Title Here

Andrei Cuceu

Supervisor: Dr. Andrew Pontzen

University College London

March 24, 2018

Abstract

Eisenstein et al. 2007

Contents

Chapter 1

Introduction

1.1 Cosmological Context

Cosmology has seen many advances in the last few decades. Piece by piece we are uncovering the history of the universe and it's components. We now know that our galaxy is just one in trillions (Conselice et al. 2016), and part of a rapidly expanding universe.

Our two main observational tools today are the Cosmic Microwave Background (CMB) and Galaxy Surveys. From the CMB we learn about the primordial universe, and Galaxy Surveys uncover the nature and structure of the Universe at late times. Each of them will be discussed in the next two sections.

Despite a wealth of knowledge, the nature of the two most important components of the Universe still eludes us. Dark Matter and Dark Energy account for about 95% of the matter-energy budget of the universe (e.g. Planck Collaboration et al. 2016).

There are also other gaps in our knowledge. Most of them are either related to the very early universe (before the CMB was emitted) or to the evolution of the Universe between recombination and the first galaxies being formed. In this project we will investigate the second gap.

1.2 The Cosmic Microwave Background

The Cosmic Microwave Background is our main observational tool for studying the early universe. It is the first light emitted in the Universe after recombination, and it encodes plenty of useful information. The CMB is the most perfect black body ever observed (White 1999). Its existence is already a strong proof in support of the Big Bang model. On the other hand, the anisotropies found in the CMB strongly support both the ΛCDM paradigm and Inflation.

CMB surveys have measured these anisotropies very accurately (Figure 1.1). They are key in understanding the matter distribution in the universe. Before recombination Baryons and Radiation were coupled, so the matter distribution at recombination was imprinted in the CMB distribution. The matter distribution then continued to evolve on its own into the large scale structure we see today. Among the key anisotropies detected in the CMB are Baryon Acoustic Oscillations. This feature is a result of oscillations in the primordial Baryon-Photon plasma. Radiation opposed the collapse of baryonic matter into the potential wells created by collapsing dark matter (which does not interact and is free to collapse). This produced acoustic waves that were imprinted in both the matter distribution (we detect it today in the galaxy distribution) and the radiation distribution (we detect it in the CMB). This probe lends strong support both to the existence of dark matter which plays a key role in their creation, and to that of dark energy (as a cosmological constant) which dictates their evolution at late times.

The CMB also has plenty of data in support of Inflation. It shows a uniform, very flat universe that contains mostly gaussian anisotropies. All of these are outcomes of inflation and are hard to explain without it. Of particular interest for us are the gaussian anisotropies. There are many models of inflation , and at the moment we do not have enough precision to distinguish between them. Most inflation models predict some primordial non-gaussianity, however constraining this is key to differentiating between them .

After the CMB was emitted we enter a period called the Dark Ages. Until the first stars and galaxies formed, the CMB was the only radiation in the Universe. This means we have little to no information about this important era. In it lie the
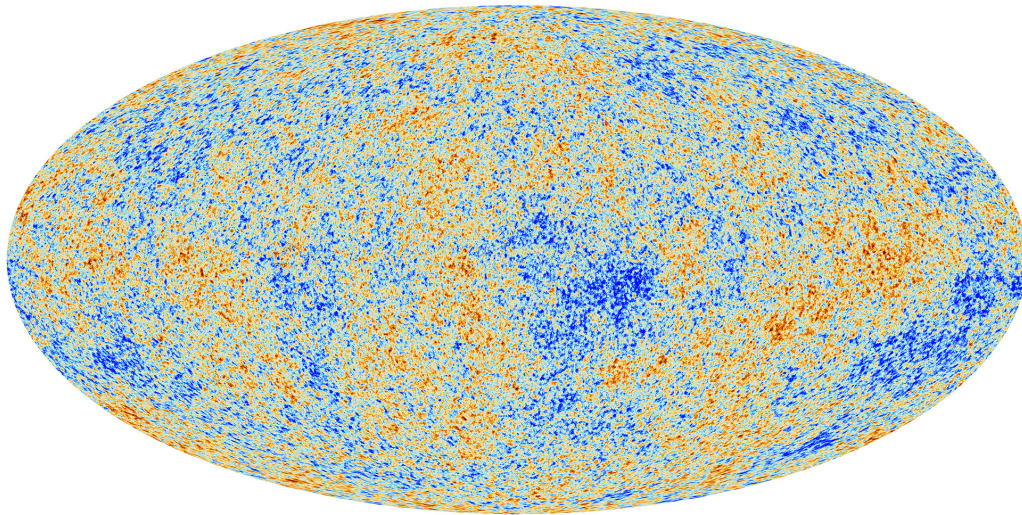
Figure 1.1: Map of the Cosmic Microwave Background acquired by the Planck Space Telescope (Planck Collaboration et al. 2016). After subtracting the effect of our own motion relative to the CMB, the anisotropies only start appearing at $10^{-5}K$. This gives us a glimpse into the primordial matter distribution.

secrets to the formation of the structure we see today in the universe. We will next discuss this Large Scale Structure and the observational tools used to study it, and only turn our attention to this gap in Section 1.4.

1.3   Large Scale Structure and Galaxy Surveys

Introduce the Structure of the Universe today and the tools used to study it.

Talk about the detection of the BAO in the galaxy distribution and its smearing due to collapse.

1.4   The Missing Link (Reconstruction)

Motivate our desire to link the two and talk about the problems we have (Dark Ages)

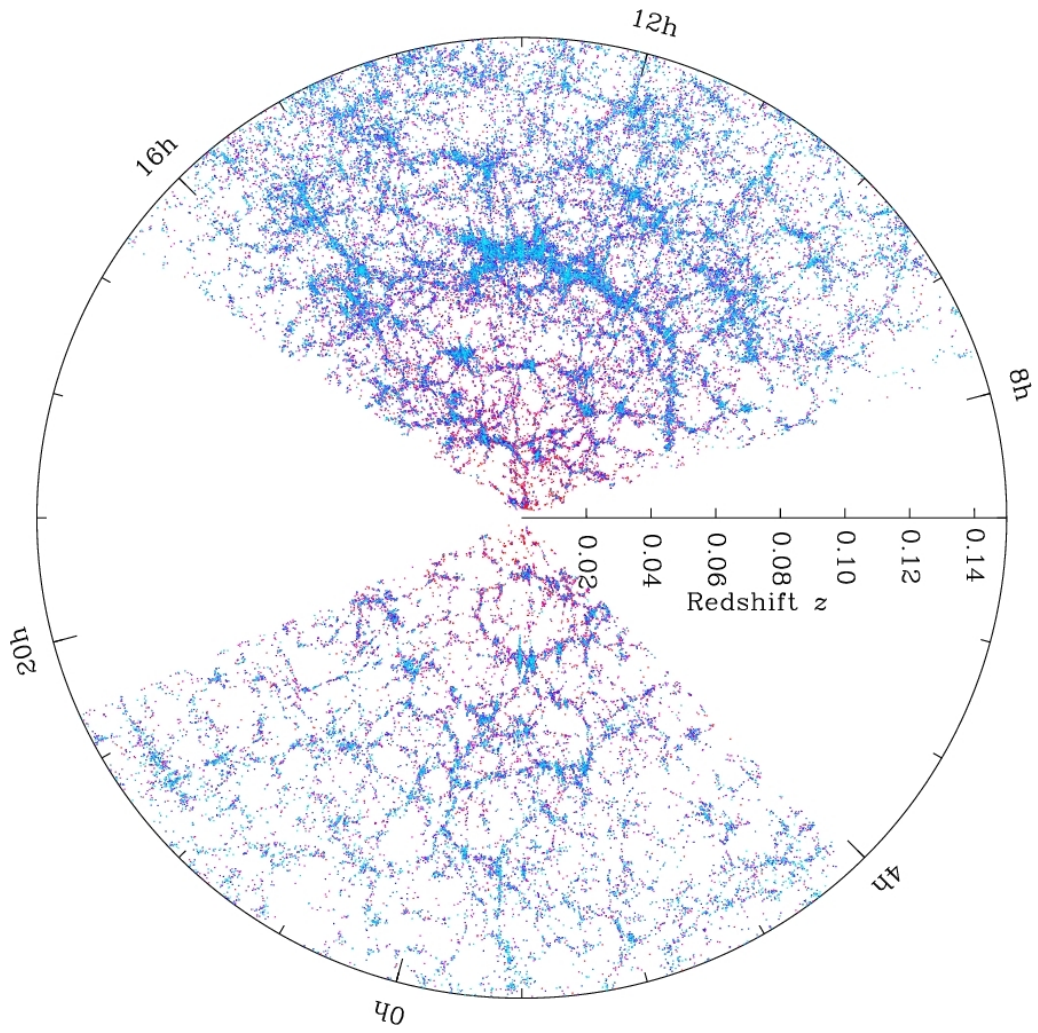Motivate the desire to reconstruct the BAO feature

Figure 1.2: Galaxy map from the Sloan Digital Sky Survey.

Chapter 2

The Growth of Structure

Introduce cosmic variance somewhere

2.1 Perturbation Theory

Give a brief introduction to the use of Perturbation theory to study the evolution of structure. Present advatages and shortcomings

This could be merged with the next section.

2.2 Linear vs Non-Linear Collapse

Talk about the Linear regime of collapse versus the non-linear regime. Present the difficulty of constructing analytical models of non-linear collapse. Motivate our use of simulations as well as our desire to get back to the linear regime for reconstruction.

Add images of the velocity field here

2.3 The Zeldovich Approximation

Introduce the theory of the Zeldovich Approximation and motivate its use (+ background).

2.4   Reconstruction (BAO)

This can be understood by thinking about the difference between looking at the large scales versus the small scales. For example, a very large (by scale) primordial underdensity will probably still be an underdensity at late times (as a huge void). However, on small scales, information tends to be destroyed. A small scale primordial underdensity might be caught in the larger bulk flows and collapse in a larger halo. In this case, the information about the small underdensity is lost.

this paragraph could go in chapter 2

Finally link everything with an overview of Reconstruction techniques and how our work fits into the modern context.

Showcase the BAO reconstruction.

Chapter 3

The Perfect Reconstruction

The first step in understanding the evolution of the universe is to look at the theoretical limits we encounter when trying to reconstruct it. As a field is defined at every point in space, any attempt at representing it with data is inherently imperfect. We would have to measure the density field at every point in the Universe in order to obtain all the information it contains. This fact already implies that no data driven reconstruction will ever succeed at perfectly recovering the primordial density field (unless we manage to make an infinity of measurements).

To show this unavoidable loss of information we performed a 'perfect' reconstruction. We call this reconstruction 'perfect', as it uses data about the primordial positions of all particles (which obviously is not available to observers). However, due to the reasons outlined above, not even this perfect reconstruction succeeds in completely recovering the primordial matter distribution.

3.1   Methods

The first step in studying reconstruction is to find a way to test its effectiveness. To do this, we use cosmological N-body simulations. These simulations give us important insights into how structure evolves in the Universe. More importantly for this project, it allows us to compare any reconstructed density field to the real starting density field. We use data from three simulations available to us. The first one is Simulation A presented in Pontzen, Slosar, et al. 2016 (also re-

ferred to as Simulation A in this work). The other two simulations are variations of the same initial setup, with smaller size and smaller resolution respectively. The details of all three simulations are presented in Table 1.

I should probably give a bit more details about the simulations

Label	Size	Number of Particles	Particle Mass (Solar Masses)
Sim A	$(200Mpc)^3$	512^3	6.59×10^9
Sim B	$(200Mpc)^3$	256^3	5.27×10^{10}
Sim C	$(100Mpc)^3$	256^3	6.59×10^9

Table 3.1: The sizes and number of particles of the three simulations used in this project.

The idea behind a perfect reconstruction is to use data about the initial state of the simulation to perform the reconstruction. We have access to multiple snapshots at various redshifts in our simulations, including the initial positions of all particles (at $z = 99$). Therefore, we used this information to reconstruct the density field. We first measured the density field of various snapshots in the redshift interval $z = 0 - 9$. The field was measured at the particle positions instead of being measured on a regular grid. This is because we want the particles to carry the density field when we move them. After that, all the particles were moved to their starting positions (taken from the initial snapshot at $z = 99$).

3.2 From Images to Statistics

3.2.1 The Reconstructed Density Field

As outlined above, the first step is to measure the density field in a snapshot. Each snapshot contains an indexed list with the positions, velocities and masses of all particles in the simulation. In this chapter, only the positions and masses are needed to perform a perfect reconstruction. To perform the first part of this analysis, we used the *pynbody*[1] package (Pontzen, Roškar, et al. 2013).

To have a visual understanding of the reconstruction, we first make some images of the density field. We use *pynbody* to import the initial snapshot and the

---

[1]https://github.com/pynbody/pynbody

snapshot at $z = 0$. The density field at the particle locations in the final snapshot is calculated and assigned as the density field of the initial snapshot. Density slices through this reconstructed field are compared to the initial and final fields in Figure .

> put image in and ref

We can already see from this comparison that the reconstruction has not recovered all the information, as it is not identical to the initial filed. However, we see the effect that we are after. The reconstruction spreads out the matter from the collapsed filaments onto a more uniform field. Also notice the large difference in the values of the density field. The reconstructed field has density values about 3 orders of magnitude larger than the initial field.

> also talk about the density distribution

This large difference is an interesting side effect of our method. At late times, most particles tend to be clumped together. Therefore, when measuring the density field at the particle positions, we will mostly get very high values. These values do not change when moving the particles, so the final field will also have very high values, but this time distributed on an almost uniform grid. This results in an apparent increase in the total mass of the simulation. As this increase is just a result of the way we represent the density field, it needs to be accounted for when analysing the results. The total mass of the simulation should be conserved.

3.2.2 Correlation with the Initial Field

> TALK ABOUT TAKING THE LOG OF THE DENSITY FIELD

In order to get a better understanding of how well this reconstruction worked, we turn to statistics. A good way to represent the reconstruction is to look at the normalized Cross-Spectrum between the initial and the reconstructed field:

$$\frac{P_{IX}(k)}{\sqrt{P_I(k) * P_X(k)}}$$

where I represents the initial field, and X the reconstructed field.

> The power-spectrum will be introduced in chapter 2

We used the GENPK code[2] (Bird 2017) to measure auto and cross power-spectra of GADGET outputs. The original normalized cross-spectra between the

---

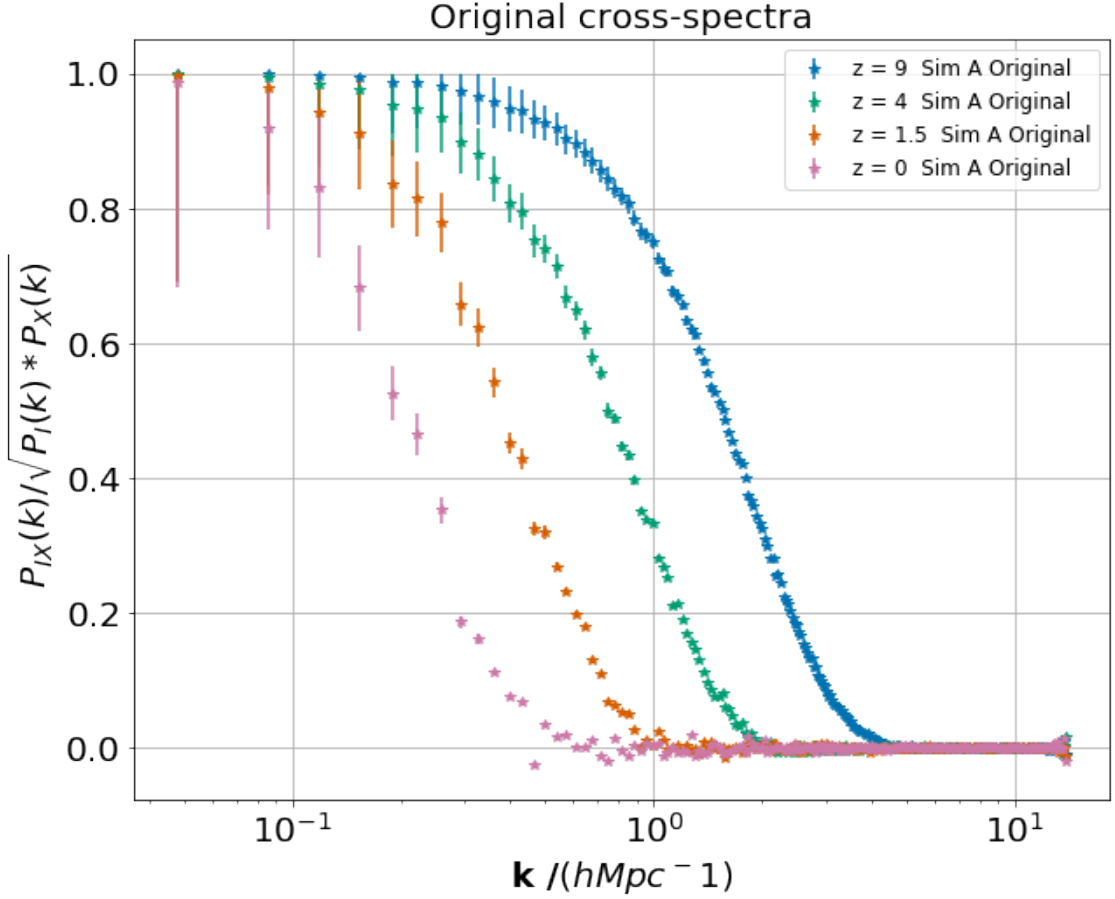[2]https://github.com/sbird/GenPK.git

Figure 3.1: Normalized cross-spectra between the initial and final density fields as a function of scale. At small wavenumbers k (large scales) the two fields are perfectly correlated because the universe is not affected by gravitational collapse over such scales (they are both uniform). On the other hand, at large k (small scales) they are decorrelated because the initial field is very uniform, while the final field is very non-uniform on such scales (it contains large empty voids, and small and massive halos). The decorrelation scale moves to smaller k as time goes on due to the progressive collapse of larger and larger overdensities.

initial and the final fields (from Sim A) can be seen in Figure 3.1. For small wavenumbers k (corresponding to large scales), the correlation is very good (converges to 1: perfect correlation). On the other hand, for large wavenumbers (corresponding to small scales), the two fields are completely decorrelated.

Talk about the binning??

The small k convergence towards perfect correlation indicates that gravitational collapse does not have a large impact over such large scales. Because of this, both the initial and the final density fields tend to be very uniform, which preserves the correlation on these scales. However, over small scales, gravity has

a large impact. This results in a large discrepancy between massive collapsed regions and mostly empty voids. This is in stark contrast to the relative uniformity of the initial field, leading to breakdown in correlation.

Figure 3.1 also shows the evolution of this correlation with redshift. The wavenumber at which the two fields decorrelate indicates the progress of gravitational collapse at that redshift. This results in the decorrelation scale moving to smaller wavenumbers with the progress of gravitational collapse. The objective of reconstruction methods is to bring this decorrelation scale to larger k (in order to recover information about the initial field).

Talk about errors here

In order to measure the power-spectra of our reconstructed fields, we modified GENPK to read the fields generated by *pynbody*. The results of the perfect reconstruction can be seen in Figure 3.2, where we compare it with the original correlation at different redshifts.

3.3   Analysis

The cross-spectra presented in Figure 3.2 show a large improvement in the correlation with the initial field. There is also an increase in the amount of information recovered for lower redshifts. This means redshift does not play a role as large in the perfect reconstruction as it originally did.

However, in order to understand this perfect reconstruction, we need to look at the key role played by the resolution of the simulation. Figure 3.3 shows a comparison of the cross-spectra across the three simulations. For the original correlations, the size of the simulation plays a larger role than the resolution. Simulation C (smaller size) shows a smaller scale of decorrelation. Simulations A and B (same size) are very close, with a slight edge for simulation B (lower resolution).

better explina-
tion here

A completely different structure can be seen once we perform the perfect reconstruction. Simulation A and C (same resolution) show identical reconstructed correlation, while the reconstruction in Simulation B (lower resolution) does not perform as well. This indicates that resolution plays the decisive role in the per-
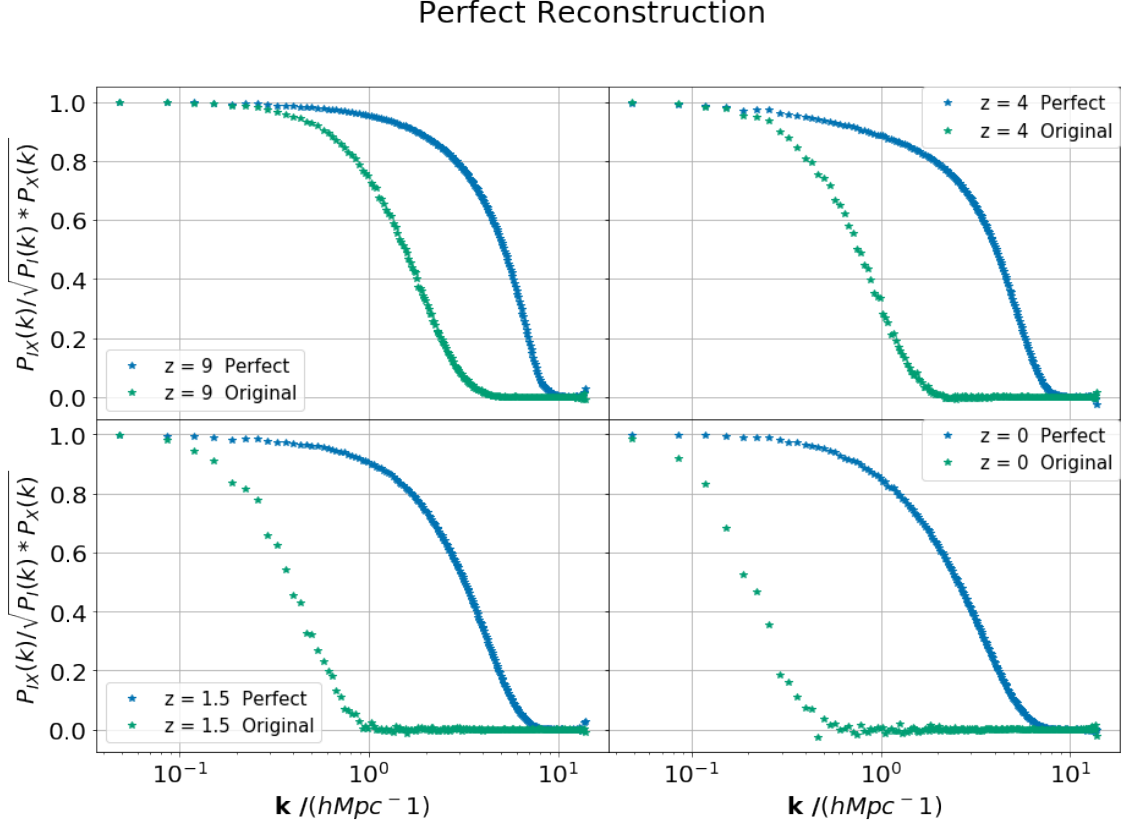
Figure 3.2: Normalized cross-spectra between the initial and the reconstructed fields compared to the original correlation. A large improvement in the correlation was achieved with the perfect reconstruction. This shows up as a shift of the decorrelation scale towards larger k (smaller scales). However, the perfect reconstruction does not lead to a perfect correlation due to the limiting resolution of our density field measurements. When comparing the reconstruction applied to fields at different redshift, we see a trend towards more information being recovered from smaller redshifts.
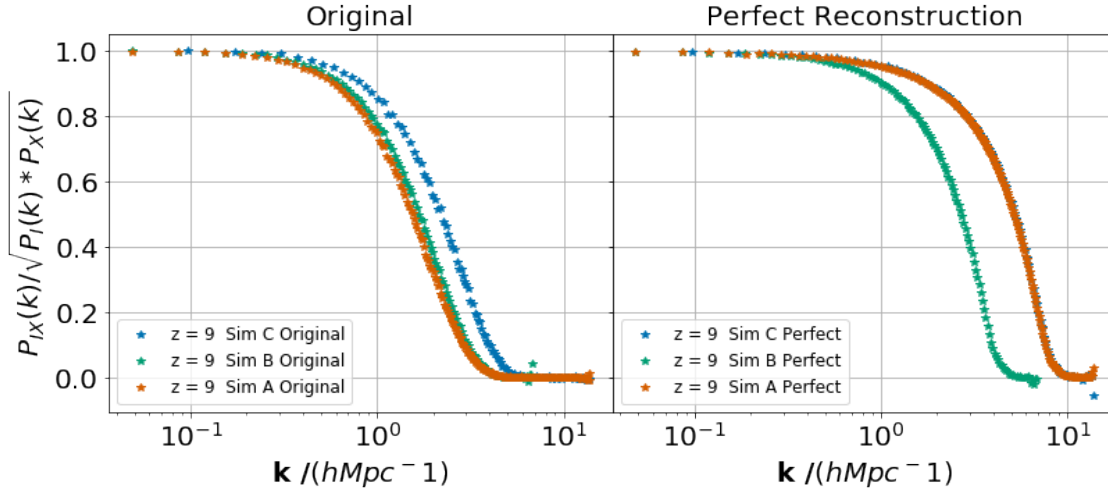
Figure 3.3: Normalized cross-spectra across the three simulations. The left plot shows the original correlation between the initial and the final fields, and the right panel shows the correlation with the reconstructed field. For the original correlations, the size of the simulation plays an important role, with the smaller simulation decorrelating on smaller scales. However, after the reconstruction, the size of the simulation does not seem to have any impact (with Sim A being almost identical to Sim C). In this case, the resolution of the simulation is the only factor that matters, with the larger resolution simulations showing a better correlation.

fect reconstruction. However, this was exactly the starting point of this chapter. The limiting factor for this reconstruction is the resolution used to measure the density field. The right panel in Figure 3.3 shows the lowest scale that can be reconstructed depending on the resolution of our simulation.

The perfect reconstructions in Figure 3.2 show that a perfect correlation cannot be achieved even in the ideal case of perfect knowledge of all the starting particle positions. However, a large improvement in the correlation can be seen, with the decorrelation scale moving to very small scales (of the order $1Mpc$). This ideal reconstruction using perfect knowledge of the particle positions serves as a theoretical upper limit to reconstruction techniques. The perfect reconstruction, along with the original correlation, will always be present in the next chapter when we look at realistic reconstructions. This can give us a better understanding to how well our techniques work.

Chapter 4

Towards Realistic Reconstructions

With the tools developed in the previous chapter, and the perfect reconstruction serving as an upper limit for reconstruction, we are ready to dive into realistic reconstruction methods. In Chapter 2 we outlined many reconstruction methods used in practice, based on both Standard Perturbation Theory and Lagrangian Perturbation Theory. In this project, we base our reconstructions on the first order approximation to LPT (the Zel'dovich Approximation).

hopefully

4.1 The Zel'dovich reconstruction

The key ingredient that we will use to perform realistic reconstructions is the density field. The idea behind the Zel'dovich approximation is to calculate a linear displacement field (we refer this field as Zel'dovich offset) based on the current peculiar velocities while taking into account the Hubble flow (see Chapter 2). *Pynbody* has tools for calculating the Zel'dovich offset using the particle velocities in a snapshot. This offset is given by:

ref???

$$\Psi_z(\mathbf{q}) = (1+z) \times \mathbf{v}(\mathbf{q}) \times \frac{D(z)}{f(z)} \tag{4.1}$$

Where $D(z)$ is the linear growth factor, $f(z)$ is the rate of linear growth and $\mathbf{v}(\mathbf{q})$ is the velocity field.

As we are interested in looking at the correlation between the reconstructed field and the initial fields in our simulations (which are at $z = 99$), we need to calculate the Zel'dovich offset up to $z = 99$ only. In order to achieve this, we

first used equation 4.1 to calculate the offset starting from the redshift z of the snapshot (Ψ_z). After that, we used the same equation to approximate this offset from $z = 99$ (Ψ_{99}) using the same velocity field. The displacement field we are after is then given by:

$$\Psi(\mathbf{q}) = \Psi_z(\mathbf{q}) - \Psi_{99}(\mathbf{q}) \tag{4.2}$$

We first start out in our investigation by performing a reconstruction using the Zel'dovich offset calculated directly from the particle velocities in each snapshot. The methodology of the reconstruction resembles the Perfect Reconstruction. We first calculate the density field at the particle positions in a snapshot, and then we apply the Displacement field $\Psi(\mathbf{q})$ to move the particles. The density field is carried along. After that, GENPK is used to measure the cross power-spectra of the reconstructed field with the initial ($z = 99$) field.

The results of this reconstruction (we call it the Zel'dovich reconstruction) can be seen in 4.1. We again look at the normalized cross-spectra between this reconstruction and the initial density field. To give us a better understanding of how well this method works, the original correlation and the perfect reconstruction are also present. The figure presents the reconstruction starting from four different snapshots in the redshift interval $z = 9$ to $z = 0$.

The reconstruction starting at $z = 9$ gives very good results, bringing the decorrelation scale to an intermediate step between the original and the perfect reconstruction. At this redshift most particles are still in the quasi-linear regime, so this result was expected. An interesting feature is the small anti-correlation obtained at large k. This effect is most likely due to particles in non-linear regimes which are past shell-crossing. To understand what gives rise to this anti-correlation, consider two fronts of matter collapsing towards each other. After shell crossing there will be a turn-around as the two evolve into a single filament. If we linearly track these velocities back, we are effectively going the wrong way. This will lead to an anti-correlation over the affected scales. For the $z = 9$ reconstruction, this effect is very small, indicating that shell crossing has only occurred on the smallest scales, and that most particle motions can be well approximated with the linear regime.

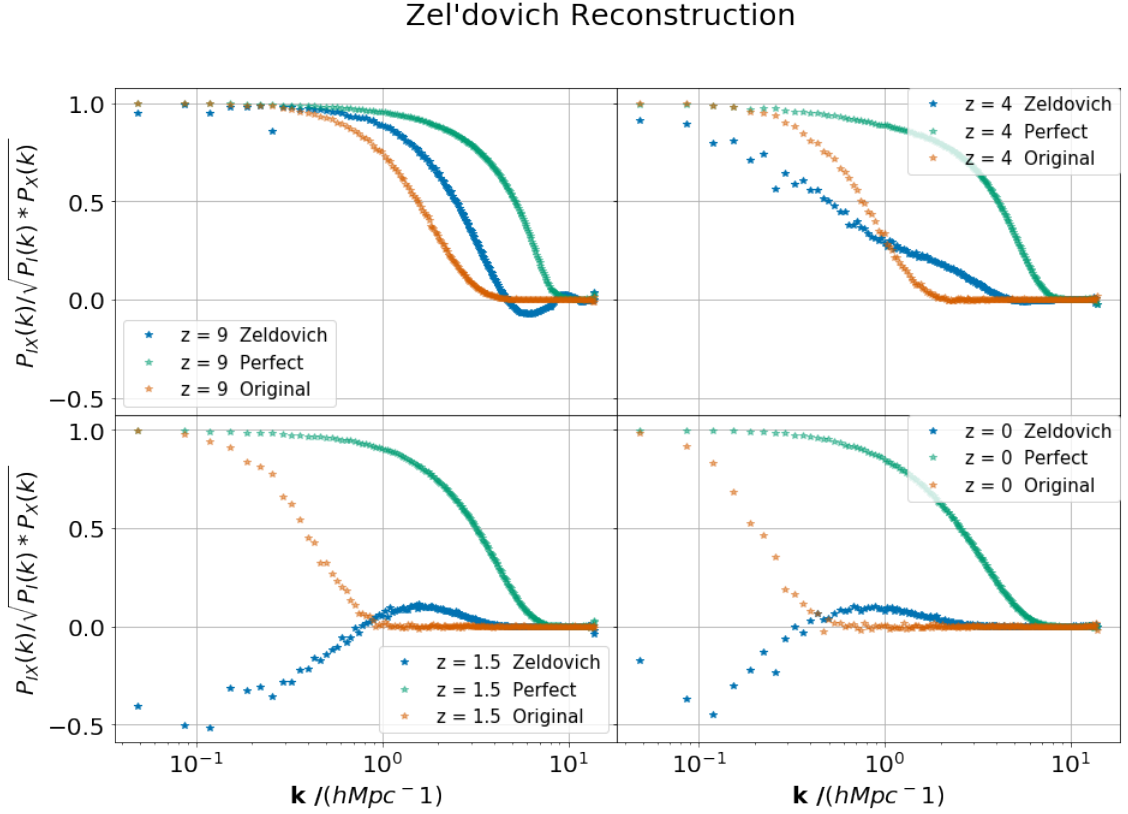this should be explained in chapter 2.

Zel'dovich Reconstruction

Figure 4.1: Normalized cross-spectra between the Zel'dovich reconstruction and the initial field. This reconstruction was performed by linearly moving the particles back in time (using the Zel'dovich approximation). As we apply the linear approximation directly to the particle velocities, this gives a good indicator of the regime we are in at that redshift. We see the reconstruction work very well when starting from $z = 9$, which indicates we are still in the quasi-linear regime there. However, as time progresses (lower redshift) the correlation breaks down even on the largest scales, indicating we are mostly in the non-linear regime.

However, starting from lower redshifts produces results much worse than even the original. This was expected, as at these redshifts most particles are now in the non-linear regime. By still treating their motions as linear we are breaking even the correlation that was there to begin with. Figure 4.1 shows the largest scales decorrelating as we move to reconstructions from lower redshifts, and even leading to anti-correlation. For the lowest redshifts, we see a small improvement in the correlation on intermediate scales, but anti-correlation on large scales. This result is much harder to understand. A possible explanation is an extension to the reasons presented above for the small scales in the $z = 9$ reconstruction. The effect of anti-correlation due to particles that are past shell crossing moving the wrong way within the Zel'dovich approximation is expected to increase with decreasing redshift. This might lead to the large scales also becoming anti-correlated. This effect should be studied further, however, we leave this for future works, as our aim in this project is to achieve a good reconstruction.

The Zel'dovich approximation is then not a good reconstruction method at low redshift when most particles have non-linear velocities. For these regimes we need higher orders of LPT to perform the reconstruction. However, at this point it is hard to justify this pursuit from an observational stand point. In this section we have used the peculiar velocities of particles in our simulations. As the ultimate goal of any reconstruction technique is to be used in practice on real data, we need to consider the feasibility of our method. Observers usually detect a few galaxies over an $8 - 10$ Mpc scale, and any peculiar velocity measurement inevitably come with errors. This means that the Zel'dovich reconstruction we just performed is very unrealistic in practice. The goal of the rest of this chapter is to modify the Zel'dovich reconstruction to make it more realistic, and also to improve its performance at low redshift.

4.2 Getting back to the linear regime

In order to make the Zel'dovich approximation work for our reconstruction, we must somehow get back into the linear regime. As discussed in Chapter 2, matter tends to be collapsed into filaments at late times. This means individual

particles have non linear velocities, but ensembles of particles might still be in the linear regime. Our solution to the two problems outlined in the previous section is to use bulk velocities to calculate the Zel'dovich offset, instead of individual particle velocities.

We smooth particle velocities over 1 Mpc and 10 Mpc scales respectively before calculating the Zel'dovich offset. This means we are now considering bulk motions instead of particle motions. These bulk motions will hopefully provide a better start point when we calculate the Zel'dovich offset. This smoothing also improves the realism of our method. With the technology we currently have, observers can maybe detect a few galaxies in a 10 Mpc bin, so by smoothing our velocity field over that scale, we simulate a more realistic scenario. The reason for attempting a separate reconstruction using velocities smoothed over 1 Mpc scales is twofold. Firstly, we want to understand the effect of the velocity smoothing scale on the reconstruction. Secondly, we use the 1 Mpc case as a test for what could be achieved with improving technology and a better handling of systematics which could be useful for the next generation of Galaxy Surveys.

To perform these reconstructions, we first split a simulation into bins of a given size: $(1Mpc)^3$ or $(10Mpc)^3$. We then use the positions of the particles to identify the bin they are in. After that, an average velocity over the particles in each bin is calculated. This average velocity is assigned to the centre of the bin. In this manner, we construct a three-dimensional grid which contains a measure of the average velocity field. Finally, we use this average velocity field to linearly interpolate the value of the velocities at the particle positions. In this manner, velocities are smoothed over the scales of interest. Using these new velocities, the Zel'dovich reconstruction is performed as outlined in the previous section.

To further improve the practicality of our method, we also attempt a second type of velocity smoothing. We perform all the steps outlined above to create an average velocity field, but this time we use smaller bins: $(0.5Mpc)^3$ in size. We then use a Gaussian Filter to smooth this field over the scales of interest (1 Mpc and 10 Mpc respectively). From here, the procedure carries on as outlined above. By first measuring the average velocity field in $(0.5Mpc)^3$ bins, we are closer to what an observer would detect. Galaxies are at most a few hundred thousand

Kpc across . Even though we may be able to measure velocities on these small [find ref] scales, we only detect a few galaxies on 10 Mpc scales. Also, on these small scales, [ref] velocities are most likely still in the non linear regime. This leads us to perform the Gaussian smoothing over larger scales.

Before we move on to the results, an interesting side effect that should be mentioned showed up during the reconstruction. The nature of our method implies that we are creating coherent movements of particles. This coherent movement leaves large gaps in our reconstructed density field (regions where the density field is equal to 0). These gaps become a problem when we want to take the logarithm of the density field as discussed in Chapter 3. When calculating the density field, *Pynbody* uses a smoothing kernel which normally fills in these gaps. However, this method uses N nearest neighbours to calculate the smoothing scale in a region. Normally, there still are a few particles even in the largest voids. These particles will have very far away neighbours, imposing a large smoothing scale. On the other hand, by creating coherent movements, completely empty regions arise. Particles on the edges of these empty regions can easily find nearby neighbours and establish a relatively small smoothing scale. Our solution is to manually find these empty regions and assign a very small value to the density field.

4.3 Results

The first step in our investigation is to understand the effect of the velocity smoothing scale on the reconstruction. In section 4.1 we found that the Zel'dovich reconstruction works very well when starting from $z = 9$. We expect our new methods to have a similar performance when starting from this quasi-linear regime.

Figure 4.2 shows the impact of the velocity smoothing on the $z = 9$ Zel'dovich reconstruction. In this case we only present the first method of calculating an average velocity over 1 Mpc and 10 Mpc scales. For brevity we refer to them as 1 Mpc reconstruction and 10 Mpc reconstruction. As expected, when starting at $z = 9$, both methods work in reconstructing the density field. However, the 1 Mpc reconstruction performs better than the 10 Mpc one. This demonstrates a further loss of information when smoothing the velocity field. The 1 Mpc re-
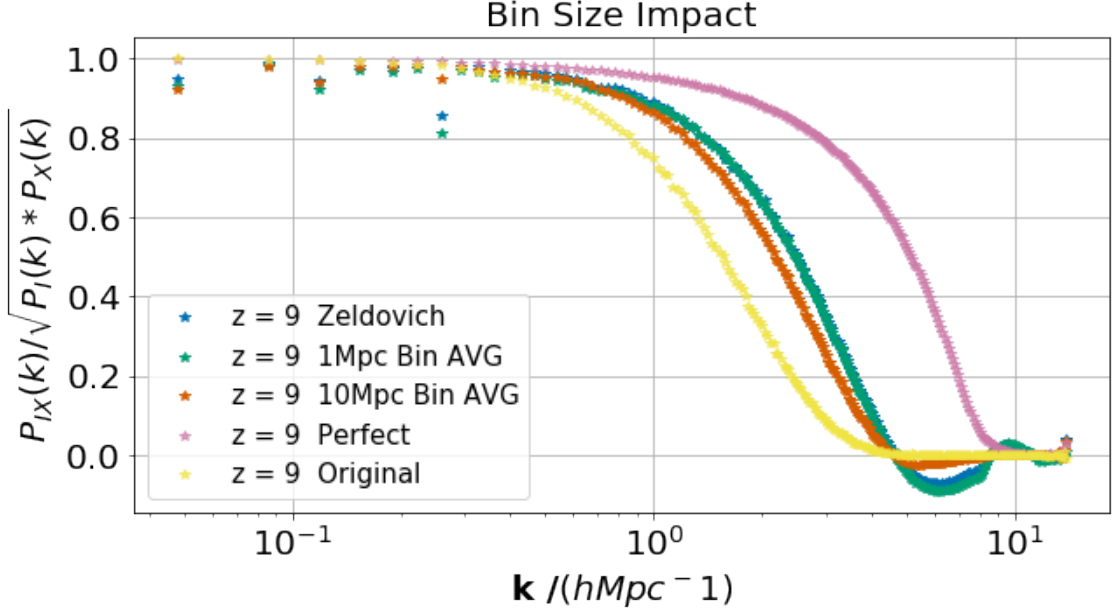
Figure 4.2: Normalized cross-spectra of realistic reconstructions with the initial field. This figure shows the impact of smoothing velocities over certain scales. Because we lose information by smoothing, the 10 Mpc reconstruction recovers less information compared to the 1 Mpc reconstruction. The Zel'dovich reconstruction is also present for comparison. As we are looking at reconstructions from $z = 9$, they all work quite well because most particles are still in the quasi-linear regime.

construction produces a correlation which is very close to the original Zel'dovich reconstruction.

Once we know the impact of the smoothing scale and that reconstruction still performs very well when starting from $z = 9$, we now want to compare their performance when starting from $z = 0$. Figure 4.3 shows the correlation between the $z = 0$ field, reconstructed with the two methods, and the initial field.

A normal velocity average over the interest scales produces results similar to what we found in section 4.1 with the Zel'dovich reconstruction. We recover some information on intermediate scales, but the large scales become anti-correlated. Here this seems to be taken to an extreme, in the sense that we have almost perfect anti-correlation on the largest scales. This result is very interesting and should be studied further in the future.

On the other hand, when we use smaller averaging bins and a Gaussian filter, we get a much better correlation. We recover quite a bit of information over intermediate to large scales. However, the largest scales still seem to decorrelate.
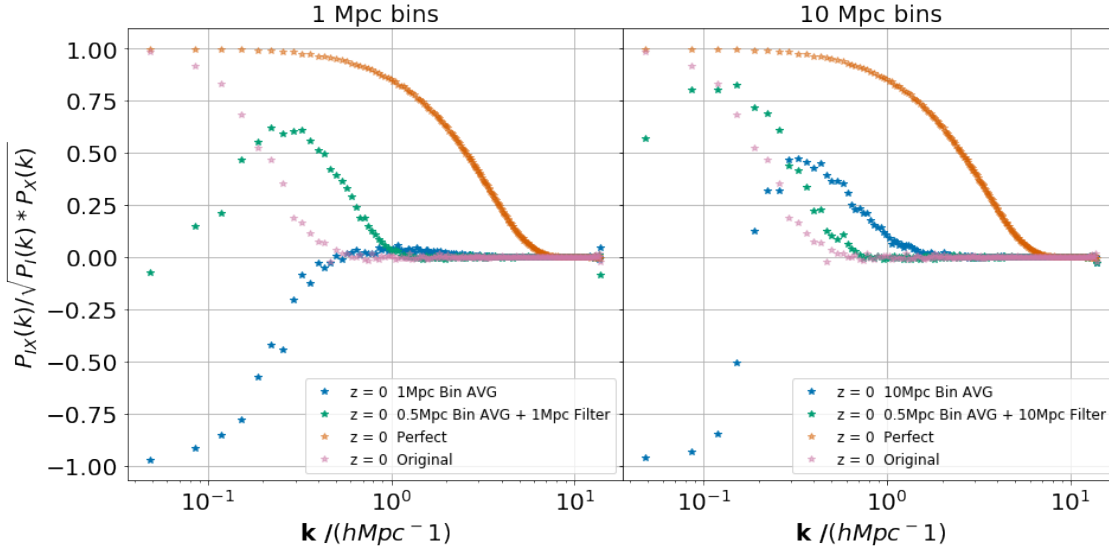
Figure 4.3: Normalized cross-spectra of fields reconstructed from $z = 0$ with the initial density field. We compare the two methods of smoothing velocities outlined in section 4.2 for the two scales of interest. The left panel shows the comparison between the performance of the two methods when velocities are smoothed over 1 Mpc scales, while the right panel shows 10 Mpc scales. If we just average velocities over those scales, we still obtain an anti-correlation over the largest scales. However, when averaging over 0.5 Mpc scales and then applying a Gaussian Filter over the required scales, we see a large improvement in the reconstruction. An interesting effect is that for the 10 Mpc reconstructions, the first method slightly outperforms the second on intermediate scales.

We will investigate this further in the next section where we compare results from different simulations.

We found that using a Gaussian filter to smooth velocities works much better than normal averaging within the Zel'dovich reconstruction. This indicates that just averaging velocities in a 1 Mpc or even a 10 Mpc bin is not enough to bring us back into the quasi-linear regime. We do see an improvement in the reconstruction of intermediate scales when going from 1 Mpc to 10 Mpc averaging, but it still leads to an anti-correlation of the large scales.

An interesting outcome of these procedures is also the fact that for a small region, on intermediate scales (for 10 Mpc reconstructions), we do obtain a better correlation using normal averaging. This is probably an indication that we are losing more information when we smooth the field with a Gaussian filter due to the tail of the function. However, this tail is most likely also the reason this method works up to much larger scales. By taking into account velocities of particles further away than the scale of interest (even though they have a small weight) we are getting closer to the quasi-linear regime.

4.4   Analysis

We now turn our attention to how these methods work when we apply them to different simulations. We have so far only looked at results from Simulation A as it's the largest one and it has the best resolution. Figure 4.4 presents the correlation of the $z = 9$ reconstructed field with the initial density field in Simulation C and Simulation A. It shows that we are recovering information up to a smaller scale in the smaller simulation (C). This is not entirely surprising, as the smaller simulation is better correlated to begin with (as discussed in Chapter 3). What is surprising is how close to the perfect reconstruction the correlation gets.

The best overall correlation when reconstructing the $z = 0$ density field was achieved in Simulation B. This result is shown in Figure 4.5. Both the 1 Mpc and the 10 Mpc reconstructions achieve a better correlation than the original on most scales. This simulation also gives us the best solution from a practical standpoint. When we calculate the average velocity field using $(0.5 Mpc)^3$) bins we have, on
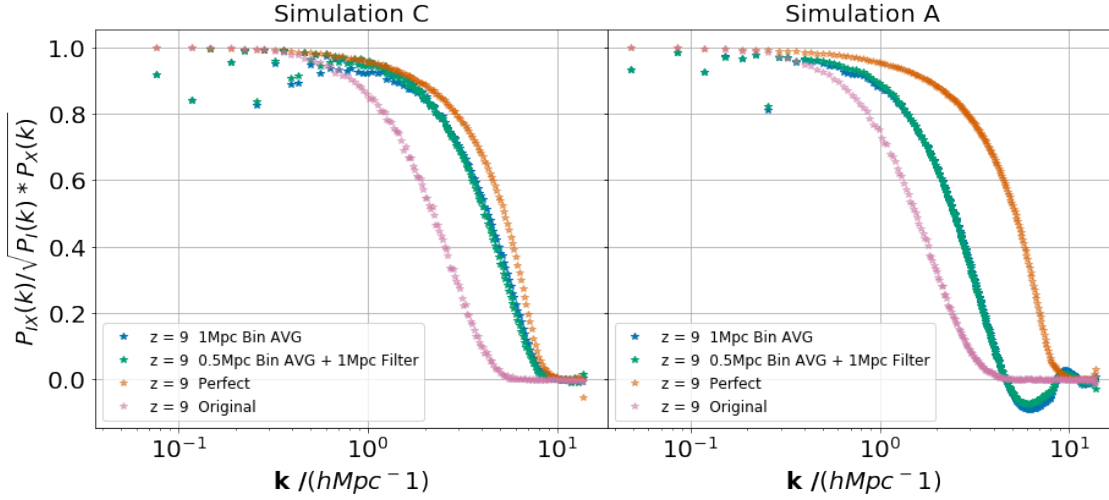
Figure 4.4: Normalized cross-spectra of reconstructed $z = 9$ fields with the initial field. We compare the same reconstruction methods (1 Mpc smoothing only) applied to two different simulations. On the left is Simulation C with a size of $(100 Mpc)^3$ and on the right is simulation A with a size of $(200 Mpc)^3$. The reconstruction in the smaller simulation recovers a lot more information, bringing the correlation very close to the perfect reconstruction.

average, about 0.25 particles per bin. This means that most of our bins are actually empty, which for such small bins is exactly the case for observers as well. This is in contrast with simulation A, which has about 2 particles per bin.

The best results for Simulation A are presented in Figure 4.6. Compared to Simulation B, we recover more information on intermediate scales. However, the largest scales decorrelate. There is also a larger difference between using a 1 Mpc and a 10 Mpc smoothing range. When using a 10 Mpc filter, we recover less intermediate scale information, but the large scales are much better correlated. This is a very good showcase for a trend we have been seeing in this chapter. When we choose to smooth the velocity field we are giving up some information in the hope of recovering the linear regime. The more velocity information we use, the more we destroy the large scale correlation by using the Zel'dovich approximation. However, using more velocity information generally leads to a better reconstruction on intermediate scales. This is an interesting information trade-off. The more intermediate scale information we recover, the more large-scale information we lose.

On the other hand, based on these results, a case could be made that by using less velocity information, we are just pushing the problem to larger scales. We
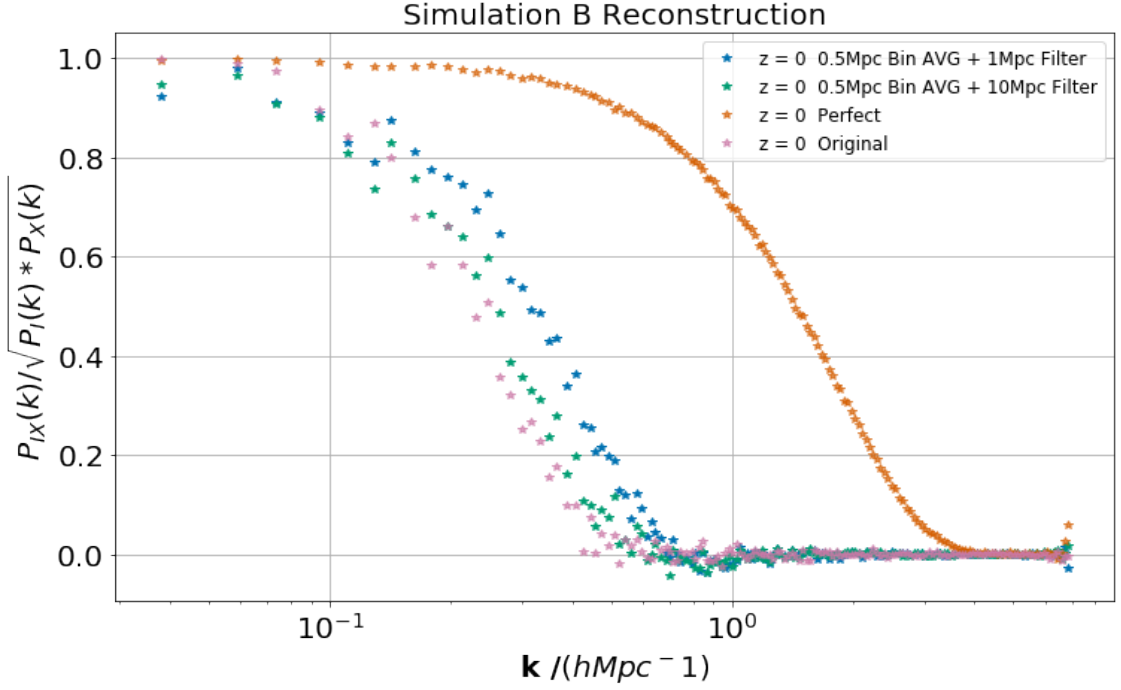
Figure 4.5: Normalized cross-spectra of the $z = 0$ reconstructed fields with the initial density field within Simulation B. Our reconstruction methods work very well in this low resolution simulation in the sense that we don't destroy the correlation that was there to begin with (the largest scales). We also recover some information on intermediate scales, however, it is very far away from the perfect reconstruction. Also, we again encounter the effect found in section 4.3: we recover less information when we smooth velocities over 10 Mpc compared to 1 Mpc.
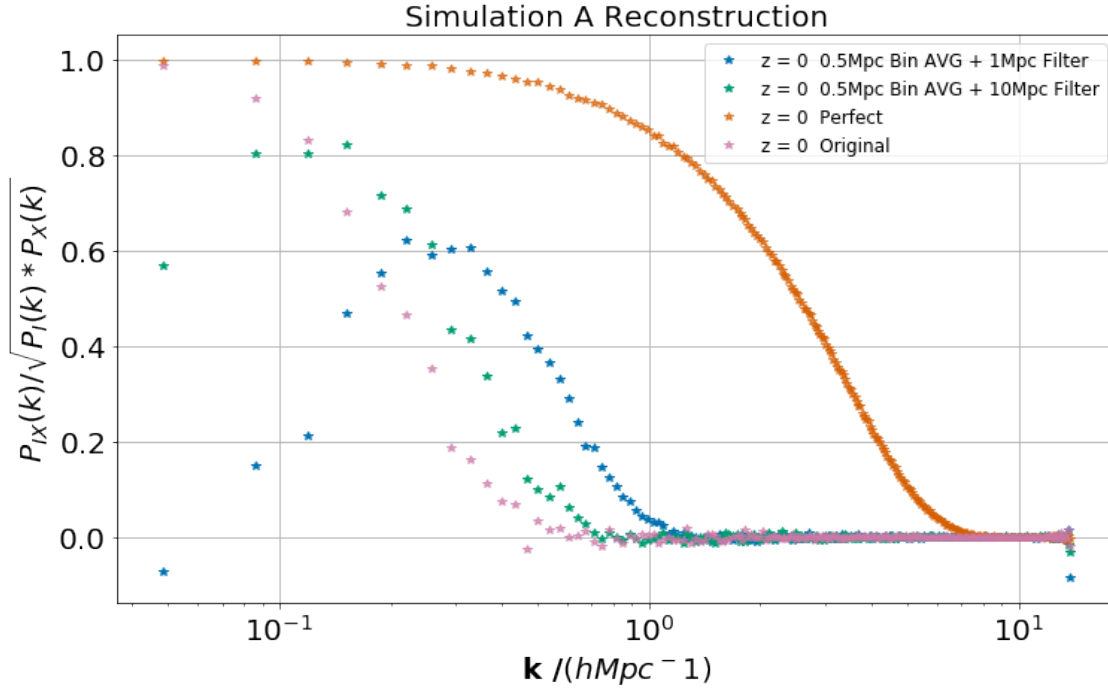
Figure 4.6: Normalized cross-spectra of the $z = 0$ reconstructed field with the initial density field within Simulation A. By smoothing velocities over 10 Mpc scales, we achieve a better correlation of the large scales, however we have a worse correlation over intermediate scales. The result for the intermediate scales was discussed in section 4.3. It is an outcome of the information lost when smoothing velocities. The interesting result is that within the Zel'dovich approximation, we manage to preserve the large scale correlation only by using less velocity information.

still see the largest scales start to decorrelate even when using the Gaussian filter. If we had a much larger simulation with the same resolution we might still see the largest scales become anti-correlated. However, our initial goal was to study reconstruction on scales smaller than the BAO. In that respect, our methods do succeed, as we see an improvement in correlation on intermediate scales.

These results are still quite far away from the correlation we obtain with the perfect reconstruction. However, we only studied reconstructions within the Zel'dovich approximation. Higher orders of Lagrangian Perturbation Theory (for example 2LPT) might be able to get us closer the perfect reconstruction, and also solve the problem with the information trade-off.

Chapter 5

Conclusions

5.1 Information loss

Talk about the inevitable information loss and the big discrepancy between the perfect and realistic reconstructions.

5.2 Future Work

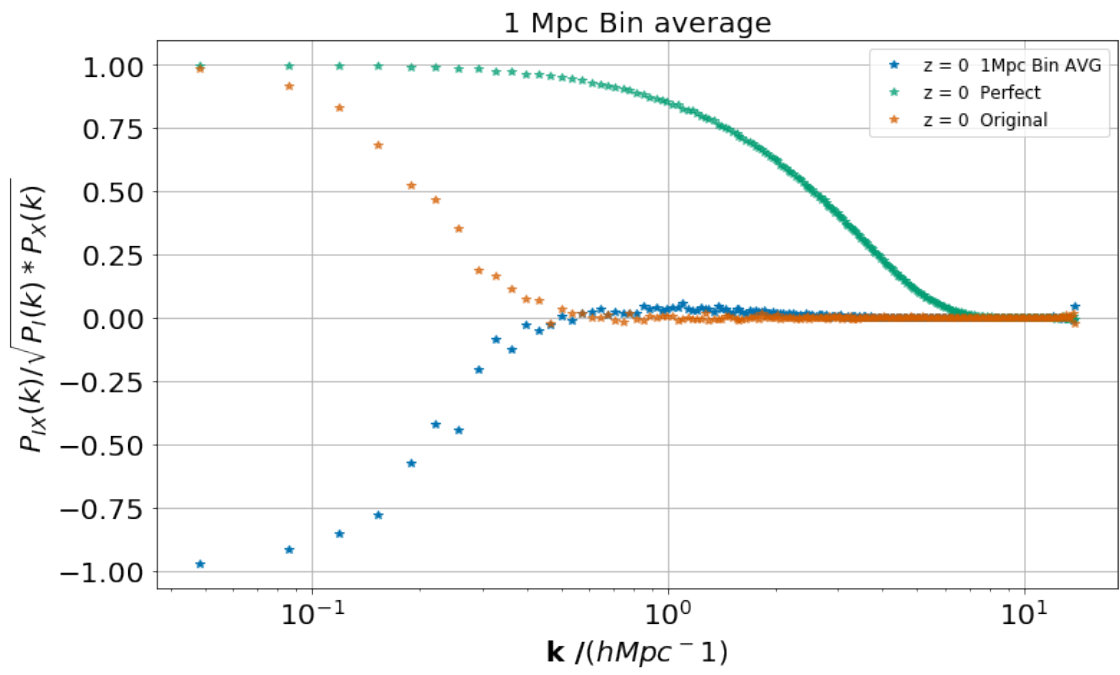Talk about the problems encountered and Future Work.

Figure 5.1: Simulation A reconstruction

Bibliography

Bird, S. (2017). *GenPK: Power spectrum generator*. Astrophysics Source Code Library. ascl: `1706.006`.

Conselice, C. J. et al. (2016). "The Evolution of Galaxy Number Density at z = 8 and Its Implications". In: ApJ 830, 83, p. 83. DOI: `10.3847/0004-637X/830/2/83`. arXiv: `1607.03909`.

Eisenstein, D. J. et al. (2007). "Improving Cosmological Distance Measurements by Reconstruction of the Baryon Acoustic Peak". In: ApJ 664, pp. 675–679. DOI: `10.1086/518712`. eprint: `astro-ph/0604362`.

Planck Collaboration et al. (2016). "Planck 2015 results. XIII. Cosmological parameters". In: A&A 594, A13, A13. DOI: `10.1051/0004-6361/201525830`. arXiv: `1502.01589`.

Pontzen, A., R. Roškar, et al. (2013). *pynbody: N-Body/SPH analysis for python*. Astrophysics Source Code Library. ascl: `1305.002`.

Pontzen, A., A. Slosar, et al. (2016). "Inverted initial conditions: Exploring the growth of cosmic structure and voids". In: Phys. Rev. D 93.10, 103519, p. 103519. DOI: `10.1103/PhysRevD.93.103519`. arXiv: `1511.04090`.

White, M., ed. (1999). *Anisotropies in the CMB*. eprint: `astro-ph/9903232`.