UNIVERSITY
OF OSLO

Master's thesis

-sometitle-

Classifying N-body simulations with and without relativistic corrections using machine learning techniques

Johan Mylius Kroken

Computational Science: Astrophysics
60 ECTS study points

Institute of Theoretical Astrophysics
Faculty of Mathematics and Natural Sciences

Autumn 2023

Johan Mylius Kroken

-sometitle-

Classifying N-body simulations with and without relativistic corrections using machine learning techniques

Supervisors:
A David Fonseca Mota
B Julian Adamek
C Francisco Antonio Villaescusa Navarro

Abstract

On large scales, comparable to the horizon, various relativistic effect will affect the observable clustering properties of galaxies. In order to solve for these effects, one need to constantly solve for the metric, velocities and densities in a particular gauge. When simulating large-scale structures we often use N-body simulations, usually performed in the Newtonian limit. However, it is not obvious that Newtonian gravity yield a good global description of an in-homogeneous cosmology when there is significant nonlinear dynamical behaviour (Jeong, Fabian Schmidt and Hirata 2012). Literature results suggest that the relativistic corrections necessary on top of realistic Newtonian cosmologies should be very small (Chisari and Zaldarriaga 2011). If this is the case, then this justifies the use of Newtonian simulations even on scales larger than the Hubble radius, whose results may be translated into relativistic cosmologies using relevant dictionaries (Green and Wald 2012).

I investigate this by running 2000 simulations with the relativistic N-body code `gevolution` by Adamek et al. 2016, with and without relativistic effects, using identical ΛCDM cosmologies. The simulations are run on a 256^3 grid each with dimension 5120 Mpc/h in order to capture the Hubble horizon. I investigate the difference between the gravities by considering the power spectra and bispectra of the gravitational potential Φ, the latter should reveal the nonlinear dynamical behaviour present in the relativistic simulation.

The whole dataset is used to train a Convolutional Neural Network (CNN), aiming to classify the two cases. If successful, the CNN may be used to analyse and understand the features separating the relativistic and Newtonian simulations. This is done through interpretability of the CNN by consider for instance saliency maps (Alqaraawi et al. 2020) or Grad-CAM (Selvaraju et al. 2020).

Contents

1 Introduction . . . 1
  1.1 Motivation . . . 1
  1.2 Outline . . . 1
  1.3 Aim . . . 1
  1.4 Nomenclature . . . 1

I Cosmological Structure Formation **3**

2 Preliminaries . . . 5
  2.1 General Relativity . . . 5
    2.1.1 Einstein's Field Equations . . . 5
    2.1.2 Riemann Connection and Covariant Derivatives . . . 5
    2.1.3 Geodesic Equation . . . 5
    2.1.4 The Stress-Energy Tensor . . . 5
  2.2 Useful Relations . . . 5
3 Background Cosmology . . . 7
  3.1 The homogeneous Universe . . . 7
    3.1.1 The Cosmological Principle . . . 7
    3.1.2 The Robertson-Walker Metric . . . 7
    3.1.3 The Friedmann Equations . . . 7
  3.2 My Universe is loaded with... . . . 7
  3.3 Thermal History of the Universe . . . 7
4 Perturbation Theory . . . 9
  4.1 Initial Conditions . . . 9
  4.2 Transfer Functions . . . 9
  4.3 Power Spectra . . . 9
  4.4 Non-linear Evolution . . . 9
  4.5 Bispectra . . . 9
5 Simulation theory . . . 11
  5.1 N-body simulations . . . 11
    5.1.1 Describing a box of particles . . . 11
    5.1.2 Forces and Fields . . . 11
    5.1.3 Mass Assignment Schemes . . . 11
    5.1.4 Validity of Box . . . 11
  5.2 Newtonian Approach . . . 11
  5.3 General Relativistic Approach . . . 11

II Machine Learning **13**
6 Fundamental Elements of Machine Learning . . . 15
6.1 Introduction . . . 15
6.2 Linear Algebra . . . 15
6.3 Probability and Information Theory . . . 15
6.4 Basic Machine Learning . . . 15
6.4.1 Estimators, Bias, Variance and Error . . . 15
6.4.2 Maximum Likelihood Estimation . . . 16
6.4.3 Bayesian Statistics . . . 16
6.4.4 Supervised Learning . . . 16
6.4.5 Unsupervised Learning . . . 16
7 Neural Networks . . . 17
7.1 Forward pass - Prediction . . . 17
7.1.1 Activation functions . . . 17
7.1.2 Loss functions. . . . 17
7.2 Backpropagation - Training . . . 17
7.2.1 Gradient descent . . . 17
7.2.2 Optimizers . . . 17
7.2.3 Regularization . . . 17
8 Convolutional Neural Networks . . . 19
8.1 Convolution . . . 19
8.2 New Layers . . . 19
8.2.1 Convolutional layers . . . 19
8.2.2 Pooling layers . . . 19
8.2.3 Dropout layers . . . 19

III Acquiring Data **21**
9 Simulations. . . . 23
9.1 Parameters . . . 23
9.1.1 Cosmological parameters . . . 23
9.1.2 Primordial power spectrum . . . 23
9.1.3 Box parameters . . . 24
9.1.4 Seeds . . . 24
9.1.5 Output . . . 24
10 Data Verification . . . 25
10.1 Slices of Datacubes . . . 25
10.2 Powerspectra from Simulations . . . 25
10.3 Powerspectra from Datacubes . . . 25
10.4 Analytical Bispectra . . . 25
10.5 Bispectra from Cube . . . 25
10.5.1 Binning . . . 25
10.5.2 Bispectra . . . 25
11 Trainable Dataset . . . 33

List of Figures

10.1 Slice 0 26
10.2 Average matter power spectra at different redshifts. 27
10.3 Average potential power spectra at different redshifts. 28
10.4 Angles in arbitrary bispectrum triangle configuration. 29

List of Tables

9.1 Cosmological parameters 23
9.2 Primordial power spectra parameters 23
9.3 Box parameters 24

Preface

Here comes your preface, including acknowledgments and thanks.

Chapter 1

Introduction

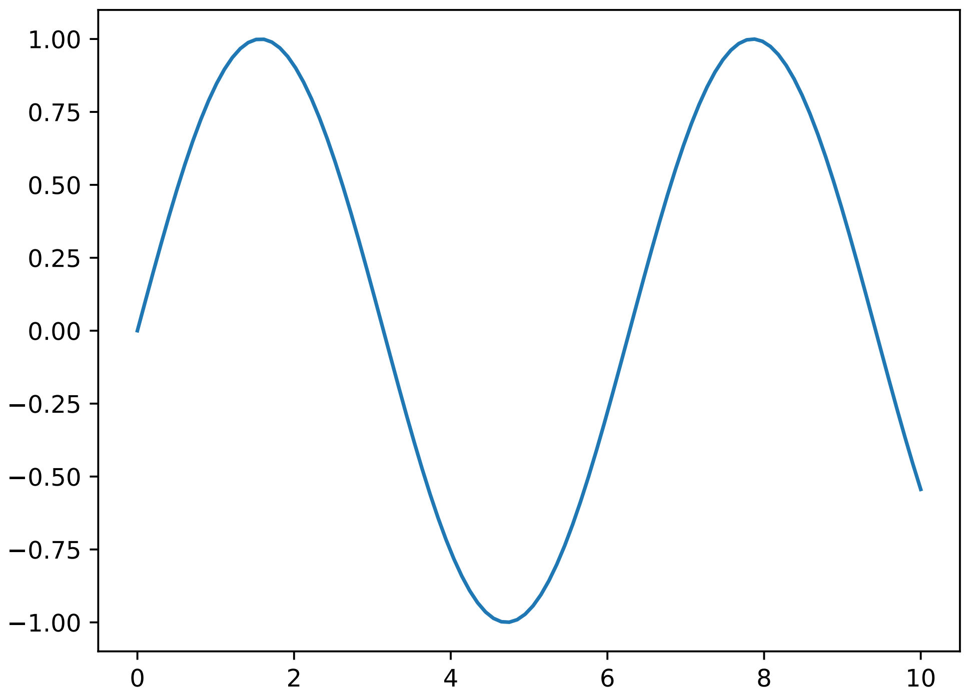
This is the introduction that will shortly be written. How fast does things change.

1.1 Motivation

1.2 Outline

1.3 Aim

1.4 Nomenclature


1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
0
2
4
6
8
10

Part I

Cosmological Structure Formation

Chapter 2

Preliminaries

2.1 General Relativity

2.1.1 Einstein's Field Equations

$$G_{\mu\nu} = 8\pi G T_{\mu\nu} \tag{2.1}$$

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R \tag{2.2}$$

$$R = g^{\mu\nu} R_{\mu\nu} \tag{2.3}$$

$$R_{\mu\nu} = \partial_\rho \Gamma^\rho_{\mu\nu} - \partial_\nu \Gamma^\rho_{\mu\rho} + \Gamma^\rho_{\mu\nu} \Gamma^\sigma_{\rho\sigma} - \Gamma^\rho_{\mu\sigma} \Gamma^\sigma_{\nu\rho} \tag{2.4}$$

2.1.2 Riemann Connection and Covariant Derivatives

$$\Gamma^\rho_{\mu\nu} = \frac{1}{2} g^{\rho\sigma} \left(\partial_\mu g_{\nu\sigma} + \partial_\nu g_{\mu\sigma} - \partial_\sigma g_{\mu\nu} \right) \tag{2.5}$$

$$\nabla_\mu T^\mu_\nu = \partial_\mu T^\mu_\nu + \Gamma^\mu_{\mu\alpha} T^\alpha_\nu - \Gamma^\alpha_{\mu\nu} T^\mu_\alpha \tag{2.6}$$

2.1.3 Geodesic Equation

$$\frac{d^2 x^\mu}{d\tau^2} + \Gamma^\mu_{\alpha\beta} \frac{dx^\alpha}{d\tau} \frac{dx^\beta}{d\tau} = 0 \tag{2.7}$$

2.1.4 The Stress-Energy Tensor

$$T_{\mu\nu} = (\rho + p) u_\mu u_\nu + p g_{\mu\nu} \tag{2.8}$$

2.2 Useful Relations

Chapter 3

Background Cosmology

3.1 The homogeneous Universe

In this chapter I will focus on explaining the background cosmology in light of a homogeneous universe. A natural place to start is the cosmological principle, followed by a description of the geometry of space itself. If not otherwise stated, the development of this chapter is based on Dodelson and F. Schmidt 2020, Weinberg 2008 and TODO: cite Baumann

3.1.1 The Cosmological Principle

3.1.2 The Robertson-Walker Metric

$$ds^2 = -dt^2 + a^2(t)\left[\frac{dr^2}{1-kr^2} + r^2 d\Omega^2\right] \tag{3.1}$$

3.1.3 The Friedmann Equations

3.2 My Universe is loaded with...

3.3 Thermal History of the Universe

Chapter 4

Perturbation Theory

4.1 Initial Conditions

4.2 Transfer Functions

4.3 Power Spectra

4.4 Non-linear Evolution

4.5 Bispectra

The bispectra are powerful tools for studying the non-linear evolution of the density field. The bispectrum is defined as the Fourier transform of the three-point correlation function, and is given by:

$$\langle \delta(\mathbf{k}_1)\delta(\mathbf{k}_2)\delta(\mathbf{k}_3) \rangle = (2\pi)^3 \delta_D(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) B(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) \tag{4.1}$$

Well this is rather awkward. Adamek et al. 2016 or (Falck et al. 2017)

Chapter 5

Simulation theory

Some theory and history as to how to conduct N-body simulations.

5.1 N-body simulations

5.1.1 Describing a box of particles

5.1.2 Forces and Fields

5.1.3 Mass Assignment Schemes

5.1.4 Validity of Box

5.2 Newtonian Approach

5.3 General Relativistic Approach

Part II

Machine Learning

Chapter 6

Fundamental Elements of Machine Learning

6.1 Introduction

In this chapter I will give a brief introduction into machine learning. This includes a mathematical description of some fundamental concepts common across numerous machine learning models. The more advanced models will be dealt with at a later stage. If not otherwise stated, the following chapter is based on Goodfellow, Bengio and Courville 2016 and Hastie, Tibshirani and Friedman 2009.

6.2 Linear Algebra

maybe

6.3 Probability and Information Theory

maybe

6.4 Basic Machine Learning

TODO: Fill more here

6.4.1 Estimators, Bias, Variance and Error

Estimators Based on the assumption that there exists some true parameter(s) $\boldsymbol{\theta}$ which remain unknown,[1] we are able to make predictions and estimations of such parameter(s). Let's say we have m independent and identically distributed (i.i.d.) random variables $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$ drawn from the same probability distribution $p(\mathbf{x})$. An *estimator* of the true values $\boldsymbol{\theta}$ is any function of the data such that $\hat{\boldsymbol{\theta}}_m = g(\mathbf{x}_1, \ldots, \mathbf{x}_m)$, where $\hat{\boldsymbol{\theta}}$ is the estimate of $\boldsymbol{\theta}$. This is known as point estimation, as we are estimating a single value. This definition does not pose any restrictions on the function g. However, a good estimator would yield an estimate $\hat{\boldsymbol{\theta}}_m$ that is close to the true value $\boldsymbol{\theta}$.

[1] This is the frequentist perspective of statistics

Function estimators Say we want to predict a variable $\mathbf{y}$ given some vector $\mathbf{x}$. We assume the true variable $\mathbf{y}$ is given by some function approxiamation $f(\mathbf{x})$ plus some error $\boldsymbol{\epsilon}$: $\mathbf{y} = f(\mathbf{x}) + \boldsymbol{\epsilon}$. The aim is then to estimate the function f with the estimator $\hat{f}$. If we then realise that $\hat{f}$ is really just a point estimator in function space, the two above concepts are equivalent.

Bias The bias of the estimator $\hat{\boldsymbol{\theta}}_m$ is defined as the difference between the expected value of the estimator and the true value of the parameter: $\text{bias}(\hat{\boldsymbol{\theta}}_m) = \mathbb{E}[\hat{\boldsymbol{\theta}}_m] - \boldsymbol{\theta}$. An unbiased estimator has zero bias, i.e. $\mathbb{E}[\hat{\boldsymbol{\theta}}_m] = \boldsymbol{\theta}$. An estimator is asymptotically unbiased if its bias approaches zero as the number of data points m approaches infinity, i.e. $\lim_{m\to\infty} \mathbb{E}[\hat{\boldsymbol{\theta}}_m] = \boldsymbol{\theta}$.

Variance

Standard Error

Mean Squared Error

6.4.2 Maximum Likelihood Estimation

6.4.3 Bayesian Statistics

6.4.4 Supervised Learning

6.4.5 Unsupervised Learning

Chapter 7

Neural Networks

7.1 Forward pass - Prediction

7.1.1 Activation functions

7.1.2 Loss functions

7.2 Backpropagation - Training

7.2.1 Gradient descent

7.2.2 Optimizers

7.2.3 Regularization

Chapter 8

Convolutional Neural Networks

8.1 Convolution

8.2 New Layers

8.2.1 Convolutional layers

8.2.2 Pooling layers

8.2.3 Dropout layers

Part III

Acquiring Data

Chapter 9

Simulations

9.1 Parameters

When performing simulations, it was import to keep all parameters fixed for all the different simulations. The only thing that was changed was the random seed.

9.1.1 Cosmological parameters

The relevant cosmological parameters are the dimensionless Hubble factor h, the baryon and cold dark matter densities Ω_b and Ω_{CDM}, the Cosmic Microwave Background temperature T_{CMB} and the effective number of ultra-relativistic neutrinos N_{ur}.

Table 9.1: Cosmological parameters

Parameter	Value	Unit
h	0.67556	-
Ω_b	0.022032	-
Ω_{CDM}	0.12038	-
T_{CMB}	2.7255	K
N_{ur}	3.046	-

9.1.2 Primordial power spectrum

The primordial power spectrum, as TODO: link to when written,contains the pivot scale k_{piv}, the primordial amplitude, $\mathcal{A}_{\mathrm{s}}$ and the spectral index, n_{s}.

Table 9.2: Primordial power spectra parameters

Parameter	Value	Unit
k_{piv}	0.05	Mpc^{-1}
$\mathcal{A}_{\mathrm{s}}$	$2.215 \cdot 10^{-9}$	-
n_{s}	0.9619	-

9.1.3 Box parameters

The relevant box parameters were the initial redshift z_{ini} where the simulations were started from. The simulations box itself was characterised by the physical length L, represented on a cube grid of size N_{grid}^3, resulting in a resolution of $\Delta_{\mathrm{res}} = L/N_{\mathrm{grid}}$. The courant factor TODO: fill and time step limit TODO: fill. This resulted in a fundamental

Table 9.3: Box parameters

Parameter	Value	Unit
z_{ini}	100	
L	5120	Mpc
N_{grid}	256	px
Δ_{res}	$20(= L/N_{\mathrm{grid}})$	Mpc px^{-1}
Courant factor	48	?
Time step limit	0.04	?

frequency of $k_{\mathrm{F}} = 2\pi/L$ and Nyquist freguency $k_{\mathrm{N}} = \pi/\Delta_{\mathrm{res}}$ TODO: what with units?

9.1.4 Seeds

In order to initialise the simulations we used random seeds, one for each simulation. This ensured that analysis performed on different simulations were of different realisations of the simulated universe, essential statistical independence. The seeds, denoted as $\mathtt{S}$ ranged from 0 to 2000, and consisted of the following set:

$$\{\mathtt{S} \in \mathbb{Z} | 0 \leq \mathtt{S} < 2000\} \tag{9.1}$$

9.1.5 Output

$$z_{\mathrm{d}} = \{20, 15, 10, 5, 1, 0\} \tag{9.2}$$

$$z_{\mathrm{p}} = \{100, 50, 20, 15, 10, 6, 5, 4, 3, 2, 1, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2, 0.1, 0\} \tag{9.3}$$

$$\mathcal{D}(\mathtt{S}, z_{\mathrm{d}}) \tag{9.4}$$

$$\mathcal{D}_{\Phi}(\mathtt{S}, z_{\mathrm{p}}) \tag{9.5}$$

$$\mathcal{D}_{\delta}(\mathtt{S}, z_{\mathrm{p}}) \tag{9.6}$$

Chapter 10

Data Verification

10.1 Slices of Datacubes

10.2 Powerspectra from Simulations

10.3 Powerspectra from Datacubes

10.4 Analytical Bispectra

$$B^{(3)}(k_1, k_2, k_3) = 2\mathcal{P}(k_1)\mathcal{P}(k_2)F_2(\mathbf{k}_1, \mathbf{k}_2) + \text{cyc} \tag{10.1}$$

$$F_2(\mathbf{k}_1, \mathbf{k}_2) = \frac{5}{7} + \frac{x}{2}\left(\frac{k_1}{k_2} + \frac{k_2}{k_1}\right) + \frac{2}{7}x^2, \tag{10.2}$$

where $x = \hat{\mathbf{k}}_1 \cdot \hat{\mathbf{k}}_2 = \cos\theta_{12}$, where θ_{12} is the angle spanned by $\mathbf{k}_1$ and $\mathbf{k}_2$. We could thus consequently write: $F_2(\mathbf{k}_1, \mathbf{k_2}) = F_2(k_1, k_2, \theta_{12})$

Given k_1 and k_2 and θ_{12} we have the following relations, with reference to Section 10.4:

$$\begin{aligned} \alpha &= \pi - \theta_{12} \\ \beta &= \pi - \theta_{23} \\ \gamma &= \pi - \theta_{31} \end{aligned} \tag{10.3}$$

From cosine rule:

$$k_3 = \sqrt{k_1^2 + k_2^2 - 2k_1k_2\cos\alpha} \tag{10.4}$$

From the rule of sines TODO: explain more?:

$$\begin{aligned} \beta &= \arcsin\left(\frac{k_1}{k_3}\sin\alpha\right) \\ \gamma &= \arcsin\left(\frac{k_2}{k_3}\sin\alpha\right) \end{aligned} \tag{10.5}$$

10.5 Bispectra from Cube

10.5.1 Binning

10.5.2 Bispectra

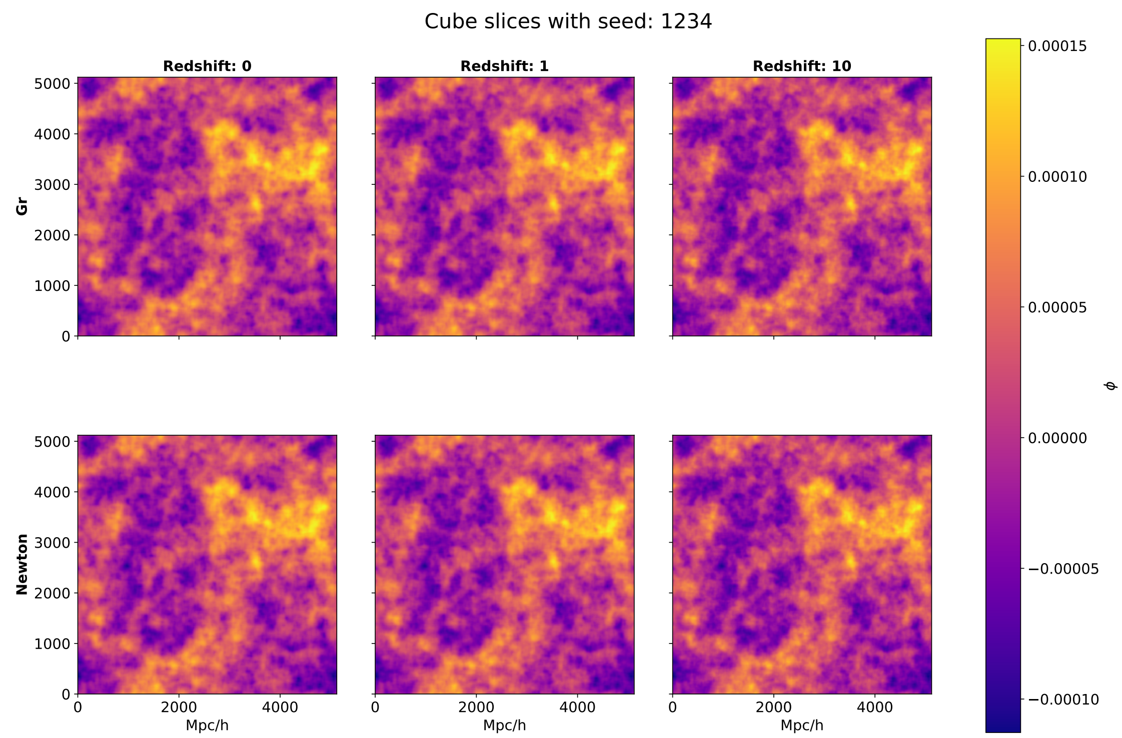


Figure 10.1: Slice 0

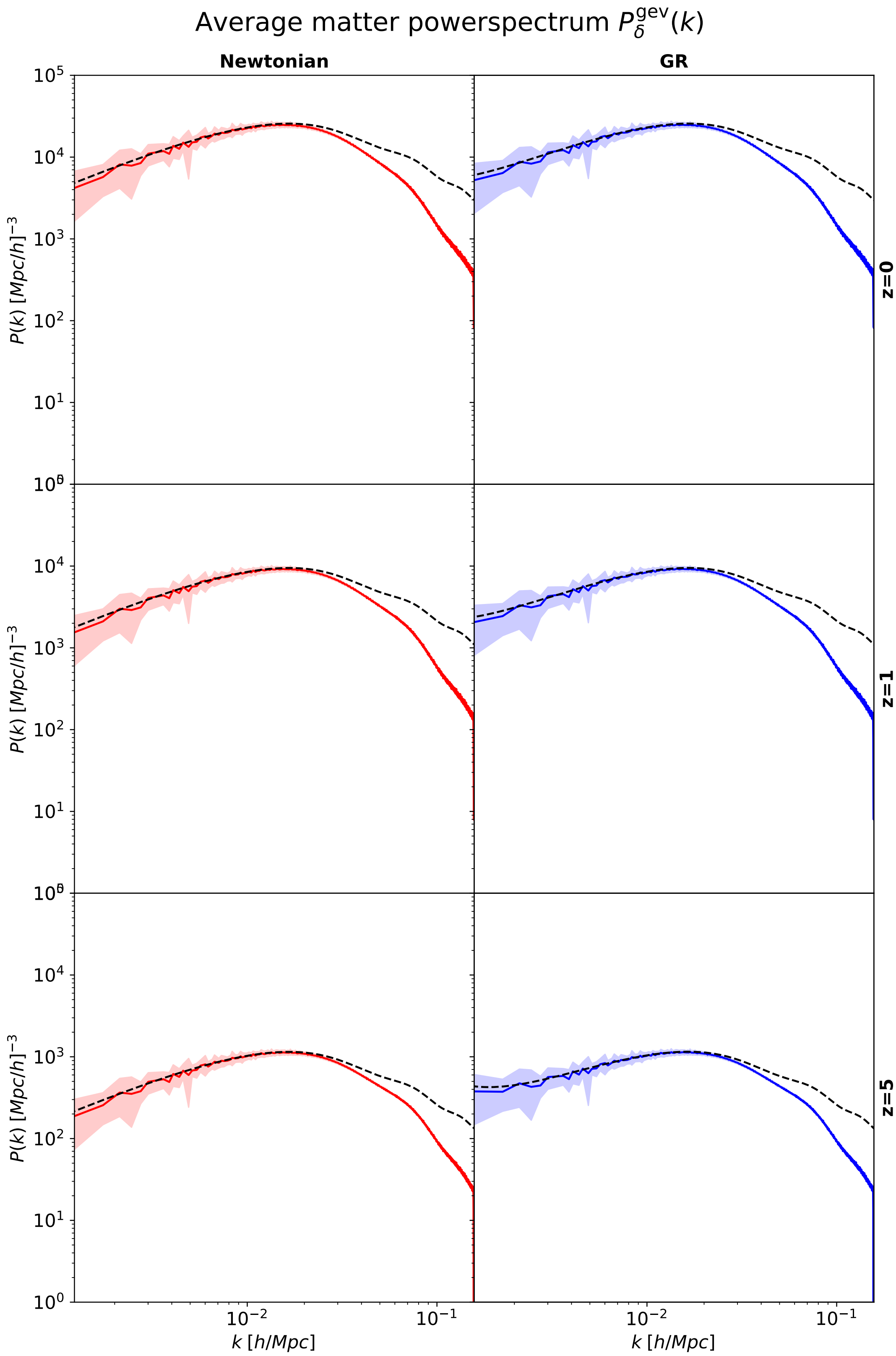


Figure 10.2: Average matter power spectra at different redshifts.

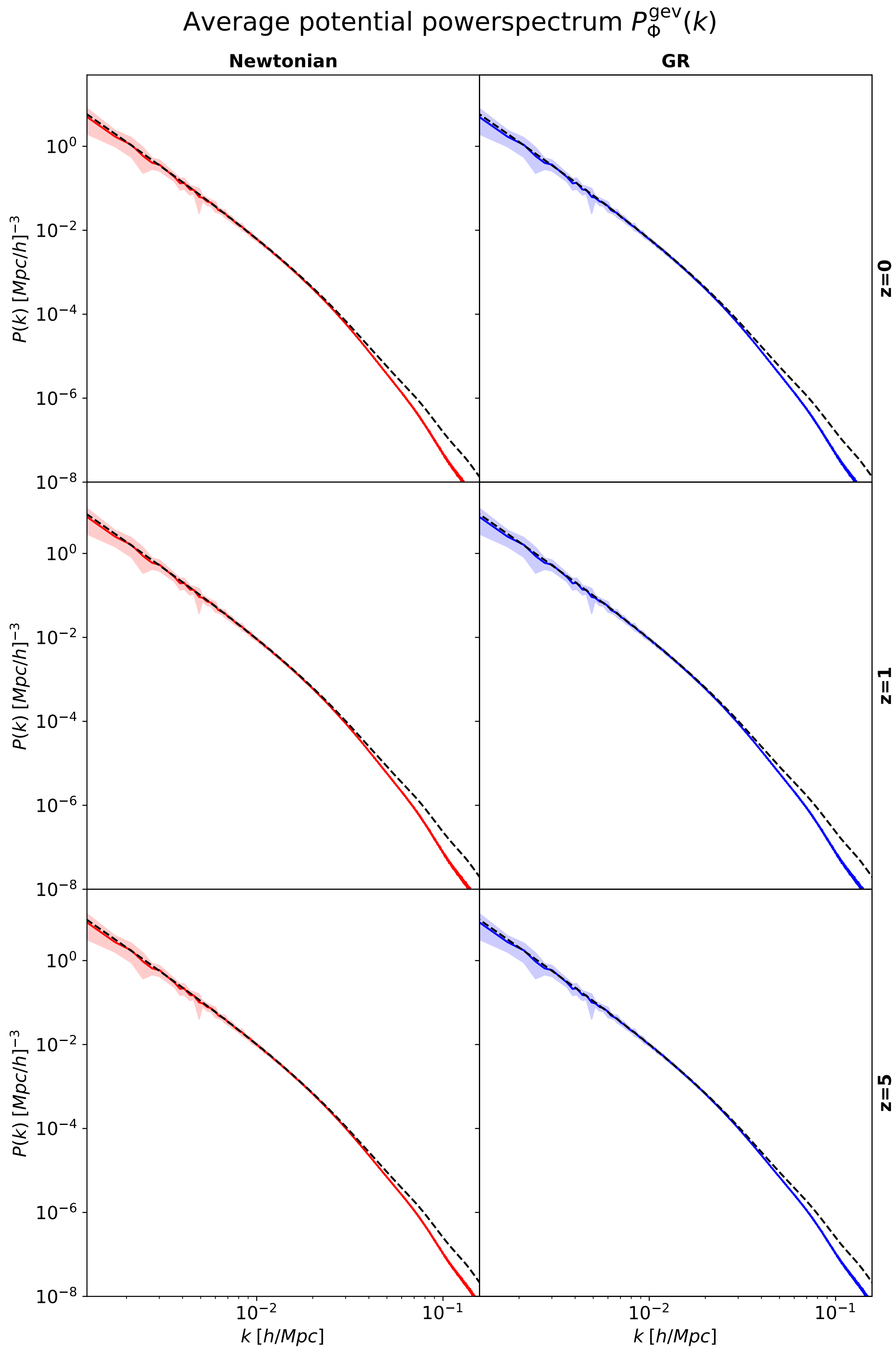


Figure 10.3: Average potential power spectra at different redshifts.

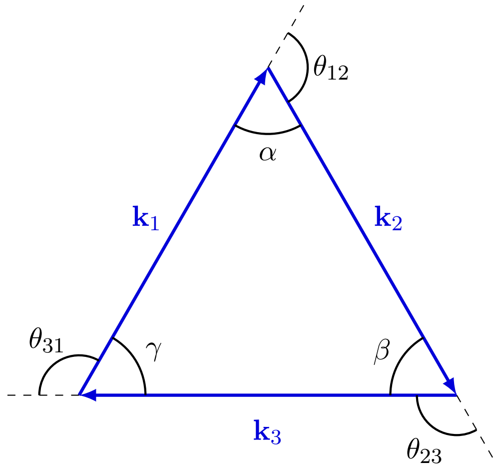
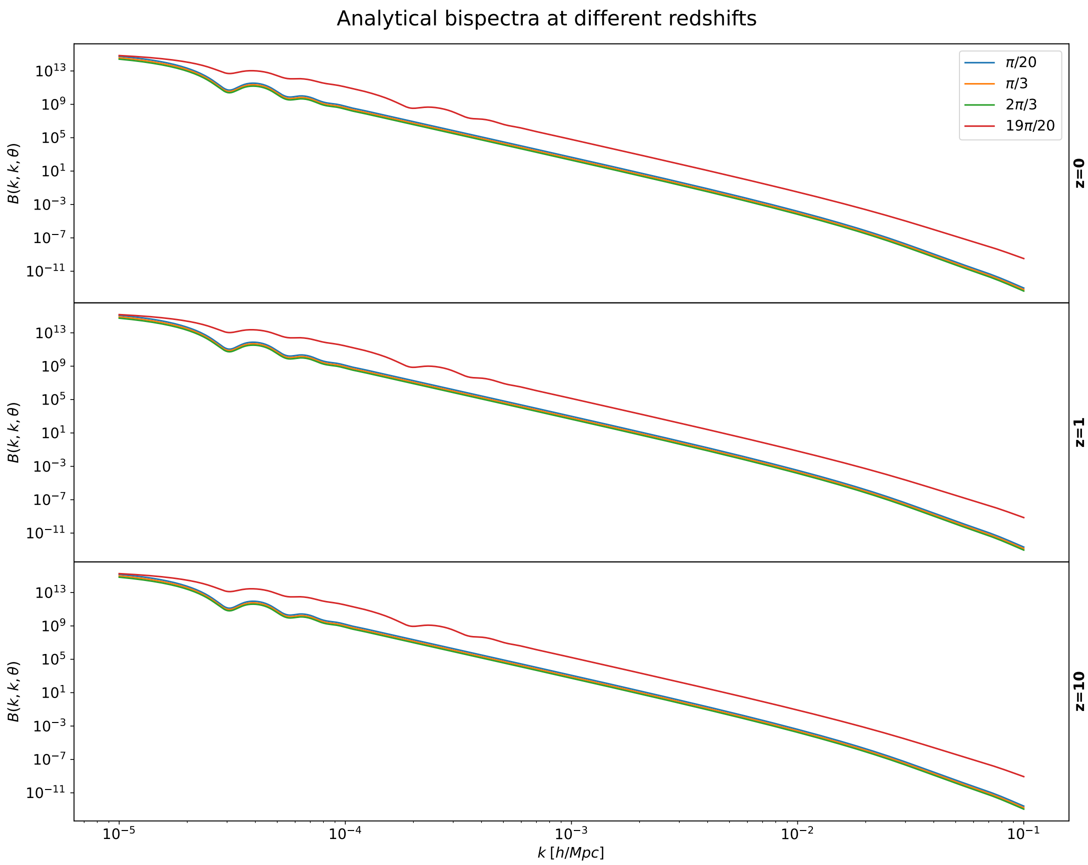
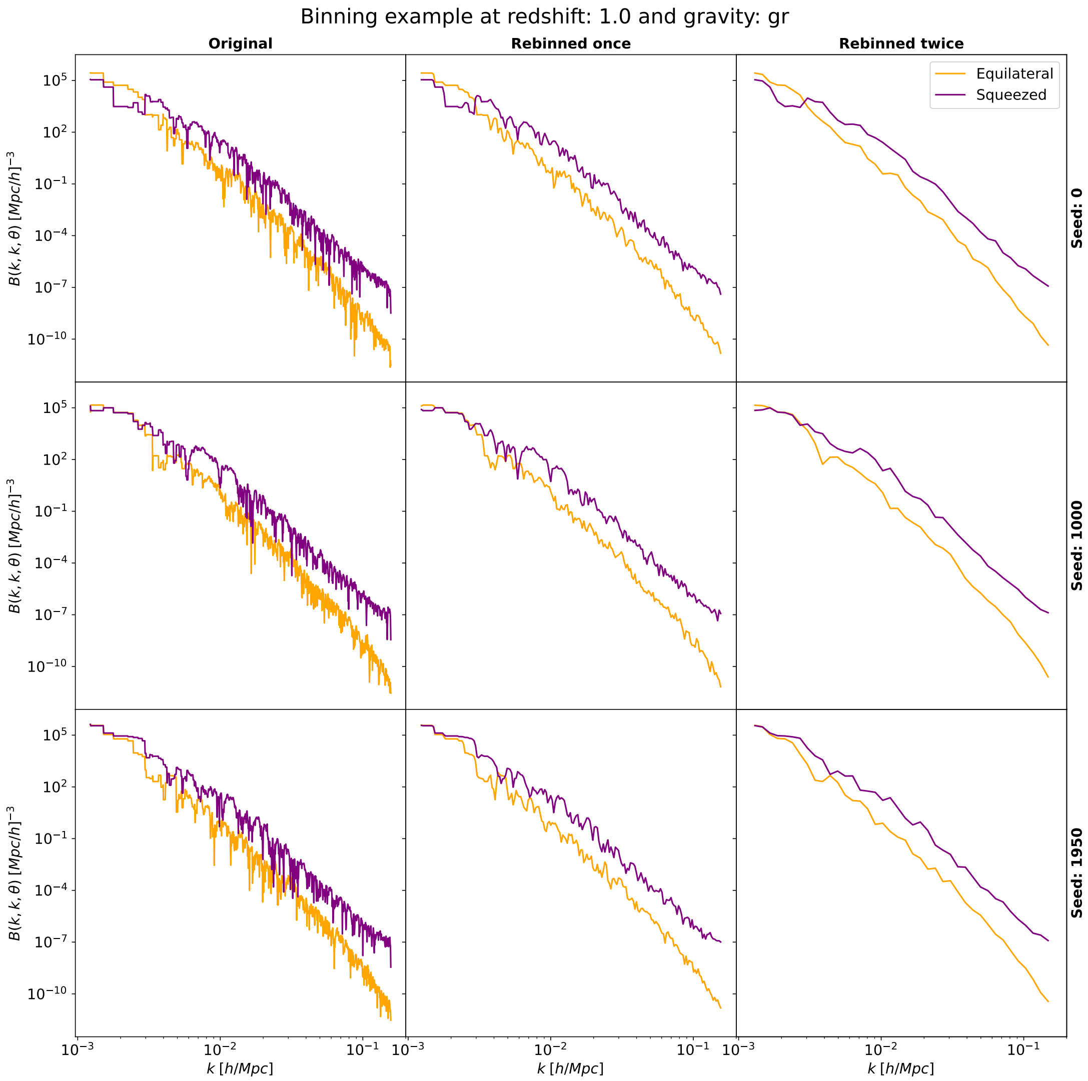
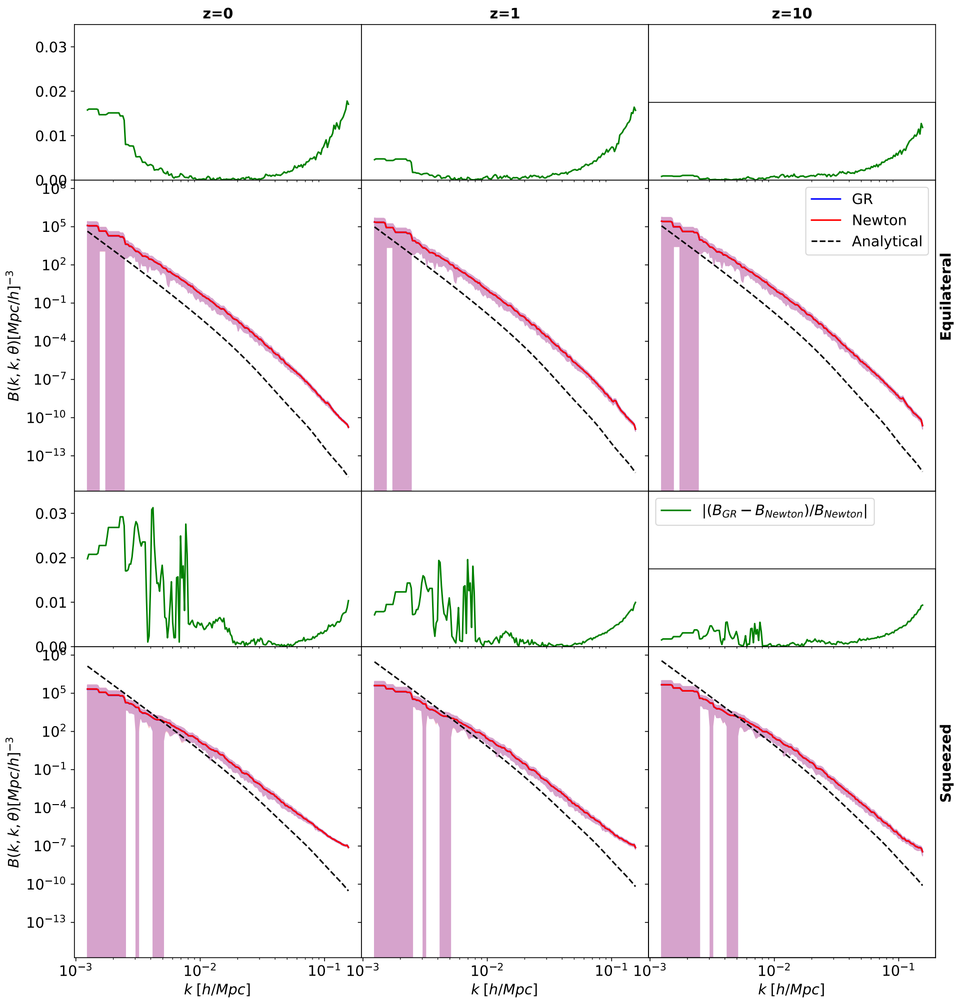


Figure 10.4: Angles in arbitrary bispectrum triangle configuration.





Binning example at redshift: 1.0 and gravity: gr
Original
Rebinned once
Rebinned twice
Equilateral
Squeezed
Seed: 0
Seed: 1000
Seed: 1950
$B(k, k, \theta)\ [Mpc/h]^{-3}$
$k\ [h/Mpc]$
10^5
10^2
10^{-1}
10^{-4}
10^{-7}
10^{-10}
10^{-3}
10^{-2}
10^{-1}


z=0
z=1
z=10
GR
Newton
Analytical
Equilateral
Squeezed
$|(B_{GR} - B_{Newton})/B_{Newton}|$
$B(k, k, \theta)[Mpc/h]^{-3}$
k [h/Mpc]

Chapter 11

Trainable Dataset

Bibliography

Adamek, Julian et al. (29th July 2016). 'gevolution: a cosmological N-body code based on General Relativity'. In: *Journal of Cosmology and Astroparticle Physics* 2016.7, pp. 053–053. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2016/07/053. arXiv: 1604.06065[astro-ph,physics:gr-qc,physics:physics]. URL: http://arxiv.org/abs/1604.06065 (visited on 23/08/2023).

Alqaraawi, Ahmed et al. (3rd Feb. 2020). *Evaluating Saliency Map Explanations for Convolutional Neural Networks: A User Study*. arXiv: 2002.00772[cs]. URL: http://arxiv.org/abs/2002.00772 (visited on 23/10/2023).

Chisari, Nora Elisa and Matias Zaldarriaga (2nd June 2011). 'Connection between Newtonian simulations and general relativity'. In: *Physical Review D* 83.12, p. 123505. ISSN: 1550-7998, 1550-2368. DOI: 10.1103/PhysRevD.83.123505. arXiv: 1101.3555[astro-ph,physics:gr-qc]. URL: http://arxiv.org/abs/1101.3555 (visited on 23/08/2023).

Dodelson, S. and F. Schmidt (2020). *Modern Cosmology*. Elsevier Science. ISBN: 9780128159484. URL: https://books.google.no/books?id=GGjfywEACAAJ.

Falck, B. et al. (16th Mar. 2017). 'The Effect of Corner Modes in the Initial Conditions of Cosmological Simulations'. In: *The Astrophysical Journal* 837.2, p. 181. ISSN: 1538-4357. DOI: 10.3847/1538-4357/aa60c7. arXiv: 1610.04862[astro-ph]. URL: http://arxiv.org/abs/1610.04862 (visited on 20/09/2023).

Goodfellow, Ian, Yoshua Bengio and Aaron Courville (2016). *Deep Learning*. http://www.deeplearningbook.org. MIT Press.

Green, Stephen R. and Robert M. Wald (15th Mar. 2012). 'Newtonian and Relativistic Cosmologies'. In: *Physical Review D* 85.6, p. 063512. ISSN: 1550-7998, 1550-2368. DOI: 10.1103/PhysRevD.85.063512. arXiv: 1111.2997[astro-ph,physics:gr-qc]. URL: http://arxiv.org/abs/1111.2997 (visited on 23/08/2023).

Hastie, T., R. Tibshirani and J.H. Friedman (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer series in statistics. Springer. ISBN: 9780387848846. URL: https://books.google.no/books?id=eBSgoAEACAAJ.

Jeong, Donghui, Fabian Schmidt and Christopher M. Hirata (4th Jan. 2012). 'Large-scale clustering of galaxies in general relativity'. In: *Physical Review D* 85.2, p. 023504. ISSN: 1550-7998, 1550-2368. DOI: 10.1103/PhysRevD.85.023504. arXiv: 1107.5427[astro-ph]. URL: http://arxiv.org/abs/1107.5427 (visited on 23/08/2023).

Selvaraju, Ramprasaath R. et al. (Feb. 2020). 'Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization'. In: *International Journal of Computer Vision* 128.2, pp. 336–359. ISSN: 0920-5691, 1573-1405. DOI: 10.1007/s11263-019-01228-7. arXiv: 1610.02391[cs]. URL: http://arxiv.org/abs/1610.02391 (visited on 23/10/2023).

Weinberg, S. (2008). *Cosmology*. Cosmology. OUP Oxford. ISBN: 9780191523601. URL: https://books.google.no/books?id=nqQZdg020fsC.