underlying one-point PDF. We stress that the last equality holds only under the assumptions we made in equations (25) and (26): namely, no correlation between v_{RT} and v_{pec} at the same point, and no spatial correlations between v_{RT} and δ_g (and v_{pec}). Furthermore, we will compare this damping with the two generic functions commonly adopted in the literature, i.e. the Gaussian and the Lorentzian damping functions (see equations 12 and 14). Here the second central moment σ is directly calculated from the one-point PDF.

In Figs 9 and 10, we show such direct comparisons for the global-peak and the red-peak cases, respectively. We show the measurements of the ratio of the two redshift-space spectra as a function of the mode parallel to the line of sight, $k_\parallel = k\mu$, colour-coding them by their absolute wavenumber, k. In addition, we plot the expected damping from the velocity PDF as implied by equation (29) (solid lines) and the two generic fitting functions with the second central moment σ of the according velocity PDF (dashed and dotted lines for the Gaussian and the Lorentzian functions, respectively). In general, we find a strong damping even on relatively large scales, $k \gtrsim 0.1\,h\,\mathrm{Mpc}^{-1}$, and strength of the damping depends on the peak detection algorithm and redshift. There is a typical redshift evolution with stronger damping at lower redshift, originating from two contributions. First, as seen in Fig. 5, the velocity distribution widens at lower redshifts, translating to a larger damping scale. This is mainly because the

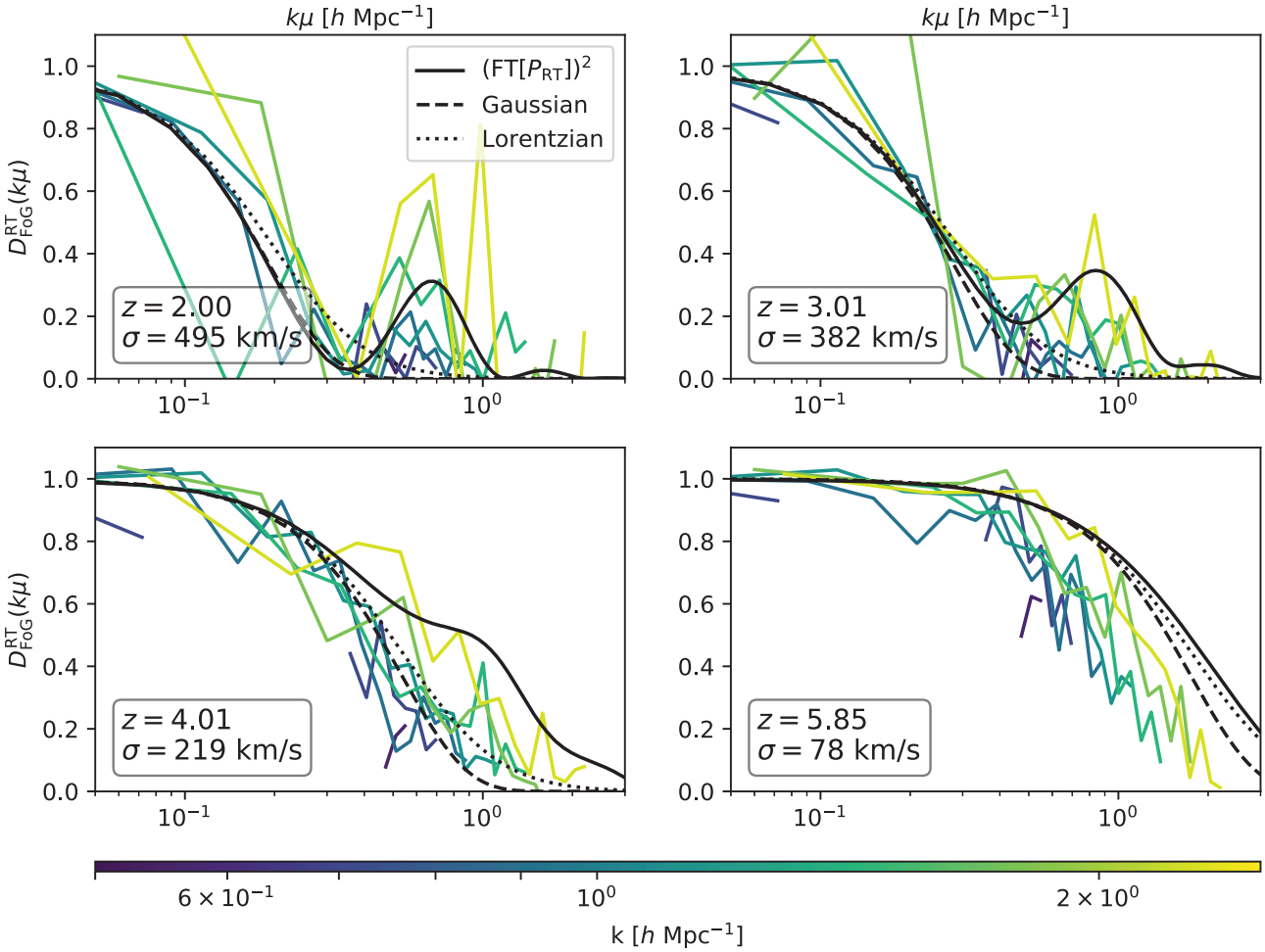

Figure 9. Damping factor as a function of the line-of-sight frequency $k_{\parallel} = k\mu$ due to Lyman α RT velocity offset $v_{\rm RT,global}$. Emitter positions are assumed to coincide with the global peak. Different coloured lines represent measurements of $D(k_{\parallel})$ from the mock catalogues at different total frequencies k. The black solid lines show the Fourier transform of the one-point PDF (see equation 29), while the dashed and dotted lines show the Gaussian and the Lorentzian forms, respectively (see equations 12/14). *Left-to-right, top-to-bottom*: $z = 2.00$, $z = 3.01$, $z = 4.01$, $z = 5.85$.

neutral hydrogen density in CGM for the threshold sample becomes larger at lower redshift (see also Behrens et al. 2017). Secondly, distance and velocity are related via the mapping from real to redshift space as given by equation (20). As the factor of aH roughly scales as $a^{-1/2}$ in the analysed redshift range, fixed velocity dispersions correspond to larger damping length scales at lower redshifts.

We also find that the measurements are in fair agreement with the squared direct FT of the one-point PDF (solid lines) at $k\mu \lesssim 1\,h\,{\rm Mpc}^{-1}$, although we see some discrepancies in detail. We discuss some possibilities to cause the discrepancies in Appendix C, but do not provide a decisive reason. Comparing the solid lines with the dashed and dotted ones, we see that all of them can qualitatively trace the damping feature as a function of scale, although there are slight differences in detail. These differences are expected given the fact that the one-point PDFs follow neither Gaussian nor exponential distributions (see Section 4.3). It is apparent that our measurements are too noisy to conclude which model works the best. Nevertheless, this result suggest that the dispersion of the one-point PDF, σ, is a good proxy for the resulting FoG damping due to RT.

Strong evidence that our measurement is consistent with the direct FT of the one-point PDF comes from the oscillatory behaviour in the solid lines. This can be understood as follows. Suppose that we model the double-peaked PDF as $P_{\rm RT}(v) = f_1(v) + f_2(v + \Delta v)$ where $f_i(v)$ is a symmetric distribution with a peak and Δv denotes the separation of the two peaks. Then we have

$$\begin{aligned} D_{\rm FoG}^{\rm toy} &= |{\rm FT}\,[P_{\rm RT}]|^2 = |{\rm FT}\,[f_1(v) + f_2(v+\Delta v)]|^2 \\ &= |{\rm FT}\,[f_1]|^2 + |{\rm FT}\,[f_2]|^2 + 2{\rm FT}\,[f_1]\,{\rm FT}\,[f_2]\cos\,[k\Delta v]\,. \end{aligned} \tag{30}$$

From this toy model we see that the first two terms give the FoG-like damping due to each of the single-peaked distributions, while the last term gives an oscillation due to Δv. In other words, the oscillatory behaviour originates from the double-peaked distribution in the global-peak case. The oscillations disappear at higher redshifts, since the second peak in the PDF at a blue end is suppressed by the attenuation due to neutral hydrogen in IGM, as we visually confirm in Section 4.3. Furthermore, one can quantify when the oscillatory term becomes dominant by approximating both peaks by a Gaussian distribution with the same dispersion σ:

$$D_{\rm FoG}^{\rm toy,Gaussian} = \underbrace{\exp\left(-\sigma^2 k^2\right)}_{D_{\rm FoG}^{\rm Gaussian}} \underbrace{\left[w_1^2 + w_2^2 + 2w_1 w_2 \cos\left(k\Delta v\right)\right]}_{D_{\rm FoG}^{\rm osci}}, \tag{31}$$

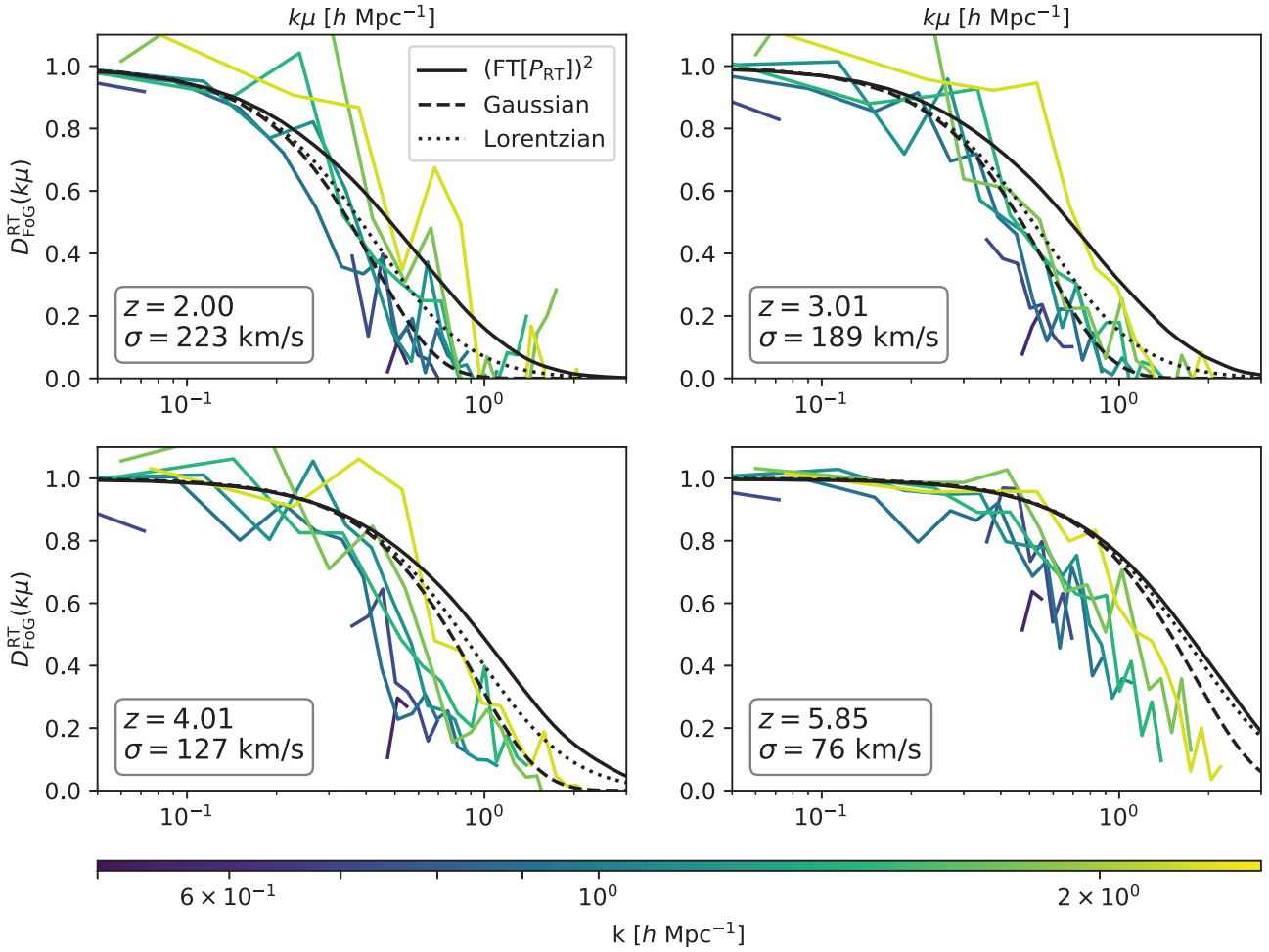

Figure 10. Damping factor as a function of the line-of-sight frequency $k_{||} = k\mu$ due to Lyman α RT velocity offset $v_{\rm RT,red}$. Emitter positions are assumed to coincide with the global peak in the red wing with respect to the halo's frame. Different coloured lines represent measurements of $D^{\rm RT}_{\rm FoG}(k_{\parallel})$ from the mock catalogues at different total frequencies k. The black lines show the Fourier transform of the one-point PDF (see equation 29), the Gaussian and the Lorentzian form (see equations 12/14). *Left-to-right, top-to-bottom*: $z = 2.00$, $z = 3.01$, $z = 4.01$, $z = 5.85$.

where w_1 and w_2 are the relative contribution of the peaks to the PDF such that $w_1 + w_2 = 1$. By looking at the leading, second-order term of $D^{\rm Gaussian}_{\rm FoG}$ and $D^{\rm osci}_{\rm FoG}$ (assuming $k\Delta v \ll 1$ and $k\sigma \ll 1$), we find that the cosine term can dominate if

$$\frac{\Delta v}{\sigma} > \frac{1}{\sqrt{w_1 w_2}}. \tag{32}$$

In addition, notice that the impact of $D^{\rm osci}_{\rm FoG}$ is limited as it has a lower bound:

$$D^{\rm osci}_{\rm FoG} \geq (2w_1 - 1)^2 . \tag{33}$$

Applying this double-peaked Gaussian to our mock data, we find reasonable fits that could be further improved by substituting the Gaussians with a better fits for the peak PDFs. At all redshifts except for $z = 5.85$, where the blue peak has completely vanished, the Gaussian dominates according to equation (32). Looking at Fig. 9, we nevertheless see that the impact of the oscillatory part is weak. This can be attributed to the strong asymmetry in the height of the two peaks (e.g. $w_1 \sim 0.06$ at $z = 4.01$), which reduces the amplitude of $D^{\rm osci}_{\rm FoG}$ according to equation (33).

5 DISCUSSION

We have shown that significant FoG-like damping arises in the large-scale LAE clustering in redshift space as the peak positions of observed LAE spectra are affected by RT. We now discuss the most important steps and caveats concerning our findings for the FoG-like damping factor, $D_{\rm FoG}(k, \mu)$, for LAEs, and also investigate possibilities to mitigate the RT damping.

In Section 5.1, we discuss the shape of spectra arising in our simulations and possible shortcomings in our modelling. Next, we discuss the line-of-sight localization method, which reduces a spectrum to a radiative velocity offset $v_{\rm RT}$ in Section 5.2. Afterwards, we seek to use additional information from the spectra to reduce the damping effect in Section 5.3.

5.1 Spectra

We obtain individual spectra in our simulations by using an aperture method with a 3 arcsec radius as introduced in Section 3. Using this simple detection algorithm, there are only three relevant parameters impacting the damping factor that are related to instrumental specifications in a real observation: The spatial resolution (aperture

size), the spectral resolution and the number density threshold. The number density threshold roughly corresponds to a flux threshold which should be determined by the signal-to-noise ratio in real observations. We show the moderate impacts of both spectral resolution and aperture size on the $v_{\rm RT}$ distribution in Appendices A1 and A2. Similarly we have already discussed the impact of the number density in Section 4.3.

In our simulations, we obtain a manifold of different spectral shapes as shown in Fig. 2. Most prominently in this work, we seem to overpredict the amount of double-peaked emitters as we mentioned in Section 4. As the original Illustris simulations do not resolve the ISM, we effectively set the post-ISM spectrum to that of the input photons which follow a Gaussian distribution. Thus, many of the photons, which are close to the line centre, will be reprocessed on the CGM (rather than the ISM) scales to exhibit the rich dataset of spectra we obtain. It is not trivial to see how the lack of ISM modelling will affect the resulting spectra other than that it most likely reduces the fraction of double-peaked profiles when an ISM model with galactic outflows (Bonilha et al. 1979) is chosen. In fact, more recent work attributes the frequency redistribution to the ISM scales rather than the CGM scales and includes IGM attenuation as an additional separated step (Inoue et al. 2018; Gurung-López et al. 2019b). We do not expect that the addition of a sophisticated ISM model will change the qualitative nature of the newly modelled FoG-like damping: An ISM model, particularly with outflows, should not introduce additional large-scale correlations, since the large-scale correlation is mainly driven by the IGM attenuation at a blue end of the input spectrum (Behrens et al. 2017). Therefore, we expect the qualitative modelling as an FoG-like damping in redshift space remains modelling even when introducing a more sophisticated spectral modelling on ISM scales.

There is a series of other modelling shortcomings in our work including the lack of dust (Laursen, Sommer-Larsen & Andersen 2009), subgrid modelling (for sub-parsec clumps) (Gronke et al. 2016, 2017) and spatial resolution. We only consider Lyman α emission from recombination in the star-forming regions. Contributions from collisional excitation can be significant and show a bluer spectral signature (Smith et al. 2019). As initial photons are spawned from a point-source within the LAEs in our simulations and thus not reflecting different physical environments, such spectral modifications cannot be captured within our framework. As we are concerned with individual LAEs' spectra, we can neglect the small, fluorescent contributions in the IGM (Dijkstra 2017).

We note that the original Illustris simulations show a growing excess of neutral hydrogen (a factor of $\sim$3 between $z = 2$ and 4) compared to observations (Diemer et al. 2019) on galaxy scales where frequency diffusion of Lyman α photons gives rise to the FoG-like damping. This leads to an overestimate of the dispersion in the $v_{\rm RT}$–PDF. For the Neufeld solution (Neufeld 1990) such an hydrogen excess would correspond to an overestimate in the peak offsets $v_{\rm RT}$ of roughly 31 per cent ($\Delta v \propto \tau^{1/3} \propto n_{\rm H\,I}^{1/3}$).

As found in Behrens et al. (2018) for the real-space clustering, large-scale correlations, and LAE mock observables significantly change with the resolution of the underlying neutral hydrogen distribution in the RT simulations. Similarly, we give a possible explanation for the lack of detection of the additional RSD in Zheng et al. (2011) due to its limited hydrodynamic resolution in Appendix B.

Overall, we thus expect the observational spectral shapes to differ from those found in our simulations. As a result, our simulations do not reproduce the observed Lyman α luminosity function as addressed in Behrens et al. (2018). This problem of reproducing observables is common when not explicitly calibrating against Lyman α specific observations such as the luminosity function [e.g. calibration against other redshifts, see Kakiichi et al. (2016) and Inoue et al. (2018)].

From the spectra themselves, one can obtain stacked surface brightness profiles $I_{\rm stacked}$ as shown in Fig. 3. However, it is important to realize that it is the *unobservable* one-point velocity PDF, $P_{\rm RT}$, that mainly determines the additional FoG-like damping along the line-of-sight direction. Since the relation between $I_{\rm stacked}$ and $P_{\rm RT}$ is non-trivial, there is no a priori way of determining the damping factor from the stacked profiles. Nevertheless, we find an empirical relationship between the square root of the second central moment of $I_{\rm stacked}$ and $P_{\rm RT}$:

$$\sigma_{\rm PDF} \approx 0.4 \times \sigma_{\rm stacked} \tag{34}$$

for $n_{\rm LAE} = 10^{-2}\ h^3\ {\rm Mpc}^{-3}$ at $z = 3$. This relation typically changes less than 10 per cent among redshifts from 3.01 to 5.85. At redshift $z = 2.0$ this proportionality factor is consistently higher by a factor of $\sim$ 25 per cent. The number density has a large impact on the resulting relation and rising from 0.4 to 0.55 when restricting the number density from 0.01 to 0.001 $h^3\ {\rm Mpc}^{-3}$. This change is caused by an increased dispersion in the velocity distribution, while the width of the stacked profiles is nearly constant independent of imposed number density threshold. This relation was found in the haloes' rest frame, which is unknown unless a secondary emission line (e.g. Hα) is measured and its radial position is identified.

5.2 Localization along the line of sight

We proposed two simple localization methods to identify LAE's radial position in redshift space for given Lyman α spectra: Either by identifying the position of the maximal spectral flux with the LAE's position (referred to as 'the global peak'), or by identifying the LAE position with the maximal spectral fluxes at wavelengths longward of the actual LAEs' radial positions as used for the $v_{\rm RT,red}$ distribution (referred to as 'the red peak only'). Combination of our simulated spectra with the choice of the localization methods gives us the one-point velocity PDF, $P(v)$, which is a key quantity to understand the resulting FoG damping. Nevertheless, let us briefly discuss potential issues in $P(v)$ in light of the two localization methods here.

The major issue with the global-peak method is closely related to the fact that we tend to overestimate the number of double-peaked spectra in our simulations. As a consequence, the velocity distribution $P_{\rm RT}(v)$ has a strong double peak feature as well, as we see in Fig. 5, which does not most likely represent real observations. Furthermore, as shown in Section 4.5, such a double-peaked velocity distribution induces a larger FoG damping as compared with the $v_{\rm RT,red}$ distribution. In a simple toy model for a double-peaked velocity distribution, we see that two additional terms occur strengthening the damping, where one is induced by the second peak's width and the second by the separation between the peaks.

The second detection method does not suffer from these additional damping contributions and can be applied to both the double-peaked spectra and single-peaked spectra. However, there remains a chance to get an oscillatory damping component as well in this localization method if some LAEs' visible peaks are blue of the line centre, which is not distinguishable from a red peak in real observations without a complementary emission line. Nevertheless, it is ad hoc to completely ignore the blue part in the red-peak only method without any physically reasonable reasons. We thus stress

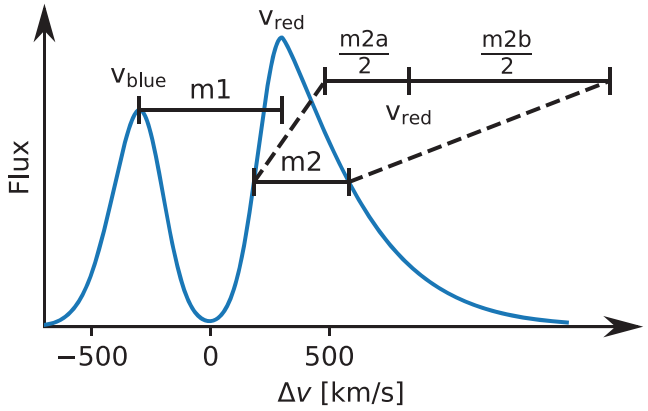

Figure 11. Sketch of the proposed correction methods (see text) applied to an example double-peaked spectrum.

that the velocity PDFs with two methods and the resultant FoG damping strength are showcases of two extreme situations.

The localization methods so far only use a single peak's frequency shift for the determination of the line-of-sight position. However, we can in principle try to use other information available through the individual spectra to determine the line-of-sight position. In the coming section, we investigate methods to further reduce the RT damping in the clustering signal based solely on additional information of the features in the Lyman α spectrum.

5.3 Correction to mitigate the impact of RT

We can try to correct for the presented Lyman α RT distortion effect by utilizing the full spectral information available for the emission line. To do so, we investigate correlations of the peak offset to other characteristics of the spectra. These include: Half the separation between red and blue peak (**m1**), the full width at half-maximum (FWHM) of the red peak (**m2**), twice the half width at half-maximum (HWHM) of the red peak towards the line centre (**m2a**) and twice the HWHM of the red peak away from the line centre (**m2b**). Hence, **m2** is the average of **m2a** and **m2b**. A visualization of these methods is shown in Fig. 11.

We choose to divide the second method (**m2**) into **m2a** and **m2b** as we expect different physical causes for their respective wing shape. The wing towards the line centre should be strongly influenced by the IGM attenuation. This has been shown to hold true even at low redshifts very close to the line centre (Laursen et al. 2011). However, the IGM is transparent for the wing away from the line centre and the red peak, and therefore should only be impacted by small-scale frequency diffusion. Indeed we will find large differences in the results from **m2a** and **m2b**.

We calculate slope f and offset $v_{\rm offset}$ for a linear regression of form

$$v_{\rm predict} = f \times v_{\rm proxy} + v_{\rm offset} \tag{35}$$

and also the corresponding Pearson correlation coefficient p. $v_{\rm proxy}$ denotes the respective velocities by **m1**, **m2**, **m2a**, and **m2b**. $v_{\rm predict}$ is the predicted correction for the corresponding peak position $v_{\rm RT}$. With such prediction, we can correct the former $v_{\rm RT}$ distribution as

$$v_{\rm RT,corr} = v_{\rm RT} - v_{\rm predict}. \tag{36}$$

When a blue peak is additionally available for **m1**, we use the detection algorithm introduced in Section A3: Peaks are identified as connected areas in $F_\lambda(\Delta\lambda)$ above such a threshold value that a

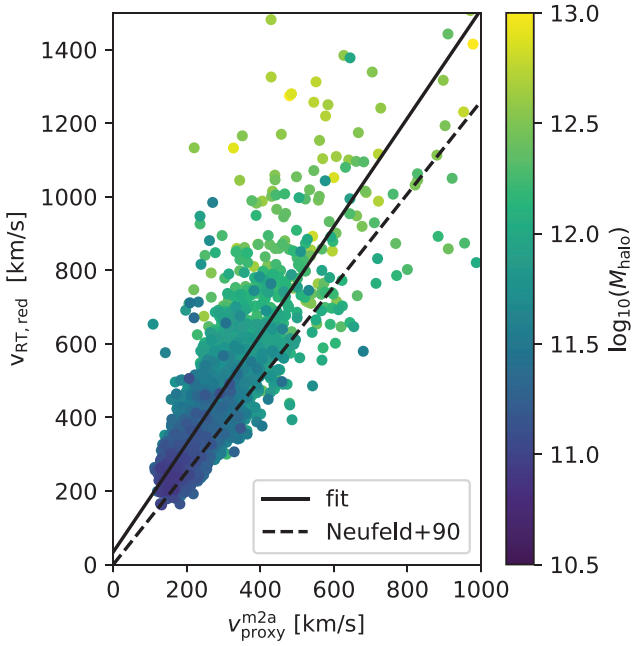

Figure 12. Scatter plot of $v_{\rm RT,red}$ and $v^{\rm m2a}_{\rm proxy}$ for emitters detected at number density threshold $n_{\rm LAE} = 10^{-2}\ h^3\,{\rm Mpc}^{-3}$ and redshift $z = 3.01$. We also show a linear fit (see equation 35), the slope expected from an optically thick spherical H I distribution (Neufeld 1990).

given number density $n_{\rm LAE}$ is reached. Additionally, we require that the maximal brightness of a peak needs to exceed 10 per cent of the maximal brightness of the brightest peak. If only one peak is available the emitter is excluded when computing m_1.

The linear regression in equation (36) is motivated by Neufeld's solution for which both peaks' FWHM (i.e. **m2**) and peak offset scale with $\tau^{1/3}$ for an optically thick spherical H I distribution. From Neufeld's derivation, a slope of 1.26 is derived for the relation between offset and FWHM (Neufeld 1990). We expect that anisotropies, dust, the velocity field and IGM interaction introduce a significant scatter as well as noticeable change to the slope parameter. Such a correlation has been found in observations (see e.g. Verhamme et al. 2018).

In general, we find **m1** to perform the best with a correlation coefficient of $p \gtrsim 0.95$ across the studied redshift range. However, most of observed LAEs are not doubly peaked, thus only methods **m2**, **m2a**, **m2b** are available. For these, we find **m2a** with $0.82 \lesssim p \lesssim 0.85$ to perform the best. **m2** performs slightly worse with $0.77 \lesssim p \lesssim 0.78$ and **m2b** significantly worse with $0.40 \lesssim p \lesssim 0.65$. Note that we restrict the emitter sample for the regression to those with $v_{\rm RT} < 800{\rm km\,s^{-1}}$ in order to allow comparisons to observational studies and hinder the most massive emitters to dominate the fit due to their increased scatter.

An example of the correlation between $v_{\rm RT}$ and $v_{\rm proxy}$ is shown in Fig. 12 for **m2a** at $z = 3.01$ and a number density threshold of 0.01 ${\rm Mpc}^{-3}\,{\rm h}^3$. The best fit in this case yields $v_{\rm predict} = 1.47 \cdot v^{\rm m2a}_{\rm proxy} + 37\ {\rm km\,s^{-1}}$ with a Pearson coefficient of 0.83. For other redshifts, we find a slope of 1.39, 1.43, 1.70 ($z = 2.0$, $z = 4.01$, $z = 5.85$) with similar or lower offsets. Note that the constant offset itself is irrelevant for the damping scale as it does not change the pairwise velocity distribution. We find that there is a strong dependence of the slope parameter on the chosen number density threshold. The slope parameter increases as the number density is decreased. This finding is related to the dependence of host halo mass: As shown in

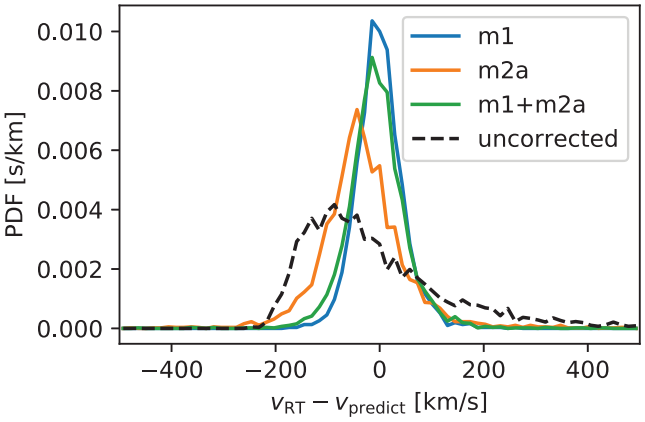

Figure 13. Radiative velocity distribution after different correction schemes at a number density threshold of $n_{LAE} = 10^{-2}\ h^3\,Mpc^{-3}$ and redshift $z = 3.01$. The mean velocity is subtracted from each distribution. The dashed line shows the uncorrected distribution. The dispersion for the three correction schemes reduces to 48, 92 , 65km s^{-1} (m1, m2b, m1 + m2a) compared to 189 km s^{-1} when not corrected.

Fig. 12, points with larger halo mass tend to have a slightly steeper slope with larger scatter.

For **m1**, we find that the slope is close to unity and in very good agreement with observations (Verhamme et al. 2018), while for methods using the peaks' width (**m2**, **m2a**, **m2b**) we find a slope above unity that makes it slightly higher than the Neufeld solution and considerably higher compared to observations and shell models (Zheng & Wallace 2014; Verhamme et al. 2018). In addition, we find a significant dependence on selection criteria and redshift, possibly explaining some of the discrepancies between our results and the literature apart from the mentioned modelling shortcomings in our simulations.

Similarly to Verhamme et al. (2018), we can correct for the systemic redshift offset by subtracting the FWHM. Doing so consistently reduces the velocity dispersion by a factor of 2–3 across the simulated redshift range and accordingly shortens the damping scale. Fig. 13 shows the corrected distribution according to equation (36) at the same redshift and number density threshold. Also the third and fourth central moments are reduced significantly: the normalized moments (skewness, kurtosis) shrink such that a Gaussian fit becomes feasible.

The fit parameters and correlation coefficient calculated here can be used as a first guess for applying the presented corrections to observations. However, variations from those in observational datasets are expected due to the modelling deficiencies discussed earlier. By measuring a second emission line for a subset of LAEs, the RT velocity offsets v_{RT} can be calculated, allowing a similar fit of different correction methods in observations that can then be applied to the entire cosmological volume of LAEs studied.

6 CONCLUSIONS

In this paper, we have studied the clustering of LAEs in redshift space using a full RT simulation on Illustris. We find a new kind of the RSD effect due to RT, and our executive summary is the following:

(i) The additional RSD stems from small-scale frequency diffusion of the Lyman α line leading to a shift of the spectral peaks. The peak shifts can be larger than those of the peculiar velocity field and thus impact the redshift space clustering signal on larger spatial scales.

(ii) We show that the peak shifts from Lyman α RT damp the power spectrum along the LOS at scales of $k_\parallel \gtrsim 0.1\,h\,Mpc^{-1}$. We also show that the shifts are mostly independent of the local density and velocity field such that an independent modelling of this shift's impact can be done. This is similar to that of the FoG effect due to random motion of galaxies, i.e. in terms of the one-point velocity PDF. However, the functional form of the damping can be more complex and even involve oscillations due to the double peak nature.

(iii) The strength of the damping depends strongly on the chosen localization method of the LAEs in its spectrum. We attempt two extreme scenarios where we find a peak from the entire spectrum (global peak) and only from the spectrum at red end (red peak).

(iv) We show that we can mitigate the impact of the distortion by applying a correction scheme of which we present two classes: In the case of double-peaked spectra, the mid-point between the two peaks is an excellent proxy for the emitter position. If only a single peak is present, the HWHM on the wing towards the Lyman α line centre can be used as a mediocre proxy.

We do not attempt to quantify the exact amplitude of the RT effect on actual BAO and RSD galaxy surveys such as HETDEX for the following reasons. First, as we often mention, our simulated LAEs do not well reproduce a variety of observables such as the luminosity function likely due to unresolved ISM physics in Illustris. Secondly, the actual impact should depend on resolving power of a spectrum. For example, HETDEX has a spectral resolution roughly corresponding to $\Delta v \sim 400\,km\,s^{-1}$ where the resultant PDF is largely smeared out. Nevertheless, we stress that the frequency shift can be larger than $\Delta v \sim 500\,km\,s^{-1}$ as seen in Fig. 4 and hence we expect that the FoG damping due to RT exists to some extent. We leave detailed assessment for future work.

Although we focus on the clustering of LAEs, it would be straightforward to extend our analysis to the intensity mapping. In fact, we visually confirm strongly elongated feature in the Lyman α intensity map in Fig. 1: For intensity mapping we expect a similar damping to that of the LAEs from their positional offset, but additionally from the width of the spectrum itself which further strengthens the damping. Observationally, Croft et al. (2016) reported the large-scale elongation along the LOS in the cross-correlation between the quasars and the Lyman α intensity map at $z \sim 2$ in the Sloan Digital Sky Survey. Even though their more recent study argues that it is due to special environment around quasars given the lack of the cross-correlation signal between the intensity map and the Lyman α forest (Croft et al. 2018), the elongation might partly be due to the RT FoG effect.

As a concluding remark, we give the following general suggestions for cosmological LAE surveys:

(i) Having a second emission line as tracer of the LAEs' RT velocity shift for a subset of emitters allows to calibrate the linear fit parameters of the correction methods based on the Lyman α spectral features presented in Section 5.3 to mitigate the distortion effect.

(ii) Furthermore, we found an empirical relationship between the second central moment for the red and blue peaks in stacked spectra in the halo's rest frame and the dispersion of those peaks in the one-point velocity distribution. Hence, the possible damping might be estimated by measuring the second central moments in the stacked spectra. Stacking these in the halo's rest frame also requires knowledge of a secondary emission line tracing the kinematics of the LAE.

Finally, concerning other target emission lines such as Hα for e.g. Euclid (Laureijs et al. 2011) and Wide-Field InfraRed Survey Telescope-Astrophysics Focused Telescope Assets (WFIRST-AFTA; Spergel et al. 2013) and [O II] for the Subaru Prime Focus Spectrograph (PFS; Takada et al. 2014) and Dark Energy Spectroscopic Instrument (DESI, DESI Collaboration et al. 2016), it is unlikely that RT makes a notable impact on the BAO and RSD measurements, as the transition must be resonant and have high optical depths in the astrophysical environment.

ACKNOWLEDGEMENTS

We thank Max Gronke for useful discussions concerning the obtained spectra in our simulations. We are also grateful to Eiichiro Komatsu for his comments on the manuscript. This work was supported in part by Japan Society for the Promotion of Science (JSPS) KAKENHI Grant Number JP16J01890, and by World Premier International Research Center Initiative (WPI Initiative), MEXT, Japan. We acknowledge use of the Python programming language (Van Rossum & de Boer 1991), the use of the NUMPY (van der Walt, Colbert & Varoquaux 2011), IPYTHON (Perez & Granger 2007), and MATPLOTLIB (Hunter 2007) packages. Our analysis benefited from the use of the following astrophysical packages: ASTROPY (Astropy Collaboration 2013) and HALOTOOLS (Hearin et al. 2016).

REFERENCES

Adams T. F., 1972, ApJ, 174, 439
Adams J. J. et al., 2011, ApJS, 192, 5
Agrawal A., Makiya R., Chiang C.-T., Jeong D., Saito S., Komatsu E., 2017, J. Cosmol. Astropart. Phys., 2017, 003
Ahn S.-H., Lee H.-W., Lee H. M., 2003, MNRAS, 340, 863
Astropy Collaboration, 2013, A&A, 558, A33
Bacon R. et al., 2015, A&A, 575, A75
Ballinger W. E., Peacock J. A., Heavens A. F., 1996, MNRAS, 282, 877
Behrens C., Niemeyer J., 2013, A&A, 556, A5
Behrens C., Byrohl C., Saito S., Niemeyer J. C., 2018, A&A, 614, A31
Beutler F. et al., 2014, MNRAS, 443, 1065
Bianchi D., Percival W. J., Bel J., 2016, MNRAS, 463, 3783
Bonilha J. R. M., Ferch R., Salpeter E. E., Slater G., Noerdlinger P. D., 1979, ApJ, 233, 649
Croft R. A. C. et al., 2016, MNRAS, 457, 3541
Croft R. A. C., Miralda-Escudé J., Zheng Z., Blomqvist M., Pieri M., 2018, MNRAS, 481, 1320
Davis M., Peebles P. J. E., 1983, ApJ, 267, 465
DESI Collaboration, 2016, preprint (arXiv:1611.00036)
Diemer B. et al., 2019, MNRAS, 487, 1529
Dijkstra M., 2017, preprint (arXiv:1704.03416)
Fisher K. B., 1995, ApJ, 448, 494
Furlanetto S. R., Schaye J., Springel V., Hernquist L., 2005, ApJ, 622, 7
Gronke M., 2017, A&A, 608, A139
Gronke M., Dijkstra M., McCourt M., Oh S. P., 2016, ApJ, 833, L26
Gronke M., Dijkstra M., McCourt M., Peng Oh S., 2017, A&A, 607, A71
Gurung-Lopez S., Orsi A. A., Bonoli S., Padilla N., Lacey C. G., Baugh C. M., 2019a, preprint (arXiv:1904.04274)
Gurung-López S., Orsi A. A., Bonoli S., Baugh C. M., Lacey C. G., 2019b, MNRAS, 486, 1882
Hamilton A. J. S., 1998, in Hamilton D., ed., Astrophysics and Space Science Library Vol. 231, The Evolving Universe. Springer, Dordrecht, p. 185
Harrington J. P., 1973, MNRAS, 162, 43
Hearin A. et al., 2017, AJ, 154, 190
Herenz E. C. et al., 2017, A&A, 606, A12
Hikage C., Yamamoto K., 2016, MNRAS, 455, L77
Hill G. J. et al., 2008, in Kodama T., Yamada T., Aoki K., eds, ASP Conf. Ser. Vol. 399, Panoramic Views of Galaxy Formation and Evolution. Astron. Soc. Pac., San Francisco, p. 115
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Inoue A. K. et al., 2018, PASJ, 70, 55
Jackson J. C., 1972, MNRAS, 156, 1P
Kaiser N., 1987, MNRAS, 227, 1
Kakiichi K., Dijkstra M., Ciardi B., Graziani L., 2016, MNRAS, 463, 4019
Kulas K. R., Shapley A. E., Kollmeier J. A., Zheng Z., Steidel C. C., Hainline K. N., 2012, ApJ, 745, 33
Landy S. D., Szalay A. S., 1993, ApJ, 412, 64
Laureijs R. et al., 2011, preprint (arXiv:1110.3193)
Laursen P., Sommer-Larsen J., Andersen A. C., 2009, ApJ, 704, 1640
Laursen P., Sommer-Larsen J., Razoumov A. O., 2011, ApJ, 728, 52
Leung A. S. et al., 2017, ApJ, 843, 130
Matsubara T., 2014, Phys. Rev. D, 90, 043537
Nelson D. et al., 2015, Astron. Comput., 13, 12
Neufeld D. A., 1990, ApJ, 350, 216
Okumura T., Seljak U., Desjacques V., 2012, J. Cosmol. Astropart. Phys., 2012, 014
Perez F., Granger B. E., 2007, Comput. Sci. Eng., 9, 21
Reid B. A., White M., 2011, MNRAS, 417, 1913
Scoccimarro R., 2004, Phys. Rev. D, 70, 83007
Sefusatti E., Crocce M., Scoccimarro R., Couchman H. M. P., 2016, MNRAS, 460, 3624
Smith A., Ma X., Bromm V., Finkelstein S. L., Hopkins P. F., Faucher-Giguère C.-A., Kereš D., 2019, MNRAS, 484, 39
Spergel D. et al., 2013, preprint (arXiv:1305.5422)
Springel V., Hernquist L., 2003, MNRAS, 339, 289
Takada M. et al., 2014, PASJ, 66, R1–1
Taruya A., Nishimichi T., Saito S., 2010, Phys. Rev. D, 82, 63522
Uhlemann C., Kopp M., Haugg T., 2015, Phys. Rev. D, 92, 063004
van der Walt S., Colbert S. C., Varoquaux G., 2011, Comput. Sci. Eng., 13, 22
Van Rossum G., de Boer J., 1991, CWI Q., 4, 283
Verhamme A., Schaerer D., Atek H., Tapken C., 2008, A&A, 491, 89
Verhamme A. et al., 2018, MNRAS, 478, L60
Vlah Z., White M., 2019, J. Cosmol. Astropart. Phys., 2019, 007
Vogelsberger M. et al., 2014, Nature, 509, 177
White M., Song Y.-S., Percival W. J., 2009, MNRAS, 397, 1348
Zheng Y., Song Y.-S., 2016, J. Cosmol. Astropart. Phys., 2016, 050
Zheng Y., Song Y.-S., Oh M., 2019, J. Cosmol. Astropart. Phys., 2019, 13
Zheng Z., Wallace J., 2014, ApJ, 794, 116
Zheng Z., Cen R., Trac H., Miralda-Escude J., 2010, ApJ, 716, 574
Zheng Z., Cen R., Trac H., Miralda-Escude J., 2011, ApJ, 726, 31

APPENDIX A: DETECTION ALGORITHM VARIATIONS

A1 Spectral resolution

In our study, we do not generally assume a spatial resolution natch to a specific observation both in the spectral and the angular resolutions. For the spectral resolution, we infer the redshift space position by adding the line shift to the real space position known from the halo catalogues, while surveys directly infer the degenerate redshift space position from the line feature's position. Nevertheless, the spectral resolution in simulations and observations should be comparable quantities. In our fiducial case, velocities are resolved to 24.7 km s^{-1} corresponding to a spectral resolution of roughly $R \sim 12\,000$. This exceeds the resolution in HETDEX by more than an order of magnitude ($R \sim 800$, see e.g. Hill et al. 2008). Even dedicated spectroscopic instruments in the search of LAEs such as MUSE only reach values up to $R \sim 3000$ (see e.g. Bacon et al. 2015).

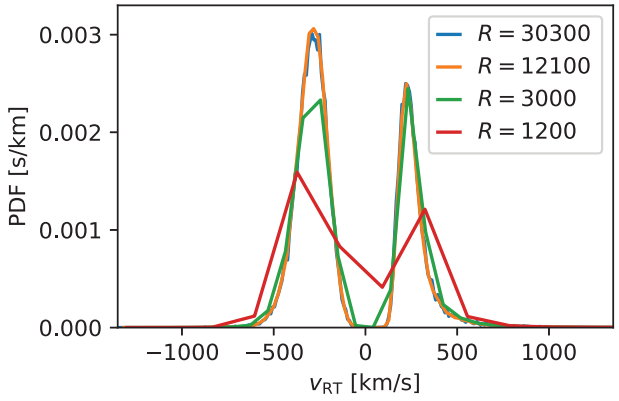

Figure A1. The radiative velocity distribution v_{RT} for varying spectral resolution in terms of $R = c/\Delta v$ at $z = 3.01$ for all LAEs (no n_{LAE} restriction).

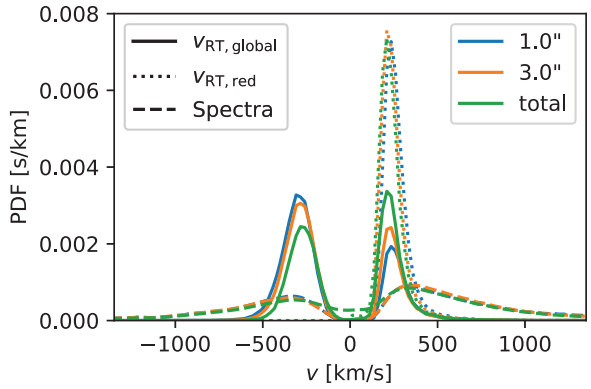

Figure A2. The radiative velocity distributions v_{RT} and stacked spectra for varying aperture sizes at $z = 3.01$ for all LAEs (no n_{LAE} restriction).

In Fig. A1 we vary the spectral resolution for the fiducial case at $z = 3.01$ and find good reproduction of the characteristics of the distribution: If existent, the double-peaked structure is conserved and the second central moment of individual peaks vary by less than 10 per cent. We have adopted the fiducial resolution of 24.7 km s^{-1} as we are interested in the physical impact of RT in theory, but simultaneously want to assure sufficient signal to noise in each bin.

A2 Aperture size

In Fig. A2, we show the aperture size dependence of the v_{RT}–probability distribution and of the stacked profiles at $z = 3.01$. The sample includes all simulated LAE rather than fixing the count to a number density threshold. This reduces the noise, while we find qualitatively similar results for a restricted sample.

Three qualitative changes occur when increasing the aperture: A larger aperture appears to favour the red peaks over the blue, the red peak slightly shifts towards the line centre and for large apertures the stacked spectra show emission in the otherwise deserted trough in the line centre. The second central moments and maxima of the red peaks change by less than 10 per cent, so that the expected change in clustering signal should be similarly small.

A3 Refined detection algorithm

In the detection algorithm applied so far, line-of-sight halo positions are inferred from the spectral peak, while the detectability is determined from the overall flux. Therefore, the same halo sample is considered as in prior real space evaluation performed in Behrens et al. (2018). However, in redshift space some emitters' spectra might significantly diffuse while others do not based on varying small and large-scale properties. We therefore expect additional distortions when emitters are detected by the specific flux. Primarily the bias is expected to change.

In order to capture the additional distortions, we refined our detection algorithm as follows: a specific flux threshold t_{flux} is imposed and only peaks above this threshold are used to identify emitters. Once a peak (red or global) is identified as the emitter's position, the surface brightness is integrated around the peak for the spectral range which reaches t_{flux}.

Fig. A3 shows an example of the refined algorithm for two emitters at $z = 3.01$ and how the bias might change: While the spectrum of HID 2262 has the lower overall flux than HID 2250, it has a larger flux around the detected peak. Therefore, flux-sorted

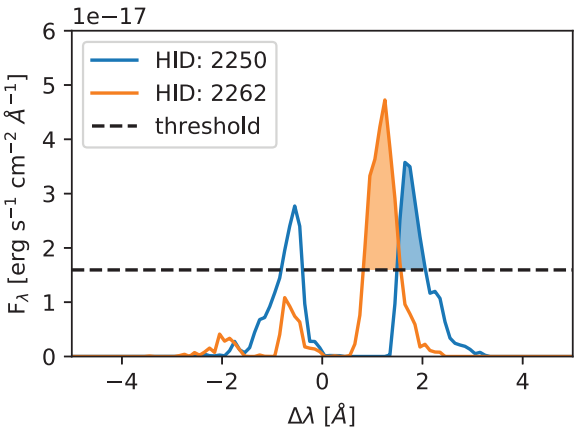

Figure A3. Example of refined detection of two individual emitters with given halo ID at $z = 3.01$. Dashed line corresponds to the specific flux threshold of 1.7 erg s^{-1} cm^{-2} Å^{-1} (for number density 0.01 h^3 Mpc^{-3}) and the shaded region the flux associated with the emitters given the refined detection algorithm. For shown emitters, the emitter with the highest total flux (which includes all peaks) is different from the emitter with the largest flux for the identified peak.

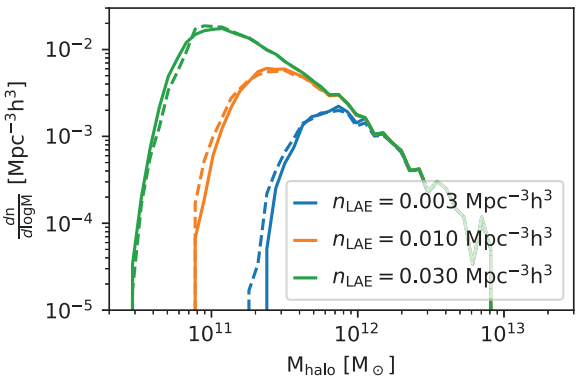

Figure A4. The halo mass function of detected LAEs at $z = 3.01$. Solid lines represent the fiducial detection method and dashed lines the refined algorithm presented. There are only marginal changes in the function at the characteristic mass cut-off imposed by setting the number density threshold n_{LAE}.

and number density limited samples will differ when sorted by total flux or flux associated with the peak.

Fig. A4 shows the halo mass function of detected LAEs for the two discussed detection algorithms. Overall, changes only occur close to a characteristic mass cut-off that is implied by the number density threshold and due to the strong correlation between observed flux and halo mass. As we only find a small change in the mass distribution of detected LAEs, we do not expect an additional change in bias/isotropic distortion from RT in redshift space. We explicitly checked this by calculating the corresponding velocity distributions and power spectra.

Nevertheless, this detection method is useful in classifying individual peaks that we will use for identifying blue and red peaks in Section 5.3 concerning the correction of the localization error leading up to the damping.

APPENDIX B: COMPARISON TO PRIOR STUDIES

While similar RT simulations have previously been performed, the shown redshift space damping has not been observed by Zheng et al. (2010)/Zheng et al. (2011). For consistency, we try to reconcile our findings with these prior simulations.

As shown, the RT RSDs significantly increase at lower redshift: At $z = 2.00$, we detect a central second moment σ of the $v_{\mathrm{RT,red}}$ distribution of 223 km s^{-1}, but only a moment of 76 km s^{-1} at $z = 5.85$. Furthermore, the significance for the redshift space clustering is reduced at higher redshifts due to the higher Hubble flow as discussed before.

Additionally, as concluded by Behrens et al. (2018), the hydrodynamic resolution has a large impact on the RT's results. In particular, the hydrodynamic resolution largely affects the photon diffusion in configuration and frequency space determining whether a selection effect is detected in mock observations or not. Similarly, one might ask whether the hydrodynamic resolution impacts the RT RSD presented here. We answer this question by comparing the results by analysing a set of RT simulations at different resolutions.

Fig. B1 shows the radiative velocity offset distribution as a function of mass at redshift $z = 5.85$ of different underlying hydrodynamic resolution for the RT post-processing. This plot can directly be compared to the results found in Zheng et al. (2010) showing a very similar plot (see fig. 6 there). Given the difference in the baryon modelling following detection method, we find a very good match between our low-resolution run with a grid spacing of $\Delta = 30.4$ pkpc and the results found by Zheng et al. (2010) across all mass bins. Both shown simulations have been imported from Behrens et al. (2018) and thus only have an initial photon count of 100 instead of 1000. Thus, minimal differences to other shown results might be expected.

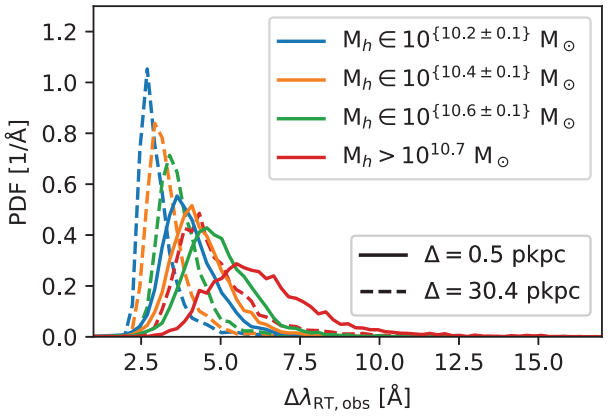

Figure B1. The radiative velocity distribution for varying halo mass range and underlying hydrodynamic resolution at redshift $z = 5.85$. $\Delta\lambda_{\mathrm{RT}}$ is the radiative velocity as observed wavelength shift at $z = 0$ with the peculiar velocity v_{pec} having been subtracted as usual.

We find a significant impact of the hydrodynamic resolution on the velocity distribution. First, the peak of the distribution is shifted towards larger spectral offsets from the line centre at higher resolutions. If no blue peak is present, as is for $z = 5.85$ due to IGM interaction, this overall shift of the distribution should not affect the damping signal. Secondly, the distributions broaden at higher resolutions and thus increases the second central moment, which enlarges the damping length scale. For the shown resolutions, the second central moment increases by roughly 20 per cent across all mass bins at higher resolution.

APPENDIX C: DISCREPANCIES IN THE DAMPING FACTOR MODELLING

As we have shown in Section 4.5, we have made an attempt to understand the FoG suppression due to RT in terms of the one-point velocity PDF. In this appendix, however, we show a hint implying that it is not sufficient to know the one-point PDF to fully model the impact of RT on the redshift-space clustering.

In Fig. C1, we present two ratios of the power spectra. The blue lines are the ratios of the total redshift-space spectra to ones only with the peculiar velocity component, i.e. $P^s_{\mathrm{g,tot}}/P^s_{\mathrm{g,pec}}$, which are the same results as we showed in Fig. 10. We also present the ratios of the redshift-space spectra only with the RT velocity component to the real-space power spectra, i.e. $P^s_{\mathrm{g,RT}}/P^r_{\mathrm{g}}$ as red lines. We did not address the latter ratio in the main text, since both the denominator and the numerator are not directly accessible in real observations. In Fig. 10, it is apparent that the red curves are systematically higher than the blue ones even though both of them are noisy.

As we stressed in deriving the last equality of equation (29), a necessary condition which makes the ratio, $P^s_{\mathrm{g,tot}}/P^s_{\mathrm{g,pec}}$, equivalent to the FT of the one-point velocity PDF is following:

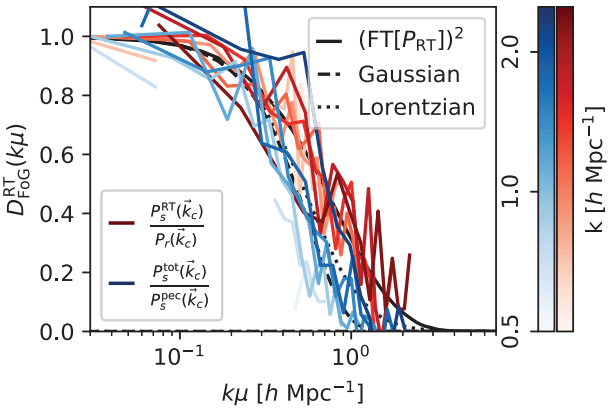

Figure C1. Damping factor as a function of the line-of-sight frequency $k_{\parallel} = k\mu$ due to Lyman α RT for $z = 3.01$. Emitter positions are assumed to coincide with the red peak only. Two different ratios P^{RT}_s/P_r and $P^{\mathrm{tot}}_s/P^{\mathrm{pec}}_s$ are shown for the damping factor, see text for an explanation. Different coloured lines represent measurements of $D(k_{\parallel})$ from the mock catalogues at different total frequencies k.

(i) $\langle v_{\rm pec}(x) v_{\rm RT}(x^{'})\rangle = 0$,
(ii) $\langle v_{\rm RT}(x) \delta_{\rm g}(x^{'})\rangle = 0$,
(iii) and $\langle v_{\rm RT}(x) v_{\rm RT}(x^{'})\rangle = 0$.

We argued from Fig. 4 that the condition (i) is satisfied at the level of one point, i.e. at $x = x^{'}$. Similarly, we did not see any strong evidence which violates the conditions (ii) and (iii) in Fig. 8. In contrast, the condition (i) is not necessary for $P^{s}_{\rm g,RT}/P^{r}_{\rm g}$, since it does not involve $v_{\rm pec}$. It is encouraging to see that the red lines are more consistent than the blue ones with the direct FT of the one-point PDF (black solid line). We find this trend holds at other redshift snapshots for both the global-peak and the red-only cases. This fact implies that the conditions (ii) and (iii) are basically satisfied. We thus may attribute the discrepancy between the red and blue lines to the residual correlation between $v_{\rm pec}$ and $v_{\rm RT}$, but a conclusive statement with more careful investigation is left for future work.

This paper has been typeset from a TEX/LATEX file prepared by the author.