

Title Here

Andrei Cuceu

Supervisor: Dr. Andrew Pontzen

University College London

March 23, 2018

Abstract

Eisenstein et al. 2007

Contents

1	Introduction	5
1.1	Cosmological Context	5
1.2	The Cosmic Microwave Background	5
1.3	Large Scale Structure and Galaxy Surveys	5
1.4	The Missing Link (Reconstruction)	5
2	The Growth of Structure	8
2.1	Perturbation Theory	8
2.2	Linear vs Non-Linear Collapse	8
2.3	The Zeldovich Approximation	8
2.4	Reconstruction (BAO)	9
3	The Perfect Reconstruction	10
3.1	Methods	10
3.2	From Images to Statistics	11
3.2.1	The Reconstructed Density Field	11
3.2.2	Correlation with the Initial Field	12
3.3	Analysis	15
4	Towards Realistic Reconstructions	17
4.1	The velocity field	17
4.2	Getting back to the linear regime	17
4.3	Results	17
4.4	Analysis	18

5	Conclusions	19
5.1	Information loss	19
5.2	Future Work	19

Chapter 1

Introduction

1.1 Cosmological Context

Introduce the context of Modern Cosmology

1.2 The Cosmic Microwave Background

Introduce the CMB and talk about the starting point. E.g. very smooth initial field with some anisotropies that will be locked in the matter distribution.

Introduce BAOs.

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)

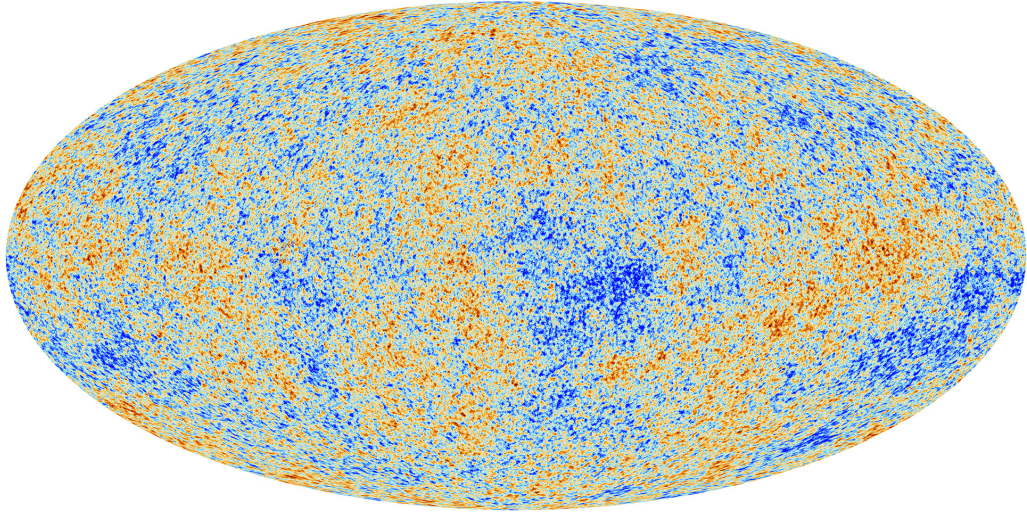


Figure 1.1: Map of the Cosmic Microwave Background acquired by the Planck Space Telescope (ref).

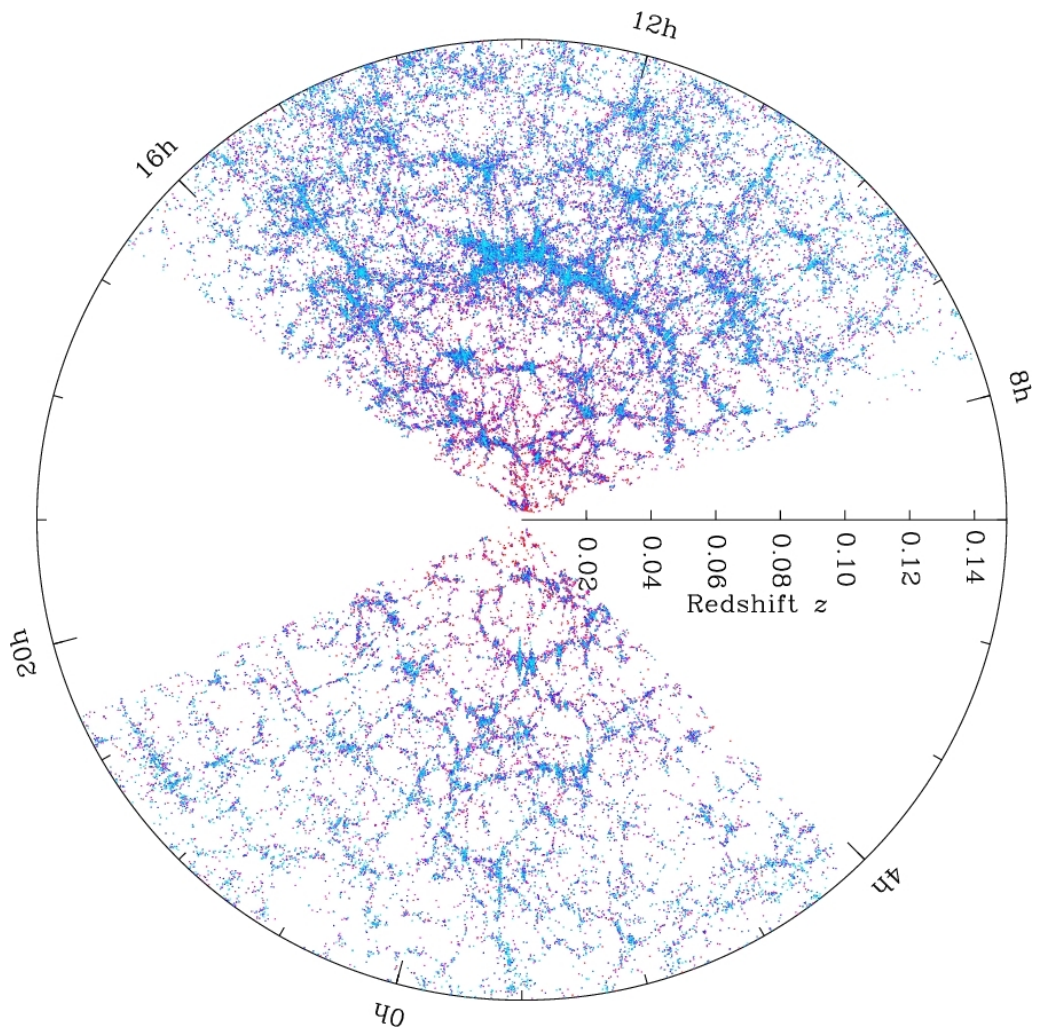


Figure 1.2: Galaxy map from the Sloan Digital Sky Survey.

Motivate the desire to reconstruct the BAO feature

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 advantages 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.

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

Showcase the BAO reconstruction.

this paragraph
could go in
chapter 2

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 hasn't recovered all the information, as it's not identical to the initial field. 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 values orders of magnitude larger than the initial field.

get number

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 don't 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

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:

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

The power-
spectrum will
be introduced
in chapter 2

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

We used the GENPK code² (Bird 2017) to measure auto and cross power-spectra of GADGET outputs. The original normalized cross-spectra between the 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 (con-

²<https://github.com/sbird/GenPK.git>

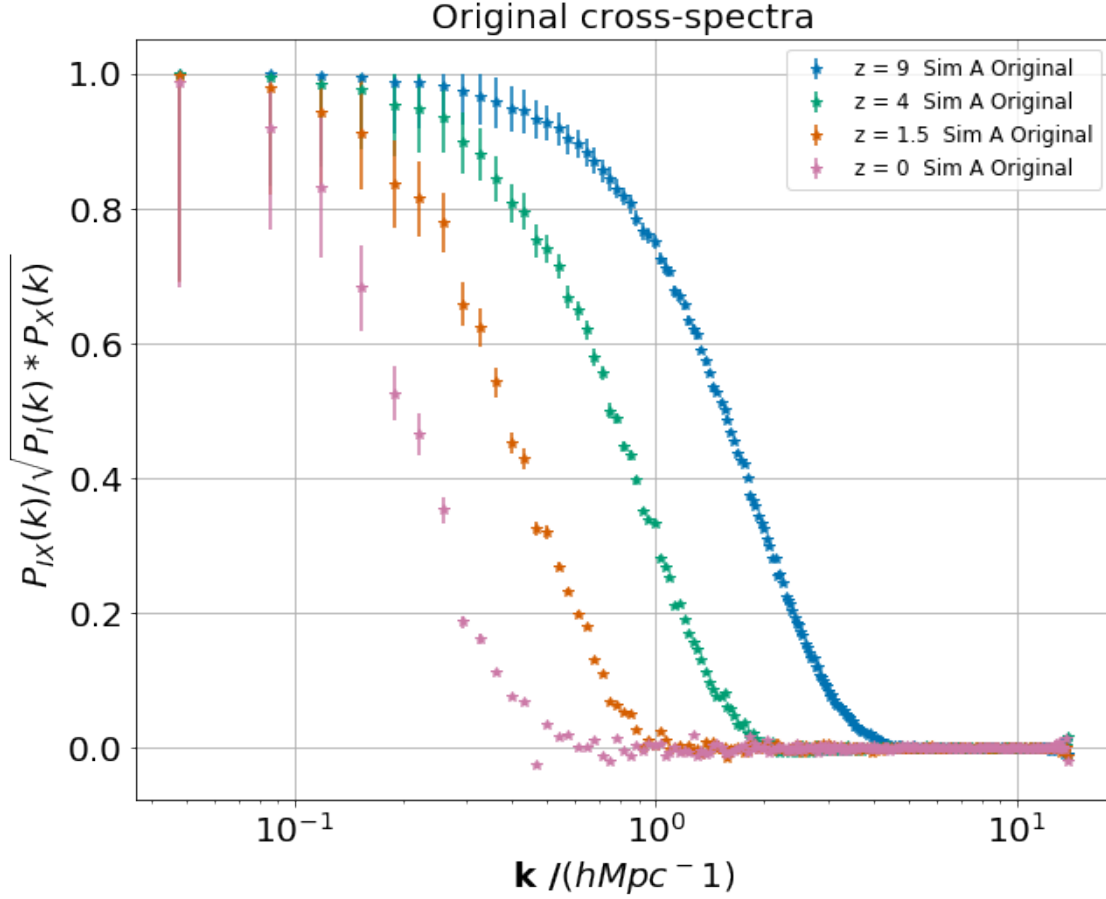


Figure 3.1: Normalized cross-spectra between the initial and final density fields as a function of scale. At small wavenumbers k (large scales) the two fields are perfectly correlated because the universe is not affected by gravitational collapse over such scales (they are both uniform). On the other 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.

verges to 1: perfect correlation). On the other hand, for large wavenumbers (corresponding to small scales), the two fields are completely decorrelated.

The small k convergence towards perfect correlation indicates that gravitational collapse doesn't 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.

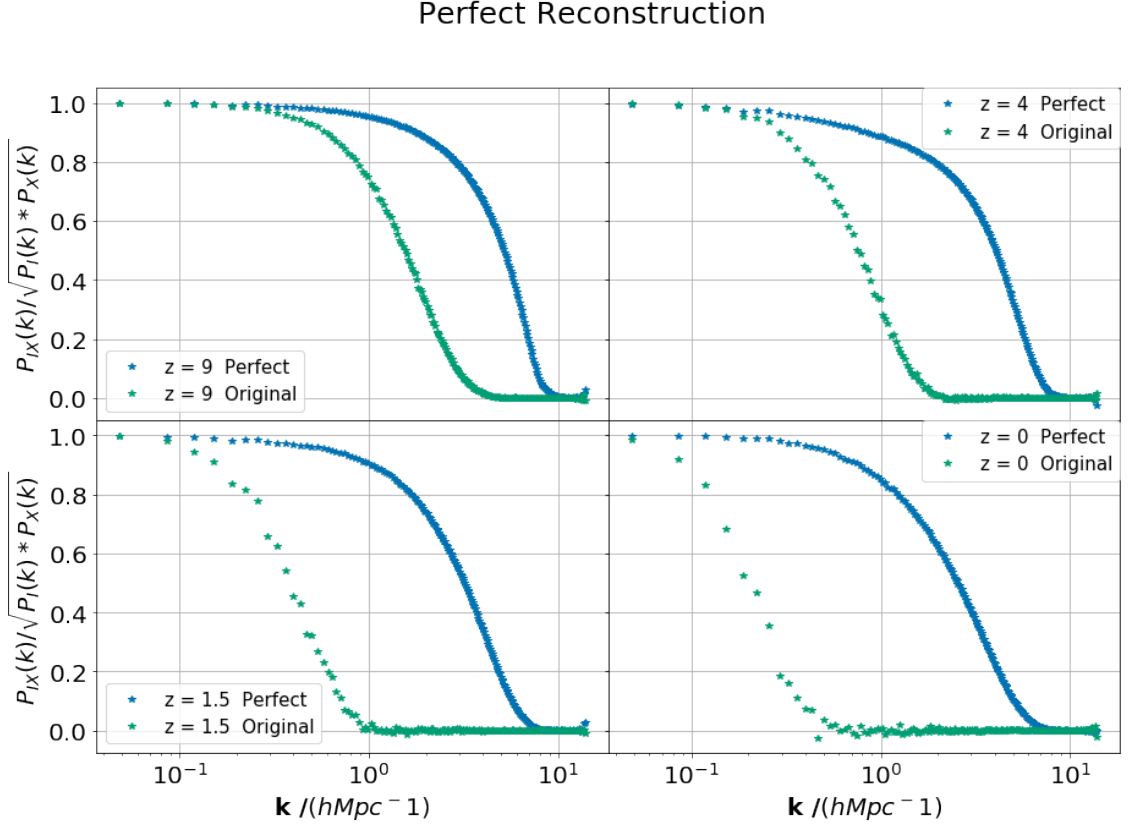


Figure 3.2: Normalized cross-spectra between the initial and the reconstructed fields compared to the original correlation. A large improvement in the correlation was achieved with the perfect reconstruction. This shows up as a shift of the decorrelation scale towards larger k (smaller scales). However, 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 3.1 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 be 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 doesn't 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).

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) doesn't perform as well. This indicates that resolution plays the decisive role in the perfect reconstruction. This was however the starting point of this chapter.

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.

better explanation here

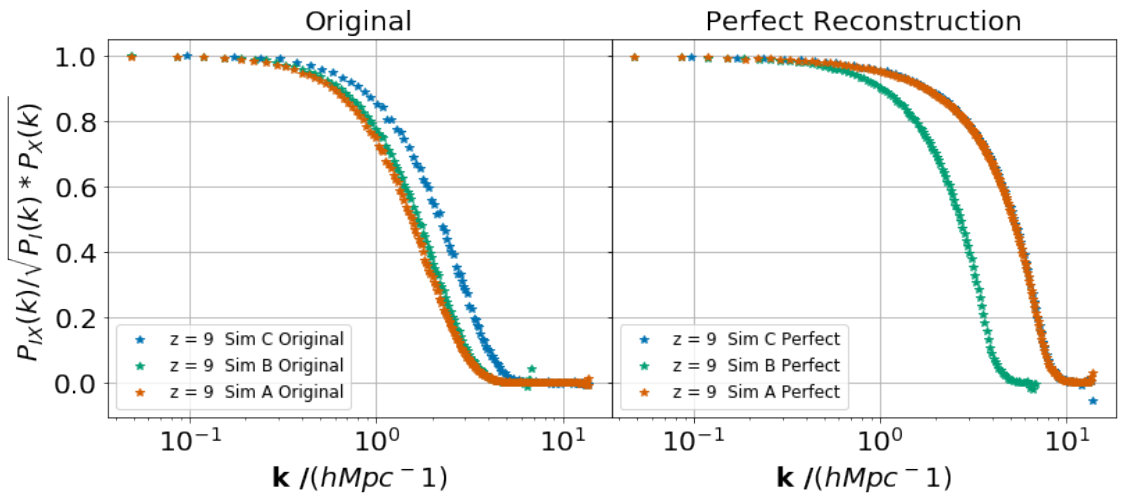


Figure 3.3: Normalized cross-spectra across the three simulations. The left plot shows the original correlation between the initial and the final fields, and the right panel shows the correlation with the reconstructed field. For the original correlations, the size of the simulation plays an important role, with the smaller simulation decorrelating on smaller scales. However, after the reconstruction, the size of the simulation doesn't 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.

Chapter 4

Towards Realistic Reconstructions

4.1 The velocity field

Introduce our use of the velocity field to do the reconstruction. Present method of obtaining this in practice and the problems encountered.

4.2 Getting back to the linear regime

Talk about how we try to get back to the linear regime. Introduce different ways of averaging the velocities

4.3 Results

Present our results - Density Slices and Cross Spectra.

4.4 Analysis

Talk about the results for different averaging methods and Bin Sizes.

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.

Bibliography

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