

# Understanding Redshift-Space Distortions: How the Pairs Move

Joseph Kuruvilla<sup>1\*</sup>† and Cristiano Porciani<sup>1</sup>

<sup>1</sup>Argelander-Institut für Astronomie, University of Bonn, Auf dem Hügel 71, D-53121 Bonn, Germany

Accepted XXX. Received YYY; in original form ZZZ

## ABSTRACT

Streaming model ... More abstract

**Key words:** keyword1 – keyword2 – keyword3

## 1 INTRODUCTION

Redshift-space distortions – growth rate – Aim of the paper – Section 2 – Section 3

## 2 N-BODY SIMULATIONS

The subsequent work in this paper was done fully using simulations. For this purpose, we use a dark matter (DM) only simulation from Pillepich 08. It was run using a lean version of *GADGET-2* (cite springel). The simulation box with a length scale of  $1200 h^{-1}\text{Mpc}$  had  $1024^3$  particles. Each DM particle in turn had a mass of  $1.246 \times 10^{11} h^{-1}\text{Mpc}$ . The structure formation was evolved using the parameters from WMAP5 cosmology (cite komatsu).

## 3 STREAMING EQUATION

Peculiar velocity distorts the true isotropic spatial distribution. Anisotropy in the spatial distribution introduced by the line-of-sight component of the peculiar velocity can be quantified in the configuration space by the 2-point correlation function. Fig 1 shows the isotropic nature of the correlation function in the real-space and the anisotropy introduced by the peculiar velocity in the redshift-space. The interesting objective now is knowing the real-space isotropic correlation function, how can we obtain the anisotropic redshift-space correlation function. One of the ways to model this is to use the streaming equation.

The streaming equation (Peebles 1993; Fisher 1995; Scoccimarro 2004) quantifies the anisotropic correlation function as a convolution of the true isotropic correlation



**Figure 1.** Correlation function in real-space and redshift-space.

function with the distribution of relative line-of-sight velocities. In the context of the distant observer approximation, it is written as

$$1 + \xi_s(s_{\parallel}, s_{\perp}) = \int_{-\infty}^{\infty} [1 + \xi_R(r)] P_{v_{\text{los}}} (s_{\parallel} - r_{\parallel} | r_{\parallel}, r_{\perp}) dr_{\parallel}, \quad (1)$$

where  $r = \sqrt{r_{\parallel}^2 + r_{\perp}^2}$  and  $v_{\text{los}}$  is the relative line-of-sight velocity. We set the line-of-sight axis along the  $z$ -axis for the remainder of this paper. The pair weighted quantity,  $v_{\text{los}}$ , can then be defined as

$$v_{\text{los}} = (\mathbf{v}_2 - \mathbf{v}_1) \cdot \hat{\mathbf{r}}_{\text{los}} = v_{\parallel} \text{sgn}(r_{\parallel}), \quad (2)$$

where

$$v_{\parallel} = v_{z_2} - v_{z_1}, \quad (3)$$

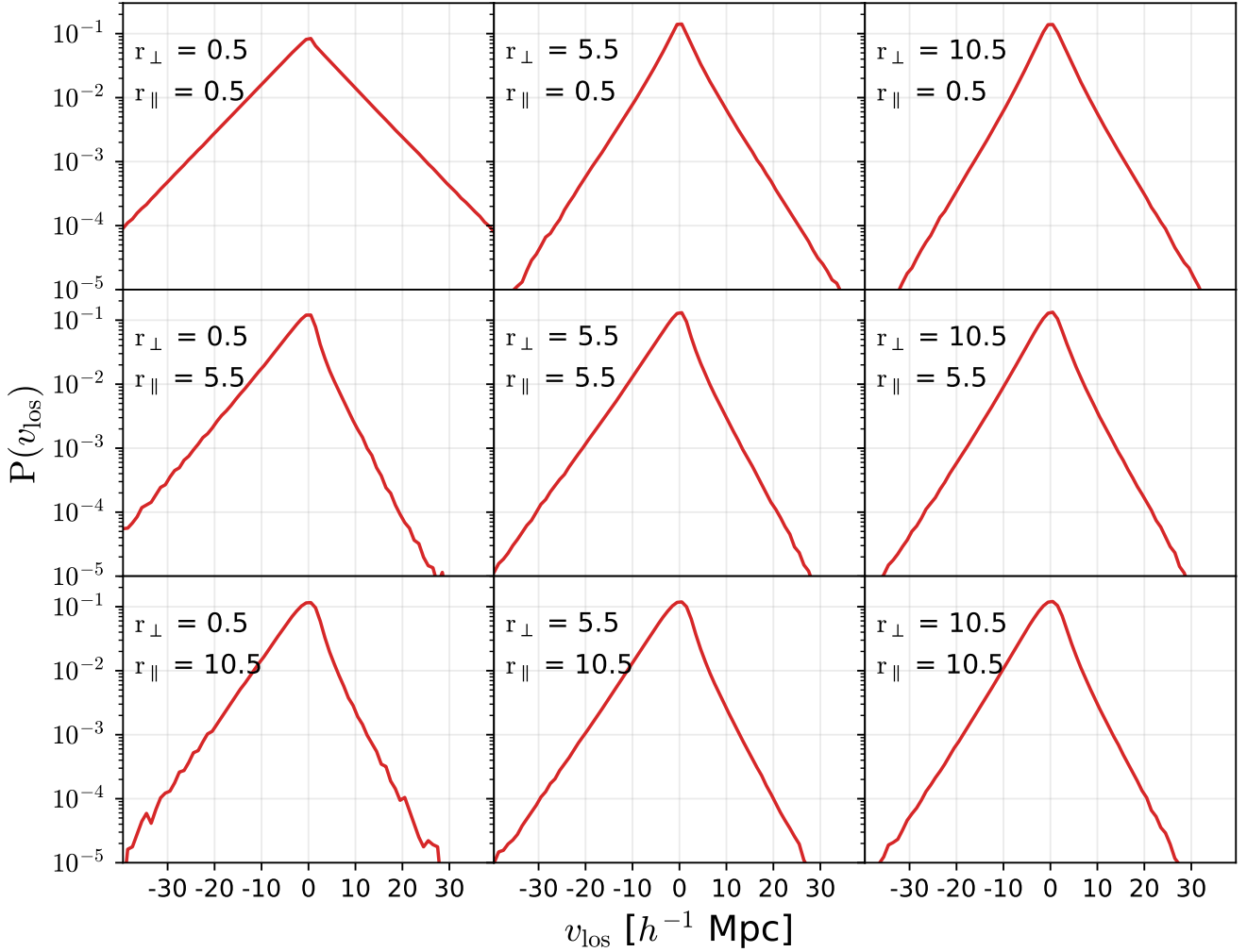
$$r_{\parallel} = r_{z_2} - r_{z_1}. \quad (4)$$

### 3.1 Ingredients

In this section, we discuss about the three ingredients (i.e. the 2D anisotropic redshift-space correlation function, 1D

\* E-mail: joseph.k@uni-bonn.de

† Member of the International Max Planck Research School (IMPRS) for Astronomy and Astrophysics at the Universities of Bonn and Cologne



**Figure 2.** Relative line-of-sight velocity distribution. The distance mentioned in each subplot refers to the mean value of each bin, having a bin width of  $1 h^{-1}$  Mpc.

isotropic real-space correlation function and the relative line-of-sight velocity distribution) entering the streaming equation. Firstly we turn our focus to the 1D real-space correlation function  $\xi_R(r)$ . The correlation functions were computed using the Landy-Szalay (LS) estimator (Landy & Szalay 1993), which is

$$\xi(r) = \frac{DD(r) - 2DR(r) + RR(r)}{RR(r)}, \quad (5)$$

where  $DD(r)$  is the normalised count of galaxy pairs,  $DR(r)$  gives the normalised count of pairs given a cross correlation between pairs and random points drawn from a random distribution and  $RR(r)$  is the normalised count of random pairs. In order to mimic the redshift-space distortion effect in the simulation, we displace the position of the particles according to

$$\mathbf{s} = \mathbf{x} + \frac{v_z}{aH(a)}\hat{\mathbf{z}}, \quad (6)$$

where  $a$  is the scale factor and  $H(a)$  is the Hubble factor. The “true” 2D redshift-space correlation function is then computed again using LS estimator on these displaced particles.

The relative line-of-sight velocity was measured using Eq. 2. For this purpose, we developed a Cython code<sup>1</sup> which is fully parallelised using MPI. The distribution was calculated from randomly down sampled selection of  $256^3$  particles. The total number of pairs we obtained from this sample is  $5.1 \times 10^{11}$  at distances from 0.5 to  $100.5 h^{-1}$  Mpc between the pairs. For the completeness, we have included the periodic boundary conditions while measuring the distance between the pairs. The convergence property of the distribution has been tested with a higher particle number and discussed further in App. A. We showcase the relative line-of-sight velocity distribution computed for  $256^3$  particles in Fig. 2. One of the immediate conclusion we can draw upon visual inspection is that the distribution changes starkly with the distance and neither an exponential nor a Gaussian will be able to model the distribution accurately.

To further understand the distribution we try to characterise the distribution using moments. We specifically look

<sup>1</sup> <https://github.com/jkuruvilla>

into the the first raw moment (mean), the second central moment (variance), the third and the fourth standardised moments (skewness and kurtosis). Fig. 3 shows these moments of the relative line-of-sight velocity distribution.. The distribution is marked by non-zero skewness, specifically it is negatively skewed at all the scales shown. Due to the negative skewness, it can be seen that the mean velocity is negative at most scales. We can interpret this as it is more probable to have in-falling pairs at these scales (Scoccimarro 2004). The distribution showcases a leptokurtic behaviour, implying it has a fatter tail than a normal distribution which has a skewness value of 3.

Having all the ingredients needed, we compute the integral of the streaming equation (Eq. 1) and compare it with the “true” redshift-space correlation function measured from the simulations. This is shown in Fig. 4. It seems that when one checks Eq. 1 with simulations there seems to be some discrepancy even though it has been quoted to be exact. The deviation from exactness is worrying as all the phenomenological streaming models are build upon the streaming equation. To understand the discrepancy, we go back to the roots from which we derived the streaming in the next section and try to understand why this deviation from exactness arises.

#### 4 CORRECTIONS TO STREAMING EQUATION

Correlation function in redshift space can be defined as the probability to find a pair at parallel and perpendicular to line-of-sight distance ( $s_{\parallel}$  and  $s_{\perp}$ ), which is given as

$$dP = 2\pi s_{\perp} n \left[ 1 + \xi_s(s_{\parallel}, s_{\perp}) \right] ds_{\parallel} ds_{\perp}, \quad (7)$$

where  $\xi_s(s_{\parallel}, s_{\perp})$  is the 2D anisotropic redshift-space correlation function. This could also written in terms of real space quantities as

$$dP = 2\pi s_{\perp} n \left[ 1 + \xi_R(r) \right] P_{v_{\text{los}}} \left( v_{\text{los}} \mid r_{\parallel}, r_{\perp} \right) \delta_D \left( (s_{\parallel} - r_{\parallel}) \text{sgn}(r_{\parallel}) - v_{\text{los}} \right) dr_{\parallel} ds_{\perp} dv_{\text{los}}. \quad (8)$$

where  $\xi_R(r)$  is the isotropic real-space correlation function and  $P_{v_{\text{los}}} \left( v_{\text{los}} \mid r_{\parallel}, r_{\perp} \right)$  is the relative line-of-sight velocity distribution function. The term inside the Dirac-delta ensures  $s_{\parallel} = r_{\parallel} + v_{\parallel}$ , which shows how the pairs are displaced.

Comparing Equations (7) and (8), and integrating over  $v_{\text{los}}$  and  $r_{\parallel}$  eliminates the delta function, one obtains

$$1 + \xi_s(s_{\parallel}, s_{\perp}) = \int_{-\infty}^{\infty} \left[ 1 + \xi_R(r) \right] P_{v_{\text{los}}} \left( (s_{\parallel} - r_{\parallel}) \text{sgn}(r_{\parallel}) \mid r_{\parallel}, r_{\perp} \right) dr_{\parallel}. \quad (9)$$

This is the streaming equation, which is exact in the distant-observer approximation. It is to be noted that Eq. 9 is different to the usual streaming equation as seen in literature till now. We can then proceed to break the integral in the

following manner

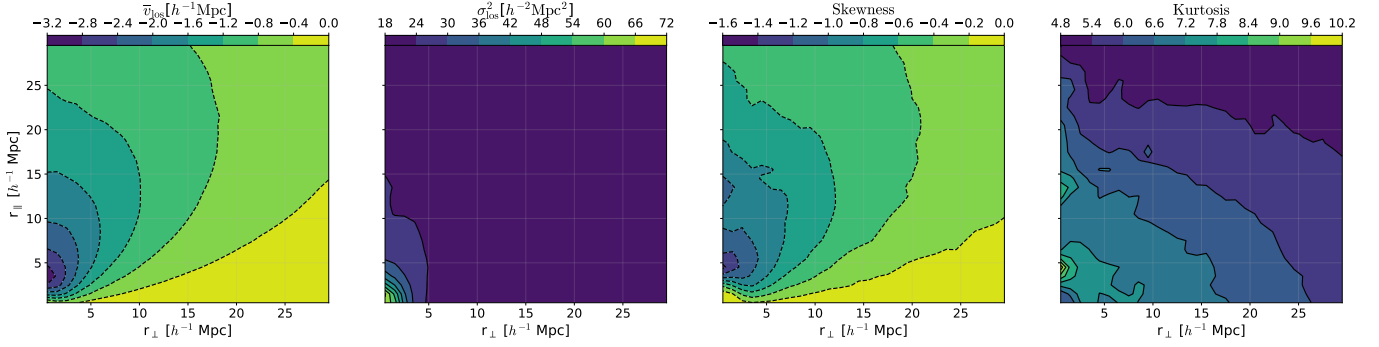
$$1 + \xi_s(s_{\parallel}, s_{\perp}) = \int_{-\infty}^0 \left[ 1 + \xi_R(r) \right] P_{v_{\text{los}}} \left( r_{\parallel} - s_{\parallel} \mid r_{\parallel}, r_{\perp} \right) dr_{\parallel} + \int_0^{\infty} \left[ 1 + \xi_R(r) \right] P_{v_{\text{los}}} \left( s_{\parallel} - r_{\parallel} \mid r_{\parallel}, r_{\perp} \right) dr_{\parallel}. \quad (10)$$

Fig. 5 shows the corrected streaming equation (Eq. 10) compared with the simulation.

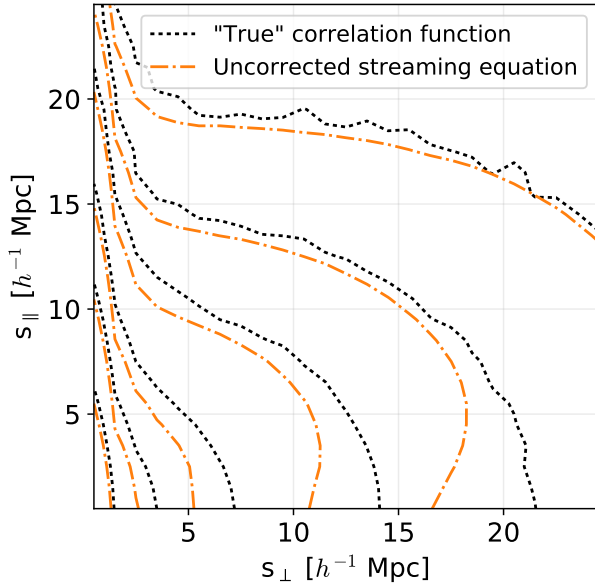
It is interesting to note that the integral runs from  $-\infty$  to  $+\infty$ . So what does the negative  $r_{\parallel}$  mean in the context of streaming model? It means that if we are to take the redshift space to be our reference frame, the negative  $r_{\parallel}$  implies pairs which have flipped their position compared to their redshift space counterparts. Fig. 8 clearly demonstrates this where the negative  $r_{\parallel}$  is the distance between the pairs which have flipped their position in real-space as compared to the position in redshift-space. This is due to the large relative in-fall velocity of the pairs. It should be noted again that the redshift-space distortions is purely a line-of-sight effect, hence the perpendicular distance between the pairs remains unchanged.

One can clearly see that with the complete knowledge of the relative line-of-sight velocity distribution and with the correction to account for the pairs which have flipped, we are able to reproduce the true redshift space correlation function accurately. To understand bit more on why there is a need to account for the pair flips and as a consistency check measure, we look at Fig. 6. The figure looks at a particular length scale in redshift-space, in this case  $s_{\parallel} = 10.5 h^{-1} \text{Mpc}$  and  $s_{\perp} = 1.5 h^{-1} \text{Mpc}$ . It is quite interesting to note that the plot features two distinct distributions. The first distribution which peaks around  $r_{\parallel} = 0$  corresponds to pairs which are in a high density region. These are the pairs which have a large relative velocity and amongst these some undergo pair flips ( $r_{\parallel} < 0$ ). So we can understand that for pairs in a high density region, it is equally likely to undergo pair flips as compared to pairs that do not. The second distribution in the plot peaks around  $s_{\parallel} = r_{\parallel}$ , these are mainly pairs which are moving away from each other according to Hubble flow (since  $v_{\parallel} = 0$ ). In a sense, by looking at the redshift-space distortion effect like this, we are able to disentangle the linear and non-linear effects. To be concise, the distribution around  $r_{\parallel} = 0$  is what causes the FoG effect while the distribution around  $s_{\parallel} = r_{\parallel}$  causes the Kaiser effect. The bottom panel of Fig. 6, shows how the integrand of the streaming equation looks like for a particular length scale in redshift-space. This panel acts as a consistency check and we can understand in a clearer way why the streaming equation would undermine the correlation function if the pair flips are unaccounted for. Due to the asymmetric nature of the velocity distribution, it is essential to pick the argument of the velocity distribution with the right parity. Otherwise we end up neglecting the flipped pairs.

We do the same exercise for a larger scale in Fig 7 and see that both formulations are equivalent since at large scales, we do not expect to see FoG effect and subsequently no pair flips. At these scales we only see one distribution which again peaks around  $s_{\parallel} = r_{\parallel}$ , telling us that most of the pairs are moving away from each other according to Hubble flow. So we can understand that these corrections are essential if we need to accurately model RSD at scales of



**Figure 3.** Characterising the relative line-of-sight velocity distribution. First panel on left shows the mean relative velocity, followed by the velocity dispersion. The third and the fourth panels show the standardised moments skewness and kurtosis respectively. Skewness is a measure of the symmetry of the distribution while kurtosis acts as a measure of the tailedness of the distribution.



**Figure 4.** When one compares Eq. 1 which is quoted as the exact equation against the “true” correlation function computed from the simulations, it can be seen that they are not exact. The formalism given in Eq. 1 leads to an undermined value.

$s < 20 h^{-1}\text{Mpc}$ . For large scales, both the formulations will be equivalent.

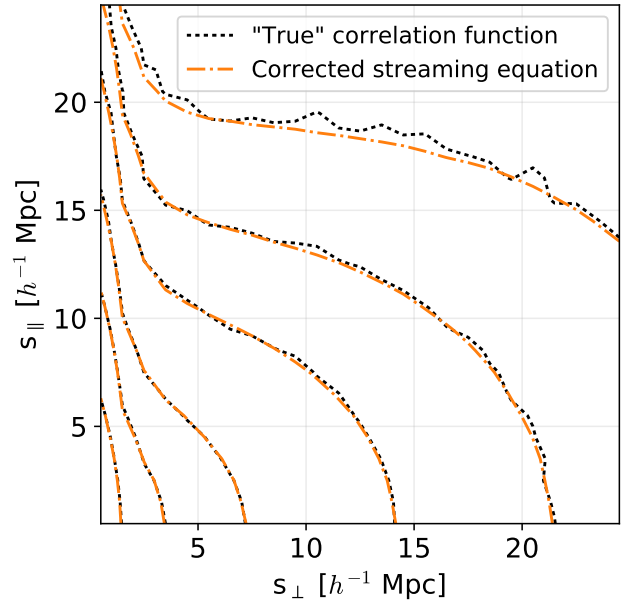
#### 4.1 Relation to radial and tangential components of pairwise velocity

It is of interest to see how the relative line-of-sight velocity is connected to the radial and tangential components of the pairwise velocity. The radial component is defined as the component which is along the separation vector, which can be mathematically written as

$$v_r = (\mathbf{v}_2 - \mathbf{v}_1) \cdot \hat{\mathbf{r}}. \quad (11)$$

However before we define the tangential component we will need to introduce few other terms as follows

$$\mathbf{n}_{12} = (\mathbf{r}_2 - \mathbf{r}_1) \times \hat{\mathbf{z}}, \quad (12)$$



**Figure 5.** Corrected streaming equation compared with the simulation. Orange line denotes the correlation function obtained using the corrected streaming equation (Eq. 10) and black dashed line denotes to the redshift space correlation function measured directly from simulation.

where  $\mathbf{n}_{12}$  is the normal vector to the plane defined by the pairwise vector and the line-of-sight axis, which in this case is the z-axis.

$$v_{\text{away}} = (\mathbf{v}_2 - \mathbf{v}_1) \cdot \hat{\mathbf{n}}_{12}, \quad (13)$$

where  $v_{\text{away}}$  is the velocity component which is in direction of the normal vector away from the plane. Knowing these quantities, we can define the tangential velocity vector as follows

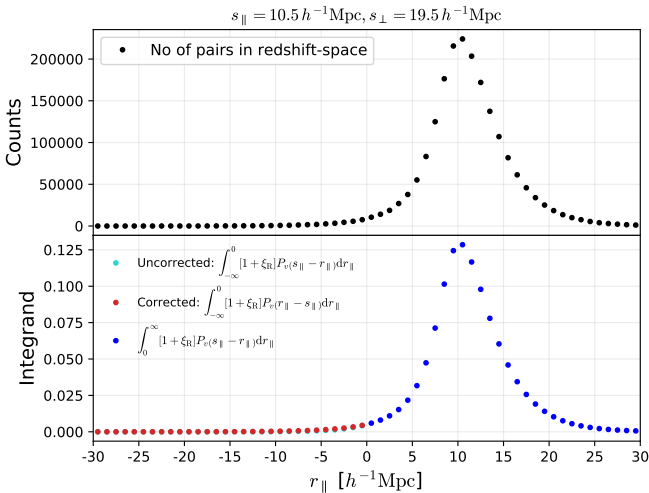
$$\mathbf{v}_t = (\mathbf{v}_2 - \mathbf{v}_1) - (v_r \cdot \hat{\mathbf{r}}) - (v_{\text{away}} \cdot \hat{\mathbf{n}}_{12}). \quad (14)$$

However we are interested in the magnitude of the tangential component which we define as

$$v_T = \text{sgn}(v_{t_z}) |\mathbf{v}_t|. \quad (15)$$



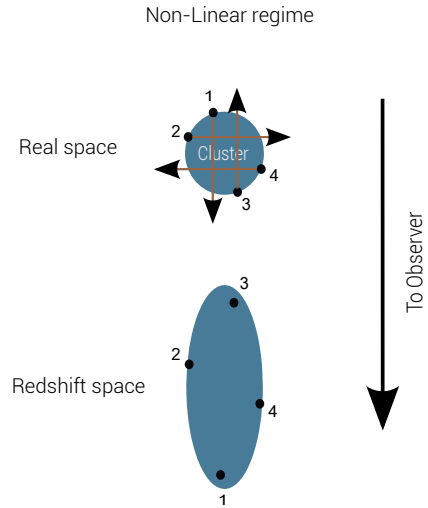
**Figure 6.** Consistency check for the streaming equation. Top panel: Shows the number of pairs in redshift-space for particular  $s_{\parallel}$  and  $s_{\perp}$ .  $r_{\parallel}$  is obtained as  $s_{\parallel} - v_{\parallel}$ . Bottom panel: Shows the integrand distribution of the streaming equations at same  $s_{\parallel}$  and  $s_{\perp}$  as top panel. Due to using the wrong argument of the PDF in the uncorrected integrand, the final value will be undermined compared to the true value.



**Figure 7.** Same as Fig. 6, but for larger scales. Both the uncorrected (Eq. 1) and corrected (Eq. 10) streaming equations are equivalent at these scales. The impact of pair flips are negligible.

The relative line-of-sight velocity can then be related to the radial and tangential components pairwise velocity as follows

$$v_{\text{los}} = \text{sgn}(r_{\parallel}) [v_r \cos \theta + v_T \sin \theta], \quad (16)$$



**Figure 8.** Pair flips happening due to the “Finger-of-God” effect which is a non-linear effect. In this figure, we take the redshift-space as our reference frame. In that framework, objects 1 and 3 here then represent the flipped pairs in real-space. Pair flips are a result of the high relative in-fall velocity of these pairs at non-linear scales.

where  $\cos \theta = r_{\parallel}/r$ . Fig. 9 shows that the equivalence of Eq. 16 indeed holds true. Thus it is exact. As we know exactly how to go from one formulation to another, it would be interesting to how they are related to each other statis-

tically, namely we show the first three cumulants and how they are related to each other in the coming subsections.

#### 4.1.1 First cumulant

To see how the first cumulant of relative line-of-sight velocity is connected to raw moments of the radial and tangential components, we need to average Eq. 16. On averaging, we obtain the following

$$\langle v_{\text{los}} \rangle = \text{sgn}(r_{\parallel}) [\langle v_r \rangle \cos \theta + \langle v_T \rangle \sin \theta]. \quad (17)$$

However due to isotropy  $\langle v_T \rangle = 0$ , implying the contribution to the mean from the tangential component vanishes. Mean of the relative line-of-sight velocity is then

$$\langle v_{\text{los}} \rangle = \text{sgn}(r_{\parallel}) [\langle v_r \rangle \cos \theta]. \quad (18)$$

The presence of  $\text{sgn}(r_{\parallel})$  term is essential as it ensures that  $\langle v_{\text{los}}(-r_{\parallel}, r_{\perp}) \rangle = \langle v_{\text{los}}(r_{\parallel}, r_{\perp}) \rangle$ , which we have verified from the simulations. In Fig. 10, we show the equivalence of Eq. 18 and show that they are equivalent.

#### 4.1.2 Second cumulant

The second cumulant of the relative line-of-sight velocity is related to its raw moments through

$$\sigma_{\text{los}}^2 = \langle v_{\text{los}}^2 \rangle - \langle v_{\text{los}} \rangle^2 \quad (19)$$

where we already know the form of  $\langle v_{\text{los}} \rangle$ .  $\langle v_{\text{los}}^2 \rangle$  can be obtained by first squaring Eq. 16 and then averaging it as below

$$\langle v_{\text{los}}^2 \rangle = \langle v_r^2 \rangle \cos^2 \theta + \langle v_T^2 \rangle \sin^2 \theta, \quad (20)$$

where we have used  $\langle v_r v_T \rangle = 0$ . This has indeed been checked from the simulation and assured that this condition holds. In fact any term involving averaging of odd power of  $v_T$  disappears due to the isotropy. The second cumulant is then given as

$$\sigma_{\text{los}}^2 = \sigma_r^2 \cos^2 \theta + \sigma_T^2 \sin^2 \theta \quad (21)$$

where  $\sigma_r^2 = \langle v_r^2 \rangle - \langle v_r \rangle^2$  and  $\sigma_T^2 = \langle v_T^2 \rangle$ . We show that the Eq. 21 is indeed exact in Fig. 10.

#### 4.1.3 Third cumulant

The third cumulant of the relative line-of-sight velocity is connected to its raw moments as

$$\gamma_{\text{los}} = \langle v_{\text{los}}^3 \rangle - 3 \langle v_{\text{los}}^2 \rangle \langle v_{\text{los}} \rangle + 2 \langle v_{\text{los}} \rangle^3 \quad (22)$$

The form of  $\langle v_{\text{los}}^3 \rangle$  can be obtained trivially by cubing Eq. 16 and averaging it thereafter as below

$$\langle v_{\text{los}}^3 \rangle = \langle v_r^3 \rangle \cos^3 \theta + 3 \langle v_r v_T^2 \rangle \cos \theta \sin^2 \theta, \quad (23)$$

It should be noted that following the isotropy argument again,  $\langle v_r^2 v_T \rangle = \langle v_T^3 \rangle = 0$ . This also has been verified from

the simulations. Using Eqs. 18, 20 and 23, Eq. 22 can be rewritten as

$$\gamma_{\text{los}} = \text{sgn}(r_{\parallel}) \left( \cos^3 \theta \left[ \langle v_r^3 \rangle - 3 \langle v_r^2 \rangle \langle v_r \rangle + 2 \langle v_r \rangle^3 \right] + 3 \cos \theta \sin^2 \theta \left[ \langle v_r v_T^2 \rangle - \langle v_r \rangle \langle v_T^2 \rangle \right] \right) \quad (24)$$

$$\gamma_{\text{los}} = \text{sgn}(r_{\parallel}) \cos \theta \left( \cos^2 \theta \gamma_r + 3 \sin^2 \theta \text{Cov} [v_r, v_T^2] \right) \quad (25)$$

Thus we see that the third cumulant of the relative line-of-sight velocity not only depends on the third cumulant of the radial component of the pairwise velocity but also the covariance between the radial and square of the tangential component of the pairwise velocity. This is similar to the form as mentioned in Uhlemann et al. 2015 and Bianchi et al. 2016, however we have explicitly derived the form and show the equivalence in Fig. 10.

## 5 DISSECTING VELOCITY DISTRIBUTION

To have a better understanding of the relative line-of-sight velocity distribution, it is important to know how halo and non-halo particles plays a role in the distribution. For this purpose, we classify the dark matter particles as either ‘halo’ particles or ‘field’ particle. This was done with the help of ROCKSTAR halo finder. Particles which were identified to belong to a halo were subsequently classified as ‘halo’ particles and the rest as ‘field’ particles.

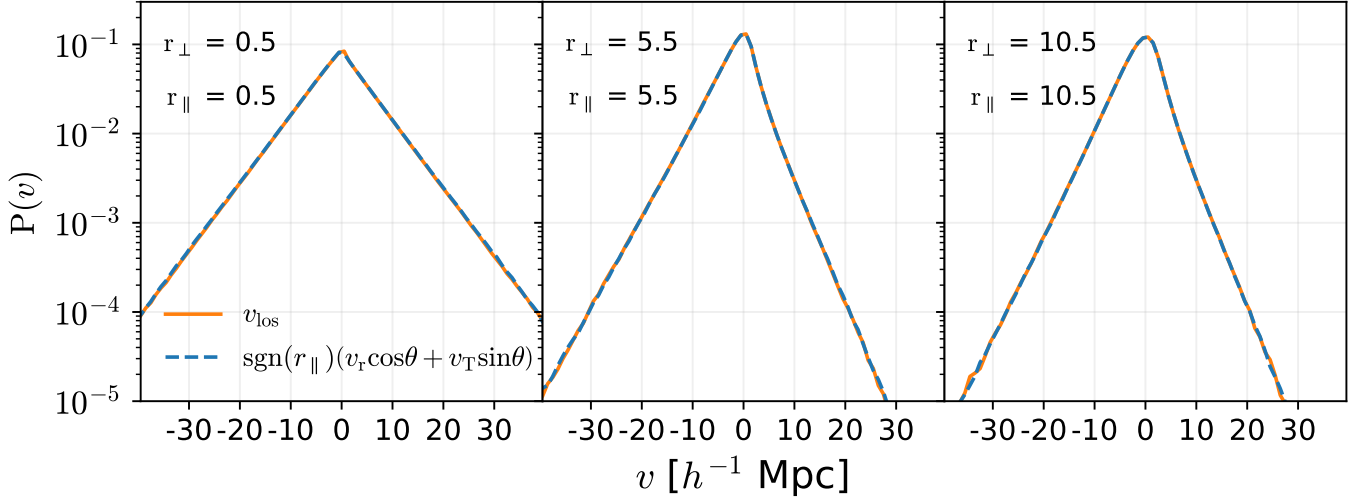
Fig. 11 shows the result of how the components contribute to the total distribution. We see that at all scales we have considered here, the field-field component seems to be the dominant factor. The mean infalling velocity of the full DM distribution hence seems to be driven by the field-field particle pairs. However it should be noted that halo-halo pairs have a more negative in-fall velocity compared to the field pairs. Narrow wings of the field-field component with the mean in-fall velocity of zero indicates that the pairs which does not belong to halos mostly undergo coherent motion. The wings of the total distribution are contributed by both the halo-halo and the field-halo pairs. The combination of distributions of these components with different mean in-fall velocity gives rise to the asymmetry along with a fatter tail than what is to be expected from a normal distribution.

Similarly we can dissect the halo-halo PDF into further subsamples defined by the halo mass. Fig. 12 shows the contribution from each of the subsamples considered which is normalised to the halo PDF. It should be noted that the PDFs of the subsamples were created using particles of same mass subsample, cross contributions are not shown here.

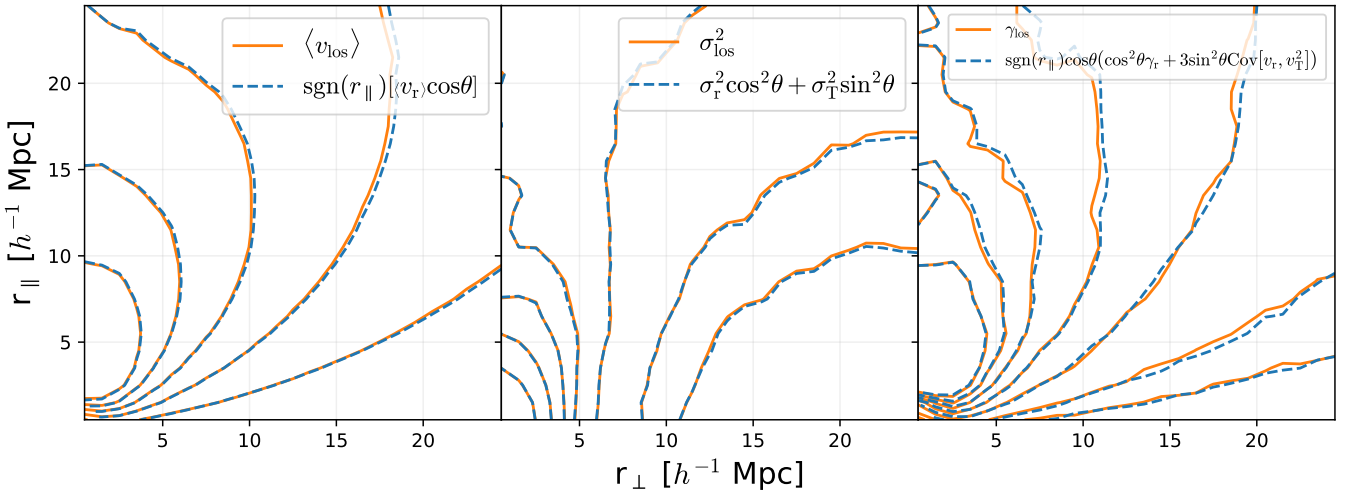
## 6 PHENOMENOLOGICAL MODEL

GH is the mixture modelling consisting of normal distribution and Generalised Inverse Gaussian distribution. It is similar to modelling pairwise velocity as mixture of the normal and log-normal density (see tinker). GIG and log-normal are very similar. Try to connect them and see. Density weighted





**Figure 9.** Shows the equivalence of relative line-of-sight velocity to the radial and the tangential component of the pairwise velocity as given by Eq. 16.



**Figure 10.** Checking the equivalence of the cumulants as shown in Eqs. 18, 21 and 25

velocity, so would make sense to connect them. Could explain why we use it and why they could be the true distribution.

## 7 CONCLUSIONS

### ACKNOWLEDGEMENTS

JK acknowledges the financial support from the Bonn-Cologne Graduate School (BCGS).

### REFERENCES

Bianchi D., Percival W. J., Bel J., 2016, *MNRAS*, **463**, 3783  
Fisher K. B., 1995, *ApJ*, **448**, 494

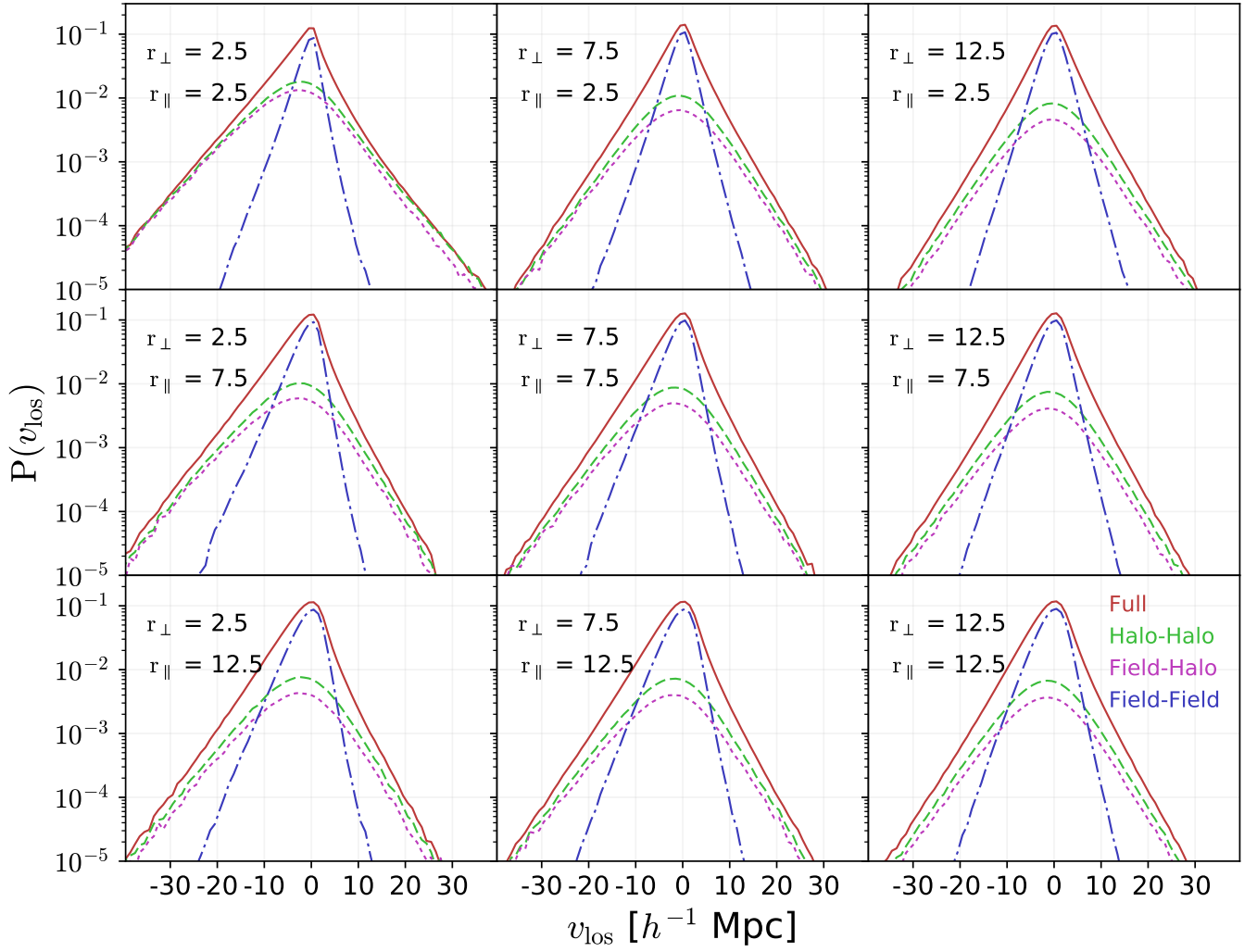
*MNRAS* **000**, 1–7 (2017)

Landy S. D., Szalay A. S., 1993, *ApJ*, **412**, 64  
Peebles P. J. E., 1993, *Principles of Physical Cosmology*  
Scoccimarro R., 2004, *Phys. Rev. D*, **70**, 083007  
Uhlemann C., Kopp M., Haug T., 2015, *Phys. Rev. D*, **92**, 063004

### APPENDIX A: CONVERGENCE

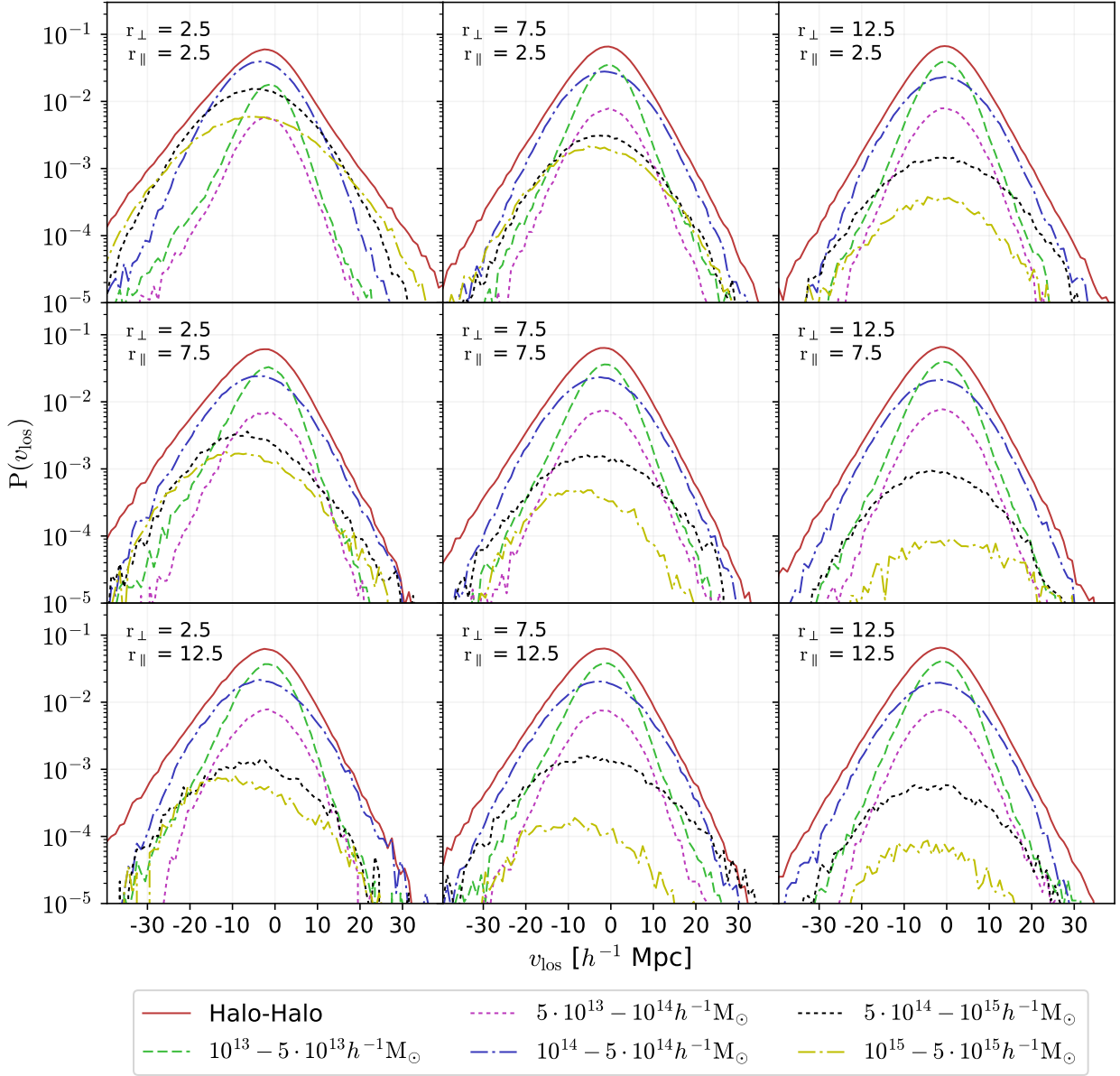
### APPENDIX B: EXTRA

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

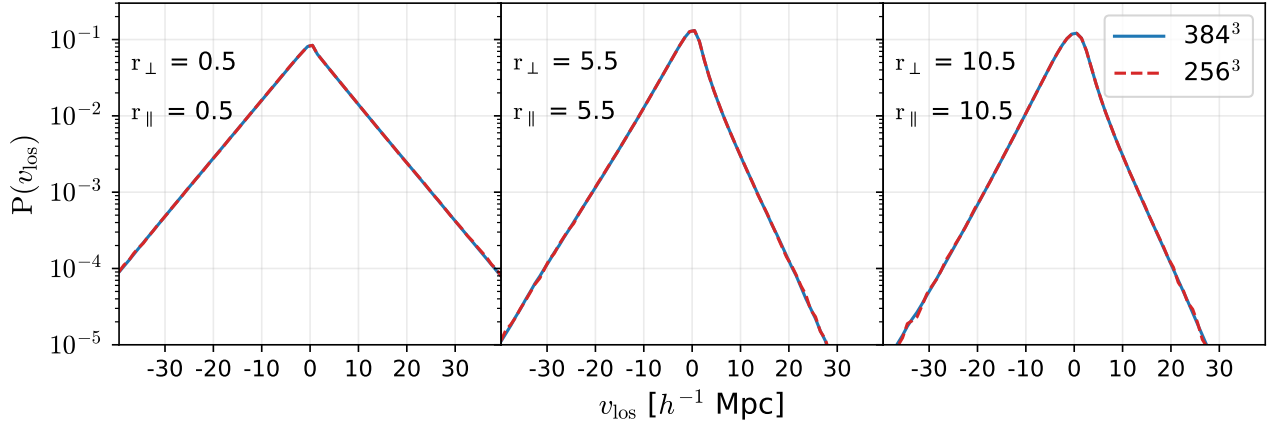


**Figure 11.** Relative line-of-sight velocity distribution for different components normalised to full PDF. Green dashed line refers to the pairs which are comprised of 2 halo particles, blue dash dotted line refers to the pair which are formed by 2 field particles and finally magenta dotted line showcases the distribution contributed by pairs in which one of the particle is a field particle and the other a halo particle. The distribution of the full particles is shown using the red solid line. The distance mentioned in each subplot refers to the mean value of each bin, having a bin width of  $1 h^{-1}$  Mpc.





**Figure 12.** Relative line-of-sight velocity distribution for different components. Green dashed line refers to the pairs which are comprised of 2 halo particles, blue dash dotted line refers to the pair which are formed by 2 field particles and finally magenta dotted line showcases the distribution contributed by pairs in which one of the particle is a field particle and the other a halo particle. The distribution of the full particles is shown using the red solid line. The distance mentioned in each subplot refers to the mean value of each bin, having a bin width of  $1 h^{-1} \text{ Mpc}$ .



**Figure A1.** Relative line-of-sight velocity distribution. The distance mentioned in each subplot refers to the mean value of each bin, having a bin width of  $1 h^{-1}$  Mpc.