The intervelocity of galaxy pairs in $\Lambda$CDM

The observed velocity peak at $\sim$130 km s$^{-1}$ is not unique to MOND

Marcel S. Pawlowski$^1$, Kosuke Jamie Kanehisa$^1$, Salvatore Taibi$^1$, and Pengfei Li$^1$,*

Leibniz-Institute for Astrophysics, An der Sternwarte 16, 14482 Potsdam, Germany
e-mail: mpawlowski@aip.de, kkanehisa@aip.de, staibi@aip.de, pli@aip.de

Received March 28, 2022; accepted July 16, 2022

ABSTRACT

Context. Observational studies of pairs of galaxies have uncovered that their differential line-of-sight velocities indicate the presence of a peak in their three-dimensional intervelocity distribution at 130 – 150 km s$^{-1}$. It had been argued that galaxy pairs in the standard model of cosmology, $\Lambda$CDM, should not exhibit such an intervelocity peak, while Modified Newtonian Dynamics (MOND) predicts such a preferred intervelocity for paired galaxies. However, no direct comparison with $\Lambda$CDM applying the same selection criteria and methodology as the observational studies has been performed yet, placing the comparison on unsure footing.

Aims. To rectify this, we investigate this potential challenge for $\Lambda$CDM by determining whether an analog of the observed intervelocity peak is present in galaxy pairs within the IllustrisTNG-300 cosmological simulation.

Methods. We identify galaxy pairs following the observational study’s selection criteria, measure their projected velocity difference, and analyse both the de-projected as well as the full velocity difference for this galaxy pair sample in the simulation.

Results. We recover a deprojected intervelocity peak at $\sim$130 km s$^{-1}$ for galaxy pairs selected from the simulation. The full three-dimensional velocity information available for the pairs in the simulation also reveals a clear preference for this intervelocity.

Conclusions. The intervelocity peak among galaxy pairs does not appear to be a feature unique to MOND, but is also present in $\Lambda$CDM. It can thus not be claimed as a unique success of either theory over the other. Developing the galaxy pair intervelocity into a test of gravity in the low acceleration regime will require more detailed studies to identify measurable differences in the models.

Key words. galaxies: groups – general – galaxies: interactions – galaxies: kinematics and dynamics – galaxies: statistics – dark matter

1. Introduction

Nottale & Chamaraux (2018a) have obtained one of the largest samples of galaxy pairs with spectroscopic velocity information, using observed galaxies from the HyperLEDA database (Makarov et al. 2014). The large sample of 13114 pairs allows to (statistically) deprojected the distribution in line-of-sight velocity difference $\Delta v$ between pair members to obtain the distribution of full three-dimensional intervelocities $v_{\text{deproj}}$ (Nottale & Chamaraux 2018b). Applying this deprojection to their pair catalog Nottale & Chamaraux (2020) have identified a robust peak at $v_{\text{deproj}} \sim 150$ km s$^{-1}$. Scarpa et al. (2022a) confirmed this intervelocity peak with an expanded input galaxy catalog, reporting its location to be at $v_{\text{deproj}} = 132 \pm 5$ km s$^{-1}$.

Scarpa et al. (2022a) proposed that galaxy pairs can be used to test gravity in the low acceleration regime, specifically to investigate the dark matter alternative Modified Newtonian Dynamics (MOND, Milgrom 1983) Famaey & McGaugh (2012). Since the MOND force at low accelerations ($a \ll a_0$, as realized for galaxy pairs) scales as $1/r$, the total mutual velocity of a pair of galaxies in circular orbit is set by their mass and independent of separation r. The absolute magnitude limit of the observed pairs ($M_0 < -18.5$) sets a typical galaxy mass scale, which then translates into a typical intervelocity. Scarpa et al. (2022b) demonstrate that a realistic B-band mass-to-light ratio of $M/L_*$ reproduces the observed peak position in MOND.

In contrast, Scarpa et al. (2022a) argued that the velocity peak is “hard to justify in the context of numerical simulations of cosmological structure formation”. This assessment appears mostly based on Moreno et al. (2013), who study galaxy pairs in the Millennium simulation (Springel et al. 2005). They find a broad distribution in intervelocities of pairs identified by three-dimensional separation. Such a pair population is possible in simulations where full phase-space information is available, but not directly comparable to the observational pair-finding algorithm that must rely on only projected positions and redshift velocities. The simulation study also allows much more extreme mass ratios between paired galaxies, while the observational study applies a stricter isolation criterion. Nevertheless, pairs in Moreno et al. (2013) do display an intervelocity peak at $\sim 200$ km s$^{-1}$ if restricted to a central-satellite subsample. Considering only pairs whose mutual binding energy dominates, as opposed to them being bound to a third object, their intervelocity distribution even reaches a maximum between 100 and 200 km s$^{-1}$. Therefore, it is not firmly established that the observed peak is absent in a $\Lambda$CDM simulation, if galaxy pairs were selected following the observational study’s approach.

We set out to determine whether an intervelocity peak between paired galaxies is indeed absent in a $\Lambda$CDM context. Were this the case, the observed peak would constitute both a serious challenge to $\Lambda$CDM and a success for MOND. We approach this task by identifying galaxy pairs in the modern hydrodynamical cosmological simulation IllustrisTNG-300 (Nelson et al. 2019, Springel et al. 2018, Nelson et al. 2018, Pillepich et al. 2018).

* Humboldt Fellow
To ensure direct comparability, we mock-observe the simulation to only consider projected positions and line-of-sight velocities when identifying pairs, follow the pair-finding algorithm of the observational studies, and also apply the same statistical deprojection algorithm to obtain the three-dimensional intervelocities.

2. Simulation, Galaxy and Pair Selection

Nottale & Chamaraux (2018a)’s galaxy sample covers those with B-band magnitudes $M_B < -18.5$ and heliocentric redshifts $2500 < v < 165000$ km s$^{-1}$, which – in the absence of peculiar velocities – corresponds to distances of $37 < d < 244$ Mpc, adopting a Hubble constant of $H_0 = 67.74$ km s$^{-1}$ Mpc$^{-1}$. To make a fair selection of simulated galaxies using the same constraints, we require a hydrodynamic simulation with a box size exceeding 200 Mpc. We elect to use IllustrisTNG-300, a hydrodynamic cosmological simulation spanning a periodic box of length $L_{\text{box}} = 300$ Mpc under a Planck Collaboration et al. (2016) cosmology.

Following Nottale & Chamaraux (2018a), we create a catalog of simulated subhalo-based galaxies with $M_B < -18.5$, and mock-observed along each of the simulation’s three orthonormal axes. For axis $a \in \{x, y, z\}$, we place an observer on the $a = 0$ plane at the edge of the simulation box and observe along the positive $a$ axis – parallax is assumed to be negligible due to the large distances involved. We calculate the redshift of each galaxy at distance $D_a$ along an axis $a$ as

$$v = H_0 d_a + \nu_{\text{pec}},$$

where $\nu_{\text{pec}}$ is the peculiar velocity of the galaxy in the simulation’s co-moving frame. By rejecting galaxies with a redshift outside $2500 < v < 165000$ km s$^{-1}$, we obtain samples between 109000 – 116000 individual galaxies depending on the axis.

Limitations of this mock-observational approach are discussed in Appendix A. In the following, we focus on the $z$-axis projection to avoid cluttered figures, but comparisons between the different projections are shown in Appendix B.

We investigate a base sample directly using the positions and velocities of the simulation’s galaxy catalog, without introducing any errors. We also generate a full sample with an apparent magnitude cut-off at $m_B = 19$ similar to the observational galaxy catalog, corresponding to an apparent magnitude of $M_B = -18$ observed at the maximum redshift of 160000 km s$^{-1}$ (see Fig. 1 in Nottale & Chamaraux 2018a). For each simulated galaxy, velocity errors are drawn with replacement from the errors in the magnitude-limited HyperLEDA galaxy database to ensure that our overall error distribution follows that of the observational catalog. The velocity $v$ of the galaxy is then displaced by a random draw from a Gaussian distribution with the error as its standard deviation. We also generate two randomized samples: shuffled velocities for which the galaxy velocities are re-assigned randomly while their positions (and thus spatial clustering) are kept, and random positions with positions randomly selected from a uniform distribution within the simulation volume.

To obtain the list of galaxy pair candidates from our mock galaxy catalogs, we closely follow the successive selection criteria of the observational study (Nottale & Chamaraux 2018a). Specifically, we identify pairs for which the two galaxies

1. have B-band magnitudes $M_B < -18.5$
2. have a projected line-of-sight velocity difference of $\Delta v = \sqrt{(\Delta x)^2 + (\Delta y)^2} < 1$ Mpc
3. fulfill a reciprocity criterion to exclude multiples or hierarchical groups.

We then require these candidate pairs to be isolated using the ratio $\rho = r_3/r_p$, where $r_3$ is the separation of the next-nearest galaxy (within 500 km s$^{-1}$ from the pair) from the pair midpoint. If no nearby galaxy with $M_B < -18.5$ is found, the magnitude range is expanded to include galaxies up to 2.5 magnitudes fainter than the faintest pair member. Only galaxies with $\rho < 2.5$ are retained, resulting in 7840 (8595) pairs for our base (full) galaxy catalog. Stricter limits on $\rho$ can be applied to look at particularly isolated galaxy pair samples.

3. Comparison to Observed Galaxy Pair Catalog

In Fig. 1 we compare the observed galaxy pairs (Nottale & Chamaraux 2018a) with our base and full galaxy pair samples, and the two randomized samples. Observed and simulated galaxy pairs follow comparable distributions in the on-sky separation, with most having $r_p < 200$ kpc. One difference is the much reduced peak at small $r_p$ for the observed sample: very small separations are over-represented in the simulated pairs. This is likely caused by an observational bias. In the simulation, all galaxy positions and velocities are known and included in the input catalogs. In contrast, in the observational galaxy catalog, galaxies that are projected on top of each other, or have only very small separations, are likely either not included, are included as one object, or do not both have measured redshift velocities due to fiber positioning constraints in surveys. To address this difference, we have confirmed that rejecting galaxy pairs with $r_p < 20$ kpc in our simulation’s pair catalog does not affect our later results. Randomizing only the galaxy velocities retains their spatial clustering and thus distribution in $r_p$, while randomized positions result in a much broader distribution.

The middle panel of Fig. 1 shows the difference in the line-of-sight velocities $\Delta v$ of the galaxies in a pair, the quantity to be deprojected to obtain the distribution of three-dimensional intervelocities of paired galaxies. The observed and the simulated galaxy pairs again show a similar behaviour, with an initially steeply declining distribution that flattens at higher $\Delta v$. This is slightly more pronounced even for the simulated full sample including velocity errors than for the observational sample, possibly related to underestimated errors or unaccounted systematics in the spectroscopic velocities, which bias towards larger velocity differences. Both randomized samples show much flatter distributions in $\Delta v$, suggesting that the shape of the distribution has a physical origin.

The right panel of Fig. 1 shows the distribution of the isolation criterion $\rho$. A very similar decreasing in the relative number of more isolated pairs is present in all samples, though pairs in the simulation show a slight excess at small $\rho$ over the observational sample. This is likely a result of the simulation having a full coverage of all galaxies, compared to some degree of incompleteness in the observational catalog. Nottale & Chamaraux (2018a) found that about 10% of the galaxies with $M_B < -18.5$ are members of fairly isolated pairs ($\rho > 5$), and that highly isolated pairs ($\rho \geq 10$) contribute 32.5% of the total number of pairs. We find 4439 fairly isolated pairs in the base sample (8.1% of 109603 galaxies) and 4892 in the full sample (8.9% of 109609 galaxies), and that 34.0% (34.5%) of
The formula is based on the fact that a randomly oriented velocity vector is a Gaussian random variable with zero mean and a given variance. The de-projected velocity distribution can then be derived by inverting the probability distribution of \( \Delta v \). Other bin widths result in very similar plots, in particular the peaks at \( \rho = 0 \), and the peak positions agree within their 1\( \sigma \) uncertainties between the simulated and observed galaxy pairs. These intervelocity peaks are approximately the same intervelocity \( v_{\text{deproj}} \) of \( 150 \) km s\(^{-1} \) reported by Nottale & Chamaraux (2020) and \( v_{\text{deproj}} = 132 \pm 5 \) km s\(^{-1} \) reported by Scarpa et al. (2022).

The peak in our base and full samples are more pronounced than the observed one. The reason might be that the simulation data is, by definition, much cleaner, while observational biases and uncertainties affect the observed sample. The result is indeed more consistent with the observed sample, the peak strength is reduced, but its position is not substantially affected.

The simulation data provides access to the full 6D phase-space of the galaxy pairs, such that we can also immediately investigate the pairs’ 3D intervelocity \( v_{3D} \) without deprojection. This is shown as the red filled histogram in Fig. 2 for the full sample (and in Fig. 2 for the base sample and various test samples). This confirms the de-projection analysis: for galaxy pairs identified in the simulation a clear intervelocity peak is present at \( v_{3D} = 132 \pm 1 \) km s\(^{-1} \), with a width of \( \sigma_{v_{3D}} = 61 \pm 2 \) km s\(^{-1} \).

5. The Intervelocity Peak as a Physical Feature

To investigate the origin of the intervelocity peak in the simulation, we performed a number of tests and additional comparisons. These are illustrated in Fig. 3, showing distributions in the three-dimensional separation of the two pairs members \( r_{3D} \) and their full three-dimensional intervelocity \( v_{3D} \).

We first compared the simulated samples to the two randomized samples (left panels in Fig. 3). While the shuffled velocities sample preserves the spatial clustering and results in a very similar distribution in \( r_{3D} \) as for the mock-observed samples, neither randomized sample reproduces the velocity distribution or its narrow peak. Instead, they display broad maxima at \( v_{3D} = 426 \pm 8 \) and \( 491 \pm 11 \) km s\(^{-1} \), respectively. This is close to the velocity limit of the pair selection algorithm and thus likely driven by that constraint. This suggests that the velocity peak identified
in the simulation is not merely an emergent feature of the typical galaxy distribution or their typical peculiar velocities, but rather has a physical origin related to galaxies associated with each other.

To investigate such a physical association, we have determined which galaxy pairs identified in the mock-observed simulation share the same halo ID. These have been identified as being physically associated by the halo finder. We find that 83.4% (82.5%) of the galaxy pairs in the base (full) sample share a common halo ID, indicating that the galaxy pair finding criteria are very good at identifying truly associated galaxies, with only one sixth being contamination by spurious pairs. We also find that 75.2% (74.4%) of the pairs share a unique halo ID not shared by any other pair, indicating that only a small fraction of pairs are associated with massive halos (likely galaxy clusters) hosting more than one pair.
Fig. 4. Three-dimensional intervelocity $v_{3\text{D}}$ and calculated circular velocity $v_{\text{circ}}$ for galaxy pairs identified in the simulation with separations $r_{3\text{D}} < 100$ kpc (blue) and $100$ kpc $< r_{3\text{D}} < 1$ Mpc. For sufficiently separated pairs, the circular velocity provides a good estimate of the 3D intervelocity, suggesting a physical origin of the intervelocity peak driven by the typical masses and separations of the selected galaxy pairs.

In the middle panels of Fig. 3 we show distributions in $r_{3\text{D}}$ and $v_{3\text{D}}$ for pairs with and without a common halo ID. Only galaxy pairs in a common halo have small $r_{3\text{D}}$, as expected for physically associated systems. They also drive the bulk of the velocity peak at $v_{3\text{D}} \sim 130$ km s$^{-1}$, while the (spurious) pairs belonging to different halos preferentially show lower intervelocities. This appears in line with their larger distance, resulting in lower mutual acceleration. Again fitting a two-Gaussian model, we find that pairs associated to the same halo display a peak position of $v_{3\text{D}} = 138 \pm 1$ km s$^{-1}$ for the full sample, while those belonging to different halos have a peak position at $v_{3\text{D}} = 88 \pm 2$ km s$^{-1}$. The corresponding positions for the base sample do not differ substantially.

Finally, we compare different ranges in the absolute magnitude $M_{B,1}$ of the more luminous galaxy of each pair. This is based on an extended galaxy pair catalog where also fainter galaxies with $M_{B,1} \leq -17$ were considered. As shown in the right panels of Fig. 3, their spatial distributions are comparable, but pairs with more luminous primaries display higher intervelocities. We find peaks at $v_{3\text{D}} = 86 \pm 1, 117 \pm 1$, and $153 \pm 1$, km s$^{-1}$ for the subsequently more luminous samples in Fig. 3. This is in line with a physical origin of the intervelocity peak as being caused by the mutual acceleration of paired galaxies, with more luminous (and thus on average more massive) primary galaxies resulting in higher accelerations, causing larger velocities. The absolute magnitude limit of the observational study then determines which galaxy masses contribute to the identified pairs most, implicitly setting the intervelocity peak of the sample.

We test the latter in our simulation by calculating the circular velocity $v_{\text{circ}} = \sqrt{\frac{GM_1 + M_2}{r_{\text{mean}}}}$ of each galaxy pair, using the total mass of each galaxy, $M_{1,2}$, obtained from the galaxy catalog of the simulation. We ignore pairs with separations exceeding 1 Mpc as these are most likely non-physical contamination. The distribution of $v_{\text{circ}}$ as well as the full 3D velocity difference $v_{3\text{D}}$ of the galaxy pairs is plotted in Fig. 4. For pairs with separations of 100 kpc to 1 Mpc, $v_{\text{circ}}$ provides an excellent match to $v_{3\text{D}}$ with both showing a pronounced peak at $\sim 130$ km s$^{-1}$, and a median $v_{3\text{D}}$ in the simulation of 161.9 km s$^{-1}$ while $v_{\text{circ}}$ for the median galaxy masses and separation is 157.5 km s$^{-1}$. However, for pairs with separations closer than 100 kpc, the circular velocity calculation over-estimates their 3D velocity difference. This is expected, as the velocities of interacting galaxies will be decreased due to dynamical friction once the halos overlap, and the overlapping halos invalidate the implicit assumption of point masses, such that the effective masses in the calculation are smaller than the total masses $M_1$ and $M_2$.

6. Discussion and Conclusions

Pairs of galaxies identified in the observational galaxy catalog HyperLEDA (Makarov et al. 2014) [Nottale & Chamaraux 2018a] have been found to display a peak in their deprojected (three-dimensional) intervelocity distribution at $v_{\text{deproj}} = 130 - 150$ km s$^{-1}$ [Nottale & Chamaraux 2020, Scarpa et al. 2022a,b].

To determine whether such a feature is expected in a cosmological context within $\Lambda$CDM, we have mock-observed galaxies from the IllustrisTNG-300 simulation [Nelson et al. 2019] and applied the observational approach to identify galaxy pairs. We recover a pronounced peak in the intervelocity of isolated pairs, both via a deprojection of line-of-sight velocity differences (Nottale & Chamaraux 2018b) at $v_{\text{deproj}} = 125 \pm 7$ km s$^{-1}$, and via the direct three-dimensional intervelocity of the identified galaxy pairs at $v_{3\text{D}} = 132 \pm 1$ km s$^{-1}$.

The agreement between the position of the intervelocity peak of the observational studies and the peak we identified in the simulation is striking, and can be considered a success of the simulation – and by extension $\Lambda$CDM – in reproducing this observed feature. It is likely of physical origin since it is not present in randomized galaxy samples, set by the majority of identified pairs for which both galaxies belong to a common dark matter halo, and its position scales with the pair primary’s luminosity and thus mass. It appears plausible that the intervelocity peak of galaxy pairs as a probe of the total mass of galaxies, is related to abundance matching or the stellar mass-halo mass (SMHM) relation. Via these the considered range of galaxy luminosities translates into a corresponding range in total masses, which in turn determine the dynamics leading to a preferred intervelocity.

That this intervelocity in $\Lambda$CDM is consistent with the one expected in MOND might then be because the TNG model was devised to reproduce the SMHM relation (Pillepich et al. 2018). Since abundance matching such as that of Behroozi et al. (2013) is consistent with the observed baryonic Tully-Fisher relation (see Fig. 4 of Li et al. 2022), which in turn matches MOND predictions, this might explain why the considered galaxy luminosities in this cosmological simulation are tied to just the right total halo mass to reproduce the observed feature of the intervelocity distribution.

Recovering the intervelocity peak with a $\Lambda$CDM simulation does not reduce MOND’s success in reproducing it, but it shows that (1) MOND is not the only model offering an explanation, and (2) $\Lambda$CDM is not immediately challenged by the observed intervelocity peak of galaxy pairs. This does not mean that the dynamics of isolated galaxy pairs cannot be developed into a useful tool to test gravitational dynamics in the low-acceleration regime. However, this will require more dedicated research to determine how the peak comes about in the cosmological context, in what details predictions of the intervelocity distributions between $\Lambda$CDM, MOND, and other models differ (e.g. are there environmental effects), and what observational data will be required to discriminate these alternatives.
Acknowledgements. We thank Mordehai Milgrom for interesting discussions, and the anonymous referee for helpful suggestions, that improved our manuscript. MSP acknowledges funding of a Leibniz-Junior Research Group (project number J94/2020) and a KT Boost Fund by the German Scholars Organization and Klaus Tschira Stiftung. PL is supported by the Alexander von Humboldt Foundation.

References
Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013, ApJ, 770, 57
Famaey, B. & McGaugh, S. S. 2012, Living Reviews in Relativity, 15, 10
Li, P., McGaugh, S. S., Lelli, F., et al. 2022, ApJ, 927, 198
Makarov, D., Prugniel, P., Terekhova, N., Courtois, H., & Vauglin, I. 2014, A&A, 570, A13
Milgrom, M. 1983, ApJ, 270, 365
Moreno, J., Black, A. F. L., Ellison, S. L., et al. 2013, MNRAS, 436, 1765
Nelson, D., Pillepich, A., Springel, V., et al. 2018, MNRAS, 475, 624
Nelson, D., Springel, V., Pillepich, A., et al. 2019, Computational Astrophysics and Cosmology, 6, 2
Nottale, L. & Chamaraux, P. 2018a, Astrophysical Bulletin, 73, 310
Nottale, L. & Chamaraux, P. 2018b, A&A, 614, A45
Nottale, L. & Chamaraux, P. 2020, A&A, 641, A115
Pillepich, A., Nelson, D., Hernquist, L., et al. 2018, MNRAS, 475, 648
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, A&A, 594, A13
Scarpa, R., Falomo, R., & Treves, A. 2022a, MNRAS, 510, 2167
Scarpa, R., Falomo, R., & Treves, A. 2022b, arXiv e-prints, arXiv:2202.13766
Springel, V., Pakmor, R., Pillepich, A., et al. 2018, MNRAS, 475, 676
Springel, V., White, S. D. M., Jenkins, A., et al. 2005, Nature, 435, 629
Appendix A: Comparisons of the Observed and Simulated Galaxy Catalogs

While our mock-observational approach and our pair finding algorithm closely follow that of the observational study, a number of differences are present: (1) our simplified approach does not cover the same volume in shape as the observational galaxy catalog, (2) the simulation is not affected by spatial biases, occlusion, or crowding issues, and (3) we employ a simplified treatment of velocity measurement errors.

The distribution of galaxy pairs in absolute B-band magnitude \( M_B \) and redshift velocity \( v \) is shown in Fig. A.1. The former shows the similarity in the covered magnitude range, though the observational sample is lacking in galaxies fainter than \( M_B = -20 \). This is likely due to insufficient completeness of the observational catalog as compared to the complete galaxy catalog available for the simulation. The distributions in velocity highlight the difference in the selection volumes of the observational study and our mock-observational approach using the IllustrisTNG-300 simulations. The observational sample rises with redshift, as a larger volume of spherical shells contributes to the galaxy distribution. Such a volume is not feasible due to the limited simulation box size. Consequently, the apparent magnitude bias resulting in less rejected pairs at higher redshifts discussed in Nottale & Chamaraux (2018a) is not as severe for our mock-observed simulation samples. We can therefore expect the latter to be less contaminated by non-isolated pairs.

Fig. A.2 compares the error of the pair velocity difference, \( \epsilon_{\mathrm{v}} \), for the observational pair catalog and the one obtained for our full sample, for which velocity errors are drawn from the errors of the HyperLEDA galaxy database. This approach typically results in similar errors for pairs: a strong and sharp peak at small \( \epsilon_{\mathrm{v}} < 10 \, \text{km s}^{-1} \), and another broader peak at higher \( \epsilon_{\mathrm{v}} \). However, there are some differences in the details, such as the relative height of the first peak and the position of the second one. We suspect that some correlation between galaxy errors in the observational data is a possible cause of this.

This appears to be confirmed by the right panel of Fig. A.2 which plots the individual velocity errors for the two members of a pair against each other. Some of the observed pairs indeed show correlation between their velocity errors. This is not unexpected, as galaxies close to each other on the sky might have likely been measured with the same technique, instrument, or survey, resulting in them having similar errors. This can result in low-error galaxies as well as high-error galaxies being predominantly associated with each other, compared to a completely random error assignment. Such effects are difficult to accurately reproduce in mock-observing the simulation. However, they are not expected to significantly alter the results, because the velocities are only relevant for the pair finding algorithm – where they are generally smaller than the allowed velocity difference of \( 500 \, \text{km s}^{-1} \) – and for the deprojected intervelocity – which however turns out to agree with the error-free, three-dimensional intervelocity peak of galaxy pairs in the simulation. This is confirmed by the absence of substantial differences between the results for our base and full input catalogs, despite the former being entirely devoid of any velocity errors.

Appendix B: Deprojected Velocities for Projections Along Other Axes

In Fig. B.1 we plot the intervelocities of galaxy pairs obtained from the deprojection approach of Sect. 4 for the mock-observed (full) galaxy samples projected along the three different axes of the simulation box. Overall, the presence of the intervelocity peak is stable and its peak position is consistent with that of the observed galaxy pairs. However, some differences in the deprojected intervelocity distributions along the three axes are present and illustrate the degree of uncertainty in the velocity deprojection approach.

To more accurately determine the position of the deprojected intervelocity peak despite the wide velocity bins required by the deprojection algorithm, Scarpa et al. (2022b) have performed the deprojection with a range of bin width and then averaged the resulting histograms by re-sampling with a bin width of \( 1 \, \text{km s}^{-1} \). For completeness, we follow their approach for our simulation data in Fig. B.2. Specifically, we apply the deprojection algorithm for bin width between 25 and \( 40 \, \text{km s}^{-1} \), re-sample the resulting de-projected histograms in bins of \( 1 \, \text{km s}^{-1} \) width, and average them. This results in a smoother deprojected velocity distribution that helps to identify the position of the intervelocity peak, but we warn that the choice of bin width and range, as well as numerical noise due to relatively small pair sample sizes, can introduce difficult to quantify biases and inaccuracies in this method. This is illustrated by showing not only the \( z \)-axis projection for the simulations, but also those for the \( x \) and \( y \) directions. The different directions of projection of the simulated galaxy distribution also result in some differences in the re-sampled deprojected intervelocity curves, the amount of which gives an idea of the level to which such curves can be trusted.
Fig. A.1. Properties of galaxy pairs for the observed galaxy pair catalog properties from Nottale & Chamaraux (2018a) (grey shaded histograms), and galaxy data from the IllustrisTNG-300 simulation. For the latter, the blue histogram corresponds to our base model without velocity errors, and the red histograms are from our full model mimicking apparent magnitude biases and errors on the observed line-of-sight velocity. Left panel: Distribution of galaxy pair members in B-band magnitude $M_B$. Right panel: Distribution of galaxy pairs in velocity $v$, consisting of redshift and peculiar velocity.

Fig. A.2. Comparison of velocity errors for the observed and simulated galaxy pair samples. Left panel: Distribution of errors $\epsilon_{\Delta v}$ on the line-of-sight velocity difference $\Delta v$ of galaxy pairs. Right panel: Errors on individual velocity measurements for the two corresponding partner galaxies (more luminous partner on the horizontal axis). Some error correlation is present for the observed sample, but absent in our mock observed simulation data.

Fig. B.1. Deprojected velocities, similar to Fig. 2 but showing the mock-observed simulation along the three different axes of the simulation box.
Fig. B.2. Deprojected intervelocities obtained from averaging deprojections with bin width between 25 and 40 km s$^{-1}$, following the approach introduced by Scarpa et al. (2022b).