TRIANGULATING THE DARKNESS
Topological dark energy differentiation

Keimpe Nevenzeel
Supervisor: prof. dr. Rien v.d. Weygaert
Kapteyn Astronomical Institute / University of Groningen

December 2013

Abstract

This study consists of a thorough application of a new mathematical formalism to analyse the topological and morphological nature of the cosmic matter distribution: homological discrete topology. As an illustration of its potential, we apply it to a set of simulations in different dark energy cosmologies. The analysis reveals that the topology of cosmic structure is not influenced intrinsically by dark energy. However, various dark energy models have different rates of structure evolution. We show that the topology of cosmic structure at a specific redshift value depends on the dark energy content. For low redshifts ($z \lesssim 1$) this effect can be used to differentiate between various dark energy models both in real and in redshift space. This paves the way for future constraints on dark energy based on the topology of the cosmic web.

Homological discrete topology has been introduced in topology only very recently. Here, we provide a reference model with an extensive description of its topology. We show that (i) the method is robust under perturbations and (ii) the topological variant of smoothing, called simplification, is a natural and better equivalent to conventional smoothing.

Finally, we provide a detailed description of both discrete topological theory within an astrophysical context and its algorithmic implementation.

PACS: 98.65.Dx and 07.05.Kf.
MSC: 57Q15, 85-08 and 85A40.

Acknowledgements

Writing this thesis would not have been possible without the support of many people. Rien, I have had the sincere honour and pleasure to have you as a supervisor. You are a scientist in the classical tradition: not only do you have an immense knowledge of astronomy, you also have an exemplary knowledge of many other disciplines including mathematics, programming, history and philosophy. Besides that, you supervise projects in an inspiring way and have a vibrant cosmology group at the frontier of mathematical cosmology. I consider myself lucky to have been part of it. Specifically from this group, I would like to thank Pratyush Pranav, Marius Cautun and Patrick Bos for sharing their code or simulations and answering questions about it. Together with Job Feldbrugge and Mathijs Wintraecken, thank you as well for the useful discussions at various stages of this thesis.

As my thesis involved running many very heavy simulations, I often ended up requiring assistance of the computer group. Eite, Martin and Wim: thanks for all the times you helped me out, on occasions even on weekend evenings.

Thank you all my friends in and around student office 134, past and present, for all the fun, discussions and support throughout the years. Due to you the office didn't feel like an office but more like a kind of living room. In general, thank you Kapteyn Institute for providing a warm and personal environment to study in. And thanks for giving the freedom for extracurricular personal development. Both scientifically in the form of many additional mathematics and physics courses, but also non-scientifically in the form of teaching, organizational and outreach activities. I am sure both will help me a lot throughout the rest of my life.

I am indebted to my parents and sister for their continuous support in many ways throughout my study years. Mum, Dad and Evelien, without your support this would not have been possible. Finally, thank you Fabienne for everything during the last months of this thesis.

CONTENTS

1 Introduction **7**

2 Cosmology and the Cosmic Web **10**
2.1 The geometry of the Universe 11
2.2 The Laws of the Universe 12
2.3 Universal ingredients 14
2.4 Dark energy in more detail 16
2.5 Cosmic structure 18
2.5.1 The seeds of cosmic structure 19
2.5.2 Linear structure formation 21
2.5.3 Non-linear structure formation: Zel'dovich 22
2.6 Cosmic web morphology 26
2.6.1 Numerical morphology identification 26
2.6.2 Statistics 26
2.6.3 Minkowski functionals 27
2.6.4 Homological discrete topology 27

3 Topology **29**
3.1 Continuous Morse Theory 31
3.2 Discretization: simplicial complexes 37
3.3 Voronoi diagrams and Delaunay triangulations 40
3.4 Discrete Morse Theory 42
3.5 Practical implementation 45
3.6 Towards the persistence diagram 47
3.6.1 Cycles and persistence 47
3.6.2 Simplification 51
3.6.3 The persistence diagram 52

4 Algorithms and software **55**
4.1 Gadet 2: model universes 56
4.1.1 N-body algorithms 56
4.1.2 The Gadget 2 code 59
4.1.3 This Gadget implementation 60

4.2 DTFE: from particles to densities . . . 60
4.2.1 From particles to densities . . . 62
4.2.2 The DTFE algorithm . . . 64
4.2.3 This DTFE implementation . . . 65
4.3 DMC and filtration builders . . . 65
4.3.1 Density DMC's and filtrations . . . 65
4.3.2 Particle DMC's and filtrations . . . 66
4.4 PHAT: computing persistence pairs . . . 66
4.4.1 Persistent homology computation . . . 67
4.4.2 Persistent homology with PHAT . . . 70
4.4.3 This PHAT implementation . . . 70
4.5 Data analysis . . . 70

5 Stability 71
5.1 Mathematical results . . . 71
5.2 Experimental stability - perturbation analysis . . . 75
5.2.1 Particle perturbation analysis . . . 76
5.2.2 DTFE perturbation analysis . . . 77
5.3 Nicing persistence diagrams . . . 78
5.3.1 Smoothing the input field . . . 79
5.3.2 Manifold simplification . . . 81
5.3.3 Persistent Betti numbers . . . 89
5.4 Summary . . . 90
5.5 Discussion and conclusions . . . 92

6 On the persistence of LCDM 94
6.1 The reference case: LCDM at $z = 0$ . . . 94
6.2 Time evolution . . . 101
6.3 Summary and conclusions . . . 105

7 Topological dark energy differentiation 106
7.1 Real space identical redshift . . . 107
7.1.1 Visual inspection . . . 108
7.1.2 The effects of simplification . . . 108
7.1.3 Statistical analysis / test setup . . . 108
7.1.4 Statistical analysis / results . . . 111
7.2 Redshift space . . . 116
7.2.1 Visual inspection . . . 116
7.2.2 The effects of simplification . . . 116
7.2.3 Statistical analysis / test setup . . . 118
7.2.4 Statistical analysis / results . . . 118
7.3 Real space identical clustering parameter . . . 119
7.3.1 Visual inspection . . . 120
7.3.2 The effects of simplification . . . 120
7.3.3 Statistical analysis / test setup . . . 122
7.3.4 Statistical analysis / results . . . 123
7.4 Summary observations . . . 123
7.5 Discussion and conclusions . . . 125

8 Discussion & conclusions **126**
8.1 Overall conclusions 126
8.2 Methodological discussion 127
8.3 Where to go from here 127
8.4 Popular scientific summary 128

A Mathematical background **129**
A.1 Topological essentials in 10 minutes 129
A.2 Chains and homology in 10 minutes 135
A.3 Formalizing persistence 142

CHAPTER 1

Introduction

"Far out in the uncharted backwaters of the unfashionable end of the Western Spiral Arm of the Galaxy lies a small unregarded yellow sun. Orbiting this at a distance of roughly ninety-eight million miles is an utterly insignificant little blue-green planet whose ape-descended lifeforms are so amazingly primitive they still think digital watches are a pretty neat idea.", or so at least Adam Douglas described the position Earth in the Milky Way. Put in this way our planet already sounds quite irrelevant, perhaps justified by the knowledge there are hundreds of billions of stars in our galaxy and there are good chances most of them have (several) planets [Petigura et al., 2013]. But even this description makes our planet, our sun and our entire galaxy inconceivable more important than they are. Our Galaxy is just another spiral galaxy somewhere in the outskirts of the Virgo cluster, in itself only a small part of the Virgo supercluster [Tully, 1982] - one of the many superclusters in the visible part of our universe...

On these largest scales of the universe every galaxy can just be represented by a dot and all these dots together form an intricate structure called the ***Large Scale Structure of the Universe*** (***LSS***). More poetically it is sometimes also called the ***cosmic web***, as it visually resembles a spider's web. Figure 1.1 shows an image of the cosmic web based on the Sloan Digital Sky Survey. It shows intricate structures throughout the universe: large filaments connecting enormous clusters and surrounding large empty voids. It is here, on the largests scales of the universe, where our story begins in chapter 2 with an overview of cosmology and the cosmic web.

As marvellous as the cosmic web may be, even more stunning is the realization that its visible components (like galaxies, gas, et cetera, in general called baryonic matter) represent only a fraction of the total energy content of the universe. On the universal stage, baryonic matter is just a side-actress lighting up the play but only marginally influencing the story. The main actors are the mysterious dark matter and dark energy, whose cosmic duet will determine the future and fate of the universe. In general, dark matter arranges the 'local' affairs: it is the main constituent of the cosmic web, whose structure and shape is strongly dependent on it. Dark energy on the contrary, is spreaded too thinly to exert local influence directly. But it is everywhere and rules the universe on global scales. However, perhaps dark energy subtly mingles locally as well. If so, it will give a distinct imprint on the detailed shape and structure of the cosmic web. Conversely, the detailed shape and structure of the cosmic web may shed some light on the mysterious nature of dark energy.

Every story is written in a language. According to the *linguistic relativity hypothesis* of cognitive linguistics, the use of a particular language influences thought and non-linguistic be-

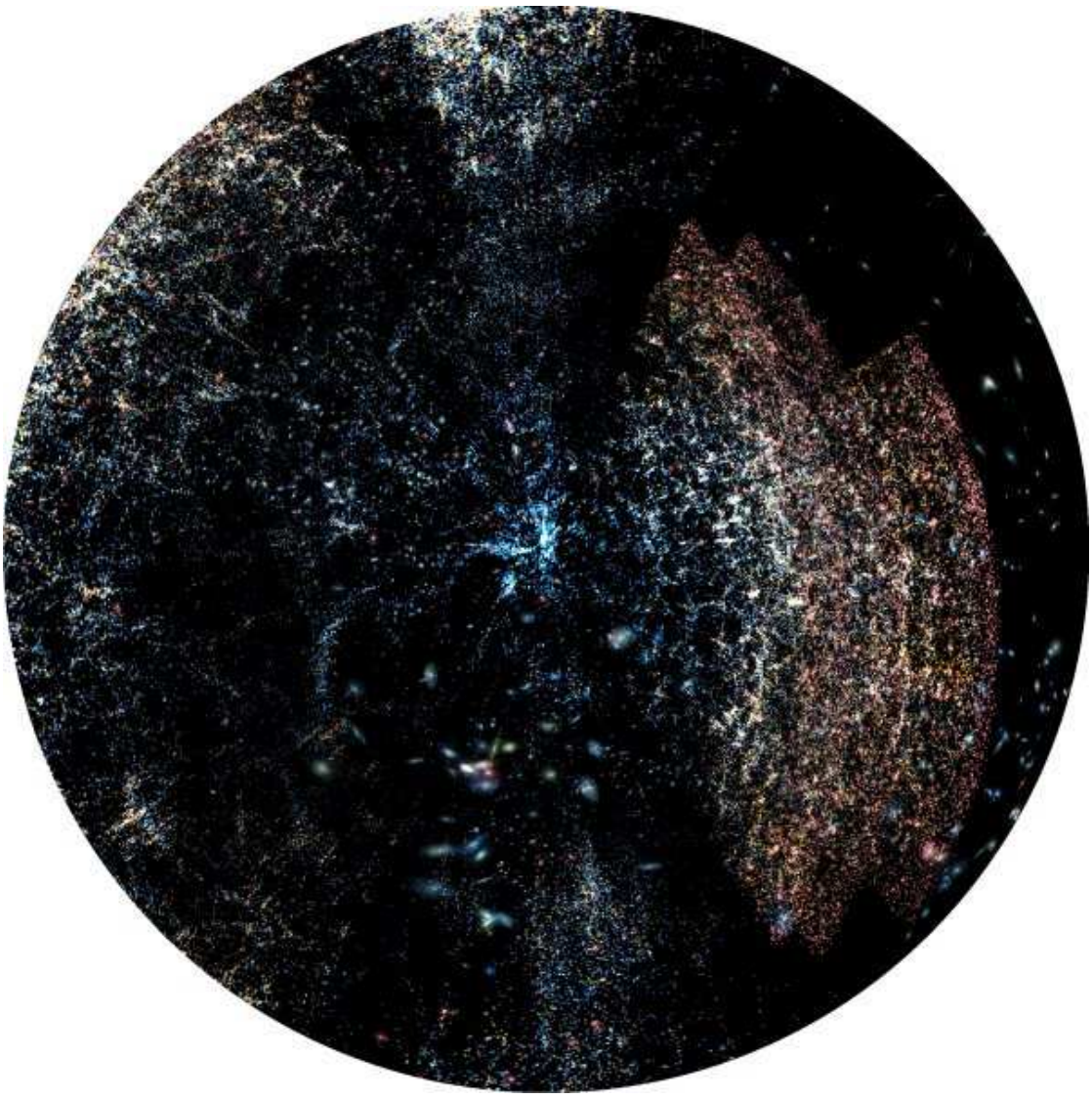


Figure 1.1: ***Illustration of the cosmic web [Subbarao et al., 2005].*** *Based inside a filament, we look back to our galaxy and see the majestic cosmic web unfolding. Image data is from the Sloan Digital Sky Survey.*

haviour (weak version) or even determines it (strong version) [Brown, 1976]. I.e. language might not just neutrally express thought, it might influence or even define it. The strong version of linguistic relativity has an acquisition problem: in order to learn a language some cognition (thoughts) are needed, so language cannot determine thought completely. The weak version is an area of active empirical research [Lucy, 1997], in general providing support for the hypothesis. Some of the most recent developments attempt to tie in the weak variant with neurology [Gilbert et al., 2006].

Nature is written in the language of mathematics. Just as the use of a certain language directs our thoughts and questions we can ask to other persons, the type of mathematics used determines the type of questions we "can ask" nature. Since the discovery of the cosmic web, many analytical and statistical methods have been employed to describe its structure. The methods employed are often genial and have a sound physical foundation but they also have a major disadvantage: they are ill-equipped to really *describe* structures and shapes. They only provide handles to work

with it. A branch of mathematics perfectly suited for describing structure and shape is topology, a generalization of geometry letting go of exact distances. Chapter 3 provides an introduction of this branch of mathematics and how it can be applied to cosmology.

One of the first applications of topology to the study of the cosmic web is by [Schaap and Weygaert, 2000] in the form of the Delaunay triangulation, which also plays a role in various stages of this thesis. Only very recent developments within topology itself allow its utilization by cosmology in full rigour and detail. As cosmological datasets are very large, manual processing of the data is unfeasible and theoretical topology needs to be translated into computational topology. Chapter 4 describes the full data pipeline.

This is one of the first times topology is applied to cosmology and as such, besides a few mathematical results for individual cases, the stability of the data pipeline of the previous chapter has not been proven for real datasets. Furthermore, the topological effects of elementary operations like smoothing the dataset is yet unknown. Chapter 5 provides this basis experimentally, showing (i) that topological measures employed are robust and (ii) smoothing has an excellent topological analogue called manifold simplification.

Having established the stability of the method, we study the topology of the cosmic web under the standard form of dark energy extensively in chapter 6. Subsequently, we do the same for several other forms of dark energy and compare the results in chapter 7. We find that at equal (and low) redshifts, the dark energy models considered here are distinguishable in real and redshift space! However, at equal σ_8 this does not seem to be the case. This suggests that dark energy does influence the shape of cosmic structure but only indirectly via its influence on the expansion of the universe.

These discoveries pave the way to put constraints on dark energy based on the shape of cosmic structure. These and other future directions of research, together with a discussion and conclusions are presented in chapter 8.

CHAPTER 2

Cosmology and the Cosmic Web

Since the dawn of mankind, in all cultures at all times, fundamental questions arise like *Where does everything come from?*, *What is our place in the universe?* and *How do we and everything around us come to an end?*. These kind of questions are studied by ***cosmology***: the study of the universe as a whole. Though such questions are seemingly remote from everyday life, the organisation of many societies ultimately depends on supposed answers to these questions and throughout history many wars have been fought over them.

The stage for modern physical cosmology was set with Copernicus' heliocentric model in 1543 [Copernicus] and Kepler's discovery of the ellipticity of planetary orbits at the first half of the 17th century [der Wissenschaften]. In the decades that followed the idea that stars might actually be other suns developed. It was around 1750 when the British astronomer Wright proposed that these stars might be ordered in a thin planar region around a centre [Hoskin, 1970][1], i.e. in a structure we call a *galaxy* today. For a long time it was unclear whether our galaxy the Milky Way constituted the entire universe or is just of many galaxies or 'island universes' within the universe. The excellent observations of Hubble [Hubble, 1929] and later by Baade [Baade, 1952] provided empirical evidence that our galaxy is just one of many, settling the issue. The realization that there are numerous galaxies in the visible universe and that our galaxy doesn't occupy a spcial position in any way heralded the birth of ***physical cosmology***. One of its cornerstones is the ***cosmological principle***, which states that on large scales ($\gtrsim 100$ Mpc) the universe is spatially homogeneous and isotropic, i.e. it looks the same everywhere and in all directions.

Besides showing the existence of other galaxies, Hubble had shown as well that they are moving away from each other with a 'velocity' proportional to their distance. The Belgian astrophysicist and priest Lemaître interpreted this correctly as meaning not that the galaxies actually moved, but that the universe is expanding. If the universe is expanding now, than it must have been smaller in the future and further backwards infinitely small, which let Lemaître to the *Big Bang cosmological model* [Lemaitre, 1931]. In this model the universe originated from a Big Bang. Several mainly agnostic scientists ridiculed the idea that the universe was created with a Big Bang[2] and proposed the *Steady State cosmological model* as alternative. The Steady

[1] More accurately, Wright proposed a thin shell or an annulus around the galactic centre, partly for religious purposes. Based on only a summary of his ideas, the famous philosopher Kant interpreted and extended Wright's ideas to a plane around the galactic centre [Hoskin, 1970].

[2] Actually the name *Big Bang* was coined by an opponent of the theory.

State model asserts the *perfect cosmological principle*: homogeneity and isotropy in space and time, but at a price. The only way to reconcile homogeneity and isotropy in time with an expanding universe is to let go energy conservation and embrace spontaneous generation of matter. The principle observation confirming the Big Bang model was the discovery of the ***Cosmic Microwave Background (CMB)***: highly uniform microwave radiation coming from every point on the sky. It was theorized by Alpher, Herman and Gamov [Alpher et al., 1948] and discovered by Penzias and Wilson [Dicke et al., 1965; Penzias and Wilson, 1965], which earned them the Nobel Prize of Physics in 1978 [Nobelprize.org, 2013a]. Within the Big Bang model, the existence of the CMB is naturally explained: when the universe is very small and dense average temperatures will be very high. Consequently, photons will continuously ionize hydrogen atoms and the universe will be filled with a photon-matter fluid. When the universe expands the average temperature decreases and at a certain moment becomes so low that photons cannot longer ionize hydrogen. Photons decouple from matter and commence a long journey through the universe. The photons which decouple at this moment form the CMB. For billions of years they travel uninterruptedly through an effectively transparent universe, carrying with them an imprint of these early moments of the cosmos. When their journey ends at a telescope detector, they reveal that early state. The Steady State model doesn't offer any reasonable explanation to account for the CMB. With the manifest victory of the Big Bang model, physical cosmology outgrew its infancy and defined the ***standard cosmological model***[3]. We note that besides the expansion of the universe and the presence of the CMB the standard cosmological model also explains the chemical abundances of light elements, the formation of structure in the universe, Olber's paradox and a host of smaller observations, making it a full-fledged scientific paradigm.

Based on this model, the first part of this chapter considers the geometry of the universe (section 2.1), its laws of physics (section 2.2) and a general overview of its ingredients (section 2.3). The main references for this part are [Weygaert, 2010], [Ryden, 2003] and [Liddle, 2009]. The second part of this chapter focusses more in detail on the cosmic ingredients with as basic references [Weygaert, 2012] and [Peacock, 2005]. We shed some light on the mysterious dark energy in section 2.4 and describe how matter forms the cosmic structures in our universe in section 2.5. In section 2.6 we consider several formalisms to describe the morphology of cosmic structure. Many of these formalisms excellently partition space in its morphological components and a few even provide some aggregate measures. But non really *describe* its morphology. It is here that discrete topology enters the stage, as it describes the cosmic web directly in its components. A general introduction to topology follows in chapter 3.

2.1 The geometry of the Universe

In a flat three dimensional space with Cartesian coordinates (a, b, c) or spherical coordinates (x, θ, ϕ) the standard Pythagorean distance metric ds (cf. definition A.1) is:

$$ds^2 = da^2 + db^2 + dc^2 = dx^2 + x^2\left(d\theta^2 + \sin^2\theta d\phi^2\right) \tag{2.1}$$

There is no a priori reason to assume we live in a flat universe. Introducing an arbitrary irregular curvature is very difficult but under the assumption of the cosmological principle only three universal curvatures need to be considered: a positive (spherical) curvature, zero (flat) curvature or negative (hyperbolic) curvature. The type of curvature is indicated by κ, which is +1, 0 and -1 for positive, zero and negative curvature respectively. For the non-zero case, the radius of

[3] A well known specific version of this model is the Concordance Model [Ostriker and Steinhart, 1995].

curvature is written as R. To generalize the metric above to include such curvatures, we have to replace the x^2 in the equation above by $R^2 \sin^2(x/R)$ for a positively or by $R^2 \sinh^2(x/R)$ for a negatively curved universe. Combining the three possible curvature options we get:

$$ds^2 = dx^2 + S_\kappa(x)^2 d\Omega^2 \tag{2.2}$$

with

$$S_\kappa = \begin{cases} R\sin(x/R) & \kappa = +1 \\ x & \kappa = 0 \\ R\sinh(x/R) & \kappa = -1 \end{cases}$$
$$d\Omega^2 = d\theta^2 + \sin^2(\theta) d\phi^2$$

Performing the coordinate transform $x \to y = S_\kappa(x)$ we can equivalently write equation 2.2 as:

$$ds^2 = \frac{dy^2}{1 - \kappa y^2/R^2} + y^2 d\Omega^2 \tag{2.3}$$

To go from a metric for space to metric for spacetime we should take time into account, adding an extra term $-c^2dt^2$ to equations 2.2 and 2.3. In time, the universe expands. Because the expansion is uniform, we can describe it with the ***expansion factor*** $a(t)$, normalized such that $a(t_{\text{today}}) = a(t_0) = 1$. Taking time and universal expansion in time into account turns equations 2.2 and 2.3 in the:

Definition 2.1 (Robertson-Walker metric)**.**

$$\begin{aligned} ds^2 &= -c^2dt^2 + a(t)^2 \left(dx^2 + S_\kappa(x)^2 d\Omega^2\right) \\ &= -c^2dt^2 + a(t)^2 \left(\frac{dy^2}{1 - \kappa y^2/R_0^2} + y^2 d\Omega^2\right) \end{aligned} \tag{2.4}$$

Here, x and y are called the ***comoving coordinate*** as they 'move along' with the expansion of the universe. The ***physical coordinates*** r can be found by multiplying the comoving coordinates with the expansion factor, i.e. $r = a(t)x$. We note that this metric is only valid under assumption of the cosmological principle. More subtle is the t in previous equation; which reference frame do we take to measure cosmic time? Here ***Weyl's postulate*** helps us out:

Postulate 2.1 (Weyl's postulate)**.** *The geodesics of all observers meet at one point in the past: cosmic time can be measured from that point.*

Due to Weyl's postulate, we can simply use a 'universal time' within our calculations. Contrary to the classical Newtonian view, the curvature of the universe and thereby its geometry are not just a passive stage on which the grand play of the universe develops. Rather, the geometry is intimately linked to the content of the universe. This link is provided by the laws governing the universe, which we will consider now.

2.2 The Laws of the Universe

Ordered from strong to weak, the four fundamental forces of the universe are (i) the strong force, (ii) the weak force, (iii) the electro-magnetic force and (iv) the gravitational force. The strong

and the weak force work on microscopic distances while the electro-magnetic and gravitational force act on macroscopic distances. On large distances the universe is electrically neutral. So slightly ironically it is the weakest force of all, gravity, which rules the universe on the largest scales. The gravitational field is described by general relativity with the gravitational field equations or ***Einstein equations***:

$$G^{\mu\nu} + \Lambda g^{\mu\nu} = -\frac{8\pi G}{c^4} T^{\mu\nu} \tag{2.5}$$

with $G^{\mu\nu}$ the Einstein tensor describing curvature, $T^{\mu\nu}$ the energy-momentum tensor describing the energy-content of the universe and Λ the *cosmological constant*. In a Robertson-Walker metric, $g^{\mu\nu}$ becomes:

$$g^{\mu\nu} = \text{diag}\left(1, -\frac{a^2(t)}{1-ky^2}, -a^2y^2, -a^2y^2\sin^2(\theta)\right)$$

It is from the Einstein equations that we see the intimate link between geometry (the left side of the equation) and the energy-content (the right side of the equation). The geometry determines how the energy-content is distributed but in turn the energy-content determines the geometry.

On cosmological scales potentials are weak ($\phi/c^2 \ll 1$) and so contraction for G^0_0 and G^1_1 gives the:

Definition 2.2 (Friedman-Robertson-Walker-Lemaître (FRWL) equations)**.**

$$\dot{a}^2 = \frac{8\pi G}{3}\rho a^2 - \kappa c^2 + \frac{1}{3}\Lambda a^2 \tag{2.6}$$

$$\ddot{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right)a + \frac{\Lambda}{3}a \tag{2.7}$$

with ρ the density, p the pressure and κ the curvature constant.

Note that in the Einstein equations, the cosmological constant term can be brought from the left to the right side. Then instead of being part of the curvature side it becomes part of the energy content of the universe and is renamed ***dark energy***. Defining the density of dark energy as $\rho_\Lambda = \frac{\Lambda}{8\pi G}$, it can be absorbed in the density term of the FRWL equations.

Multiplication of equation 2.6 with $2\dot{a}$ and taking the time derivative of equation 2.7 we can equate both, giving the:

Definition 2.3 (Fluid Equation)**.**

$$\dot{\rho} + 3\left(\rho + p/c^2\right)\frac{\dot{a}}{a} = 0 \tag{2.8}$$

Essentially, this equation states the universe expands adiabatically.

The FRWL equations and the fluid equations can be obtained in an insightful quasi-Newtonian way as well [see Ryden, 2003, chap. 4]:

1. Integrate Newton's law of gravity over the surface of a sphere and equate the integration constant with the curvature term. This gives the first FRWL equation.

2. The fluid equation can be derived by applying the first law of thermodynamics on an adiabatically expanding sphere.

3. Combining the time derivative of the first FRWL equation with the fluid equation equation gives the second FRWL equation.

As becomes clear from the derivations above (both the relativistic and Newtonian version), the FRWL equations and the fluid equation are only two independent equations. As they contain three variables, a, ρ and p, another independent equation is required. Here the equation of state comes in, which for a given cosmic ingredient relates its density and pressure:

Definition 2.4 (Cosmic equation of state)**.**

$$p(\rho) = \omega(a)\rho c^2 \tag{2.9}$$

with $\omega(a)$ a parameter determined by the cosmic ingredient under consideration.

Plugging the equation of state in the fluid equation gives:

$$\dot{\rho} + 3\rho\left(1 + \omega_{DE}\right)\frac{\dot{a}}{a} = 0 \tag{2.10}$$

After integration this becomes:

$$\rho(a) = \rho_0 a^{-3(1+\omega_{DE})} \tag{2.11}$$

Equations 2.10 and 2.11 show that the rate of density decrease as function of a is determined by the parameter ω, i.e. by the cosmic ingredient under consideration. The next section considers these ingredients in more detail.

2.3 Universal ingredients

So far we spoke about density without actually specifying a density of what. Globally, our universe has four main ingredients: (i) matter, (ii) radation, (iii) dark energy and (iv) curvature. We shortly discuss each of these in more detail below, for a complete and thorough overview of all ingredients we refer to [Fukugita and Peebles, 2004].

Matter is often subdivided in two types. Baryonic matter is the stuff humans, stars and interstellar gas is made of. But baryonic matter is only like the lights in a Christmas tree: clearly visible but almost negligible in terms of total mass. A far greater mass contribution comes from dark matte: a pressureless form of matter which doesn't interact with radiation. We refer to [Bertone et al., 2005] for a recent review on candidate species. Galaxies consist largely of dark matter and therefore the N-body simulation we use later on (cf. section 4.1) use only dark matter to describe large scale structure. In good approximation, for near pressureless baryonic matter and pressureless dark matter we have $\omega(a) = 0$. Plugging this in equation 2.11 shows that matter density decreases proportionally to the increase of volume:

$$\rho_m(a) \propto a^{-3}$$

Radiation consists of two types as well: photons and neutrino's. Although radiation particles outnumber matter particles by orders of magnitude, their current contribution to the total energy budget of the universe is near negligible. This can be understood by considering their equation of state. Thermodynamics gives $\omega_{rad}(a) = 1/3$. Plugging this in equation 2.11 results in:

$$\rho_{rad}(a) \propto a^{-4}$$

We see that radiation density decreases a factor a faster than matter density. Intuitively, this additional factor a can be understood from stretching of light waves due to expansion of the universe.

Dark energy (DE) is responsible for the largest contribution to the energy budget of the universe. Despite abundant experimental evidence for its existence it's exact nature remains a mystery. Many models have been proposed, we will review the most important classes in section 2.4. The idea of dark energy originated from Einstein, who introduced the Λ term is his equations (cf. equation 2.5). Einstein was a firm believer of the Steady State cosmological model and the Λ term was meant to prevent such a model to collapse under its own gravity. When Hubble showed the universe expanded, Einstein withdrew the term and named it his *biggest blunder*.

The famous discovery of an accelerated expansion of the universe by two independent supernova redshift surveys [Perlmutter et al., 1999; Riess et al., 1998] heralded the return of the Λ term, as it can provide the required negative pressure to explain the acceleration. The Nobel Prize of Physics 2011 was awarded to the lead scientists of previous papers for this unexpected and ground-breaking discovery [Nobelprize.org, 2013b]. Later observations based on amongst others the CMB [Collaboration, 2013; Komatsu et al., 2011], large scale structure [Dodelson et al., 2002; Percival et al., 2001] and the Hubble constant [Freedman et al., 2001] confirmed their findings.

Although the equation of state of dark energy is unknown, the fact that dark energy exerts negative pressure can be used to obtain some general constraints. Consider the second FRWL-equation 2.7 with the equation of state 2.9 plugged in:

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho_{DE} + \frac{3p_{DE}}{c^2}\right) = -\frac{4\pi G}{3}\rho_{DE}\left(1+3\omega\right)$$

To create negative pressure the term between brackets on the right hand side should change sign. This gives the restriction $\omega_{DE} < -\frac{1}{3}$. Furthermore, it follows from the equation of state 2.9 that $\omega_{DE} < -1$ implies the DE density increases when the universe expands, an highly unlikely (although not completely disproved [Caldwell et al., 2003]) scenario. Thus, reasonable constraints for ω_{DE} are $-1 \leq \omega_{DE} < -\frac{1}{3}$.

Curvature can also be seen as a contribution to the total energy budget of the universe, as according to the Einstein equations 2.5 the geometry and energy content of the universe are intimately linked. Detailed measurements of the CMB [Collaboration, 2013; Komatsu et al., 2011] show the curvature parameter $\Omega_\kappa = -\kappa c^2/\dot{a}^2 = 0$ within one percent, indicating the universe is (nearly) flat.

The exact energy density value required to obtain a flat universe can be found by setting $\kappa = 0$ in the first FRWL equation 2.6, defining the *critical density*

$$\rho_c = \frac{3}{8\pi G}\left(\frac{\dot{a}}{a}\right)^2 \tag{2.12}$$

Evaluated at today, we get $\rho_{c,0} = (9.2 \pm 1.8) \cdot 10^{-27}\text{kg m}^{-3}$ [see Ryden, 2003, chap. 4]. As the universe is nearly flat, densities are often expressed in terms of the dimensionless density parameter $\Omega(t) = \rho(t)/\rho_c(t)$. A graphical illustration of the density contributions of matter, radiation and DE is shown in figure 2.1.

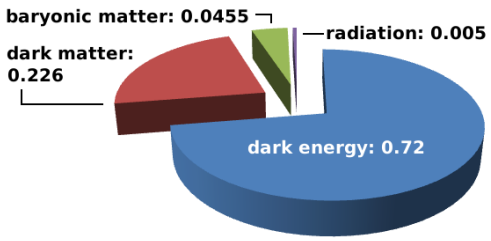


Figure 2.1: ***The cosmic energy inventory.*** *The values shown are from [Komatsu et al., 2011]. Very recent measurements [Collaboration, 2013] suggest* $\Omega_{DE} = 0.6825$ *and* $\Omega_{DM} + \Omega b = 0.3175$ *instead, nonetheless the general pictures remains the same.*

2.4 Dark energy in more detail

Figure 2.1 clearly shows that today, DE provides by far the largest contribution to the energy budget. Furthermore, it plays a vital role in the standard cosmological model. Many ideas on the nature of dark energy have been put forward. Here, we give a short overview of the most important classes. More extensive reviews can be found in [Caldwell and Kamionkowski, 2009; Frieman et al., 2008].

The cosmological constant Λ as introduced by Einstein was a constant meant to keep the universe from collapsing under its own gravity. As mentioned in section 2.2, we can associate the Λ component with a density $\rho_\Lambda = \frac{\Lambda}{8\pi G}$. To keep the density constant, the fluid equation 2.8 tells us that we than have $p_\Lambda = -\frac{\Lambda c^2}{8\pi G}$ and consequently $\omega_\Lambda = -1$. Although observed values converge in this direction, the cosmological constant has one problem: it has no theoretical underpinning.

Vacuum energy attempts to provide such an underpinning by bringing the Λ term from the geometrical to the energy-momentum side and relating it with vacuum energy from quantum mechanics. According to quantum field theory empty space is filled with virtual particles have been measured in the shifts of atomic lines and particle masses. Consequently, the vacuum adds a term to the energy-momentum tensor. As the density of this effect is constant, we again get $\omega(a) = -1$. Unfortunately, attempts to calculate the resulting vacuum energy density fail dramatically: they require the vacuum energy density to be 120 orders of magnitude larger than the critical density ρ_c! This very large discrepancy is known as the cosmological constant problem [Weinberg, 1989]. Supersymmetry helps to reduce this to 60 orders of magnitude due to partial cancelling of zero-point energy contributions of fermions and bosons, but 60 orders of magnitude is still a lot.

Quintessence hypothesises the existence of a yet unmeasured scalar field ϕ that permeates the universe. Introducing such a scalar field makes vacuum energy effectively dynamical, allowing it to vary in time[4]. In its standard form, the scalar field has Lagrangian $\mathcal{L} = \frac{1}{2}\partial^\mu\phi\partial_\mu\phi - V(\phi)$,

[4]Technically the scalar field can also vary in space, but often it is chosen to keep it spatially homogeneous.

resulting in:

$$\begin{aligned} \rho &= \dot{\phi}^2/2 + V(\phi) \\ p/c^2 &= \dot{\phi}^2/2 - V(\phi) \\ &\Leftrightarrow \\ \omega &= \frac{\dot{\phi}^2/2 - V(\phi)}{\dot{\phi}^2/2 + V(\phi)} = \frac{-1 + \dot{\phi}^2/2V}{1 + \dot{\phi}^2/2V} \end{aligned} \tag{2.13}$$

For slowly evolving scalar fields, i.e. $\dot{\phi}^2/2V \ll 1$ we get $\omega \approx -1$ and the scalar field behaves like a slowly changing vacuum energy with $\rho_{VAC}(t) \simeq V[\phi(t)]$. A visual illustration of the scalar field potential and how it gives rise to dark energy is illustrated in figure 2.2. Scalar fields are interesting because they offer degrees of freedom, contrary to vacuum energy. Furthermore, it might an explanation for dark energy and inflation in one go. They have some downsides as well: (i) the cosmological constant problem isn't addressed and (ii) the introduction of new forces may be necessary.

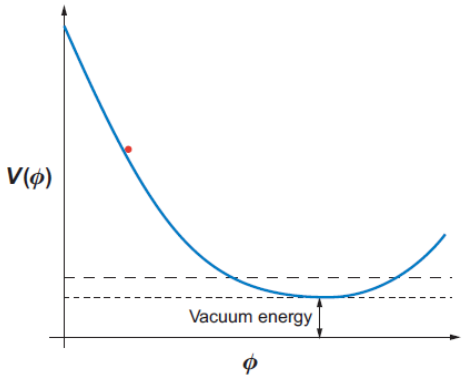


Figure 2.2: ***Generic scalar field potential [Frieman et al., 2008].*** *The scalar field rolls down the potential and settles eventually in a minimum, corresponding to the vacuum.*

Modified gravity changes the left hand side of the Einstein equations instead of the right hand side. If 4D spacetime can still be described by a metric, such a theory needs to change two things: (i) the FRWL equations 2.6 and 2.7; (ii) the equations that describe the growth of density perturbations in the early universe that evolved into the cosmic structure we see today. Different authors have different starting points to modify gravity, some originating from higher dimensional theories, some from string theories and some are purely phenomenological. Changing the laws of gravity is an attractive approach as it doesn't require new scalar fields. However, no consistent theory has been put forward so far.

Inhomogeneous cosmologies suggest the universe is inhomogeneous up to far larger scales than normally assumed. For example, if our galaxy resides near the middle of a very large underdense region, the acceleration effects of DE could be mimicked. To be in agreement with the isotropy of the CMB, this region would have to be nearly spherical as well. Although the idea is interesting, it requires quite some coincidences. Furthermore, it is unknown whether such cosmologies can be made consistent with all cosmological observations.

In this thesis, we will consider and compare the following dark energy models:

- **LCDM** assumes the classical $\omega(a) = -1$, i.e. the cosmological constant or vacuum energy[5].
- **RP** is a classical quintessence model [Ratra and Peebles, 1988] and has a scalar field potential $V_{RP}(\phi) = \frac{\Lambda^{4+\alpha}}{\phi^{\alpha}}$.
- **SUGRA** extends the RP model by including supergravity corrections [Brax and Martin, 2000; Freedman et al., 1976], which extends the potential to $V_{SUGRA}(\phi) = \frac{\Lambda^{4+\alpha}}{\phi^{\alpha}} \exp(4\phi G\phi^2)$.

The DE models above are relatively uncomplicated. Much more complicated models exist, each with their own set of parameters. That many models have their own set of parameters makes it rather difficult to compare them directly. Therefore, it has become customary in literature to parametrize the DE equation of state as:

$$\omega(a) = \omega_0 + \omega_a(1 - a) \tag{2.14}$$

This parametrization allows good comparison of many different models [Linder, 2003] up to observing accuracy [Linder, 2010]. For the three models under consideration here, the values of ω_0 were set at $z = 0$ and their ω_a values determined by a χ^2 fit. These values are tabulated in table 4.1 and the resulting ω-evolution is shown in figure 2.3.

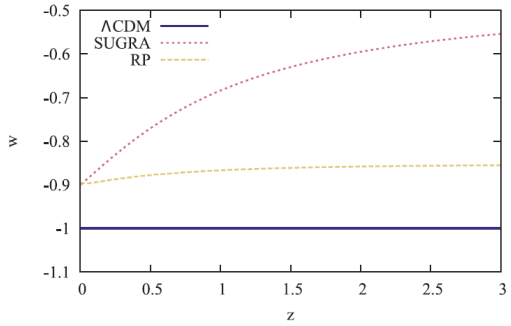


Figure 2.3: *ω evolution as function of z [Bos et al., 2012] (adapted). By definition LCDM's equation of state parameter ω remains constant. RP's and SUGRA's ω value increases slightly respectively significantly in time.*

The rich variety of DE models stands in sharp contrast with their spatial behaviour, which is very monotone. Due to its negative pressure any density difference is directly washed out and consequently its density is the same everywhere. Matter on the contrary can form structures. Exactly how is discussed in the next section.

2.5 Cosmic structure

The magnificent cosmic web we see today developed from the very earliest moments; its imprint can already be found in the CMB. No formalism exists which describes the cosmic web from its very earliest moments to nowadays, but especially for the earlier stages of structure development some excellent models have been developed. By the cosmological principle, the universe is homogeneous and isotropic on large scales ($\gtrsim$ 100 Mpc), giving rise to a universal background density ρ_u, background potential Φ_u and background gravitational force g_u. Furthermore, expansion of the universe causes all galaxies to 'move away' from each other with

[5]It can be either, as fro the point of view of particle simulations, the exact cause of $\omega(a) = -1$ doesn't change any outcomes.

'velocity' $v_H(r,t)$. For structure formation, it is more useful to describe physical quantities in terms of their perturbations with respect to the background. These quantities are called the density perturbation $\delta(x)$, peculiar potential $\phi(x,t)$, peculiar gravitational force $g(x,t)$ and peculiar velocity $v(x,t)$ and are defined as:

$$\begin{aligned} \delta(x,t) &= \frac{\rho(x,t) - \rho_u(t)}{\rho_u(t)} \\ \phi(x,t) &= \Phi_u(r,t) - \tfrac{1}{2}a\ddot{a}x^2 \\ g(x,t) &= -\nabla\phi(x,t)/a(t) \\ v(x,t) &= u(r,t) - v_H(r,t) = a(t)\dot{x} \end{aligned} \tag{2.15}$$

with $u(r,t)$ the physical velocity.

Structure is formed by dark and baryonic matter under influence of radiation. On the Megaparsec scales of the universe we study here, matter and radiation may be seen as a continuous fluid [Weygaert, 2012]. Consequently, we can use fluid dynamics[6] to describe the cosmic evolution of matter and radiation. For each cosmic ingredient j (matter, radiation or dark energy[7]) the three main equations, rewritten in terms of the general relativistic comoving perturbation quantities are the:

- **Continuity equation** formalizes mass conservation: the amount of mass flowing in a volume equals the increase in mass within that volume:

$$\frac{\partial \delta_j}{\partial t} = -\frac{1+\omega_j}{a}\nabla_x \cdot (1+\delta_j)v \tag{2.16}$$

- **Euler equation** describe the forces on the mass which result into mass flows:

$$\frac{\partial v}{\partial t} = -\frac{1}{a}(v\cdot\nabla_x)v - \frac{\dot{a}}{a}v - \frac{1}{a}\nabla_x\phi \tag{2.17}$$

- **Poisson equation** specifies the gravitational potential from which the sources originate:

$$\nabla_x^2\phi = 4\pi G a^2 \left(\sum_j (1+3\omega_j)\rho_{j,c}\delta_j\right) \tag{2.18}$$

These equations will play a pivotal role in analysing the development of large scale structure, which we will do in three stages: (i) the initial conditions; (ii) the linear phase and (iii) the non-linear phase.

2.5.1 The seeds of cosmic structure

The seeds of cosmic structure can already be found in the CMB in the form of small temperature perturbations, see figure 2.4. For larger angular scales[8] these temperature perturbations correspond inversely to density fluctuations: at places the density was higher, photons required

[6] Familiarity with fluid dynamics is assumed. The reader who requires additional background is referred to the excellent treatise [Veldman and Velicka, 2010].

[7] Although for dark energy the perturbation quantities are zero.

[8] For smaller angular scales the relation between photon temperature and underlying density is more complex, we refer to [Hu and Dodelson, 2002] for an overview.

more energy to escape the gravitational pull, hence lowering their temperature. In the CMB the temperature differences are very small: $\Delta T/T \approx 10^{-5}$ and analogous figures are true for the density differences. Under these circumstances structure formation can excellently be approximated linearly, as we will see in the next section. In this section we will see how these density perturbations came into being. Within standard theory, the density perturbations visible

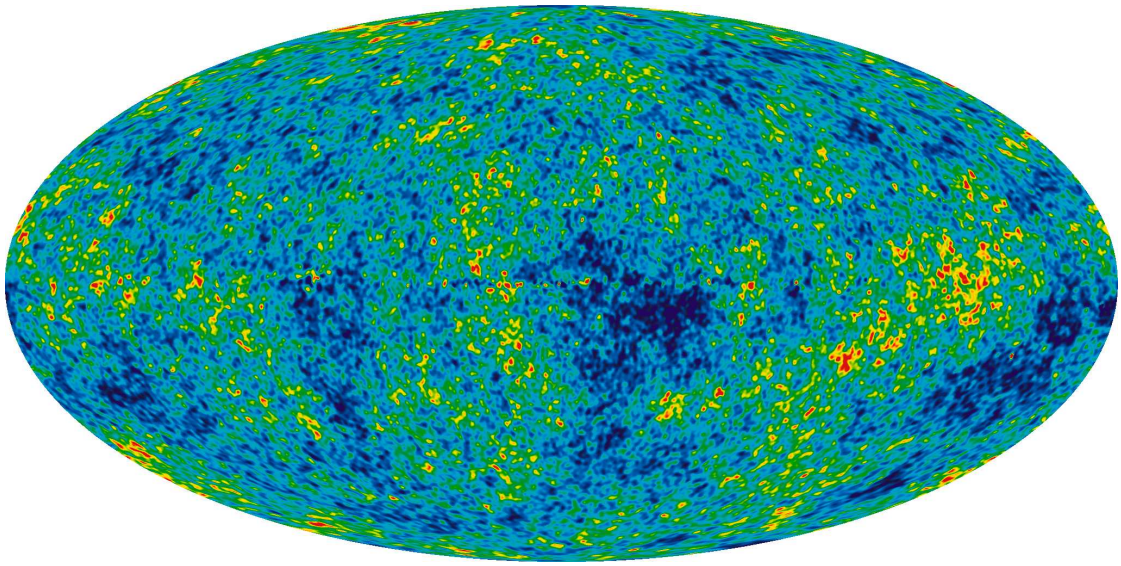


Figure 2.4: ***The CMB [Team]****, based on 9 years of data from the Wilkinson Microwave Anisotropy Probe (WMAP). The fluctuations are extremely small:* $\Delta T/T \approx 10^{-5}$*, showing that the universe was already isotropic at its infancy.*

in the CMB actually come from the earliest moments after the Big Bang, when the universe was incredibly small. Because the universe was so small it was in thermal equilibrium, which explains the near equilibrium of temperature and density. That there are differences at all is due to *Heisenberg's Uncertainty Principle*. Around 10^{-36} seconds after the Big Bang the universe expanded with the extraordinary factor factor e^{100}. This expansion is called ***inflation*** [Guth, 1981; Linde, 1990] and it blew up the Heisenberg's microscopic energy fluctuations to macroscopic scale. The power spectrum of the CMB supports this hypothesis: it is (nearly) perfectly Gaussian [Collaboration, 2013; Komatsu et al., 2011], exactly as one would expect from quantum fluctuations.

Besides sowing the seeds of structure, inflation also solves several other problems in cosmology, most notably the:

1. *The flatness problem*: as we saw in section 2.3 the curvature of the universe is (almost) zero. Inflation explains why: even if the universe was curved, the extreme expansion would reduce this curvature to nearly zero.

2. *The horizon problem*: the CMB shows the matter-radation soup at its time of last scattering was nearly isotropic, even between regions which couldn't have been in causal contact. Causal contact can be reestablished if the universe expanded fast from a much smaller state.

3. *The monopole problem*: we don't observe magnetic monopoles, whereas Grand Unified Theories predict they should be formed around the moment of inflation. Blowing up the universe by a factor of e^{100} descreases the monopole density to 1 in our observable universe.

Although the exact physical mechanism giving rise to inflation has yet to be determined, it is expected that inflation arises from the breaking of symmetry between the strong force and electro-magnetic & weak forces.

With the seeds of structured sowed, the next subsection describes their development.

2.5.2 Linear structure formation

In the first epochs of real structure formation, density and velocity perturbations are small so structure formation can be described with a linear approximation. Concretely, the linear regime is valid when $\delta \ll 1$ and $(vt_{exp}/d)^2 \ll \delta$ with d the coherence length for spatial variations of δ and t_{exp} the expansion time. In the linear regime higher order terms are negligible and the fluid equations 2.16, 2.17 and 2.18 simplify to:

$$\begin{aligned} \frac{\partial \delta_j}{\partial t} &= -\frac{1+\omega_j}{a}\nabla_x \cdot v \\ \frac{\partial v}{\partial t} &= -\frac{\dot{a}}{a}v - \frac{1}{a}\nabla\phi \\ \nabla^2\phi &= 4\pi G a^2 \sum_j (1+\omega_j)\rho_{j,u}\delta_j \end{aligned} \tag{2.19}$$

The stage of linear structure formation takes place at redshifts at which the universe hasn't reached the DE dominated epoch yet. Furthermore, as most of the matter in the universe consists of collisionless dark matter, in good approximation pressure effects can be neglected. This gives us the linearised fluid equations for matter perturbations:

$$\begin{aligned} \frac{\partial \delta}{\partial t} &= -\frac{1}{a}\nabla_x \cdot v \\ \frac{\partial v}{\partial t} &= -\frac{\dot{a}}{a}v - \frac{1}{a}\nabla\phi \\ \nabla^2\phi &= 4\pi G a^2 \rho_u \delta \end{aligned} \tag{2.20}$$

Taking the divergence of the Euler equation, substituting the v-terms by δ-terms using the continuity equation and substituting the ϕ-term by a δ-term using the Poisson equation we obtain a second order partial differential equation for δ:

$$\frac{\partial^2 \delta}{\partial t^2} + 2\frac{\dot{a}}{a}\frac{\partial \delta}{\partial t} = \frac{3}{2}\Omega_0 \dot{a}_0^2 \frac{1}{a^3}\delta \tag{2.21}$$

This equations allows us to draw two conclusions regarding the linear stage of structure formation: (i) as second order partial differential equation it has two independent solutions and (ii) the time and space variables can be separated, i.e. $\delta(x,t) = D(t)\Delta(x)$. We can write the general solution as:

$$\delta(x,t) = D_+(t)\Delta_+(x) + D_-(t)\Delta_-(x) \tag{2.22}$$

In general, the growing mode solution D_+ will become stronger in time whereas the decreasing mode D_- will decrease in time. Quite fast, the D_- solution will become negligible and we only need to consider the D_+ solution to describe linear structure evolution. A general analytic formula $D_+(a)$ is not available, but for spatially flat models a good fitting formula is

[Lahav and Suto, 2004]:

$$
\begin{aligned}
D(a) &= \frac{5a\lambda_m(a)/2}{\lambda_m^{4/7}(a) - \lambda_{DE}(a) + [1 + \lambda_m(a)/2][1 + \lambda_{DE}/70]} \\
\lambda_m(a) &= \frac{\Omega_{m,0}}{\Omega_{m,0} + \Omega_{DE,0}a^3} \\
\lambda_{DE}(a) &= \frac{\Omega_{DE,0}a^3}{\Omega_{m,0} + \Omega_{DE,0}a^3}
\end{aligned}
\tag{2.23}
$$

Although the above formula could be used directly to study the evolution of structure formation, usually the evolution of $D(a)$ is expressed in terms of the *dimensionless linear velocity growth factor* f, defined as:

$$
f \equiv \frac{d\ln(D)}{d\ln(a)} = \frac{a}{D}\frac{dD}{da} \tag{2.24}
$$

The following approximation turns out to be extremely accurate [Linder, 2005]:

$$
\begin{aligned}
f(\Omega_m) &= \Omega_m^{\gamma} \\
\gamma &= 0.55 + 0.05[1 + \omega(z = 1)]
\end{aligned}
\tag{2.25}
$$

Both formulae 2.23 and 2.25 for the growing mode of the density perturbation term $\delta(x,t)$ show that structure evolution mainly depends on matter, but a subtle influence of DE is visible as well. The second of these equations also shows that the magnitude of the DE influence depends on its equation of state. To illustrate the dependence of structure formation on DE, figure 2.5 shows the evolution of δ for $-1 \leq \omega_{DE} < -1/3$. Under the influence of gravity, density perturbations

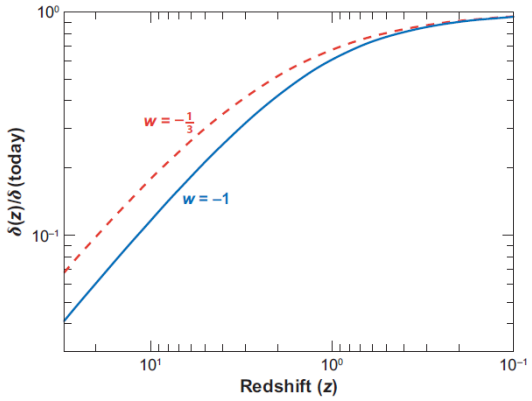


Figure 2.5: ***The influence of dark energy on structure formation [Frieman et al., 2008].***

grow. When $\delta(x,t) \approx 1$ linear theory is not longer an appropriate description of structure evolution and we will need to resort to non-linear theory.

2.5.3 Non-linear structure formation: Zel'dovich

Whereas the linear stage of structure formation has been completely worked out analytically, no complete analytic theory is available for the non-linear case. The most straightforward approach is to take the non-linearised fluid equations 2.16, 2.17 and 2.18 and continue from there.

Extensive amounts of literature exist in this direction and complicated higher-order perturbation analyses has been done [Bernardeau et al., 2002]. However, increasing the order of the approximation quickly leads to a drastic increase in complexity, resulting swiftly in the loss of structural insight and computational achievability. Fluid dynamics as described so far uses a metric fixed in space over which the fluid moves, called the ***Eulerian approach***. Alternatively, fluid dynamics can also be described on a metric that moves along with the fluid, called the ***Lagrangian approach***. As example, the simplest Eulerian approach uses a coordinate system with equidistant points in terms of space whereas the simplest Langrangian approach uses a coordinate system with equidistant points in terms of mass. A transformation from Eulerian to Lagrangian comoving coordinates is achieved with:

$$\frac{d}{dt} \equiv \frac{\partial}{\partial t} + \frac{1}{a} v \cdot \nabla \tag{2.26}$$

Thereby the Lagrangian fluid equations (in terms of general relativistic comoving perturbation quantities) become:

$$\begin{aligned} v &= a\frac{dx}{dt} \\ \frac{dv}{dt} &= -\frac{\dot{a}}{a} v - \frac{1}{a}\nabla\phi \\ \nabla^2\phi &= 4\pi G a^2 \rho_c \delta \end{aligned} \tag{2.27}$$

The ***Zel'dovich approximation*** [Zel'dovich, 1970] is a first order approximation of the inverse of mapping 2.26 to describe density evolution. Consider a mapping from initial Langrangian coordinate q to its Eulerian coordinate $x(q,t)$, i.e. $q \rightarrow x(q,t)$. If we consider the perturbation quantity $s(q,t) = x(q,t) - q$, previous mapping becomes $q \rightarrow s(q,t)$.

The mass originally contained in the infinitesimal volume dq is transported to the infinitesimal volume dx. Naturally, the density in Lagrangian space $q(t)$ is simply the average cosmic density $\rho_c(t)$, giving us:

$$\begin{aligned} \rho(x,t)dx &= \rho_c(t)dq \\ &\Rightarrow \\ 1 + \delta(x,t) &= \frac{\rho(x,t)}{\rho_c(x,t)} = \|\frac{\partial x}{\partial q}\|^{-1} \end{aligned} \tag{2.28}$$

with $\|\ldots\|$ the Jacobian determinant. To evaluate the Jacobian determinant, note that we can write $x(q,t)$ in terms of an ordered sequence of moments of displacement:

$$x(q,t) = q + x^{(1)}(q,t) + x^{(2)}(q,t) + \mathcal{O}(h.o.t.)$$

where $x^{(m)}$ corresponds to the the m-term of the relative displacement $|\partial(x-q)/\partial q|$. Taking the derivative to q and limiting ourselves to first order terms only we obtain:

$$\|\frac{\partial x}{\partial q}\| = 1 + \nabla_q \cdot x + \mathcal{O}(h.o.t.) \tag{2.29}$$

Combining equations 2.28 and 2.29 gives:

$$\delta(x,t) = -\nabla_q \cdot x \tag{2.30}$$

Analogously with the linear case above, plugging the Poisson and continuity equation in the Euler equation (restricting ourselves to first order terms and assuming only longitudinal contributions) we get:

$$\frac{d^2x}{dt^2} + 2\frac{\dot{a}}{a}\frac{dx}{dt} = 4\pi G\rho_c x \tag{2.31}$$

Again a growing and decaying mode exist and solutions for $x(q,t)$ can be found, allowing to compute particle displacement as function from its initial position and time.

Using the Zel'dovich approximation, we can also derive another feature of the cosmic web: its ***anisotropic collapse***. Consider again equation 2.28. The Zel'dovich approximation computes how the mass in the infinitesimal volume dq is transported to the infinitesimal volume $dx(q,t)$. During this transport the volume can be deformed, allowing ut to write:

$$\begin{aligned} 1 + \delta(x,t) &= \|\frac{\partial x}{\partial q}\|^{-1} = \|\delta_{mn} - a(t)\psi_{mn}\|^{-1} \\ &= \frac{1}{[1-a(t)\lambda_1][1-a(t)\lambda_2][1-a(t)\lambda_3]} \end{aligned} \tag{2.32}$$

with ψ_{mn} the Zel'dovich deformation tensor and λ_i its eigenvalues. Note that when $a(t) \rightarrow \lambda_i$ we get $\delta \rightarrow \infty$, implying gravitational collapse in one of the dimensions. Here we see the anisotropic nature of the cosmic web. Assume $\lambda_1 < \lambda_2 < \lambda_3$, than with increasing expansion factor $a(t)$ the following happens:

- While $a(t) < \lambda_1$ none of the dimensions has collapsed and the structure under consideration is a 3D space, a *void*.
- While $\lambda_1 < a(t) < \lambda_2$ one dimension collapsed. The resulting 2D structure is called a *wall*.
- While $\lambda_2 < a(t) < \lambda_3$ two dimensions collapsed, giving us a 1D *filament*.
- After $\lambda_3 < a(t)$ all dimensions collapsed, resulting in a 0D *node* or *cluster*.

The resulting weblike pattern forms the cosmic web and is illustrated in figure 2.6 below. Writing out equation 2.32 explicitly leads to a second order partial differential equation in terms of $\delta(x,t)$, analogously to the linear case. The Zel'dovich approximation remains valid up to surprisingly high $\delta(x,t)$ values, but when sheets start to cross the approximation breaks down. Up to that point, the Zel'dovich approximation provides valuable physical insights and explains the anisotropic nature of the cosmic web. If we could prevent sheets from crossing the Zel'dovich approximation remains valid into even stronger non-linear regimes. The adhesion approach does so by introducing an artificial viscosity that models the self-gravity of bound structures. Due to that, structures stick together instead of cross. An extensive discussion of adhesion is beyond the scope of this thesis, we refer the interested reader to [Hidding, 2010; Shandarin, 2009] for an overview. Although the adhesion approach brings us even further in the non-linear regime than Zel'dovich did, so far strongly non-linear regimes are still out of reach. Several approaches exist to continue from here, each with their own merits and shortcomings:

- The classical Press-Schechter formalism [Press and Schechter, 1974] describes the formation of objects out of the perturbed density field $\delta(x,t)$. It assumes isolated spherical densities and provides support for the *hierachical buildup* of the cosmic web: small structures form first and larger structures form later. The excursion set formalism [Bond et al.,

1991] improves upon the Press-Schechter formalism by filtering over variance of the density field. It is more physically intuitive than Press-Schechter and deals with the cloud-in-cloud problem (smaller objects which form in larger objects). However, both approaches give only a local description of overdensities. In reality the overdensities aren't isolated but part of the global density field.

- Amongst others the spherical model or the homogeneous ellipsoidal model [Weygaert, 2012; Peacock, 2005] assume a special simple configuration and attempt to follow its full non-linear evolution. Although such models provide insight in the mechanisms in play, they only work for the idealized cases.

- N-body codes (cf. section 4.1) allow simulation of structure formation all the way in general configurations. The results are impressive, see for example the snapshot of an N-body code in figure 2.6. But contrary to the formalisms above, they don't provide any physical insight.

When gas falls in the highly non-linear potential wells, dissipative effects kick in, resulting in energy loss due to cooling. The pressure drops and further infall and compression of baryons follows. In time this leads to the formation of galaxies and stars, a fascinating process with many open questions. We refer to the classic [Formation] for details. Ultimately, galaxies form the cosmic web [Bond et al., 1996].

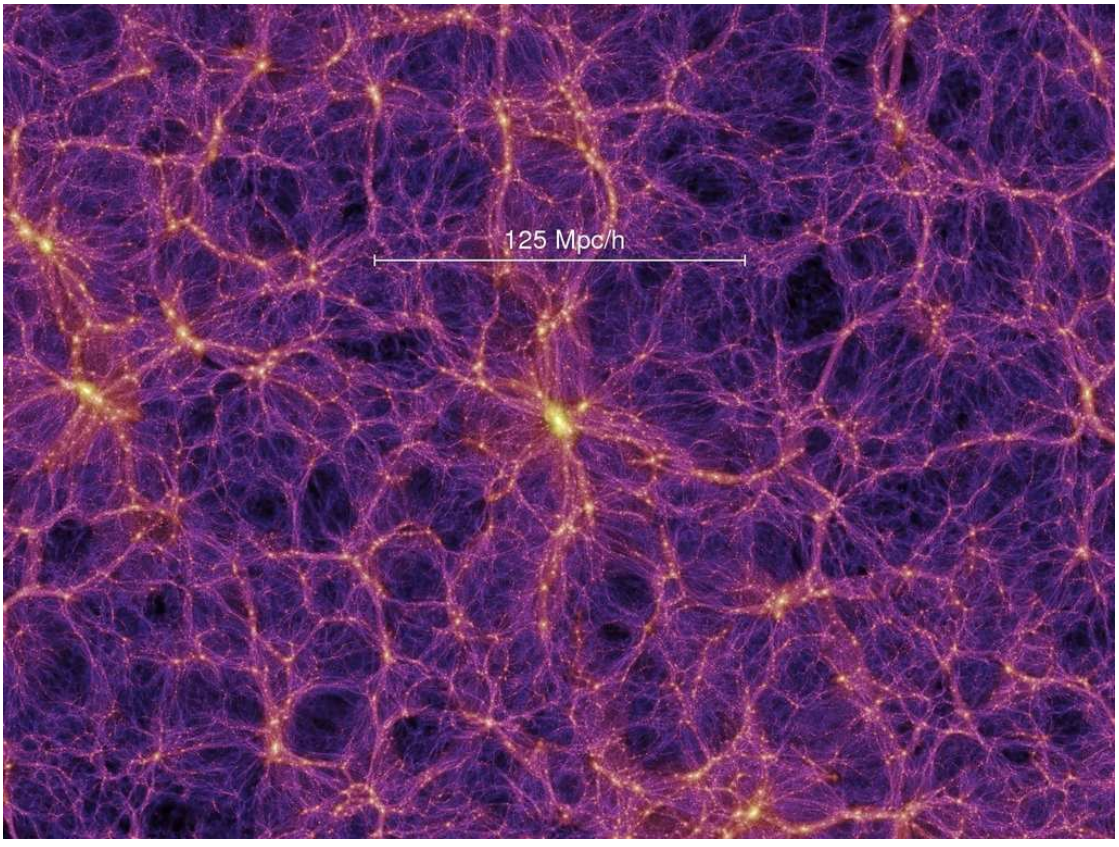


Figure 2.6: ***Snapshot of the Millennium Simulation [consortium].*** *The Millennium Simulation is one of the largest cosmological simulations done to date. It involved over 10 billion particles distributed in a cubic universe of* $2 \cdot 10^9$ *light-years in each dimension. The weblike pattern is clearly visible.*

As N-body codes are the only means to evolve a general density field in the strongly non-linear regime, they are frequently used and we will use them in this thesis. But as they don't provide any physical insights, how to check what comes out is what should come out and whether it makes sense? An intuitive way is to compare the shape and connectedness, i.e. the morphology, of the structure formed in a simulation with the structure observed in the real universe. Visual inspection is capable of detecting large deviations but is not precise enough to find more subtle differences. Therefore, several attempts have been undertaken to describe the morphology of the cosmic web. We will review them in the next section.

2.6 Cosmic web morphology

The cosmic web consists of very dense nodes or clusters, connected to each other by an intricate network of filaments. Filaments in turn are bounded by walls, large sheets who surround the gigantic empty large regions called cosmic voids. Several of these morphological components can be readily appreciated from figure 2.6. Yet although these components follow directly from the Zel'dovich formalism and are recognizable by the human mind directly, mathematical identification and handling is a challenge. Several approaches in this direction exist, we discuss the most important classes below.

2.6.1 Numerical morphology identification

For each of these components and in particular nodes and voids, many methods exist to identify them in simulations or observations. Examples of node finders are SUBFIND [Springel et al., 2001] and VOBOZ [Neyrinck et al., 2005] and some excellent void finders are the (MULTISCALE) WATERSHED TRANSFORM [Dries, 2013; Platen et al., 2007] or ZOBOV [Neyrinck, 2008]. Some outstanding methods like NEXUS [Cautun et al., 2013] even present a coherent framework to identify all morphological components. Besides the insights such component identifiers give in the structure of the cosmic web, they are also a valuable tool for cosmologists in general. For example [E.G.P., 2010] uses void ellipticity in to distinguish different DE models. But identification and numerically designating a part of space as being part of a certain component is something else then really *describing the morphology* of the cosmic web.

2.6.2 Statistics

Another approach to get a handle on the morphology of the LSS is provided by statistics. The most basic approach in this direction is perhaps measuring the amount of clustering. Traditionally, this is done by measuring σ_8: the rms-density variantion averaged over 8 h^{-1} Mpc spheres. If we 'fill space' with a set of spheres $\mathcal{S}_i$ with masses $m_{\mathcal{S}_i}$, than:

$$\sigma_8 = \frac{\sum_i \left(m_{\mathcal{S}_i} - \bar{m}_{\mathcal{S}}\right)^2}{\bar{m}_{\mathcal{S}}} \tag{2.33}$$

with $\bar{m}_{\mathcal{S}}$ the average mass contained in a sphere. σ_8 is one of the basic parameters of the standard cosmological model. A more elaborated statistical approach is the ***two point correlation function*** [Peacock, 2005; Martinez and Saar, 2001]. It gives the joint probability that both infinitesimal volumes dV_1 and dV_2 at distance r contain a galaxy:

$$dP(r) = \bar{n}^2 dV_1 dV_2 \left(1 + \xi(r_{12})\right) \tag{2.34}$$

with $\bar{n}$ the number density of galaxies and $\xi(r_{12})$ the auto-correlation function. Variations like a continuous and angular two-point correlation function exist. However, in general their approximations are crude: they don't contain the information provided by the phases of the Fourier field and completely different spatial patterns could display the same two-point correlation function [Martinez et al., 1990]. Extensions to higher order correlation functions exist but in the background a problem remains: a set numbers from n-point correlation functions might describe the general galaxy distibution, they still don't *describe the morphology*.

2.6.3 Minkowski functionals

Major progress in handeling cosmic morphology mathematically was made with the introduction of *Minkowski functionals* [Kerscher et al., 1997; Mecke, 1994; Schmalzing et al., 1995]. Minkowski integrals give a full morphological description in terms of both topological and geometrical descriptors. In d-dimensional space there are $d+1$ Minkowski functionals, thus in 3D Euclidean space there are four, being:

1. the volume $V = \int dV$;
2. the surface area $S = \oint dS$;
3. the integrated mean curvature $C_m = \frac{1}{2} \oint (R_1^{-1} + R_2^{-1})$;
4. the integrated Gaussian curvature $C_G = (1/2\pi) \oint (1/R_1 R_2) dS$.

These integrals can be computed using the following formalism: consider a point set $\{x_i, i = 1, ..., N\}$ of galaxies in 3D Euclidean space and introduce a set of closed 3D balls $\{B_r(x_i), i = 1, , , .N\}$ with radius r around these points, where in this context a closed ball is defined as (cf. definition A.3):

$$B_r(a) = \{x \in \mathbb{R}^3 | \|x - a\| \leq r\}$$

We start with a very low value for r, such that all balls are disjoint. When r is increased, the balls grow and balls close to another connect. Figure 2.7 shows a set of balls for three different radia, showing a clear difference in connectivity. The Minkowski formalism studies how the four Minkowski functionals on the structure of balls changes as function of r. Such a formalism has several nice properties, like robustness and invariance under translation and rotation.

Minkowski functionals turn out to be good descriptors of different (idealized) structures, see figure 2.8. Furthermore, they really give a description of the geometry under consideration. But they have limitations as well. First, their topological characterization of structures (in essence the fourth integral) is limited. Second, although the formalism can discriminate between different structures, it doesn't really capture the essence of cosmic structure in terms of anisotropic structures. A formalism that goes beyond these limitations is discrete topology.

2.6.4 Homological discrete topology

A mathematical language that describes the morphology of the cosmic structure in the way humans do intuitively, in terms of nodes, filaments, walls and voids, is homological discrete topology. Although the fundaments are around since the beginning of the 20th century, only very recent developments both theoretically and algorithmically makes the formalism usable for large cosmological datasets. An introduction to the this field of mathematics is given in chapter 3 and its algorithmic implementation is discussed in chapter 4. Here we will use this formalism

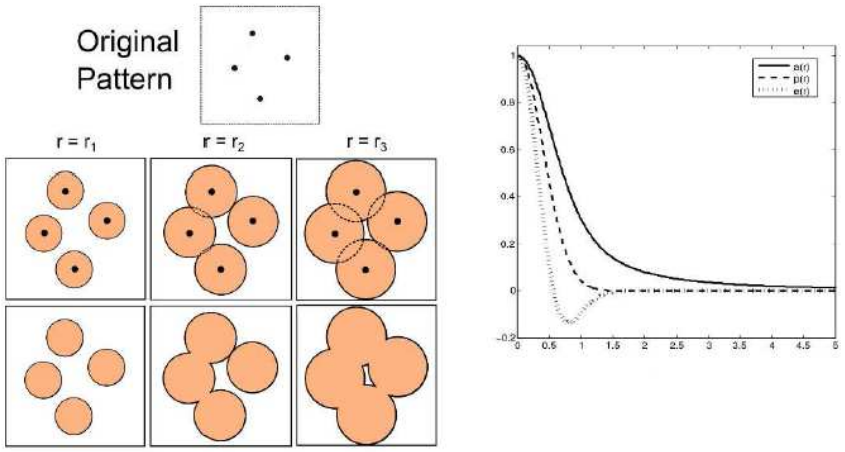


Figure 2.7: ***Minkowski analysis of a point set with growing balls [Parker et al., 2013].*** *A set of points in $\mathbb{R}^2$ with balls of increasing radius (top left). The resulting structure (bottom left) is analysed using Minkowski functionals (right), which in 2D are the area $a(r)$ (solid curve), the perimeter $p(r)$ (dashed curve) and the Euler characteristic χ.*

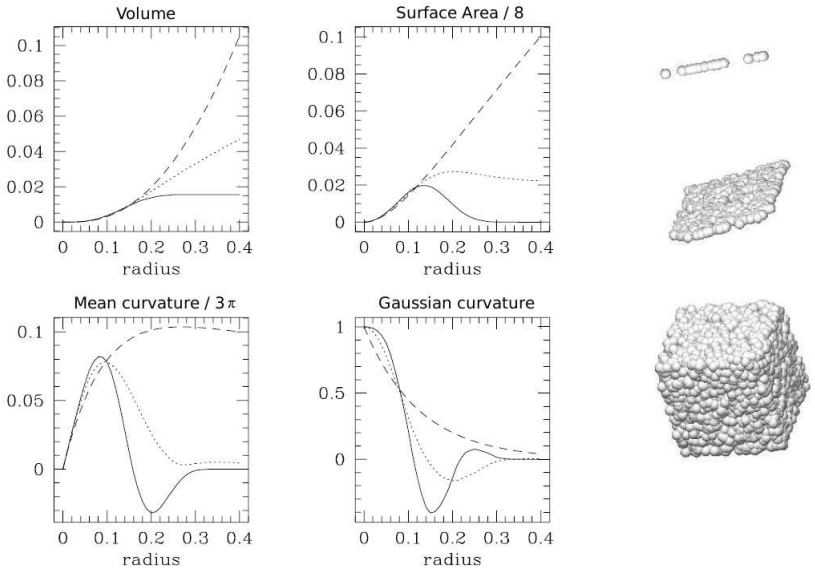


Figure 2.8: ***Morphology discrimination via Minkowski functionals [Schmalzing et al., 1995].*** *Minkowski functionals for idealized structures are good discriminators, as is shown by comparing a Poisson process on a filament (dashed curve), a wall (dotted curve) and a cubic void (solid curve). The division of the geometric quantities by constants is due to normalization, etc.*

to distinguish between the three DE models described in section 2.4. As closing note we remark that the seeds of homological topology are already present in the fourth Minkowski functional C_G. Betti numbers β_0, β_1 and β_2 are the ranks of homology classes (cf. definition A.25 and will be defined in chapter 3. It turns out they give the number of components, tunnels and shells of a structure. Writing χ for the Euler characteristic and using subsequently the Gauss-Bonnet theorem and the Euler-Poincare formula, we can write:

$$C_G = \chi = \beta_0 - \beta_1 + \beta_2 \tag{2.35}$$

I.e. the Minkowski functionals gave a taste of what is to come!

CHAPTER 3

An introduction to topology

Mathematics forms an integral part of natural sciences for several reasons. First, mathematically formulated natural laws allow deeper understanding of phenomena. Second, in absence of such laws mathematics can help data analysis by indentifying relevant features and surpress others. But what exactly qualifies as ‘a feature’ and what as ‘relevant’? Consider the slightly ominous looking mountains on the photo below. With the misty clouds hiding most of the valleys, the mountain peaks feature prominently on the picture. With a bit of fantasy some of the mountains contain large hollow caves in which dragons live... Suppose the mountain landscape to be a holiday picture, how would you describe the landscape to friends and family back home? Probably you would describe the whimsicality of the landscape, the central peak in the middle and the large peaks on the background. Just as some of the most prominent caves. If the clouds would hang a bit lower and some of the mountain passes or valleys would be revealed, you might have described some of these as well. The smaller peaks surrounding the central one are already less relevant. The irregular structure on the peaks themselves is hardly worth mentioning in detail. Not so important as well are the exact locations of the peaks or their exact height. The last of which is even impossible to say as the mountain roots are hidden deep beneath the cloud layer. What is true for the ominous mountains above is true for many scientific ‘mountain landscapes’ as well: to describe their structure, the exact location or height of features is not always very important. It is more their shape, the type of feature (peaks, caves, valleys) and the amount of them what matters. Unfortunately, the normally used mathematical field of analysis is very good in describing analytical measures like location or height but is ill-equiped to describe shapes. The mathematical field of ***topology***, more or less a generalization of geometry letting go of exact distance measures, is much more suited for the latter. This chapter gives a short introduction in topology and its application to cosmological density fields. One of the most important concepts we need from general topology is the concept of a smooth manifold: a topological object which is everywhere locally equivalent with simple Euclidean space. As such it combines the generality of topology and the exactness of regular calculus. Readers less familiar with topology can consult appendix A.1 for a quick overview of the most important concepts. Consider a function describing a manifold. Starting from its largest value, we slowly decrease the function value and so reveal more and more of the manifold. Much like descending clouds would reveal more and more of the mountain landscape above. First, only the highest peaks are visible, but slowly more and more peaks emerge and later mountain passes will connect the peaks. In essence, the appearance of new features means the manifold changes topologically. We begin this chapter with an

introduction to Morse Theory in section 3.1, which describes in detail how a manifold changes when new features emerge. Many astronomical datasets and in particular galaxy catalogues contain discrete data. Sections 3.2 introduces simplices, the basic building blocks of discrete topological space. Using simplices a manifold can be triangulated and as such their shapes can be described in a both concise and computable manner. An instructive example of a triangulation is the Delaunay triangulation. It is based on the Voronoi diagram and will be used extensively in the next chapters. We describe the Delaunay triangulation in detail in section 3.3. By defining functions over simplices Morse Theory can be extended to discrete topology, which is done in section 3.4. Previous sections are combined and some issues regarding practical implementation are given in section 3.5. Triangulated spaces describe manifolds very well but for scientific analysis some aggregated measures are useful. How many features like separate components or mountain peaks does a manifold have, how many tunnels and how many caverns where dragons can hide? By defining equivalence relations on groups of simplices, these features can be extracted from the triangulated topological space in the form of cycles. This allows us to answer the first question posed in the beginning of this section: what is a feature? The details of these equivalence relations is relatively involved, we give their mathematical background in appendix A.2. As we saw in the picture of the mountain landscape, not all cycles are equally important: some are main features but other are just small ripples on the waves of larger structures. How to distinguish? Suppose we have a descending cloud layer. In essence a feature is important if it remains a separate feature for a long time after its peak was revealed by the clouds, whereas it is not so signficant if it melts with another feature fast. The previous is captured mathematically in the concept of persistence. Using persistence, we can concretely answer which features are relevant and which not. Cycles and persistence will be defined in section 3.6, the culmination of this chapter.

The material discussed in this chapter encompasses substantial material from a wide range in mathematics, but throughout this chapter two sources will be used repeatedly: [Edelsbrunner and Harer, 2010] conceived the idea of homological persistence and persistence diagrams, the summit of this chapter. [Sousbie, 2011] was the first to translate and apply [Edelsbrunner and Harer, 2010]'s ideas fully to astrophysics. We cite them here as major references for this entire chapter. Although most of the topological theory expounded below is applicable to a wide range of

manifolds, for ease of exposition and because we don't require more in several cases we restrict ourselves to $\mathbb{R}^3$. A note will be made wherever this is done.

Let no one unversed in geometry enter here.
Entrance words of Plato's academy

3.1 Describing topological changes: continuous Morse Theory

The most important topological space we will encounter in this thesis is a (smooth) manifold. Critical points of a function defined over a manifold allow us to analyze the manifold's topology. Exactly how is excellently described by Morse Theory [Milnor, 1973; Morse, 1960] developed by the famous Marston Morse. By some Morse Theory is called *perhaps the single greatest contribution of American mathematics* [O'Connor and Robertson, 2003]. To understand the rationale behind Morse theory, let's look at a simple and classical example in mathematics, a torus M tangent to the plane V as shown in figure 3.1. Analogously to the mountains above, suppose

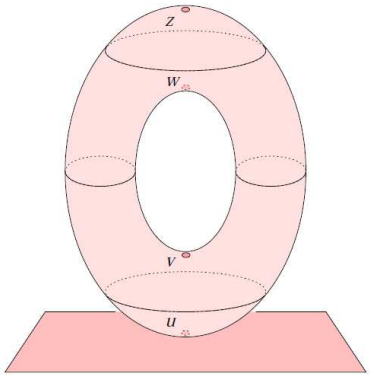


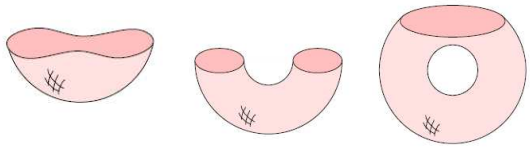
Figure 3.1: ***Torus*** M ***tangent to the plane*** V ***[Edelsbrunner and Harer, 2010]***.

the torus is completely covered in clouds up to the ground V. Instead of standing above the clouds, we go stand below it and while the clouds rise to a certain height a, a part of the torus M_a is revealed. The part of the torus becoming visible is defined as the:

Definition 3.1 (Sublevel set)**.** *A sublevel set* M_a *of function* $f : M \to \mathbb{R}$ *is the subset* $M_a = M|f^{-1}(-\infty, a]$.

We define $f : M \to \mathbb{R}$ as the height above V and consider the topology of M_a:

1. If $a < 0 = f(u) \Rightarrow M_a = \varnothing$.
2. If $f(u) < a < f(v) \Rightarrow M_a$ is topologically equivalent to a 2-ball, see figure 3.2 left.
3. If $f(v) < a < f(w) \Rightarrow M_a$ is topologically equivalent to a cylinder, see figure 3.2 middle.
4. If $f(w) < a < f(z) \Rightarrow M_a$ is topologically equivalent to a torus with a disc removed, see figure 3.2 right.
5. If $f(z) < a \Rightarrow M_a$ is topologically equivalent to a full torus.



(a) A 2-ball. *(b) A cylinder.* *(c) A torus with a circle missing.*

Figure 3.2: M_a ***with increasing level of*** a ***[Edelsbrunner and Harer, 2010].***

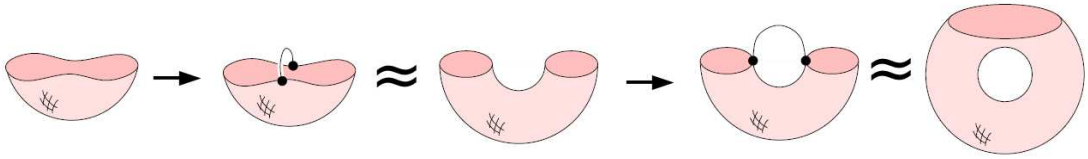


Figure 3.3: ***'Constructing'*** M_a ***from bottom to top, based on [Edelsbrunner and Harer, 2010] .***

As we saw above the topology of M_a changes at the points u, v, w and z, i.e. the critical points of f. This is intuitively understandable: the shape of M_a only 'really' changes when the rising clouds reveal a new structure ($1 \to 2$), a "mountain pass" ($2 \to 3$ and $3 \to 4$) or rise above the complete structure ($4 \to 5$). What is true in this specific example of a height function on a torus is true in general, bringing us to:

Theorem 3.1 (Classical Morse Theory, part I)**.** *Let $f : M \to \mathbb{R}$ be smooth and let $a < b$ such that $f^{-1}[a, b]$ is compact and contains no critical points of f. Then M_a is diffeomorphic to M_b.*

What happens at the critical points $f(u)$, $f(v)$, $f(w)$ and $f(z)$? As visualized in figure 3.3, this depends on the critical point in question:

$1 \to 2$: 'attaching' a 0-ball to the (till than empty) M_a.

$2 \to 3$: attaching a 1-ball to M_a.

$3 \to 4$: attaching a 1-ball to M_a.

$4 \to 5$: attaching a 2-ball to M_a.

For simplicity of exposition and because we don't need more we confine ourselves to $M = \mathbb{R}^d$. The 'type' of critical point influences how the topology changes. A first step in formalizing this is explicitely defining a:

Definition 3.2 (Critical point)**.** *For function f over $\mathbb{R}^d$ and point $p \in \mathbb{R}^d$, p is critical if $\nabla_x f(p) = 0$.*

Remembering that the *Hessian matrix* $\mathcal{H}_f$ is the matrix of second derivates $\mathcal{H}_f(x) = d^2 f / dx_i dx_j(x)$, we can write:

Definition 3.3 (Order of a critical point)**.** *Critical point p has order k if $\mathcal{H}_f(p)$ has exactly k negative eigenvalues. This means intuitively that there are k directions to go from the critical point in descending direction.*

Example 3.1. *for an n-dimensional space minima have order 0, saddle points order 1 till $n-1$ and maxima order n.*

In 2D, like the surface of our torus, there is one saddle of order 1. Look again at the list of changes when the critical point of our example torus was encountered. It seems that when the rising cloud level meets a critical point of order k, a k-ball is added to the topological structure. But this rule is too simple, consider for example the continuously deformed torus shown in figure 3.4. Near the critical point w' the torus is locally approximated by $f(x,y) \sim (x-w'_x)^3 + (y-w'_y)^2$. From the figure we can already see that the topology doesn't change, which is confirmed by a completely zero Hessian determinant. Of course this is independent of coordinates used, as elementary matrix operations on the Hessian can't change the determinant value. A zero Hessian can also occur when critical points aren't isolated from each other. This

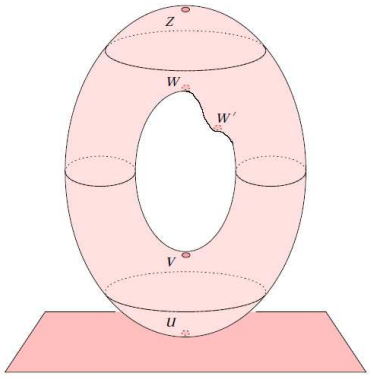


Figure 3.4: ***Continuously deformed torus, based on [Edelsbrunner and Harer, 2010].***

happens for example if we give the torus above a slight push such that it falls on its side. On this fallen torus there is an entire circle of minima and maxima with zero Hessian. This can be seen by changing to polar coordinates and noting that $\partial f(r,\theta)/\partial\theta = 0 \,\forall\, \theta$, creating a zero column in the determinant and thus a zero Hessian. Although the topology changes it is impossible to say at which point. Therefore, it makes sense to look only at ***non-degenerate*** critical points, meaning that $\det(\mathcal{H}_f(p)) \neq 0$. Later on, also another property will be required: all critical points need to have distinct function values. Functions with these desirable properties are called:

Definition 3.4 (Morse functions)**.** *A Morse function is a smooth function on a manifold, $f : M \to \mathbb{R}$ such that (i) all critical points are non-degenerate and (ii) the critical points have distinct function values*[1].

Morse functions have several nice properties:

1. they lead to the complete classical Morse Theory (part II follows shortly);

2. they form an open dense set in the space of all proper smooth functions with appropriate topology, so any proper smooth function can be approximated by a Morse function;

3. if a function f' is close enough to f is some topology, than f and f' have the same smooth topological type.

[1] Sometimes the second condition is dropped but in this thesis it will always be required.

Critical values obtained by Morse functions describe or and how the topology of a manifold changes:

Theorem 3.2 (Classical Morse Theory, part II)**.** *Let f be a Morse function with $a < b$ such that $f^{-1}[a, b]$ is compact and contains one critical point p of f with order k. Then M_b is homotopy equivalent to M_a with a k-ball attached along its boundary.*

Let us pause here for a moment and realize the importance of Morse Theory: for a very general class of functions f it describes precisely when the topology is constant as well as when and how it changes. In the context of large scale structure the cosmological density ρ is (certainly for all practical purposes) a Morse Function. Consequently, Morse Theory describes exactly its topology i.e. the shape of cosmological structure! For the rest of this section we assume f to be Morse. At any non-critical point the gradient defines a preferred direction and by following this direction we can define:

Definition 3.5 (Integration lines (or field lines))**.** *An integral line (or field line) is a curve $\mathbf{L}(t) \in \mathbb{R}^d$ such that*

$$\frac{d\mathbf{L}}{dt} = \nabla_x f \tag{3.1}$$

Integral lines are defined for all t and their origin and destination are written as $\lim_{t\to-\infty} \mathbf{L}(t)$ and $\lim_{t\to+\infty} \mathbf{L}(t)$.

Example 3.2. *The integral lines of the torus are shown in figure 3.5.*

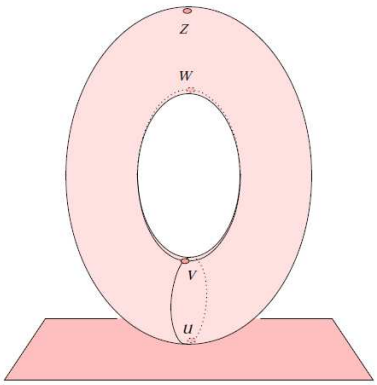


Figure 3.5: *[Edelsbrunner and Harer, 2010] Our torus with the four integral lines that end at the two saddles.*

For the application of our thesis, we are interested in astronomical density fields and not so much in a torus. Therefore, the top-left panel of figure 3.6 shows the integral lines on an astronomical density field. Integral lines have several convenient properties:

Properties 3.1 (Integral lines of Morse functions)**.**

1. *The origin and destination of integral lines are critical points.*
2. *Integral lines are ordinary differential equations. From their existence and uniqueness theorems it follows that two integral lines passing through different points p and q are either the same of fully disjunct, except perhaps at their origin and destination.*

3. *It follows from previous properties that $\mathbb{R}^d$ is completely covered by distinct integral lines, except perhaps at critical points.*

These properties allow us to divide space in regions whose integral lines have the same origin or destination:

Definition 3.6 (Ascending/descending k-manifold)**.** *The stable manifold of a critical point p of f is the point together with all non-critical points whose integral lines end at p. As $f(p) \geq f(x) \forall x$ in the stable manifold, it is often called the descending manifold. Analogously, the unstable manifold of a critical point p of f is the point together with all non-critical points whose integral lines originate at p. As $f(p) \leq f(x) \forall x$ in the unstable manifold, it is often called the ascending manifold.*

If f is defined over $\mathbb{R}^d$ and it has critical point p of order k, than the Hessian has k negative and $d-k$ positive eigenvalues thus the descending manifold has dimension k and the ascending manifold dimension $d-k$.

Example 3.3. *Consider a Morse function f defined over $\mathbb{R}^3$ and a critical point p of order k, than:*

- *If $k = 0$, an ascending manifold of p consist of a 3D space (a 'void'), whereas a descending manifold of p only consists of the 0D point itself (a 'node').*
- *If $k = 1$, an ascending manifold of p consist of a 2D subspace (a 'walls'), whereas a descending manifolds of p consist of 1D subspaces (a 'filament').*
- *$k = 2$ resembles the $k = 1$ case with the dimensions of the ascending and descending manifolds interchanged.*
- *$k = 3$ resembles the $k = 0$ case with the dimensions of the ascending and descending manifolds interchanged.*

The set of ascending or descending manifolds together divide the entire d-dimensional manifold in 0 till d-dimensional regions called the:

Definition 3.7 (Morse complex)**.** *The Morse complex of Morse function f is the set of ascending or descending manifolds.*

The top-right and bottom-left part of figure 3.6 show the descending respectively ascending 2-manifolds of the 2D density field. Here we see how Morse Theory applies to cosmology. The Morse complex of the ascending manifolds naturally divides space in large voidy 2D regions, which intersect at the 1D medium density filaments, which in turn connect 0D high density nodes. The extension to 3D space is straight forward. Thus, if the cosmological density field would be a Morse function the morphological division in voids, walls, filaments and nodes follows automatically. Of course there is no guarantee the cosmological density function is Morse, but as the set of Morse functions is dense we know for sure it can be approximated arbitrarily well by a Morse function.

It should be noted that the Morse complex in some exceptional cases behaves bad. A mathematical complex[2] must satisfy that every subspace of an element is contained in the complex. For Morse functions this doesn't necessarily have to be the case. Consider for example the descending 1-manifold of the upper saddle w in figure 3.5. It reaches down to the lower saddle v

[2]Formally defined in definition 3.15.

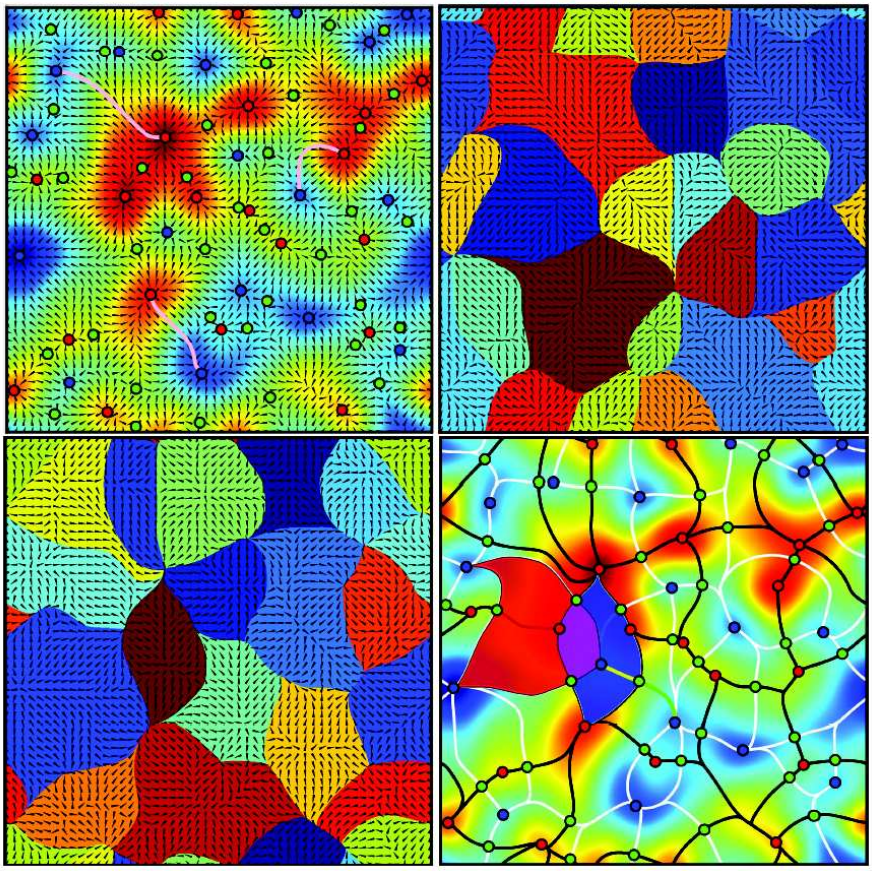


Figure 3.6: ***From density field to its Morse complex [Sousbie, 2011].*** *A 2D density field with its gradient (top left), its descending 2-manifolds (top right), its ascending 2-manifolds (bottom left) and its Morse-Smale complex (bottom right). Critical points are indiceated by red (maxima), green (saddle points) or blue (minima). The three pink lines on the top left are three integral lines. On the bottom right the descending/ascending 2-manifolds are bounded the white/black lines. The intersection between a red descending and blue ascending 2-manifold is a 2-cell shown in purple and the intersection between the green descending 1-manifold and previous blue ascending 2-manifold is a 1-cell shown in yellow.*

but the latter in not a stable 0-manifold and thus (one) of its endpoints is not part of the set of descending manifolds. The problem is caused by a degeneracy in the gradient flow: an integral line both begins and ends at a saddle of the same order or equivalently, the integral line between v and w belongs both to the descending 1-manifold of w and the ascending 1-manifold of v. To avoid these situations, integral lines of ascending and descending manifolds should only intersect transversely, where 'transverse' is the opposite of 'tangent'. I.e. when two transverse lines cross each other they penetrate and do not only touch, making clear for both lines in which direction they continue. Adding this conditions to Morse functions, we get:

Definition 3.8 (Morse-Smale functions)**.** *A Morse-Smale function is a Morse function whose ascending and descending manifolds intersect only transversely.*

Under the assumption of transversality, the intersection of an ascending p-manifold and a descending q-manifold has dimension $\min(p, q)$ or is void and we can define a:

Definition 3.9 (Morse-Smale k-cell)**.** *A Morse-Smale k-cell is the non-void intersection of a p-ascending manifold with a q-descending manifold with dimension $k = \min(p, q)$. Integral lines*

from descending/ascending manifolds all have the same origin/destination, thus integral lines of a Morse-Smale k-cell have all the same origin and destination.

The set of all k-cells together defines the:

Definition 3.10 (Morse-Smale complex)**.** *The Morse-Smale complex of Morse-Smale function f is the set of all k-cells of f.*

A Morse-Smale complex is a proper mathematical complex and solves the problem we encountered in the torus above: the boundary of every descending manifold is a union of stable manifolds of lower dimension. The Morse-Smale complex of a slightly perturbed torus is shown in figure 3.7 and the Morse-Smale complex of the density field is displayed in the bottom-right of figure 3.6.

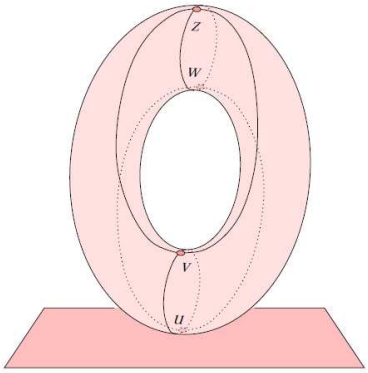


Figure 3.7: ***The Morse-Smale complex of our slightly disturbed torus[Edelsbrunner and Harer, 2010].***

3.2 Discretization: simplicial complexes

Many datasets in astrophysics, under which galaxy catalogues studied here, are not continuous but discrete data sets. The simplest way to transform the ideas developed above for the continuous case to the discrete case is to switch from continuous to discrete topology. There are many ways to represent discrete topological spaces, an intuitive one is a simple decomposition in pieces which form a ***simplicial complex***, the main object of study in this section. On simplicial complexes a discrete version of Morse Theory can be defined, which can be used analogously with continuous Morse Theory and will be developed in the next section.

We begin where we should begin: with a set of discrete points $u_0, u_1, ..., u_k \in \mathbb{R}^d$, each with weight $\lambda_0, \lambda_1, ..., \lambda_k \in \mathbb{R}$. Combining the set of discrete points with weights gives a point $x = \sum_{i=0}^{k} \lambda_i u_i$, which is called an *affine combination* of u if $\sum_i \lambda_i = 1$. The set of all affine combinations is the *affine hull*. Depending on the choices of the coefficients, some affine combinations might be analogous whereas other might really differ. To make this more specific we define:

Definition 3.11 (Affine independence)**.** *Consider two affine combinations $x = \sum \lambda_i u_u$ and $y = \sum \mu_i u_i$. Two affine combinations $x = \sum \lambda_i u_u$ and $y = \sum \mu_i u_i$ are the same $\Leftrightarrow$ $\lambda_i = \mu_i$ for all i. The $k+1$ points u_i are affinely independent $\Leftrightarrow$ the k vectors $u_j - u_0$ with $j \in \{1, ..., k\}$ are linearly independent.*

In $\mathbb{R}^d$ we have at most d linearly independent vectors so at most $d+1$ affinely independent points. An affine combination is called a *convex combination* if all $\lambda_i \geq 0$. The *convex hull* (conv) is the set of all convex combinations. We use the convex hull to define the basic elements of this section, the:

Definition 3.12 (k-simplex)**.** *A k-simplex is the convex hull of $k+1$ affinely independent points $\sigma = conv\{u_0, u_1, ..., u_k\}$. In other words, a k-simplex σ is the convex hull spanned by u_i.*

Just as in linear algebra a set of vectors spans up a vector space, a set of affinely independent points spans up simplex. Consequently, simplices can be seen as the elementary building blocks of topological structure. Some important properties of simplices are:

Properties 3.2 (Simplex related definitions)**.**

1. *A non-empty subset of vectors spans a subspace and analogously a subset of points τ span an subsimplex called a face. A face is proper is the subset isn't the entire set. Conversely σ is the (proper) coface of τ, sometimes denoted as $\tau \leq \sigma$ with a strict inequality in case of a proper (co)face.*

2. *As we will need them often, faces and cofaces of one dimension lower or higher have obtained their own name: facets respectively cofacets.*

3. *A set of $k+1$ elements has 2^{k+1} subsets including the empty set (all elements can be switched on or off) thus σ has $2^{k+1}-1$ faces. It is said that σ has dimension k.*

4. *The boundary of σ, bd(σ), is the union of all proper faces. The interior of σ, int(σ), is everything else, i.e. int$(\sigma) = \sigma -$ bd(σ).*

Example 3.4 (Basic simplices)**.** *Figure 3.8 shows some basic k-simplices. From left to right: a vertex ($k=0$), a line ($k=1$), triangle ($k=2$) and a tetrahedron ($k=3$). Clearly, a line has 2^2-1 proper faces (its vertices and $\varnothing$), a triangle 2^3-1 (three vertices, three lines and $\varnothing$), etc. The boundary of a tetrahedron consists of all the vertices, lines and triangles. The remaining 'inside' is its interior.*

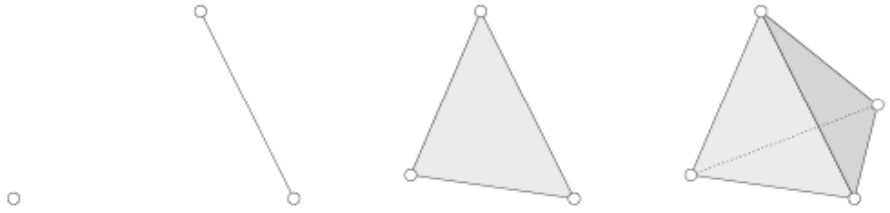


Figure 3.8: ***Basic simplices.***

A set of simplices can form a:

Definition 3.13 (Simplicial complex)**.** *A simplifical complex is a finite collection of simplices K such that:*

1. *if $\sigma \in K$ and $\tau \leq \sigma \Rightarrow \tau \in K$;*
2. *if $\sigma_0, \sigma_1 \in K \Rightarrow \sigma_0 \cap \sigma_1$ is either empty or a face of both.*

Example 3.5. *Figure 3.9 shows some valid and invalid simplicial complexes. The invalid complexes have either incorrect intersections or missing faces.*

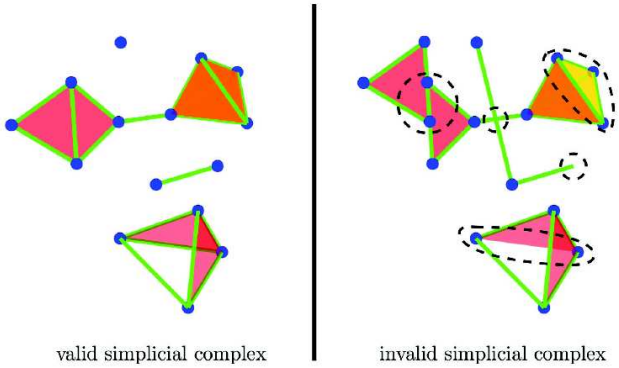


Figure 3.9: ***Example of valid (left) and invalid (right) simplicial complexes [Sousbie, 2011].***

Some important properties of a simplicial complex are:

Properties 3.3 (Simplicial complex related definitions)**.**

1. *A subcomplex L is a simplicial complex $L \subseteq K$. Three specially named subcomplexes are:*
   - *A full subcomplex: all simplices in K spanned by vertices in L.*
   - *A j-skeleton: all simplices of dimension j or less, i.e. $K^{(j)} = \{\sigma \in K | dim \leq j\}$. The 0-skeleton is sometimes named a vertex set.*
   - *A star subset of simplex τ is its set of all cofaces, $St\tau = \{\sigma \in K | \tau \leq \sigma\}$. The closed star of τ is the smallest subcomplex that contains the star, which is obtained by adding all missing faces from the star.*
2. *The dimension of K is the maximum dimension of any of its simplices*
3. *The underlying space of K, denoted with $|K|$, is the union of its simplices together with topology inherited from the ambient Euclidean space.*

This can give us a:

Definition 3.14 (Triangulation of topological space $\mathbb{X}$)**.** *A simplicial complex K together with a homeomorphism between $\mathbb{X}$ and $|K|$. A topological space is triangulable if it has a triangulation.*

For mathematical purposes it is sometimes easier first to construct a complex abstractly and consider how to put it in Euclidean space later, if at all. This warrants the following definition:

Definition 3.15 (Abstract simplifical complex)**.** *An abstract simplicial complex is a finite collection of sets A such that $\alpha \in A$ and $\beta \subseteq \alpha \Rightarrow \beta \in A$.*

The *Geometric Realization Theorem* [Edelsbrunner and Harer, 2010] guarantees that every abstract simplicial complex of dimension d can be realized in $\mathbb{R}^{2d+1}$. A nice example of a simplicial complex is the Delaunay triangulation, often constructed via the Voronoi diagram. Both to illustrate the concept of triangulations and given their importance later on in this thesis, the next section is devoted to them.

3.3 Voronoi diagrams and Delaunay triangulations

Suppose we have a set of points and we want to know the region closest to it. In 2D these regions can be obtained by first connecting all points and subsequently drawing the perpendicular bisectors of these lines, stopping at the intersection with another perpendicular bisector. This gives us a diagram as shown in figure 3.10. This process can be generalized to higher dimensions, giving us a:

Definition 3.16 (Voronoi diagram [Okabe et al., 2000; Voronoi, 1907, 1908])**.** *Let $u_0, ...u_n$ be a finite set of points $S \subseteq \mathbb{R}^d$. Each point u_i has a Voronoi cell V_{u_i} associated with it, defined as that part of the space which is closest to it, i.e. $V_{u_i} = \{x \in \mathbb{R}^d | \|x - u\| \leq \|x - v\|, v \in S\}$. The set of all Voronoi cells associated with the points u_i, $i \in \{1, n\}$ is the Voronoi diagram.*

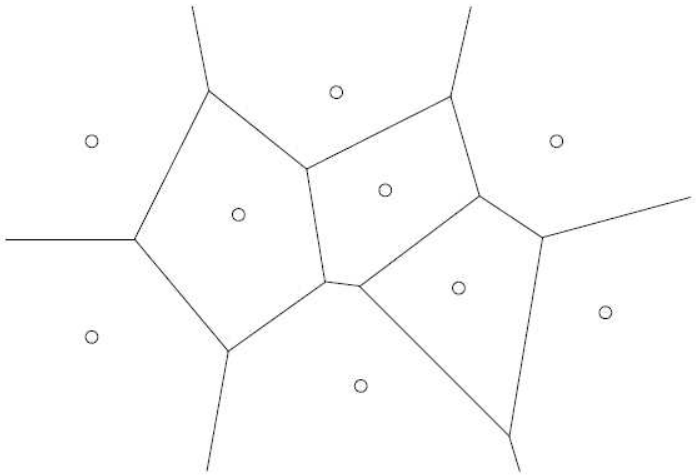


Figure 3.10: ***Example of a 2D Voronoi diagram [Edelsbrunner and Harer, 2010].*** *The area inside each cell is by definition closest to its center point.*

The above can be generalized by giving each point u_i a real weight w_{u_i} and defining a distance function $\pi_u(x)$. For example, when the points represent galaxies their weight can represent their mass and the distance function their domain of strongest gravitational attraction. These can be used to define the ***weighted Voronoi diagram***, whose cells are defined as the regions of space closest to a $u_i \in S$ using the distance function instead of Euclidian distance. Some useful properties of weighted Voronoi diagrams are:

Properties 3.4.

1. *Every point in $\mathbb{R}^d$ lies closest to some point in S, thus the Voronoi diagram covers the entire space.*

2. *The d-dimensional inside of a Voronoi cell is defined by the region closest to one of the points $u_i \in S$. The $(d-1)$-dimensional boundary of two cells is equally close to two points u_i and $u_j \in S$. Two boundaries intersect at a $(d-2)$-dimensional region closest to three points u_i, u_j and $u_k \in S$, etc. In case of the 2D example of figure 3.10, two cells define a Voronoi segment at their intersection and three cells a Voronoi vertex.*

The last of the properties above states that a d-dimensional Voronoi region lies closest to 1 point, a $(d-1)$ dimensional Voronoi region lies closest to 2 points, and so on till a 0-dimensional region is defined by $(d+1)$ points. Consequently, in general no $(d+2)$ or more Voronoi regions have non-empty intersection; it only happens when $(d+2)$ or more points lie on a common $(d-1)$-sphere. In this case the point set is called *degenerate*. Probabilistically, the change of this happening is zero, as a $(d-1)$-sphere has measure zero in d-dimensional space. When a point set is non-degenerate the points are said to lay in *general position*, which we will assume for the remainder of the section unless mentioned otherwise.

A dual triangulation can be obtained by connecting all d-dimensional points whose Voronoi cells meet along a $(d-1)$-dimensional boundary, giving us the:

Definition 3.17 (Delaunay triangulation [Okabe et al., 2000; Delone, 1934])**.** *Let $u_0, ...u_n$ be a finite set of points $S \subseteq \mathbb{R}^d$ in general position with associated Voronoi cells V_{u_i}. Connecting all the d-dimensional points whose Voronoi cells meet along a $(d-1)$-dimensional boundary gives the Delaunay triangulation. This triangulation is the dual of Voronoi diagram.*

The Delaunay triangulation for the Voronoi diagram above is shown in figure 3.11. Named after

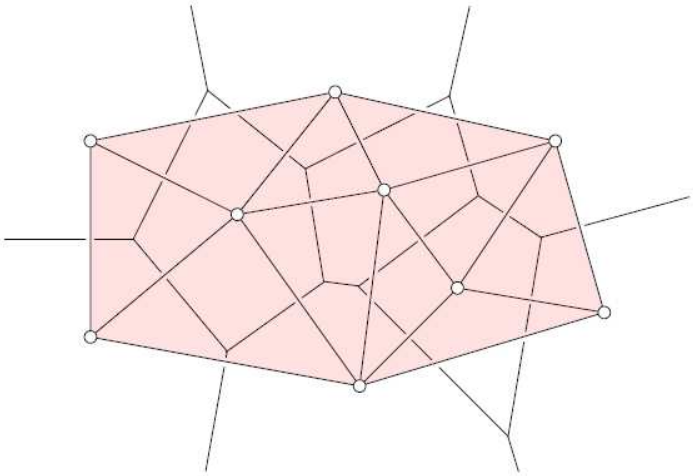


Figure 3.11: ***Example of a Delaunay triangulation [Edelsbrunner and Harer, 2010].*** *The triangulation is the dual of the Voronoi diagram of figure 3.10.*

its inventor, the Russian mathematician Delone, the triangulation has some useful properties:

Properties 3.5.

1. *Delaunay cells are clearly a convex combination of points and thus a simplex, thus the Delaunay triangulation is a simplicial complex.*

2. *Even stronger, amongst all possible triangulations, the Delaunay cells have minimal size and elongation.*

3. *It follows directly from the definition that the circumcircle of the Delaunay cell doesn't contain any other points besides its vertices. The center of the circumcircle is a vertex of the Voronoi diagram.*

Here, it is important that the point set is non-degenerate. To understand why, consider the 2D figure 3.12. On the left $(d + 2) = 4$ dotted Voronoi segments meet at a common vertex, resulting in an ambiguity in the definition of the Delaunay triangulation. An arbitrary small perturbation of any of the points brings the point set in general position and solve the issue. Just as the Voronoi diagram, the Delaunay triangulation can be generalized to a weighted variant.

3.4 Discrete Morse Theory

In previous sections we developed and examplified discrete topology, we now continue with discretizing Morse Theory. A fundamental part of continuous Morse theory are continuous Morse functions, so our first discretization step is discretizing them:

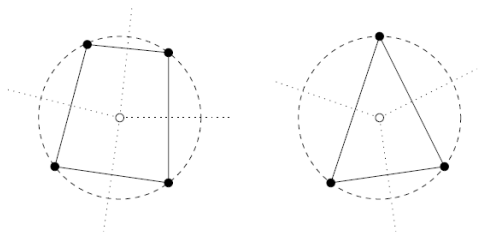


Figure 3.12: ***The importance of a non-degenerate point set [Vegter, 2012].*** *The degenerate point set (left) results in a degenerate Delaunay triangulation. A non-degenerate point set (right) results in a valid Delaunay triangulation. The dotted lines are Voronoi segments.*

Definition 3.18 (Discrete Morse Function)**.** *Let K be a simplicial complex defined over $\mathbb{R}$. A discrete function $f : K \to \mathbb{R}$ assigns a real value $f(\sigma_k)$ to each simplex $\sigma_k \in K$. The function f is a discrete Morse function if and only if for all $\sigma_k \in K$:*

1. *there exists at most one facet α_{k-1} of σ_k such that $f(\sigma_k) \leq f(\alpha_{k-1})$;*

2. *there exists at most one cofacet β_{k+1} of σ_k such that $f(\sigma_k) \geq f(\beta_{k+1})$;*

3. *all critical points have distinct function values*[3].

Example 3.6. *Figure 3.13 shows a simple example of a Morse and not-Morse function.*

In words, the non-zero Hessian conditions is replaced with a condition on the value of the simplices: a simplex has a higher value than its facets and a lower values than its cofacets. Only one exception is allowed in each case. [Forman, 1998] showed that only condition 1 or condition 2 of discrete Morse functions can be satisfied for a given simplex. I.e. actually condition 1 and 2 can be seen as a single condition with an 'either-or' construction. Although perhaps not directly obvious, these conditions assure a preferential discrete gradient flow (to be defined shortly), just as the non-zero Hessian did in the continuous case. Clearly, when no preferential direction can be defined the simplex is critical:

[3] Sometimes the third condition is dropped but in this thesis it will always be required.

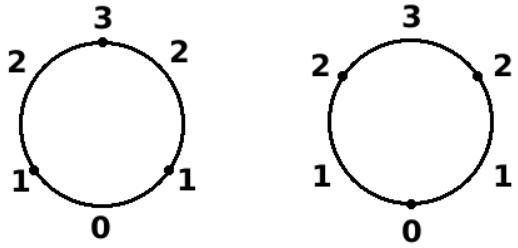


Figure 3.13: ***Two discrete functions [Forman, 1998].*** *The left function is not Morse as $f^{-1}(0)$ violates rule 2 $f^{-1}(3)$ violates rule 1. The right function is Morse.*

Definition 3.19 (Critical k-simplex)**.** *A k-simplex σ_k is a critical k-simplex of discrete Morse function f if:*

1. *there exists no facet α_{k-1} of σ_k such that $f(\sigma_k) \leq f(\alpha_{k-1})$; and*
2. *there exists no cofacet β_{k+1} of σ_k such that $f(\sigma) \geq f(\beta_{k+1})$.*

*A critical k-simplex has **order** k.*

Note that this definition introduces an asymmetry not existing in the continuous case. In the continuous case a minimum and a maximum are both points in space. Here in the discrete case, a minimum is a critical vertex (0-simplex) whereas in 3D a maximum is critical tetrahedron (3-simplex). As a direct consequence, when f is a discrete Morse function it doesn't follow automatically that $-f$ is a discrete Morse function as well. However, $-f$ over the dual complex is Morse and defines the same topology.

If a Morse function isn't critical, there are two possibilities left: simplex σ_k has (i) a facet with a larger value or (ii) a cofacet with a smaller value. In either case, a preferential relation is established. Using this, we define a:

Definition 3.20 (Discrete gradient vector field)**.** *Let f be a discrete Morse function of simplicial complex K. Then a discrete gradient vector field is defined by coupling simplexes in gradient arrows (also called gradient pairs) in the following way:*

1. *If simplex σ_k has a lower valued cofacet α_{k+1}, then $[\sigma_k, \alpha_{k+1}]$ form a gradient arrow.*
2. *If simplex σ_k has a higher valued facet β_{k-1}, then $[\sigma_k, \beta_{k-1}]$ form a gradient arrow.*

By convention the lowest-valued simplex is the tail and the highest-valued simplex is the head. Therefore, a discrete gradient actually point in opposite direction of its continuous counterpart.

Example 3.7. *An uncomplicated example of a discrete gradient is shown in figure 3.14 and a more elaborate example is shown on the top of figure 3.16.*

To define discrete Morse theory in analogy with its continuous counterpart we still need the discrete equivalence of a sublevel set:

Definition 3.21 (Discrete sublevel set)**.** *Let K be a simplicial complex with discrete morse function f and denote for simplex σ_k its facets with β_{k-1}. Consider a threshold level $a \in \mathbb{R}$, then the discrete sublevel set is defined as:*

$$K(c) = \bigcup_{\sigma_k \in K, f(\sigma_k) \leq c} \bigcup_{\beta_{k-1} \leq \sigma_k} \beta_{k-1} \tag{3.2}$$

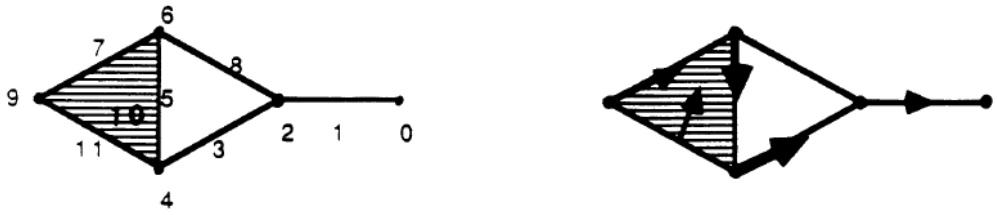


Figure 3.14: ***An example of a discrete gradient [Forman, 1998].*** *An uncomplicated density field over a simplicial comples (left) and its discrete gradient (right).*

In words: $K(c)$ is the set of all simplices on which $f \leq c$, as well as all of their faces. $K(c)$ is a full subcomplex of K.

Nicely, [Forman, 1998] showed that to check whether a simplex β with $f(\beta) > c$ might be in $K(c)$ because there is a simplex α with $\beta < \alpha$, it is enough to consider all facets of α.

The first part of continuous Morse theory states when the manifolds of rising sublevel sets are diffeomorphic, i.e. 'not essentially different'. For an equivalent discrete notion we require:

Definition 3.22 (Simplicial collapse)**.** *Let K be a simplicial complex with simplex σ_k and one of its cofacets α_{k+1} in K. If σ_k is not the face of any other cell besides α_{k+1} and simplicial complex $L = K - (\sigma_k \cup \alpha_{k+1})$, we say K can be collapsed onto L. If K can be transformed into L by a finite sequence of such operations, we write $K \searrow L$.*

In essence, simplicial collapse allows simplices to be reduced to their most essential components. Figure 3.15 illustrates simplicial collapse: on the left the simplex consisting of triangle with boundary is collapsed to a point. By analogous operatios, the figure on the right can collapsed to a circle.

(a) Stepwise simplicial collapse from triangle to point.

(b) This complex can be collapsed to a circle.

Figure 3.15: ***Illustrations of simplicial collapse [Forman, 1998].***

Using, simplicial collapse, we can state the main theorems of discrete classical Morse Theory:

Theorem 3.3 (Discrete classical Morse Theory, part I)**.** *Let K be a simplicial complex and $a < b \in \mathbb{R}$ in the range of a discrete Morse function $f(K)$. If $[a, b]$ contains no critical values of f, then $M(a) \searrow M(b)$.*

Theorem 3.4 (Discrete classical Morse Theory, part II)**.** *Consider the same setup as in part 1 and let σ_k be a critical simplex of order k with $f(\sigma_k) = c$. If $a < c < b$ and $f^{-1}([a, b])$ contains no other critical simplicies besides σ_k, than $M(b)$ is homotopy equivalent to $M(a)$ with a k-simplex attached to its boundary.*

Analogously with the continuous case, at every non-critical simplex the discrete gradient describes a preferred direction. By following this direction, we obtain the discrete variant of the integral line:

Definition 3.23 (Discrete integral line or V-path)**.** *A discrete integral line, in literature called a V-path, is a strictly decreasing alternative sequence of $(k+1)$-simplices a^i_{k+i} and k-simplices β^j_k of the form:*

$$\alpha^0_k, \beta^0_{k+1}, \alpha^1_k, \beta^1_{k+1}, ..., \alpha^n_k, \beta^n_{k+1}$$

where each pair $\{\alpha^i_k, \beta^i_{k+1}\}$ form a gradient pair and α^{i+1}_k is a facet of β^i_{k+1}.

Intuitively, a V-path just consists is following the arrows of the discrete gradient, see figure 3.16 bottom left. This allows us to divide discrete space in regions which have the same origin or destination:

Definition 3.24 (Discrete ascending/descending k-manifold)**.** *The discrete stable or ascending manifold of critical k-simplex σ_k of discrete Morse function f over simplicial complex K is the set of k-simplices that belong to at least one V-path with destination σ_k. The discrete unstable or descending manifold is the set of k-simplices reached by at least one V-path with origin σ_k.*

Note that a discrete k-manifold only contains k-simplices, which leaves holes between them and makes it to define manifold intersections. Therefore, the following definition is of more practical use:

Definition 3.25 (Extended discrete ascending/descending k-manifold)**.** *An extended discrete ascending/descending k-manifold discrete is an ascending/descending k-manifold together with its cofacets/facets and their extended discrete ascending/descending k-manifolds.*

The bottom right of figure 3.16 shows an (extended) ascending manifold for our example gradient field. The set of ascending/descending manifolds forms the:

Definition 3.26 (Discrete Morse complex (DMC))**.** *The discrete Morse complex (DMC) of Morse function f is the set of its extended ascending/descending manifolds.*

Contrary to the continuous case no transversality conditions need to be imposed, as this is automatically taken care of by the tessellation itself. This allows us to define directly the discrete Morse k-cell and makes the introduction of a discrete 'Morse-Smale' complex superfluous.

Definition 3.27 (Discrete Morse k-cell)**.** *A discrete Morse k-cell is the intersection of two extended ascending and descending discrete manifolds.*

3.5 Practical implementation

How to use the above practically? The point set of a N-body simulation or of a density data cube can be triangulated with a Delaunay triangulation, giving a simplicial complex. For N-body simulations the density at each point can be computed using a scheme like DTFE (cf. section 4.2). In the density case the density at each point is already given. In this section we will discuss how to go from a density at each point to a discrete Morse density function over the whole complex. And how to go from there to a valid discrete gradient.

Discrete Morse functions have three restricting properties (cf. definition 3.18). (i+ii) for each simplex at most one facet may have a higher function value or at most one cofacet may have a

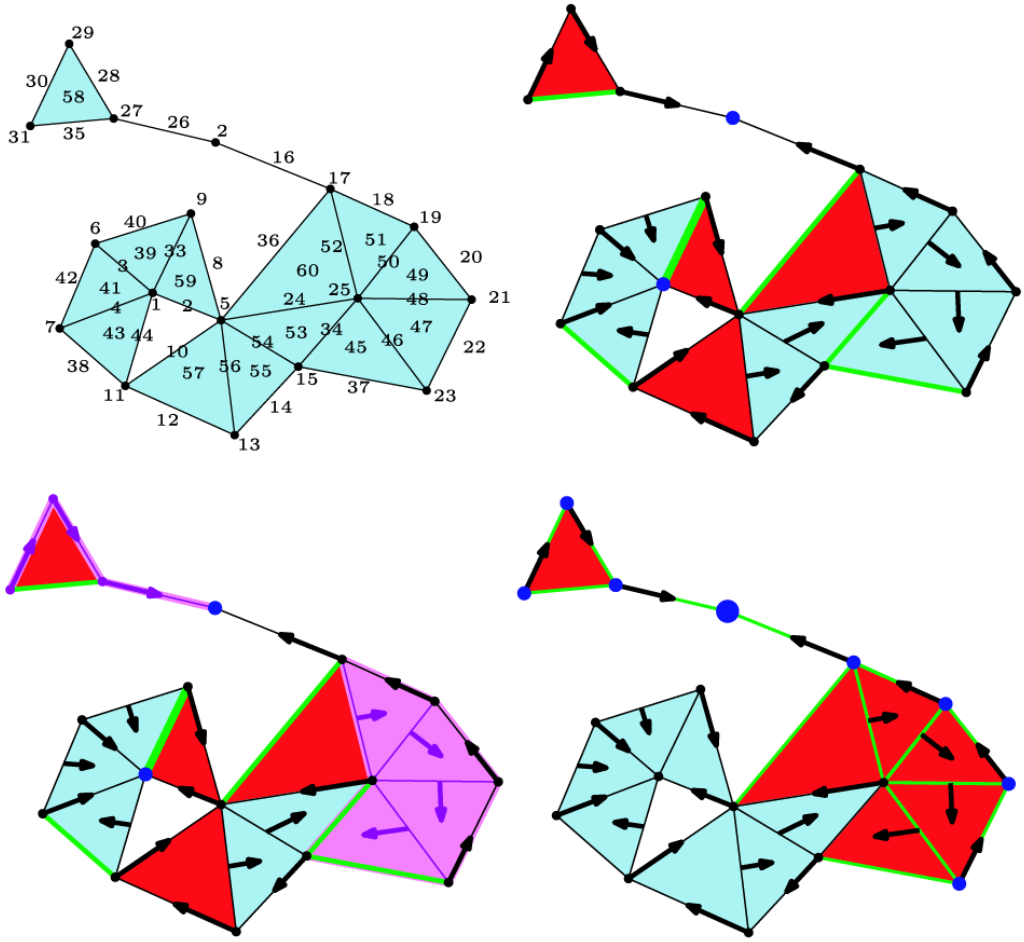


Figure 3.16: ***Example of a discrete simplicial complex with discrete gradients and extended manifolds [Sousbie, 2011].*** *A 2D simplicial complex with a Morse function (top left) and its corresponding discrete gradient (top right), with the critical simplices color coded red (triangles), green (segments) and blue (vertices). The gradient results in two V-paths, two of them are shown in pink (bottom left). A corresponding ascending 2-manifold is formed by the blue vertices (bottom right), combined with the green segments and red triangles it becomes an extended ascending 2-manifold.*

lower function value and (iii): all critical simplices should have different function values. Note that the last constraint is in practise always satisfied for cosmological density fuctions. Even if the function itself doesn't satisfy the constraint, an infinitesimal change in density values will make it satisfy the constraints without affecting the physics. To satisfy the first constraint, the easiest and surest way to implement a discrete Morse function is to assure all facets have a lower value and all cofacets a higher value. I.e. to make all simplices critical. To do so we start with known densities at the vertices ($\rho(\sigma_0)$) and extend the discrete function to higher simplices by giving it the density value of its highest valued face plus a little bit extra. To still qualify as a Morse function the exact value of a little bit extra is not even important, it can simply be an infinitesimal amount ϵ. Concretely, this gives us the following receipt to build a discrete Morse function ρ_D:

$$\rho_D(\sigma_k) = \begin{cases} \rho(\sigma_k) & \text{for } k = 0 \\ \max\left(\rho_D(\text{facet}[\sigma_k])\right) + \epsilon^k \sum \rho_D(\text{vertex}[\sigma_k]) & \text{for } k > 0 \end{cases} \tag{3.3}$$

Note that the receipt guarantees that if every vertex has a different density, the same holds for all higher dimensional simplices in the triangulation.

A discrete Morse function built by the definition above can be said to be strictly increasing along increasing chains of cofaces. Under this assumption, the definition of a discrete sublevel set (cf. definition 3.21) can be simplified: when a simplex is in the discrete sublevel set its faces are in automatically as well. This simplified discrete sublevel set is called a:

Definition 3.28 (Filtration)**.** *Let f be a strictly increasing Morse function on simplex K with m simplices. Starting with an empty set, we add the simplices of K one by one in order increasing function value. This gives the following chain of nested subcomplexes:*

$$\varnothing = K_0 \subseteq K_1 \subseteq ... \subseteq K_m = K \tag{3.4}$$

where due to the condition that the Morse function is increasing, each K_i is a full subcomplex of K_{i+1}. Such a chain of subcomplexes is called a filtration.

A filtration will play an important role in building the DMC as we will see in section 3.6. For the moment, we set the filtration aside and continue with defining a suitable discrete gradient. Actually, from the function definition above a discrete gradient follows quite naturally: many simplices have facets or cofacets which differ only by an arbitrarily small amount ϵ. As this is an arbitrarily small amount, the function values at such a facet-cofacet pair can be interchanged without changing the physics. Interchanging the density values automatically makes the pair non-critical and allows them to be combined by a gradient pair. Doing so for all simplices in the triangulation gives (quite probably) many gradient pairs, of which some of them will form V-paths. The procedure above is illustrated in figure 3.17. It should be noted that there are many ways to implement a discrete Morse function and gradient field [Lewiner, 2002; Ni et al., 2004]. Although the method explained here is not necessarily mathematically optimal, it does agree rather well with the physics and is computationally achievable.

3.6 Towards the persistence diagram

3.6.1 Cycles and persistence

The DMC provides us with a natural way to divide a point set (of galaxies) in a cosmic structure of voids, walls, filaments and nodes; providing us with a mathematical handle on these morphological components. Large galaxy catalogues contain millions of galaxies and their Morse complexes will be enormous. For such large complexes, a simple description of the DMC is not very insightful and a more systematic way to explore the structures of the DMC is required. This section focusses on two topics: (i) identifying the relevant features within a DMC and (ii) identifying their relevance. Regarding the first, ***homology*** helps out by defining equivalence relations on chains of simplices. Homology is mathematically rather involved, therefore in this section we will largely gloss over it. The reader who wants to know the details is referred to appendix A.2. Using homological equivalence relations, chains of simplices which make up the same structure in the DMC are identified. Such chains of simplices are named:

Definition 3.29 (k-cycles)**.** *Let K be a simplicial complex of dimension d. A k-dimensional feature within such a complex is called a k-cycle, with $0 \leq k < d$. In particular a:*

0-cycle *is an independent component, i.e. a set of simplices not linked to the rest of the complex.*

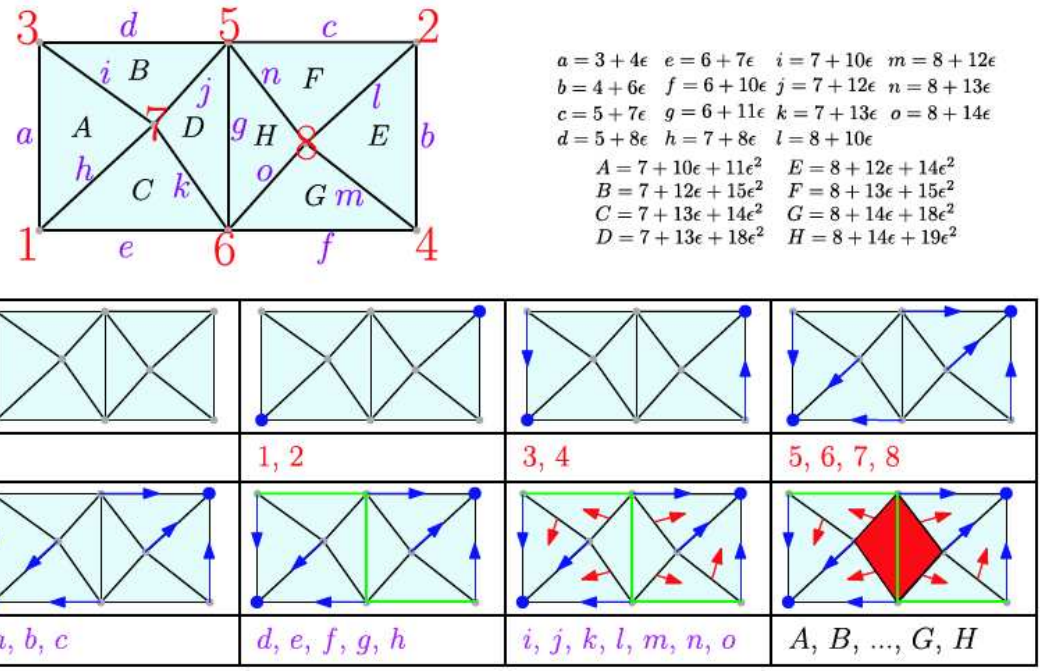


Figure 3.17: ***An example of a discrete gradient [Sousbie, 2011].*** *A simplex with values for the vertices (top left), the resulting values for the higher dimensional simplices following the description of the discrete Morse function presented here (top right) and its stepwise creation of a discrete gradient throughout the simplex (bottom). Stepwise explanation of the discrete gradient creation: (i) initial field; (ii) the low-valued not connected vertices 1 and 2 are initially critical (blue dots); (iii) these vertices differ an ϵ amount with vertices a and b and thus are paired, vertices 4 and 4 don't differ an ϵ amount with any segment and thus remain critical; (iv) vertices 5, 6, 7 and 8 are all paired with unpaired segments with which they differ the smallest ϵ amount; (v) segments a, b and c are already paired; (vi) of the segments d till h segments d, f and g turn out to be critical (green lines), the other were already paired; (vii) from the segments i till o, those who weren't paired yet are paired with the triangle with which they differ the smalleset ϵ amount; (viii) all triangles safe for D and H were already paired, whereas D and H turn out to be critical. The remaining gradient field shows nice V-paths from triangle D to vertex 1 and from triangle H to vertex 2.*

1-cycle *is a tunnel, i.e. a set of simplices forming a loop with a hole in the middle.*

2-cycle *is a shell, i.e. a set of simplices bounding an emptier 3D volume.*

An example of each of the k-cycles named explicitly in the definition is shown in figure 3.18. As this figure also shows, k-cycles can be created and destroyed. To understand cycle creation and destruction, let's consider the simple 3D simplex K shown in figure 3.19. It consists of vertices s, t, u, v and w together with some of its cofacets. Suppose f is a Morse function defined over K following the function prescription of section 3.5. In this prescription, a given simplex has always a higher value than its facets and a lower value than its cofacets. Using f we can build a filtration in which the simplices enter one by one as we take an increasing sublevel set of the function[4]. The figure shows tile by tile the

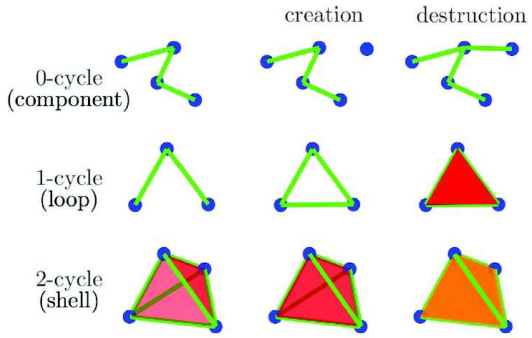


Figure 3.18: ***Cycles, their creation and their destruction [Sousbie, 2011].***

[4]This description is called the *Incremental Betti algorithm*, cf. [Delfinado and Edelsbrunner, 1993].

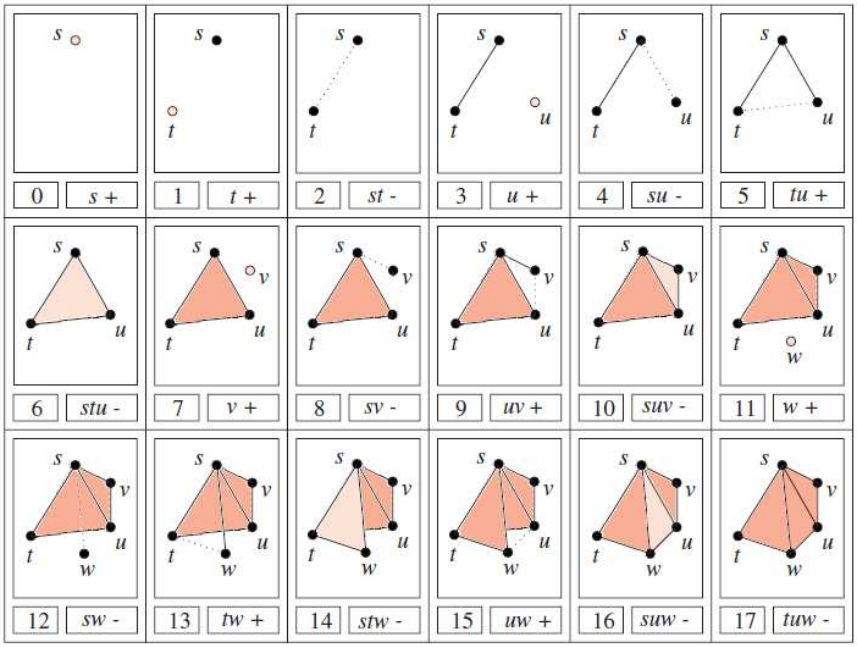


Figure 3.19: ***Construction of a simplicial complex, based upon [Edelsbrunner et al., 2002].*** *Using the filtration of a Morse function f defined over simplicial complex K, the complex is constructed simplex by simplex. The introduction of every new simplex either creates a cycle or destroys it. The text below each tile states the index of the simplex entering the filtration and the vertices of the simplex, followed by a plus for cycle creation or a minus for cycle destruction. The last step is the addition of triangle tuv which is hidden from view by the other triangles.*

construction of the simplex. The text below each tile states the index of the simplex entering the filtration, the vertices of the simplex and a plus for cycle creation or a minus for cycle destruction. In the figure we observe the following:

1. Each vertex creates a new component (snapshots 0, 1, 3, 7 and 11) and thus increases the number of 0-cycles.

2. Each segment either merges two components thus decreasing the number of 0-cycles (2, 4, 8 and 12) *or* creates a tunnel and thus increases the number of 1-cycles (snapshots 5, 9, 13 and 15).

3. Each triangle either fills a tunnel and thus decreases the number of 1-cycles (snapshots 6, 10, 14 and 16) *or* creates a shell and thus increases the number of 2-cycles (snapshots 17).

4. And although not visible: each tetrahedron would fill a shell and thus decreases the number of 2-cycles.

A critical k-simplex has order k (cf. definition 3.19), so the above can be summarized as follows. When a simplex of order k enters the filtration either (i) a k-cycle is born or (ii) a $(k-1)$-cycle is destroyed. Formulated inversely: each m-cycle is created by a m-simplex and destroyed by a $(m+1)$-simplex. The m-simplex which creates the component and the $(m+1)$-simplex which destroys it are together called a ***persistence pair*** and their difference in fuction value is their ***persistence*** [Edelsbrunner et al., 2002]. We are now in a position to concretely answer the first question posed in the introduction of this chapter: *what is a feature?*. A feature of a smooth manifold is a k-cycle. To answer the second question, *which features are relevant?*, we need to

take a closer look at the concept of persistence. We note that in the discussion above persistence is defined in an intuitive manner. Using homology, persistence can be defined more rigorously. We refer the reader interested in a more formal definition to appendix A.3. An essential part of persistence is pairing. Which $(m+1)$-simplex destroyed a cycle is obvious but it is not directly clear which m-cycle gave birth to it. For example, when a segment merges two components there are two vertices involved and when a triangle fills a tunnel it involves three segments, et cetera. The most elementary rule is the:

Definition 3.30 (Elder Rule)**.** *Consider an m-cycle destroyed by an $(m+1)$-simplex. The $(m+1)$-simplex is paired with the cycle's unpaired m-simplex which entered the filtration latest. I.e. the m-simplex which latest entered the filtration is paired first, thus the lowest (eldest) simplices live longest.*

Example 3.8. *Using the Elder Rule, persistence has a very intuitive interpretation. Consider the the 1D manifolds A and B and their discretizations A' and B' on the top of figure 3.20, together with a Morse density function defined on them. In 1D there are only two types of non-degenerate critical points: minima and maxima. Their pairings via the Elder Rule are indicated at the top of the bottom figure. The persistence of each pair is shown by the green arrows. We see that the persistence of a pair clearly correlates with the relevance of a feature.*

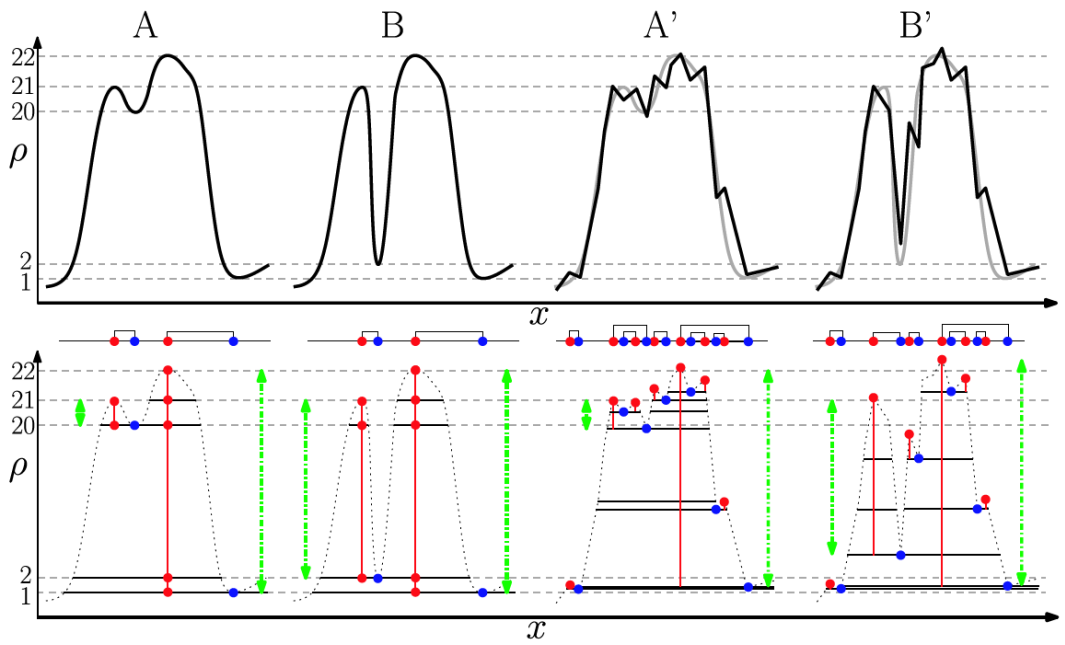


Figure 3.20: ***Visual illustration of persistence [Sousbie, 2011]****. Top: 1D functions A and B with their discretizations A' and B'. Bottom: Their minima (blue) and maxima (red), paired via the Elder Rule. The length of the green arrows indicates the persistence of the feature.*

The Elder Rule has one caveat: it allows pairing of completely disconnected simplices, one on one side of the density field and one on the other. In physical scenarios often an additional constraint is added: the simplices to be paired should be connected via a gradient path.

Now we can also answer the second question posed in the introduction: *which features are relevant?*. The relevance of a feature is directly related with its persistence. Low persistent features are not so relevant whereas high persistent features are very relevant.

3.6.2 Simplification

To obtain a global view on the topology of the manifold less persistent features can be taken out by cancelling low persistence pairs, a process called ***simplification***. Persistence pairs are formed by critical simplices with order difference 1. Thus in 3D three types of persistence pairs can be found: (i) a minimum with a 1-saddle forms a 0-cycle; (ii) a 1-saddle with a 2-saddle forms a 1-cycle and (iii) a 2-saddle with a maximum forms a 2-cycle. The three types of persistence pairs require different cancellation procedures [Gyulassy and Natarajan, 2005; Gyulassy et al., 2006], which we will discuss now.

Actually, the minimum-1-saddle and 2-saddle-maximum persistence pairs are symmetrical to each other and they can be cancelled in an analogous way. Their cancellation procedure is shown in figure 3.21. At image (a) the twelve 3-simplices of the complex affected by the cancellation of the saddle-maximum pair are colored. Image (b) shows the 2-saddle s with its surrounding maxima m and n, where s and m are scheduled for cancellation. Effectively, the cancellation process can be visualized as merging m and s into n. The resulting complex is shown in image (c). We see that the yellow cell died in the process. All arcs and discs that originally flowed to maximum m are rerouted to maximum n. This is examplified by the arcs a_1, a_2 and a_3 in image (b) rerouted as arcs b_1, b_2 and b_3 in image (c). Rerouting is practically achieved by inversing the gradient flow between m and s. Image (d) shows that all other discs and arcs of the original 12 cells are rerouted as well, making three more cells to disappear.

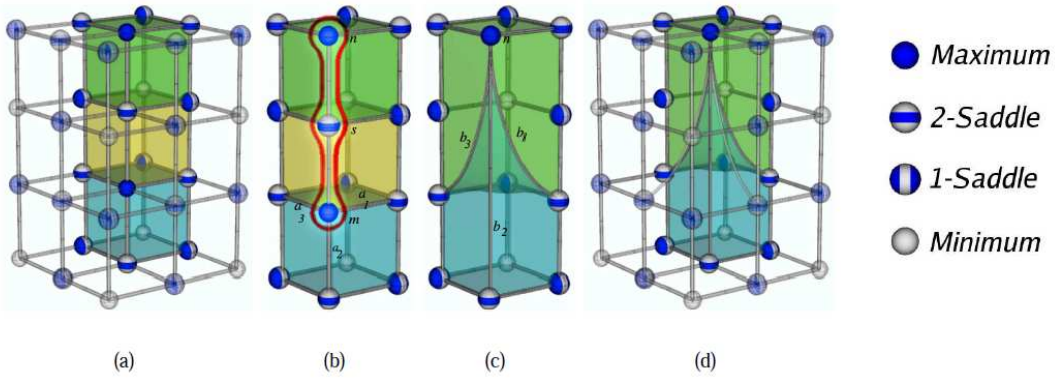


Figure 3.21: ***A saddle-max cancellation [Gyulassy and Natarajan, 2005].*** *See text for explanation.*

Saddle-saddle cancellations are a bit more cumbersome, we illustrate it using figure 3.22. In analogy with the saddle-max case four regions are effected, but for clarity of presentation we show only one in image (a). The other regions behave analogously. s_1 and s_2 are the two saddles which are paired and simplified. This is achieved by rerouting all arcs and discs which flow to s_2 to t_1 and those flowing to s_1 to t_2. In the process some additional arcs are created, for example the arc b between t_1 and t_2. Due to simplification some existing cells are reshaped, as shown in image (c) and (e) and (f). Due to the addition of new arcs, new cells are created: in image (a) we had 3 cells but in image (b) we have 4 cells. Image (d) shows the newly created cell. In principle, saddle-saddle cancellations can extend the number of cells in a complex significantly. Nevertheless the manifold is simplified as the number of critical points decreases. Furthermore, these newly created cells will be simplified out in a later stage, when their extrema are paired with a saddle and cancelled.

It should be noted that not all persistence pairs can be cancelled. In the saddle-extremum case problems occur when two integral lines beginning at a 2-saddle flow to the same maxi-

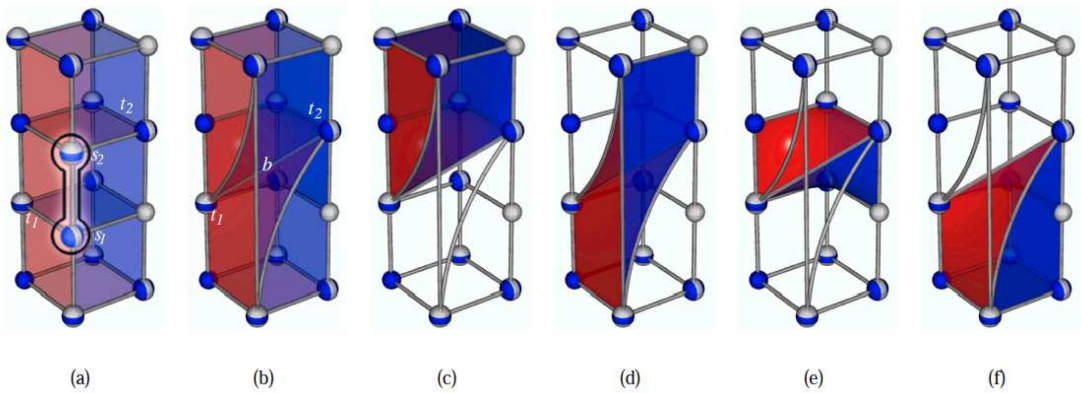
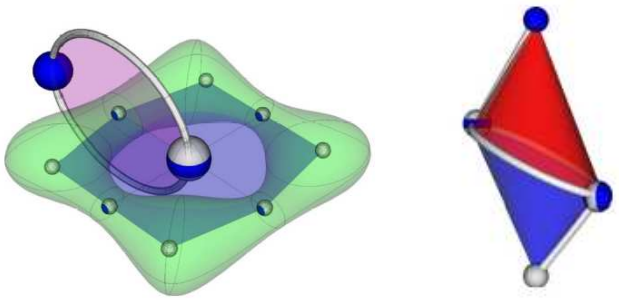


Figure 3.22: ***A saddle-saddle cancellation [Gyulassy and Natarajan, 2005]****. See text for explanation.*

mum or two integral lines from a minimum flow to the same 1-saddle. Cancellation of the first pair will leave the integral lines flowing to the 2-saddle without final destination. In the second case integral lines flowing away from a 1-saddle won't have an origin. Such impossible cancellations are called *strangulations* and an example is given in the left tile of figure 3.23. In the saddle-saddle case another type of impossible cancellation might occur, see the right tile of previous figure. Impossible saddle-saddle cancellations occur when a cell has exactly two boundaries, one connecting the saddles to a minimum and the other connecting the saddles to a maximum. Cancellation of the saddles will directly connect the minimum with the maximum. If it would be done a cell called *pouches* is created and such cells are not allowed by the DMC. Persistence pairs with impossible cancellations do occur in practise, but their number is quite



(a) Cancellation of the maximum 2-saddle pair will lead to a strangulation [Gyulassy and Natarajan, 2005]

(b) Cancellation of a 1-saddle 2-saddle pair will lead to a pouch [Gyulassy et al., 2006].

Figure 3.23: ***Impossible cancellations.***

limited. Forcing their cancellation and reconfiguring the manifold accordingly often works in practise [Sousbie, 2011].

3.6.3 The persistence diagram

For a large simplicial complex like cosmic structure, many persistence pairs will be found. For each cycle dimension, the persistence pairs indicating the birth and death of the cycle can be

grouped together in a:

Definition 3.31 (Persistence diagram)**.** *Let K be a simplicial complex of dimension d where a strictly increasing Morse function f has paired cycles in persistence pairs. Consider a 2D diagram with on the horizontal axis the birth of i-cycles and on its vertical axes the death of i-cycles, with $0 \leq i \leq d$. Than each persistence pair related to a i-cycle can be represented by a dot in the diagram with its x-coordinate the cycle birth-value and its y-coordinate the cycle death-value.*

Persistence diagrams have several interesting properties:

Properties 3.6.

1. *As every cycle is born on a lower value than it dies, persistence diagrams have only dots in the upper left half of the plane.*
2. *The diagonal is a birth=death diagonal: cycles born on their die at the same moment, i.e. they actually don't exist.*
3. *The orthogonal distance from a point to the diagonal corresponds with the persistence of the cycle.*

Example 3.9. *The left panel of figure 3.24 shows the persistence diagrams of discretized 1D density functions A' and B' of figure 3.20. In a single glance it can be seen that A' has only one relevant feature and B' two.*

A persistence diagram gives much information at a glance, but 2D diagrams are difficult to compare visually. Therefore, sometimes a 1D summary can be very useful. Three 1D summaries that will be often used are:

1. **Betti numbers** indicate the amount of cycles alive at a certain sublevel set. To be counted as cycle living at a certain threshold level α of Morse function ρ, it has to be born before $\rho = \alpha$ and die after. In the persistence diagram, this corresponds exactly to the square resting on the diagonal at the threshold level. In 3D space there are three cycles dimensions thus three Betti numbers: β_0, β_1 and β_2. Physically, they give respectively the number of components, tunnels and shells as function of α.
2. **Lifetimes** are directly related to persistence. A curve showing the number of cycles as function of their persistence provides relevant insight in the presence of long lived (important) and short lived (less important) features. Such a curve corresponds to considering an axis orthogonal to the diagonal. Mathematically, if we have a cycle σ_k with birth density ρ_b and death density ρ_d, its lifetime $L(\rho_b(\sigma_k), \rho_d(\sigma_k))$ is given by:

$$L(\rho_b(\sigma_k), \rho_d(\sigma_k)) = \frac{\rho_d - \rho_b}{\sqrt{2}}$$

Given the large range of density values we will usually work in log-log diagrams. In such diagrams the distance to the diagonal corresponds with the birth-death fraction:

$$\begin{aligned} L_{\log}(\rho_b(\sigma_k), \rho_d(\sigma_k)) &= \frac{\log_{10}(\rho_d) - \log_{10}(\rho_b)}{\sqrt{2}} \\ &= \frac{1}{\sqrt{2}} \log_{10}\left(\frac{\rho_d}{\rho_b}\right) \end{aligned} \tag{3.5}$$

3. **(Product) mean density curves** are relevant in the context of structure formation and show the number of cycles as function of their density. This provides information about the number of low density and high density cycles. In the persistence diagram it corresponds to considering an axis parallel to the diagonal. Mathematically, using the same notation as with the lifetime curves, we get for the mean density $MD(\rho_b(\sigma_k), \rho_d(\sigma_k))$:

$$MD(\rho_b(\sigma_k), \rho_d(\sigma_k)) = (\rho_b + \rho_d)/2$$

In the log-log diagrams we work with here, we use the product mean density instead:

$$\begin{aligned} MD_{\log}(\rho_b(\sigma_k), \rho_d(\sigma_k)) &= \tfrac{1}{2}(\log_{10}(\rho_b) + \log_{10}(\rho_d))/2 \\ &= \log_{10}\left(\sqrt{\rho_b \cdot \rho_d}\right) \end{aligned} \tag{3.6}$$

The three 1D summaries are visualized on the right panel of figure 3.24. In cosmological datasets

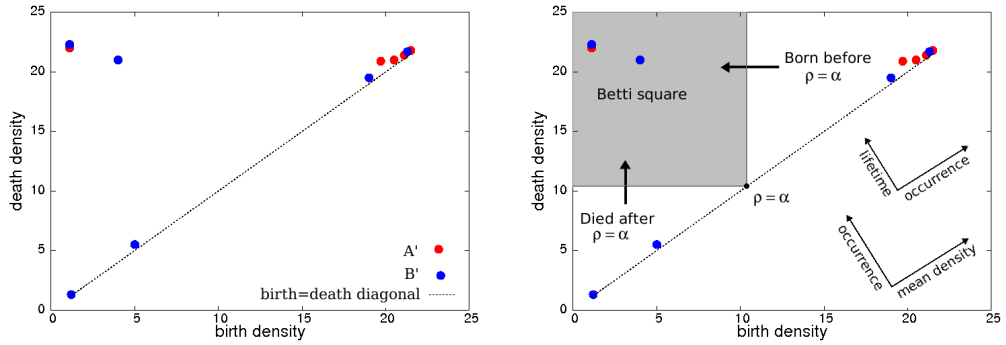


Figure 3.24: ***A persistence diagram.*** *Left the persistence diagram of figure 3.20 and right its three 1D persistence measures.*

hundreds of thousands of cycles will be found. Putting them all in a diagram will give a computationally very demanding point set. Therefore, in many situations below persistence diagrams will be transformed to 2D persistence histograms. We obtain histograms by putting a raster of boxes over the diagram and place each persistence point in its corresponding box. Subsequently the boxes are transformed to densities, i.e. to the number of cycles in each box per Mpc^3. It should be noted that the exact density depends on the chosen box-size. Changing from persistence diagram to a persistence histogram has an additional advantage: persistence histograms can be box-wise substracted from each other. This results in a persistence difference diagram, showing where and how much two persistence diagrams differ. Physically, the power of such diagrams is that it shows at a single glance which of the two density fields underlying the persistence diagram has more persistent structure.

Persistence diagrams are the culmination of this chapter. Founded on a mathematically sound basis, they show all physical features and their relevance in a single glance. Throughout the rest of this thesis, we will use persistence diagrams and their 1D summary curves to analyse the topology of cosmic structure.

CHAPTER 4

Algorithms and software

Previous chapters described the cosmic web and a topological formalism to capture the morphology of the web. Building on the concepts developed there, this chapter will explain how the topological formalism is technically implemented. The main output will be a Discrete Morse Complex (DMC, cf. definition 3.26): a mathematical structure in which all the voids, walls, filaments and nodes are identified in the form of simplices, the building blocks of discrete topological space. The DMC simplices can be ordered based on density value, such an ordered set is called a filtration (cf. definition 3.28). The filtration serves as basis to determine which simplices form a persistence pair (cf. section 3.6), i.e. a feature like a component or tunnel. Persistence pairs are the points of persistence diagrams, the main topological tool this thesis uses for analysis.

In the end we are after the influence of DE on the morphology of cosmic structure. Therefore, before topology comes into play, a first step is the creation of model universes which are ruled by different forms of dark energy. Using the galaxy distributions (i.e. point sets) from such simulations, two different approaches exist:

1. Based on the particle positions a DMC can be computed directly.

2. The particle positions can be used to compute the density field, which is sampled on a regular grid. The DMC is computed on this regular grid.

The first method is fully adaptive to the density field: high density regions are sampled in more detail. The second method samples space at regular intervals and thus has more samples in low density regions. In general the output of both will be analogous, but not similar. Both methods will be considered at various stages in this thesis but with a preference for the second, as low density regions will show the largest difference between DE models.

The sequence of steps above, from data generation to persistence diagram, form a data pipeline. A schematic illustration of the pipeline is shown in figure 4.1. In that figure from left to right, the following algorithms and software packages are mentioned:

1. Physically realistic particle distributions are generated by V. Springel's GADGET-2, with dark energy models by P. Bos.

2. To generate a density field on a regular grid, we use DTFE version 1.2 by R. v.d. Weygaert, W. Schaap and M. Cautun.

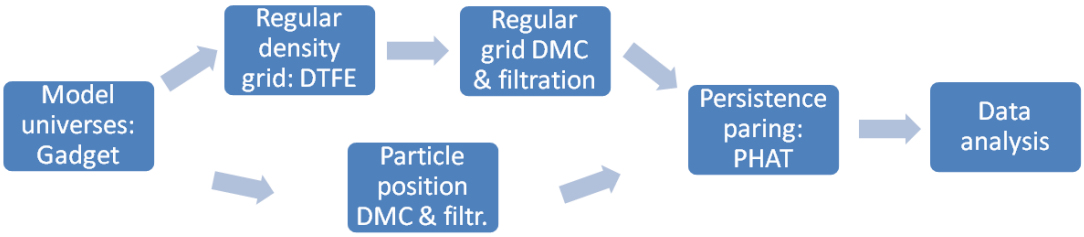


Figure 4.1: ***Graphical overview of the data pipeline.***

3. The DMC directly based on the particle positions is computed using P. Pranav adapted version of D. Morozov's Dionysus. The DMC of the density case follows from P. Pranav's implementation of P. Bendich's ideas.

4. The persistence pairs are identified using U. Bauer's PHAT, with adaptations of P. Pranav.

5. Data analysis is performed with custom methods or Matlab.

In the sections below, each element of the data pipeline is considered in more detail. For each element we (i) explain the general idea behind the algorithm or software; (ii) give a description of the specific implementation of these ideas in the algorithm and (iii) mention the settings or input we used.

4.1 Gadet 2: model universes

Creation of model universes can be done in two ways: using a fluid approximation or with N-body simulations. The first are computationally less demanding but loose their accuracy in highly non-linear cases (cf. section 2.5). N-body simulations on the other hand are much more computationally demanding but retain their accuracy also in the non-linear domain. At low redshifts structure is highly non-linear. As differences between various DE models might be small, we choose N-body simulations for optimal accuracy. Based on the disquisition of [Hockney and Eastwood, 1981], the general ideas behind N-body simulations are set forth in subsection 4.1.1. Subsection 4.1.2 explains how these ideas are implemented in the N-body code Gadget 2. Finally, subsection 4.1.3 gives the details of the Gadget 2 settings used in this thesis.

Within the data pipeline: the input of this step is (i) a law modeling gravitational effects of some dark energy model and (ii) a random number as input for the initial conditions. The output are data cubes at various z. Each data cube contains the positions of n particles who evolved under the gravity of matter and the DE model assumed as input.

4.1.1 N-body algorithms

N-body simulations are used to solve the N-body problem. In a gravitational context the N-body problem can be stated as: suppose we have n particles indexed by i with initial positions x_i, initial velocities v_i and masses m_i, how do the positions and velocities of these particles evolve in time? The simplest and most intuitive solution is the *particle-particle* (PP) method. In this method, time variable quantities are assumed constant on a small interval dt. On each such dt interval, the force of each particle on all other particles is computed. Although conceptually

straightforward, computation of the force of every particle on every other particle requires $\mathcal{O}(n^2)$ operations. This is very inefficient and consequently many alternatives have been developed. The simplest of these is the *particle-mesh* (PM) method. This method overlays space with an array of m mesh points, with $m << n$, and approximates the force or potential values with the mesh values. Poisson's equation is used to relate the force to the mass density of the mesh. Interpolation of particle properties to mesh points goes as $\mathcal{O}(n)$ and solving Poisson's equation on the mesh goes as $\mathcal{O}(m\log(m))$. Thus for some $\alpha, \beta \in \mathbb{R}$ the number of required operations goes as $\alpha n + \beta m log(m) \propto n$, as $m << n$. A huge improvement but at significant cost: for nearby particles the mesh approximation results in very limited resolution and gives large errors. For far away particles on the contrary the mesh approximation can be computed with Fourier based methods, which are near exact.

An often used approach combines the PP and the PM method to a *P^3M method*, whichs splits the total force F_{ij} in a short range component which uses the PP method and a long range components which uses the PM method. Splitting is done by overlaying an additional coarse mesh which determines which particles are closeby and which far away. Using linked lists, the overhead of storing which particles are where is minimal.

The P^3M method overcomes the inaccuracy of the PM method by splitting the simulation box in a short range and long range part. Another way to deal with the relative inaccuracy of the PM method is by making the grid finer on places where this is necessary. A finer grid is obtained by dividing a cell in daughter cells, and these daughter cells where needed again in daughter cells, and so forth. This is the basic idea behind *Tree algorithms* [Barnes and Hut, 1986], which apply the following rules:

1. Every cell is considered a (pseudo-)particle with the mass of all particles inside located at the centre-of-mass of these particles.

2. If a pseudo-cell is far away enough, it doesn't need to be resolved.

A tree is constructed as follows:

1. Begin with an empty cubical root cell that contains the system. One by one, load the particles in the root cell.

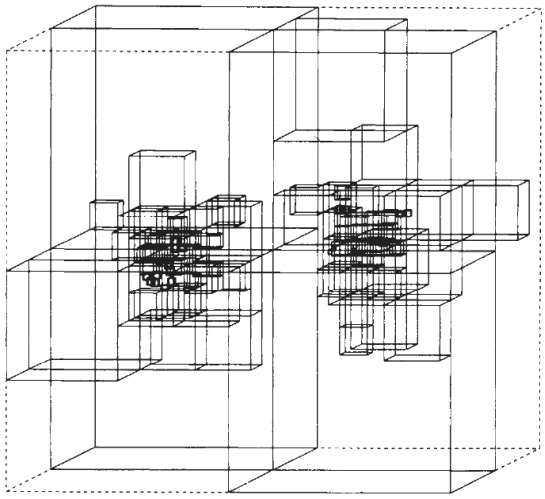
2. If any two particles fall in the same cell, divide that cell in 8 daughter cells. Continue until all particles are loaded.

3. Each cell has a mass, a centre-of-mass coordinate and a link to its daughter cells.

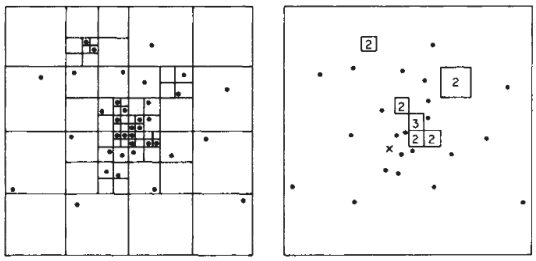
A graphical illustration of a tree is shown in figure 4.2(a). Instead of placing all particles in a different box, we can keep a few together if that grouping has only a small effect on the accuracy. The accuracy is determined by a parameter θ, which determines or cell j is divided in its daughter cells. Let l we the length of cell j and D the distance from its centre-of-mass to particle i, than the cell is resolved in its daughters if $l/D < \theta$ and used unresolved otherwise. A graphical illustration is shown in figure 4.2(b).

The efficiency of Tree algorithms is determined by two things: (i) the tree and (ii) force computations. Let h be the number of subdivisions required, i.e. the height of the tree, and B the total box size. Than we can write two equations for the average size of a particle bearing cell:

- based on the number of divisions we can write $l = (\frac{1}{2})^h B$;



(a) The Tree structure boxing of a 3D encounter of two $N = 64$ systems.



(b) The Tree structure boxing of a 2D encounter (left) and the force computation on particle x (right). The number in the boxes indicate the number of unresolved particles inside.

Figure 4.2: ***Three structures boxings [Barnes and Hut, 1986].***

- using that the average size of a particle bearing cell is of the order of the interparticle spacing we get $l = N^{-1/3}B$.

Equating both gives:

$$(\tfrac{1}{2})^h \propto N^{-1/3} \Leftrightarrow h \propto \log_2(N^{1/3})$$

This equation shows that tree construction as well as propagating mass and centre-of-mass information through a tree of size h goes as $N \log(N)$.

To find how force computations depend on N, suppose we increase the number of particles by a factor of 8. In essence this means we add 7 analogous trees to the current tree. The 7 new trees will give a modest number Δj extra terms to the force computation. But very importantly, Δj depends on θ and not on N or the size of the system. A linear increase in Δj while the number of particles increased with a factor 8 corresponds with an efficiency $\propto \log(N)$. Thus, the total efficiency of a Tree algorithm goes as $N \log(N)$.

Tree methods can be very accurate. Furthermore, they have well understood errors in terms of multipole moments of the mass distribution of the unresolved cells. Exactly that is what *multipole algorithms* [Greengard and Rokhlin, 1997] use for an even more accurate and fast force computations on special configurations. The essence of multipole algorithms is that Poisson's

equation can be rewritten as a series expansion in complex space. Consider two groups of separated particles, for example two groups of galaxies with stars. Both creating and evaluating a multipole expansion of the forces of one group on the other group are linear operations. Thus for groups of separated particles the multipole method is linear. In the most modern N-body codes, combinations of all methods described before are used to provide the optimal tradeoff between accuracy and minimal system requirements.

So far we only discussed collisionless (dark matter) particles moving under the gravitational force. Modern N-body software packages like Gadget 2 can also deal with gas particles, whose motion is determined by hydrodynamics. Traditionally gas can be dealt with in two ways, as we saw also in subsection 2.5.3. In the Eulerean method space is discretized and gas is represented using hydrodynamics on a (possibly adaptive mesh). The Lagrangian method divides the gas into a set of discrete elements, 'particles'. These particles have a smoothing length, over which their properties are smoothed. Inversely, this means the properties of a particle can be found by considering the properties of all other particles within the smoothing range. This gives us Smoothed Particle Hydrodynamics (SPH) and allows relatively similar treatment between particles and the gas (particles). In the cosmic web simulations under study here, no gas is used. Therefore, we refrain from an overview of computational hydrodynamics and refer the interested reader to [Veldman and Velicka, 2010] for an excellent introducationary treatise on the subject.

4.1.2 The Gadget 2 code

With the general ideas behind N-body simulations discussed in the previous subsection, we now turn our attention to the details of the N-body code used in this thesis: Gadget 2 [Springel, 2005][1]. Gadget 2 is a Tree-SPH code and an upgrade of the earlier Gadget 1 code. The public release of Gadget 2 contains two types of physics: (i) collisionless dynamics and (ii) hydrodynamics.

The collisionless dynamics uses a multipole extended tree code, optionally with a fast PM method approximation for large distances. The simulation space can both be a static Newtonian space as an expanding space, in the last case comoving coordinates and the peculiar potential are used (cf. equation 2.15). Three important properties in which tree codes differ are (i) the tree structure; (ii) which multipole moments are computed and (iii) the cell opening criterium used.

For grouping Gadget 2 uses the standard octonal tree structure as explained in above. Although alternatives to octonal trees exist (amongst others binary trees), octonal trees have relatively little internal cells and thus a low memory consumption. Also, octonal trees fit well with the parallelization strategy employed in Gadget 2.

Memory consumption is further limited by only evaluating monopole moments, instead of higher order multipoles. Besides the obvious advantage of reducing total memory it also makes computations relatively fast: many nodes can be kept in the processor cache and as such are quickly accessible.

Writing M and l for the mass respectively size of a cell a distance r away from a particle whose force is computed, Gadget 2 uses the following opening angle θ:

$$\frac{GM}{r^2}\frac{l}{r} \leq \theta |a|$$

with a the size of the total acceleration obtained in the previous time step. Physically, this criterium compares an estimate of the truncation error (left side of the equation) with the total

[1] Gadget 2 can be downloaded from `http://www.mpa-garching.mpg.de/gadget/`.

expected acceleration (right side of the equation). In this way, a limit is set to the absolute force error introduced in each particle-node interaction. For the first time step, the ordinary opening criterium is used. For very closeby nodes, the opening criteria above can give very large or even unbounded errors. To prevent this from happening, an extra opening criterium of

$$|r_j - c_j| \leq 0.6l$$

is implemented, with $j \in \{1, 2, 3\}$, r_j the particle coordinates and c_j the geometric centre of the node.

Walking deeper in a tree requires more computational resources. To prevent deep walks, space can be splitted in a closeby region, for which the tree is walked, and a far away region that can be approximated by the PM-method. Not only does this decrease the total computational cost, it also increases the accuracy of very long-range force contributions as is shown in figure 4.3.

A SPH approach to hydrodynamics is taken, with two technical details. First, to assure entropy conservation the internal state of each particle is defined in terms of the entropy per unit mass instead of the 'default' thermal energy. Second, as usual with the Lagrangian approach an artificial viscosity needs to be introduced to inject entropoy in shock fronts. Gadget 2 uses the viscosity derived by [Monaghan, 1997], combined with an additional viscosity-limiter. The last prevents unphysical viscosity induced excess acceleration and makes time integration more stable.

4.1.3 This Gadget implementation

The astrophysical input of the Gadget 2 N-body simulation consists of two elements: (i) the initial conditions and (ii) the DE models used. For the initial conditions, a primordial Gaussian random field (cf. subsection 2.5.1) is used, evolved to $z = 60$ using the Zel'dovich approximation (cf. subsection 2.5.3). The three dark energy models used are LCDM, RP and SUGRA (cf. section 2.4 for details). An overview of the DE model specific parameters is given in table 4.1. The DE models are normalized at the cosmic microwave background using:

$$\sigma_{8,\text{DE}} = \sigma_8 \frac{D_{\Lambda\text{CDM}}(z_{\text{CMB}})}{D_{\text{DE}}(z_{\text{CMB}})} \tag{4.1}$$

with D the linear growth factor (cf. subsection 2.5.2) and $z_{\text{CMB}} = 1089$.

The simulations involve 256^3 dark matter particles (no other particles) with identical masses of $0.443 \cdot 10^{10} h^{-1} M_\odot$, living in a cubic box with axes of $300h^{-1}$ Mpc and periodic boundary conditions. The general cosmological parameters are from the WMAP 3 data: $\Omega_m = 0.268$, $\Omega_\Lambda = 0.732$, $\Omega_b = 0.044$, $h = 0.74$, $\sigma_8 = 0.776$ and $n = 0.947$. Snapshots are available at 8 redshifts: $z = 3.80$, $z = 2.98$, $z = 2.05$, $z = 1.00$, $z = 0.51$, $z = 0.25$, $z = 0.10$ and $z = 0.0$. For more details on the simulations, we refer to [Bos et al., 2012; E.G.P., 2010]. For each dark energy model five different realizations are used, namend run 14 - 18. The different runs of the same DE model differ in random seed number used for the primordial Gaussian field.

4.2 DTFE: from particles to densities

In many situations the density field underlying a particle distribution is needed. Based on the exposition of [Schaap, 2007], subsection 4.2.1 explains how to obtain it. The ***Delaunay Tesselation Field Estimator (DTFE)*** method [Schaap and Weygaert, 2000; Schaap, 2007] turns out

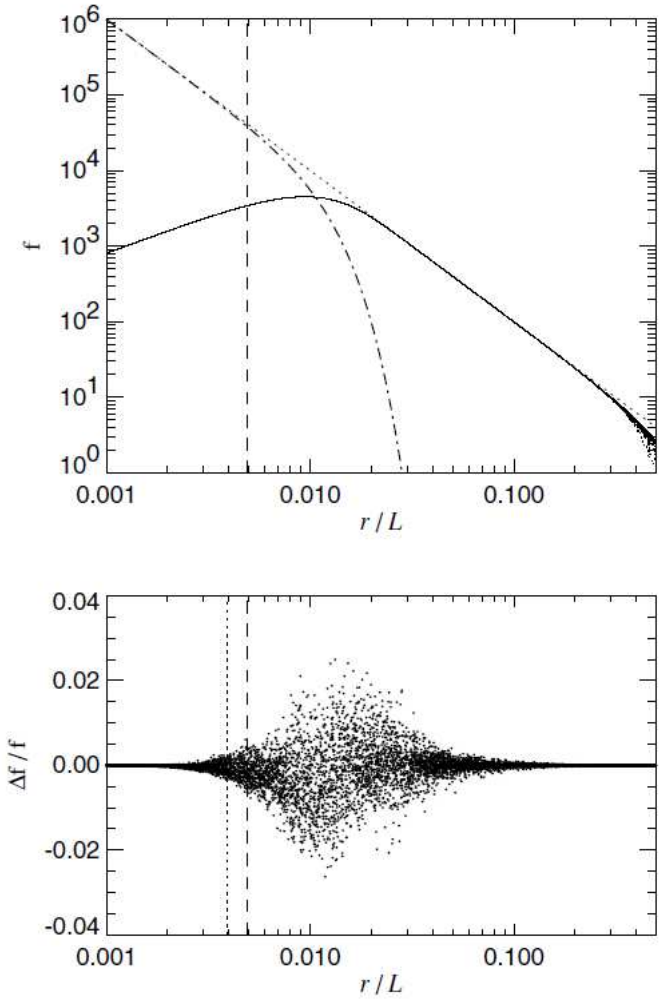


Figure 4.3: ***Force f decomposition (top) and relative force error $\Delta f / f$ (bottom) of the Gadget 2 Tree-PM scheme [Springel, 2005].*** *The plots are for several test particles as function of distance r from a unit mass particle in a box of size L. In the top panel, the short range Tree force is shown as dotted-dashed line and the long range PM force as solid line. As one can see, for spatial split scales r_s (vertical dashed line) of the same order and slightly larger than the mesh size (vertical dotted line), the relative error is only in the order of percents. The relative error can be descreased by increasing r_s.*

Model	α	w_0	w_a	σ_8
ΛCDM	-	-1.0	0.0	0.776
RP	0.347	-0.9	0.0564	0.746
SUGRA	2.259	-0.9	0.452	0.686

Table 4.1: ***DE models parameters used in the Gadget N-body simulations*** *σ_8 is fixed at $z = 0$ and is the same for all models at $z_{CMB} = 1089$. w_0 is also fixed at $z = 0$ and w_a is determined using a χ^2 fit. Cf. section 2.4 for a conceptual overview of the DE models.*

to be a natural choice for a density field reconstruction. The algorithmic implementation of the DTFE method is dicussed in subsection 4.2.2. Subsection 4.2.3 elaborates on the specific DTFE setting used within this thesis. The DTFE code used here returns the interpolated densities at a regular grid. Compared with interpolation to the original particle positions, this gives better

sampling of low density regions.

Within the data pipeline: the input of this step is a data cube containing the positions of n particles. The output is a data cube with the interpolated density field values on a regular grid.

4.2.1 From particles to densities

Consider a continuous field $f(x)$, observed at positions x_i with values $\tilde{f}_i$, $i\{1, ..., N\}$. An often used method to reconstruct $f(x)$ based on the observed $\{\tilde{f}_i\}$ is to smooth the observed values out. Concretely, this is done by filtering the data with a filter fuction $W(x)$, giving the approximated field:

$$\hat{f}(x) = \frac{\sum_{i=1}^{N} \tilde{f}_i W(x - x_i)}{\sum_{i=1}^{N} W(x - x_i)} \tag{4.2}$$

Often chosen filters are circular or spherical Gaussians. Although conceptually simple, both from a practical as well as from a theoretical point of view the method is plagued with problems. Practically, filtering removes features smaller than the filtering function used. Consequently, the geometry of anisotropic features with length scales smaller than the filter might change in some directions. Theoretically, applying a filter means implicitely using a mass-weighted reconstruction, which is best appreciated by simply writing out formula 4.2:

$$\begin{aligned} \hat{f}(x) &= \frac{\sum_{i=1}^{N} \tilde{f}_i W(x - x_i)}{\sum_{i=1}^{N} W(x - x_i)} \\ &= \frac{\int dy f(y) W(x - y) \sum_{i=1}^{N} \delta_D(y - x_i)}{\int dy W(x - y) \sum_{i=1}^{N} \delta_D(y - x_i)} \\ &= \frac{dy f(y) W(x - y) \rho(y)}{dy W(x - y) \rho(y)} \end{aligned} \tag{4.3}$$

with $\rho(y)$ the volume density of the sampling points. I.e. the reconstructed field $\hat{f}(x)$ is the real field $f(x)$ filtered by $W(x)$ and weighted by the mass of the sampling points. It can be seen as a mass-weighted estimate of $f(x)$. Unfortunately analytical calculations mostly involve a volume-weighted estimate because stochastic integrals are volume-weighted:

$$f_{\text{volume}}(x) = \frac{\int dy f(y) W(x - y)}{\int dy W(x - y)} \tag{4.4}$$

The difference between equations 4.3 and 4.4 is due to fact we only have knowledge of the field at the sampling points. In essence, using a filter to reconstruct the density field confuses a mass-weighted with a volume-weighted estimate, disqualifying this approach.

Volume-weighted reconstruction is possible using interpolation. The interpolated field value at any arbitrary point can be computed using a linear combination of field values at the N sampling points. The linear coefficients have the two constraints: (i) they should sum up to one and (ii) at the location of the sampling points the interpolated values should equal the measured values.

A huge number of interpolation schemes are available. Conceptually the simplest is zeroth-order interpolation: divide the space in regions closest to each sampling point, i.e. a Voronoi diagram (cf. definition 3.16), and assume a constant density within each cell. Of course this gives discontinuities at the boundaries, which can be avoided by using a first-order interpolation scheme. An illustration of zero and first-order interpolation is shown in figure 4.4. In one

dimension first-order interpolation is straightforward linear interpolation. In more dimensions this approach can be generalized by covering the space with a triangulation which has the sampling points as its vertices. Inside the hyper-triangles one can interpolate between the values at the vertices: take an arbitrary point x inside the d-dimensional hyper-triangle j with vertices $x_0, ..., x_d$. Than the function value $f(x)$ can be approximated with:

$$\bar{f}(x) = \hat{f}(x_0) + \nabla \hat{f}|_j \cdot (x - x_0) \tag{4.5}$$

To solve this equation we still need the the gradient function $\nabla \hat{f}$. It can be found by solving equation 4.5 at the d vertices, where all parts except the gradient function are known.

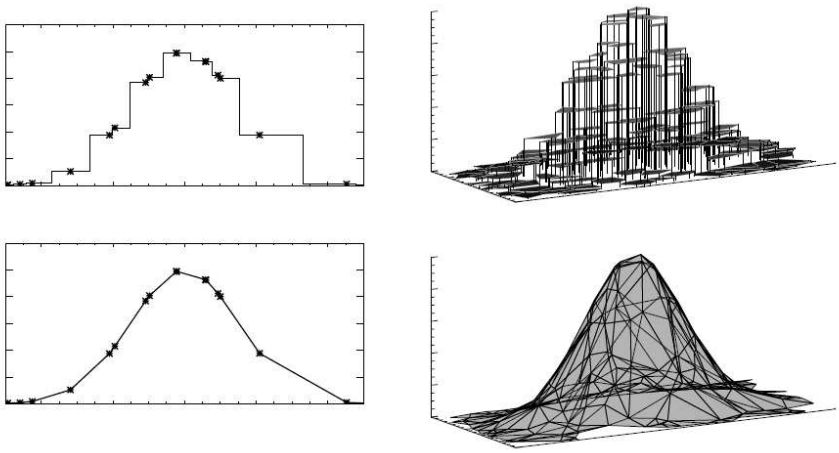


Figure 4.4: ***Visual illustration of interpolation schemes [see Schaap, 2007, chap. 4].*** *Overview of zeroth-order interpolation (top row) and first-order interpolation (bottom row) in 1D (left) and 2D (right).*

The choice of triangulation determines the accuracy of the triangulation. [Bernardeau and Weygaert, 1996], [Schaap and Weygaert, 2000] and [Pelupessy et al., 2003] have shown that a Delaunay triangulation (cf. definition 3.17) is the preferred triangulation, because:

- It is fully adaptive, i.e. it automatically probes regions at maximum resolution.
- Delaunay triangles have minimal size and elongation, preferable properties for a relatively local field.
- Linear interpolation requires the definition of neighbor intervals. For the Delaunay triangulation such a definition rols naturally out of its dual, the Voronoi diagram.

For fields of which the measurements directly provide field values, for example when measuring velocities or temperatures, we are done: interpolation with a Delaunay triangulation provides an optimal approximation of the underlying field. Things are more complicated when trying to recover the density field (kg m^{-3}) based on particle masses (kg). Then, using the particle masses and the particle positions, density values have to be estimated. An intuitive choice for the density estimate of particle i might be $\rho(x_i) = \frac{m}{V_{\text{Vor},i}}$ with V_{Vor} the volume of the Voronoi cell corresponding to the particle. Unfortunately, it turns out such a definition doesn't conserve mass [Schaap and Weygaert, 2000] and the *continguous Voronoi cell* should be used instead. The continguous Voronoi cell is defined as the agglomerate of all K Delaunay tetrahedra containing

point i as one of its vertices. It has volume $W_{\mathrm{Vor},i} = \sum_{j=1}^{K} V_{\mathrm{Vor},i}$ and is used $d+1$ times in a d-dimensional field, giving the density estimate:

$$\rho(x_i) = m(d+1)/W_{\mathrm{Vor},i} \tag{4.6}$$

Figure 4.5 illustrates the difference between the volumes given by the Voronoi cell and by the continguous Voronoi cell. Using the point sample as density estimator puts constraints on the sampling process. The procedure is only valid if the points constitute a fair sample of the underlying distribution, i.e. if the points form a Poisson point process of the density field [Weygaert and Schaap, 2009]. Therefore, DTFE density estimates are quite sensitive to Poisson noise.

Summarizing the above we obtain the DTFE method, which estimates the underlying (density) field of a set of discrete observations (of particle positions). Concretely, DTFE consists of the following steps:

1. Construct the Delaunay triangulation corresponding to the particle distribution.
2. Use the Delaunay triangulation to estimate the density values at the particle positions.
3. With the previous two steps as input for linear interpolation, a representation of the underlying field can be computed.
4. The field can be outputted to a regular grid.

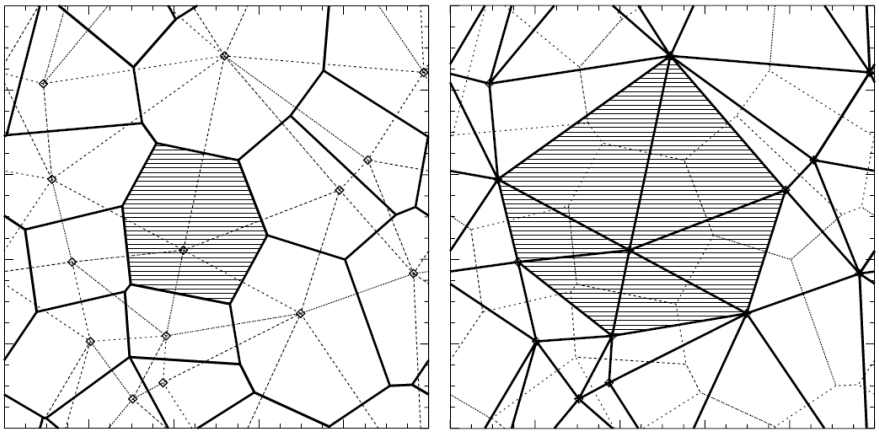


Figure 4.5: ***The Voronoi cell (left) and the continguous Voronoi cell (right) of a point [Schaap and Weygaert, 2000].***

4.2.2 The DTFE algorithm

An excellent and efficient algorithm to obtain the DTFE field is the publicly available code of [Cautun and Weygaert, 2011][2]. The code is written in C++ using the CGAL (Computational Geometry Algorithms Library) library for geometrical algorithms, OpenMP for parallelization, the Boost C++ libraries and the GNU Scientific Library. The code allows periodic boundary conditions and more detailed interpolations in interesting subregions. Input boxes can be square, cubic, redshift cone and user defined. Many scalar and gradient scalar fields are available. The output is a regular grid file with DTFE obtained field estimates at each grid point.

[2] Cautun et al's DTFE code can be downloaded from `http://www.astro.rug.nl/~voronoi/DTFE/dtfe.html`

Interpolation to a regular grid can be done in two ways: (i) each grid cell gets the interpolated density value of the grid-centre and (ii) a volume average is used. The second method is much less affected by Poisson noise, especially in the high density regions. But volume averaging is very computationally demanding. Therefore, volume averaging is implemented by random sampling, using either Monte Carlo integration inside the Delaunay cells or Monte Carlo integration inside the grid cells.

4.2.3 This DTFE implementation

For this thesis, we used DTFE version 1.2 with periodic boundary conditions and interpolation option (ii). A regular output grid with as many grid cells as particles was chosen.

4.3 DMC and filtration builders

Gadget gave us a particle distribution, which was possibly transformed in a regularly sampled density field using DTFE. The first step in the topological part of the data pipeline is the computation of the DMC and subsequently its filtration. For these steps the density and particle scenarios use different algorithms. We will discuss both.

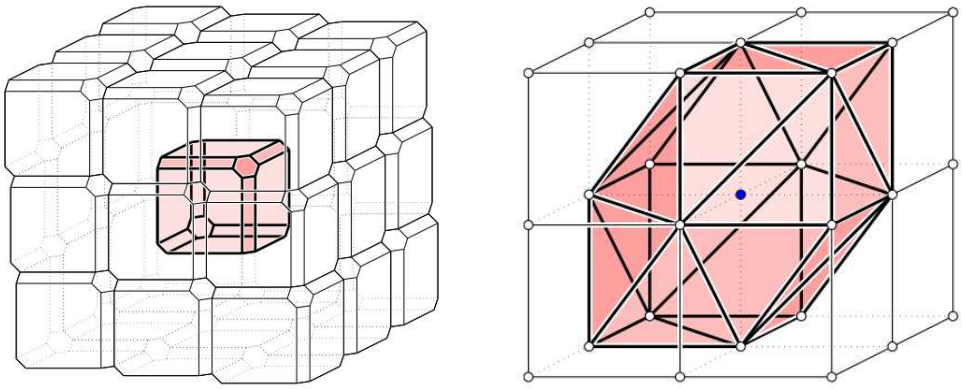
4.3.1 Density DMC's and filtrations

The output of Cautun's DTFE program is a regularly sampled grid, where each voxel has a density value associated to it. To build the DMC of a regular grid and compute its filtration, an algorithm based on the ideas of [Bendich et al., 2010] and concretely implemented by [Pranav et al., 2013a] is used. The algorithm consists of three steps:

1. By iterated division of the datacube in 8 smaller datacubes a tree is created, much akin the procedure described in section 4.1.1 above. Although a stopping criterion can be set, here subdivision continues until individual voxel level is reached. For the $256^3 = (2^8)^3$ datacubes considered here, this requires 8 subdivisions.

2. Each voxel of the tree can be seen as the set of points 'closest' to voxel center, much alike Voronoi cells (cf. definition 3.16). Considering the set of voxel centers as vertices, its dual Delaunay complex can be created. Note that doing this directly on the regular grid runs into problems because each vertex has six nearest neighbors. This makes it impossible to unambigously define the correct trianguluation, much alike the situation sketched in figure 3.12. Therefore, all voxels are slightly perturbed in the direction of the main diagonal, transforming the cubic cells in simple polyhedra. Concretely, the voxel grid coordinates (i, j, k) are shifted in the following way: $(i, j, k) \longmapsto (i - \epsilon m, j - \epsilon m, k - \epsilon m)$, with $\epsilon > 0$ but very small and $m = i + j + k$. The effect of such a transformation is shown in the left panel of figure 4.6. The resulting triangulation looks the same at all voxels and is illustrated in the right panel of figure 4.6. By assigning function values to edges, triangles and tetrahedra as described in section 3.5, a DMC is obtained.

3. The last step is ordering all the simplices by function value. In terms of computing power this is the most expensive step of the three. It can be simplified by first ordering only the vertices. Than, for each vertex p the simplices who have p as its highest valued vertex can be added. In the filtration, they come directly after p.



(a) Perturbed cubes of a regular grid. *(b) The closed star (cf. properties 3.3) of a vertex.*

Figure 4.6: ***Illustration of a perturbed regular grid complex [Bendich et al., 2010].***

4.3.2 Particle DMC's and filtrations

The particle case is analogous to the density case [Pranav et al., 2013b] but simpler. At the position of each particle, the density is estimated by its continguous Voronoi cell (cf. figure 4.5). The Delaunay triangulation of the particle positions gives rises to a simplicial complex. Extension of the density function over the higher dimensional simplices can be done as described in section 3.5 and results in a DMC. Subsequently, all simplices can be ordered by function value.

The Gadget particle simulations had periodic boundary conditions. Therefore, the DMC and the resulting filtration is computed assuming periodic boundary conditions as well. Computationally, this is achieved by glueing a copy of the data cube to all its 6 faces, 12 edges and 8 vertices and performing computations on the triangulation on the 27 data cubes together. Later on, the triangulation is updated by identifying simplices from the opposite faces, after which the 26 copies are discarded.

We note that computation of the DMC and its filtration is the bottleneck of the data pipeline is terms of hard disc memory. Although the exact size of a DMC/filtration file depends on the data cube under consideration, for the 256^3 element data cubes considered here a single file can exceed the 10 GB when stored as ASCII data, the default program output. For 8×5 runs $\times$ 3 DE models this requires about 1.2 TB of data storage. In one of the earlier stages of this thesis it was attempted to compute the DMC and filtration of uniform random noise distribution. This attempted failed when the hard disc used got out of memory after 1.8 TB of data was written to it in a few hours. In the end, using excessive amounts of hard disc memory was circumvented by feeding the filtration files directly to PHAT without storing them.

4.4 PHAT: computing persistence pairs

In the previous section we obtained a filtration, here we use the filtration to compute the cycles and Betti numbers. First, we give a general overview of computational persistent homology: what is the general idea and which options are there? Subsequently we focus on the software package PHAT (Persistent Homology Algorithm Toolbox) used here and describe it in detail.

We end this section with a short note one some issues of the implementation of PHAT used within this thesis.

4.4.1 Persistent homology computation

Edelsbrunner [Edelsbrunner and Harer, 2010] more or less defined the field of computational persistent homology with the introduction of the first computational homology algorithm. Their algorithm is based on matrix reduction and very illustrative. It is the first algorithm we describe. Unfortunately, its explanatory power goes at the cost of its efficiency, as for n simplices it scales as $\mathcal{O}(n^2)$. Consequently, several refinements or alternative approaches are developed, we describe them afterwards.

The 'standard' algorithm [Edelsbrunner and Harer, 2010]

Let K be a simplicial complex with n simplices and suppose we have its filtration. We denote the simplex on the i-th index in this filtration with σ_i. Than we define an $n \times n$ ***boundary matrix*** δ which stores for all simplices its facets and cofacets in a very structured way:

$$\delta[i,j] = \begin{cases} 1 \text{ if } \sigma_i \text{ is a facet of } \sigma_j \\ 0 \text{ otherwise} \end{cases} \tag{4.7}$$

Formulated in words: each colum indicates the facets of σ_j and each row indicates the cofacets of σ_i. A simple simplicial complex is shown in ifgure 4.7. Its filtration is $\sigma_1 < \sigma_2 < ... < \sigma_9$, giving the boundary matrix shown for clarity in table 4.2. Note that the boundary matrix is an upper triangular matrix, as by construction the faces of a simplex are at lower indices than the simplex itself. We define $low(j)$ as the row index of the lowest 1 in column j. If the

Simplices	σ_1	σ_2	σ_3	σ_4	σ_5	σ_6	σ_7	σ_8	σ_9
σ_1			1		1	0	0		
σ_2			1		0	1	0		1
σ_3					0	0	1		
σ_4					1	1	0		
σ_5							1		
σ_6							1		
σ_7									
σ_8									1
σ_9									

Table 4.2: ***Boundary matrix for the simplicial complex shown in figure 4.7.*** *Empty matrix entries indicate zeros.*

entire colum consists of zeros $low(j)$ is undefined. In our example, we directly see that $low(1)$, $low(2)$, $low(4)$ and $low(8)$ are undefined and $low(3) = 2$. The only operation allowed on δ is subtraction of columns from left to right. Starting at the leftmost column j_1, we apply this operation as often as necessary to make sure that $low(j_1) \neq low(j)$ for all $j \neq j_1$. Previous is repeated for the second column j_2, the third column j_3 up till the last column j_n, after which the matrix δ is transformed to the *reduced matrix* R. In our example matrix reduction is very easy: only column 6 can be reduced by subtracting columns 3 and 5. The reduced boundary matrix is

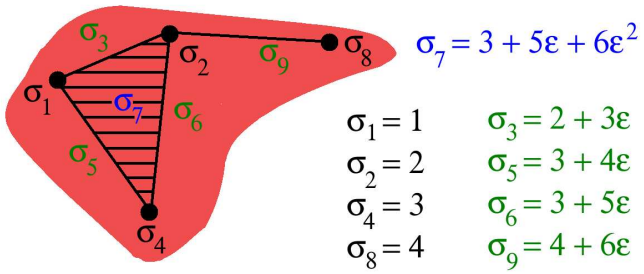


Figure 4.7: ***A simplicial complex to illustrate the concept of boundary matrix.*** *0-simplices are indicated in black, 1-simplices in green and the 2-simplex in blue. The boundary matrix is shown in table 4.2 and its reduced variant in table 4.3.*

Simplices	σ_1	σ_2	σ_3	σ_4	σ_5	σ_6	σ_7	σ_8	σ_9
σ_1			1		1	0	0		
σ_2			1		0	0	0		1
σ_3					0	0	1		
σ_4					1	0	0		
σ_5							1		
σ_6							1		
σ_7									
σ_8									1
σ_9									

Table 4.3: ***Reduced boundary matrix for the simplicial complex shown in figure 4.7.*** *Empty matrix entries indicate zeros.*

shown in table 4.3. Algorithmically, previous looks like:

Algorithm 1 Basic persistent homology computation algorithm.

```
R = δ
for c = 1 : n do
    for i = c + 1 : n do
        if low(j_i) = low(j_c) then
            j_i += j_c
        end if
    end for
end for
```

Alternatively, if we let matrix C denote the column additions required to reduce δ to R, we can write previous in matrix notation as $R = \delta C$. Reduced matrix R has a straightforward interpetation:

1. If column j of R is zero, simplex σ_j has no boundary and thus represents a new cycle which enters the filtration at index j. Indeed, in our example columns 1, 2, 4 and 8 correspond to vertices and thus new 0-cycles and their columns are zero from the start. By reduction column 6 became zero as well. A short look at its simplicial complex in figure 4.7 shows that when column 6 enters the filtration, a tunnel is created i.e. a new one-cycle

is born.

2. If column j of R is non-zero simplex σ_j has facets and thus connects two components or fills a hole. Either way, a cycle dies. Let row $i = low(j)$ indicate its highest valued facet, than all other faces of σ_j entered the filtration at an earlier stage than σ_i. Then, pairing via the Elder Rule (cf. definition 3.30) implies that simplices σ_i and σ_j are paired. Thus for non-zero column j the cycle born at $i = low(j)$ dies. In our example simplex corresponding to the non-zero column σ_3 unites components σ_1 and σ_2 and is paired with the latter. Analogously, σ_5 unites $\{\sigma_1, \sigma_2\sigma_3\}$ with σ_4 and is paired with the latter, etc.

Persistent homology (cf. appendix A.2) follows directly: the number of zero columns that correspond to k-simplices gives the rank of Z_k and the number of non-zero columns that correspond to k-simplices gives the rank of B_k. Using equation A.5 their difference gives β_k. The birth, death and persistence of cycle-classes follows directly from the $(low(j), j)$ pairs.

Although conceptually very simple, the double for-loop shows the algorithm is $\mathcal{O}(n^2)$, which is not very efficient. Furthermore, the memory requirements also scale as $\mathcal{O}(n^2)$. Both become quickly prohibitive for large datasets. Some direct efficiency upgrades were provided by [Edelsbrunner and Harer, 2010] themselves. For large datasets δ contains many zeros. Thus a sparse matrix representation in the form of an array of simplices with each element containing linked lists to its facets saves both space and time. In some situations using specific knowledge about the simplicial complex allows further simplifications. Nonetheless, for large general simplices the effect of these upgrades is only marginal. For large datasets more efficient algorithms are a necessity. So far, three such algorithms have been developed.

The 'twist' algorithm [Chen and Kerber, 2011]

The twist algorithm improves upon the standard algorithm by a key observation: if element $i = low(j)$ appears as pivot in a reduced column, simplex j is paired with simplex $i < j$. Thus, simplex i creates a cycle and its column is zero. For example, in the simplicial complex of figure 4.7 we have amongst others $low(3) = 2$ and $low(7) = 6$ and indeed, 2 & 3 and 6 & 7 for persistence pairs. Wee see in table 4.3 that columns 2 and 6 are zero.

In the standard algorithm this doesn't save any operations because column $i < j$ and thus column i was already zero. Yet by changing the order of reduction such that higher dimensions are reduced first (still from left to right) many operations can be saved. In large datasets, for up to half of the columns reduction is avoided by this technique. However, in principle this algorithm still scales as $\mathcal{O}(n^2)$.

The 'row' algorithm [De Silva et al., 2011]

The basic operations above were column operations but in principle using row operations are possible as well, resulting in the same output [De Silva et al., 2011]. Using row operations is in particular efficient if *cohomology* instead of homology is used. Cohomology is the dual of homology and uses the maps on simplices instead of the simplices itself. For example, if c is a chain of simplices and $\varphi(c)$ a function that maps c to 0 or 1 (for c is not present respectively present in a simplicial complex), than φ is a basic element of cohomology just as groups of simplices are the basic elements of homology. Cohomology is homeomorphic (cf. definition A.11) to homology, i.e. in the end they give rise to equivalent structure. For some datasets the row algorithm is remarkably faster than the standard column variant. Why this is and whether it can be predicted a priori is an area of active research [Chen and Kerber, 2011].

The 'chunck' algorithm [Bauer et al., 2013]

The chunck algorithm is basically a parallelization of the twist algorithm. It consists of three steps:

1. The matrix is divided in local blocks and these blocks are reduced.
2. The already reduced rows and columns are taken out.
3. The remaining (relatively small) matrix with global persistence pairs is reduced to its final form.

The chunck algorithm incorporates the key idea of the twist algorithm. Furthermore, it implements another efficiency trick based on the following observation: if simplex j was paired with simplex $i < j$, simplex j cannot be paired again thus the rest of the row can be set to zero directly. An important choice within the chunck algorithm is which chunck size to use. Generically, for a dataset with n simplices the authors choose a chunck size of $\sqrt{n}$, giving $\sqrt{n}$ separate chuncks.

On multi-core computers, the chunck algorithm is orders of magnitude faster than the other methods above.

4.4.2 Persistent homology with PHAT

All the methods above have been implemented in the publicly available library PHAT[3]. The code is written in C++, requires a boundary matrix as input and gives the persistence pairs as output. Which of the algorithms discussed above is used is up to the user. In case of the chunck algorithm, the OpenMP library is used for parallelization.

4.4.3 This PHAT implementation

The dramatic increase in processing speed of the chunck algorithm compared with the standard algorithm has been noted while processing the astronomical datacubes considered in this thesis. For a 256^3 particles or density gridpoints dataset, the standard algorithm required more than a month. (The exact time is unknown, as power failures, disc errors and network failures never resulted in complete computation of a 256^3 data cube with the standard algorithm.) The chunck algorithm used later computed the persistence pairs of the same datasets in less than three hours.

We note that the persistent homology computation is the bottleneck of the data pipeline in terms of computational power and RAM. Although the chunck algorithm required less than three hours, this was on a 160 hyperthreading CPU machine. The RAM requirements can exceed 150 GB.

4.5 Data analysis

The persistence diagrams obtained from PHAT are analysed using code written in C++ and compiled with Unix Red Hat gcc version 4.4.7. Some code snippets from [Press et al., 2007] were incorporated. On occasions, Matlab 2013 was used in addition. All software, algorithms and code were run on Unix machines of the Kapteyn Astronomical Instite with Scientific Linux. Data cube visualizations are generated with VisIt[4].

[3] PHAT can be downloaded via `https://code.google.com/p/phat/` .

[4] VisIt can be downloaded via `https://wci.llnl.gov/codes/visit/` .

CHAPTER 5

Stability of persistence diagrams

Experimental data contains measurement errors and simulations might also contain uncertainties. For example, the Monte Carlo integration used at the DTFE implementation here introduces small uncertainties. Homology and persistence are only useful if they are not influenced by these kind of uncertainties too much. This chapter will look at the way persistence diagrams are influenced by small uncertainties in the input. Subsequently, various methods to get uncertainties out will be compared. First, section 5.1 expounds the mathematical results obtained in this direction. As this is a new field of mathematics the mathematical results are only quite general: small perturbations result only in small perturbations of each point of the persistence diagram. However, a persistence diagram consists of many points. How small perturbations of each point will influence the diagram as a whole is unknown. Therefore, a perturbation analysis is carried out: varying amounts of particles are perturbed by various distances and the obtained persistence diagrams are compared with the unperturbed case. The details of the analysis and the resulting persistence diagrams are discussed in section 5.2.

Persistence diagrams have quite some low persistent features, either noise or insignificant ripples in the large density waves of the field. Several ways to take out these features thus nicing[1] the persistence diagram are explored in section 5.3. The results of previous sections are combined in section 5.4 and a discussion and conclusions are presented in section 5.5.

5.1 Mathematical results

For persistence diagrams to be useful as analytical description of observational (and even simulated) data they have to be ***stable***: small changes in input data should result in only small changes in the persistence diagram. The concept of stability is visually illustrated in figure 5.1. [Cohen-Steiner et al., 2007] showed stability for each point in persistence diagrams for a broad class of tame functions (defined below). for each point in the persistence diagram the distance the point is moved by a perturbation is bounded by a supremum-norm. In this section we work towards their result following [Edelsbrunner and Harer, 2010, chap. VIII]. The material in the sections uses quite some topology. For users less acquainted with topology we advise to read the chapter 3 on topology first.

[1] Within this thesis, the term *nicing* will be used as an aggregate term for various ways of smoothing or simplifying.

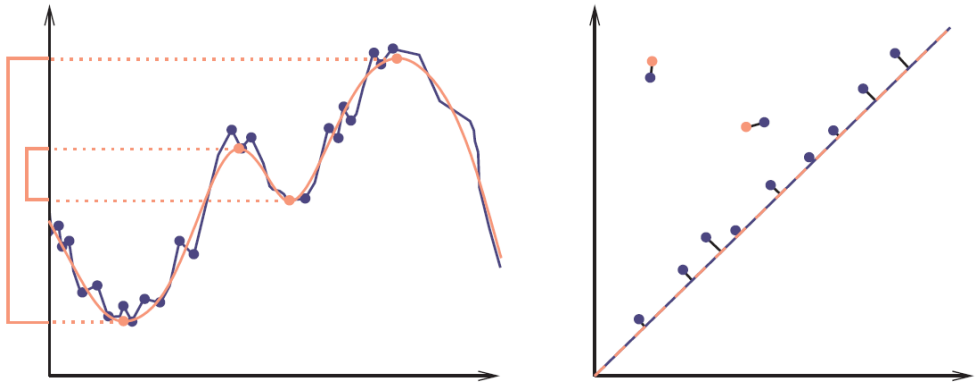


Figure 5.1: ***Visual illustration of stability of persistence diagrams [Cohen-Steiner et al., 2007].*** *A small change in density (due to a small change in particle positions, left) results in a small change in the corresponding persistence diagram (right).*

To explore whether persistence diagrams are stable, we need a way to measure distances between two diagrams. Let X and Y be two persistence diagrams with two points $x = (x_1, x_2) \in X$ and $y = (y_1, y_2) \in Y$. In topology, a well known distance measure between two point sets is the:

Definition 5.1 (Hausdorff distance d_H)**.**

$$d_H(x, y) = \max\{\sup_{x \in X} \inf_{y \in Y} \|x - y\|_\infty, \sup_{y \in Y} \inf_{x \in X} \|y - x\|_\infty\} \tag{5.1}$$

Informally, the Hausdorff distance is the largest distance from a point in X to the closest point in Y. So, two persistence diagrams are close in Hausdorff space if for all points in X there is a point in Y closeby, and vice versa.

with $\|x - y\| = \max\{x_1 - y_1, x_2 - y_2\}$ the supremum-norm. Adding the restriction that we only look at bijections between the persistence diagrams $\gamma(x) : X \to Y$, the Hausdorff distance can be refined to the:

Definition 5.2 (Bottleneck distance)**.**

$$d_B(X, Y) = \inf_{\gamma} \sup_{x \in X} \|x - \gamma(x)\|_\infty \tag{5.2}$$

Essentially this is the Hausdorff distance with the additional constraint that $y = \gamma(x) \in Y$ for some $x \in X$.

Example 5.1. *Figure 5.2 shows a superposition of two persistence diagrams, one consisting of white and the other of black points. The bottleneck distance is half the side length of the squares illustrating the bijection.*

We remark that by definition of the supremum-norm the bottleneck distance satisfies:

1. $d_B(X, Y) \geq 0$
2. $d_B(X, Y) = 0 \Leftrightarrow X = Y$
3. $d_B(X, Y) = d_B(Y, X)$
4. for another persistence diagram Z we can write $d_B(X, Y) \leq d_B(X, Z) + d_B(Z, Y)$

Consequently, the Bottleneck distance defines a metric (cf. definition A.1) and thus really is a distance.

Suppose that our persistence diagram comes from functions f, g defined over a simplicial complex K. We let $f, g : K \to \mathbb{R}$ be monotonicly increasing along the filtration of K (cf. definition 3.28). This allows us to define the straight-line homotopy (cf. definition A.10) $F : K \times [0, 1] \to \mathbb{R}$ by:

$$F(\sigma, t) = (1 - t)f(\sigma) + tg(\sigma) \tag{5.3}$$

which interpolates between f and g. The homotopy changes function values and thus changes the values of critical points. Consequently, it changes the persistence of persistence pairs. With increasing time, a persistence pair $[\sigma_k, \sigma_{k+1}]$ will start to wander around in the persistence diagram because its birth and death value change. When the function values change enough, persistence pairs might even be paired differently. Alternatively critical simplices might become non-critical, after which their persistence pairs disappear. The last happens easiest with points who are quickly smoothed out, i.e. which have low persistence and are close to the diagonal.

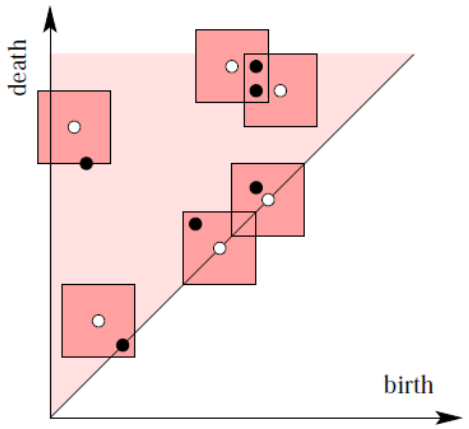


Figure 5.2: ***Visual illustration of the bottleneck distance [Edelsbrunner and Harer, 2010].*** *A superposition of two persistence diagrams, one with black and one with white pionts. The bottleneck distance is half the side length of the squares illustrating the bijection. The maximum distance is given by the black and white point top-left.*

Every time t_i that the pairing is reordered or persistence pairs disappear, a new filtration order can be established for $F(\sigma, t)$. Furthermore, for a finite simplicial complex K there is only a finite number of times $0 = t_0 \leq t_1 \leq \ldots \leq t_n \leq t_{n+1}$ on which the filtration is reorderd. Within each interval (t_i, t_{i+1}) the pairing is constant. Stacking persistence diagrams for each new ordering on top of each other, we get a three dimensional persistence diagram with the third axis corresponding to time. An illustration of such a persistence diagram is shown in figure 5.3. For a given persistence pair, we can drawn lines between their values at t_i and t_{i+1}. In case the persistence pair changes partners with another persistence pair, a line ending at t_{i+1} will continue in a different direction but it will continue. In case the persistence pair ceases to exist, it is connected with the diagonal. Diagonal points are not continued.

Thus, each persistence pair gives rise to a path monotonically increasing in t. The path is either born on X or on some interpolated persistence diagram F_{t_i} in between X and Y. Paths either survive to Y or die on the diagonal at an earlier interpolated persistence diagram F_{t_j}. Each path element is called a ***vine*** and a multiset of path elements a ***vineyward***.

The fact that the paths are connected from birth to death is very important, as it allows us to see how far points wander on the diagram. To so exactly so, we quantify the path of a persistence pair by:

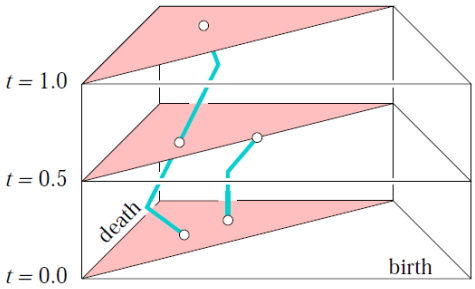


Figure 5.3: ***A vineyard [Edelsbrunner and Harer, 2010].*** *Two paths of the straight-line homotopy between the monotonic function f and g. One vine lines the entire time, while another vine dies at $t = 0.5$.*

$$p(t) = (1-t)\,(f(\sigma_k), f(\sigma_{k+1}), 0) + t\,(g(\sigma_k), g(\sigma_{k+1}), 0) \tag{5.4}$$

Its change in time is given by:

$$\partial p/\partial t(t) = (g(\sigma_k) - f(\sigma_k), g(\sigma_{k+1}) - f(\sigma_{k+1}), 1) \tag{5.5}$$

If we project the endpoints of each part of the path back on $\mathbb{R}^2$, we can measure how far the pair has wandered off:

$$\|p(t_{i+1}) - p(t_i)\| = (t_{i+1} - t_i) \max\{g(\sigma_k) - f(\sigma_k), g(\sigma_{k+1}) - f(\sigma_{k+1})\} \tag{5.6}$$

We use this to measure the distance between the functions f and g using the L_∞ norm:

$$\|f - g\|_\infty = \max_\sigma |f(\sigma) - g(\sigma)| \tag{5.7}$$

Clearly this is also an upper bound on the slope of any line segment in the vineyard and therefore an upper bound on the L_∞ norm between the end-points of any vine. This gives us the:

Theorem 5.1 (Stability theorem for filtrations)**.** *Let K be a simplificial complex and $f, g : K \to \mathbb{R}$ two monotonic functions along the filtration of K. Let X and Y be two persistence diagrams. For each dimension d, the bottlenck distance between them is bounded above by the L_∞-norm. I.e. $d_B(X, Y) \leq \|f - g\|_\infty$.*

Although for the purpose of this thesis the stability theorem for filtrations suffices, we shortly continue with a generalization of theorem for a much broader class of functions defined as:

Definition 5.3 (Tame functions)**.** *Let $\mathbb{X}$ be a triangulable manifold. A function $f : \mathbb{X} \to \mathbb{R}$ is tame if it has a finite number of critical values[2] and the homology groups $H_k(f^{-1}(-\infty, a])$ are finite-dimensional for all $k \in \mathbb{Z}$ and $a \in \mathbb{R}$.*

Example 5.2. *Morse functions on compact manifolds and piece-wise functions on finite simplicial complexes are both tame functions.*

Consider a rising sublevel set of a tame function $f : \mathbb{X} \to \mathbb{R}$. Each critical value changes the topology of the manifold. In analogy with filtrations of a complex, the increasing sublevel sets of the manifold create a sequence of homology group connected by maps. Denote for each dimension d the map $f_d^{a,b} : \mathsf{H}_d(\mathbb{X}_a) \to \mathsf{H}_d(\mathbb{X}_b)$ as the map from the d-th homology group at sublevel set a to sublevel set $b > a$. Than in a manner comparible with the construction of the vineyards above we can create sequences of homology classes over increasing sublevel sets of $\mathbb{X}$. From there we can again relate back to the L_∞ norm and obtain the:

Theorem 5.2 (Stability theorem for tame functions)**.** *Let $\mathbb{X}$ be a triangulable topological space with continuous tame functions $f, g : \mathbb{X} \to \mathbb{R}$. The functions f and g define persistence diagrams X respectively Y. The persistence diagrams satisfy $d_B(X, Y) \leq \|f - g\|_\infty$.*

Basically, the proof follows the same lines as the exposition prior to the stability theorem for filtrations. We don't state the proof here as it is technically involved and not directly relevant for this thesis. We refer the interested reader to [Cohen-Steiner et al., 2007]. Both stability theorems above are based on the L_∞ norm. Stability results of persistence diagrams for general L_p norms are found as well [Cohen-Steiner et al., 2010]. Unfortunately they require Lipschitz functions,

[2]Technically, a finite number of *homological* critical values is required. For a Morse function f on a smooth manifold, as we consider here, the homological critical values of f coincide with its classical critical values.

i.e. functions that satisfy $|f(x) - f)y)| \leq C\|x - y\|$ for some constant $C \in \mathbb{R}$. In general density fields don't satisfy this condition, so we refrain from further investigation at this point.

When the stability theorems are applied to this thesis, we can state that uncertainty in a measurement value will have a bounded effect on the persistence diagram. The data pipeline as such should be stable. However, previous doesn't say anything about the group behaviour of the thousands of points of which our persistence diagrams consist. As in this direction no mathematical results exist, we continue with an experimental perturbation analysis to identify possible group behaviour.

5.2 Experimental stability - perturbation analysis

Previous section tells that uncertainties in galaxy positions, be it measurement or analysis induced, will have a limited effect on each of the points in the persistence diagram. However, it doesn't tell how all these tiny uncertainties change the diagram as a whole. Therefore, a perturbation analysis is performed. We take a density field and perturb a certain random fraction f_p of the particles by a predefined amount Δp in a random direction. We put the perturbed fields through the data pipeline and compare the resulting persistence diagrams with each other and the unperturbed case. To distinguish between variations due to intrinsic randomness and difference in $(f_p, \Delta p)$ pair values, for each $(f_p, \Delta p)$ pair the procedure is repeated several times. A Mersenne Twister random number generator [3] is used to determine which particles to move in which direction.

Concretely, the following specific experimental values were chosen:

- The data pipeline is computationally very intensive, so instead of a full 300 Mpc^3 density field a 75 Mpc^3 subbox is considered. As the influence of perturbations might depend on density, we choose a high and a low density subbox. The high and low density subboxes are selected by visual inspection from a Gadget LCDM simulation (run 14).

- The influence of perturbations might also depend on the density variance within the subbox. The density distribution variance is monotically increasing with time, thus for both density subboxes the earliest ($z = 3.8$) and latest ($z = 0.0$) available times are considered. The resulting four density fields are shown in figure 5.4.

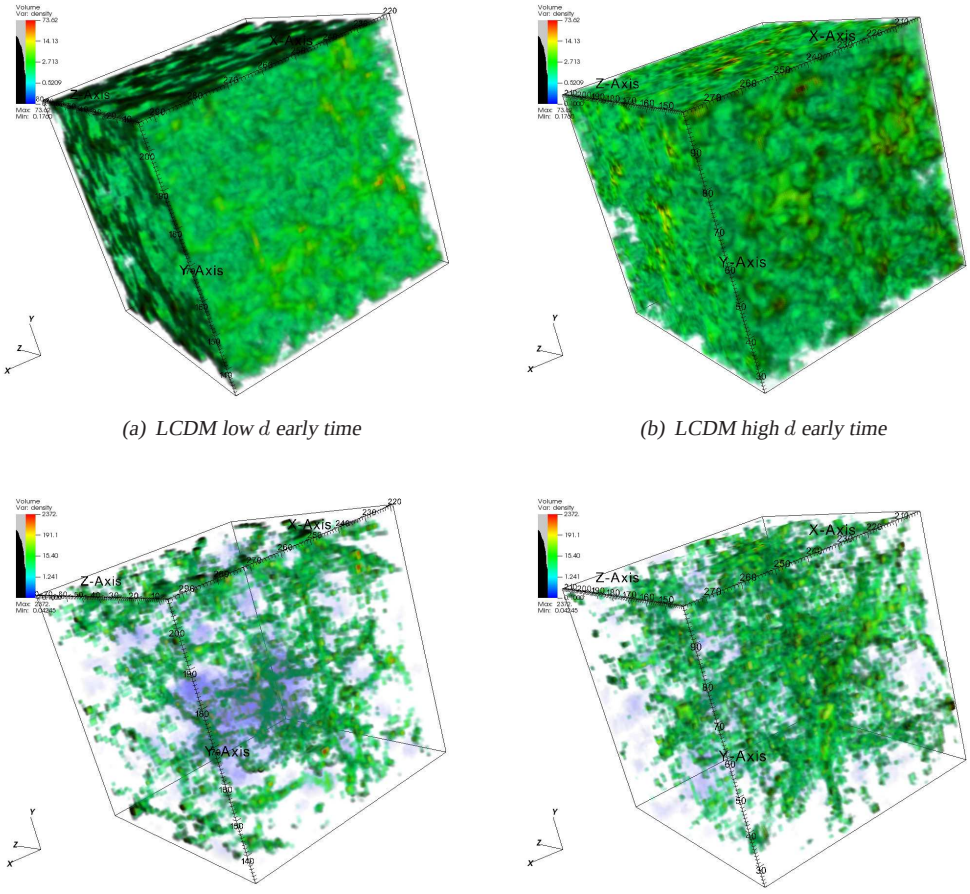
- Furthermore, there might (and is) a difference between the particle and density case, so we analyse both.

- For all density fields we take $f_p \in \{10^0, 10^{0.5}, 10^1, 10^{1.5}, 10^2\}$ and for Δp the same values times the interparticle distance with s_{ipd}.

- For each $(f_p, \Delta p)$ pair ten realizations are computed.

Summarizing the variables above, we get a total of 2 density subboxes $\times$ 2 times $\times$ 2 for particles/DTFE $\times$ 5 f_p's $\times$ 5 Δ_p's $\times$ 10 repetitions = 2000 experiments. The resulting 2000 persistence diagrams can be analysed in several ways. We use Betti curves as 1D summaries because they are both directly embedded in mathematical theory and represent physical features.

[3] The default C++ Mersenne Twister 19937 with a seed coming from the C++ standard random number generator, whose seed in turn came from the system time.



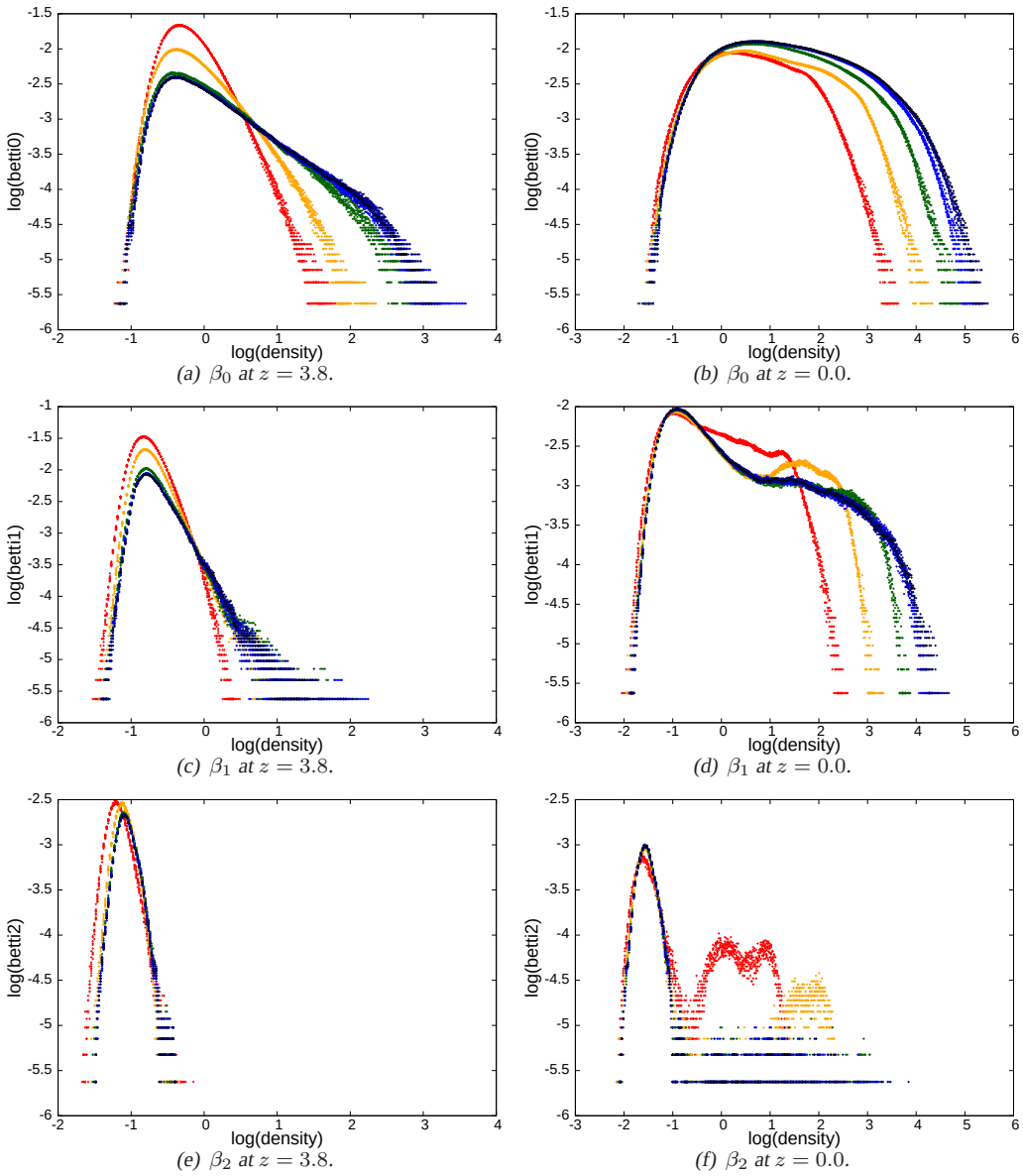
(a) LCDM low d early time *(b) LCDM high d early time*

(c) LCDM low d late time *(d) LCDM high d late time*

Figure 5.4: ***3D rendering of the LCDM subboxes.***

5.2.1 Particle perturbation analysis

Figure 5.5 shows β_0, β_1 and β_2 for LCDM at both times when all particles are perturbed. Interestingly, increasing perturbation magnitude doesn't increase the spread around the Betti curve of the unperturbed case. Instead the curves are systematically shifted to a different position, such that the difference between two highly perturbed curves is smaller than the difference of either with the unperturbed case. With increasing perturbation magnitude, the early time curves increase at lower density regions and decrease at smaller density regions. Probably because particle spreading destroys dense structures but consequently creates structures at lower densities. The later time curves also show the near disappearance of Betti curves at the highest densities but behaviour at lower densities depends on Betti number. For β_0 (shown here) an overall decrease is visible: apparently for more evolved density fields particle spreading doesn't create additional components at lower densities but the particles are just absorbed in the existing ones. β_1 increases for middle ranged densities and β_2 shows a drastic increase at middle ranged densities for strong perturbations.



(a) β_0 *at* $z = 3.8$. *(b)* β_0 *at* $z = 0.0$.

(c) β_1 *at* $z = 3.8$. *(d)* β_1 *at* $z = 0.0$.

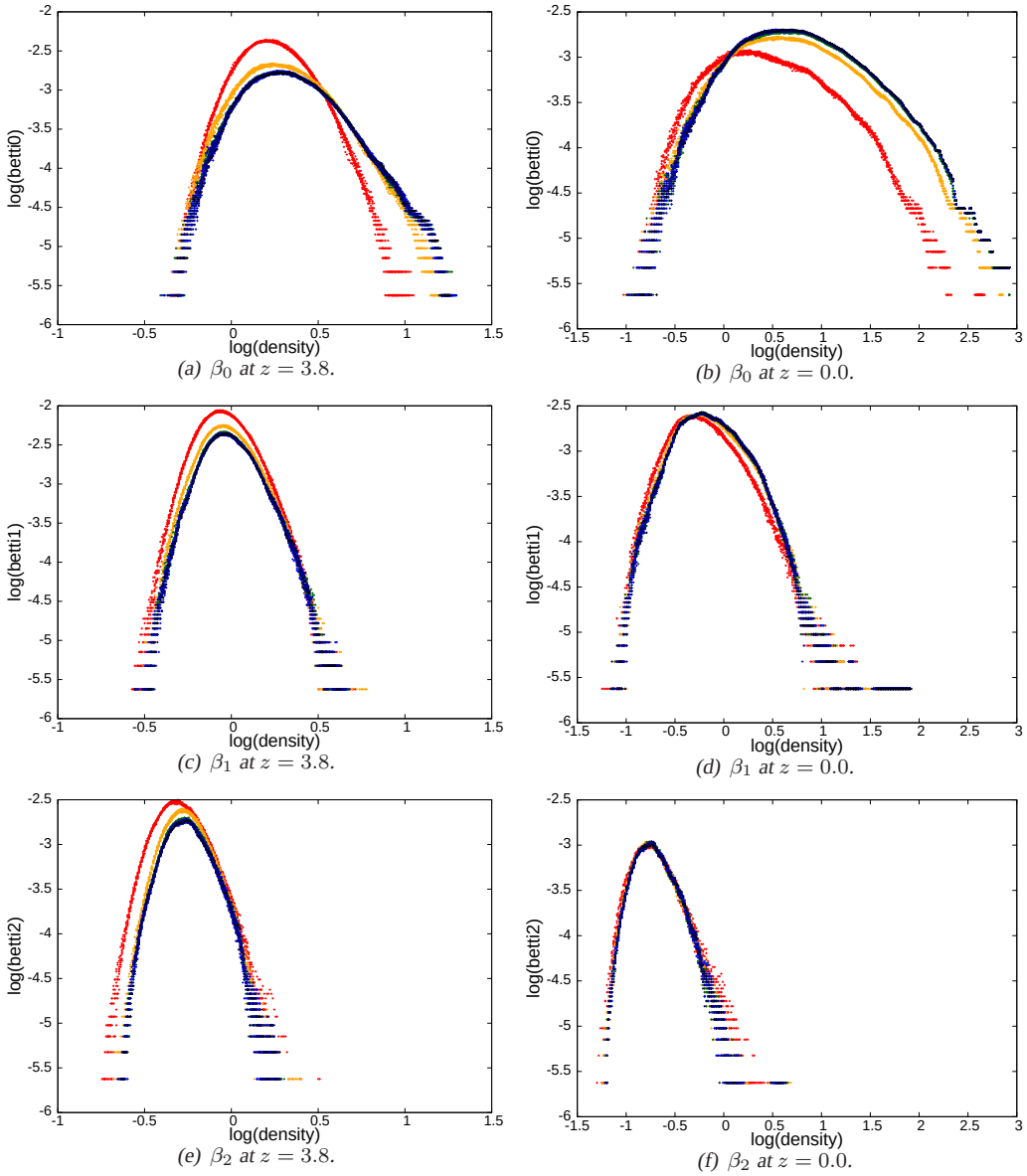
(e) β_2 *at* $z = 3.8$. *(f)* β_2 *at* $z = 0.0$.

Figure 5.5: ***Betti curves for particle perturbation analysis.*** *In order of priority in case of overlapping points: Betti curves for non-perturbed case (black) and perturbed cases, with perturbation magnitudes of* $1 \cdot s_{ipd}$ *(dark blue),* $10^{0.5} \cdot s_{ipd}$ *(sky blue),* $10^{1} \cdot s_{ipd}$ *(green),* $10^{1.5} \cdot s_{ipd}$ *(yellow) and* $10^{2} \cdot s_{ipd}$ *(red). For each of the Betti curves above, all ten realizations are shown for the high density field. The Betti curves for the low density field aren't show, as their behaviour is analogous to the high density case. Left to right: increasing time and top to bottom: increasing Betti dimension.*

5.2.2 DTFE perturbation analysis

The Betti curves of the density field are much less influenced by perturbations. Probably the method which generates the density field absorbs quite some of the changes in individual particle position. Globally, where there is an effect it is analogous to the particle case, as can be seen in figure 5.6. From these graphs we see that a perturbation up to around 10 percent doesn't visibly perturb the diagram. Perturbations of $10^{1.5}$ or larger do give a different curve. The influence of the magnitude of the perturbation goes down with increasing Betti number. The increase mentioned for β_1 and β_2 for middle ranged densities in the particle case disappears. For β_0 and

$\Delta p < 1 \cdot s_{\text{ipd}}$ there is a high density knee, but this disappears for higher Betti numbers and larger perturbations.



(a) β_0 *at* $z = 3.8$. *(b)* β_0 *at* $z = 0.0$. *(c)* β_1 *at* $z = 3.8$. *(d)* β_1 *at* $z = 0.0$. *(e)* β_2 *at* $z = 3.8$. *(f)* β_2 *at* $z = 0.0$.

Figure 5.6: ***Betti curves for density perturbation analysis.*** *In order of priority in case of overlapping points: Betti curves for non-perturbed case (black) and perturbed cases, with perturbation magnitudes of* $1 \cdot s_{ipd}$ *(dark blue),* $10^{0.5} \cdot s_{ipd}$ *(sky blue),* $10^{1} \cdot s_{ipd}$ *(green),* $10^{1.5} \cdot s_{ipd}$ *(yellow) and* $10^{2} \cdot s_{ipd}$ *(red). For each of the Betti curves above, all ten realizations are shown for the high density field. The Betti curves for the low density field aren't show, as their behaviour is analogous to the high density case. Left to right: increasing time and top to bottom: increasing Betti dimension.*

5.3 Nicing persistence diagrams

For modest perturbations the results of previous section are encouraging, such perturbations don't influence the Betti curves very much. For larger perturbations Betti curves systematically shift to other positions. Furthermore, persistence diagrams or Betti curves count cycles, 'but

not all cycles are created equal': some live very long whereas others are around just for a short while. Short-lived cycles are not that relevant: they represent small features or noise and for proper analysis can better be taken out. This section discusses several ways to do so, i.e. to ***nice*** a persistence diagram. A classical way to take out small features is ***smoothing*** the diagram. Technically, this often means convolving diagram with an approriately chosen function. We smooth the input field and discuss how this influences the topology in subsection 5.3.1. One of the huge advantages of a persistence diagram is that the relevance of a point can be inferred directly from the diagram. A more topological approach is ***manifold simplification***, discarding all points whose persistence pair $q_k = [\sigma_k \sigma_{k+1}]$ (cf. definition A.27) has a density ratio or difference smaller than a certain threshold value and rearranging the manifold accordingly. We consider this approach in subsection 5.3.2. Betti numbers are one of the most useful topology measures. Their definition can be generalized to *persistent Betti numbers*, which also gives a topological measure of importance of features. In subsection 5.3.3 we use this definition to nice the persistence diagrams.

5.3.1 Smoothing the input field

The classical way to nice a density field is by convolving it with a smoothing function. A natural and often used smoothing function is a Gaussian, which we will use here as well. Naturally particle positions can't be smoothed, so this approach only works for the density case. If the density field would be continuous, a Gaussian smoothing function has some unique and attractive properties:

1. It is the solution of the diffusion equation $\frac{\partial f}{\partial d} = \frac{1}{2}\frac{\partial^2 L}{\partial x^2}$ with initial condition $L(x;0) = f(x)$.
2. Increment of scale space parameter σ will not lead to additional local extrema or additional zero crossings.
3. Causality: for $\sigma_2 > \sigma_1$ than $L(x;\sigma_2)$ depends only on $L(x;\sigma_1)$.
4. It is shift invariant and doesn't depend on image values.

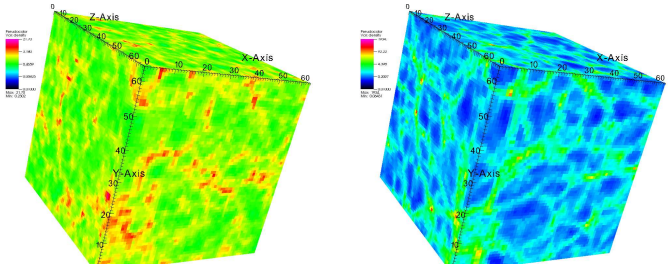
The numerical density field is computed on grid points. Although it is an approximation of the continuous density field, it consists of a set of discrete values. An intuitive approach might be to sample the Gaussian at grid point centres, but the attractive properties above won't be retained. To retain them, the *discrete Gaussian kernel* has to be used [Lindeberg, 1990], defined as:

Definition 5.4 (Discrete Gaussian kernel)**.** *At index n and scale space parameter σ, the discrete Gaussian kernel is given by:*

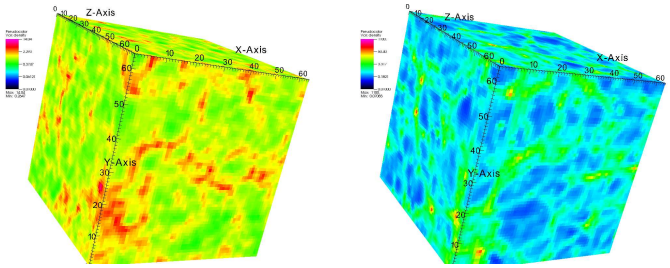
$$T(n,\sigma) = e^{-\sigma} I_n(\sigma) \tag{5.8}$$

with $I_n(\sigma)$ the modified Bessel function of integer order n.

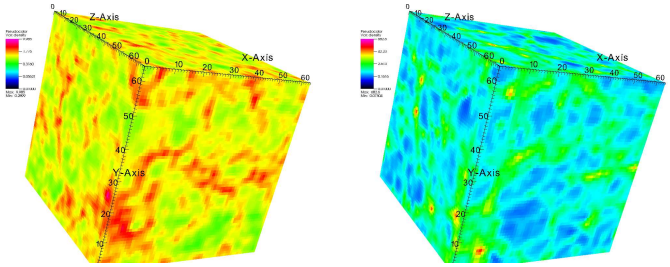
Here, we use a 3D normalized discrete Gaussian filter to smooth the density field. The filter is cut off in each dimension when the size of the filter reaches the limit of numerical accuracy. Figure 5.7 shows the high density field for various values of σ. The low density field is not shown as it behaves analogously. Observing that $\sigma = 0.25$ only marginally smooths the density field and that $\sigma = 2.0$ almost homogenizes it, we choose smooth $\sigma \in \{0.25, 0.50, 1.0\}$.



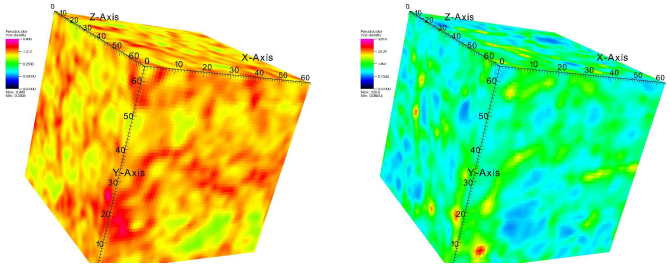
(a) Unsmoothed. *(b) Unsmoothed.*



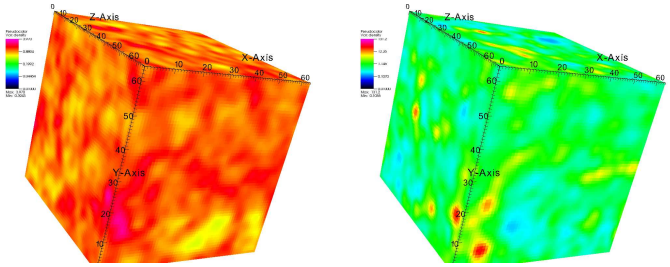
(c) Smoothing $\sigma = 0.25$. *(d) Smoothing $\sigma = 0.25$.*



(e) Smoothing $\sigma = 0.50$. *(f) Smoothing $\sigma = 0.50$.*



(g) Smoothing $\sigma = 1.0$. *(h) Smoothing $\sigma = 1.0$.*

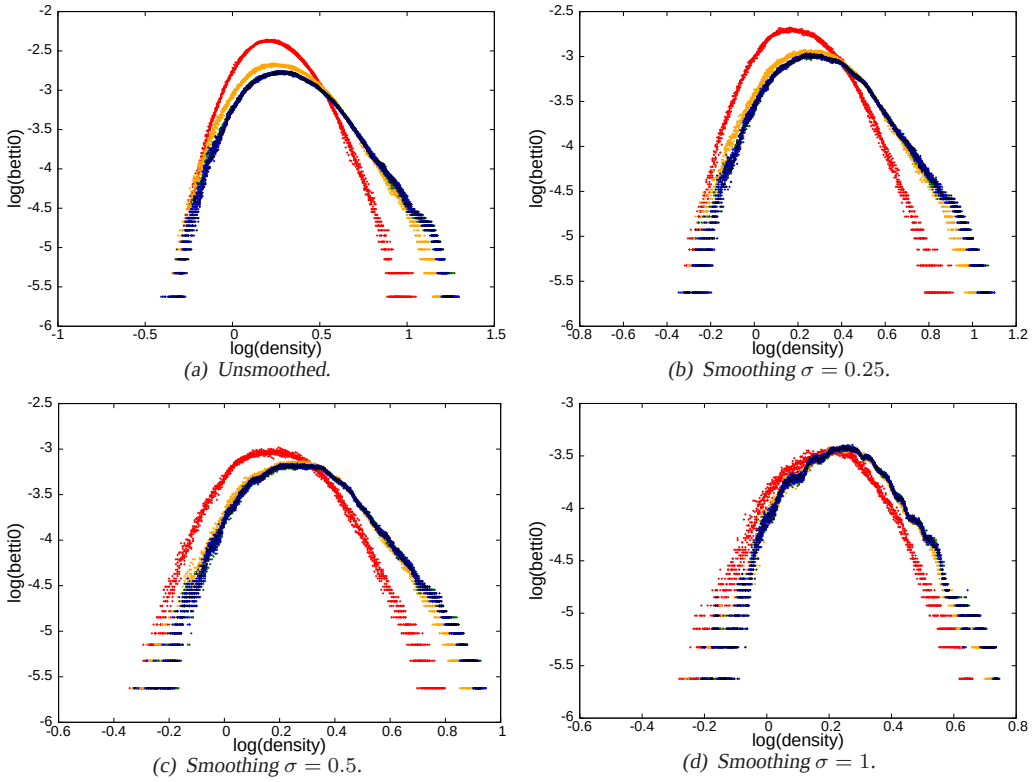


(i) Smoothing $\sigma = 2.0$. *(j) Smoothing $\sigma = 2.0$.*

Figure 5.7: ***The effect of smoothing on the density field.*** *The high density field smoothed with a discrete Gaussian with various values for σ. Left: earliest available time ($z = 3.8$) and right: today ($z = 0$).*

In this subsection, we taken the LCDM 0-cycles for the high density field at the earliest time as reference case. Figure 5.8 shows its Betti curves. Smoothing lowers the top of the Betti curve and decreases its range. Probably because a lot of smaller cycles are averaged out. The difference between the non-perturbed and perturbed cases remains about the same for small and moderate smoothing σ's. Only for large σ's, the most perturbed Betti curves come closer to their less perturbed brethren. The last could be explained by noting that increasing σ also increases the 'threshold level' of persistence pairs being taken out. As increasing perturbation magnitude generates more low persistence pairs (cf. section 5.2), highly perturbed fields will be stronger influenced by smoothing. Interestingly, increased smoothed makes the Betti curves less smooth.

Compared with this reference case, 1-cycles and 2-cycles show the same effect although less pronounced. The low density and later time Betti curves respond analogous to smoothing.



(a) Unsmoothed. *(b) Smoothing $\sigma = 0.25$.*

(c) Smoothing $\sigma = 0.5$. *(d) Smoothing $\sigma = 1$.*

Figure 5.8: ***The effect of smoothing on density Betti curves.*** *Betti curves of the LCDM high density field at $z = 3.8$ for various smoothing σ's. Line colors as in figure 5.6.*

5.3.2 Manifold simplification

In a persistence diagram a line perpendicular to the birth=death diagonal indicates the lifetime of a feature (cf. section 3.6). An intuitively nice way to delete low-persistent features is simply to remove all points in the persistence diagram closer to birth=death diagonal than a certain minimum $\mathcal{L}_{p,min}$. This corresponds to simplifying the manifold such that all features whose lifetime is smaller than this value are cancelled out. For a persistence pair $q_k = [\sigma_k \sigma_{+1}]$, in a loglog persistence diagram the distance from the birth=death diagonal corresponds to the persistence ratio

$$r(q_k) = \rho_D(\sigma_{k+1})/\rho_D(\sigma_k) \tag{5.9}$$

In a 'normal' axes persistence diagram, the distance from the birth=death diagonal corresponds to the persistence difference:

$$d(q_k) = \rho_D(\sigma_{k+1}) - \rho_D(\sigma_k) \tag{5.10}$$

Both might have their uses so we compute both, starting with the peristence ratio.

Ratio simplification

Cosmic structure has a large density range so within this thesis we mainly use a log-log representation of persistence diagrams. This makes ration simplification the most intuitive simplifcation to consider. If we know the ratio distribution function of $r(q_k)$, a certain significance-threshold can be set and all points below this threshold ignored. In the particle case this is non-trivial from the start. In the density case we might derive the ratio distribution function starting from the DTFE density distribution, which is excellently approximated by [Schaap, 2007]:

$$dp(\tilde{\lambda}) = \tfrac{1944}{5}\tilde{\lambda}^{-8}e^{-6/\tilde{\lambda}}d\tilde{\lambda} \tag{5.11}$$

Defining X and Y as random DTFE density variables for σ_k and σ_{k+1} respectively, we can write for their their joint distribution $f(x, y)$:

$$f(x,y) = P(X = x, Y = y) = P(X = x)\frac{P(Y = y|X = x \geq y)}{\int_0^x P(Z = z)dz} \tag{5.12}$$

where in the second equality (i) we took into account that for a persistence pair q_k we have $r(q_k) \geq 1$; (ii) we assumed the amount of points is very large such that taking out one point doesn't change the distribution and (iii) the denominator is for normalization. Plugging in equation 5.11 multiple times gives:

$$f(x,y) = \tfrac{1944}{5}x^{-8}e^{-6/x}\frac{y^{-8}e^{-6/y}}{\int_0^x z^{-8}e^{-6/z}dz} \tag{5.13}$$

with $x \in [0, \infty)$ and $y \in [x, \infty)$. Now we can define the ratio distribution $R = Y/X$ with cumulative distribution $F(r)$:

$$\begin{aligned} F(r) &= P(r \leq R) = P(Y \leq uX|X > 0) \\ &= \int_0^\infty \int_0^{ux} f(x,y)dxdy \\ &= \tfrac{1944}{5}\int_0^\infty \int_0^{ux} x^{-8}e^{-6/x}\frac{y^{-8}e^{-6/y}}{\int_0^x z^{-8}e^{-6/z}dz}dxdy \end{aligned} \tag{5.14}$$

This is a very nasty expression and quite likely incorrect: the only assumption we took above was $r(q_k) \geq 1$ but pairings are not that random. In general, very low density points are more likely to get paired with very high density points, and slightly low density points are more likely to get paired with slightly high density points. The exact pairing, however, depends on the detailed characteristics of the density field. There has been some preliminary research in this direction in the case of Gaussian fields (cf. [Feldbrugge et al.]) but this has so far been inconclusive and doesn't apply to strongly nonlinear fields like we discuss here. Therefore, the best way to procede might simply be obtaining $F(r)$ from experiment.

Removing small features actually serves two goals: (i) removing little persistent cycles from the persistence diagram and (ii) get out noise. The DMC is based on Delaunay triangulations and these are in particular sensitive to Poisson noise (cf. 4.2). Poisson noise will mainly have low persistent features. Therefore, when we proceed with some consideration, manifold simplification allows at the same time in an analytical way. First, we create a random Poisson particle distribution, compute its persistence diagram and determine its $F(r)$ experimentally. Then, we set a given small probability that a cycle is actually Poisson noise. The persistence value corresponding to this probability can be found from $F(r)$ and all cycles with smaller a smaller persistence value can be simplified out. Hence, we have simplified and de-noised the persistence diagram in one go!

Following this route, we create mock Poisson distributions for 75 Mpc ^{3} boxes with on average 64^3 particles, determine their corresponding to $F(r)$ and use this to set a simplification threshold level. The mock catalogues are created in the following way:

1. Draw from the Poisson distribution with $\lambda = 64^3$ a random variable which determines the number of particles in the box.

2. The particles are distributed within the box using a random uniform distribution.

Technically, the C++ Boost library Poisson distribution and uniform distribution are used, again with a Mersenne Twister random number generator. Of course the persistence diagram with a random amount of randomly placed particles is very random. To get an idea of the intrinsic variation 100 mock Poisson catalogues were made. Their cumulative ratio distributions $F(r)$ are displayed in figure 5.9 showing the following features:

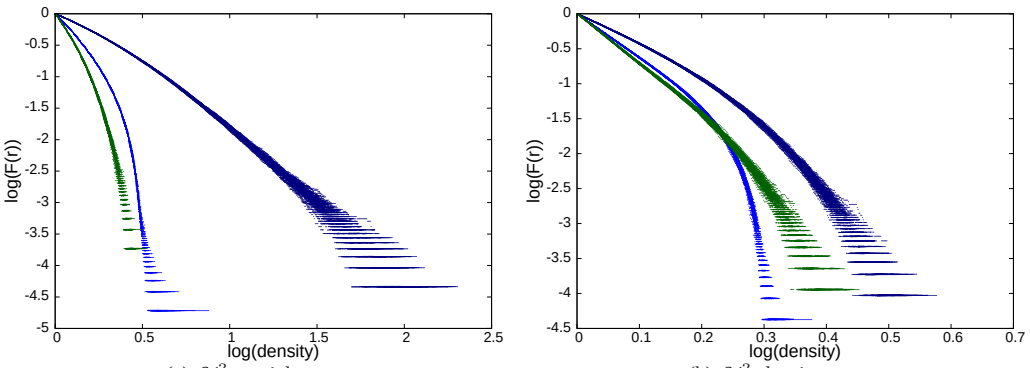
- There is huge difference between the density and particle case. The slope of $F_{\text{particle}}(r)$ is much lower than the slope for $F_{\text{dtfe}}(r)$ and consequently the first reaches up to an order of magnitude higher ratio's. A probable cause is that the particle case is much more sensitive in the high density regions and so several very high density peaks might exist. With DTFE these tiny very high density patches are smeared out.

- In the density case the 1-cycle and 2-cycle curves cross whereas this doesn't happen for the particle case. We have no intuitive explanation to offer at this point.

- Although there is some spread in $F(r)$, especially at large r values, the spread is modest.

- Differences aside, both the particle and density graphs clearly show a large excess of low persistent points.

We set three significance values: $F_1(r) = 0.1, F_2(r) = 0.01$ and $F_3(r) = 0.001$. The corresponding threshold values of r are tabled in table 5.1. Using these thresholds, all persistence pairs with a smaller ratio are cancelled. For completeness, we note that [Sousbie, 2011] followed this approach as well and defines thresholds 'in analogue with the Gaussian case' by using an inverse error function. However there is not a Gaussian anywhere in sight, so we refrain from this approach. The resulting Betti curves for the particle and density case are shown for a representative case in figures 5.10 and 5.11 respectively. We directly see that increasing simplification threshold lowers the Betti curves as many low-persistent cycles are taken out. Also directly visible is the changing difference between the non-perturbed and slightly perturbed cases and the more perturbed cases with increasing simplification threshold levels. Both for the particle and density case highly perturbed fields are (much stronger) influenced by simplification than less



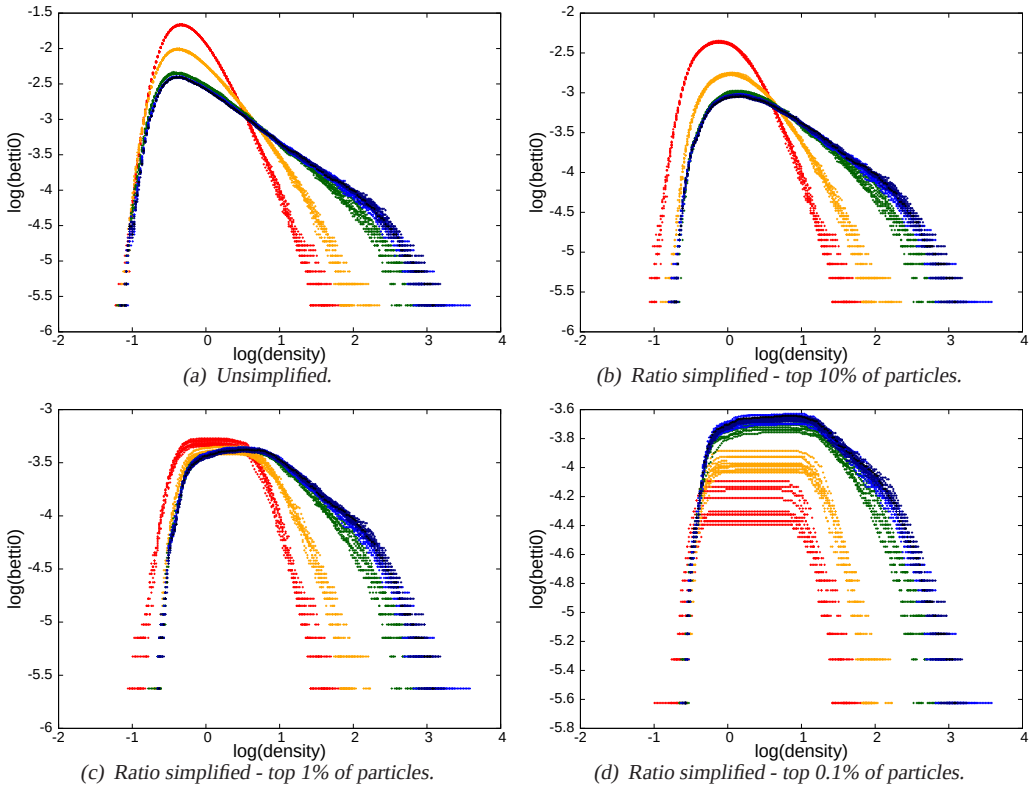
(a) 64^3 particle case. *(b) 64^3 density case.*

Figure 5.9: ***Cumulative probability*** $F(r)$**.** *The $F(r)$ graphs show the probability of the existence of a persistence pair with ratio equal or larger than r. Mazarine: 0-cycles; sky-blue: 1-cycles and green: 2-cycles; for 100 mock Poisson distributions.*

Sign. values	density			particles		
	0-cycles	1-cycles	2-cycles	0-cycles	1-cycles	2-cycles
0.1	0.212 ± 0.002	0.156 ± 0.002	0.142 ± 0.002	0.628 ± 0.004	0.282 ± 0.002	0.202 ± 0.002
0.01	0.344 ± 0.004	0.252 ± 0.002	0.262 ± 0.004	1.08 ± 0.002	0.424 ± 0.002	0.320 ± 0.006
0.001	0.430 ± 0.010	0.286 ± 0.002	0.340 ± 0.010	1.50 ± 0.034	0.480 ± 0.002	0.396 ± 0.010

Table 5.1: r ***threshold values.*** *Threshold values of persistence ratio r for several cycle dimensions and both the particle and density case at various significance values.*

perturbed fields. As possible explanation we offer that increasing perturbation magnitude generates more low persistence pairs (cf. section 5.2). Thus highly perturbed fields will be stronger influenced by simplification, analogously as with smoothing. β_1 and β_2 show similar behaviour as β_0. The increase of β_1 and β_2 at middle ranged densities disappear at high simplification thresholds.



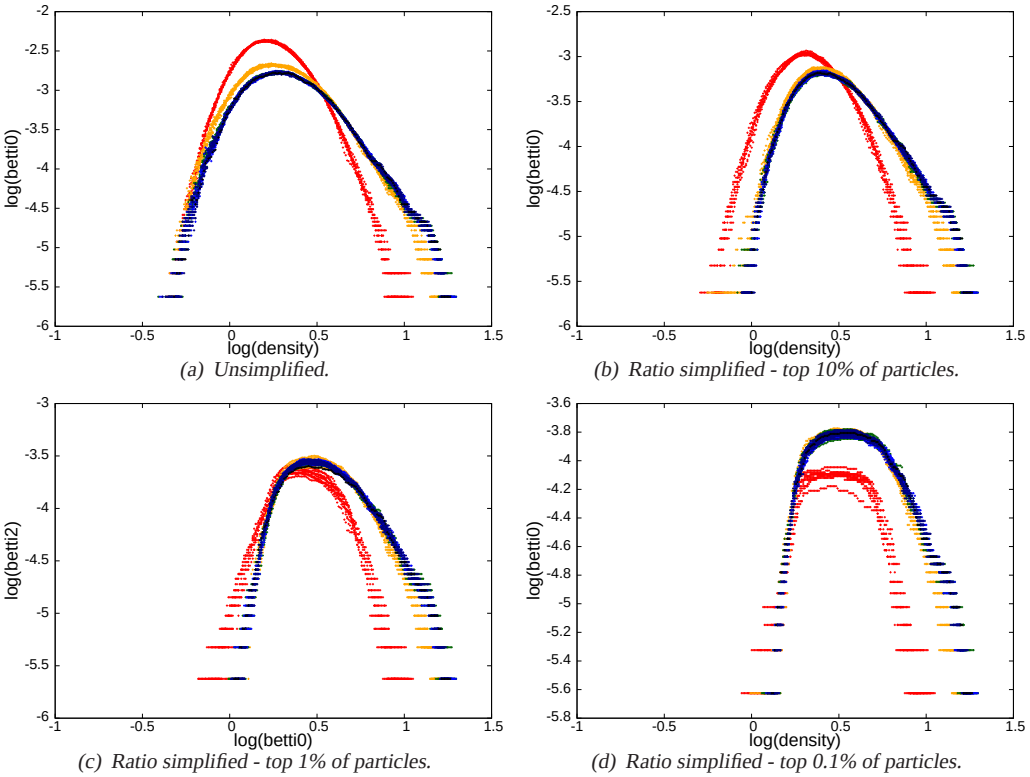
(a) Unsimplified.

(b) Ratio simplified - top 10% of particles.

(c) Ratio simplified - top 1% of particles.

(d) Ratio simplified - top 0.1% of particles.

Figure 5.10: ***The effect of ratio simplification on particle Betti curves.*** *Betti curves of the LCDM high density field at $z = 3.8$ for various simplification ratios, particle case. Line colors as in figure 5.5.*



(a) Unsimplified.

(b) Ratio simplified - top 10% of particles.

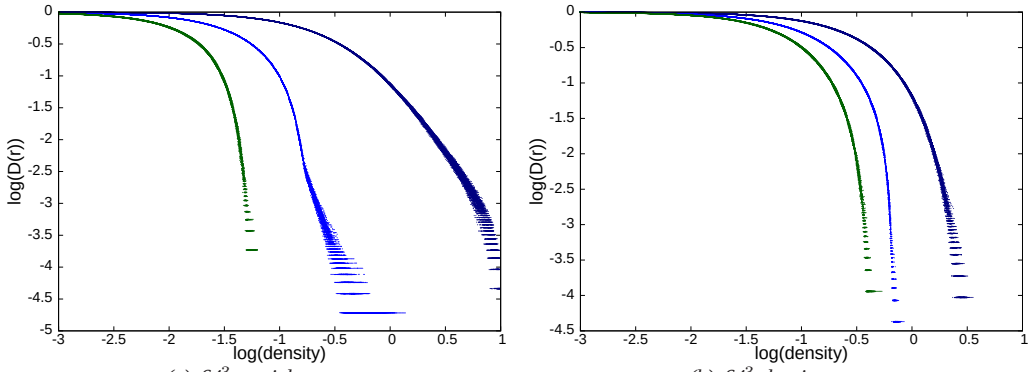
(c) Ratio simplified - top 1% of particles.

(d) Ratio simplified - top 0.1% of particles.

Figure 5.11: ***The effect of ratio simplification on density Betti curves.*** *Betti curves of the LCDM high density field at $z = 3.8$ for various simplification ratios, density case. Line colors as in figure 5.6.*

Difference simplification

The persistence ratio follows swiftly from the log-log persistence diagram but in terms of physics the difference instead of the ratio is more intuitive. Therefore, the analysis above is repeated for the cumulative difference distribution $D(d)$, which is shown for the Poisson mock distributions in figure 5.12. Compared with $F(r)$ we see that $D(d)$ decreases much steeper and the graphs for different cycles don't cross. The values corresponding to the same thresholds as before are



(a) 64^3 particle case. *(b) 64^3 density case.*

Figure 5.12: ***Cumulative probability*** $D(d)$**.** *The $D(d)$ graphs show the probability of the existence of a persistence pair with difference equal or larger than d. Mazarine: 0-cycles; sky-blue: 1-cycles and green: 2-cycles; for 100 mock Poisson distributions.*

tabled in table 5.2 and the corresponding Betti curves are shown in figures 5.13 and 5.14. At first glance the similarity with the ratio-simplified Betti curves is most striking. Considering the graphs in more detail some differences emerge between both methods of simplification can be found. Compared with ratio-simplification, difference-simplification:

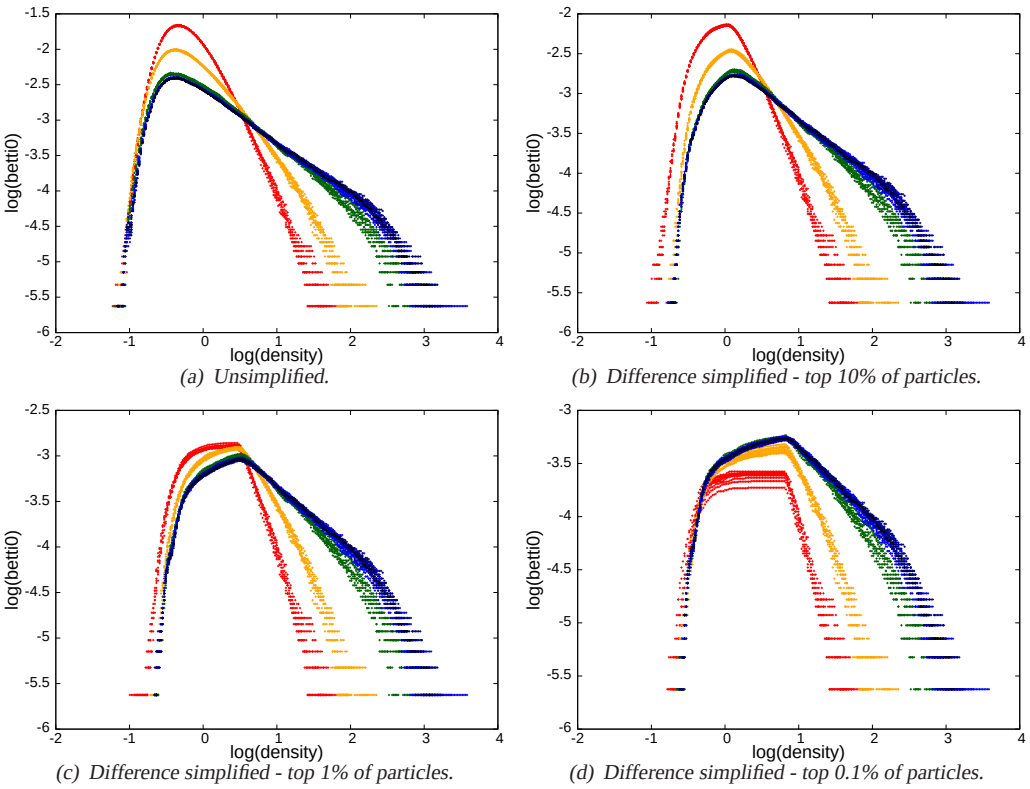
- is 'more effective' on small densities and 'less effective' on high densities;
- consequently, spurious high density features are dealt with less effectively and inversely for spurious low density features;
- is less sensitive to perturbations, i.e. the higher magnitude perturbation curves lie closer to the smaller magnitude perturbation curves;
- lowers the Betti curve less with increasing simplification threshold for β_0, about the same for β_1 and much more for β_2.

Sign. value	density			particles		
	0-cycles	1-cycles	2-cycles	0-cycles	1-cycles	2-cycles
0.1	-0.070 ± 0.004	-0.456 ± 0.004	-0.726 ± 0.004	-0.078 ± 0.005	-1.004 ± 0.002	-1.520 ± 0.004
0.01	0.202 ± 0.008	-0.252 ± 0.004	-0.510 ± 0.006	0.418 ± 0.002	-0.820 ± 0.002	-1.368 ± 0.006
0.001	0.350 ± 0.016	-0.194 ± 0.004	-0.428 ± 0.008	0.790 ± 0.026	-0.666 ± 0.012	-1.304 ± 0.008

Table 5.2: d ***threshold values.*** *Threshold values of persistence difference d for several cycle dimensions and both the particle and density case at various significance values.*

For both simplification methods, compared with smoothing a few important things can be noticed:

- simplification increases the minimum density value of Betti curves, whereas smoothing mainly decreases their maximum value;

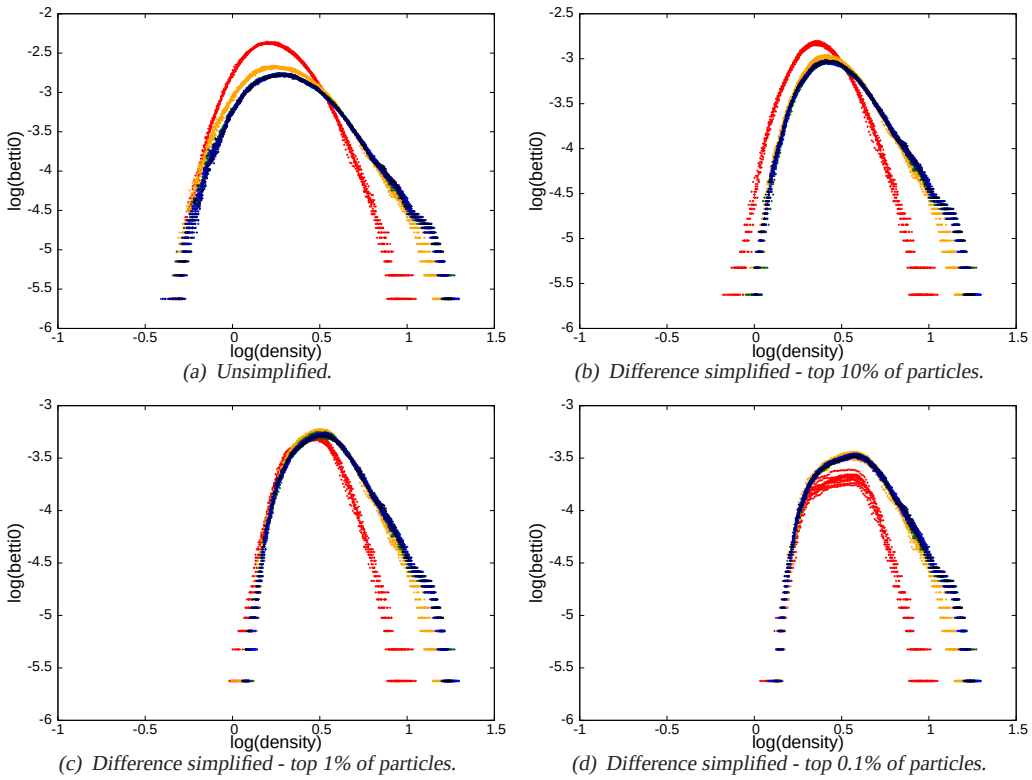


(a) Unsimplified. *(b) Difference simplified - top 10% of particles.*

(c) Difference simplified - top 1% of particles. *(d) Difference simplified - top 0.1% of particles.*

Figure 5.13: ***The effect of ratio simplification on particle Betti curves.*** *Betti curves of the LCDM high density field at $z = 3.8$ for various simplification differences, particle case. Line colors are as before. Line colors as in figure 5.5.*

- simplification results in much smoother Betti curves than smoothing;
- especially difference simplification has a much stronger effect on higher Betti numbers.



(a) Unsimplified.

(b) Difference simplified - top 10% of particles.

(c) Difference simplified - top 1% of particles.

(d) Difference simplified - top 0.1% of particles.

Figure 5.14: ***The effect of ratio simplification on density Betti curves.*** *Betti curves of the LCDM high density field at $z = 3.8$ for various simplification differences, density case. Line colors are as before. Line colors as in figure 5.6.*

5.3.3 Persistent Betti numbers

In a log-log persistence diagram ratio-simplification corresponds with ignoring all points whose distance to the birth-death diagonal is smaller than some value. For Betti curves this corresponds graphically to removal of the lowest persistent corner of a Betti square, see the green triangles in figure 5.15. Specifically for Betti numbers, persistent Betti numbers (cf. definition A.26) give another method to indicate their longevity. Graphically, persistent Betti numbers are obtained by moving the Betti square left-upwards. To see how this subtle change in persistence measure influences Betti curves the analysis above is repeated for persistent Betti curves. Given the limited differences between ratio and difference simplification, we only consider the ratio case. Using the threshold values found for ratio simplification we plot the persistent Betti curves in figures 5.16 and 5.17 for the particle and density case respectively.

Look again at figure 5.15. In persistence diagrams the difference between ratio simplified Betti curves and persistent Betti curves are a horizontal and vertical ‘rectangular’ region. The horizontal region corresponds with low-birth low-death points and the vertical region with high-birth high-death points. That persistent Betti numbers leave these regions out is clearly reflected in the persistent Betti curves. They resemble these of ratio simplification neatly for a low simplification threshold but with increasing threshold ‘loose’ the low and high density tails. Furthermore, the height of the Betti curves drops sharper. These effects become even stronger for β_1 and β_2, the last almost completely disappears for a high simplification threshold.

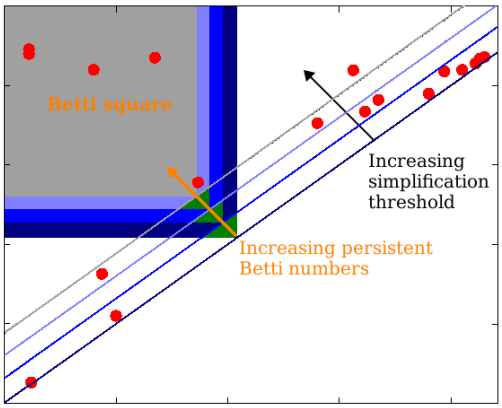
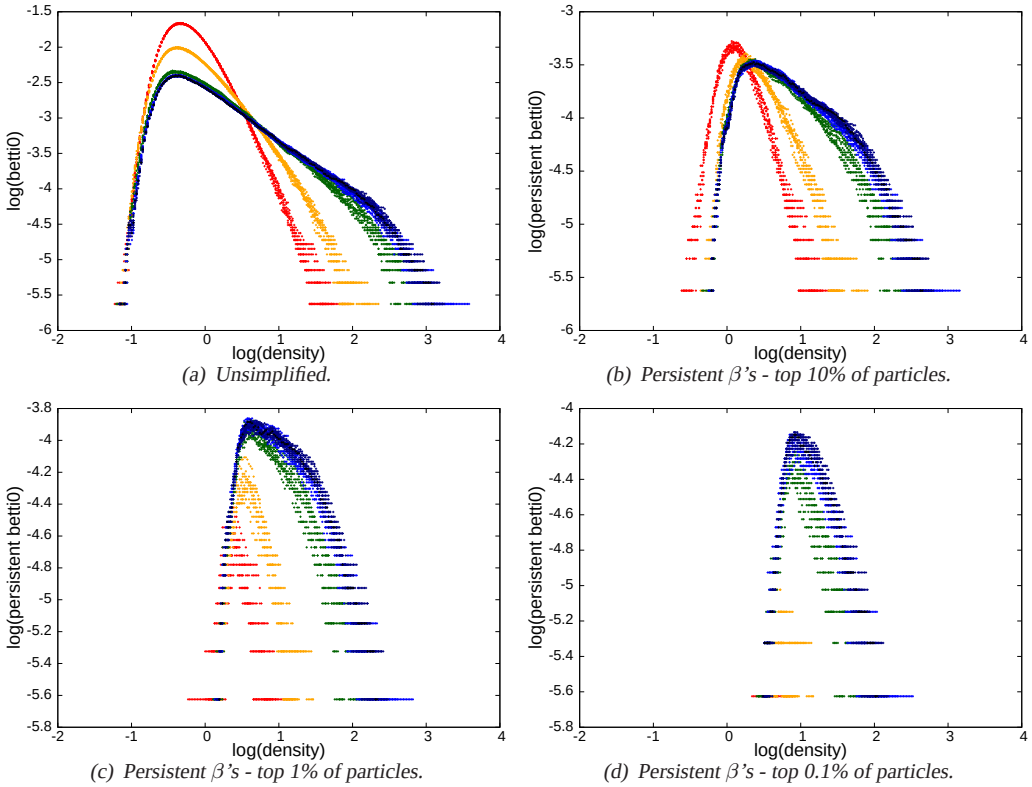


Figure 5.15: ***Simplification vs persistent Betti numbers****. Graphically, Betti numbers given by the number of points in the Betti square in a persistence diagram. Simplification takes out the bottom-right corner of this square, here indicated with the green triangles. Considering persistent Betti numbers corresponds with moving the square upwards. The difference between simplification and persistent Betti numbers are the colored 'rectangles' corresponding with low-birth low-death and high-birth high-death points.*



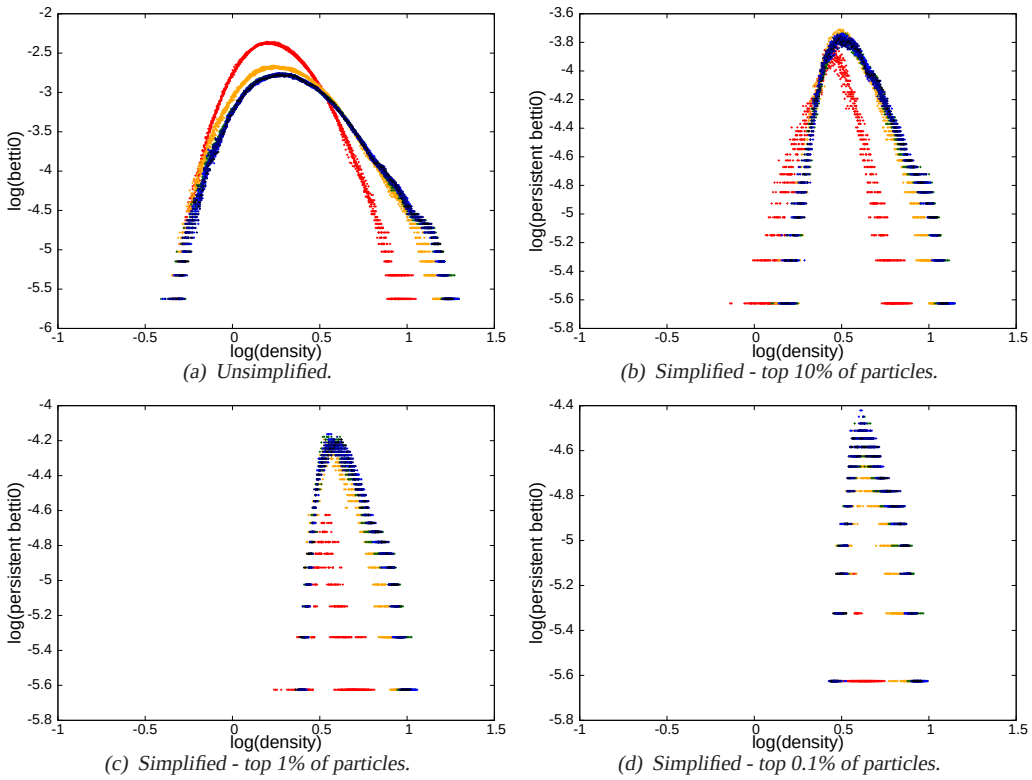
(a) Unsimplified. *(b) Persistent β's - top 10% of particles.*

(c) Persistent β's - top 1% of particles. *(d) Persistent β's - top 0.1% of particles.*

Figure 5.16: ***The effect of ratio simplification on particle Betti curves.****Persistent Betti curves of the LCDM high density field at $z = 3.8$ for various simplification ratios, particle case. Line colors as in figure 5.5.*

5.4 Summary

Data used for persistence diagrams may contain errors, be it due to observational uncertainties or due to data-processing methods. For tame mathematical functions it is known how uncertainties will change individual points in the persistence diagram. Yet it is until now unknown how systematic uncertainties in many or all data points will influence the persistence diagram as a



(a) Unsimplified. *(b) Simplified - top 10% of particles.*

(c) Simplified - top 1% of particles. *(d) Simplified - top 0.1% of particles.*

Figure 5.17: ***The effect of ratio simplification on density Betti curves.*** *Persistent Betti curves of the LCDM high density field at $z = 3.8$ for various simplification ratios, density case. Line colors as in figure 5.6.*

whole. To get an handle on this both for the particle and density case a high-density and low-density region were selected. At an early and late time the particles inside were perturbed with various magnitudes in arbitrary directions. The following observations can be made:

1. The mathematical results of the influence of errors on individual points are valid for the densities under consideration. (But these results don't tell how the diagram as whole is influenced.)

2. Betti curves are relatively robust against perturbations up to 10 %.

3. Perturbations of predefined magnitude in random directions don't lead to a spread of the Betti curves around the unperturbed case but to a systematic shift. In general this shift is downwards for high densities and upwards for (very) low densities.

4. For the particle case, large perturbations can lead to irregular changes in the Betti curves. The density case seems much more robust against irregularities.

5. Besides these irregularities, the effect of perturbations decreases with increasing Betti number.

Persistence diagrams contain a wealth of information. Raw diagrams contain large amounts of noise as well in the form of low persistent features and noise. Several methods are available to nice the diagrams: (i) the classical method of smoothing; (ii) ratio-simplification; (iii) difference-simplification; and in case of Betti curves (iv) persistent Betti curves. For each of these methods, the persistence diagrams were niced and the analysis above was repeated. This shows us:

1. All nicing methods have the same quantative effect: irrelevant features are cancelled, lowering the height of the curve and influencing its lower and upper density boundaries.

2. None of the nicing methods decreases the difference between the unperturbed and highly perturbed cases systematically. Often, nicing the diagram even increases the difference.

3. Interesingly, nicing diagrams has the same quantative effect on perturbations as time evolution: the Betti curves of highly perturbed cases are reduced much more than those of little perturbed cases. This probably means that both have the same effect: strong structure is enhanced and small structure surpressed.

4. Classical smoothing influences both the lower and upper densities boundaries, the latter more than the first. Persistent Betti numbers also strongly influence both boundaries. Ratio-simplification mainly influences the lower densities and difference simplification even more.

5.5 Discussion and conclusions

From the summary observations above, two things need further attention. First, the choice of perturbation type might influence the resulting changes in the persistence diagram. I.e. the observed systematic shift of the Betti curves might be caused by the choice for perturbing particles a set size in random direction. This type of perturbation will disperse closely clustered groups of particles and thus mainly destroy structure. We've seen that persistence diagrams are very sensitive to the amount of structure. Consequently, the systematic shift under increasing perturbations might mainly reflect the loss of structure rather than anything else. Further support to this idea is lended by the enhanced decrease under increasingly stringest nicing of large perturbation magnitude Betti curves. For a general analysis as here the current type of perturbations sufficies. It reflects where persistence is most senstive for (structure) and illustrates that up to quite large perturbations the persistence diagrams are not much influenced. For specific and strong observational biases in future applications, it might be worthwhile to repeat this analysis perturbing the particles in specific way.

Second, which nicing method is to be used? The ideal nicing method will have the following properties:

1. it cancels out the points corresponding to 'irrelevant features';

2. by doing so it enhances the differences between features we are interested in;

3. it decreases the effect of perturbations;

4. is not very computationally demanding.

Let's consider these points one by one.

1. All methods cancel out 'irrelevant features' but each methods defines an irrelevant feature in a different way. Simplification defines 'relevant' intuitively as having a birth/death ratio or difference above a certain threshold. Persistent Betti numbers define 'relevant' based on the density ratio after which a point seizes to be in the Betti square. In classical smoothing 'relevant' depends in a non-trivial way on the environment.

2. Strikingly, these different definitions of relevant give analogous results but also some major differences can be found. In particular, classical smoothing and persistent Betti numbers cancel out the low and high density tail of Betti curves quite fast. Later on, we will see that it is especially this lower tail which is of interest.

3. Unfortunately none of the methods decrease the effect of perturbations. Mostly, nicing even increases the effect of perturbations, as perturbations generate more low-persistent structure which is easier niced out.

4. Smoothing is done on the density field and only afterwards the filtration can be computed and the persistence diagram created. As particularly the last one is the computational bottleneck, with smoothing it is difficult to 'play around' with various nicing thresholds. The other nicing methods use the persistence diagram itself which makes playing around easier. Furthermore, smoothing is only applicable to the density case.

Since simplification: (i) performs in general terms as good as the other methods; (ii) has an intuitive definition of 'relevant'; (iii) doesn't cancel out density tails too fast; (iv) is computationally efficient and (v) is applicable to both the particle and density case, in the rest of this thesis simplification will be used as the default nicing method. In this thesis we mainly consider log-log persistence diagrams, making ratio simplification the logical choice.

Things brings us to the conclusions of this chapter:

1. Persistence diagrams (as represented by Betti curves) are - as expected - very sensitive to the presence of structure.

2. Persistence diagrams are robust against uncertainties up to at least 10 %.

3. Persistence simplification is an excellent, intuitive and computationally efficient method to take out irrelevant features and noise. Whether ratio or difference simplification is used depends on the specifics of the situation under consideration. As we mainly consider log-log persistence diagrams here, ratio simplification is the logical choice.

CHAPTER 6

On the persistence of LCDM

In chapter 3 topology was introduced. Topology is an abstraction of geometry allowing us to describe the cosmic web in terms of its morphological components: voids, walls, filaments and nodes. The main tool introduced there for topological analysis is the *persistence diagram* (cf. definition 3.31). A persistence diagram is a set of points indicating the birth and death densities of physical features in a density field. The difference between the birth and death density values of a feature is a measure for its relevance and is called its *persistence* (cf. section 3.6).

Chapter 4 explained the step from topology to computational topology and chapter 5 showed that the data pipeline developed in chapter 4 is stable. Furthermore, chapter 4 showed the well-behaved-ness of *simplification* (cf. section 5.3): a topological analogue of smoothing that cancels low persistent features.

In this chapter we apply the topological machinery developed so far on a full LCDM simulation (cf. subsection 4.1.3 for details). In section 6.1 we describe the results in the following way:

1. a visual impression and interpretation of the persistence diagrams is given;
2. two 1D summary curves are parametrized;
3. the effects of simplification are considered.

Subsequently we investigate how the persistence diagram changes with z in section 6.2. We do so by computing it at several other snapshots of the simulation, each at a different z value. We results are described in the same way as above. We round up this chapter with summary observations and conclusions in section 6.3.

6.1 The reference case: LCDM at *z = 0* (run 15).

In total this thesis uses 3 DE models $\times$ 5 realizations $\times$ 8 redshifts is 120 files. This large amount of data precludes investigation of each individual file. Therefore we apply a different strategy and compare all results with respect to a reference case. As reference case we choose LCDM run 15 at $z = 0$ because:

- LCDM is the standard cosmological model (cf. the introduction of chapter 2);

- computation of the structure clustering parameter σ_8 (cf. equation 2.33) for LCDM shows that run 15 has the median σ_8 of all runs. This holds for all z.
- at lower z structure is more evolved and thus a possible dark energy imprint on cosmic structure will be largest.

A visual impression

Figure 6.1 gives an impression of the cosmic structure in LCDM (run 15) at $z = 0$. Large filaments connecting massive nodes are clearly visible, just as enormous empty regions devoid of anything. Figure 6.2 shows the basic persistence measures for LCDM at $z = 0$ (run 15). The persistence diagrams are displayed on the left. All diagrams show many short lived features near the birth=death diagonal and more long lived features further away. With increasing dimension the cycle density shifts from late-born late-died to early-born and early-died. This is understandable: the 0-cycles are related with components whereas the 1-cycles and 2-cycles are related with tunnels and shells, lower density regions inside components. The extendedness of the 0-cycle persistence diagram might incorrectly suggest there are more 0-cycles, but this is not the case. There are 288731 0-cycles, 587153 1-cycles and 215241 2-cycles in the persistence diagrams below. I.e. on average every component has around two tunnels and one shell.

On the right of figure 6.2 three 1D summary-graphs of the persistence diagrams (cf. figure 3.24 and the text above the figure) are shown:

1. The Betti curves (top) visualize the number of components, tunnels and shells alive at a certain density threshold.
2. The lifetime ratio curves (middle) give the amount of cycles as function of their lifetime ratio.
3. The product mean density curves (bottom) show the amount of cycles as function of mean density.

The Betti curves behave qualitatively analogous to the Betti curves found in previous chapter. With increasing dimension the curves shift to lower densities and become shallower. Increasing dimension shifts the lifetime curves strongly to lower rations and the mean density curves to lower densities.

Parametrization of 1D summary curves

On the eye, the shape of the mean density curve of figure 6.2 resembles that of a log-normal distribution quite well. This can be connected with known properties of the cosmic density field. [Coles and B., 1991] found that the large scale matter distribution is quite well approximated by a lognormal distribution. Interestingly, the shape of the Betti curves resembles the shapes of the mean density curves and are thus lognormally distributed as well. Betti curves show the amount of cycles alive at a certain density threshold in the rising sublevel set. There is no fundamental reason known to us why the amount of cycles should follow the same behaviour as the density field.

A lognormal distribution is a distribution of a variable whose exponent is normally distributed. The lognormal probability density function $f_{\ln}(x|\mu,\sigma)$ can be obtained in a straightforward way from the normal distribution using variable transformation:

$$f_{\ln}(x|\mu,\sigma) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right) \tag{6.1}$$

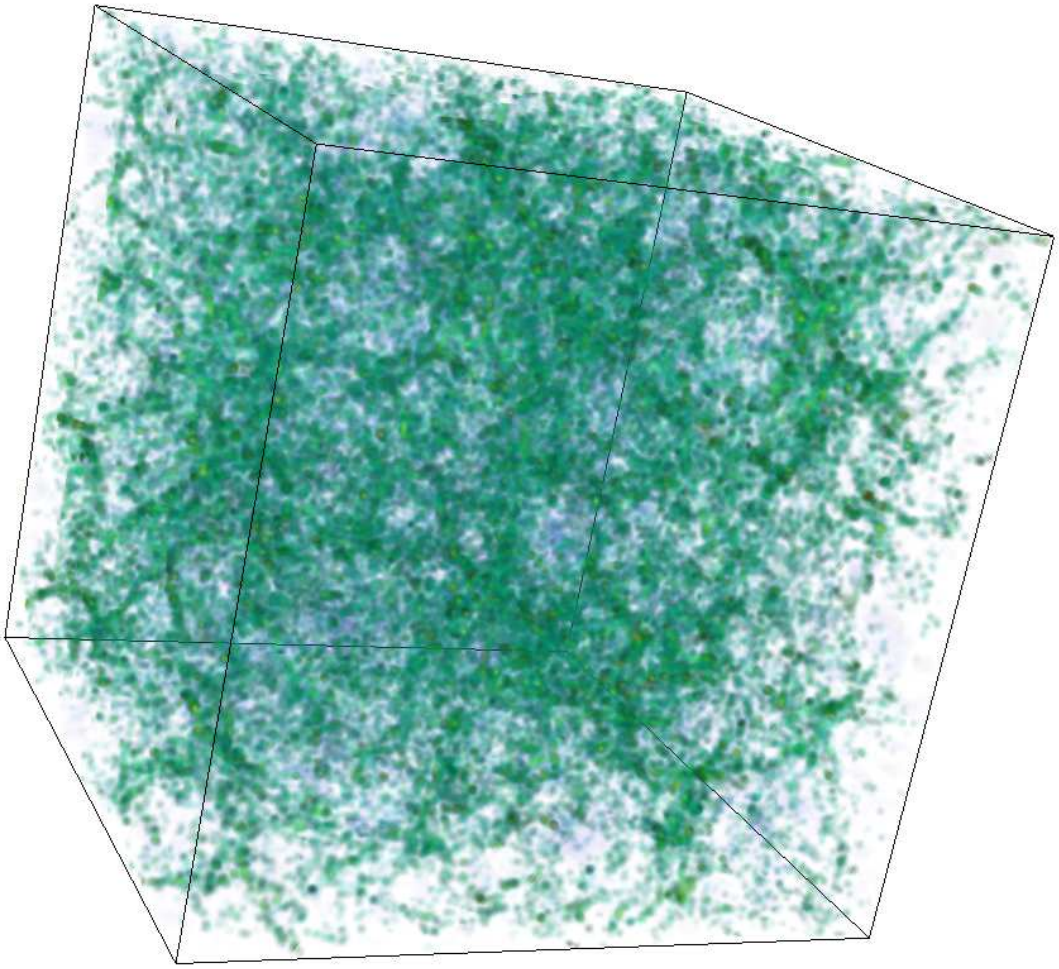
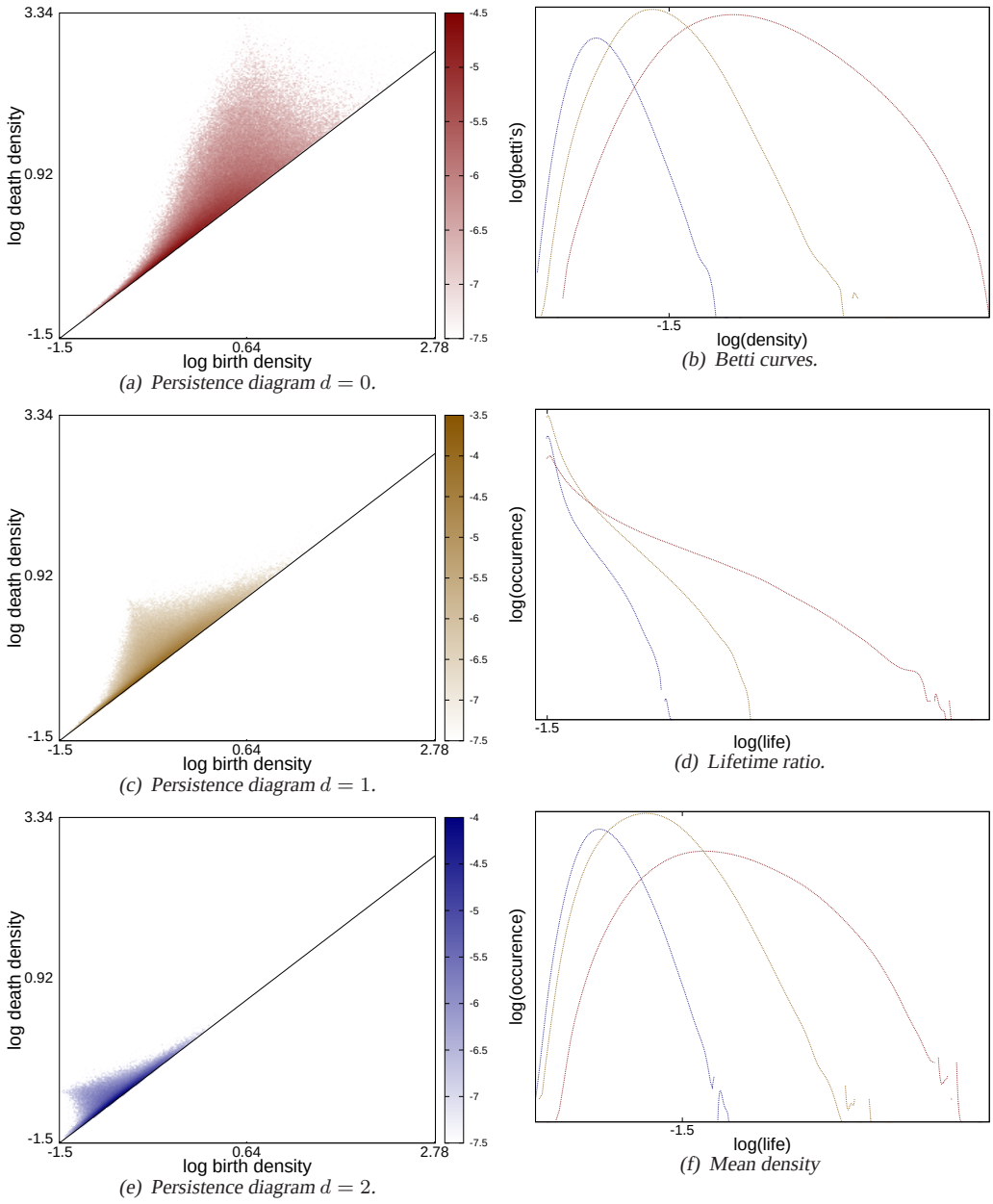


Figure 6.1: ***Density field of LCDM at*** $z = 0$ ***(run 15).***

with μ the mean and σ the standard deviation. It has both mathematically and physically nice properties: (i) for small fluctuations it approximates arbitrarily close a Gaussian distribution; (ii) the density remains always positive; and (iii) many nice properties of the Gaussian distribution can be computed analogously for the lognormal distribution. If the Betti curves and mean density curves are resembling a lognormal distribution, than their exponents follow a normal distribution. To check how strong the lognormality is, we fit both exponents with a normal distribution. The two panels of figure 6.3 show the Gaussian fits of the Betti curve and mean density curve exponents. Although the Gaussian fits agree quite well, they seem a bit skewed with respect to the Betti and mean density curves. This suggests to try a fit with a skewed normal distribution, defined as:

$$f(x|\mu,\sigma,\alpha) = 2\phi(x|\mu,\sigma)\int_{-\infty}^{\alpha x}\phi(t,|\mu,\sigma)dt \tag{6.2}$$

with $\phi(x|\mu,\sigma)$ the normal distribution and α the skewness parameter. The skewed normal distribution gives a visibly better fit, see figure 6.4. For the Betti curves the fit is even near exact.



(a) Persistence diagram $d = 0$.

(b) Betti curves.

(c) Persistence diagram $d = 1$.

(d) Lifetime ratio.

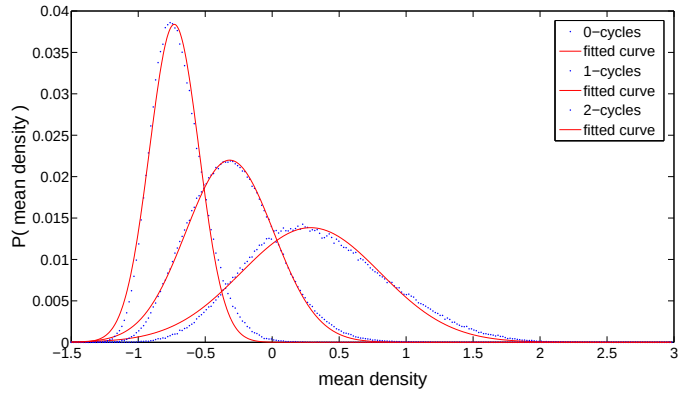
(e) Persistence diagram $d = 2$.

(f) Mean density

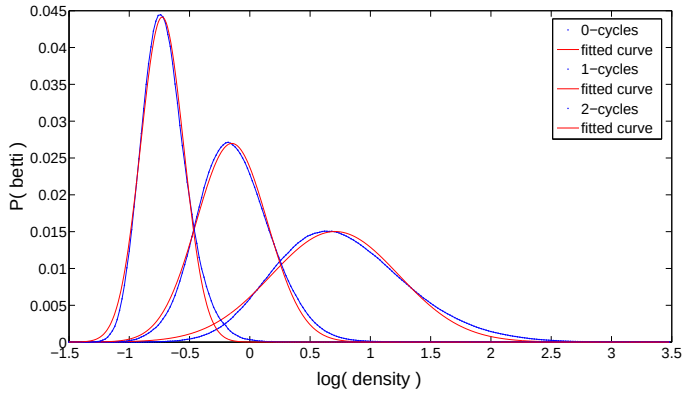
Figure 6.2: ***Persistence measures for LCDM at*** $z = 0$ ***(run 15).*** *Left: persistence diagrams for various dimensions. Right: 1D summaries of the previous. Colors indicate different cycle dimensions: red (0-cycles), yellow (1-cycles) and blue (2-cycles).*

The effects of simplification

Many features close to the birth=death diagonal might be Poisson noise and if not, they are low persistent features which are of limited relevance (cf. subsection 5.3.2). Often, density fields are smoothed to remove such low persistent features. The topological analogue of smoothing is simplification (cf. section 5.4). To investigate the effect of simplification on the density field, we simplify the manifold according to the prescription of subsection 5.3.2. The first step in the prescription is setting a significance value indicating the maximum probability that a feature is actually Poisson noise. Thresholds have been computed for 64^3 particle distributions but as the



(a) Gaussian fit of mean density curves.



(b) Gaussian fit of Betti curves.

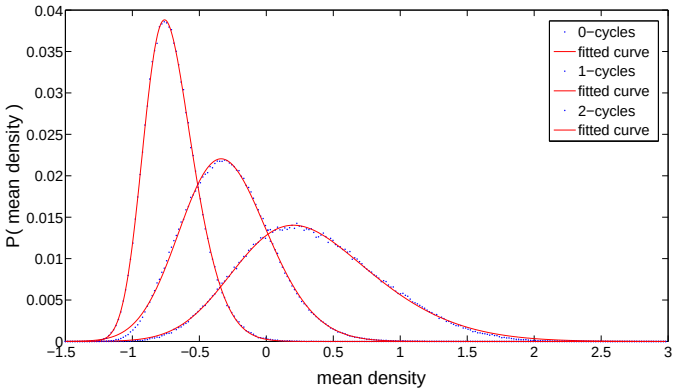
Figure 6.3: ***Gaussian fits to LCDM (run 15) at*** $z = 0$***.*** *In each diagram from right to left i.e. from high to low densities: 0-cycles, 1-cycles and 2-cycles.*

threshold depends on the amount of particles, we have to recompute the thresholds for the 256^3 case considered here. To do so, ten Poisson mock catalogues with $\lambda \sim 256^3$ are created. Their cumulative ratio distribution $F(r)$ and cumulative difference distribution $D(r)$ are computed. The thresholds for the significance values 0.1, 0.01 and 0.001 can be found in table 6.1 and $F(r)$ and $D(r)$ are plotted in figure 6.5. For comparison, both the Poisson mock catalogues and all LCDM runs at $z = 0$ are shown. The figure unambiguously shows that the density fields extend up to far larger densities than the Poisson mock catalogues and thus contain a lot more structure. The differences between the LCDM realizations themselves is minimal.

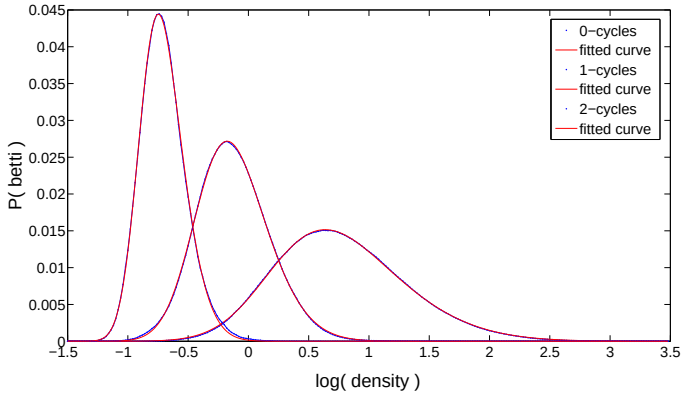
Sign. values	persistence ratio r			persistence difference d		
	0-cycles	1-cycles	2-cycles	0-cycles	1-cycles	2-cycles
0.1	0.212 ± 0.000	0.156 ± 0.000	0.142 ± 0.000	-0.072 ± 0.004	-0.460 ± 0.002	-0.726 ± 0.000
0.01	0.344 ± 0.000	0.252 ± 0.000	0.262 ± 0.000	0.202 ± 0.004	-0.250 ± 0.000	-0.508 ± 0.000
0.001	0.436 ± 0.002	0.288 ± 0.000	0.342 ± 0.002	0.354 ± 0.002	-0.194 ± 0.000	-0.429 ± 0.004

Table 6.1: ***Simplification threshold values.*** *Threshold values of persistence ratio r (left) and persistence difference d (right) for several cycle dimensions for 256^3 particles, density case.*

Figures 6.6 and 6.7 show how the persistence diagrams, Betti curves, lifetime curves and mean density curves evolve with increasing simplification threshold. Ratio simplification (cf.

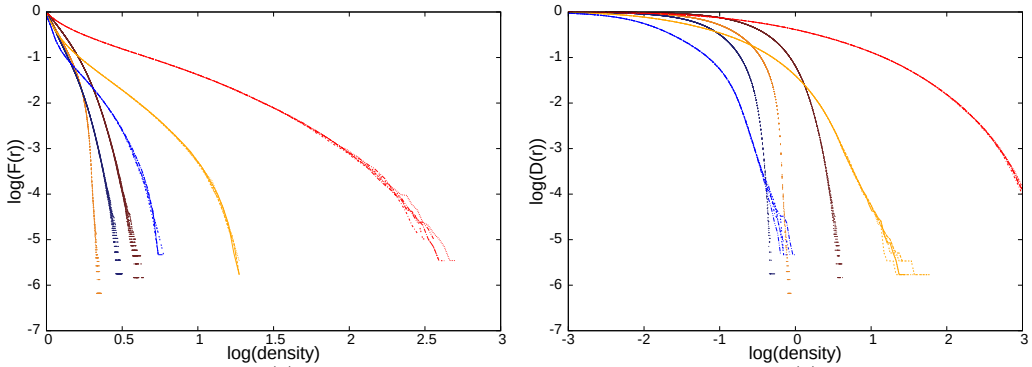


(a) Skewed Gaussian fit of mean density curves.



(b) Skewed Gaussian fit of Betti curves

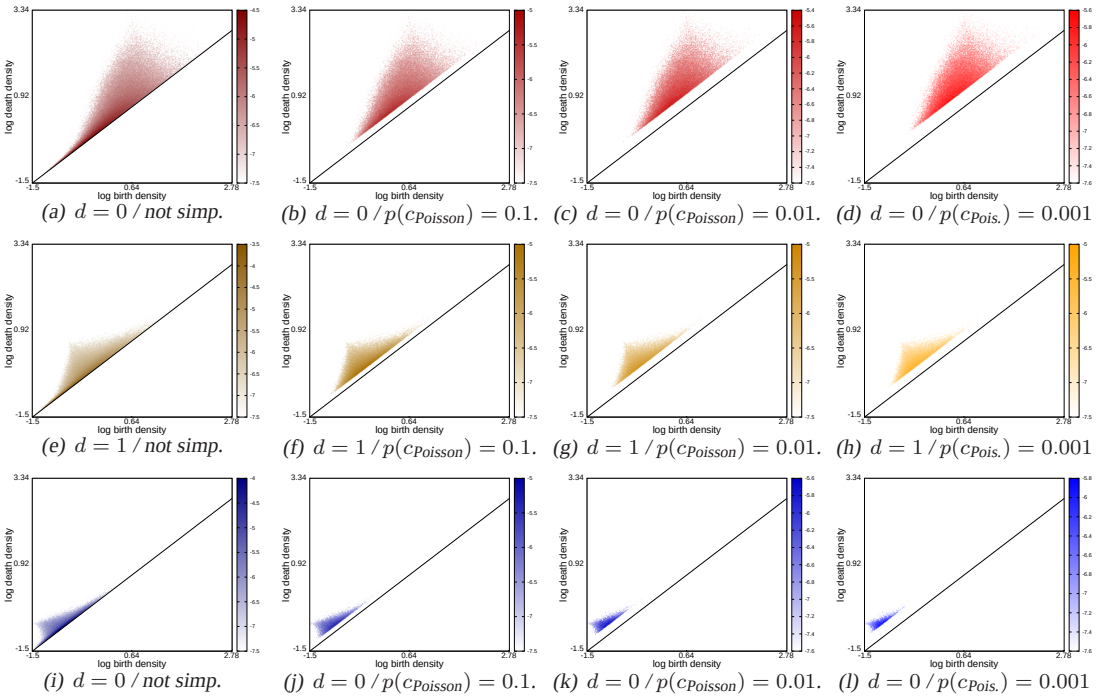
Figure 6.4: ***Skewed Gaussian fits to LCDM (run 15) at*** $z = 0$. *The skewed Gaussians are visibly a better fit than the ordinary Gaussians shown in figure 6.3. In each diagram from right to left i.e. from high to low densities: 0-cycles, 1-cycles and 2-cycles.*



(a) $F(r)$ *(b)* $D(r)$

Figure 6.5: ***Cumulative density distributions****. The graphs show the probability of the existence of a persistence pair with a ratio or difference equal or larger* r *respectively* d*. LCDM simulations at* $z = 0$*: red (0-cycles), yellow (1-cycles) and blue (2-cycles). The Poisson mock catalogues of* 256^3 *particles: dark red (0-cycles), orange (1-cycles) and dark blue (2-cycles). For all dimensions, the LCDM curves show significantly more persistent structure. The differences amongst the LCDM curves themselves is negligible.*

subsection 5.3.2) does exactly what is was meant to do: shifting the birth=death diagonal upwards. Due to the shape of their distributions, for β_0 this goes mainly at the expense of the low density tail; for β_1 the burden is spread relatively evenly between the low and high density tail

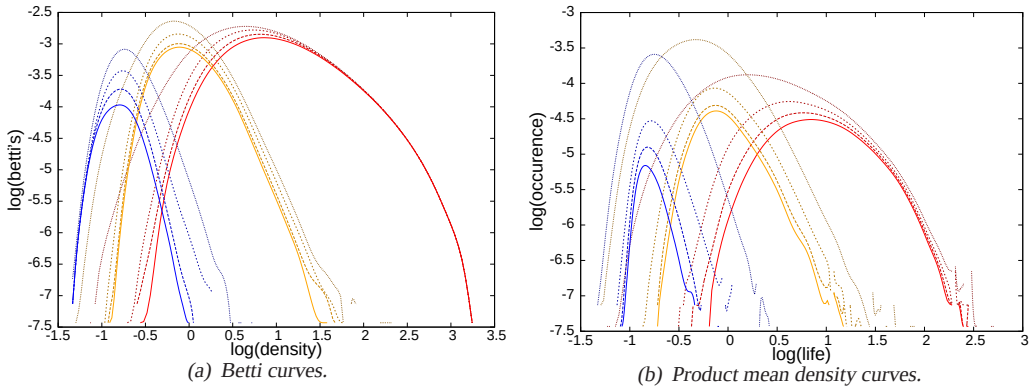


(a) $d = 0$ / *not simp.* (b) $d = 0$ / $p(c_{Poisson}) = 0.1$. (c) $d = 0$ / $p(c_{Poisson}) = 0.01$. (d) $d = 0$ / $p(c_{Pois.}) = 0.001$

(e) $d = 1$ / *not simp.* (f) $d = 1$ / $p(c_{Poisson}) = 0.1$. (g) $d = 1$ / $p(c_{Poisson}) = 0.01$. (h) $d = 1$ / $p(c_{Pois.}) = 0.001$

(i) $d = 0$ / *not simp.* (j) $d = 0$ / $p(c_{Poisson}) = 0.1$. (k) $d = 0$ / $p(c_{Poisson}) = 0.01$. (l) $d = 0$ / $p(c_{Pois.}) = 0.001$

Figure 6.6: ***The effect of ratio simplification on persistence diagrams.*** *Persistence diagrams for LCDM at $z = 0$ (run 15). From left to right: increasing simplification threshold. Here, $p(c_{Poisson}) = x$ indicates that the probability of a cycle to be generated by Poisson noise is equal or smaller than x. Top to bottom: 0-cycles, 1-cycles and 2-cycles.*



(a) *Betti curves.* (b) *Product mean density curves.*

Figure 6.7: ***The effect of ratio simplification on 1D persistence summary curves.*** *1D persistence summary curves for LCDM at $z = 0$ (run 15). The colors indicate different cycle dimensions: red (0-cycles), yellow (1-cycles) and blue (2-cycles). Brighter means higher simplification threshold, with the followings steps: (i) not simplified, (ii) $p(c_{Poisson}) = 0.1$, (iii) $p(c_{Poisson}) = 0.01$ and (iv) $p(c_{Poisson}) = 0.001$. Here, $p(c_{Poisson}) = x$ indicates that the probability of a cycle to be generated by Poisson noise is equal or smaller than x.*

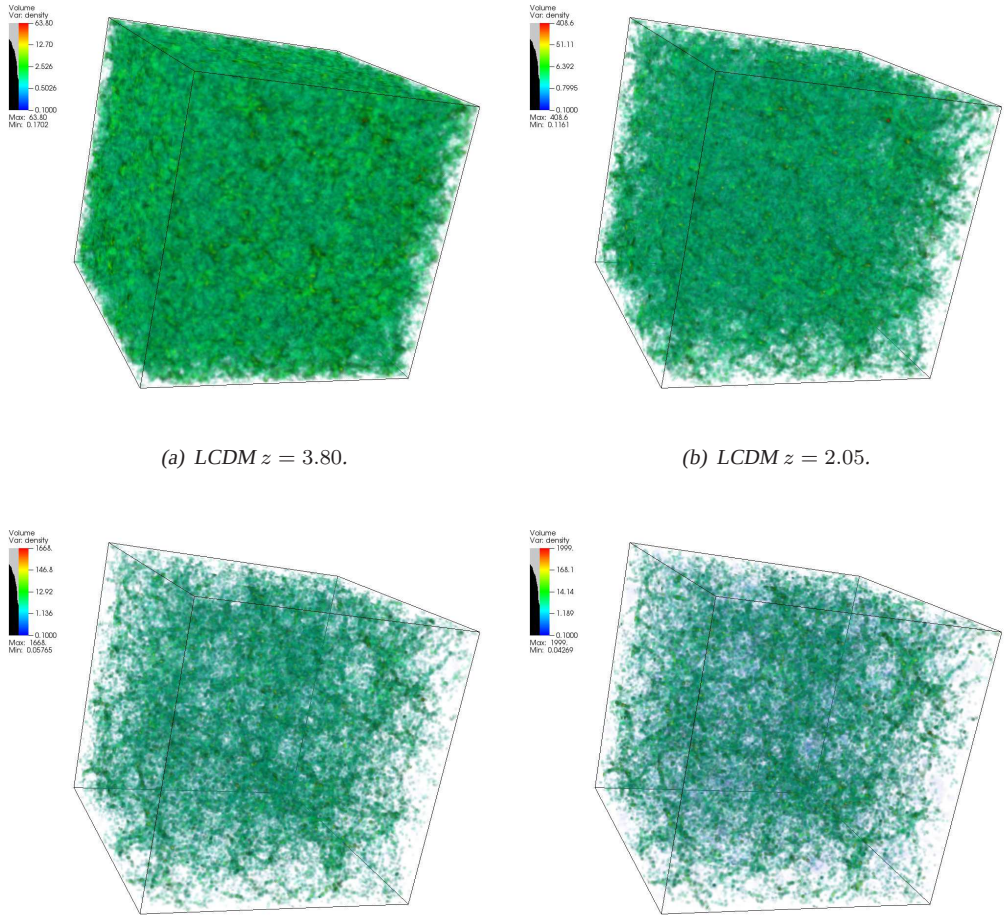
and; for β_2 the high density tail is severely diminished. Also the Betti curves behave as was expected from previous chapter: with increasing threshold they lower slightly and partly loose their tails. Similar effects can be seen in the mean density curves, where simplification has a much stronger effect. The direct effect of ratio simplification has a trivial effect on the lifetime curves (not shown here): the graphs get cut off at threshold level.

6.2 Time evolution

Cosmic structure evolves with time, so for different values of z we expect to see differences in the persistence measures. To investigate the influence of time evolution, we consider LCDM run 15 at eight redshifts: $z = 3.80$, $z = 2.98$, $z = 2.05$, $z = 1.00$, $z = 0.51$, $z = 0.25$, $z = 0.10$ and $z = 0.0$.

A visual impression

Figure 6.8 illustrates the evolution of the density field with z. The top-left tile shows the density field at $z = 3.80$. Although the seeds of cosmic structure are discernible, structures are diffuse and the density differences limited. The bottom-right tile showing $z = 0$ gives a completely different view. Pronounced structures with a large density difference are discernible. To see how this structure evolution is reflected in persistence diagrams, figure 6.9 shows how persistence diagrams and its 1D summary graphs evolve in time. A distinct time evolution is visible: with



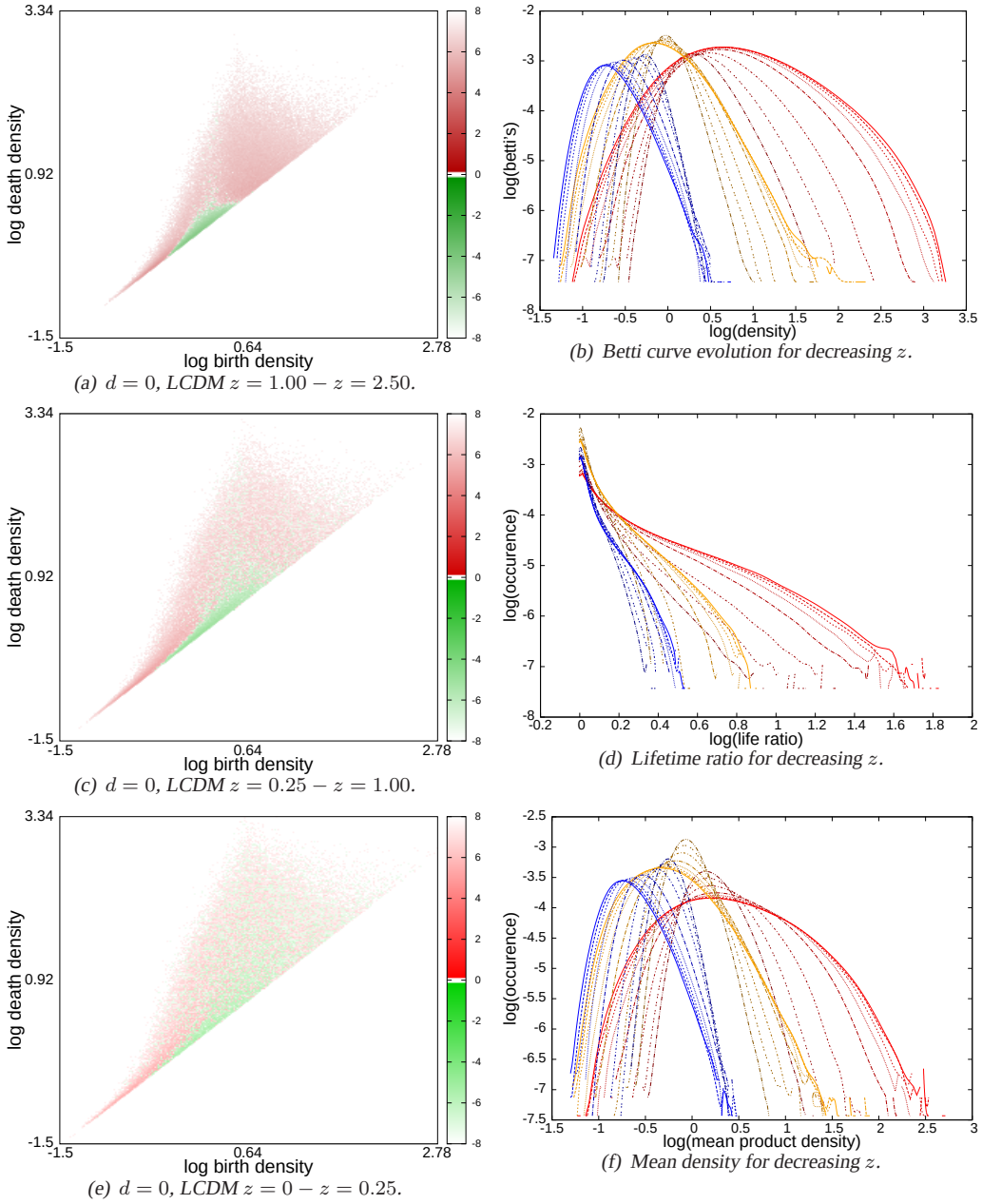
(a) LCDM $z = 3.80$. *(b) LCDM $z = 2.05$.*

(c) LCDM $z = 0.507$. *(d) LCDM $z = 0.00$.*

Figure 6.8: ***LCDM time evolution.*** *Density field of LCDM 15 for various z. Structure clearly evolves when z decreases.*

decreasing z the persistence diagrams fan out from close to the middle and near the diagonal to

all directions. The effect seems strongest at higher z but it could be the effect of time sampling. The 0-cycles mainly migrate upwards along the diagonal, contrary to 2-cycle which mainly move downwards; 1-cycles spread quite evenly upwards and downwards. Naturally, the spread is also reflected in the mean density curves and (to a lesser degree) in the Betti curves. The lifetime ratio increases in time to higher densities, quite strong for 0-cycles and only slightly for 2-cycles.



(a) $d = 0$, *LCDM* $z = 1.00 - z = 2.50$.

(b) Betti curve evolution for decreasing z.

(c) $d = 0$, *LCDM* $z = 0.25 - z = 1.00$.

(d) Lifetime ratio for decreasing z.

(e) $d = 0$, *LCDM* $z = 0 - z = 0.25$.

(f) Mean density for decreasing z.

Figure 6.9: ***Persistence diagram time evolution.*** *Persistence diagrams for LCDM at* $z = 0$ *(run 15). Left: 0-cycle time evolution, color indicates logarithmic excess of later time (red) or earlier time (green). Right: 1D persistence summaries with red (0-cycles), yellow (1-cycles) and blue (2-cycles). Brighter means lower* z.

Parametrization of 1D summary curves

Throughout time evolution a skewed lognormal remains an excellent fit for the mean density and Betti curves. Fits for the latter curves for various values of z are shown in figure 6.10. The plotted skewed Gaussians have 4 parameters: mean μ, standard deviation σ, the skewness

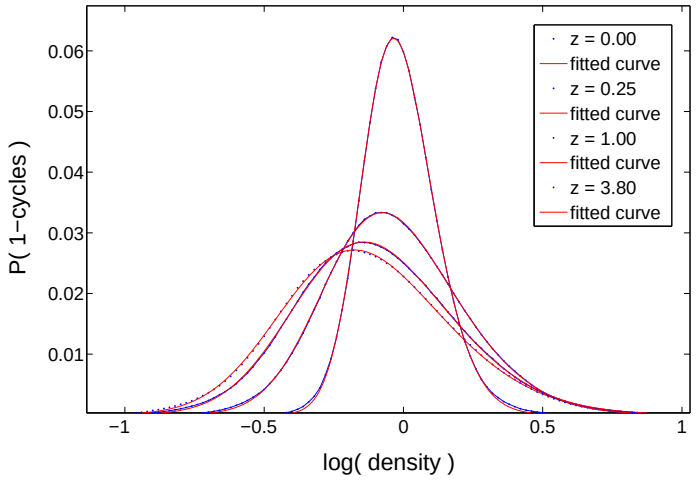


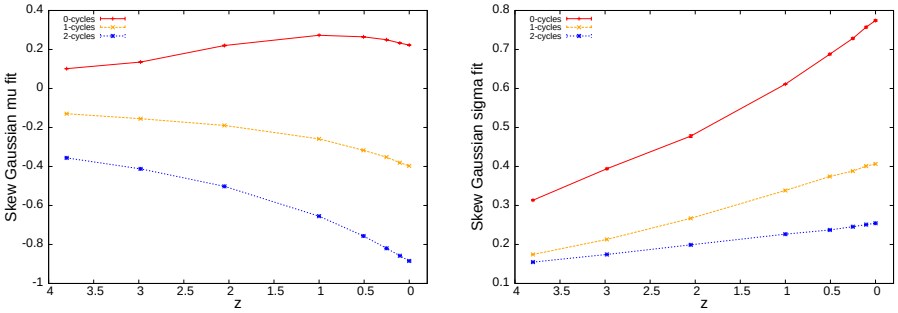
Figure 6.10: ***Skew Gaussian fits to Betti curves for various redshifts.*** *Above 1-cycles are shown, fits for 0-cycles and 2-cycles are analogous. The height of the curves descreases with decreasing z. Note the excellent fit for all z and all dimensions.*

parameter α and the scaling parameter c. (The last is a pre-multiplication factor to correct for bin size.) With time the skewed Gaussians evolve and these parameters change. To assess whether there is any trend in the evolution of these parameters figure 6.11 shows the evolution of the best-fit parameters in time. For μ and σ a trent corresponding with observations above is evident. With time σ increases for all dimensions, whereas μ increases for 0-cycles, decreases for 2-cycles and remains about the same for 1-cycles. More interesting perhaps is the behaviour of the skewness parameter α, which surprisingly doesn't increase with time but shows rather eccentric behaviour. As for all z the number of bins is the same, c doesn't show any evolution, as expected.

The effects of simplification

Ratio simplification might influence the persistence diagrams differently for various epochs. As ratio simplification is an important tool to nice the manifold, it is important to understand if this is indeed the case and by how much. Figure 6.12 shows Betti curve time evolution for increasing simplification threshold, with brighter color indicating lower z. If z goes up the Betti curve peaks go down, an effect that becomes stronger with increasing simplification threshold. For the highest simplification threshold (bottom right panel) the difference between the lowest and highest z peak has become almost an order of magnitude. This suggests that at lower z structure is less pronounced, as it simplifies away easier.

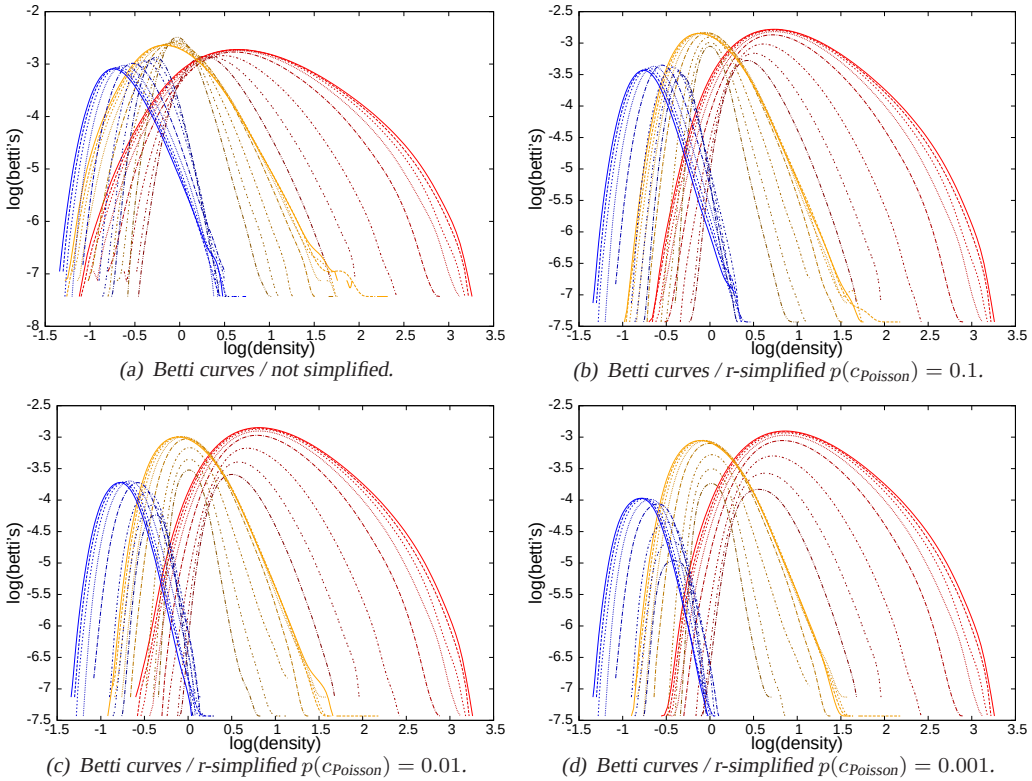
Note also that in the unsimplified case the Betti curve peaks of all dimensions have about the same magnitude. On the contrary, at the highest simplification level considered here the 2-cycle peak decreased about an order of magnitude with respect to the 0-cycle peak. This can be explained by the lower range of densities of 2-cycles, which makes their persistence relatively low as well. The mean density curves and lifetime curves (not displayed) show the same pattern.



(a) Mean μ. *(b) Standard deviation σ.*

(c) Skewness parameter α. *(d) Scaling parameter $c/2$.*

Figure 6.11: ***Parameter time evolution of skewed normal fits to Betti curves.*** *All four parameters are shown. The error lines correspond with the 95% confidence intervals.*



(a) Betti curves / not simplified. *(b) Betti curves / r-simplified $p(c_{Poisson}) = 0.1$.*

(c) Betti curves / r-simplified $p(c_{Poisson}) = 0.01$. *(d) Betti curves / r-simplified $p(c_{Poisson}) = 0.001$.*

Figure 6.12: ***Betti curve time evolution under increasing simplification threshold.*** *Colors indicate different cycle dimensions: red (0-cycles), yellow (1-cycles) and blue (2-cycles). Brighter means lower z and $p(c_{Poisson}) = x$ indicates that the probability of a cycle to be generated by Poisson noise is equal or smaller than x.*

6.3 Summary and conclusions

First, as reference case LCDM (run 15) at $z = 0$ is described in detail:

1. The 1-cycle persistence diagram has an extended inverse v-shape along the birth=death diagonal. The 0-cycle diagram fans out towards higher births and higher deaths, whereas the 2-cycle diagram goes in opposite direction.

2. The Betti curves and product mean density curves follow nearly exact a skewed log-normal pattern. The lifetime curves are monotonically decreasing, especially the 1-cycle and 2-cycle curves decrease quite fast.

3. Ratio simplification effects the persistence diagrams as expected by 'pushing up' the diagonal. Consequently, the lifetime curves abruptly loose their lowest densities. 1-cycle Betti and mean density curves mainly shift downwards. Their 0-cycle and 2-cycle variants also loose their low respectively high density tail.

With time structure evolves. For LCDM run 15, the persistence diagrams of all available epochs (from $z = 3.8$ to $z = 0$) are computed and compared, showing us:

1. In time persistence diagrams fan out. 0-cycles mainly migrate to higher densities, 2-cycles mainly to lower densities and 1-cycles spread quite evenly in both directions. This spread also clearly shows from the mean density curves and to a lesser degree from the Betti curves, in the form of shifting μ values and increasing σ. Interestingly, no trend seems to be visible in the skewness parameter α. In time, the lifetime ratio increases to higher densities, quite strongly for 0-cycles and only slightly for 2-cycles.

2. Simplification enhances the effect of time evolution, as earlier epochs are thinned out much more than later epochs. This suggests that at earlier epochs structures are less persistent.

We conclude:

1. Persistence diagrams are a great tool to describe structure evolution: they are very sensitive to structure and structural evolution.

2. The skewed normal distribution is an excellent model for Betti curves of cosmological matter distributions. μ and σ have clear patterns, the skewness parameter α not.

CHAPTER 7

Topological dark energy differentiation

Dark energy determines the global evolution of the universe. (Dark) matter is more powerful on smaller scales and runs local affairs. As such it is dark matter that determines the detailed shape of cosmic structure. But perhaps the tentacles of dark energy run deeper than we think and it mingles in local affairs as well. If so, it will not leave the shape of cosmic structure untouched. Different flavors of dark energy might leave a characteristic imprint on cosmic structure. In the future, such an imprint could than be used to put constraints on dark energy models. The previous chapter gave an extensive overview of the shape of LCDM persistence diagrams and how they are influenced by simplification and time evolution. Based on the intuition gained there, in this chapter we will or different DE models leave a different imprint on the shape of the cosmic web.

The three dark energy models considered here are the three models explored in section 2.4: LCDM, RP and SUGRA. For each dark energy model five realizations are available, named run 14 to run 18. Each dark energy model is considered at eight redshifts: $z = 3.80$, $z = 2.98$, $z = 2.05$, $z = 1.00$, $z = 0.51$, $z = 0.25$, $z = 0.10$ and $z = 0.0$. Combining previous gives 3 DE models $\times$ 5 realizations $\times$ 8 redshifts is 120 files. This large amount of data precludes investigation of each individual file. Therefore we apply a different strategy and compare all results with the reference case LCDM (run 15) at $z = 0$ (cf. section 6).

Three questions will lead us throughout this chapter:

1. Using persistence, can we distinguish between various dark energy models at the same z in real space?

2. If the answer to the question of the previous point is positive, do the differences remain visible if we transform to readshift space? If so, we might be able to use persistence to *observationally* distinguish between various dark energy models. In the future, this could lead to topology induced constraints on dark energy models.

3. Using persistence, can we distinguish in real space between various dark energy models at the same structure clustering parameter σ_8 (cf. equation 2.33)? If so, this suggests DE has a 'local' effect on structure formation. Various DE models will result in structure with *intrinsically* different topology.

The first and third question both consier structure in real space. Their essential lies in the fact that for different DE models structure evolves differently in time, i.e. the function $\sigma_8(z)$ depends on DE model: $\sigma_8(z) = \sigma_8(z, \mathrm{DE})$. Therefore, topological differences between two models at some z might be the combination of two separate effects: (i) different rates of structure evolution and (ii) intrinsic topological differences between the structure.

Of course, considering z and σ_8 effects separately is only useful if $\sigma_8(z)$ differs notably between various models. To see whether this is the case, $\sigma_8(z, \mathrm{DE})$ w.r.t. $\sigma_8(z, \mathrm{LCDM})$ is shown for the three dark energy models in figure 7.1. The figure clearly shows that for smaller z ($z \lesssim 1$) the models differ from each other significantly, making a separate treatment of the z and σ_8 effects necessary. Furthermore, we observe that for high z SUGRA evolves fastest, but it is quickly overtaken by RP. Later LCDM takes over from both models and becomes the fastest evolving model. We note that for larger z the σ_8 differences between the models become smaller. This is understandable as at z_{CMB} structure is identical for all models (cf. subsection 4.1.3).

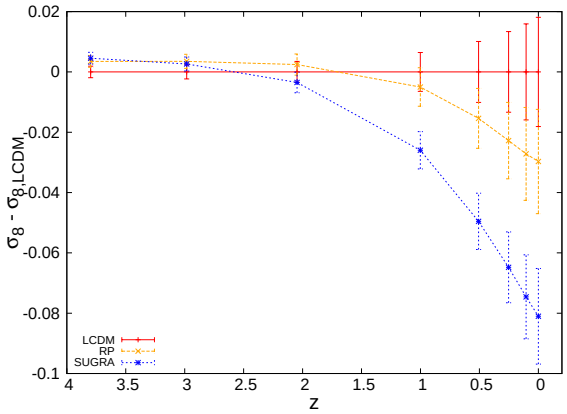


Figure 7.1: $\sigma_8(z, DE)$ ***with respect to*** $\sigma_8(z, LCDM)$ ***for the three dark energy models considered here.*** *The values are averaged over the five realizations, with the error bars indicating the variance.*

For each of the three questions stated above, the different DE models will be compared in the following way:

1. a visual comparison of the similarities and differences is given;
2. the effects of simplification, i.e. topological nicing of the density field (cf. section 5.3), are considered;
3. we describe the statistical test setup used to determine whether differences (if any) are statistically signicant;
4. the results from the statistical tests are presented.

An overall summary is given in section 7.4 and a discussion and conclusions are presented in section 7.5.

7.1 Real space identical redshift

For the three dark energy models under consideration snapshots of the matter distribution in real space at the same z are directly available (cf. subsection 4.1.3). In subsection 7.1.1 we

inspect their similarities and differences visually, followed by an investigation of the effect of simplification in subsection 7.1.2. The statistical test procedure is explained in 7.1.3 and the test results are presented in subsection 7.1.4.

7.1.1 Visual inspection

Plotting the persistence diagrams of various models and runs directly is not very illuminating, as the differences turn out to be small. Instead, we use persistence difference diagrams (cf. section 3.6): the bin-wise difference of a diagram with another diagram (the reference case). Persistence difference diagrams are a great tool to investigate the pattern of differences. For several SUGRA models at $z = 0$, their persistence difference diagrams are shown on the left of figure 7.2. Within the persistence difference diagrams red indicates an excess of the SUGRA model and green of the LCDM reference case. As we see from the figure, for all dimensions LCDM has clearly more low-birth low-death cycles. Especially for 2-cycles this effect is consistent across several runs, cf. the bottom three persistence difference diagrams.

The persistence difference diagrams show the difference pattern, but just as important is the magnitude of the difference. This is better shown using the 1D summary curves (in particular the Betti and mean density curves), shown on right of figure 7.2. They portray that the difference between LCDM and SUGRA might have a pattern, but in general is very small compared to the function values. Only careful examination shows a systematic shift of LCDM towards lower densities at the low density part of the curves. Furthermore, for 0-cycles LCDM seems shifted towards higher densities at the highest tail end. SUGRA has an opposite behaviour: at lower densities it is slightly shifted to higher densities whereas the high density 0-cycle tail it is shifted slightly to lower densities. RP lies nicely in between both models.

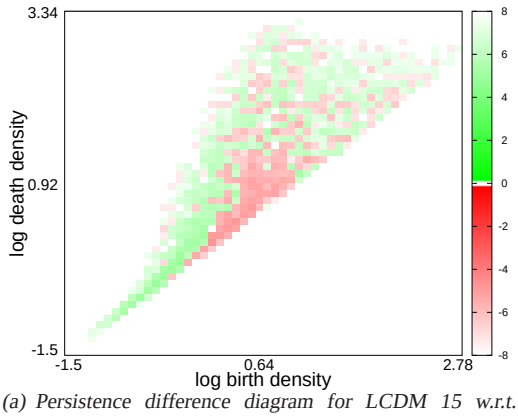
The relative shift of LCDM to lower densities might be attributable to a different stage of structure evolution. At $z = 0$ structure is most evolved in LCDM and we saw in section 6.1 that structure evolution causes persistence diagrams to fan out, exactly what we observe. Physically we can interpret it as follows: structure evolution makes empty regions emptier and the most dense peaks much denser. This explains both the shift of the low density LCDM tail to even lower densities and the shift of its high density tail in the opposite direction.

7.1.2 The effects of simplification

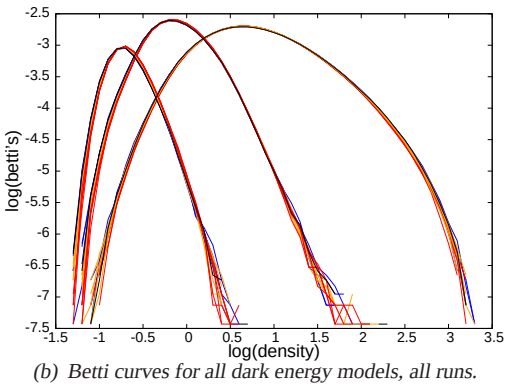
Simplification of the manifold supresses noise or small features, thus if the differences described above are mostly small and insignificant simplification will take them out. By doing so simplification might enlarge the effects of more persistent features. To see how the observed differences abide under simplification, figure 7.3 shows the Betti curves with increasing simplification threshold. The differences observed in the non-simplified case seem robust: they neither become more nor less prominent.

7.1.3 Statistical analysis / test setup

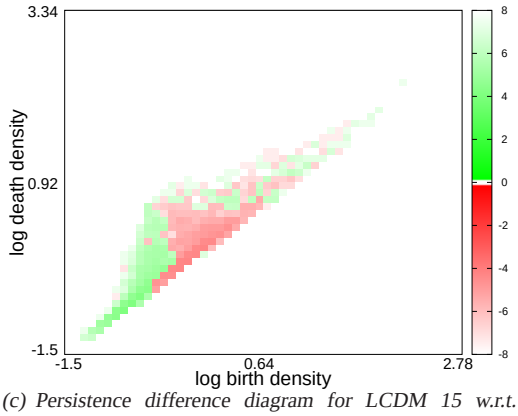
To test whether the small but apparently systematic differences between the various DE models are significant, we perform a two-tailed ***Kolmogorov-Smirnov (KS) two-sample test*** on each pair of dark energy models and runs. The KS test is chosen because it is non-parametric and sensitive to shape as well as location. For completeness we note that the Anderson-Darling test, more or less a weighted L^2-norm variant of the Kolmogorov-Smirnov test, is more sensitive to local differences [Feigelson and G., 2012]. However, as the KS test has enough statistical



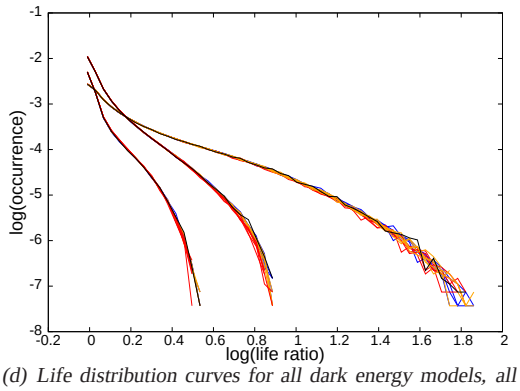
(a) Persistence difference diagram for LCDM 15 w.r.t. SUGRA 16, 0-cycles.



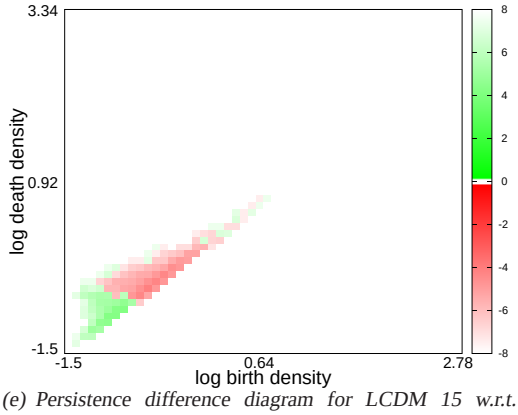
(b) Betti curves for all dark energy models, all runs.



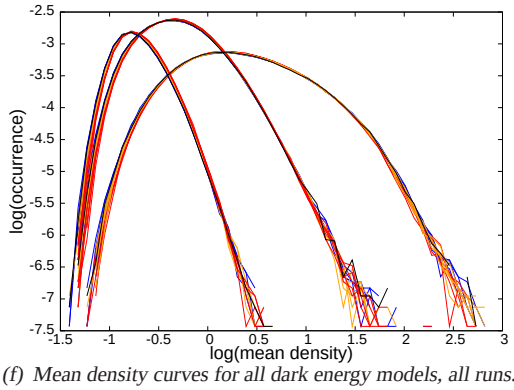
(c) Persistence difference diagram for LCDM 15 w.r.t. SUGRA 16, 1-cycles.



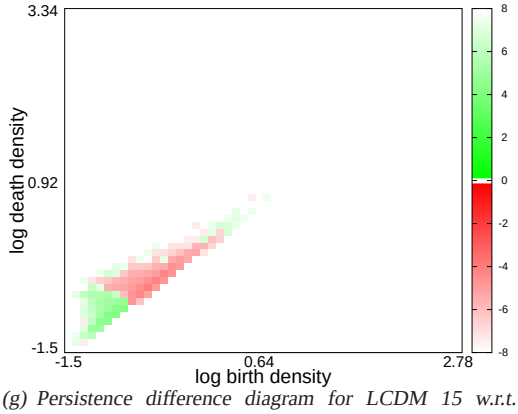
(d) Life distribution curves for all dark energy models, all runs.



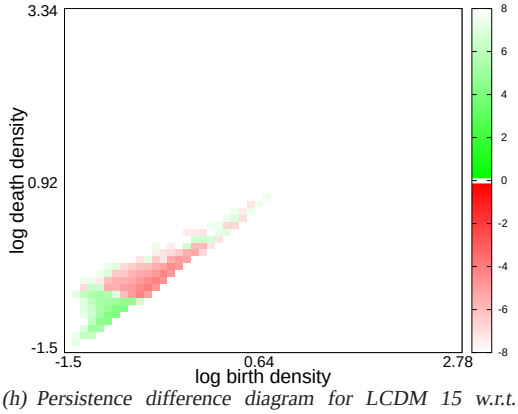
(e) Persistence difference diagram for LCDM 15 w.r.t. SUGRA 16, 2-cycles.



(f) Mean density curves for all dark energy models, all runs.

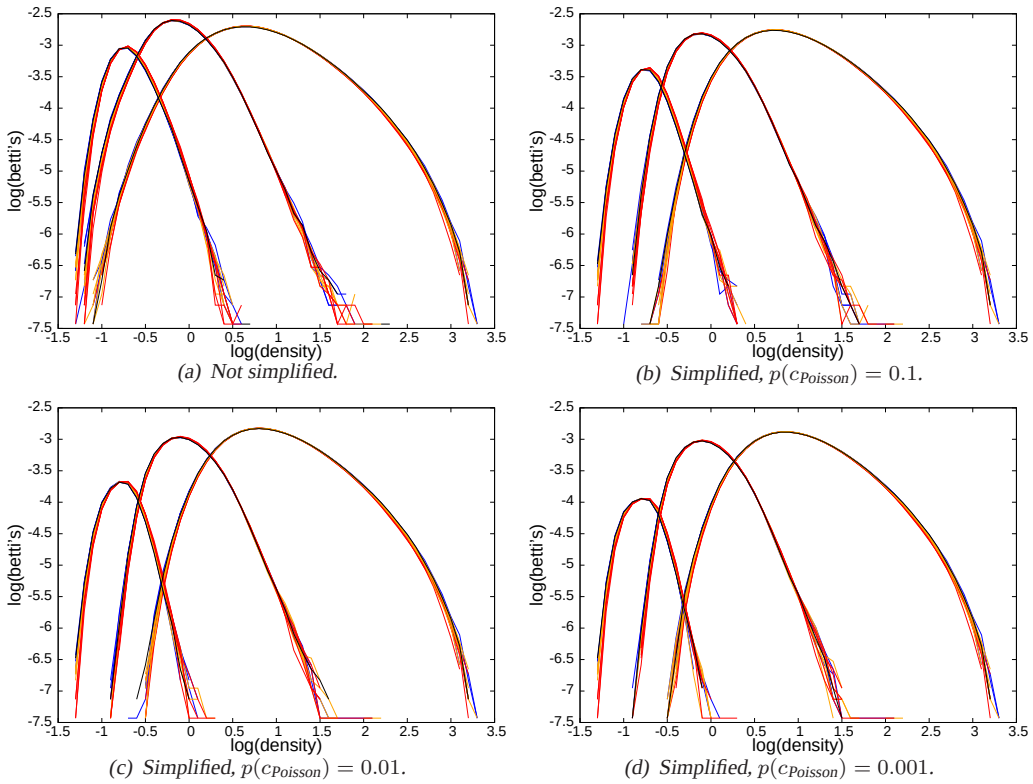


(g) Persistence difference diagram for LCDM 15 w.r.t. SUGRA 17, 2-cycles.



(h) Persistence difference diagram for LCDM 15 w.r.t. SUGRA 18, 2-cycles.

Figure 7.2: ***Comparison of persistence diagrams at*** $z = 0$**.** *Left: representative persistence difference diagrams, with green an excess of the reference LCDM run 15 and red an excess of another dark energy model. Stronger colors means larger excess, the numbers next to the color bar indicate the (negative) power of the excess, so note that the green color bar is inverted. Right: 1D summary curves for LCDM (blue and LCDM run 15 black), RP (yellow) and SUGRA (red). As can be seen from the 1D summary curves presented earlier: cycle dimension decreases from left to right curve set.*



(a) *Not simplified.* (b) *Simplified,* $p(c_{Poisson}) = 0.1$. (c) *Simplified,* $p(c_{Poisson}) = 0.01$. (d) *Simplified,* $p(c_{Poisson}) = 0.001$.

Figure 7.3: ***Betti curves at*** $z = 0$ ***with increasing simplification threshold.*** *Betti curves for LCDM (blue and LCDM run 15 black), RP (yellow) and SUGRA (red), with increasing simplification threshold.* $p(c_{Poisson}) = x$ *indicates that the probability of a cycle to be generated by Poisson noise is equal or smaller than* x*. As can be seen from the 1D summary curves presented earlier: cycle dimension decreases from left to right curve set.*

power for our purposes and is much better known within the astronomical community, we use it here. As zero hypothesis we take that each DE pair has the same underlying distribution and as alternative hypothesis we take the opposite.

Unfortunately, the KS test can't work with binned data as it is based on creating an *empirical distribution function*. An empirical distribution function requires an ordering of data points based on their function value. The points we consider here are the points in a persistence diagram. Each of these points has a lifetime and mean density, thus we can test lifetime and mean density in the KS test. In other words: we can test the difference between lifetime curves or mean density curves of various DE models, as both curves represent a summary of individual points properties. Betti curves, however, do not. A Betti curve gives the number of cycles alive at a certain density threshold, but this is a global property of the diagram and does not relate directly back to individual points. Individual points don't have a 'Betti value' associated to them (other then their dimension) and thus cannot be ordered based on their 'Betti value'. Consequently, on Betti curves no KS test can be performed.

It would be interesting to continue with a Kruskal-Wallis (KW) test [1] to identify any group differences between the DE models. Unfortunately the KW tests requires all the distributions to be identically shaped and scaled, a requirement not satisfied here. Therefore, unfortunately, we cannot proceed in this direction.

[1] The Kruskal-Wallis test is a non-parametric variant of ANOVA.

7.1.4 Statistical analysis / results

We test the significance of the difference between each pair of DE models and runs at $z = 0$ using (i) lifetime curves and (ii) mean density curves. For both, we give (i) the test outcomes (H0 acceptation or rejection) for a p-value of 0.01 and (ii) the more nuanced test statistic values.

Testing lifetime curves

A visual representation of the KS test outcomes is shown in figure 7.4. The cycle dimension increases from left to right and the simplification threshold increases from top to bottom. Rejection of H0 (yellow) means the lifetime curves originate from significantly different models, where non-rejection of H0 (black) means they do not.

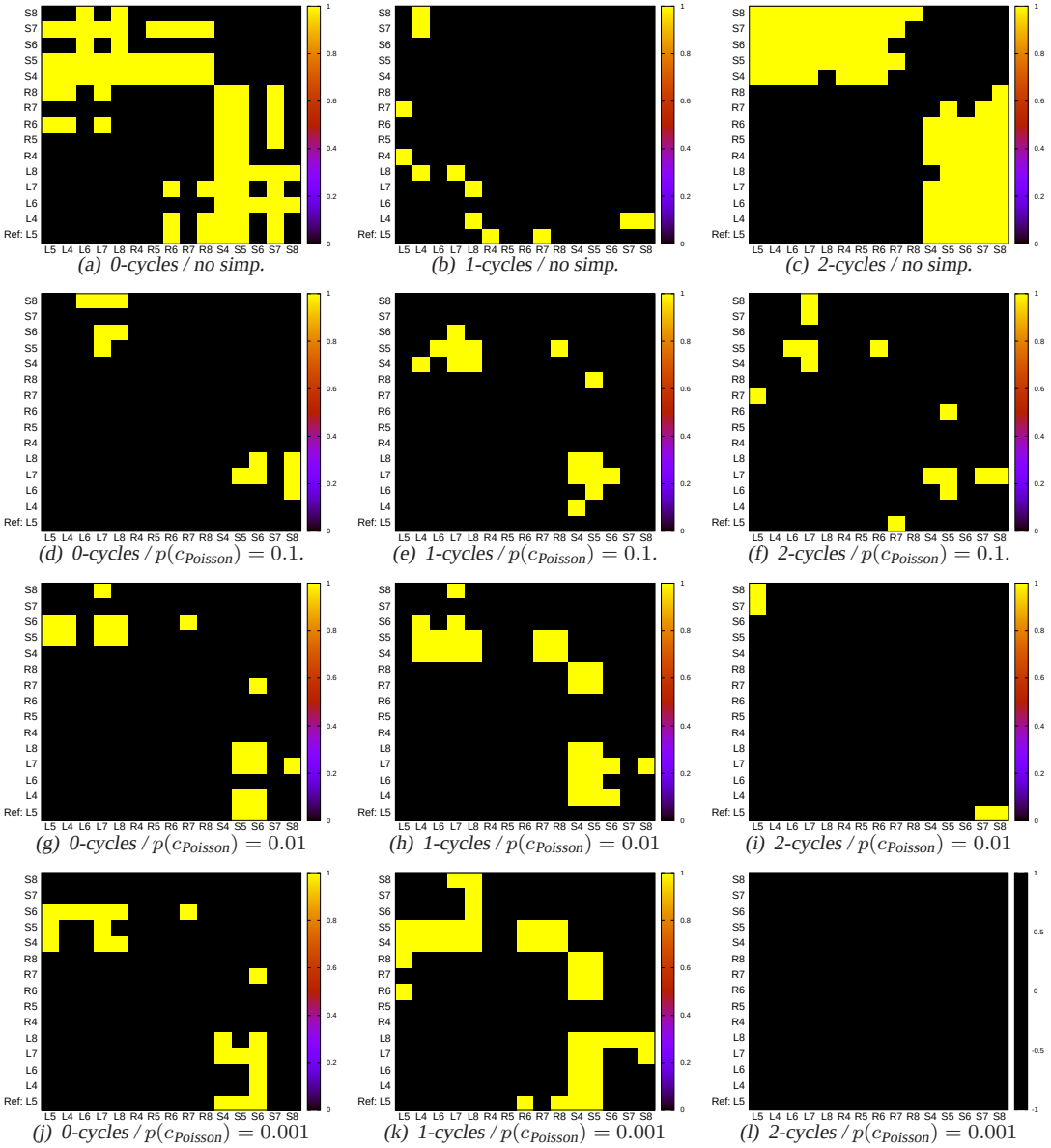
Let's first consider the unsimplified case, i.e. the top row. For 0-cycles and in particularly for 2-cycles a clear pattern is visible: in general SUGRA differs significantly from LCDM as well as from RP. A few RP runs also differ significantly from LCDM runs. Different runs of the same model are never significant. Interestingly, the 1-cycles shown another pattern: only a few significant results are found, mainly between different runs of the same model.

The statistical significance changes in a non-trivial way with increasing simplification threshold. For 2-cycles, the statistical significance of the pattern decreases with increasing simplification threshold, whilst for 1-cycles the statistical significance of the pattern increases. For 0-cycles the effects of simplification are more ambiguous.

More nuanced outcomes than acceptation/rejection are provided by the test statistic values itself, shown in figure 7.5. The results are analogous to those of the test outcomes, but patterns are more easily found.

Testing mean density curves

We repeat the analysis above for the mean density curves, the results are shown in figure 7.6 for H0 acceptation/rejection and in figure 7.7 for the corresponding test statistic values. Results reveal again an unmistakable and significant difference between the various dark energy models, even stronger than the lifetime curves did. Without simplification, almost all runs differ significantly from each other, also those of the same underlying DE model. With rising simplification threshold the contrast between the various DE models increases. Simplification levels of $p(c_{\text{Poisson}}) = 0.1$ and $p(c_{\text{Poisson}}) = 0.01$ (middle rows) show a near perfect model differentation. The contrast between the various DE models also strenghtens with increasing dimension (left to right).



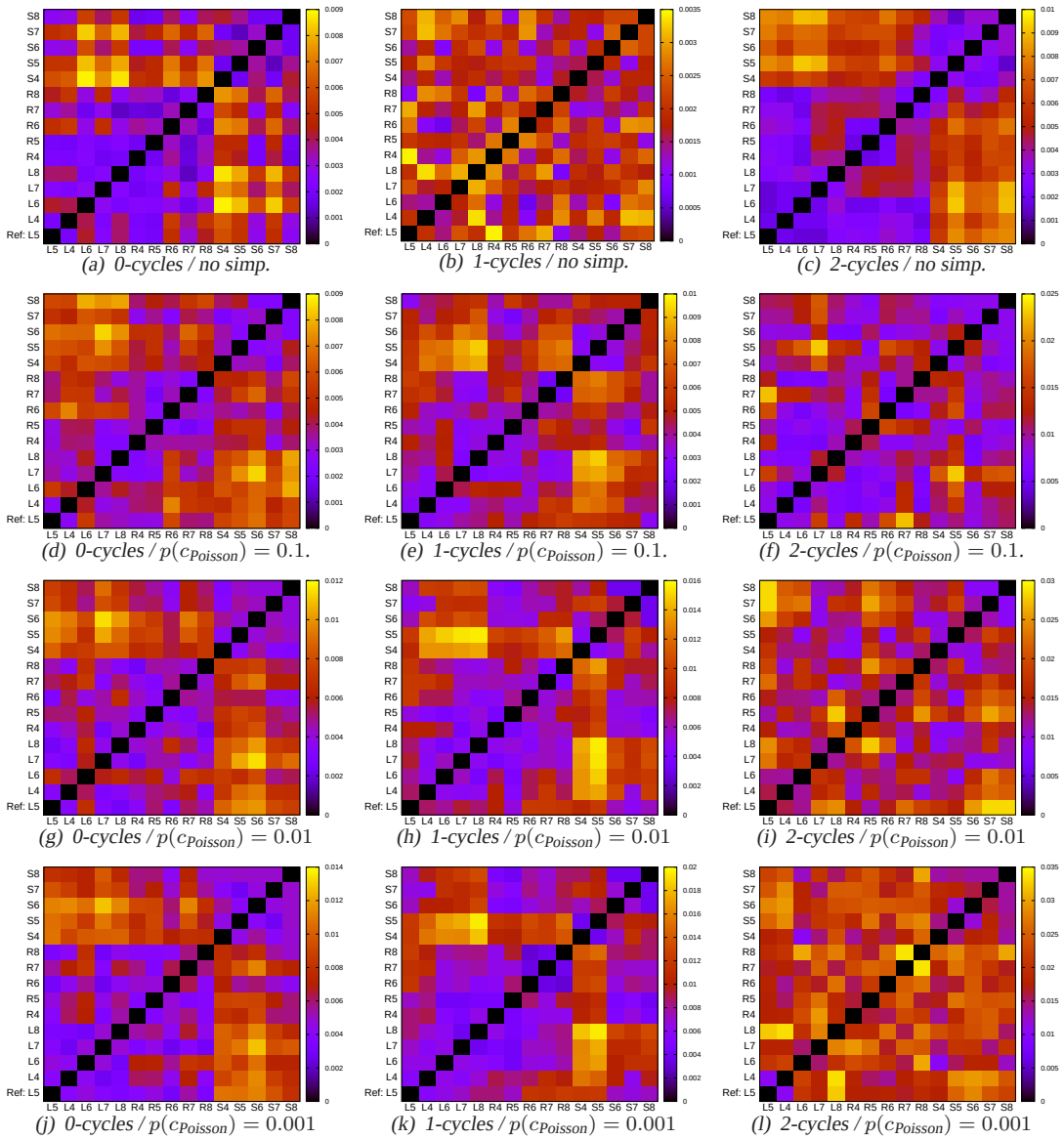
(a) 0-cycles / no simp. *(b) 1-cycles / no simp.* *(c) 2-cycles / no simp.*

(d) 0-cycles / $p(c_{Poisson}) = 0.1$. *(e) 1-cycles / $p(c_{Poisson}) = 0.1$.* *(f) 2-cycles / $p(c_{Poisson}) = 0.1$.*

(g) 0-cycles / $p(c_{Poisson}) = 0.01$ *(h) 1-cycles / $p(c_{Poisson}) = 0.01$* *(i) 2-cycles / $p(c_{Poisson}) = 0.01$*

(j) 0-cycles / $p(c_{Poisson}) = 0.001$ *(k) 1-cycles / $p(c_{Poisson}) = 0.001$* *(l) 2-cycles / $p(c_{Poisson}) = 0.001$*

Figure 7.4: ***Two-tailed Kolmogorov-Smirnov two-sample test outcomes for lifetime curves / real space identical z.*** *For each pair of dark energy models and runs we test whether the curves at $z = 0$ have statistically different persistence diagrams and thus topology. If a pattern between the various DE models is visible, it suggests DE models result in different topological structure at $z = 0$. The zero hypothesis states the underlying topology is identical, the alternative hypothesis that they are different. A p-value of 0.01 is used. From left to right: increasing cycle dimension. From top to bottom: increasing simplification threshold. $p(c_{Poisson}) = x$ indicates that the probability of a cycle to be generated by Poisson noise is equal or smaller than x. Table indicators: Ln, Rn and Sn stand for LCDM respectively RP or SUGRA run n+10.*



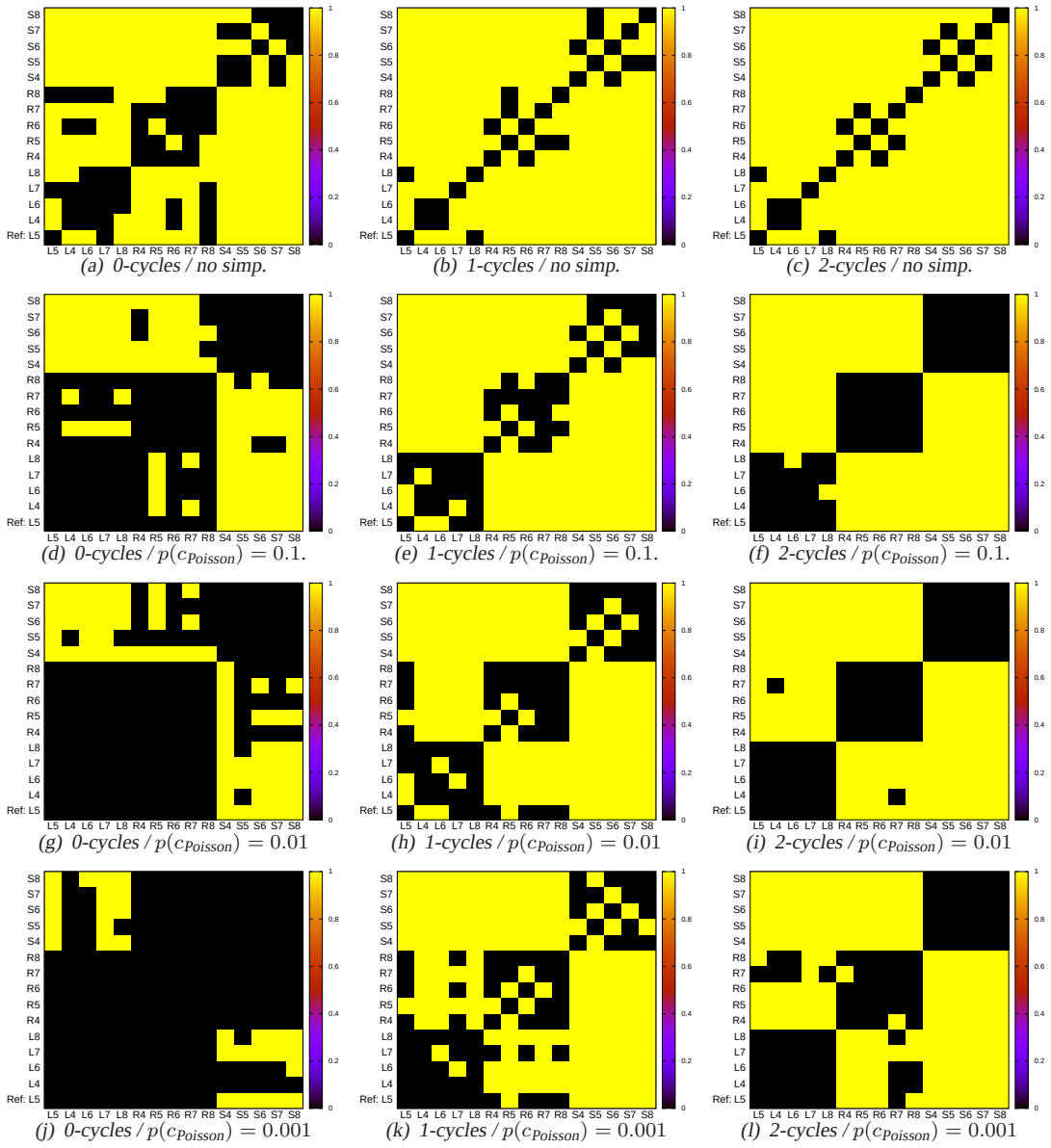
(a) 0-cycles / no simp. *(b) 1-cycles / no simp.* *(c) 2-cycles / no simp.*

(d) 0-cycles / $p(c_{Poisson}) = 0.1$. *(e) 1-cycles / $p(c_{Poisson}) = 0.1$.* *(f) 2-cycles / $p(c_{Poisson}) = 0.1$.*

(g) 0-cycles / $p(c_{Poisson}) = 0.01$ *(h) 1-cycles / $p(c_{Poisson}) = 0.01$* *(i) 2-cycles / $p(c_{Poisson}) = 0.01$*

(j) 0-cycles / $p(c_{Poisson}) = 0.001$ *(k) 1-cycles / $p(c_{Poisson}) = 0.001$* *(l) 2-cycles / $p(c_{Poisson}) = 0.001$*

Figure 7.5: ***Two-tailed Kolmogorov-Smirnov two-sample test statistic values for lifetime curves / real space identical z.*** *The test setup is analogous to that described in the text or in the caption of figure 7.4. From left to right: increasing cycle dimension. From top to bottom: increasing simplification threshold. Table indicators: Ln, Rn and Sn stand for LCDM respectively RP or SUGRA run n+10.*



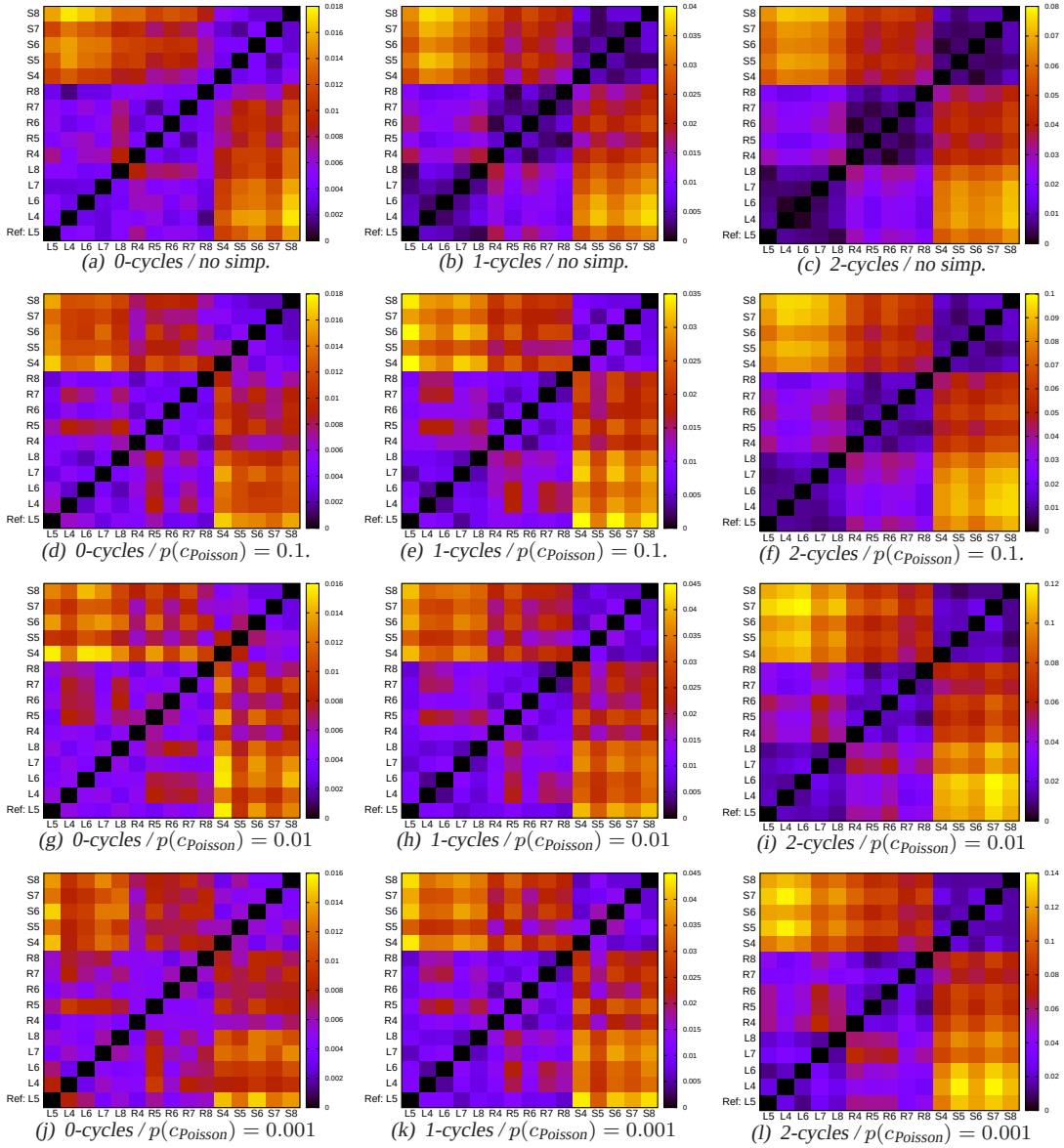
(a) 0-cycles / no simp. *(b) 1-cycles / no simp.* *(c) 2-cycles / no simp.*

(d) 0-cycles / $p(c_{Poisson}) = 0.1$. *(e) 1-cycles / $p(c_{Poisson}) = 0.1$.* *(f) 2-cycles / $p(c_{Poisson}) = 0.1$.*

(g) 0-cycles / $p(c_{Poisson}) = 0.01$ *(h) 1-cycles / $p(c_{Poisson}) = 0.01$* *(i) 2-cycles / $p(c_{Poisson}) = 0.01$*

(j) 0-cycles / $p(c_{Poisson}) = 0.001$ *(k) 1-cycles / $p(c_{Poisson}) = 0.001$* *(l) 2-cycles / $p(c_{Poisson}) = 0.001$*

Figure 7.6: ***Two-tailed Kolmogorov-Smirnov two-sample test outcomes for mean density curves / real space identical*** z***.*** *The test setup is analogous to that described in the text or in the caption of figure 7.4. From left to right: increasing cycle dimension. From top to bottom: increasing simplification threshold. Table indicators:* Ln*,* Rn *and* Sn *stand for LCDM respectively RP or SUGRA run* n*+10.*



(a) 0-cycles / no simp. *(b) 1-cycles / no simp.* *(c) 2-cycles / no simp.*

(d) 0-cycles / $p(c_{Poisson}) = 0.1$. *(e) 1-cycles / $p(c_{Poisson}) = 0.1$.* *(f) 2-cycles / $p(c_{Poisson}) = 0.1$.*

(g) 0-cycles / $p(c_{Poisson}) = 0.01$ *(h) 1-cycles / $p(c_{Poisson}) = 0.01$* *(i) 2-cycles / $p(c_{Poisson}) = 0.01$*

(j) 0-cycles / $p(c_{Poisson}) = 0.001$ *(k) 1-cycles / $p(c_{Poisson}) = 0.001$* *(l) 2-cycles / $p(c_{Poisson}) = 0.001$*

Figure 7.7: ***Two-tailed Kolmogorov-Smirnov two-sample test statistic values for mean density curves / real space identical*** z***.*** *The test setup is analogous to that described in the text or in the caption of figure 7.4. From left to right: increasing cycle dimension. From top to bottom: increasing simplification threshold. Table indicators:* Ln*,* Rn *and* Sn *stand for LCDM respectively RP or SUGRA run* n*+10.*

7.2 Redshift space

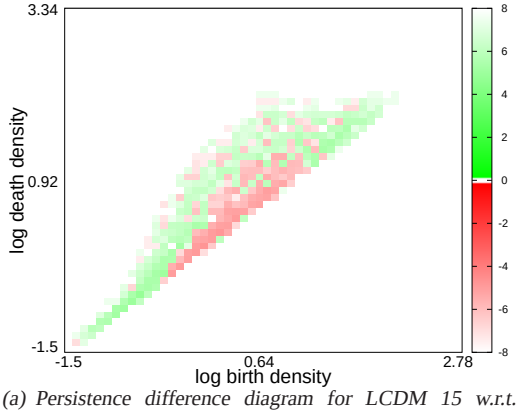
Observations are made in redshift space, which produces distortions mainly in the line of sight direction, cf. [Hamilton, 1998] for an overview. Such distortions might influence the topology and thereby the pattern we found in the statistical tests. To see whether this is indeed the case, we repeat the analysis of previous section in redshift space. For this, instead of using the simulations we described in subsection (cf. section 4.1.3) we use an analogous run (run 6) for which linear redshift distorted boxes are readily available. Run 6 is identical in all respects, safe for (i) other initial conditions and (ii) slightly different z values. Although the latter prohibits direct comparison of run 6 simulations with the other simulations of this thesis, we can compare the different DE models of run 6 with each other. For each of the DE models, 4 redshift distorted simulation boxes are avaiable. The boxes differ in 'observational position', being: (a) $(75, 75, 75)$ Mpc; (b) $(225, 75, 75)$ Mpc; (c) $(200, 0, 100)$ Mpc and (d) $(50, 75, 200)$ Mpc; with $(0, 0, 0)$ Mpc one of the box corners. This gives us a total of 3 DE models $\times$ 4 observational positions $= 12$ snapshots. For each of these snapshots, the persistence diagram is computed. In subsection 7.2.1 we inspect the similarities and differences of these persistence diagrams visually, followed by an investigation of the effect of simplification in subsection 7.2.2. The statistical test procedure is explained in 7.2.3 and the test results are presented in subsection 7.2.4.

7.2.1 Visual inspection

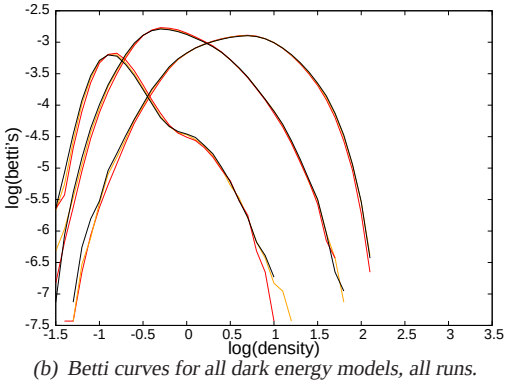
Proceeding along the same lines as in the real space case, several persistence diagrams have been plotted on the left of figure 7.8. A clear distinction between LCDM and SUGRA is visible, consistent throughout dimensions (the top three persistence diagrams) and between different distorted positions (bottom three persistence diagrams). In addition to the low-birth low-death tail also visible in real space, LCDM seems to have a more pronounced high-birth high-death tail as well. Whether this comes due to redshift distortions or slightly different z values cannot be answered at this point. Because differences in observational position naturally introduce some extra spreading, the 1D summary curves on the right show only the curves based on distortions as observed from position (a). Comparing (only the general) characteristics of the curves with their normal space variants, we see that (i) the redshift curves here reach up to much less high densities and (ii) skewed lognormality is lost due to 'bumps'. As we saw in section 6.2 lognormallity is preserved under z evolution, thus the loss of lognormality is a redshift distortion effect. Redshift distortions spread the highest density regions in the line of sight direction, the characteristic 'fingers of God'. This probably explains both observations. A small difference between the curves of various DE models is visible at low densities.

7.2.2 The effects of simplification

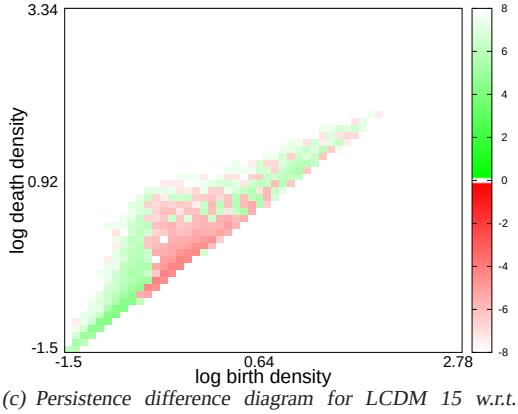
The simplified Betti curves for all DE models all positions are displayed in figure 7.9. The spread due to observations from different positions as mentioned above is clearly visible. More careful inspection shows that the low density differences between DE models are robust under simplification. Interestingly, the 'bumps' in the Betti curve are not. Whereas in particular the 2-cycles show a bump in the higher density region, this bump is almost completely simplified out at the highest simplification threshold considered here. This directly implies that the cycles responsible for this bump are mainly of low persistence.



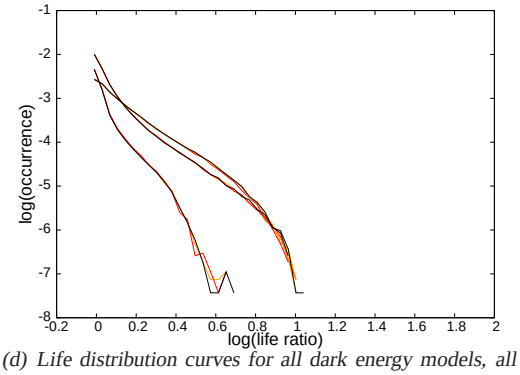
(a) Persistence difference diagram for LCDM 15 w.r.t. SUGRA 16, 0-cycles.



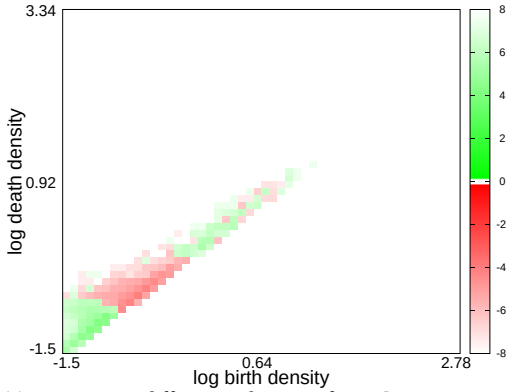
(b) Betti curves for all dark energy models, all runs.



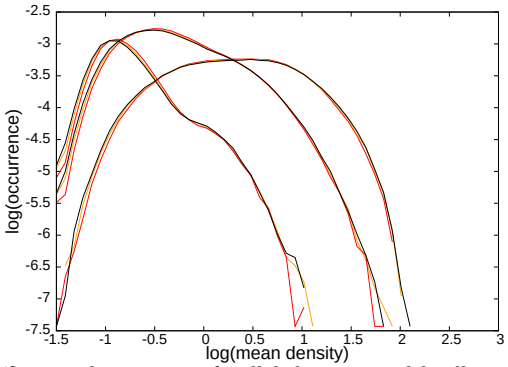
(c) Persistence difference diagram for LCDM 15 w.r.t. SUGRA 16, 1-cycles.



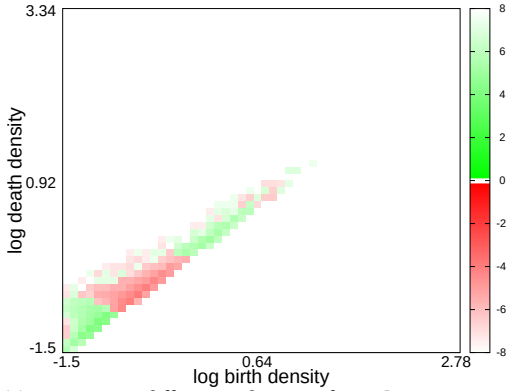
(d) Life distribution curves for all dark energy models, all runs.



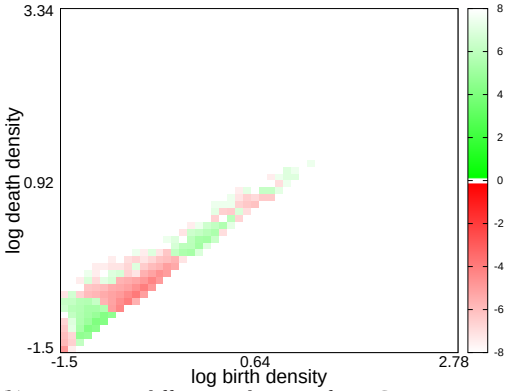
(e) Persistence difference diagram for LCDM 15 w.r.t. SUGRA 16, 2-cycles.



(f) Mean density curves for all dark energy models, all runs.

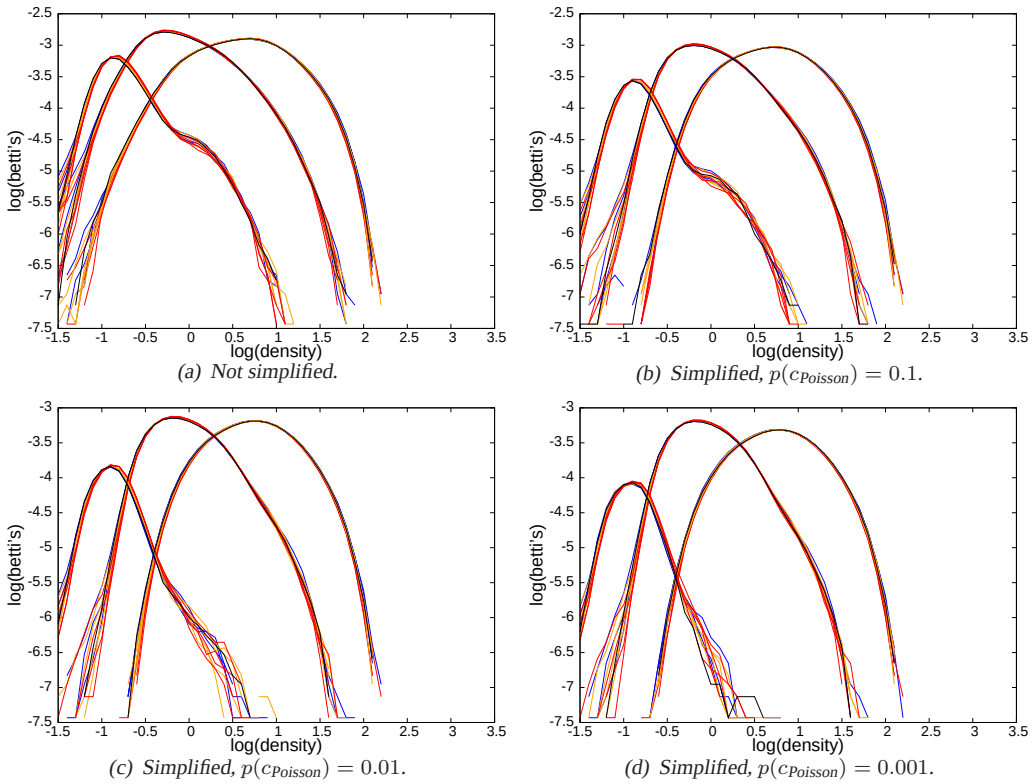


(g) Persistence difference diagram for LCDM 15 w.r.t. SUGRA 17, 2-cycles.



(h) Persistence difference diagram for LCDM 15 w.r.t. SUGRA 18, 2-cycles.

Figure 7.8: ***Comparison of persistence diagrams in redshift space at*** $z = 0$**.** *Left: representative persistence difference diagrams, with green an excess of the reference LCDM run 15 and red an excess of another dark energy model. Stronger colors means larger excess, the numbers next to the color bar indicate the (negative) power of the excess, so note that for green the color bar is inverted. Right: 1D summary curves for LCDM (black), RP (yellow) and SUGRA (red), all at position (a). As can be seen from the 1D summary curves presented earlier: cycle dimension decreases from left to right curve set.*



(a) Not simplified. *(b) Simplified, $p(c_{Poisson}) = 0.1$.*

(c) Simplified, $p(c_{Poisson}) = 0.01$. *(d) Simplified, $p(c_{Poisson}) = 0.001$.*

Figure 7.9: ***Betti curves in redshift space, at identical z with increasing simplification threshold.*** *Betti curves for LCDM (blue and LCDM run 15 black), RP (yellow) and SUGRA (red), with increasing simplification threshold. $p(c_{Poisson}) = x$ indicates that the probability of a cycle to be generated by Poisson noise is equal or smaller than x. As can be seen from the 1D summary curves presented earlier: cycle dimension decreases from left to right curve set.*

7.2.3 Statistical analysis / test setup

The test setup for the statistical analysis is identical to the setup for the real space case described in subsection 7.1.3. Instead of using several runs of the same DE model, we use one run of each DE model, distorted from several positions. In the real space case the mean density curve turned out to be the most sensitive tracer. Furthermore, it turned out that the test statistic outcomes and test statistic values showed an analogous pattern, but using the latter a pattern was more easily distinguishable. Therefore, for compactness we restrict ourselves to the testing the difference between mean density curves and only show the test statistic values.

7.2.4 Statistical analysis / results

The test statistic values are shown in figure 7.10 with increasing from left to right and simplification factor increasing from top to bottom. Focussing first on the unsimplified case (top row) we see the various DE models are clearly distinguishable, specifically 1-cycles and 2-cycles. Now taking into account the effects of simplification, we notice that it affects the statistical pattern much stronger than in the unsimplified case (cf. figure 7.7). This is in line with previous subsection, where we saw complete bumps disappearing under simplification. For 0-cycles simplification almost completely destroys any pattern, whereas for 1-cycles and 2-cycles the pattern becomes stronger. For 2-cycles some extra strong lines at the rows and columns associated with position d are visible. Only this position gives extra strong lines, so the extra strenght is probably

due to a different point of view (literally) rather than anything else.

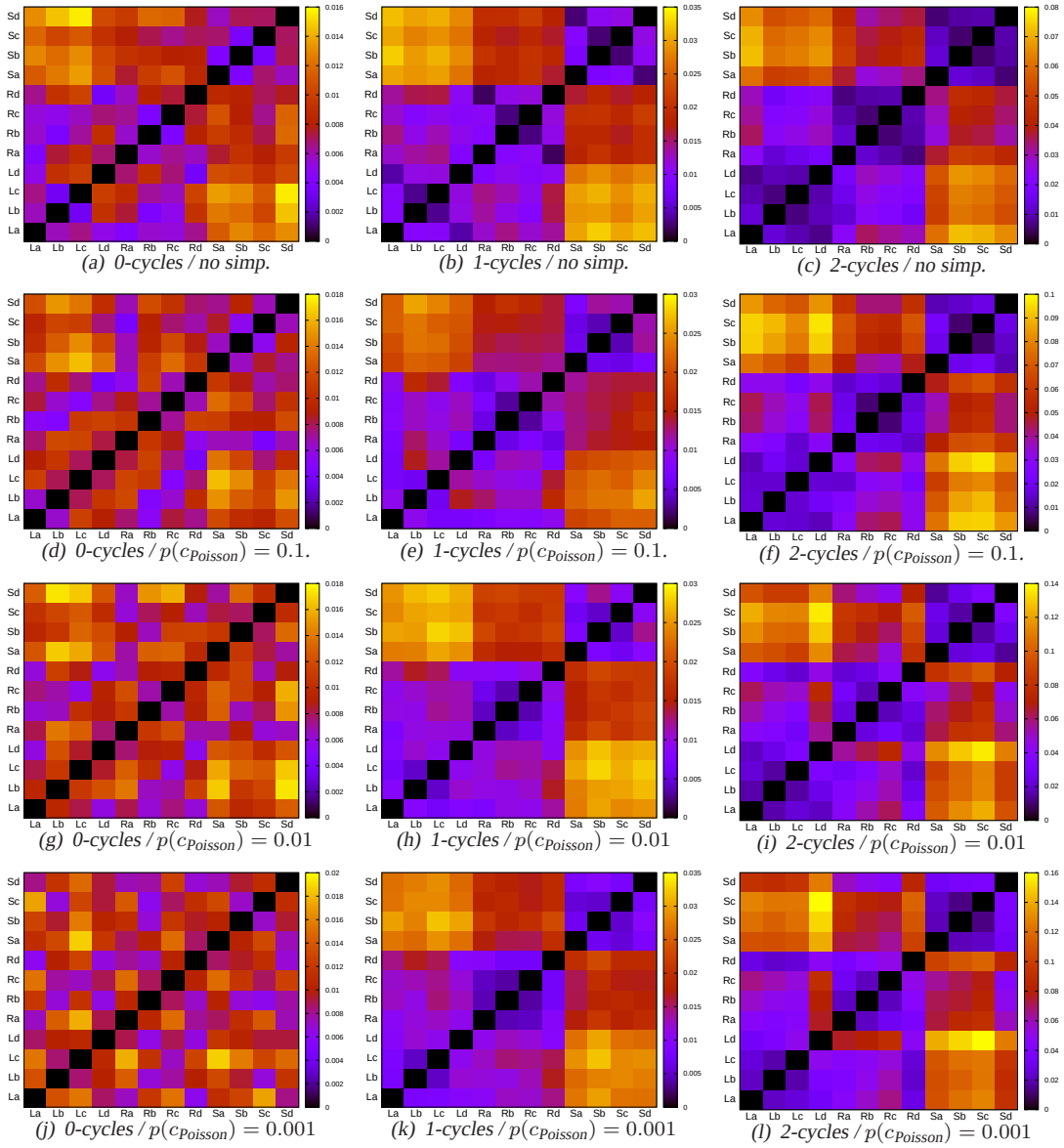
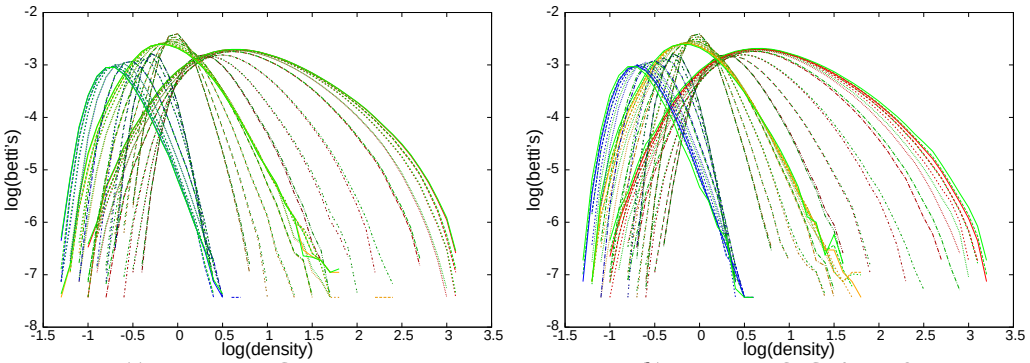


Figure 7.10: ***Two-tailed Kolmogorov-Smirnov two-sample test statistic values for mean density curves / redshift space identical*** z***.*** *The test setup is analogous to that described in the text or in the caption of figure 7.4. From left to right: increasing cycle dimension. From top to bottom: increasing simplification threshold. Table indicators: Ln, Rn and Sn stand for the LCDM respectively RP or SUGRA run at observational positions* n*, with* $n \in \{a, b, c, d\}$*.*

7.3 Real space identical clustering parameter

In section 7.1 we found a significant difference between DE models at the same z. But this doesn't imply the models also differ significantly at the same stage of structure evolution. To consider the latter, we need to compare the persistence diagrams from all models all runs at identical σ_8 values. Unfortunately, the available Gadget snapshots are at fixed z values. We therefore use binwise interpolation or extrapolation to obtain the persistence diagrams at the σ_8 values of

LCDM run 15, using a 2nd degree polynomial. For LCDM run 17 (σ_8's are very similar to LCDM run 15) and SUGRA run 18 (σ_8's are very different from LCDM run 15) the interpolated/extrapolated Betti curves for all epochs are shown in figure 7.11. At very low redshifts some of the extrapolated curves start to behave irregular, probably because the extrapolation breaks down there. For LCDM run 17 the differences between the z-based original and interpolated Betti curves are more or less negligible. For SUGRA run 18 with small z on the contrary, the differences between original and interpolated Betti curves are larger than the differences between successive time steps of one of them.



(a) Betti curves LCDM run 17. *(b) Betti curves SUGRA run 18.*

Figure 7.11: ***Interpolated Betti curves.*** *Betti curves of at fixed z are directly avaible from the Gadget snapshots, colored in red (0-cycles), yellow (1-cycles) and blue (2-cycles). Using binwise interpolation or extrapolation, the Betti curves at fixed σ_8 are determined, colored green (for all cycles). Brighter means lower z.*

In analogy with the real space identical redshift case in section 7.1, we first inspect differences visually. Than, we turn to statistical analysis.

7.3.1 Visual inspection

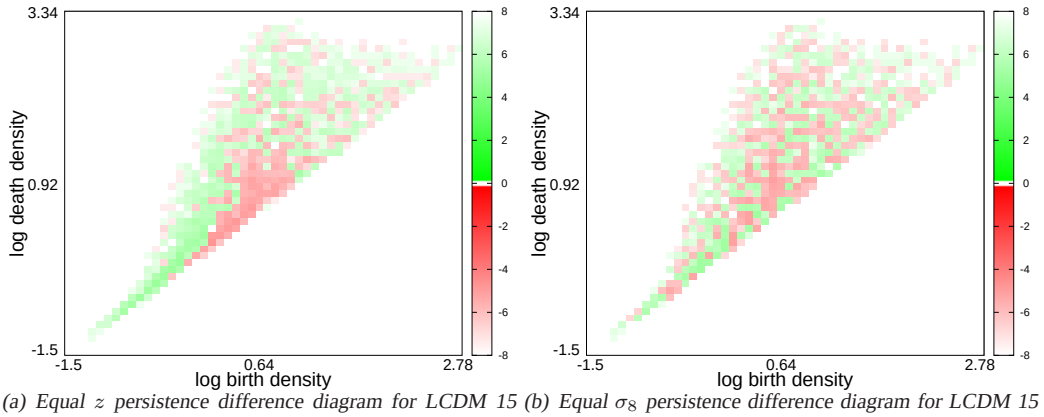
A selection of representative persistence difference diagrams compared at identical σ_8 values is shown *in the right column* of figure 7.12. The same difference diagrams compared at identical z values are shown *in the left column* of the figure. We observe the following:

- For the 0-cycle diagram (top row), the pattern visible in the identical-z case disappears in the identical-σ_8 case.
- For the 2-cycle diagrams (rows 2-4) the identical-z case shows a consistent pattern, but the consistency of the pattern is lost in the identical-σ_8 case. In the diagram on the second row (right column) we see a SUGRA excess at lower birth and lower deaths, whereas in the diagram below SUGRA displays an excess at intermediate densities. In the bottom row SUGRA shows an excess at higher births and higher deaths.

The identical-σ_8 1D summary curves (not shown here) agree with previous observations: they are nearly indistinguishable one their complete density range.

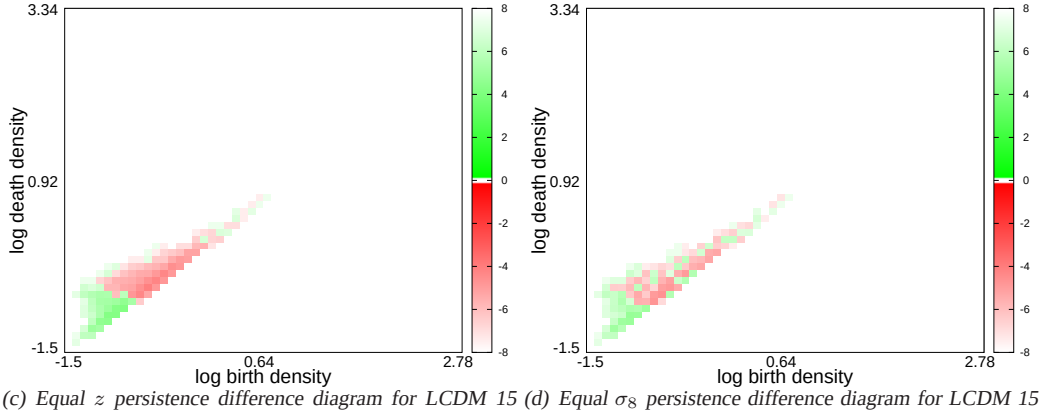
7.3.2 The effects of simplification

To investigate the effects of simplification, figure 7.13 shows its effect on the Betti curves for increasing simplification threshold. In analogy with the equal-z case, simplification doesn't seem to have any effect on the presence of patterns.



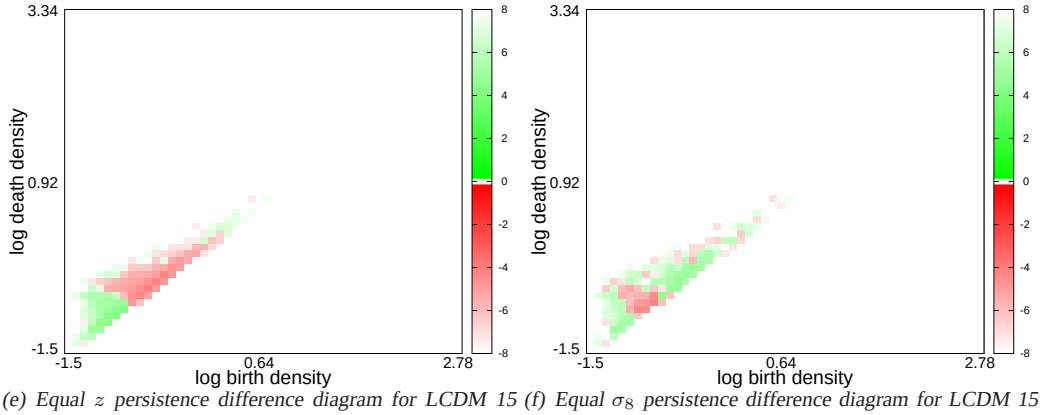
(a) Equal z persistence difference diagram for LCDM 15 w.r.t. SUGRA 16, 0-cycles.

(b) Equal σ_8 persistence difference diagram for LCDM 15 w.r.t. SUGRA 16, 0-cycles.



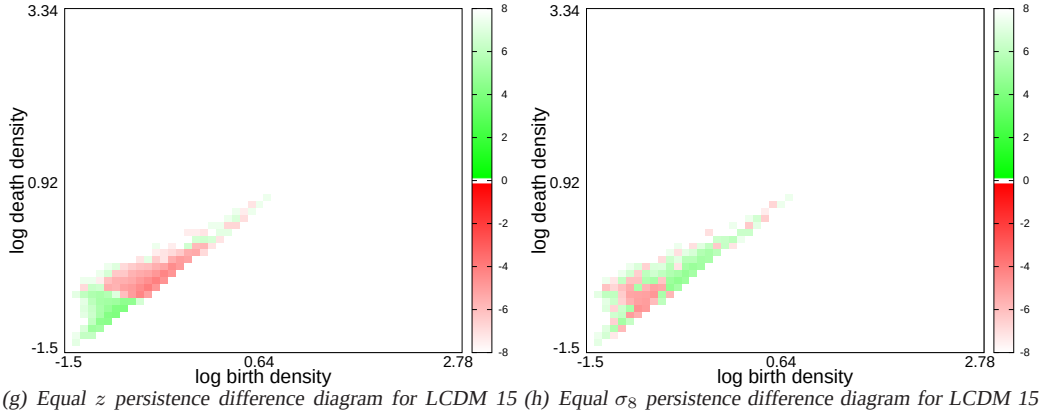
(c) Equal z persistence difference diagram for LCDM 15 w.r.t. SUGRA 16, 2-cycles.

(d) Equal σ_8 persistence difference diagram for LCDM 15 w.r.t. SUGRA 16, 2-cycles.



(e) Equal z persistence difference diagram for LCDM 15 w.r.t. SUGRA 17, 2-cycles.

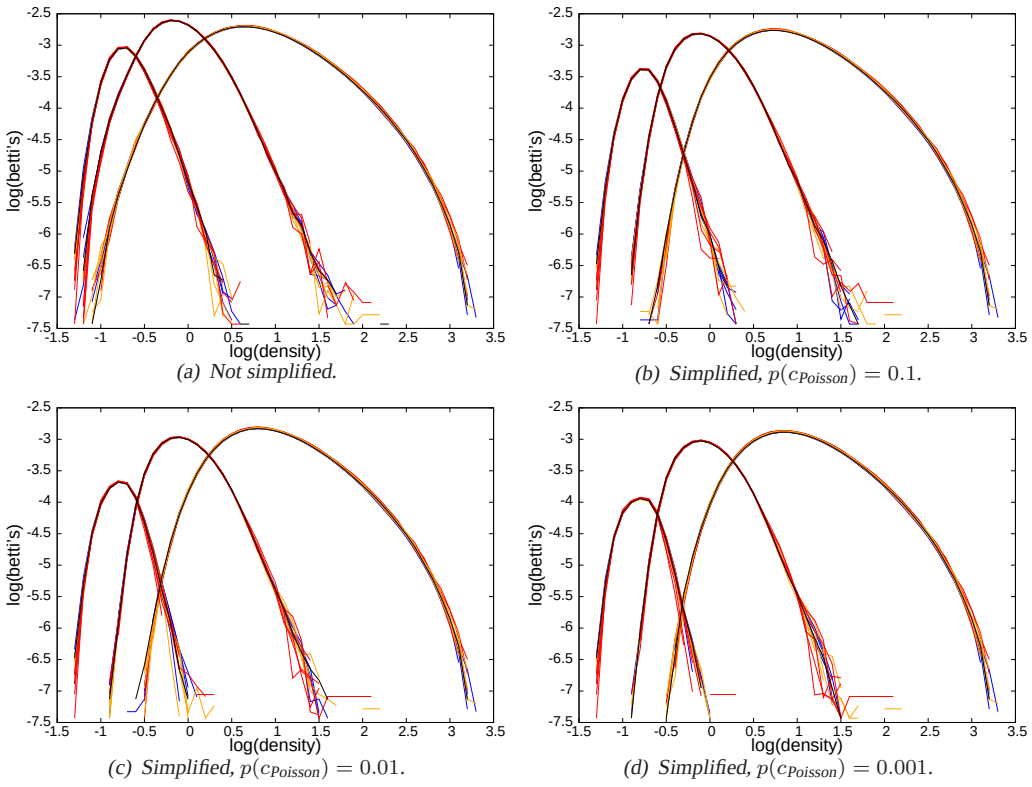
(f) Equal σ_8 persistence difference diagram for LCDM 15 w.r.t. SUGRA 17, 2-cycles.



(g) Equal z persistence difference diagram for LCDM 15 w.r.t. SUGRA 18, 2-cycles.

(h) Equal σ_8 persistence difference diagram for LCDM 15 w.r.t. SUGRA 18, 2-cycles.

Figure 7.12: ***Persistence difference diagrams at identical z (left) vs identical σ_8 (right).*** *Representative persistence difference diagrams, with green an excess in the reference LCDM run 15 and red an excess another dark energy model. Stronger colors means larger excess, the numbers next to the color bar indicate the (negative) power of the excess, so note that for red the color bar is inverted. Left: difference diagrams at identical z; right: the same difference diagrams at identical σ_8. In case of the identical z structural differences between LCDM and SUGRA are visible. These structural differences disappear in the identical-σ_8 case.*



(a) Not simplified. *(b) Simplified, $p(c_{Poisson}) = 0.1$.*

(c) Simplified, $p(c_{Poisson}) = 0.01$. *(d) Simplified, $p(c_{Poisson}) = 0.001$.*

Figure 7.13: ***Betti curves at identical σ_8 with increasing simplification threshold.*** *Betti curves for LCDM (blue and LCDM run 15 black), RP (yellow) and SUGRA (red), with increasing simplification threshold. $p(c_{Poisson}) = x$ indicates that the probability of a cycle to be generated by Poisson noise is equal or smaller than x. As can be seen from the 1D summary curves presented earlier: cycle dimension decreases from left to right curve set.*

7.3.3 Statistical analysis / test setup

The interpolated Betti curves have all the same σ_8. Any statistically significant differences between them are solely attributable to DE induced intrinsic topological differences in structure. Unfortunately, binning makes it impossible to define an empirical distribution function and thus the KS test (or related measures) cannot be used. We therefore switch to (Pearson's) χ^2 test. The χ^2 test has three main disadvantages: (i) it is in general less sensitive than the KS test [Wall and Jenkins, 2003]; (ii) test outcomes might depend on bin size[2] and; (iii) the number of empty cells should be limited. To mediate the second disadvantage, we note that for Betti curves the power of the density is approximated reasonably well by a normal distribution (cf. section 6.1). Therefore, we can use Heald's optimal bin width δx [Heald, 1984], defined as:

$$\delta x = \sigma \left(\frac{20}{n} \right)^{1/5}$$

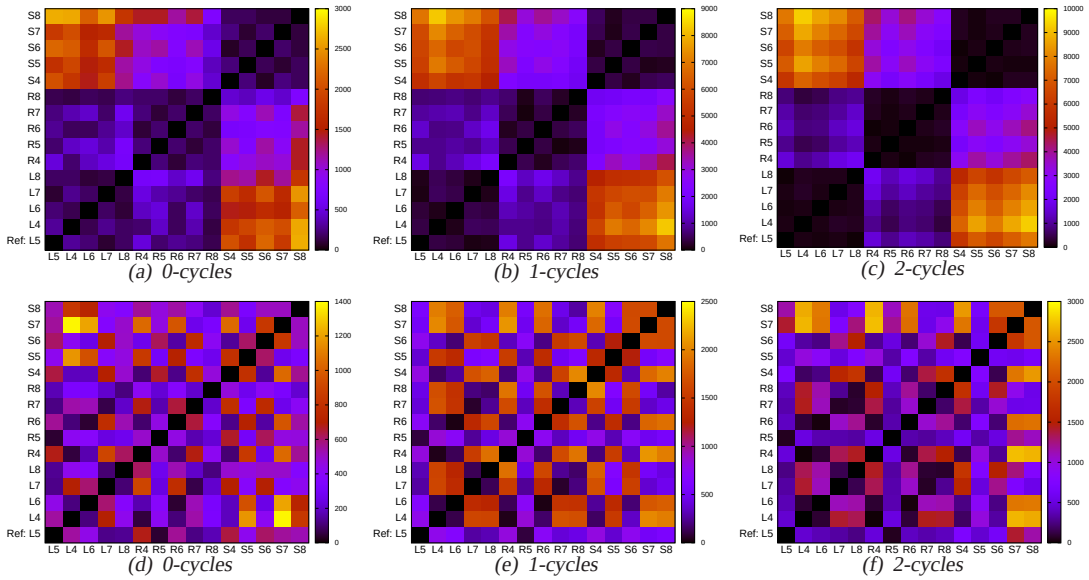
with σ the cycle density standard deviation and n the number of cycles. To counteract the third disadvantage we cut all Betti curves near the end of their tails. This will make results more trustworthy, but at the price of further reduced sensitivity.

[2]The shear size of our sample makes this effect probably limited, though.

7.3.4 Statistical analysis / results

To see how well the χ^2 test detects differences, we first apply it to a more known case. The top row of figure 7.14 shows the χ^2 values at the same z, analogously with the tests performed in section 7.1. Again the models are unmistakably distinguishable: SUGRA differs visibly with LCDM as well as with RP. Differences between LCDM and RP are smaller and differences between several runs of the same model are very small. Apparently, although the χ^2 test might not be sensitive as KS tests, it is still sensitive enough to detect the different models.

We now proceed with a comparison between the interpolated Betti curves, as on the bottom of figure 7.14. Note that the values of the χ^2 statistics decrease by a factor 2 till 3. Hardly any pattern is left. Only for a few specific combinations of LCDM and SUGRA runs, the 0-cycles show slightly larger χ^2 values, but they are far from convincing. We therefore conclude that it is unlikely that different DE models generate intrinsically different structure topology.



(a) 0-cycles *(b) 1-cycles* *(c) 2-cycles*

(d) 0-cycles *(e) 1-cycles* *(f) 2-cycles*

Figure 7.14: χ^2 ***two-sample test statistic values for Betti curves at identical*** z ***(top row) and identical*** σ_8 ***(bottom row).*** *For each pair of dark energy models and runs we test whether the curves have statistically different persistence diagrams and thus topology. In the top row we compare curves at $z = 0$, in the bottom row we compare at $\sigma_8 = 369.591$. This σ_8 value corresponds to the σ_8 value of LCDM run 15 at $z = 0$. To top row shows a pattern whereas the bottom row does not. This suggests that the DE models result in intrinsically identical structure topology. The differences at the identical-z case are solely attributable to different rates of structure evolution. Table indicators: Ln, Rn and Sn stand for LCDM respectively RP or SUGRA run n+10.*

We note that interpolation/extrapolation creates additional uncertainty. Therefore, in this section we make no statements about the statistical significance of the disappearance of a pattern.

7.4 Summary observations

Dark energy rules the global evolution of the universe, but it might mingle in local affairs as well. If so, it may leave a distinct imprint on the topology of cosmic structure. To see whether such imprint exists the persistence diagrams of the three DE models considered in this thesis are computed and compared.

Differences between the topology of cosmic structure induced by different DE models consists of two independent factors: (i) the rate of structure evolution varies between DE models and (ii) DE models might each result in intrinsically different structure topology. The first and

the second effect can be found by comparing persistence diagrams of different DE models at the same z respectively at the same σ_8.

Comparison of persistence diagrams of several DE models at the same z learns us:

1. Some structural differences are visible in the low birth low death tail of the persistence diagrams: LCDM is shifted consistently to slightly lower densities than RP, whilst SUGRA is shifted to slightly higher densities. The differences are quite small compared to function value. Physically, they might be caused by the difference in the rate of structure evolution. The observed differences are robust under simplification.

2. The lifetime curves and in particular the mean density curves show a statistically significant difference between the various DE models. Mean density 2-cycles at intermediate simplification thresholds give the most pronounced difference.

Thus, in real space at a given z it is in principle possible to distinguish between DE models based on persistence diagrams. To see or whether these differences are also *observable*, we repeat part of the analysis above in redshift space:

1. Again LCDM persistence diagrams are systematically shifted to the low birth low death tail of persistence diagrams, in particular for 2-cycles. Contrary to the real space case, a high density effect seems visible as well. The observed differences are robust under simplification.

2. The transformation from real space to redshift space severely distorts the topology of cosmic structure, as is visible from in particular the Betti curves and mean density curves. The most prominent differences between the redshift space versions of these curves compared with their normal space counterparts are: (i) the excess of low density cycles and (ii) the presence of 'bumps'. With increasing simplification threshold these bumps become smaller, implying that they are caused by low persistent cycles.

3. The mean density curves show a statistically significant difference between the various DE models, especially for 1-cycles and 2-cycles. The difference is robust under simplification.

Hence, using persistence it is in principle possible to observationally differentiate between various DE models!

The observed dissimilarities of the cosmic structure might be due to one or both of the following: (i) different structure formation rates and (ii) intrinsic topological differences. To determine the magnitude of the latter, the persistence diagrams of several DE models at the same σ_8 are compared:

1. Visually, the structural differences between persistence diagrams of various DE models largely disappears and isn't recovered by simplification.

2. Statistical tests seem to confirm previous.

The DE models here don't result in intrinsically different structure topology.

7.5 Discussion and conclusions

Three aspects of the analysis above require some more thought:

1. Throughout this chapter, for computational efficiency all computations and visualizations of complete persistence diagrams were done on a 2D histogram instead of the original point diagram. Mathematically this is not shown to be equivalent, but some investigation suggests neither the presence nor the magnitude of any patterns are influenced by the specific choice of any reasonable bin size.

2. Of more pressing concern are the χ^2 tests performed on Betti curves in section 7.3. Technically, the χ^2 test demands that every point is put in one bin and that the choice of bin is independent for all points. These conditions are not met for Betti curves: a persistent cycle will contribute in many Betti curve bins and consequently these bins are not independent. The χ^2 test performs quite well in the equal-z case though. Furthermore, a short experiment with mean density curves shows they exhibit the same pattern as the Betti curves. Thus not strictly meeting the mathematical requirements doesn't seem to have very large effects.

3. In the discussion of chapter 5 we claimed that specific types of perturbations might influence persistence diagrams in different ways. In that chapter we consider random perturbations of predetermined magnitude and here we studied redshift perturbations. A short comparison of the resulting Betti curves learns us that these different perturbations influence the persistence diagrams indeed in different ways, supporting the claim above. Consequently, perturbation results for one type of perturbation have only limited value for other types of perturbations.

With the previous thoughts in mind, we conclude:

1. The first and second questions posed in the introduction of this chapter can be answered with a definite *yes*. At fixed low z different DE models result in significantly different structure topology, both in real and redshift space. In principle, this makes possible to observationally distinguish between dark energy models based on their persistence diagrams!

2. The effects above can be found both from lifetime curves and mean density curves, although these curves are not sensitive at all cycle dimensions and simplification thresholds. Mean density 2-cycles distinguish the DE models nearly flawless, especially for intermediate ratio simplification thresholds.

3. At fixed σ_8, the persistence diagrams of the DE models investigated here are (nearly) indistinguishable. Thus LCDM, RP and SUGRA don't result in intrinsically different structure topology.

CHAPTER 8

Discussion & conclusions

8.1 Overall conclusions

1. Homological discrete topology is a natural choice to describe the structure, shape and connectedness of the cosmic web. It visualizes topology using a Discrete Morse Complex and provides analytical handles via persistence diagrams (cf. section 3.6).

2. Persistence diagrams are (i) very sensitive to the presence of structure and (ii) robust against uncertainties up to at least 10 %.

3. Persistence simplification is an excellent, intuitive and computationally efficient method to take out irrelevant features and noise.

4. Persistence diagrams of the cosmic web in an LCDM universe at $z = 0$ have the following characteristics:

   (a) The 1-cycle persistence diagram has an extended inverse v-shape along the birth=death diagonal. The 0-cycle diagram fans out towards higher births and higher deaths, whereas the 2-cycle diagram goes in opposite direction.

   (b) The Betti curves and product mean density curves follow nearly exact a skewed log-normal pattern. The lifetime curves are monotonically decreasing, especially the 1-cycle and 2-cycle curves decrease quite fast.

5. With time structure evolves. For the persistence diagrams of the cosmic web of an LCDM universe, this is visible as follows:

   (a) In time persistence diagrams fan out. 0-cycles mainly migrate to higher densities, 2-cycles mainly to lower densities and 1-cycles spread quite evenly in both directions. This spread also clearly shows from the mean density curves and to a lesser degree from the Betti curves, in the form of shifting μ values and increasing σ of their skewed lognormal fits. Interestingly, no trend seems to be visible in the skewness parameter α.

   (b) Simplification enhances the effect of time evolution, as earlier epochs are thinned out much more than later epochs. This suggests that at earlier epochs structures are less persistent.

6. Comparison of the cosmic web topology of LCDM, RP and SUGRA universes learns us:

   (a) At fixed low z ($z \lesssim 1$) the different DE models result in significantly different structure topology, both in real and redshift space. In principle, this makes possible to observationally distinguish between dark energy models based on their persistence diagrams!

   (b) Mean density 2-cycles at intermediate ratio simplification thresholds are the best DE model differentiators.

   (c) At fixed σ_8 the persistence diagrams of the DE models investigated here are (nearly) indistinguishable. Thus LCDM, RP and SUGRA don't result in intrinsically different structure topology.

8.2 Methodological discussion

Two aspects regarding the topological analysis within this thesis warrant some attention:

1. In section 3.6 we discussed impossible cancellations (cf. figure 3.23 in particular). The topological algorithms used within this thesis (cf. sections 4.3 and 4.4) provide insufficient information to determine whether a simplification is possible (i.e. allowed). Therefore, all simplification are carried out whether possible or not. Although the percentage of impossible cancellations is expected to be low and their relevance limited [Sousbie, 2011; Pranav, 2013], a more detailed study of the effects of carrying out impossible cancellations would further substantiate the conclusions of this thesis.

2. Of more fundamental nature is the structure the Discrete Morse Complex (cf. definition 3.26) *imposes* on the Cosmic Web: every 1-simplex connects two 0-simplices, every 2-simplex is surrounded by 1-simplices, etc. From observations we know this is not necessarily the case in the real universe, for example some filaments just end in the middle of a void. As such the Discrete Morse Complex is an idealized structure. Artificially excluding simplices of the Discrete Morse Complex from analysis might be an option, but current algorithms don't provide such options yet. (Note that artificially excluding simplices is not the same as simplification, as after simplification the manifold is restructured such that all conditions imposed by the Discrete Morse Complex remain satisfied.)

8.3 Where to go from here

As direct follow-up of this thesis, we suggest the following:

1. Here we showed that series of pair-wise statistical tests display patterns indicating the difference between various DE models. A non-parametric group test to show the statistical significant difference between a set of runs of one DE model and a set of runs from another DE model could not be performed, as the conditions for the Kruskal-Wallis test are not met (cf. subsection 7.1.3). In section 6.1 we noted the good fit of a lognormal to the Betti curves. It would be interesting to fit the Betti curves of all DE models all runs with lognormals and perform a classical parametric ANOVA test on these fits.

2. Very recently, the branch of statistical topology developed discrete topology hypothesis testing [Robinson and Turner, 2013; Turner, 2013]. Such hypothesis tests compare two

persistence diagrams using the *Wasserstein metric* (also called the Earth mover's distance). Informally, the metric can be described as a 'minimum effort' required to change one persistence diagram in the other. Unfortunately, the algorithms for such a test are not publicly available yet (and writing them is beyond the scope of a master thesis). When the algorithms become publicly available, we recommend to repeat the analysis above using the complete persistence diagram instead of a 1D summary curve.

3. In this thesis we based our analysis on the density field but other cosmic fields like the velocity field can be used as well [Cautun et al., 2013; Hoffman et al., 2012]. Repeating the analysis of this thesis using the velocity field instead of the density field would be useful.

More fundamentally, how a specific (density) function gives rise to a specific persistence diagram is poorly understood. That is, we can calculate the persistence diagram of the function algorithmically with the data pipeline described in chapter 4 but we cannot predict it or calculate it from more basic principles. [Feldbrugge et al.; Pranav et al., 2013c] provide first steps in this direction, but a lot of work remains to be done.

From a visualization perspective, several ways to visualize the cosmic web exist (cf. section 2.6). Comparison of the visual output of these methods with the visual output delivered by discrete topology (in the form of the Discrete Morse Complex, cf. definition 3.26) would be insightful.

8.4 Popular scientific summary

A popular scientific summary of this thesis will be published in the Periodiek [Nevenzeel, 2014].

APPENDIX A

Mathematical background

A.1 Topological essentials in 10 minutes

Since the dawn of civilization geometry is closely related to astronomy. In antiquity, the Greek used geometry to describe heavens and as such developed very accurate astronomical models far beyond their time. In the many centuries that followed geometry remained closely connected to astronomy. At the beginning of the 17th century geometrical astronomy again provided an astronomical breakthrough: following classical tradition, Kepler stated his laws of planetary motion in geometrical terms. His work not only provided strong support for heliocentric model, Newton also used it as evidence for his general laws of motion. In modern times, Einstein's general relativity basically asserts a relation between the geometry of spacetime and the mass distribution. General relativity was proved to be correct by astronomical observations.

Basically a geometric property is a property unaffected by translation or rotation of a coordinate system, such as the orbit of planets around the sun, the influence of mass on spacetime or the height of a mountain. Here we directly see one of the limitations of geometry: an (implicite) coordinate system (technically: a distance metric) is required. Topology relaxes this constraint by letting go the requirement of a distance metric and only requiring a notion of 'near' or 'far', without specifying exactly how near or far. As such, topology is sometimes described as 'rubber sheet geometry': objects may be twisted, deformed or stretched without influencing topology. Contininuing the examples above: the presence of a peak in a mountain landscape is a topological property of the height function, just as the presence of a potential well in the graviational field function. Here we see at once both the advantage of as well as the extra layer of abstraction added by topology over geometry. The geometric property of peak height will in general change: in the winter snow might heighten the mountain peak and in time erosion will lower it. However, the topological property not concered exact distances remains invariant: there is a mountain peak. Below follows a very short crash course of the most important topological concepts required for this thesis.

Our story begins [Katok and Sossinsky; Sutherland, 2005] as many stories in mathematics begin, with a definition:

Definition A.1 (Metric space)**.** *A metric space (M, d) is a non-empty set M on which a distance function d is defined. Take $x, y \in M$, than $d(x, y)$ must satisfy:*

1. $d(x, y) \geq 0 \; \forall \; x, y \in X$;

2. $d(x, y) = 0 \Leftrightarrow x = y$;

3. $d(x, y) = d(y, x)$;

4. $d(x, z) \leq d(x, y) + d(y, z)$

Example A.1 (Euclidean space)**.** *$\mathbb{R}^n$ with the normal Pythagorean distance $d(x, y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$ is a metric space.*

Example A.2. *$\mathbb{R}^2$ with the distance function $d(x, y) = \cos(x)\cdot\sin(y)$ is not a metric space as the distance function violates all demands: (1) it can become negative; (2) $d(x, y) = 0 \nLeftrightarrow x = y$; (3) it is not symmetric in x and y and (4) is violated as well, take for example $n = 1$ and $(x, y, z) = (0, 0, \pi/2)$.*

In an n-dimensional metric we can define an:

Definition A.2 (Open n-ball $B^o_\epsilon(a)$)**.** *Given metric space (M, d), a point $x \in M$ and $\epsilon \in \mathbb{R}_{>0}$, the open n-ball $B^o_\epsilon(a)$ is the set*

$$B^o_\epsilon(a) = \{x \in M | d(x, a) < \epsilon\}$$

Note: sometimes the number in front of the ball denotes the ϵ instead of the dimension n. Here we use n as for our purpose specific values of ϵ are not so relevant, but the dimension is.

Analogously, one can define a:

Definition A.3 (Closed n-ball $B_\epsilon(a)$)**.** *Given metric space (M, d), a point $x \in M$ and $\epsilon \in \mathbb{R}_{>0}$, the closed n-ball $B_\epsilon(a)$ is the set*

$$B_\epsilon(a) = \{x \in M | d(x, a) \leq \epsilon\}$$

In essence, a closed ball is an open ball with boundary. From now on, unless mentioned otherwise with the term n-ball we will refer to a closed n-ball.

Example A.3. *Consider $\mathbb{R}^n$. An open 1-ball is $B^o_\epsilon(a) = (a - \epsilon, a + \epsilon)$ and a (closed) 1-ball is $B_\epsilon(a) = [a - \epsilon, a + \epsilon]$. I.e. the open and closed balls are open respectively closed line segments. An open 2-ball is the interior of a disc with radius ϵ centered at a and a (closed) 2-ball the disc with boundary. A (closed) 3-ball is a sphere with boundary, etc. A 0-ball is a point and an open 0-ball doesn't exist.*

Subsequently, we define an:

Definition A.4 (Open subset $U \subset M$)**.** *A subset U of n-dimensional metric space M is open in M if for any $y \in U$ there exists an $\epsilon(y) > 0$ such that an open n-ball with $B^o_{\epsilon(y)}(y) \subset M$.*

Example A.4. *Any open ball $B^o_\epsilon(x) \subset M$ is open M, which one can see by taking a smaller open ball $B^o_\delta(y)$ and choose $\delta = \epsilon - d(x, y)$. The situation is visually illustrated in figure A.1.*

Using open sets, we can let go of a distance metric and define a general:

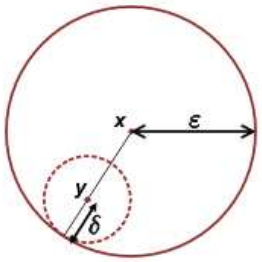


Figure A.1: ***An open ball in an open ball.***

Definition A.5 (Topological space)**.** *A topological space (T, s) is a non-empty set T and a fixed collection of subsets s satisfying:*

1. *$T \in s$ and $\varnothing \in s$;*
2. *the intersection of any two sets in s is again in s;*
3. *the union of any collection of sets is again in s.*

The collection s is called a topology for T and the members of s are called open sets of s. Thus '$U \in s$' and 'U open in s' mean the same.

Example A.5 (Metrizable spaces)**.** *Any metric space (M, d) gives rise to a topological space (T, s) by setting $T = M$ and choosing for s the open subsets on the metric space as defined via definition A.4. A topological space on which a metric can be defined is metrizable. Note that a metrizable space doesn't necessarily have to allow a unique metric.*

Just like a simple vector space, a topological space can have a basis:

Definition A.6 (Topological basis)**.** *Let (T, s) be a topological space. A basis for s is a subcollection $\mathcal{B} \subset s$ such that every set in s is a union of sets from $\mathcal{B}$.*

Naturally, between topological spaces we can define functions which map an element of one space to another space. Topology is used to define closeby and far away without worrying about exactly how closeby or how far away, and as such allow continuous deformations of space. Therefore, we need to have a notion of continuity:

Definition A.7 (Continuity at an element)**.** *Let (T_1, s_1) and (T_2, s_2) be two topological spaces, $f : T_1 \to T_2$ a map between them and a an element of s_1. Then f is continuous at $a \in s_1$ if for any $U_2 \in s_2$ such that $f(a) \in U_2$, there exists $U_1 \in s_1$ such that $a \in U_1$ and $f(U_1) \subset U_2$.*

Just as in real analysis, continuity at an element can be extend to continuity of a function:

Definition A.8 (Continuous topological functions)**.** *Consider the situation of continuity at a point. f is continuous (with respect to s_1 and s_2) $\Leftrightarrow$ it is continuous at every point (every element) of T_1.*

Above we mentioned that topology is concerned with the general shape of figures, we now make that a bit more precise using:

Definition A.9 (Homotopic functions)**.** *Let T_1 and T_2 be two topological spaces, $f, g : T_1 \to T_2$ two continuous functions and $I(t)$ over $[0, 1] \subset \mathbb{R}$. We say that f and g are homotopic if there exists a continuous function $F : X \times I \to Y$ such that $F(x, 0) = f(x)$ and $F(x, 1) = g(x)$. Intuitively, the previous means that two functions are homotopic if they can be continuously transformed in each other. The interval I is used to indicate how far in the transformation proces we are. Homotopy is illustrated in figure A.2.*

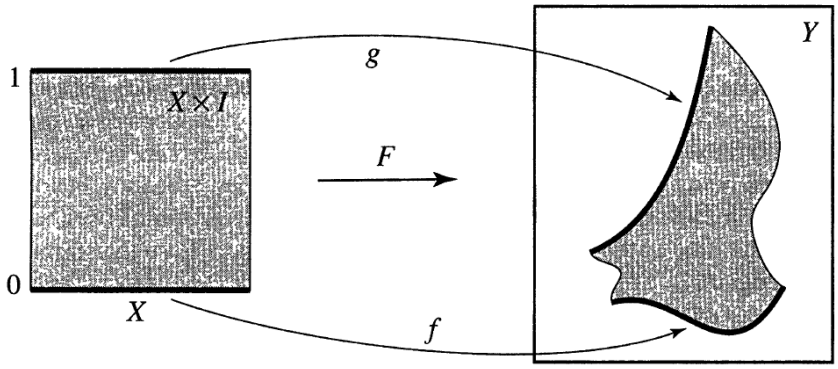


Figure A.2: ***A homotopy transforms f in g [Adams and Franzosa, 2009].***

Example A.6. *Define the continuous function $F : \mathbb{R} \times I \to \mathbb{R}$ by $F(x,t) = x - t$. Then $F(x,t)$ is a homotopy between the identity map $f(x) = x$ and the map $g(x) = x - 1$, which translates the entire graph downward by one element. As t increases from 0 to 1 $F(x,t)$ slowly translates the identity map downwards.*

Example A.7. *Consider the unit disc $\mathbb{D}^2 \subset \mathbb{R}^2$. The identity map on the unit disc $id_\bullet : \mathbb{D}^2 \to \mathbb{D}^2$ and the constant map $c_0 : \mathbb{D}^2 \to 0$ are homotopic. Using polar coordinates, a possible homotopy between them is given by $F((r,\varphi),t) = ((1-t)r, \varphi)$.*

Homotopic functions can be continuously transformed in each other. Using such functions we say when topological spaces are continuously transformable in each other. Spaces for which this is possible are called:

Definition A.10 (Homotopy equivalent spaces)**.** *Two topological spaces T_1 and T_2 are homotopy equivalent if there exists two continuous maps $f : T_1 \to T_2$ and $g : T_2 \to T_1$ such that $g \circ f$ is homotopic to the identity map id_{T_1} on T_1 and $f \circ g$ is homotopic to the identity map id_{T_2} on T_2.*

Example A.8. *Any n-ball is homotopy equivalent with $\mathbb{R}^n$ and with any point p. To see why p is homotopy equivalent with $\mathbb{R}^n$, define the zero-map $f : p \to 0 \in \mathbb{R}^n$ and the point map $g : \mathbb{R}^n \to p$. Then $g \circ f = id_p$ is the identity on the point p whereas the map $f \circ g : \mathbb{R}^n \to \mathbb{R}^n$ is homotopy equivalent to the identity of $\mathbb{R}^n$ by the homotopy $F((r,\varphi),t) = ((1-t)r, \varphi)$.*

As the example above shows, homotopy can lead to 'loss of information', as a point clearly contains less information than a n-dimensional space. To prevent this, the idea of homotopic functions and homotopy equivalence can be sharpened by demanding a one-one relation:

Definition A.11 (Homeomorphisms)**.** *A homeomorphism is a one-one correspondence $f : T_1 \to T_2$ such that both f and f^{-1} are continuous.*

Definition A.12 (Topological equivalence)**.** *Two topological spaces T_1 and T_2 are topologically equivalent if there exists an homeomorphism between them.*

Here we see why continuous deformations don't change topology. As long as every open set is mapped one-one to another open set in a structure preserving way the topology doesn't change. Such mappings are not influenced by deformations, stretching or twisting and so topology is neither.

Example A.9. *Clearly, any topological equivalent spaces are homotopy equivalent as well, whereas the converse is not necessarily true. As counterexample of the latter statement: in example A.8 was shown that $\mathbb{R}^n$ is homotopy equivalent with a point. But as $\mathbb{R}^n$ and a point don't allow a one-one mapping they are not topologically equivalent.*

Example A.10. *Let $a, b, c, d \in \mathbb{R}$ with $a < b$ and $c < d$. Any two open intervals (a, b) and (c, d) in $\mathbb{R}$ are homeomorphic. To see this, consider the function $f : (a, b) \to (c, d)$ given by:*

$$f(x) = c + \{(x - a) \cdot (d - c)/(b - a)\}$$

which maps any open interval in (a, b) to an open interval in (c, d). Since both f and f^{-1} are continuous, (a, b) and (c, d) are homeomorphic.

Example A.11. *A famous example: a donut and a coffee cup are topologically equivalent, as can be seen by the continuous transformation of the one in the other illustrated in figure A.3.*

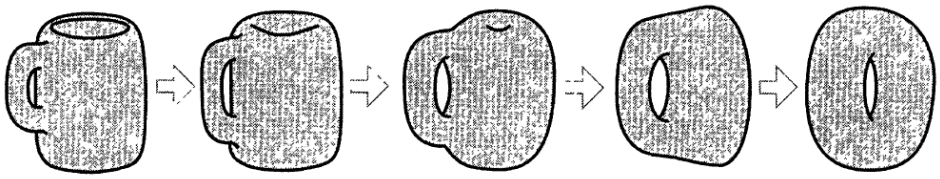


Figure A.3: ***Topological equivalence of a coffee cup and a donut [Adams and Franzosa, 2009].***

Example A.12. *In fact, all of the objects in figure A.4 are topologically equivalent to a coffee cup or donut and thus to each other.*

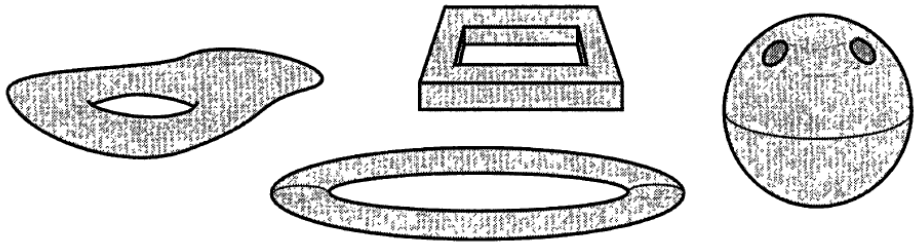


Figure A.4: ***All objects above are topologicall equivalent to a coffee cop or donut [Adams and Franzosa, 2009].***

Example A.13. *The objects in figure A.5 are not topologically equivalent. By continuous deformation it is impossible to change a sphere in a donut, as continues transformations can't create the central hole inside. Analogously, neither a sphere nor a donut can be transformed in a brezeln, which has two holes.*

As the saying goes: you can't always have your cake and eat it too. Going to topological spaces was motivated by the wish to look at certain properties of space without considering detailed values. But this has an important downside: almost all calculus techniques are lost in the process. Therefore, in physics in general and in this thesis we will mainly look at n-manifolds: topological spaces that locally resemble n-dimensional Euclidean space [Giblin, 2010]. For example, a 1-manifold locally resembles a line, a 2-manifold a plane, a 3-manifold 3D space, et cetera. Then for the full space we have the generality allowed by topology but locally our intuition and calculus techniques for n-dimensional Euclidean space can be used. To formalize previous we define an:

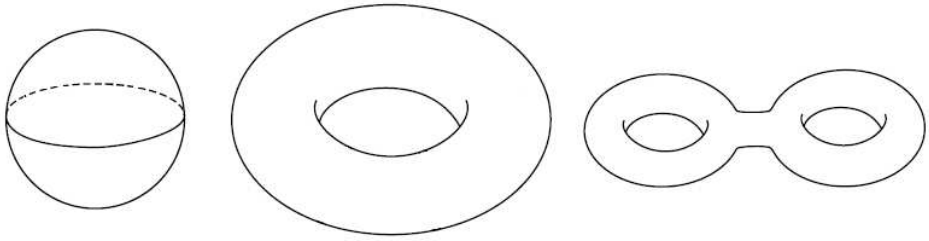


Figure A.5: ***Topologically inequivalent objects [Giblin, 2010].*** *The objects are topologically equivalent because they can't be continuously deformed in each other.*

Definition A.13 (n-manifold)**.** *An n-manifold M is a topological space*[1] *with a countable basis such that every point of M has a neighborhood that is homeomorphic to an open n-ball. A map ϕ between the neighborhood on the manifold and the open n-ball is called a* ***coordinate chart****.*

Actually, the idea is rather intuitive and a prime example we encounter in everyday life is the surface of the Earth. The 2D-surface of the Earth is spherical, but locally it seems flat and can be approximated by a plane. To give some more mathematical examples:

Example A.14. *Any surface of two variables in 3D Euclidean space is a 2-manifold. An example is given in figure A.6.*

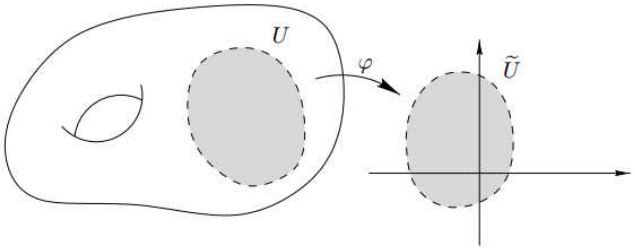


Figure A.6: The torus as an example of a 2D surface in 3D space [Lee, 2003]. The neighorhood U on the torus is homeomorphic to a disc, with φ indicating the coordinate chart.

Example A.15. *The n-sphere $\mathbb{S}^n$, the n-torus $\mathbb{R}^n$ and the real projective plane $\mathbb{R}P(n)$ are n-dimensional manifolds.*

Example A.16. *Let $F : \mathbb{R}^n \to \mathbb{R}$ be a continuously differentiable function and let c be a non-critical value of F. Then $F^{-1}(c)$ is a manifold of dimension $n-1$, if it is non-empty.*

The existence of local coordinates on manifolds allows local use of calculus. Global analysis is still out of reach if there are no differentiable coordinate transformations between the local patches. This leads to the following definition:

Definition A.14 (Smooth n-manifold)**.** *A smooth manifold is a manifold with a collection of charts that (i) cover the entire manifold and (ii) at their intersection are smooth i.e. infinitely differentiable.*

[1] Formally: a topological Hausdorff space, meaning that for every pair of points $p, q \in M$ there are disjoint open subsets $U, V \subset m$ such that $p \in U$ and $q \in V$. I.e. p and q are in some sense separable. In (astrophysical) practise, the Hausdorff condition is almost always satisfied.

Example A.17. *Following example A.14, figure A.7 shows the idea behind a smooth manifold. Two different neighborhoods U and V induce two different charts φ and ψ. If at their intersection φ and ψ are smooth and the entire torus is covered with neighborhoods inducing such charts, the manifold is smooth.*

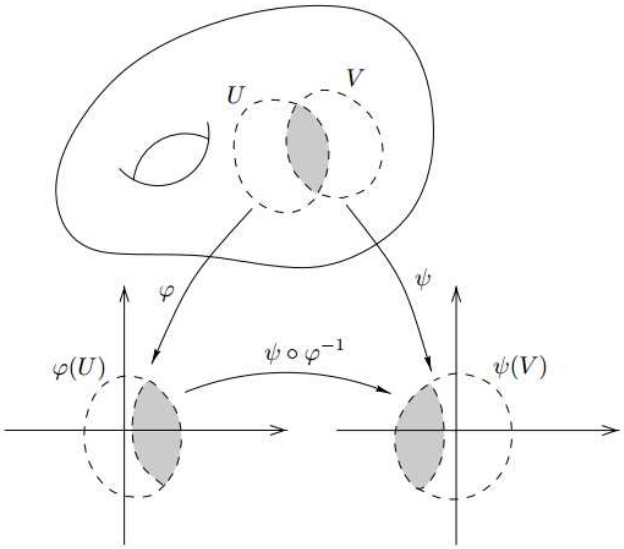


Figure A.7: An illustration of a smooth manifold [Lee, 2003]. Two neighorhoods U and V of the torus are shown, together with their charts φ and ψ. If the entire manifold is covered with such charts and at their intersection both are smooth, the manifold is a smooth manifold.

Sometimes, it is convenient to map properties of one smooth manifold to another. This can be done using a:

Definition A.15 (Diffeomorphism)**.** *A diffeomorphism is a function between two manifolds which is a bijective smooth map with a smooth inverse.*

Example A.18. *Obviously every diffeomorphism is a homeomorphism, but the inverse doesn't have to be true.*

Here we developed enough basic topology for our purposes below. Interested readers with a taste for more are referred to any of the references within this appendix.

A.2 Chains and homology in 10 minutes

The DMC provides us with a natural way to divide a point set (of galaxies) in a cosmic structure of voids, walls, filaments and nodes; providing us with a mathematical handle on these morphological components. Large galaxy catalogues contain millions of galaxies and their Morse complexes will be enormous. For such large complexes, a simple description of the DMC is not very insightful and a more systematic way to explore the structures of the DMC is required. Here homology [Giblin, 2010; Poincare, 1904] helps out by defining connectedness using equivalence relations. This appendix gives an elementary introduction to the field, focussed on homology on triangulated spaces. Homological equivalence relations are based on group theory, therefore

it was unavoidable to use some basic ideas from group theory. Readers unfamiliar with group theory are referred [Jones, 1998] for a physics focussed introduction to this field of mathematics.

We start with the building blocks of the DMC: simplices. To describe structures of simplices we need a way to describe large groups of them, which we can do using:

Definition A.16 (p-chains)**.** *Let K be a simplicial complex of dimension d. A p-chain c is a formal sum of p-simplices in K, written as*

$$c = \sum a_i \sigma_i \tag{A.1}$$

Here, σ_i are the simplices and a_i the coefficients. In principle the coefficients can be real numbers or even more general elements, but here we restrict them to modulo 2 integers: 0 (off) or 1 (on)[2]. This allows us to think of a chain as the set of p-simplices with coefficients $a_i = 1$. Naturally, for $p < 0$ or $p > d$ the p-chain consists only of the neutral element.

p-chains have intuitive properties:

Properties A.1.

1. *Just as with polynomials, addition is defined component-wise. Specifically, if $c = \sum a_i \sigma_i$ and $d = \sum b_i \sigma_i$ than*
$$c + d = \sum (a_i + b_i)\sigma_i \tag{A.2}$$
Because the coefficients are integers modulo 2, adding two chains actually gives their difference.

2. *The neutral element $0 = \sum 0\sigma_i$*

3. *Due to addition modulo 2 the inverse element of c is actually c itself, i.e. $-c = c$.*

To relate chains of different dimensions to each other we define the:

Definition A.17 (Boundary of a p-simplex)**.** *The boundary of a p-simplex is the sum of its facets. More formally, let $\sigma_p = [u_0, ..., u_p]$ be a p-simplex spanned by the vertices $u_0, ..., u_p$, then its boundary $\partial_p \sigma$ is defined as:*

$$\partial_p \sigma = \sum_{j=0}^{p} [u_0, ..., \hat{u}_j, ..., u_p] \tag{A.3}$$

with the hat indicating that vertex $\hat{u}_j$ is omitted.

Example A.19. *Consider the 1-simplex $\sigma = [u_0, u_1]$. Its its boundary is the 0-chain given by $\partial\sigma = u_1 - u_0$.*

Example A.20. *Consider the 2-simplex $\sigma = [v_0, v_1, v_2]$ as shown in figure A.8. Its boundary is the 1-chain given by $\partial\sigma = [v_0, v_1] + [v_1, v_2] + [v_2, v_0]$.*

[2] Restricting the coefficients to modulo 2 gives the convenient property that a p-chain is its own inverse, see below. In a more general setting, an simplex needs to be given an orientation, i.e. an 'arrow' indicating in which direction to add the simplices.

From previous definition follows directly:

Definition A.18 (Boundary of a p-chain)**.** *The boundary of a p-chain is the sum of the boundaries of its simplices. More formally, let $c = \sum a_i \sigma_i$ be a p-chain, than its boundary is*

$$\partial_p c = \sum a_i \partial_p \sigma_i \tag{A.4}$$

Clearly, the boundary maps a p-chain to a $(p-1)$-chain, so it can be thought of as a function $\partial_p : \mathsf{C}_p \to \mathsf{C}_{p-1}$. Taking the boundary commutes with addition, i.e. for two p-chains c and d we have

$$\partial_p(c+d) = \sum (a_i + b_i)\partial_p \sigma_p = \left(\sum a_i + \sum b_i\right) \partial_p \sigma_p = \partial_p c + \partial_p d$$

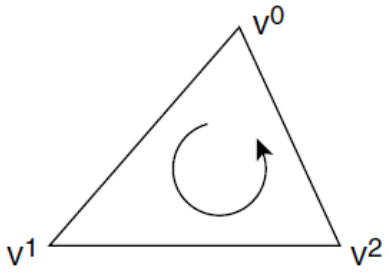


Figure A.8: ***The simplicial complex of example A.19 [Giblin, 2010].*** *A 2-simplex, with the arrow indicating the direction in which we take the boundary.*

From group theory we know this is the defining property of a homomorphism, so ∂_p is a homomorphism. We define the the boundary of a vertex to be 0.

Example A.21. *Consider the simplicial complex T consisting of two triangles t^1 and t^2 shown in figure A.9. We have:*

$$\partial_2(T) = \partial_2(t^1 + t^2) = (a + e + d) + (b + c - e) = a + b + c + d$$

This corresponds to the geometrical idea that the boundary of the simplicial complex are their outer segments. Taking the boundary of the diamond $(a + b + c + d)$, we get:

$$\partial_1(a + b + c + d) = (q - p) + (r - q) + (s - r) + (p - s) = 0$$

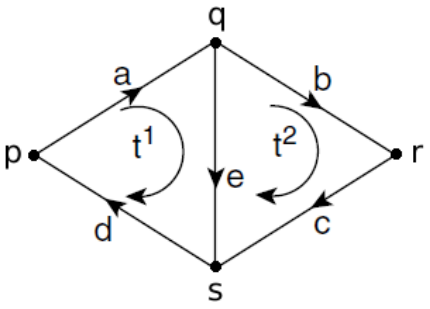


Figure A.9: ***The simplicial complex of example A.21 [Giblin, 2010].*** *A simplicial complex, with the arrows indicating the direction in which we take the boundary.*

Using the boundary operator, we can define a:

Definition A.19 (p-boundary)**.** *A p-boundary is a p-chain c that is the boundary of a $(p+1)$-chain d, i.e. $c = \partial d$ with $d \in \mathsf{C}_{p+1}$.*

Example A.22. *In previous example A.21, the diamond $(a + b + c + d)$ is a 1-boundary of simplicial complex T. The boundary of the diamond is zero.*

Chains with zero boundary turn out to play an important role in homology and have been given their own name:

Definition A.20 (p-cycle)**.** *A p-cycle c is a p-chain with empty boundary, i.e. $\partial c = 0$.*

Example A.23. *In previous example A.21 the 2-chain $t^1 + t^2$ is not a 2-cycle, as its boundary $\partial_2(t^1 + t^2) = a + b + c + d \neq 0$. The diamond $(a + b + c + d)$ is a one-cycle however, as $\partial_1(a + b + c + d) = 0$.*

In the examples above we sawe that applying the boundary operator twice gave zero, something which turns out to be true in general:

Lemma A.1 (Fundamental Lemma of Homology)**.** *$\partial_p \partial_{p+1} c = 0$ for every p and every $(p + 1)$-chain d.*

This lemmia is one of the fundamental properties that makes homology work. Given the importance of the lemma, the proof is amazingly simple.

Proof. The boundary $\partial_{p+1}c$ consists of all p-facets of c. Every $(p - 1)$-face of c belongs to exactly two p-facets, which cancel each other out. Therefore $\partial_p(\partial_{p+1}c) = 0$. □

Consider again the properties of p-chains: (i) the sum of two p-chains gives another p-chain; (ii) there is a neutral 'zero p-chain'; and (iii) every p-chain has an inverse p-chain (here actually being the chain itself). This are the defining properties of a group, showing that p-chains actually form the:

Definition A.21 (Group of p-chains $\mathsf{C}_p(K)$)**.** *The group $(\mathsf{C}_p(K), +)$ is the group of p-chains together with the addition operation. In notation, addition is usually understood implicitely. The group is associative because addition is associative and Abelian because addition modulo 2 is Abelian.*

p-boundaries are a special kind of chains, thus the set of p-boundaries $\mathsf{B}_p(K) \subseteq \mathsf{C}_p(K)$. Observing that ∂ commutes with addition, we can say that B_d is not just a subset of C_p but it is a subgroup:

Definition A.22 (Group of p-boundaries $\mathsf{B}_p(K)$)**.** *The subgroup of p-boundaries $\mathsf{B}_p(K) \subseteq \mathsf{C}_p(K)$ is the image of $(p + 1)$-st boundary homomorphism, i.e. $\mathsf{B}_p = im(\partial_{p+1})$. Since C_d is Abelian, so is B_p.*

Analogously the subset of p-cycles Z_p also form a subgroup:

Definition A.23 (Group of p-cycles $\mathsf{Z}_p(K)$)**.** *The subgroup of p-boundaries $\mathsf{Z}_p(K) \subseteq \mathsf{C}_p(K)$ is the kernel of the p-th boundary homomorphism, i.e. $\mathsf{B}_p = ker(\partial_{p+1})$. Since C_p is Abelian, so is Z_p.*

Different p-chain, p-cycle and p-boundary groups are related via the boundary operator and together form the ***chain complex***, which is illustrated in figure A.10.

The chain complex follows from a triangulated. It describes the structure but doesn't give direct information on general topological properties. But we can use the chain complex to extract such information. To intuitively understand why, we let go mathematical rigor for a moment and consider figure A.11 below. The 1-cycles a, a' and d bound an area of the surface which is shaded in grey, although it would be possible to use the complementary region as well. By continuously deforming the cycles without leaving the surface, it would be possible to shrink them to a point i.e. a 0-cycles. On the contrary, such a thing isn't possible with the 1-cycles b, b'

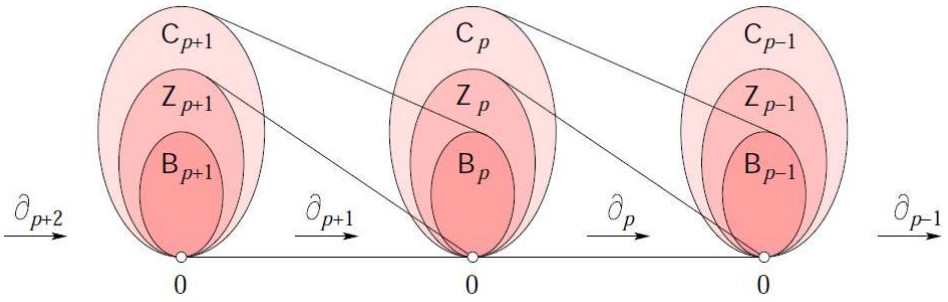


Figure A.10: ***The chain complex [Edelsbrunner and Harer, 2010].*** *It consists of chain, cycle and boundary groups connected via the boundary map.*

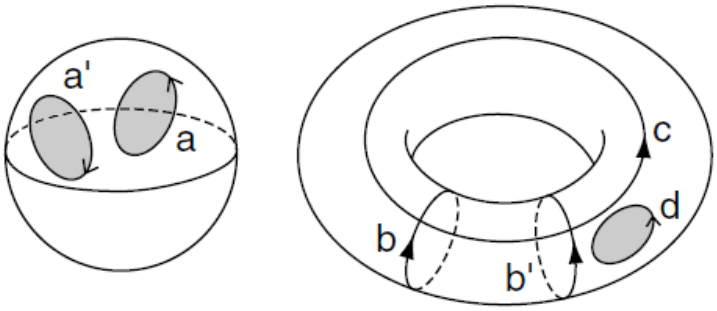


Figure A.11: ***A sphere and torus with several cycles, based upon [Giblin, 2010].***

and c, which reveals the presence of 2D tunnels. For c the tunnel is the center of the torus (the 'donut hole') and for b and b' the tunnel is the 'torus tube'.

By continuously moving the shrunken point a over the surface of the circle, we can let it coincide with shrunken point a' but not with shrunken point d, revealing that a and a' lie on the same surface and bind the same object, while d lies on another surface and binds another object. Because the 0-cycles a and a' bind the same object we can say they are in a certain way equivalent, whereas a and d bind another object and thus are not equivalent. I.e.: cycles which bind the same object are equivalent and the other way around: the number of non-equivalent cycles gives the number of different objects. This suggests that by looking at the quotient group $\mathsf{H}_0 = \mathsf{Z}_0/\mathsf{B}_0$, we can use 0-chains to reveal an intrinsic topological property: the number of different objects.

By continuous deformation and continuous movement over the surface of the torus, segment b can be transformed in segment b', showing that they form a boundary for the same tunnel. On the contrary, b cannot be continuously transformed into c showing that b and c don't bind the same tunnel. So in a certain manner, b and b' are equivalent and b and c are not. I.e: cycles which bind the same tunnel are equivalent and the other way around: the number of non-equivalent cycles gives the number of tunnels. Again, it seems the quotient group $\mathsf{H}_1 = \mathsf{Z}_1/\mathsf{B}_1$ allows us to use 1-chains to reveal an intrinsic topological property: the number of tunnels.

The above can be generalized to arbitrary dimensions: the p-th quotient group $\mathsf{H}_p = \mathsf{Z}_p/\mathsf{B}_p$ tells something about the number of p-dimensional tunnels also called ***holes***. An exception is $p = 0$, here it gives the amount of objects. Denoting 'A is generated by' with $A = \{\text{generating elements}\}$ we formalize previous by defining the:

Definition A.24 (Homology group H_p)**.** *The p-th homology group $\mathsf{H}_p = \mathsf{Z}_p/\mathsf{B}_p$, i.e. the p-th quotient group.*

If K consists of n simplices, the number of elements in $\mathsf{C}_p(K)$ is 2^n and its rank (the number of elements in the smallest generating set) is n. Both Z_p and B_p follow the same structure, giving us a concrete way to compute the rank of the p-th homology group, defined as its:

Definition A.25 (p-th Betti number β_p)**.**

$$\beta_p = \mathit{rank}(\mathsf{H}_p) = \mathit{rank}(\mathsf{Z}_p/\mathsf{B}_p) = \mathit{rank}(\mathsf{Z}_p) - \mathit{rank}(\mathsf{B}_p) \tag{A.5}$$

For $p = 0$ the Betti number gives the number of components and for $p > 0$ it gives the number of p-dimensional holes in the manifold under consideration.

Example A.24. *Consider again figure A.11 above. There are two separate components, thus $H_0 = \{a, d\}$ and $\beta_0 = 2$. The sphere doesn't have any tunnels but the donut has two seperate tunnels enclosed by respectively b and c, giving us: $H_1 = \{b, c\}$ and $\beta_1 = 2$. Note that the surface of the sphere S encloses its 3D volume, so S is actually the boundary of the volume of the sphere. Analogously, the surface of the torus T is the boundary of the volume of the torus. Thus $H_2 = \{S, T\}$ and $\beta_2 = 2$.*

Example A.25. *Figure A.12 shows a 2D manifold approximated by a simplicial complex. Following [Feldbrugge et al.] we compute the chain, cycle and boundary groups explicitely, followed by its homology groups. Directly from the figure, we can see that the chain groups are*

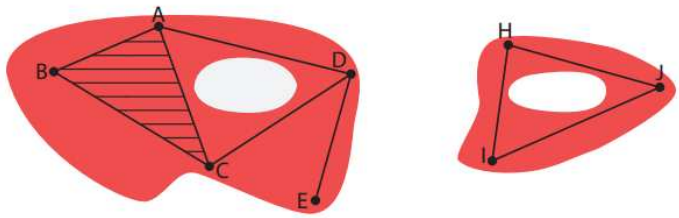


Figure A.12: ***A manifold approximated by a simplicial complex, based upon [Feldbrugge et al.].***

generated by:

$$\begin{aligned}
\mathsf{C}_0 &= \{A, B, C, D, E, I, J, H\} \\
\mathsf{C}_1 &= \{AB, AC, AD, BC, CD, DE, HI, IJ, JI\} \\
\mathsf{C}_2 &= \{ABC\}
\end{aligned}$$

Which of the chains are cycles? By definition any point has zero boundary and any closed loop of segments as well. For a 2D surface to have zero boundary it needs to enclose a 3D volume, obviously impossible in a 2D manifold. Thus we get:

$$\begin{aligned}
\mathsf{Z}_0 &= \mathsf{C}_0 \\
\mathsf{Z}_1 &= \{AB + BC + CA, AC + CD + DA, HI + IJ + JH\} \\
\mathsf{Z}_2 &= \{0\}
\end{aligned} \tag{A.6}$$

Any combination of two vertices which bind a segment form a 0-boundary and analogously any combination of three segments binding a triangle form a 1-boundary. Again, 2-boundaries need to bind a 3D volume which is impossible in a 2D manifold. This gives us:

$$\begin{aligned}\mathsf{B}_0 &= \{A+B, A+C, A+D, B+C, C+D, D+E, H+I, H+J.I+J\}\\ \mathsf{B}_1 &= \{AB+BC+CA\}\\ \mathsf{B}_2 &= \{0\}\end{aligned} \tag{A.7}$$

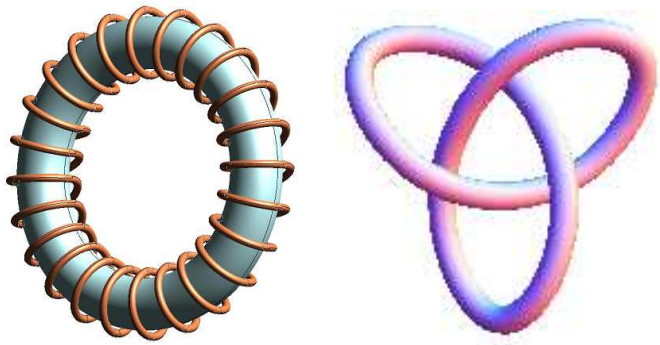
When two elements differ by an boundary element $b \in \mathsf{B}_i$, we can continuously transform an element over b to another element. When two elements don't differ by a boundary element there is a hole somewhere and a continuous transformation is not possible. Therefore, we partition all elements of Z_i in subsets which differ by an element of B_i. Obviously, all points in the same structure can be reached via boundaries. One dimension higher, $(AB+BC+CA)$ and $(AC+CD+DA)$ clearly differ by the first, itself a boundary element. Or more intuitively, the 1-cycles $(AC+CD+DA)$ and $(AB+BC+CD+DA) = (AB+BC+CA)+(AC+CD+DA)$ can be continuously transformed into each other. The above leads to:

$$\begin{aligned}\mathsf{H}_0 &= \{A, H\}\\ \mathsf{H}_1 &= \{AC+AD+CD, HI+HJ+IJ\}\\ \mathsf{H}_2 &= \{0\}\end{aligned}$$

Counting generating elements we see that the Betti numbers are given by $\beta_0 = 2$, $\beta_1 = 2$ and $\beta_2 = 0$. We could have seen this intuitively by observing that the manifold consists of two components with two 1D holes (tunnels).

Before we continue to the next section a note about mathematical rigor. In the intuitive explanation of homology classes and Betti numbers, we wrote things like "*by continuous deformation and continuous movement over the surface of the torus, segment b can be transformed in segment b' [and therefore b is homologous to b']*". Actually, mathematically this comes closer to a description of homotopy (cf. definition A.10) than of *homology*. Two cycles are homotopy equivalent if they can be continuously transformed in each other and homologous if there exists a manifold of higher dimension of which they are both the boundary. For smooth compact surfaces embedded in 3D Euclidean space the differences between homotopy and homology are limited. As we consider the first much more intuitive than the latter, we used a bit of homotopy in our intuitive description. Differences do occur, however. Consider for example a toroidal coil with n windings, as shown in the left part of figure A.13. On the torus, the coil is homotopy equivalent with $n \cdot b + c$ of the cycles of figure A.11 but the coil still has $\beta_1 = 1$. To see why, note that if you cut the torus along b or c, they both bind the torus surface. Analogously, if you cut the torus along the coil, the coil also binds the torus surface. In more complicated spaces the differences between homology and homotopy become much more important, in particular when it comes to measuring holes. For example homotopy misses the 2D hole given by the inside of the torus, whereas we saw above it was picked up by homology. On the contrary, homology doesn't detect the varying structures of complements on knots in 3D, whereas homotopy does [Weisstein, 2013a]. For $\mathbb{R}^d$ the intuitive picture of homology measuring d-holes is correct [Edelsbrunner and Harer, 2010, par IV.1].

If homotopy is more intuitive, why didn't we use this equivalence relation instead of homology? The most important reason is computational efficiency. Homology has compared with homotopy much faster algorithms [Edelsbrunner and Harer, 2010, chap. 4] allowing its practical use for cosmological datasets.



(a) *[Brandisky] A toroidal coil with the endpoints connected. Although for illustrative purposes the coil is lifted as a thick line above the torus surface, it should be seens as a 1-cycles wrapped around it.*

(b) *[Weisstein, 2013b] The trefoil knot is one of the simplests knots and the unique prime knot with three crossings. The homology on its complement doesn't detect the varying structure in the knott.*

Figure A.13: ***Examples where homotopy and homology differ.***

A.3 Formalizing persistence

In the main text persistence was defined intuitively. Here we formalize persistence using homology, building upon the definitions developed in appendix A.2.

Let f be a strictly increasing Morse function on simplex K with m simplices. Starting with an empty set, we add the simplices of K one by one in order increasing function value. This gives us the filtration (cf. definition 3.28):

$$\varnothing = K_0 \subseteq K_1 \subseteq ... \subseteq K_m = K$$

Each of the subcomplexes will give rise to several homology groups H_p (cf. definition A.24), whose ranks gave us the number of p-dimensional holes. To consider how the homology changes while the subcomplex increases in size, we can define an inclusion map from the underlying space $|K_i|$ to $|K_j|$, with $i \leq j$. The inclusion between the underlying spaces induces a homomorphism between the homology groups:

$$f_p^{i,j} : \mathsf{H}_p(K_i) \to \mathsf{H}_p(K_j) \tag{A.8}$$

Connecting these homomorphisms between all stages of the filtration, we get a sequence of homomorphisms:

$$0 = \mathsf{H}_p(K_0) \to \mathsf{H}_p(K_1) \to \ldots \to \mathsf{H}_p(K_m) = \mathsf{H}_p(K) \tag{A.9}$$

At each stage the homology might change because some new classes are born and some other classes become trivial or merge. For example, let γ be a cycle class which is born at stage i of the filtration and dies at stage j because it merges at j with an older component. Then the sequence of homomorphism and the life of γ can be illustrated as in figure A.14. (Note that merging of γ with an older component results in the death of γ is in accordance with the Elder Rule, cf. definition 3.30.) Collecting the classes that are born and which die at or before a certain stage in the filtration gives us the:

Definition A.26 (p-th persistent homology groups $\mathsf{H}_{i,j}$)**.** *Let K be a simplicial complex on which a filtration is defined and let $f_p^{i,j}$ be a homomorphism between the p-th homology group of stage i and stage j in the filtration. Then the p-th persistent homology groups $\mathsf{H}_{i,j}$ are the images of the homomorphisms induced by inclusion:*

$$\mathsf{H}_p^{i,j} = im\ f_p^{i,j} \tag{A.10}$$

*for $0 \leq i \leq j \leq n$. The corresponding p-**th persistent Betti numbers** $\beta_p^{i,j}$ are the ranks of these groups:*

$$\beta_p^{i,j} = rank\ \mathsf{H}_p^{i,j} \tag{A.11}$$

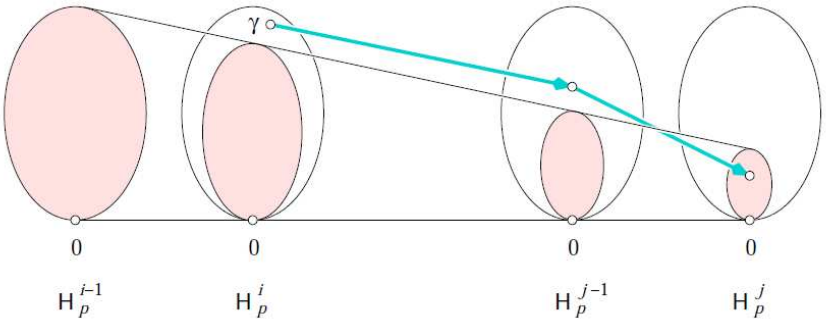


Figure A.14: ***Persistent homology groups [Edelsbrunner and Harer, 2010]****. The cycle class γ is born in subcomplex K_i. At stage j of the filtration its image merges for the first time into the image of $\mathsf{H}_p(K_{i-1})$. When this happens the class is paired and dies.*

Now we can define:

Definition A.27 (Persistence)**.** *Consider the setup of definition A.26. Denote for each stage in the filtration the function value of the simplex being added with a_i, this gives us a sequence of function values $a_1 < a_2 < \ldots < a_m$. We set the value corresponding to the empty set $a_0 = -\infty$. Then, if γ is born at K_i and dies at K_j, its persistence is:*

$$persistence\ (\gamma) = a_j - a_i \tag{A.12}$$

Intuitively, persistence is the difference between the birth and death value of the critical simplicies composing the cycle. I.e. persistence gives the lifetime of the cycle with respect to a rising sublevel set of the function.

BIBLIOGRAPHY

Erik A. Petigura, Andrew W. Howard, and Geoffrey W. Marcy. Prevalence of earth-size planets orbiting sun-like stars. *Proceedings of the National Academy of Sciences*, 110(48):19273–19278, 2013.

R. B. Tully. The local supercluster. *Astrophysical Journal*, 257:389 – 422, 1982.

M. Subbarao, D. Surendran, and R. Landsberg. Sloan galaxies dome 300, 2005. URL `http://astro.uchicago.edu/cosmus/projects/sloangalaxies`. Consulted at December 7, 2013.

R. Brown. Reference in memorial tribute to eric lenneberg. *Cognition*, 4:125 – 153, 1976.

J.A. Lucy. Linguistic relativity. *Annu. Rev. Anthropol.*, 26:291 – 312, 1997.

Aubrey L. Gilbert, Terry Regier, Paul Kay, and Richard B. Ivry. Whorf hypothesis is supported in the right visual field but not the left. *Proceedings of the National Academy of Sciences of the United States of America*, 103:489–494, 2006.

W.E. Schaap and R.v.d. Weygaert. Continuous fields an discrete samples: reconstruction through delaunay tesselations. *Astronomy and Astrophysics*, 363:L29 – L32, 2000.

N. Copernicus. *On the revolutions: Nicolas Copernicus Complete Works*. Foundations of Natural History. Johns Hopkins University Press. ISBN 978-0801845154. Translation and commentary by E. Rosen.

Bayerische Akademie der Wissenschaften. Kommission zur herausgabe der werke von johannes kepler. URL `http://www.kepler-kommission.de/index.html`. Consulted at December 10, 2013.

M. Hoskin. The cosmology of thomas wright of durham. *Journal for the History of Astronomy*, 1:44 – 52, 1970.

E. Hubble. A relation between distance and radial velocity among extra-galactic nebulae. *PNAS*, 15:168 – 173, 1929.

W. Baade. A revision of the extra-galactic distance scale. *Trans. IAU*, 8:397 – 398, 1952. [Reprinted with commentary in Lang, K. R. and Gingerich, O., eds., A Source Book in Astronomy and Astrophysics, 1900-1975 (Harvard Univ. Press, Cambridge, MA, 1979), 750-52.].

A.G. Lemaitre. Contributions to a british association discussion on the evolution of the universe. *Nature*, 128:704 – 706, 1931.

R. A. Alpher, H. Bethe, and G. Gamow. The origin of chemical elements. *Phys. Rev.*, 73: 803–804, 1948.

R. Dicke, P. Peebles, P. Roll, and D. Wilkinson. Cosmic black-body radiation. 142:414 – 419, 1965.

A. Penzias and R. Wilson. A measurement of excess antenna temperature at 4080 mc/s. *Astrophysical Journal*, 142:419 – 421, 1965.

Nobelprize.org. The nobel prize in physics 1978. 2013a. URL `http://www.nobelprize.org/nobel_prizes/physics/laureates/1978/`. Consulted at December 7, 2013.

JP Ostriker and Paul J Steinhart. The observational case for a low-density universe with a non-zero cosmological constant. *Nature*, 377:19, 1995.

R.v.d. Weygaert. Lecture notes cosmology, 2010. URL `http://www.intra.astro.rug.nl/~weygaert/cosmo2010.html`. University of Groningen.

B. Ryden. *Introduction to Cosmology*. Addison Wesley, 2003. ISBN ISBN 0-8053-8912-1.

A. Liddle. *An introduction to modern cosmology*. Wiley, 2nd edition, 2009. ISBN 978-0-470-84835-7.

R.v.d. Weygaert. Lecture notes large scale structure, 2012. URL `http://www.astro.rug.nl/~weygaert/lss2012.html`. University of Groningen.

J. Peacock. *Cosmological Physics*. Cambridge University Press, 2005. ISBN 0-521-42270-1.

Masataka Fukugita and P. J. E. Peebles. The cosmic energy inventory. *The Astrophysical Journal*, 616:643, 2004.

Gianfranco Bertone, Dan Hooper, and Joseph Silk. Particle dark matter: evidence, candidates and constraints. *Physics Reports*, 405:279 – 390, 2005.

S. Perlmutter, G. Aldering, G. Goldhaber, R. A. Knop, P. Nugent, P. G. Castro, S. Deustua, S. Fabbro, A. Goobar, D. E. Groom, I. M. Hook, A. G. Kim, M. Y. Kim, J. C. Lee, N. J. Nunes, R. Pain, C. R. Pennypacker, R. Quimby, C. Lidman, R. S. Ellis, M. Irwin, R. G. McMahon, P. Ruiz-Lapuente, N. Walton, B. Schaefer, B. J. Boyle, A. V. Filippenko, T. Matheson, A. S. Fruchter, N. Panagia, H. J. M. Newberg, W. J. Couch, and The Supernova Cosmology Project. Measurements of and from 42 high-redshift supernovae. *The Astrophysical Journal*, 517: 565, 1999.

Adam G Riess, Alexei V Filippenko, Peter Challis, Alejandro Clocchiatti, Alan Diercks, Peter M Garnavich, Ron L Gilliland, Craig J Hogan, Saurabh Jha, Robert P Kirshner, et al. Observational evidence from supernovae for an accelerating universe and a cosmological constant. *The Astronomical Journal*, 116:1009, 1998.

Nobelprize.org. The nobel prize in physics 2011. 2013b. URL `http://www.nobelprize.org/nobel_prizes/physics/laureates/2011/`. Consulted at December 7, 2013.

Planck Collaboration. *Astronomy and Astrophysics*, 2013. Submitted.

Eiichiro Komatsu, KM Smith, J Dunkley, CL Bennett, B Gold, G Hinshaw, N Jarosik, D Larson, MR Nolta, L Page, et al. Seven-year wilkinson microwave anisotropy probe (wmap) observations: cosmological interpretation. *The Astrophysical Journal Supplement Series*, 192:18, 2011.

Scott Dodelson, Vijay K. Narayanan, Max Tegmark, Ryan Scranton, Tamas Budavari, Andrew Connolly, Istvan Csabai, Daniel Eisenstein, Joshua A. Frieman, James E. Gunn, Lam Hui, Bhuvnesh Jain, David Johnston, Stephen Kent, Jon Loveday, Robert C. Nichol, Liam OConnell, Roman Scoccimarro, Ravi K. Sheth, Albert Stebbins, Michael A. Strauss, Alexander S. Szalay, Istvan Szapudi, Michael S. Vogeley, Idit Zehavi, James Annis, Neta A. Bahcall, Jon Brinkman, Mamoru Doi, Masataka Fukugita, Greg Hennessy, Zeljko Ivezic, Gillian R. Knapp, Peter Kunszt, Don Q. Lamb, Brian C. Lee, Robert H. Lupton, Jeffrey A. Munn, John Peoples, Jeffrey R. Pier, Constance Rockosi, David Schlegel, Christopher Stoughton, Douglas L. Tucker, Brian Yanny, Donald G. York, and (fortheSDSSCollaboration). The three-dimensional power spectrum from angular clustering of galaxies in early sloan digital sky survey data. *The Astrophysical Journal*, 572:140, 2002.

Will J. Percival, Carlton M. Baugh, Joss Bland-Hawthorn, Terry Bridges, Russell Cannon, Shaun Cole, Matthew Colless, Chris Collins, Warrick Couch, Gavin Dalton, Roberto De Propris, Simon P. Driver, George Efstathiou, Richard S. Ellis, Carlos S. Frenk, Karl Glazebrook, Carole Jackson, Ofer Lahav, Ian Lewis, Stuart Lumsden, Steve Maddox, Stephen Moody, Peder Norberg, John A. Peacock, Bruce A. Peterson, Will Sutherland, and Keith Taylor. The 2df galaxy redshift survey: the power spectrum and the matter content of the universe. *Monthly Notices of the Royal Astronomical Society*, 327:1297–1306, 2001.

Wendy L. Freedman, Barry F. Madore, Brad K. Gibson, Laura Ferrarese, Daniel D. Kelson, Shoko Sakai, Jeremy R. Mould, Jr. Robert C. Kennicutt, Holland C. Ford, John A. Graham, John P. Huchra, Shaun M. G. Hughes, Garth D. Illingworth, Lucas M. Macri, and Peter B. Stetson. Final results from the hubble space telescope key project to measure the hubble constant. *The Astrophysical Journal*, 553:47, 2001.

Robert R. Caldwell, Marc Kamionkowski, and Nevin N. Weinberg. Phantom energy: Dark energy with w smaller than -1 causes a cosmic doomsday. *Phys. Rev. Lett.*, 91:071301, 2003.

Robert R. Caldwell and Marc Kamionkowski. The physics of cosmic acceleration. *Annual Review of Nuclear and Particle Science*, 59:397–429, 2009.

Joshua A. Frieman, Michael S. Turner, and Dragan Huterer. Dark energy and the accelerating universe. *Annual Review of Astronomy and Astrophysics*, 46:385–432, 2008.

S. Weinberg. Cosmological constant problem. *Rev. Mod. Phys.*, Jan 1989.

Bharat Ratra and P. J. E. Peebles. Cosmological consequences of a rolling homogeneous scalar field. *Phys. Rev. D*, 37:3406–3427, 1988.

Philippe Brax and Jérôme Martin. Robustness of quintessence. *Phys. Rev. D*, 61:103502, 2000.

Daniel Z. Freedman, P. van Nieuwenhuizen, and S. Ferrara. Progress toward a theory of supergravity. *Phys. Rev. D*, 13:3214–3218, 1976.

Eric V. Linder. Exploring the expansion history of the universe. *Phys. Rev. Lett.*, 90:091301, 2003.

Eric V. Linder. Constraining models of dark energy. 2010.

E.G.P. Bos, R v.d. Weygaert, K. Dolag, and V. Pettorino. The darkness that shaped the void: dark energy and cosmic voids. *Monthly Notices of the Royal Astronomical Society*, 426:440 – 461, 2012.

A.E.P. Veldman and A. Velicka. Stromingsleer, 2010. URL `http://www.math.rug.nl/~veldman/Colleges/stromingsleer/Stromingsleer1011.pdf`. University of Groningen.

Wayne Hu and Scott Dodelson. Cosmic microwave background anisotropies. *Annual Review of Astronomy and Astrophysics*, 40:171–216, 2002.

NASA / WMAP Science Team. Nine year microwave sky. URL `http://map.gsfc.nasa.gov/`. Consulted at December 10, 2013.

Alan H. Guth. Inflationary universe: A possible solution to the horizon and flatness problems. *Phys. Rev. D*, 23:347–356, 1981.

Andrei D Linde. *Inflation and quantum cosmology*. Elsevier, 1990.

Ofer Lahav and Yasushi Suto. Measuring our universe from galaxy redshift surveys. *Living Reviews in Relativity*, 7, 2004.

Eric V. Linder. Cosmic growth history and expansion history. *Phys. Rev. D*, 72:043529, 2005.

F. Bernardeau, S. Colombi, E. Gaztaaga, and R. Scoccimarro. Large-scale structure of the universe and cosmological perturbation theory. *Physics Reports*, 367:1 – 248, 2002.

Y. Zel'dovich. Gravitational instability: An approximate theory for large density perturbations. *Astronomy and astrophysics*, 5:84–89, 1970.

J. Hidding. Adhesion - a sticky way of understanding large scale structure. Master's thesis, 2010. URL `http://www.astro.rug.nl/~hidding/go/report.pdf`.

S. Shandarin. Tessellating the universe: the zel'dovich and adhesion tiling of space. 2009. URL `http://arxiv.org/abs/0912.4520`.

W.H. Press and P. Schechter. Formation of galaxies and clusters of galaxies by self-similar gravitational condensation. *Astrophysical Journal*, 187:425 – 438, 1974.

JR Bond, S Cole, G Efstathiou, and Nick Kaiser. Excursion set mass functions for hierarchical gaussian fluctuations. *The Astrophysical Journal*, 379:440–460, 1991.

Galaxy Formation. *Galaxy formation*. Astronomy and Astrophysics Library. Springer, 2nd edition.

J.R. Bond, L. Kofman, and D. Pogosyan. How filaments of galaxies are woven into the cosmic web. *Nature*, 380:60 – 606, 1996.

Virgo consortium. Slices of the dark matter distribution. URL `http://www.mpa-garching.mpg.de/galform/virgo/millennium/`. Consulted at December 10, 2013.

Volker Springel, Simon D. M. White, Giuseppe Tormen, and Guinevere Kauffmann. Populating a cluster of galaxies i. results at. *Monthly Notices of the Royal Astronomical Society*, 328: 726–750, 2001.

Mark C. Neyrinck, Nickolay Y. Gnedin, and Andrew J. S. Hamilton. voboz: an almost-parameter-free halo-finding algorithm. *Monthly Notices of the Royal Astronomical Society*, 356:1222–1232, 2005.

M. Dries. A hierarchy of voids. Master's thesis, 2013. URL `http://www.astro.rug.nl/~dries/Thesis.pdf`.

Erwin Platen, Rien Van De Weygaert, and Bernard JT Jones. A cosmic watershed: the wvf void detection technique. *Monthly notices of the royal astronomical society*, 380:551–570, 2007.

Mark C. Neyrinck. zobov: a parameter-free void-finding algorithm. *Monthly Notices of the Royal Astronomical Society*, 386:2101–2109, 2008.

Marius Cautun, Rien van de Weygaert, and Bernard JT Jones. Nexus: tracing the cosmic web connection. *Monthly Notices of the Royal Astronomical Society*, 429:1286–1308, 2013.

Bos. E.G.P. Voids as probes of the nature of dark energy. Master's thesis, 2010. URL `http://egpbos.nl/`.

V.J. Martinez and E. Saar. *Statistics of the Galaxy Distribution*. Chapman and Hall/CRC, 2001. ISBN 978-1584880844.

Vicent J Martinez, Bernard JT Jones, Rosa Dominguez-Tenreiro, and Rien van de Weygaert. Clustering paradigms and multifractal measures. *The Astrophysical Journal*, 357:50–61, 1990.

Martin Kerscher, Jens Schmalzing, J&oumrg Retzlaff, Stefano Borgani, Thomas Buchert, Stephan Gottlöber, Volker Müller, Manolis Plionis, and Herbert Wagner. Minkowski functionals of abell/aco clusters. *Monthly Notices of the Royal Astronomical Society*, 284:73–84, 1997.

T.; Wagner H. Mecke, K. R.; Buchert. Robust morphological measures for large-scale structure in the universe. *Astronomy and Astrophysics*, 288:697 – 704, 1994.

J. Schmalzing, M. Kerscher, and T. Buchert. Minkowski functionals in cosmology. In *Proc. Int. School of Physics Enrico Fermi, Course CXXXII*, 1995.

Joshua Parker, Eilon Sherman, Matthias van de Raa, Devaraj van der Meer, Lawrence E. Samelson, and Wolfgang Losert. Automatic sorting of point pattern sets using minkowski functionals. *Phys. Rev. E*, 88:022720, 2013.

H. Edelsbrunner and J.L. Harer. *Computational Topology, an introduction*. American Mathematical Society, 2010.

T. Sousbie. The persistent cosmic web and its filamentary structure - i. theory and implementation. *Monthly Notices of the Royal Astronomical Society*, 414:350 – 383, 2011.

J. Milnor. *Morse Theory*. Princeton University Press, 5th edition, 1973. ISBN: 0-691-08008-9.

M. Morse. *The Calculus on Variations in the Large*, volume 18 of *Colloquium publications*. American Mathematical Society, 1960. ISBN: 9780821874554.

J. O'Connor and E. Robertson. Harold calvin marston morse, 2003. MacTutor History of Mathematics archive.

A. Okabe, B. Boots, K. Sugihara, and S. Chiu. *Spatial Tesselations: concepts and applications of Voronoi Diagrams*. John Wiley and Sons, 2nd edition, 2000. ISBN: 0-471-98635-6.

G. Voronoi. Nouvelles applications des parametres continus a la theorie des formes quadratiques. premier memoire. *Journal fr die reine und angewandte Mathematik*, 133:97 – 178, 1907.

G. Voronoi. Nouvelles applications des parametres continus a la theorie des formes quadratiques. deuxieme memoire. *Journal fr die reine und angewandte Mathematik*, 134:198 – 287, 1908.

B. Delone. Sur la sphere vide. *Izv. Akad. Nauk SSSR, Otdelenie Matematicheskii i Bestestvennyka Nauk*, 7:793 – 800, 1934.

G. Vegter. Lecture notes for applied geometry, 2012.

R. Forman. Morse theory for cell complexes. *Advances in Mathematics*, 134:90–145, 1998.

T. Lewiner. Constructing discrete morse functions. Master's thesis, PUC-Rio, Rio de Janeiro, 2002.

X. Ni, M. Garland, and J. Hart. Fair morse functions for extracting the topological structure of a surface mesh. *ACM Trans. Graph.*, 23:613–622, 2004.

Cecil Jose A. Delfinado and Herbert Edelsbrunner. An incremental algorithm for betti numbers of simplicial complexes. In *Proceedings of the Ninth Annual Symposium on Computational Geometry*, SCG '93, pages 232–239. ACM, 1993. ISBN 0-89791-582-8.

H. Edelsbrunner, D. Letscher, and A. Zomorodian. Topological persistence and simplification. *Discrete and computational geometry*, 28:511 – 533, 2002.

A. Gyulassy and Vijay Natarajan. Topology-based simplification for feature extraction from 3d scalar fields. In *Visualization, 2005. VIS 05. IEEE*, pages 535–542, 2005.

A. Gyulassy, Vijay Natarajan, V. Pascucci, P.-T. Bremer, and B. Hamann. A topological approach to simplification of three-dimensional scalar functions. *Visualization and Computer Graphics, IEEE Transactions on*, 12:474–484, 2006.

R.W. Hockney and J.W. Eastwood. *Computer simulations using particles*. McGraw-Hill, 1981. ISBN 0-07-029108-X.

J. Barnes and P. Hut. A hierarchical o(n log n) force-calculation algorithm. *Nature*, 324:446 – 449, 1986.

L. Greengard and V. Rokhlin. A fast algorithm for particle simulations. *Journal of Computational Physics*, 135:280 – 292, 1997.

V. Springel. The cosmological simulation code gadget-2. *Monthly Notices of the Royal Astronomical Society*, 364:1105 – 1134, 2005.

J.J. Monaghan. Sph and riemann solvers. *Journal of Computational Physics*, 136:298 – 307, 1997.

W.E. Schaap. *DTFE: the Delaunay Tessellation Field Estimator*. PhD thesis, University of Groningen, 2007.

F. Bernardeau and R.v.d. Weygaert. A new method for accurate estimation of velocity field statistics. *Monthly Notices of the Royal Astronomical Society*, 279:693 – 711, 1996.

F. I. Pelupessy, W.E. Schaap, and R.v.d. Weygaert. Density estimators in particle hydrodynamics - dtfe versus regular sph. *Astronomy and astrophysics*, 403:389 – 398, 2003.

R. vd. Weygaert and W. Schaap. The cosmic web: Geometric analysis. In Vicent J. Martinez, Enn Saar, Enrique Martinez Gonzales, and Maria Jesus Pons-Borderia, editors, *Data Analysis in Cosmology*, volume 665 of *Lecture Notes in Physics*, pages 291–413. Springer Berlin Heidelberg, 2009. ISBN 978-3-540-23972-7.

M. Cautun and R.v.d. Weygaert. The dtfe public software, 2011.

P. Bendich, H. Edelsbrunner, and M. Kerber. Computing robustness and persistence for images. *Visualization and Computer Graphics, IEEE Transactions on*, 16:1251–1260, 2010.

P. Pranav, R. vd. Weygaert, H. Edelsbrunner, J. Feldbrugge, M. Kerber, B. Jones, and G. Vegter. Persistence landscapes of gaussian random fields. *Monthly Notices of the Royal Astronomical Society*, 2013a. To be published.

P. Pranav, H. Edelsbrunner, R. vd. Weygaert, M. Kerber, B. Jones, G. Vegter, and M. Wintraecken. On the betti of the universe, and her persistence. *Monthly Notices of the Royal Astronomical Society*, 2013b. To be published.

Chao Chen and Michael Kerber. Persistent homology computation with a twist. In *Proceedings 27th European Workshop on Computational Geometry*, 2011.

V. De Silva, D. Morozov, and M. Vejdemo-Johansson. Dualities in persistent (co)homology. *Inverse Problems*, 27:124003, 2011.

U. Bauer, M. Kerber, and J. Reininghaus. Clear and compress: computing persistent homology in chuncks. 2013.

W. Press, S. Teukolsky, W. Vetterling, and B. Flannery. *Numerical recipes 3rd edition: the art of scientific computing*. Cambridge University Press, 3rd edition, 2007. ISBN 978-0521880688.

D. Cohen-Steiner, H. Edelsbrunner, and J. Harer. Stability of persistence diagrams. *Discrete and Computational Geometry*, 37:102 – 120, 2007.

D. Cohen-Steiner, H. Edelsbrunner, J. Harer, and Y. Mileyko. Lipschitz functions have lp-stable persistence. *Foundations of Computational Mathematics*, 10:127–139, 2010.

T. Lindeberg. Scale-space for discrete signals. *IEEE Transactions of Pattern Analysis and Machine Intelligence*, 12:234–254, 1990.

J. Feldbrugge, M. Van Engelen, R.v.d. Weygaert, P. Pranav, and G. Vegter. Stochastic homology of random fields using betti numbers and persistence diagrams. To be published in Monthly Notices of the Royal Astronomical Society, 2013.

P. Coles and Jones. B. A lognormal model for the cosmological mass distribution. *Monthly Notices of the Royal Astronomical Society*, 248:1–13, 1991.

E. Feigelson and Babu. G. *Modern Statistical Methods for Astronomy*. Cambridge University Press, 2012. ISBN 9780521767279.

A.J.S. Hamilton. Linear redshift distortions: A review. In Donald Hamilton, editor, *The Evolving Universe*, volume 231 of *Astrophysics and Space Science Library*, pages 185–275. Springer Netherlands, 1998. ISBN 978-0-7923-5074-3.

J. Wall and C Jenkins. *Practical statistics for astronomers*. Cambridge University Press, 2003. ISBN 0-521-45616-9.

Mark A. Heald. On choosing the bin width of a gaussian histogram. *American Journal of Physics*, 52:254–255, 1984.

P. Pranav, 10 2013. Personal communication.

A. Robinson and K. Turner. Hypothesis testing for topological data analysis. 2013.

K. Turner. Means and medians of sets of persistence diagrams. 2013.

Yehuda Hoffman, Ofer Metuki, Gustavo Yepes, Stefan Gottlber, Jaime E. Forero-Romero, Noam I. Libeskind, and Alexander Knebe. A kinematic classification of the cosmic web. *Monthly Notices of the Royal Astronomical Society*, 425:2049–2057, 2012.

P. Pranav, R.v.d. Weygaert, H. Edelsbrunner, J. Feldbrugge, M. Kerber, B.J.T. Jones, and G. Vegter. Persistence landscapes of gaussian random fields. *Monthly Notices of the Royal Astronomical Society*, 2013c. to be published.

K.J. Nevenzeel. Triangulating the darkness. *Periodiek*, 2, 2014. To be published.

A. Katok and A. Sossinsky. Introduction to modern topology and geometry. Available via http://www.personal.psu.edu/axk29/TOPOLOGY/.

W. Sutherland. *Introduction to Metric and Topological Spaces*. Oxford Science Publications, 2005. ISBN 978-0-19-853161-6.

C. Adams and R. Franzosa. *Introduction to topology, pure and applied*. Pearson Prentice Hall, indian edition, 2009. ISBN: 978-81-317-2692-1.

P. Giblin. *Graphs, surfaces and homology*. Cambridge University Press, 3th edition, 2010. ISBN: 978-0-521-15405-5.

J. Lee. *Introduction to smooth manifolds*, volume 218 of *Graduate texts in mathematics*. Springer, 2003. ISBN: 0-387-95448-1.

H. Poincare. Cinquieme complement a l'analysis situs. *Rendiconti del Circolo Matematico di Palermo*, 18:45 – 110, 1904.

H.F. Jones. *Groups, Representations and Physics*. Taylor and Francis, 2nd edition, 1998. ISBN 978-0750305044.

E. Weisstein. Hole, 2013a. From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/Hole.html.

Brandisky. Coil on torus. Retrieved from http://fa.tu-sofia.bg/te/Brandisky/research.html at November 11, 2013.

E. Weisstein. Trefoil knot, 2013b. From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/TrefoilKnot.html.