Universidade Estadual de Londrina

EDUARDO JOSÉ BARROSO

LARGE SCALE STRUCTURE FORMATION: FROM PRIMORDIAL BLACK HOLES IN BOUNCING COSMOLOGY TO GALAXY CLUSTERS ABUNDANCE

LONDRINA
2024

EDUARDO JOSÉ BARROSO

LARGE SCALE STRUCTURE FORMATION: FROM PRIMORDIAL BLACK HOLES IN BOUNCING COSMOLOGY TO GALAXY CLUSTERS ABUNDANCE

Dissertação apresentada ao Programa de Doutorado em Física da Universidade Estadual de Londrina para obtenção do título de Doutor em Física.

Orientador: Prof. Dr. Sandro Dias Pinto Vitenti

LONDRINA
2024

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração
Automática do Sistema de Bibliotecas da UEL

Barroso, Eduardo José.
Large Scale Structure Formation: from Primordial Black Holes in Bouncing Cosmology to Galaxy Clusters Abundance / Eduardo José Barroso. - Londrina, 2024.
126 f.

Orientador: Sandro Dias Pinto Vitenti.
Tese (Doutorado em Física) - Universidade Estadual de Londrina, Centro de Ciências Exatas, Programa de Pós-Graduação em Física, 2024.
Inclui bibliografia.

1. Primordial Black Holes - Tese. 2. Cosmology - Tese. 3. Galaxy Clusters - Tese. 4. APES: Approximate Posterior Ensemble Sampler - Tese. I. Vitenti, Sandro Dias Pinto. II. Universidade Estadual de Londrina. Centro de Ciências Exatas. Programa de Pós-Graduação em Física. III. Título.

CDU 53

EDUARDO JOSÉ BARROSO

LARGE SCALE STRUCTURE FORMATION: FROM PRIMORDIAL BLACK HOLES IN BOUNCING COSMOLOGY TO GALAXY CLUSTERS ABUNDANCE

Dissertação apresentada ao Programa de Doutorado em Física da Universidade Estadual de Londrina para obtenção do título de Doutor em Física.

BANCA EXAMINADORA

Orientador: Prof. Dr. Sandro Dias Pinto Vitenti
Universidade Estadual de Londrina - UEL

Prof. Dr. Rodrigo Corso Baptista dos Santos
Universidade Estadual de Londrina - UEL

Prof. Dra. Mariana Penna Lima
Universidade de Brasília – UNB

Prof. Dr. Miguel Boavista Quartin
Universidade Federal do Rio de Janeiro – UFRJ

Dr. Nelson Pinto Neto
Centro Brasileiro de Pesquisa em Física – CBPF

Londrina, 16 de julho de 2024.

AGRADECIMENTOS

I want to thank my family: Amanda Barroso, Sergio Barroso and Shirlene Barroso for all the love and support during these 4 years.

I want to thank my friends from the physics department: Julio Toledo, Luiz Demétrio, Caio Oliveria, Henrique Leitieri, and Cinthia Lima, for the all the discussions and coffees. I also want to thank my Friends from France: Constantin Payerne, Manon Ramel, Damien Cherouvrier, Carolina Arata, Marina Masson, Céline Combet, who received me with open arms during my exchange program.

I thank my advisor Sandro Dias Pinto Vitenti, for guiding me with patience during this period and helping me grow as a researcher and person.

Specially, I thank my brother Henrique Barroso for being in my life, a person who was always there for me and gave me all kinds of support while writing my thesis. Last, I thank CAPES for the financial support.

BARROSO, E. J. . **Formação de Estruturas em Larga Escala: de Buracos Negros Primordiais em Cosmologica com Ricochete até Abundância de Aglomerados de Galáxias**. 2024. 156f. Tese Doutorado em Física – Universidade Estadual de Londrina, Londrina, 2024.

RESUMO

Esta tese é constituída por três projetos diferentes: a formação de buracos negros primordiais num universo em expansão; a abundância de aglomerados dentro da colaboração LSST-DESC; e o desenvolvimento do APES, um algoritmo de ensemble sampler. No que diz respeito à formação de buracos negros primordiais, as perturbações cosmológicas escalares lineares têm espectros crescentes na fase de contração dos modelos de expansão. Estudamos as condições para as quais estas perturbações podem colapsar em buracos negros primordiais e a hipótese de que estes objetos constituem uma fração da matéria escura. Calculamos o contraste de densidade crítico que descreve o colapso das perturbações de matéria no modelo de expansão plana com uma solução paramétrica, obtida a partir da métrica de Lemaitre-Tolman-Bondi que representa o colapso esférico. Discutimos a incapacidade de calibre newtoniano para descrever perturbações em modelos de contração, uma vez que a hipótese perturbativa não se mantém nesses casos. Realizamos os cálculos para uma escolha de calibre diferente e computamos os espectros de potência das perturbações numericamente. Por fim, assumindo uma distribuição gaussiana, calculamos a abundância de buracos negros primordiais com o formalismo de Press-Schechter e comparamos com as restrições observacionais. A partir da nossa análise, concluímos que a formação de buracos negros primordiais numa fase de contração dominada por poeira não leva a uma fração significativa de massa de buracos negros primordiais na matéria escura atualmente.

Ainda em relação aos objetos colapsados das perturbações da matéria, podemos usar aglomerados de galáxias para investigar a cosmologia. Os aglomerados de galáxias são os maiores objetos gravitacionais limitados e são formados pelo colapso de sobre-densidades de matéria. Devido ao seu tamanho e conteúdo de matéria, são usados para estudar a formação da estrutura em larga escala no universo. Especificamente, a abundância de aglomerados de galáxias está fortemente relacionada à densidade de matéria do universo Ω_m e à flutuação de massa rms σ_8. Consequentemente, sua abundância pode ser prevista teoricamente e depois comparada com dados para ajustar parâmetros cosmológicos e assim investigar modelos cosmológicos. Neste contexto, o Legacy Survey of Space and Time (LSST) fornecerá uma grande quantidade de dados de aglomerados para serem usados em análises computacionais.

O segundo projeto consiste em desenvolver o pipeline de cluster dentro da Dark Energy Science Collaboration (DESC). Durante o doutorado, trabalhamos no Firecrown, um pacote Python para implementar verossimilhanças dentro do DESC. Implementamos uma verossimilhança de contagem de número de clusters, utilizando previsões teóricas com base na riqueza de clusters de galáxias como um proxy de massa. Após completar o código do Firecrown, testamos em vários conjuntos de dados, incluindo o catálogo RedMaPPer, que faz parte do Data Challenge 2 (DC2), um catálogo simulado de aglomerados. Validamos a verossimilhança de abundância de clusters do Firecrown usando o catálogo RedMaPPer ajustando parâmetros cosmológicos e de relação de massa-proxy e preparamos os dados combinando o catálogo de clusters de galáxias com seu catálogo de halo de simulações N-body para atribuir massas médias em cada bin de riqueza de redshift. A análise mostrou que os valores verdadeiros do catálogo estavam dentro da região de 2 sigmas.

Por último, o algoritmo APES propõe uma abordagem inovadora para gerar amostras de distribuições-alvo que são difíceis de amostrar usando métodos de Monte Carlo Markov Chain (MCMC). Algoritmos MCMC tradicionais frequentemente enfrentam uma convergência lenta devido à dificuldade em encontrar propostas que se adequem ao problema em questão. Para resolver esse problema, introduzimos o algoritmo Approximate Posterior Ensemble Sampler (APES), que emprega estimativa de densidade de kernel e interpolação de base radial para criar uma proposta adaptativa, levando a uma rápida convergência das cadeias. A escalabilidade do algoritmo APES para dimensões mais altas o torna uma solução prática para problemas complexos. O método proposto gera uma probabilidade posterior aproximada que se aproxima de perto da distribuição desejada e é fácil de amostrar, resultando em tempos de autocorrelação menores e uma maior probabilidade de aceitação pela cadeia. Comparamos o desempenho do algoritmo APES com o ensemble sampler de invariância afim com o stretch move em vários contextos, demonstrando a eficiência do método proposto. Por exemplo, na função Rosenbrock, o APES apresentou um tempo de autocorrelação 140 vezes menor do que o ensemble sampler de invariância afim. A comparação destaca a eficácia do algoritmo APES na geração de amostras de distribuições desafiadoras.Este trabalho apresenta uma solução prática para gerar amostras de distribuições complexas, abordando o desafio de encontrar propostas adequadas. Com novas pesquisas cosmológicas preparadas para lidar com muitas novas sistemáticas, este método oferece uma solução prática para a próxima era de análises cosmológicas.

Palavras-chave: Cosmologia, Buracos Negros Primordiais, Modelos de Ricochete, Aglomerados de Galáxias, Monte Carlo Markov Chain.

BARROSO, E. J. . **Large Scale Structure Formation: from Primordial Black Holes in Bouncing Cosmology to Galaxy Clusters Abundance**. 2024. 157p. Ph.D. Thesis – State University of Londrina, Londrina, 2024.

ABSTRACT

This thesis is constituted by three different projects: the primordial black hole formation in a bouncing universe; the cluster abundance inside the LSST-DESC collaboration; and the development of APES, an ensemble sampler algorithm. Regarding the primordial black hole formation, linear scalar cosmological perturbations have increasing spectra in the contracting phase of bouncing models. We study the conditions for which these perturbations may collapse into primordial black holes and the hypothesis that these objects constitute a fraction of dark matter. We compute the critical density contrast that describes the collapse of matter perturbations in the flat-dust bounce model with a parametric solution, obtained from the Lemaitre-Tolman-Bondi metric that represents the spherical collapse. We discuss the inability of the Newtonian gauge to describe perturbations in contracting models as the perturbative hypothesis does not hold in such cases. We carry the calculations for a different Gauge choice and compute the perturbations' power spectra numerically. Finally, assuming a Gaussian distribution, we compute the primordial black hole abundance with the Press-Schechter formalism and compare it with observational constraints. From our analysis, we conclude that the primordial black hole formation in a dust-dominated contracting phase does not lead to a significant mass fraction of primordial black holes in dark matter today.

Still regarding collapsed objects from matter perturbations, we can use galaxy clusters to probe cosmology. Galaxy clusters are the biggest gravitational bounded objects and are formed through the collapse of matter over-densities. Due to their size and matter content, they are used to study the large-scale structure formation in the universe. Specifically, the abundance of galaxy clusters is strongly related to the universe's matter density Ω_m and the rms mass fluctuation σ_8. Consequently, their abundance can be predicted theoretically and then compared with data to fit cosmological parameters and thus probe cosmological models. In this context, the Legacy Survey of Space and Time will provide a great amount of cluster data to be used in computational analyses.

The second project consists in developing the cluster pipeline inside the Dark Energy Science Collaboration (DESC). During the Ph.D, we worked on Firecrown, a Python package for implementing likelihoods within DESC. We implemented a cluster number counts likelihood, using theoretical predictions based on the richness of galaxy clusters as a mass-proxy. After completing the Firecrown code, we tested it on various datasets, including the RedMaPPer catalog, which is part of Data Challenge 2 (DC2), a simulated catalog. We validated the Firecrown cluster abundance likelihood using the RedMaPPer catalog by fitting cosmological and proxy-mass relation parameters and prepared the data by matching the galaxy cluster catalog with its halo catalog from N-body simulations to assign mean masses in each redshift-richness bin. The analysis showed that true catalog values were within the 2-sigma region.

Lastly, the APES algorithm proposes a novel approach to generate samples from target distributions that are difficult to sample from using Markov Chain Monte Carlo (MCMC) methods. Traditional MCMC algorithms often face slow convergence due to the difficulty in finding proposals that suit the problem at hand. To address this issue, we introduce the Approximate Posterior Ensemble Sampler (APES) algorithm, which employs kernel density estimation and radial basis interpolation to create an adaptive proposal, leading to fast convergence of the chains. The APES algorithm's scalability to higher dimensions makes it a practical solution for complex problems. The proposed method generates an approximate posterior probability that closely approximates the desired distribution and is easy to sample from, resulting in smaller autocorrelation times and a higher probability of acceptance by the chain. We compare the performance of the APES algorithm with the affine invariance ensemble sampler with the stretch move in various contexts, demonstrating the efficiency of the proposed method. For instance, on the Rosenbrock function, the APES presented an autocorrelation time 140 times smaller than the affine invariance ensemble sampler. The comparison showcases the effectiveness of the APES algorithm in generating samples from challenging distributions. This work presents a practical solution to generating samples from complex distributions while addressing the challenge of finding suitable proposals. With new cosmological surveys set to deal with many new systematics, this method offers a practical solution for the upcoming era of cosmological analyses.

Keywords: Cosmology, Primordial Black Holes, Bouncing Models, Galaxy Clusters, Monte Carlo Markov Chain.

SUMÁRIO

Parte I

Theoretical Cosmology

1 INTRODUCTION

Primordial Black Holes (PBHs) are believed to have been formed in the early universe through the collapse of density fluctuations [1, 2, 3, 4]. Due to their formation mechanisms, in the inflationary scenario, PBHs can have a wider range of masses from around $M \sim 5 \times 10^{-29} M_{\odot}$ if formed at Planck time $t = 10^{-43} s$ and $M \sim 10^{5} M_{\odot}$ if formed at $t \sim 1s$, resulting in several different effects. The study of primordial black holes has yielded insights into various phenomena. They could have contributed to the cosmological density parameter [5, 6]. Their influence in the Cosmic Microwave Background Radiation (CMB) was studied in Ref. [7] and some recent works explored the connection between evaporating PBHs and gravitational waves [8]. There also have been discussions about whether some measurements of gravitational waves could be attributed to primordial black holes [9]. Undoubtedly, the possibility that primordial black holes constitute a significant fraction of cold dark matter remains a focal point in current research [10]. Recent observations of merging binary black holes with unexpected mass ranges by the LIGO/Virgo collaborations indicate that they could have originated from PBHs as their mass do not correspond to black holes formed by the death of stars [11, 12].

In the inflationary scenario [13, 14, 15, 16], several studies have analyzed the formation of PBH at the end of inflation [17, 18, 19, 20, 21] and in the reheating phase [22, 23] either as a probe for the inflationary model or to analyze the abundance of PBH in dark matter. When confronting the results with combined probes from observational data, it was found that the only acceptable PBH mass range that allows these objects to entirely constitute dark matter is $10^{-16} M_{\odot} - 10^{-12} M_{\odot}$ [24, 25]. Nonetheless, other mass ranges that lead to a smaller fraction of dark matter have intriguing impacts to be considered. For instance, in the mass range $10^{6} M_{\odot}$ to $10^{10} M_{\odot}$ where PBHs only represent 0.1% of DM, they could play a role in generating supermassive black holes. In [26] it was also concluded that the possibility of dark matter being constituted by supermassive primordial black holes is viable. However, these results led researchers to study PBH formation in other cosmological models to check the dependency of the mass constraints on the chosen models.

As is well known, the inflationary paradigm alleviates the initial condition problems of the Λ-CDM model but does not completely solve them [14, 27, 28]. For instance, one still needs to assume that a small patch of spacetime was of FLRW type and evolved to become the universe that we now observe, which is modeled via perturbation theory [29]. Alternative models to inflation have been considered, mainly bouncing models (see [30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40] for extensive reviews on bouncing cosmology), which focus on solving the singularity problem by introducing a contracting phase connected to the present expanding phase trough a minimal scale factor: a bounce. Classical bouncing models, where exotic matter or modifications of GR are considered, have been analyzed extensively in the literature, mainly in Refs. [40, 28]. In these works, one sees that classical bounces are non-trivial to implement and lead to undesired features such as instabilities and new singularities, which are associated with the violation of the null energy condition.

One may also consider quantum bounces, where the quantization of gravity itself eliminates the primordial singularity. In particular, since quantum bounces do not make use of exotic matter such as the inflaton, there is no need for a reheating phase. This has been achieved in previous works in the framework of Canonical Quantum Gravity [41]. Other relevant proposals are Loop Quantum Cosmology [42, 43] and String Gas Cosmology [28, 40].

Some recent studies have analyzed the formation of PBH in bouncing models [44, 45, 46, 47, 48, 49, 50]. In this context, it is believed that the long duration of the contracting phase plus the diminishing scale factor may lead to an enhancement of PBH formation, resulting in a more significant contribution to the DM density. In [48] it is concluded that a dust-dominated bouncing universe ($w > 1/3$) is robust against the formation of such primordial black holes. In contrast, for a matter-dominated universe ($w \ll 1$), their formation becomes relevant. A similar conclusion is achieved in [47], where they got an enhanced production of PBHs near the bouncing point and utilized the abundance of PBHs in DM to constrain the bouncing model. However, in the same work, it is stated that these constraints are still not well understood due to a lack of precision in numerical computations.

Furthermore, it is still not quite clear how to properly define the critical threshold for PBHs in bouncing cosmology. Different from the inflationary scenario, the contracting phase dynamics lead to growing perturbations that will collapse before becoming super-

Hubble. To circumvent this problem, in [48] it is used the argument that the black holes must be larger or equal to the Schwarzschild radius for the given formation mass seeded by the perturbations. However, there is still the need for a more precise definition of the critical contrast in the bouncing scenario. Furthermore, the gauge definition used in previous works may not be the best choice as it leads to large increasing spectra and thus a miscalculation of the energy-density perturbations [51].

Still regarding collapsed objects from matter perturbations, we can use galaxy clusters to probe cosmology. Galaxy clusters are the biggest gravitationally bounded objects and are formed through the collapse of matter over-densities. Due to their size and matter content, they are used to study the large-scale structure formation in the universe [52, 53]. Specifically, the abundance of galaxy clusters is strongly related to the universe's matter density Ω_m and the rms mass fluctuation σ_8 [54, 55]. Consequently, their abundance can be predicted theoretically and then compared with data to fit cosmological parameters and thus probe cosmological models.

In this context, the Legacy Survey of Space and Time (LSST) [56], conducted by the Rubin observatory during a ten-year interval, will provide a gigantic dataset on a variety of topics, ranging from solar systems to cosmological scales. The survey will provide a great amount of cluster data to be used in computational analyses, such that the codes have been developed by the members of the Dark Energy Science Collaboration (DESC), an international collaboration preparing the upcoming cosmological analysis with Rubin data. One of the goals is to analyze the variety of data using Markov Chain Monte Carlo (MCMC) methods with posterior distributions and find constraints on the ΛCDM cosmological parameters.

Markov Chain Monte Carlo (MCMC) methods have become a prominent tool in Bayesian inference since their inception by Metropolis in 1953 [57]. MCMC has proven to be a versatile and effective method for generating samples from complicated and high-dimensional probability distributions, making it popular among scientists across various disciplines. In particular, MCMC methods have found widespread application in cosmology and astrophysics for estimating model parameters from observational data, as evidenced by several studies.

Obtaining samples from the posterior distribution, which is computed using

Bayes' theorem with prior distributions and the likelihood function of the data, can be computationally demanding, especially for high-dimensional problems [58]. However, MCMC methods offer a viable solution to this issue by enabling the generation of samples from the posterior distribution, which can then be used to infer properties of the underlying physical process. The effectiveness of MCMC in this context has been instrumental in advancing our understanding of complex physical systems.

In this thesis, we study both theoretical and observational effects in cosmology. Chapters 2 and 3 are devoted to review the standard cosmological model and the dust bouncing model for contracting universes. Then, we delve in three different projects: the formation of PBH in a quantum dust non-singular bounce and the hypothesis that these structures constitute a fraction of DM today (Chapter 4) [59]; the cluster number counts analysis in LSST-DESC in Chapter 5, where we discuss the development of the cluster pipeline inside the collaboration and our contribution to it; and in Chapter 6 we have APES: Approximate Posterior Ensemble Sampler [60], an MCMC sampler algorithm that uses kernel density estimation and radial basis interpolation to construct posterior distributions, developed by me and my advisor. We conclude our thesis in Chapter 7.

2 STANDARD COSMOLOGICAL MODEL

The standard cosmological model describes the universe as spatially homogeneous and isotropic. This hypothesis is well-known and accepted among the cosmology community and it is referred to as the cosmological principle. This portrait of space-time was first introduced by Friedman [61] and later by Lemaître [62] when studying some applications of General Relativity (GR). It was then proved by Robertson [63] and Walker [64] that the metric used by Friedman and Lemaître was the only possibility for a homogeneous and isotropic universe. We want to study this model and understand why it is a good description of the universe. We start this chapter with Sec. 2.1 by introducing basic concepts about cosmology and the notation used throughout this work. In Sec. 2.2, we discuss the cosmological principle and make a brief review of the isotropic and homogeneous universe model. Lastly, in Sec. 2.3, we discuss the conformal description of cosmological quantities. For extensive reviews on cosmology models and the cosmological principle, we point the reader to Refs. [65, 66, 67, 68].

2.1 Basic Concepts and Notation

Let us define some notations. In this work, we use SI units and make no simplifications. As we will later see, the universe is considered to have a spatially homogeneous and isotropic background with small perturbations around it. Quantities with an over-bar sign (for example $\bar{g}_{\mu\nu}$) will refer to background quantities defined in the perturbed manifold. The absence of the overbar sign refers to the physical/perturbed quantities around the background ones (check App.A for a review on the perturbation theory). We are also using the metric signature $(-,+,+,+)$ and the Einstein summation convention. Let us now define some needed concepts from differential geometry.

2.1.1 Differential Geometry

When describing a universe in the presence of gravitational interaction, one common approach is to introduce General Relativity (GR). GR depends on differential geometry as a tool to define the structure of the universe. In such theory, the space-time is viewed as a differential manifold $\mathcal{M}$, a topological m-dimensional space whose set of points are space-time points. There are some key concepts needed in this context and we shall devote this section to reviewing them, mostly focusing on the explicit notation of some quantities rather than their demonstrations. For a more complete revision on GR and differential geometry, check Refs. [69, 70, 71].

We will review some concepts to properly describe the space-time manifold $\mathcal{M}$, such as the time direction, spatial vectors, derivatives, distances, and others. We first define the time direction. Suppose that at a time t_0, we choose a spatial surface represented by $\Sigma_{t_0} \in \mathcal{M}$ such that all tangent vectors related to curves in this hypersurface are space-like. In this context, we define a vector field n^u that is orthogonal to the hypersurface Σ_{t_0} with negative norm $n^\mu n_\mu = -1$. The split of the manifold in a time direction and spatial surfaces allows us to create a sectioning in our manifold and to define time and space-like vectors.

The next natural concept that we need is a notion of distance in the space-time manifold $\mathcal{M}$. For that, we use the metric tensor $g_{\mu\nu}$, such that the line element between two vectors is given by $ds^2 = g_{\mu\nu}dx^\mu dx^\nu$. For example, the three-dimensional Euclidean metric tensor in Cartesian coordinates is given by

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{2.1}$$

which is basically how we measure the distance between physical quantities in classical physics. On another hand, we have the Minkowski space-time metric

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{2.2}$$

The metric tensor is a very useful and necessary quantity in our manifold. Since in general relativity there is no preferable coordinate system, the metric gives us an absolute and observer-independent quantity, which will be used to construct physical quantities. The metric tensor also gives us a connection between different tangent spaces as it is defined in the entire tangent bundle, the set of all tangent spaces related to the manifold. With the notion of time and distance, let us now move to derivatives.

We will now introduce the Lie derivative, which is going to be needed in our calculations. Given a vector field on a manifold $\mathcal{M}$, the Lie derivative of a tensor field[1] gives the speed with which the tensor field changes in the direction of the flow generated by the vector field. For a vector v^{μ}[2], its Lie derivative regarding another vector l^{μ} is represented by $\pounds_l v^{\mu}$. We will mostly use this concept when analyzing the change of some quantities in the time direction. We will refer to the time derivative of a vector $\frac{\partial}{c\partial t}$, which will be represented by

$$\dot{} \equiv \partial_{ct} \equiv \pounds_n, \tag{2.3}$$

where $\pounds_n$ is the Lie derivative in the time direction and c is the speed of light. Note that in the coordinate system where n^{μ} coincides with the temporal direction, the Lie derivative in this direction is equivalent to the partial time derivative, as stated above.

The next concept will be the covariant derivative, which introduces the notion of parallel transport and curvature in a manifold[3]. A vector v^{μ} is said to be parallel transported along a curve defined by w^{μ} if

$$w^{\mu}\nabla_{\mu}v^{\nu} = 0, \tag{2.4}$$

meaning that v^{μ} remains unchanged. In the above, we are defining the covariant derivative as the operator $\nabla \equiv \nabla_{\nu}dx^{\nu}$[4]. In this context, we are now able to define a geodesic as curves whose tangent vector is parallel propagated along itself, which is the case of n^{μ} by construction. Mathematically, if

$$\nabla_n n^{\nu} = n^{\mu}\nabla_{\mu}n^{\nu} = 0, \tag{2.5}$$

---

[1] We will soon not write explicitly the field description for general tensors and vectors fields, which shall be sub-intended.

[2] Throughout this work, we use abstract index notation, such that we represent a general (m, n) tensor $T = \sum_{\nu_1,...,\nu_m,\mu_1,...,\mu_n} T^{\nu_1,...,\nu_m}{}_{\mu_1,...,\mu_n}\partial_{\nu_1}...\partial_{\nu_m}dx^{\mu_1}...dx^{\mu_n}$ by its components $T^{\nu_1,...,\nu_m}{}_{\mu_1,...,\mu_n}$.

[3] For more information about covariant derivatives, check Ref. [69]

[4] For a given metric, there is only one covariant derivative compatible so that $\nabla_{\alpha}g_{\mu\nu} = 0$, which is the one that we use everywhere in this work.

then the vector n^μ defines a geodesic. Also, in the above, we used $\nabla_n = n^\mu \nabla_\mu$[5]. The intrinsic curvature of the manifold is expressed by the Riemann curvature tensor R as

$$\left[\nabla_\alpha, \nabla_\beta\right] v_\mu = \nabla_\alpha \nabla_\beta v_\mu - \nabla_\beta \nabla_\alpha v_\mu \equiv R_{\alpha\beta\mu}{}^\nu v_\nu. \tag{2.6}$$

Having the time-space splitting, a distance notion and the derivative definitions in our manifold, we now discuss some properties of the metric tensor.

A useful tool of the metric is that it can be used to interchange vectors for co-vectors and vice-versa. Explicitly,

$$g_{\mu\nu} v^\mu = v_\nu \quad \text{and} \quad g^{\mu\nu} w_\mu = w^\nu. \tag{2.7}$$

We can also use the metric to split our manifold into a set of spatial hypersurfaces. For a metric $g_{\mu\nu}$ in a manifold $\mathcal{M}$, we can define

$$\gamma_{\mu\nu} = g_{\mu\nu} + n_\mu n_\nu \tag{2.8}$$

where n^μ is the vector field normal to the spatial hypersurfaces. The tensor $\gamma_{\mu\nu}$ is a positive-definite metric in the spatial hypersurfaces and can be used as a projector. Given a generic (m, n) tensor T, its contraction with this metric will lead to its spatial counterpart, given by

$$\gamma[T^{\nu_1,\dots,\nu_m}{}_{\mu_1,\dots,\mu_n}] \equiv \gamma^{\mu_1}{}_{\alpha_1} \dots \gamma^{\mu_m}{}_{\alpha_m} \gamma^{\nu_1}{}_{\beta_1} \dots \gamma^{\nu_n}{}_{\beta_n} T^{\alpha_1,\dots,\alpha_m}{}_{\beta_1,\dots,\beta_n}. \tag{2.9}$$

The covariant derivative applied in the spatial sections is written as

$$D_\alpha T^{\nu_1,\dots,\nu_m}{}_{\mu_1,\dots,\mu_n} \equiv \gamma[\nabla_\alpha T^{\nu_1,\dots,\nu_m}{}_{\mu_1,\dots,\mu_n}]. \tag{2.10}$$

Lastly, the contraction of a vector in the normal direction is given by

$$T_n = T_\mu n^\mu. \tag{2.11}$$

We now move to some concepts in general relativity.

---

[5] Bear in mind that when applied to a scalar quantity f, $\nabla_n f \equiv \partial_{ct} f$

2.1.2 Einstein's Field Equations

A key element in GR is Einstein's field equations as they connect the geometry of space-time to its matter content. These equations are given by

$$G_{\mu\nu} + g_{\mu\nu}\Lambda = \frac{8\pi G}{c^4} T_{\mu\nu}. \tag{2.12}$$

Let us analyze term by term. First, $G_{\mu\nu}$ is the Einstein tensor which describes the curvature of space-time. It is given in terms of the Ricci tensor $\mathcal{R}_{\mu\nu} = R_{\mu\alpha\nu}{}^{\alpha}$ and the scalar curvature R, that is,

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu}, \quad \text{with} \quad R = g^{\mu\nu} R_{\mu\nu}. \tag{2.13}$$

The second term is the cosmological constant Λ. It can be understood as the dark energy contribution or the vacuum energy of the universe leading to an accelerated expansion [72]. Lastly, on the right, the stress-energy tensor $T_{\mu\nu}$ describes the continuous matter distribution and G is Newton's gravitational constant. Eq. (2.12) is the base of general relativity as it connects the geometry of space-time on the left to the distribution of matter on the right.

If we consider some regions with different symmetries, the solutions to Einstein's equation lead to important results in GR. In the next section, we will study in detail the case of a homogeneous and isotropic universe, which is viewed as a suitable description of the real universe.

2.2 Isotropic and Homogeneous Universe

All modern cosmology starts with two main assumptions: the universe is spatially homogeneous (invariant under spatial translations) and isotropic (invariant under rotations). The combination of both plus a trivial topology is called the *cosmological principle* and it is considered to be an acceptable approximation for the universe on large scales. This model is the simplest one to study and there are some arguments on why it should be the real case. Isotropy implies that the universe is the same in all directions and there is strong evidence that supports this affirmation on large scales around the Earth, including the distribution of galaxies, radio sources, etc. Satellites as WMAP [73] and PLANCK [74] observed the isotropy on the Cosmic Microwave Background (CMB) and large sky surveys as the Sloan digital sky survey [75] and the 2-degree field galaxy redshift survey [76] also support the same hypothesis.

Assuming that the universe is isotropic around Earth, the Copernican principle can be used to easily generalize this concept. It states that neither the Earth nor the solar system occupies any privileged position in the universe, which seems intuitively true among physicists since there is no reason to choose any preferable . If this principle is valid, all observers at different places in the universe should also experience the isotropy property. From there, isotropy in all points also implies homogeneity in the universe.

Of course, locally the universe is not isotropic and homogeneous and we can consider perturbations around such a model, as done in App. A. However, the cosmological principle is a good global approximation. In this type of model, the only possible metric that takes into account all these symmetries is the Friedmann–Lemaître–Robertson–Walker (FLRW) metric, given by

$$ds^2 = -c^2dt^2 + \bar{a}^2(t)\left(\frac{dr^2}{1-\tilde{K}r^2} + r^2\left(d\theta^2 + \sin^2\theta d\phi^2\right)\right), \tag{2.14}$$

where t is the cosmic time $\bar{a}(t)$ is the dimensionless scale factor that dictates the expansion rate of the universe, and $\tilde{K}$ is a constant with dimension of one over length squared that defines the universe's spatial curvature. The cosmic time is the proper time of observers that are free falling in directions tangent to the geodesics defined by n^μ and we shall refer to observers in this frame as comoving observers. In other words, it is the time measured by observers that are in a coordinate system that expands with the same velocity as the universe and it is constant in each hypersurface. Regarding the curvature, it can assume different values with the cases $\tilde{K} = 1, 0, -1$ specifying the spatial curvature of the universe as spherical, flat, and hyperbolic respectively.

In Eq. (2.14), the coordinates r, θ, and ϕ are spherical comoving coordinates, implying that comoving observers are free-falling and remain at rest in this frame. To measure physical distances, one must take into account the scale factor when comparing the difference between vectors in this coordinate system.

Additionally, another quantity that can be used to measure the universe's expansion is the redshift, a measurement of the stretch of the light's wavelength from distant sources.

This quantity is represented by the letter z and is given by

$$1 + z = \frac{\lambda_{obs}}{\lambda_{em}} = \frac{\bar{a}(t_{obs})}{\bar{a}(t_{em})}, \tag{2.15}$$

where λ_{obs} and λ_{em} are the wavelengths received by the observed and emitted by the source respectively. The redshift can be seen both as a measure of time and distance in a cosmological model, as it will be higher for sources far in the past or for great distances separating the source from the observer[6,7]. Having the metric to describe the space-time, we also want to analyze the dynamics of the universe. In Eq. (2.14), we are only left with the scale factor as the dynamic degree of freedom and we will now evaluate its change in time.

2.2.1 Friedman Equations

The Friedmann equations are a set of equations derived from Einstein's field equations in the context of an isotropic and homogeneous universe. They describe the evolution of the scale factor, which determines the expansion rate of the universe. Luckily, imposing the cosmological principle, the universe's symmetries lead to a surprisingly simple form to deduce these equations. In this scenario, the stress-energy tensor in Eq. (2.12) can only have the form [77]

$$\bar{T}_{00} = \bar{\rho}(t)c^2, \;\; \bar{T}_{i0} = 0, \;\; \bar{T}_{ij} = \bar{p}(t) \tag{2.16}$$

Therefore the tensor must take the form of a perfect fluid, i.e.,

$$\bar{T}_{\mu\nu} = \bar{\rho}(t) n_\mu n_\nu + \bar{p}(t)\gamma_{\mu\nu}, \tag{2.17}$$

such that $\bar{\rho}$ is the energy density and $\bar{p}$ is the fluid's isotropic pressure defined by the comoving observers.

Plugging Eq. (2.17) plus the FLRW metric into Eq. (2.12) results in the Friedmann equations

$$\bar{H}^2 = \left(\frac{c\dot{\bar{a}}}{\bar{a}}\right)^2 = \frac{c^2\kappa}{3}\bar{\rho} - \frac{\tilde{K}c^2}{\bar{a}^2} \tag{2.18}$$

---

[6] Since in an expanding universe the scale factor always increases when approaching the present time, $z > 0$ when comparing the present with sources in the past.

[7] In this work, we will denote the scale factor and other quantities evaluated in the present time t_0 with a zero subscript, that is, $\bar{a}_0 = \bar{a}(t_0)$.

and

$$c\dot{\bar{H}} + \bar{H}^2 = \frac{c^2\ddot{\bar{a}}}{\bar{a}} = -\frac{c^2\kappa}{2}\left(\bar{\rho} + 3\bar{p}\right). \tag{2.19}$$

In the above, $\kappa = \frac{8\pi G}{c^4}$ and we introduced the Hubble function $\bar{H}(t) = \frac{c\dot{\bar{a}}(t)}{\bar{a}(t)}$. Note that Eq. (2.18) is not a dynamical equation. It works as a constraint for the initial conditions of the system given by the dynamical equation in (2.19). It is also convenient to write them as

$$\bar{\Theta}^2 = 3\kappa\bar{\rho} - 3\bar{K} \tag{2.20}$$

and

$$\dot{\bar{\Theta}} = -\frac{3\kappa}{2}(\bar{\rho} + \bar{p}) + 3\bar{K}, \tag{2.21}$$

where $\bar{\Theta} = \frac{3\bar{H}}{c}$ and $\bar{K} = \left(\frac{\bar{K}}{\bar{a}^2}\right)$.

With these relations, one can analyze how the scale factor evolves in an isotropic and homogeneous universe in the presence of a perfect fluid. We shall now discuss the type of fluids in our cosmological model.

2.2.2 Universe Energy Density

In the FLRW model, the fluids present in the universe are taken to follow an equation of the state of the form

$$w = \frac{\bar{p}}{\bar{\rho}}, \tag{2.22}$$

where w is a constant. We call dust-like fluids the ones with $w = 0$, valid for fluids with $mc^2 \gg k_B T$, where m is the fluid's mass, k_B the Boltzmann constant, and T the fluid's temperature. We call radiation fluids with $mc^2 \ll k_B T$ and they have $w = \frac{1}{3}$.

For fluids described by Eq. (2.22) in our model, we can assume a thermodynamic equilibrium in the universe, which is seen in the CMB and we assume it to be true in the early universe as well. We can use the energy conservation equation $\nabla_\mu T^{\mu\nu} = 0$ to obtain an explicit form of the fluids' density,

$$\bar{\rho} = \bar{\rho}_0 x^{3(1+w)}, \tag{2.23}$$

with $x = \frac{\bar{a}_0}{\bar{a}}$. Let us now discuss the different matter components.

It is useful that we introduce dimensionless densities to describe the types of matter in the universe. From Eq. (2.18), we define a critical density at a time t as the universe density for a flat universe ($\tilde{K} = 0$). Explicitly,

$$\bar{\rho}_c(t) = \frac{3\bar{H}^2}{c^2\kappa}, \tag{2.24}$$

such that we can now define dimensionless densities $\Omega_i = \frac{\bar{\rho}_i}{\bar{\rho}_c}$, where i refers to different fluids. In our model, we consider a matter contribution Ω_m constituted by cold dark matter and baryons, a radiation contribution Ω_r from photons and neutrinos, a cosmological constant contribution Ω_Λ to explain the accelerated expansion of the universe as given by a fluid with negative pressure ($w = -1$). Thus,

$$\Omega_i = \frac{\bar{\rho}_{i0}}{\bar{\rho}_c} x^{3(1+w)}, \text{ for } i, w = (m, 0), (\gamma, \frac{1}{3}), (\Lambda, -1) \tag{2.25}$$

and the total density of the universe is

$$\Omega = \Omega_m + \Omega_\gamma + \Omega_\Lambda. \tag{2.26}$$

In this context, we can rewrite Eq. (2.18) by defining a normalized variable $E(t)$ as

$$\bar{E}(t) \equiv \frac{\bar{H}(t)^2}{\bar{H}_0^2} = \Omega + \Omega_k = \Omega_{\gamma 0}x^4 + \Omega_{m0}x^3 + \Omega_{k0}x^2 + \Omega_{\Lambda 0}, \tag{2.27}$$

where $E(t_0) = 1$, $\Omega_{i0} = \frac{\bar{\rho}_{i0}}{\bar{\rho}_c}$ and $\Omega_k = -\bar{K}\bar{R}_H^2$ is the curvature density with $\bar{R}_H = \frac{c}{\bar{H}_0}$. In the case of the cosmological constant, we define its value today as $\Omega_{\Lambda 0} \equiv \frac{\Lambda \bar{R}_H}{3}$. This description of the matter content is known as the Λ Cold Dark Matter (ΛCDM) model since it includes the cold dark matter contribution in Ω_m and the dark energy in through the cosmological constant. For different times, we may consider that the universe is dominated by one of the fluids components and to analyze its behavior, although all contributions may be taking into account in Eq. (2.27). For an expanding universe, x decreases as we approach t_0, meaning that the universe was first dominated by radiation, followed by matter and we are now at a cosmological constant epoch. As we will discuss in the next chapter, the universe is believed to always have been flat, and thus there was no curvature epoch. The last topic in this chapter is devoted to the conformal description.

2.3 Conformal Description

In some cases, it may be interesting to rescale the time coordinate in Eq. (2.14) to simplify some calculations. Therefore we define the conformal time η as

$$\eta - \eta_0 \equiv c \int_{t_0}^{t} \frac{dt}{a(t)}, \tag{2.28}$$

such that the metric is given by

$$ds^2 = \bar{a}^2(\eta) \left(d\eta^2 + \frac{dr^2}{1 - \tilde{K}r^2} + r^2 \left(d\theta^2 + \sin^2 \theta d\phi^2 \right) \right). \tag{2.29}$$

In this coordinate system, we can define a constant spatial metric $\tilde{\gamma}_{\mu\nu} = \frac{\gamma_{\mu\nu}}{\bar{a}^2}$ that is the same in all the spatial hypersurfaces. Also, for a flat universe, the resultant metric is conformal to the Minkowsky metric in these coordinates, that is,

$$ds^2 = \bar{a}^2(\eta) \left(d\eta^2 + dx^2 + dy^2 + dz^2 \right). \tag{2.30}$$

Thus, any geometric properties of objects defined in the Minkowsky space are preserved in the flat FLRW model.

The conformal time derivative of a general quantity f can be related to the cosmological time derivative through

$$\dot{f} = \frac{\partial f}{\bar{a} \partial \eta}. \tag{2.31}$$

Very often we will utilize the Hubble function in conformal time $\mathcal{H}$ and its respective Friedman equations, given by

$$\bar{\mathcal{H}} \equiv \frac{H\bar{a}}{c} = \frac{\frac{\partial \bar{a}}{\partial \eta}}{\bar{a}}, \tag{2.32}$$

$$\bar{\mathcal{H}}^2 = \frac{\bar{a}^2 \kappa}{3} \bar{\rho} - \bar{a}^2 \tilde{K} c^2 \tag{2.33}$$

and

$$\frac{d\bar{\mathcal{H}}}{d\eta} = -\frac{\bar{a}^2 \kappa}{2} \left(\bar{\rho} + 3\bar{p} \right). \tag{2.34}$$

With the basic concepts developed in this chapter, we now begin to study Bouncing cosmology. We are interested in analyzing the Primordial Black Hole formation in a bouncing cosmology. For that, we first need to study the bounce model for the universe. In the next chapter, we start with well known cosmological problems and see how the inflation model may solve them and then move to the bounce model.

3 BOUNCING COSMOLOGY

Due to physical constraints such as the cosmological horizon debilitating the observable universe or technology restrictions, we are not able to access data from the beginning of the Universe to define its formation. Therefore we must propose models that can describe the early Universe while being able to predict its state in the present time. In this context, the Standard Big Bang (SBB) model with an inflation phase is the most used among the cosmology community [78]. However, despite its prediction capability, the SBB plus inflation model does not completely describe the Universe formation. For example, it leads to a singularity when we extrapolate the model in the past beyond the validity of classical GR [79]. Hence it is worth studying other models for the Universe formation, either for their own capability of describing the Universe or to compare with the inflation predictions.

We want to explore the Bounce model, which proposes the existence of a contraction phase with a decreasing scale factor before the current expansion phase [80, 81]. Although there are Bounce Inflation models based on a contraction phase followed by an inflating phase(see [82, 83]), in this work we focus on a Bounce with only a contracting phase, as discussed in Refs. [40, 28]. The latter provides a well-suited explanation for the universe formation, while the addition of inflation to the model would add more uncertainty since it needs the inflaton to be achievable. In the following sections, we first analyze some cosmological problems used as guidelines for the formation models and analyze how Inflation deals with them. Next, we discuss the Bounce model and how it approaches the same problems.

3.1 Cosmological Problems

Before the proposal of inflationary models, the cosmology community was searching for ways of justifying some properties of the universe that were not well described by the Standard Big Bang model, which we may conceive as "problems"that should be explained by any theory of the beginning of the Universe. In this work, we focus on the Flatness, Horizon, and Large Scale Structures problems [84, 85, 79], that we shall now discuss.

3.1.1 Flatness

Let us analyze the curvature of the universe. To conclude if the Universe is spatially flat or not, we can use Friedman‘s equation to define the curvature density, given by

$$\Omega_k = \frac{-\tilde{K}}{\bar{a}^2\bar{H}^2}, \tag{3.1}$$

where Ω_k is the critical curvature density of the universe and $\tilde{K}$ is the conformal curvature parameter. Assuming that the universe is governed by one fluid with an equation of state defined by w, if we take the first derivative of Eq. (3.1), we have

$$\dot{\Omega}_k = \bar{H}\Omega_k(1+3w). \tag{3.2}$$

Thus, for fluids with $w > -\frac{1}{3}$, if $\Omega_k < 0$ the curvature will always decrease with time and if $\Omega_k > 0$, the curvature will always increase with time. The case where $\Omega_k = 0$ is an unstable solution, that is, any perturbation will break the equilibrium and the system will culminate in one of the other two cases.

The results from the Planck satellite show that the universe is nearly flat today, that is, $\Omega_{k0} \approx 0$ [74]. Thus, since the universe is basically flat in the present time, it had to be fine-tuned to a very small value in the initial of the universe, otherwise it would have grown to higher scales and dominated over other parameters. For example, the universe is estimated to have started 13.787 Gyrs ago and so at the Planck scale ($t \approx 10^{-43}s$) for a matter-dominated universe, we need $|\Omega_k| \lessapprox 10^{-63}$ to assure that the curvature would not have evolved to higher values today [84].

This fine-tuning is not per se a problem, however, it is odd that we must set a parameter small as 10^{-63} to explain the evolution of the Universe. Also, it is possible that initially $\Omega_k = 0$, although there is no evidence of it. Let us now move to the Horizon problem.

3.1.2 Horizon

There are regions of the universe that had never been in causal contact, yet they have similar properties. The Big Bang model describes the universe as isotropic and homogeneous, but how is this possible in this scenario? This is called the horizon problem. Let us look at the Cosmic Microwave Background (CMB) radiation to better understand the scenario.

According to standard cosmology, photons had decoupled from the matter content present in the Universe at recombination time when the universe was close to $t_{dec} \approx 380,000$ years old [86]. After their decoupling, these photons have been traveling through the universe and are responsible for the CMB radiation that we see today. Thus the CMB is a snapshot of the universe at that time. Connecting to the horizon problem, we have seen that the CMB coming from multiple directions separated by large angular apertures on the sky has a homogeneous temperature up to fluctuations around 10^{-5}K [87]. We will now study if these regions had been in causal contact.

The particle horizon, the maximum distance that a photon could have traveled since the origin of the universe, is given by

$$\begin{aligned} \tilde{d}_{dec} &= c\int^{t_{dec}} \frac{\mathrm{d}t_1}{\bar{a}\left(t_1\right)} \\ &= c\int^{a_{dec}} \frac{\mathrm{d}\bar{a}}{\bar{a}\bar{H}\left(\bar{a}\right)}. \end{aligned} \tag{3.3}$$

Assuming that the universe was dominated by radiation until the recombination, we can compute this integral to be

$$\tilde{d}_{dec} = 2c\frac{H_0^{-1}}{\sqrt{\Omega_{r0}}}\bar{a}_{dec}^2 \approx 2cH_0^{-1} \times 10^{-2}. \tag{3.4}$$

where we used $\bar{a}_{dec} = 10^{-3}$ and $\sqrt{\Omega_{r0}} \approx 9.6 \times 10^{-3}$. Thus, at recombination time, any regions separated by a radius of more than $\tilde{d}_{dec}$ could not have had a causal contact. We can also measure the comoving distance between us and the recombination time, which is given by

$$\tilde{d}_{dec-t0} = 2c\frac{\bar{H}_0^{-1}}{\sqrt{\Omega_{m0}}}\left(\sqrt{\bar{a}_0} - \sqrt{\bar{a}_{dec}}\right) \approx 2c\bar{H}_0^{-1}, \tag{3.5}$$

where we set $a_0 = 1$ and consider the universe to be dominated by dust in this interval. Thus, we are seeing a homogeneous volume element with a comoving radius of $\tilde{d}_{dec-t0}$. However, there are about 10^5 disconnected regions in this volume, that is, regions that had never interacted between themselves. Why do these regions have the same temperature and properties? This is a major issue not explained by the classical cosmological approach. Lastly, we are going to discuss the large-scale structures problem.

3.1.3 Large Scale Structures

For several years, many scientists studied the origin of large-scale structures in the universe, such as a galaxy, galaxy clusters, and other cosmological objects, which led to the conclusion that the universe itself is not actually homogeneous and isotropic [88]. This homogeneous hypothesis shall be interpreted as a first approximation, such that the real universe is explained by a perturbation theory around it (which can be seen in details in App. A). In this context, we want to analyze how these fluctuations evolve.

The dynamics of the primordial perturbations in the matter content of the universe can be expressed in the Fourier space as

$$\delta\ddot{\phi}_{\mathbf{k}} + \frac{3\bar{\mathcal{H}}}{c}\delta\dot{\phi}_{\mathbf{k}} + \frac{\mathbf{k}^2}{a^2}\delta\phi_{\mathbf{k}} = 0, \tag{3.6}$$

where $\delta\phi_{\mathbf{k}}$ is a scalar field perturbation in the Fourier space and $\mathbf{k}$ are the modes. For perturbations with a wavelength smaller than the Hubble radius ($\frac{\mathbf{k}^2\dot{\phi}_{\mathbf{k}}}{\bar{a}^2} \gg \frac{\phi_{\mathbf{k}}\bar{\mathcal{H}}}{c}$), different modes evolve differently, following a harmonic oscillator solution. For perturbations with ($\frac{\mathbf{k}^2\dot{\phi}_{\mathbf{k}}}{\bar{a}^2} \ll \frac{\phi_{\mathbf{k}}\bar{\mathcal{H}}}{c}$) and in a de Sitter universe, we have a constant mode solution and a decreasing contribution. A decreasing contribution would imply that the perturbations were large enough at the beginning of the universe such that they were reduced to small values today. This would not agree with any linear perturbation theory where we require the perturbations to be small [89]. On the other hand, the lack of concrete evidence of such large perturbations leads to the discard of such modes.

If we analyze both limits, we can see that for large wavelengths ($\frac{\mathbf{k}^2\dot{\phi}_{\mathbf{k}}}{\bar{a}^2} \ll \frac{\phi_{\mathbf{k}}\bar{\mathcal{H}}}{c}$), the modes are constant and are considered to be frozen, while for wavelengths smaller than the Hubble radius ($\frac{\mathbf{k}^2\dot{\phi}_{\mathbf{k}}}{\bar{a}^2} \gg \frac{\phi_{\mathbf{k}}\bar{\mathcal{H}}}{c}$), each mode evolves differently in time. For the frozen modes, considering a fluid with $w > -\frac{1}{3}$, we can trace back their evolution to the origin of the universe and conclude that these modes have always been frozen [90]. This behavior is not predicted by the standard big bang model and we would have to place these correlations at the initial condition of the universe. Also, there is no explanation for why the modes behave differently for different scales. In the next section, we briefly going to see how inflation approaches the problems above.

3.2 Inflation Approach and its Solution to Cosmological Problems

The theory of inflation was first proposed in the 80's to improve the Big Bang model as an attempt to solve the cosmological problems [91, 92, 93]. The usual inflation model consists in a accelerated expansion phase in the early universe seeded by an scalar field such that the Hubble function is almost constant and the universe is governed by a scale factor that growths exponentially. To achieve this regime, some models use the slow-roll approach based in an attractive potential that forces the scalar field to decay slowly assuring that the Hubble function is almost constant. Note that there are several inflation models with different assumptions. We will not get into much details in this section since our goal is only to analyze how inflation solves the problems above. For a more complete review, see Refs. [94, 84].

Let us go back to the flatness problem and analyze Eq.(3.1). In inflation, the Hubble is function is considered to be constant and the scale factor follows an exponential relation, that is,

$$\bar{a} = \bar{a}_{\lambda} \exp(Ht), \tag{3.7}$$

where $\bar{a}_{\lambda}$ is a constant. In this model, the scale factor is rapidly increasing and from Eq.(3.1) we can see that the curvature density will decrease in this regime. At the end of inflation, the curvature begin to grow in the standard Big Bang model. As long as inflation endures a sufficient amount of time, the curvature density will be close to zero at the end of this regime and despite its growth in the following epochs, it will remain negligible if set to smaller enough values initially.

Regarding the horizon problem, inflation solves the issue in a similar manner. We need that the particle horizon at the end of inflation to be bigger than the event horizon from recombination until today. Explicitly, we need

$$\begin{gathered} \tilde{d}_{dec} \geq \tilde{d}_{dec-t0} \text{ or} \\ c\int^{t_{dec}} \frac{\mathrm{d}t_1}{\bar{a}\left(t_1\right)} \geq c\int_{t_{dec}}^{t_0} \frac{\mathrm{d}t_1}{\bar{a}\left(t_1\right)}. \end{gathered} \tag{3.8}$$

The second integral remains the same since inflation takes place in earlier times. For the integral on the left, we can divide into three intervals, the beginning of the universe until the start of inflation, the inflationary period, and the time between the end of inflation and recombination. Thus the integral is given by

$$\tilde{d}_{dec} = c\int_0^{t_{t_i}} \frac{\mathrm{d}t_1}{\bar{a}\,(t_1)} + c\int_{t_i}^{t_f} \frac{\mathrm{d}t_1}{\bar{a}\,(t_1)} + c\int_{t_f}^{t_{dec}} \frac{\mathrm{d}t_1}{\bar{a}\,(t_1)}, \tag{3.9}$$

where t_{t_i} is the beginning of inflation and t_{t_f} is the end. Consider that all intervals are radiation dominated, the first and third integrals are negligible and we can compute the particle horizon to be

$$\begin{aligned}\tilde{d}_{dec} &= c\int_{t_i}^{t_f} \frac{\mathrm{d}t_1}{\bar{a}\,(t_1)}\\ &= c\int_{t_i}^{t_f} \frac{\mathrm{d}t_1}{\bar{a}_\lambda e^{\bar{H}t_1}\,(t_1)} = \frac{c}{\bar{H}}\left(\frac{1}{\bar{a}_{t_i}} - \frac{1}{\bar{a}_{t_f}}\right)\\ &= \frac{c}{\bar{H}}\frac{1}{\bar{a}_{t_i}},\end{aligned} \tag{3.10}$$

such that we used Eq.(3.7) in the second passage and considered that inflation lasted a good amount of time, that is, $t_i \ll t_f$. We also used the fact that H is constant during inflation. We can replace the Hubble constant's value given by $\bar{H} = \bar{H}_0\sqrt{\Omega}_{r0}\frac{\bar{a}_0^2}{\bar{a}_{t_f}^2}$ in Eq.(3.10) such that

$$\tilde{d}_{dec} = \frac{c}{\bar{H}_0}\frac{x_{t_i}}{x_{t_f}^2}. \tag{3.11}$$

Going back to Eq. (3.8), we need

$$\begin{aligned}\frac{\tilde{d}_{dec}}{\tilde{d}_{dec-t0}} = 10^{-2}\times\frac{x_{t_f}^2}{x_{t_i}} \geq 1 \;\text{ or}\\ N \geq \ln\left(10^{-2}\times x_{t_f}\right)\end{aligned} \tag{3.12}$$

such that we define the number of *e-folds* N with $x_{t_f} = e^N x_{t_i}$. This variable dictates the necessary rate of expansion of the universe during inflation such that the cosmological problems are solved. Supposing that inflation ends close to the Big Bang Nucleossintesis ($x_b \approx 10^{11}$), $N \geq 21$. The more in the past we put the end of inflation, the bigger N has to be.

For the large scale structures problem, regarding the inflation scenario where the Hubble radius is constant, during the exponential expansion of the universe the perturbation's wavelengths increase until they reach the size of the Hubble radius while its amplitudes decrease. Then, their amplitude is frozen while the wavelengths continue to growth with the expansion rate of the universe. At the end of inflation, these perturbations reenter the horizon with the same amplitude that they had when leaving the Hubble radius. Therefore this mechanism predicts the behaviors expected by the solutions discussed in Sec. 3.1.3(see Ref.[95] for a more detailed review).

Although inflation solves the above problems, the model predicts some behaviors that are not well explained by the standard Big Bang model. The most outstanding problem is that if we extrapolate the inflation model to the very beginning of the universe, the scale factor becomes infinitesimally small and the universe is seeded by a primordial singularity. In the next section, we will see how a bouncing cosmology can be an alternative to inflation.

3.3 Bounce Model

In this section, we briefly discuss our adopted quantum bouncing background model, for which we follow mostly Refs. [41, 27]. The primary motivation behind considering quantum bounces lies in the necessity to evade the primordial singularity at the end of the contracting phase. Classical General Relativity, as described by the Penrose-Hawking singularity theorems, requires the violation of the null energy condition $\rho + p \geq 0$ to avoid the singularity, a condition that often leads to instabilities in classical bouncing models.

Our quantum bouncing model is established through the canonical quantization of General Relativity. The flat-dust bounce is thus a contracting universe model with a negative effective energy density that dominates near the bounce, while classical behavior prevails on larger scales. Nonetheless, the quantum contribution is sufficient to avoid the singularity. This approach represents a conservative method in addressing quantum gravity, as it applies the usual Dirac quantization techniques of constrained systems to General Relativity [96].

To apply canonical quantization to GR, we adopt the standard Hamiltonian formalism. The complete Wheeler-De Witt (WdW) equation that will arise from the Dirac quantization poses several challenges that can be solved by assuming some additional hypothesis [97, 98, 27, 99], which we now discuss. However, we should emphasize that the actual bouncing mechanism is irrelevant to our results since the PBHs with relevant scales are formed way before the quantum phase.

3.3.1 Additional Hypothesis for Quantization

The first problem arises because the WdW equation is formulated on superspace, which represents the space of all possible metrics modulo diffeomorphisms and remains poorly understood [98, 100]. Therefore, we shall only perform quantization on a well-behaved sub-space that possesses the required symmetries, a procedure that is known as a mini-superspace quantization. Thus, we quantize only the sub-space of all possible flat FLRW geometries whose line element is given by the form

$$ds^2 = -N^2 dt^2 + \bar{a}^2 \delta_{ij} dx^i dx^j, \tag{3.13}$$

where $\bar{a}$ is the scale factor, N is the lapse function and δ_{ij} is the Kronecker delta. In this case, all information about the metric is stored in only one degree of freedom, the scale factor $\bar{a}(t)$ [27].

Another problem is the lack of an explicit time evolution in the Hamiltonian of the theory. This can be seen by attempting to write the conventional Schrödinger equation for the total Hamiltonian and stating the lack of a clear time evolution parameter. This fact is commonly referred to as the Problem of Time in Quantum Gravity [101, 41, 102]. To define a non-trivial propagator, we shall use an intrinsic variable of the system whose classical evolution is monotonic. Subsequently, we require that the classical concept of time emerges from this variable in the classical limit. Let us now apply such considerations and discuss the solutions for this quantization procedure.

3.3.2 Wheeler-DeWitt Equation Solutions

We start by applying the aforementioned considerations to a flat, homogeneous, and isotropic universe containing a single perfect fluid characterized by the equation of state $p = w\rho$, with a constant w. Subsequently, we specialize in the dust case where $w \approx 0$. In this framework, the total Hamiltonian takes the form

$$H = N\left(-\frac{\Pi_{\bar{a}}^2}{4\bar{a}} + \frac{\Pi_T}{\bar{a}^{3w}}\right), \tag{3.14}$$

where

$$\Pi_{\bar{a}} \equiv \frac{\partial L}{\partial \dot{\bar{a}}}, \tag{3.15}$$

$$\Pi_T \equiv \frac{\partial L}{\partial \dot{T}}, \tag{3.16}$$

are the canonical momenta associated with the scale factor $\bar{a}$ and T, the latter being a variable that describes the fluid's evolution. Also, we denote the cosmic time derivative as $\frac{\partial q}{c\partial t} \equiv \dot{q}$, which has a dimension of one over length. Remarkably, Π_T appears linearly in the Hamiltonian, and its classical equation of motion is given by $\dot{T} = N/\bar{a}^{3w}$, implying $\dot{T} > 0$ and making T a monotonic function of classical time. Hence, T can be used as a time variable, allowing us to write a Schrödinger-like equation by choosing $N = \bar{a}^{3w} \implies \dot{T} = 1$. For the details, see Ref. [27].

Canonical quantization is then performed by promoting classical variables to operators satisfying the canonical commutation relations. This process yields the following Wheeler-DeWitt equation for the wavefunction of the universe $\Psi(\bar{a}, T)$ [41]

$$i\frac{\partial}{\partial T}\Psi(q, T) = \frac{1}{4}\frac{\partial^2}{\partial q^2}\Psi(q, T), \tag{3.17}$$

such that a specific operator factor ordering was chosen to preserve the symmetries of the classical system [98]. Also, we introduced the change of variable

$$q \equiv \frac{2\bar{a}^{\frac{3}{2}(1-w)}}{3\,(1-w)}. \tag{3.18}$$

Eq. (3.17) resembles a time-reversed free particle Schrödinger equation and, with appropriate boundary conditions, its solutions are wavefunctions of the scale factor $\bar{a}(T)$. We turn to the De Broglie-Bohm interpretation [103, 97], such that assuming a Gaussian wavefunction $\Psi(\bar{a}, T)$, the Bohmian trajectory solution reads

$$\bar{a}(T) = \bar{a}_B\left[1 + \left(\frac{T}{T_B}\right)^2\right]^{\frac{1}{3}\frac{1}{(1-w)}}, \tag{3.19}$$

where $\bar{a}_B$ is an integration constant that represents the minimum scale factor value and T_B is a small arbitrary constant related to the time scale of the bounce. Remarkably, $\bar{a}(t) \neq 0$ for all t, which means that this model is non-singular and represents an eternal universe [41]. For a detailed derivation, see [27].

Given that we focus solely on a contracting universe model filled with dust, we set $w \approx 0$, simplifying the time variable T to the conventional cosmic time t. With the scale factor obtained in Eq. (3.19), we derive its associated Hubble function

$$\bar{H}(t) \equiv \frac{1}{\bar{a}}\frac{d\bar{a}}{dt} = \frac{2}{3}\frac{t}{\left(t^2 + t_b^2\right)}, \tag{3.20}$$

and invert (3.19) to obtain

$$t(\bar{a}) = \pm t_b \sqrt{\left(\frac{\bar{a}}{\bar{a}_b}\right)^3 - 1}. \tag{3.21}$$

Eliminating the time t using (3.21) in (3.20), we find

$$\bar{H}^2 = \frac{4}{9t_b^2}\left(\frac{\bar{a}_b^3}{\bar{a}^3} - \frac{\bar{a}_b^6}{\bar{a}^6}\right), \tag{3.22}$$

which is equivalent to the Friedmann equation

$$\bar{H}^2 = \frac{\kappa c^2}{3}\rho - \bar{H}_0^2 \Omega_{q0} \bar{a}^{-6}, \tag{3.23}$$

where $\kappa = \frac{8\pi G}{c^4}$ and G is the gravitational constant. In the above, there is an additional term $-\bar{H}_0^2\Omega_{q0}\bar{a}^{-6}$ when compared to the usual Friedmann equation for a classical universe. In our model, the scale factor dynamics mirrors a typical Friedmann equation with a total energy density $\rho_T = \rho_d + \rho_q$, where ρ_d represents the dust energy density and $\rho_q = -\Omega\rho_c\bar{a}^{-6}$ denotes an effective negative energy density accounting for quantum effects that circumvent the primordial singularity [51]. For large-scale factors $\bar{a} \gg \bar{a}_b$, Eq. (3.23) and Eq. (3.19) reduce to the classical dust dominated flat FLRW universe. Thus, far from the bounce, we still have a classical behavior, and the quantum corrections only become relevant when $t \approx t_b$. With this, we conclude the analysis of our background quantum bouncing model and move to its associated perturbations. In the next sections, we will consider the above quantum bouncing model with $w \approx 10^{-10}$ and $x_b = 10^{35}$ and we solve the equations numerically without any approximation. We will also analyze our results for different equations of state.

3.3.3 Solution to Cosmological Problems

We now analyze how the same cosmological problems can be solved via bounce. To facilitate comparison, we focus on a bounce phase dominated by radiation, enabling a direct analogy to previous computations. This implies that

$$E(t) = \begin{cases} -\Omega_{r_0} x^4 & \text{for } t_i < t < t_b \\ \Omega_{r_0} x^4 & \text{for } t > t_b \end{cases}, \tag{3.24}$$

where t_i is the time at the beginning of the universe before the contracting phase and t_b is the bounce time.

Let us first examine how this model resolves the flatness problem. According to the standard cosmological model, we know that

$$\Omega_k = \Omega_{k_0} x^2 \tag{3.25}$$

and

$$E^2 = \Omega_{r_0} x^4 + \Omega_{m_0} x^3 + \Omega_{k_0} x^2 + \Omega_{\Lambda_0}. \tag{3.26}$$

At t_i, even if there is a significant contribution from Ω_k, it quickly becomes negligible compared to radiation and matter as x increases during the contracting phase. Thus we expect a small curvature density at the beginning of the expansion phase.

For the horizon problem, we use a similar approach. Going back to Eq. (3.8), we must re-evaluate the left side for the contracting phase, that is

$$\tilde{d}_{bounce} = \int_{t_i}^{t_b} \frac{\mathrm{d}t_1}{\bar{a}(t_1)}. \tag{3.27}$$

Considering that $t_i \gg t_b$, we have

$$\tilde{d}_{bounce} = \frac{R_H}{\sqrt{\Omega_{r_0}}} \left(\frac{1}{x_i} - \frac{1}{x_b} \right) \approx \frac{R_H}{\sqrt{\Omega_{r_0}}} \frac{1}{x_i}. \tag{3.28}$$

In the above expression, we consider that the bounce occurs at smaller scales, such that $x_b \gg x_i$. Assuming that the interval between the beginning of the universe and the bounce is large enough, $t_i = -\infty \to a_i = \infty$, such that $\tilde{d}_{bounce} \to \infty$. Therefore the distance that a particle could have travel since the beginning of the universe is large enough such that the inequality in Eq. (3.8) is satisfied and there is causality effect between the disconnected regions that we see today in the CMB.

Lastly, regarding the large scale structures problem, we need to analyze the evolution of the wavelength of the perturbations in the bounce model. From Eq. (3.6), the wavelength evolve with the inverse of the scale factor. Thus, for $t = t_i$, all the large scales will remain smaller than the Hubble radius since $\bar{a}_i = \infty$ and $\frac{\mathbf{k}^2 \dot{\phi}_{\mathbf{k}}}{\bar{a}^2} \gg \frac{\phi_{\mathbf{k}} \bar{\mathcal{H}}}{c}$ and thus there is no prediction of frozen modes. In the next chapter, we begin the study of primordial black holes. We later intend to analyze the formation of these objects in the bounce scenario.

4 PRIMORDIAL BLACK HOLE FORMATION IN A DUST BOUNCING MODEL

In this chapter, we consider the single barotropic fluid quantum bouncing model developed in [41, 97, 27] using Canonical Quantum Gravity. This is a conservative approach to the quantization of General Relativity and should hold as an effective theory up to a certain energy scale [27]. We shall focus on the critical collapse of matter perturbations characterized by the density contrast δ, such that perturbations collapse to form black holes when they achieve a given threshold ($\delta > \delta_c$). To perform this analysis, we need to compute the spectra of the perturbations and the critical threshold needed for the distribution of PBHs. We compute the critical threshold in a more detailed approach, using the Tolman-Bondi-Lemaitre metric [104, 105] in a similar way as done in [23]. The perturbations' power spectra will be obtained through an algorithm that computes the valid interval for the adiabatic approximation and solves the dynamics of the perturbations. With these results, we will compute the abundance of PBHs and its fraction that can constitute dark matter.

4.1 Primordial Black Holes

Primordial Black Holes (PBH) are black holes formed in the early Universe through the collapse of density fluctuations. The investigation of these objects began in 1966 with Zeldovich and Novikov [1]. In their work, it was discussed that for a PBH to be formed at a time t after the Big Bang where the Universe density was $\rho \sim 1/Gt^2$, it would be necessary a density of $\rho \sim c^6/(G^3M^2)$ for an object with horizon mass M to fall into the Schwarzschild radius. Thus, the black holes formed in such times would have masses of the order

$$M \sim \frac{c^3 t}{G} \sim 5 \times 10^{-19} \left(\frac{t}{10^{-23} s}\right) M_{\odot}. \tag{4.1}$$

In contrast to stellar black holes, which can only present masses above $3M_{\odot}$ due to the Tolman–Oppenheimer–Volkoff limit [106, 107], the formation mechanism of PBHs allows for a wider range of masses.

Initially, Zeldovich and Novikov concluded that the existence of PBHs was improbable. Considering Brandi's accretion, PBH of the size of the horizon would have grown at the same rate as the horizon and would reach a mass of $10^{17} M_\odot$ by the end of the radiation-dominated era. However, the absence of evidence or detection of such massive black holes led to the rejection of the existence of PBHs. The PBH formation was later revisited by Hawking in 1971 [2]. He explored the possibility of primordial black holes of the size of elementary particles, which motivated Hawking's investigation on quantum corrections in black hole formation and eventually led to the discovery of Hawking radiation in 1974 [3]. In the same year, Carr and Hawking demonstrated that there is no self-similar solution in which black holes would grow as fast as the horizon in an expanding Universe [4]. This rekindled interest in the study of primordial black holes.

The existence of Hawking radiation has set a lower bound for the detection of primordial black holes in the present era. Black holes with masses $M \leq 10^{-17} M_\odot$ would have completely evaporated due to Hawking radiation, making their direct detection impossible[1]. Consequently, attention has shifted towards PBHs with $M > 10^{-17} M_\odot$, which could exhibit observable gravitational effects.

The study of primordial black holes has yielded insights into various phenomena. They could have contributed to the cosmological density parameter [5, 6]. Furthermore, their influence in the cosmic microwave background was studied in Ref. [7] and some recent works explored the connection between evaporating PBHs and gravitational waves [8]. There also have been discussions about whether some measurements of gravitational waves could be attributed to primordial black holes [9]. Undoubtedly, the possibility that primordial black holes constitute a significant fraction of the cold dark matter remains a focal point in current research, which we will elaborate on in the next section.

4.2 PBH as a probe for dark matter

The nature of Dark Matter (DM) remains a fundamental mystery in modern cosmology. Extensive measurements conducted by Planck and WMAP have revealed that DM accounts for up to 85% of the total matter density of the present Universe [110],

---

[1] References such as [108] and [109] have extensively studied the effect of this evaporation.

demonstrating its importance. Over the years, various candidates, including axions, inert Higgs Doublet, and others, have been proposed for DM [111]. Among these, primordial black holes emerged as an intriguing suggestion for DM, first proposed by Chapline in 1975 [10]. Since then, numerous researchers, such as Carr and Kühnel [24] and Villanueva et al. [25], have extensively explored this idea.

As primordial black holes are formed before the Big Bang Nucleosynthesis (BBN), when particles condensed into protons and neutrons, they are not constrained by the fact that baryonic matter can only account for up to 5% of the total critical density of the Universe [112]. Unfortunately, there is no concrete evidence that PBHs constitute DM. The search for massive halo objects (MACHOs) through microlensing suggested that objects with mass $M \sim 0.5 M_{\odot}$ could constitute cold dark matter [113], which points to the possibility of PHBs. Additionally, recent observations of merging binary black holes by the LIGO/Virgo collaborations indicate that they could have originated from PBHs [11, 12]. They detected black holes in the unexpected mass range ($10 M_{\odot} - 50 M_{\odot}$), which can be explained by PBH.

Several studies have investigated the fraction of dark matter that could have been constituted by PBHs ($f(M)$) depending on their formation mass. Not all ranges are relevant to this hypothesis and according to Carr et al. [114], the interest relies on the asteroid to sublunar mass range ($2 \times 10^{-17} - 2 \times 10^{-11} M_{\odot}$), the intermediate range ($10 - 10^{2} M_{\odot}$) and the large range ($M > 10^{11} M_{\odot}$).

PBH has the potential to impact various areas of cosmology, resulting in diverse observational effects, which can be utilized to constrain their possible contribution to dark matter. In a recent study by Villanueva [25], a comprehensive analysis was conducted considering a variety of effects such as the evaporation of PBH, microlensing, gravitational waves, dynamical constraints, and others, aiming to obtain bounds for the fraction of PBH in dark matter $f = \Omega_{PBH}/\Omega_{c}$. When combining the probes, it was found that the only acceptable range that allows PBHs to entirely constitute dark matter is within the mass range $10^{-16} M_{\odot} - 10^{-12} M_{\odot}$, as we can see in Fig. 1. Nonetheless, even for mass ranges in which PBHs represent smaller fractions of DM, there remain intriguing impacts to consider. For instance, in the mass range $10^{6} M_{\odot}$ to $10^{10} M_{\odot}$ where PBHs only represent 0.1% of DM, they could play a role in generating supermassive black holes. Finally, it is worth noticing that their mass depends on the formation mechanism. In this work, we assume the formation through spherical collapse, which will be further explored in the subsequent section.

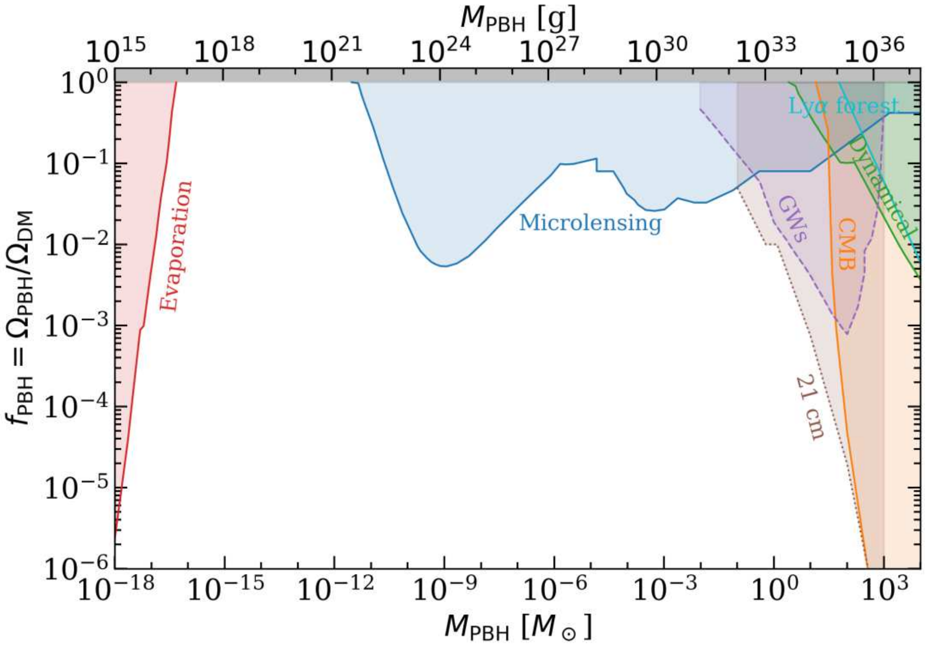


Figura 1 – Constraints on the PBH fraction constituting dark matter. Each color correspond to a different probe considered to constrain PBH abundance. The uncolored area of the graph correspond to the possible fraction values that are in agreement with all probes. Figure from Ref. [25].

4.3 Gauge Invariant Perturbations

We follow mainly the perturbation theory developed in Refs. [29, 38] and studied in App. A and B. In this section, we summarize some relations between our perturbation variables that will be needed and discuss our Gauge choice. Assuming only scalar perturbations, the total metric of the physical space-time is given by

$$ds^2 = -(1-2\phi)c^2dt^2 + \bar{a}\bar{D}_i\mathcal{B}cdtdx^i + \bar{a}^2(1-2\psi)\delta_{ij}dx^idx^j - \bar{D}_i\bar{D}_j\mathcal{E}dx^idx^j. \quad (4.2)$$

Here, $\phi, \psi, \mathcal{B}$ and $\mathcal{E}$ denote the scalar metric perturbations that we will assume to be much smaller than one [51]. Also, $\bar{D}_i$ is the spatial covariant derivative in the i-th direction. Note that our barotropic fluid has no anisotropic pressure.

The metric perturbations are gauge-dependent variables. In cosmology, a gauge can be seen as the freedom in how we connect, or map, the background and physical manifold, and how we choose our coordinate system [51]. Since GR is a covariant theory, this freedom may then be interpreted as a gauge, which may lead to an ambiguous description of perturbations. Depending on the foliation that characterizes the manifold hyper-surfaces and how we define the perturbations around our background, the physical quantities may have different values [29]. Thus, it is recommended that we work with gauge independent variables to carry out our computations and go back to the physical variables at the end when necessary [103].

We define gauge invariant quantities by analyzing their transformations under gauge transformations, such that we can combine different gauge-dependent quantities to form new invariant ones [29]. However, this leads to a freedom on the variable definitions since many combinations of variables may lead to gauge invariant quantities. With this in mind, we define the Bardeen gauge invariant variables [115]

$$\Phi \equiv \phi + \dot{\delta\sigma}\,, \tag{4.3}$$

$$\Psi \equiv \psi - \frac{\bar{\Theta}\delta\sigma}{3}, \tag{4.4}$$

with

$$\delta\sigma \equiv -(\dot{\mathcal{E}} - \mathcal{B}) + \frac{2\bar{\Theta}\mathcal{E}}{3}. \tag{4.5}$$

It is important to notice that the new variables in Eqs. (4.3)-(4.4), and other gauge invariant variables do not have a physical meaning unless a gauge is chosen. For instance, in the case of the Newtonian gauge, $\delta\sigma = 0$ and Φ represents the Newtonian potential. Hence we must define our gauge invariant quantities such that they represent our desired physical variables when we assume a particular gauge choice.

We are mostly interested in the energy density perturbations that collapse to form PBHs. These perturbations can be examined using the density contrast, defined as

$$\delta \equiv \frac{\delta\rho}{\bar{\rho} + \bar{p}}. \tag{4.6}$$

This variable provides a normalized measurement of the energy density perturbation $\delta\rho$ around the background matter density field. However, once again we are interested in its gauge invariant form

$$\widetilde{\delta\rho}^{F} \equiv \delta\rho - \mathcal{V}\dot{\bar{\rho}}, \tag{4.7}$$

where $\mathcal{V}$ is the velocity perturbation of the fluid, which has its gauge invariant form

$$\tilde{\mathcal{V}} \equiv \mathcal{V} + \delta\sigma. \tag{4.8}$$

Using the background Einstein equations, we can relate the density contrast to the Bardeen variables and the gauge invariant curvature perturbation ζ through the following relations [116, 39]

$$\Psi = \Phi, \tag{4.9}$$

$$-\frac{2\bar{D}^2\Phi}{3\kappa(\bar{\rho}+\bar{p})} = \frac{\widetilde{\delta\rho}^{F}}{3(\bar{\rho}+\bar{p})}, \tag{4.10}$$

$$\zeta \equiv \Psi + \frac{\tilde{\mathcal{V}}\bar{\Theta}}{3} \tag{4.11}$$

$$\zeta = \frac{3\bar{a}^3}{N^2 z^2 c_s^2}\left[\frac{\partial}{c\partial t}\left(\frac{\Phi}{\bar{\Theta}}\right) + \frac{\Phi}{3}\right], \tag{4.12}$$

$$\bar{D}^2\Phi = -\frac{z^2\bar{\Theta}}{3\bar{a}^3}\dot{\zeta} = -\frac{\bar{\Theta}}{6\bar{a}^3}\Pi_\zeta. \tag{4.13}$$

For completeness, we also write its relation to the usual Mukhanov-Sasaki variable v, that is,

$$v \equiv -\zeta z\sqrt{\frac{2}{\kappa c}}, \tag{4.14}$$

with $\kappa = \frac{8\pi G}{c^4}$. We will establish a connection between PBH formation and the density contrast in the next section. For now, it suffices to understand that the excess energy density associated with the perturbations leads to black hole formation, and we can measure such excess through the density contrast. Thus our definition in Eq. (4.7) is extremely important and we will now analyze it.

Note from Eq. (4.7) that if we choose a gauge where $\mathcal{V} = 0$, $\widetilde{\delta\rho}$ becomes the physical density perturbation, which leads to an easier interpretation of this quantity. Also, the choice of $\mathcal{V} = 0$ will be well suited to connect our perturbation theory with the

Lemaitre-Toman-Bondi metric discussed in App. A.5.4. Other works have defined the gauge invariant density perturbation as

$$\tilde{\delta\rho}^N \equiv \delta\rho - \delta\sigma\dot{\bar{\rho}}, \tag{4.15}$$

so that this quantity becomes the physical variable in the Newtonian Gauge, where $\delta\sigma = 0$ and $\Phi = \phi$. However, in Ref. [51], it was shown that for bounce models with long contracting phases, regardless of the bounce type, the Bardeen perturbation Φ grows larger than one and invalidates the perturbative series, as ϕ also grows larger than one in the Newtonian Gauge. Additionally, the definition in Eq. (4.15) leads to

$$\zeta = \Phi + \frac{2\bar{D}^2\Phi}{3\kappa(\bar{\rho}+\bar{p})} + \frac{\tilde{\delta\rho}^N}{3(\bar{\rho}+\bar{p})}, \tag{4.16}$$

which implies that $\tilde{\delta\rho}^N$ grows with Φ and thus has larger spectra as well in this gauge. Hence, the Newtonian gauge is not a valid choice for bounce models as it would lead to miss-calculation of the physical quantities whose values would be inflated. We shall avoid this choice in this work and stick with the definition in Eq. (4.7). We now compute the density contrast modes, which require the perturbative equations of motion of the model, displayed in Sec. B.1.1.

4.3.1 Numerical Solution

The equations of motion for the curvature perturbation modes are solved numerically using the Numerical Cosmology library (NumCosmo) [117]. Specifically, the NcmCSQ1D and the NcHIPertAdiab algorithms are employed for this purpose. The equations are split into two parts: one representing a harmonic oscillator with a mass m_A and the other representing the time evolution of the conjugate momenta $\Pi_{k_A,A}$ of the modes. Explicitly,

$$\Pi_{k_A,A} = m_A\dot{\zeta}_{k_A,A} \tag{4.17}$$

and

$$\dot{\Pi}_{k_A,A} = -m_A\omega_A^2\zeta_{k_A,A}, \tag{4.18}$$

The adiabatic vacuum prescription is considered, and the initial conditions for the modes are set according to Eqs. (B.34) up to the fourth order[2]. These initial conditions are used as inputs for the numerical algorithm to compute the Fourier modes of the curvature perturbation modes $\zeta_{k_A,A}$ and their conjugate momenta $\Pi_{k_A,A}$. Also, the numerical code computes the validity of the adiabatic approximation for different intervals of time. To assure the veracity of the code, the algorithms have been validated with unit testing and we compared them with analytical solutions (see GitHub)[3]

4.3.2 Spectra

To compute the spectra, we use Eq. (B.37), i.e.,

$$P_\zeta(k) \equiv \frac{k^3|\zeta_k(t_i)|^2}{2\pi^2}. \tag{4.19}$$

With the above expression and the numerical code described in the last section, we can plot the spectra of the theory in Fig. 2. Once again we have different initial times for the adiabatic limit for distinct modes. We also notice an increasing spectra for all the modes, such that they peak at the bounce time.

As known from the literature, the dust bouncing model has an approximate scale-invariant spectrum, that is

$$P_\zeta(k) \approx A k^{n_s - 1}, \tag{4.20}$$

where n_s is the spectral index. This is a general feature of single barotropic fluid quantum bouncing models, and the spectral index is related to the equation of state parameter w by[4]

$$n_s(w) = 1 + \frac{12w}{1+3w}, \tag{4.21}$$

which is nearly scale invariant for $|w| \ll 1$. However, the usual positive values of w lead to a blue tilted spectrum $n_s > 1$, which differ from CMB observations [119], consistent with the inflationary scenario prediction of a red tilted spectrum with $n_s \approx 0.96$. Although

---

[2] This is because the adiabatic approximation leads to an asymptotic series, whose precision drops as one goes up to a certain definite order [118]. In particular, the code determines the order of optimal precision.

[3] Also, see the Jupyter Notebook for an example on the usage of the code.

[4] For an explicit semi-analytic derivation, see [41].

the initial power spectrum is not consistent with observations, one must recall that our model is considering a pure dark matter-dominated universe, neglecting the effects of radiation. In particular, it has been suggested [27] that the inclusion of radiation may lead to a red almost scale-invariant spectrum. Therefore, this model must be understood as a first approximation to a more complete model, which is still under development. Furthermore, since most of the modes that influence our universe have crossed the Hubble horizon during dust domination, this means that this model, even with its simplicity, may still cover relevant information for the future complete model.

This result should not advocate for the prediction failure of the contracting scenario since some works also suggest a blue-tilted spectrum for inflationary models [120, 121]. For instance, in Ref. [122] they discuss the possibility of explaining the recent NANOGrav results with a blue-tilted specrta [123]. Also, the spectra are only dictated by a power law for narrow intervals of momenta, which implies that corrections to Eq. (4.20) must be applied. Nonetheless, we see that the contracting phase produces a nearly scale-invariant spectrum. In the next section, we begin to analyze the primordial black hole formation seeded by the scalar perturbations in our cosmological background.

The resulting plots in Fig. 2 depict the power spectra for a single momentum mode ($k_a = 0.1$) of the curvature perturbation, the Bardeen field variable, the density contrast, and the evolution of the scale factor over time. For visualization purposes, we use the same time parameter as the one in the numerical code, namely,

$$\frac{1}{\cosh(\tau_a)^2} \equiv \left[1 + \left(\frac{t}{t_b}\right)^2\right], \tag{4.22}$$

such that τ_a is dimensionless by definition. The plots show an increasing power spectrum for all fields, peaking at the bounce. Note that Φ has highly divergent spectra, as discussed previously, which shows the inapplicability of the Newtonian Gauge in this model. Additionally, the scale factor decreases with time until reaching its minimum value at the bounce.

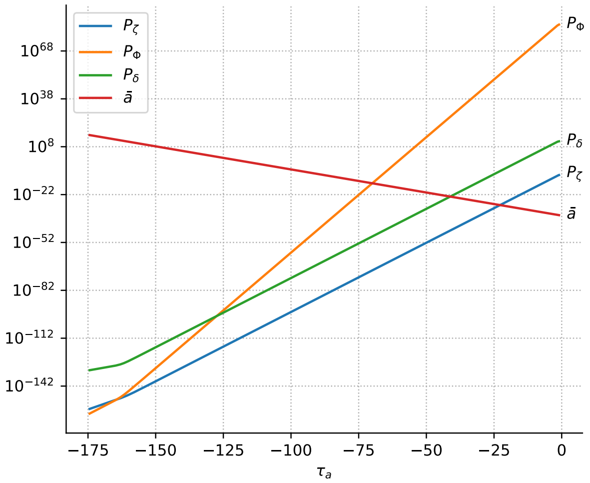


Figura 2 – Plot of the power spectrum of the curvature perturbation modes ζ_k, the Bardeen field variable Φ_k and the density contrast δ computed with the NumCosmo library for $w = 10^{-10}$ and $x_b = 10^{35}$. The red plot represents the scale factor evolution in time. The time parameter in the x-axis is given by Eq. (4.22).

4.4 Formation Criteria

In our context, critical collapse is the collapse of matter perturbations when they achieve a given threshold. The energy-density perturbation is characterized by the density contrast δ defined in Eq. (4.6). We assume that if $\delta > \delta_c$, where δ_c is the critical threshold, the perturbations will collapse into a black hole. This critical threshold relies both on the cosmological model and the shape of perturbations (see Ref. [124, 125] for a comprehensive analysis of this effect) and therefore must be carefully studied as it will heavily impact the formation of black holes.

The critical collapse model assumes the existence of a region of radius r with an overdensity δ. To compute the probability of the existence of this overdense region, we will need to compute the variance of our random variable δ. Instead of working with the original density contrast, we work with its filtered version δ_r, defined by

$$\delta_r(t) = \int d^3\mathbf{x}'\delta(t,\mathbf{x}')W_r\left(\mathbf{x}-\mathbf{x}'\right), \tag{4.23}$$

where r is a smoothing scale in comoving units related to the mean density by $M = \frac{4\pi r^3\bar{\rho}_m}{3}$

and W is the top-hat filter

$$W_r(\mathbf{x}) = W_r\,(|\mathbf{x}| = x) = \begin{cases} \frac{1}{4\pi r^3} & \text{if } x \leq r \\ 0 & \text{otherwise} \end{cases}. \tag{4.24}$$

The filter function is used to select a desired scale such that we would only analyze collapsed objects that lie in this interval[5]. In this work, we are using a simple top-hat windowed function for the filter. We usually work on the Fourier space with comoving wave-number k, such that the filter is given by

$$W_r(k) = \frac{3}{(kr)^2}\left(\frac{\sin(kr)}{kr} - \cos(kr)\right) = \frac{3}{kr} j_1(kr). \tag{4.25}$$

We will see that the filter is necessary to analyze scales where the PBH formation is most likely to happen.

Coming back to δ_r, we know from our quantum fields that the density contrast is a random field following a Gaussian distribution with a zero mean. This is also supported by significant research on the statistics of peaks in random Gaussian fields, which can give rise to collapsed objects, as explored in Ref. [126]. Mathematically, the probability of a region with radius r having an overdensity δ_i is given by

$$P(\delta_i) = \frac{1}{\sqrt{2\pi}\sigma_r} \exp\left(-\frac{\delta_i^2}{2\sigma_r^2}\right). \tag{4.26}$$

This Gaussian is completely defined by its variance

$$\sigma_r^2 \equiv \left\langle \delta_r^2(\mathbf{x}) \right\rangle = \frac{1}{2\pi^2} \int_0^\infty \frac{\mathrm{d}k}{k} \left[P_{\bar{\delta}}(k) |W_r(k)|^2 \right], \tag{4.27}$$

where $P_{\bar{\delta}}(k)$ is given by Eq. (4.19) with the density contrast modes computed with Eqs. (4.6) and (4.7). Unfortunately, for our bounce model, this integral does not converge. Before we perform this calculation, we will analyze the critical threshold for PBH formation, which will lead to scale constraints that will help us solve this problem.

---

[5] Frequently we will refer to such scale in terms of the mass M instead of the comoving radius r that encloses this mass.

4.5 Critical Threshold

In inflationary models, the frozen super-Hubble density perturbations re-enter the Hubble horizon at the end of the potential decay and collapse into a black hole if they have values above the critical threshold [127]. Hence the Hubble horizon is viewed as a characteristic scale for their formation and the threshold (δ_c) is obtained via the Minster-Sharp equations (see Ref. [125]). However, in the bounce scenario, the perturbations constantly evolve in time and there is not only one characteristic scale for the formation, since any perturbation above a critical threshold may collapse as they are not frozen. Thus, we must carefully analyze how to obtain this critical value starting with a local metric for the collapse.

In App. C, we find the solutions for Einstein's equations of a critical collapse represented with a local metric, which led to the LTB set of solutions

$$R(\theta, r) = \frac{r(1+\delta_{ini})}{\delta_{ini}} \sin^2\left(\frac{\theta}{2}\right), \tag{4.28}$$

$$t(\theta, r) = t_1(r) + \frac{1+\delta_{ini}}{2\bar{H}_{ini}\delta^{\frac{3}{2}}(t_{ini})} \left(\theta - \pi - \sin\theta\right). \tag{4.29}$$

In this section, we wish to impose some constraints on those solutions to find the exact point at which the perturbations will form a black hole. Consequently, this will lead to a critical value for the density contrast. First of all, let us analyze the parameter θ. From Eq. (C.39),

$$\theta_{ini} = \pm 2 \arcsin\left(\sqrt{\frac{\delta_{ini}}{1+\delta_{ini}}}\right). \tag{4.30}$$

To choose between the positive or negative value, we use our initial condition in Eq. (C.35) and Eq. (C.34) as a test, such that replacing Eq. (4.30)

$$\left.\frac{\dot{R}}{R}\right|_{ini} = \left.\frac{\frac{\partial R}{\partial \theta}}{R\frac{\partial t}{\partial \theta}}\right|_{ini} = \pm \bar{H}_{ini}. \tag{4.31}$$

From the above, we note that the positive sign in Eq. (4.30) leads to the right definition of our initial conditions. If one chooses the negative sign, and as $\bar{H}_{ini}$ is negative in the contracting phase, we would end up with an initially expending patch. This feature can be seen in Fig. 3, where Eqs. (4.28) and (4.29) were solved analytically with Wolfram

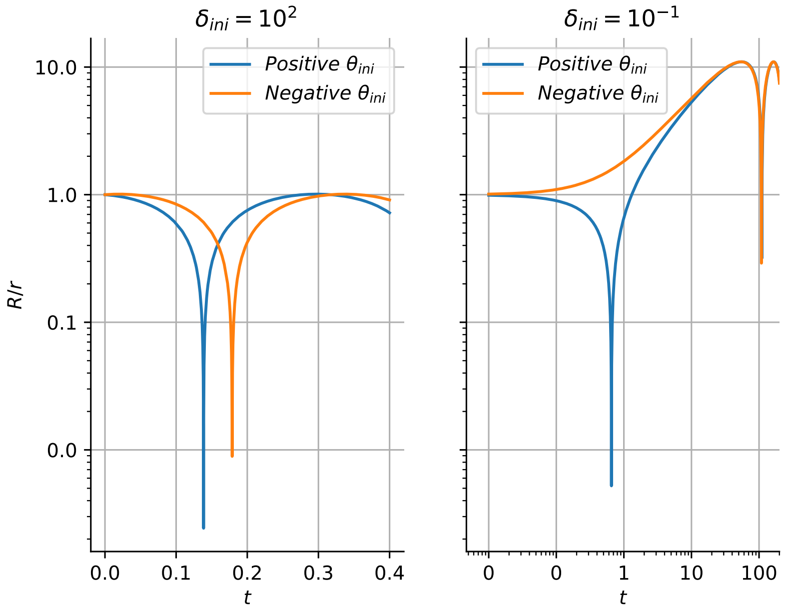


Figura 3 – Plot of R vs θ. We have chosen a dimensionless Hubble function, $\bar{H}_{ini} = -1$, for simplification purposes, which implies that t is in units of $1/\bar{H}_{ini}$. On the left side, we see the case for a large initial density contrast while the right side indicates a small initial density contrast. The blue and the orange lines indicate the positive and negative initial conditions for θ_{ini} respectively. The blue line represents the right choice and a contracting model while the orange line indicates an expansion model. We chose different values of δ_{ini} for both graphs as the difference between both choices becomes more evident for a smaller initial density contrast. The data was computed with Wolfram Mathematica [128] and the graphs were generated in Python.

Mathematica [128]. The blue lines represent the positive choice that leads to the LTB collapse model in a contracting universe, while the orange lines describe an LTB collapse model with an initially expanding patch. On the left side of the figure, for a larger value of δ_{ini}, we see that both solutions are similar as the orange plot has an insignificant initial expansion. For a smaller value of δ_{ini} on the right, the model in orange first goes on an expansion phase followed by the collapse, while the blue plot is already in a contracting phase and collapses much earlier.

Let us now analyze the PBH formation. The collapsed object will form a singularity when $\frac{\partial R}{\partial r} = R' = 0$ [129]. From Eq. (4.28) we see that

$$R' = \frac{(1+\delta_{ini})}{\delta_{ini}} \sin^2\left(\frac{\theta}{2}\right) = 0 \tag{4.32}$$

if $\theta = 0$. Consequently, we want to analyze the necessary time for a black hole to form, which takes place at $\theta = 0$. From Eq. (4.29) we have that the final time t_f, i.e., the time at formation is

$$t_f = t(0,r) = t_1(r) - \frac{\pi(1+\delta_{ini})}{2\bar{H}_{ini}\delta^{\frac{3}{2}}(t_{ini})}. \tag{4.33}$$

We want to compute the formation time $\Delta t = t_f - t_i$, such that $t_i = t(\theta_{ini}, r)$ with

$$\theta_{ini} = 2\arcsin\left(\sqrt{\frac{\delta_{ini}}{1+\delta_{ini}}}\right). \tag{4.34}$$

Hence, from Eq. (4.29), we have that

$$\begin{aligned}\Delta t &= t_f - t_i \\ &= \frac{(1+\delta_{ini})}{2\bar{H}_{ini}\delta_{ini}^{\frac{3}{2}}} \left(-\theta_{ini} + \sin\theta_{ini}\right), \\ &= \frac{(1+\delta_{ini})}{2\bar{H}_{ini}\delta_{ini}^{\frac{3}{2}}} \left(-2\arcsin\left(\sqrt{\frac{\delta_{ini}}{1+\delta_{ini}}}\right) + \sin\left(2\arcsin\left(\sqrt{\frac{\delta_{ini}}{1+\delta_{ini}}}\right)\right)\right).\end{aligned} \tag{4.35}$$

Assuming that $\delta_{ini} \ll 1$, we can expand the above relation such that

$$\Delta t = -\frac{2}{3\bar{H}_{ini}} + \frac{2\delta_{ini}}{15\bar{H}_{ini}} + O^2(\delta_{ini}). \tag{4.36}$$

We now know how to compute the time interval with the right side of the above equation. We shall now study the left side of Eq. (4.36) as we want to find a constraint on the final formation time t_f. Supposing that the PBH is formed way before the bounce[6], we can approximate Eq. (3.20) to

$$\bar{H} = \frac{2}{3t}. \tag{4.37}$$

---

[6] We do not make this approximation in the numerical code. This approximation is only used to demonstrate our calculations.

Plugging this result for t_i on the left side of Eq. (4.36) leads to

$$t_f = \frac{2\delta_{ini}}{15\bar{H}_{ini}}. \tag{4.38}$$

Let us rewrite the above expression in terms of the redshift function x defined as

$$x(t) \equiv \frac{\bar{a}_0}{\bar{a}(t)}, \tag{4.39}$$

such that the Hubble function can be written as

$$\bar{H} = -\bar{H}_0\sqrt{\Omega_w}x^{3/2} \tag{4.40}$$

and thus Eq. (4.38) becomes

$$t_f = -\frac{2\delta_{ini}}{15\bar{H}_0\sqrt{\Omega_w}x_{ini}^{3/2}}. \tag{4.41}$$

We need to impose a limit on the formation time to find the critical values of delta. For collapsed objects larger than the Hubble radius, their dynamics will be dominated by the FLRW metric and not the LTB approximation. The supper-Hubble perturbations will be frozen and thus there is no collapse for these scales. This sets the Hubble length as an upper cut-off, i.e., their formation time must be at most the time for the perturbation to achieve the Hubble radius, labeled as t_H. The radius is related to the perturbation's wavelength λ and comoving wavenumber k by [48]

$$r = \frac{\lambda}{2}, \tag{4.42}$$

$$\lambda = \frac{2\pi}{k}. \tag{4.43}$$

We want to analyze when a perturbation with wavenumber k has the same size as the Hubble radius, that is,

$$k_H = \frac{1}{xc}\left|\bar{H}\right|. \tag{4.44}$$

From Eq. (4.44) and Eq. (4.40), we have that the redshift function associated with this scale is given by

$$x_H = \frac{k_a^2}{\Omega_w}, \tag{4.45}$$

where k_a is the comoving dimensionless wave number and the subscript H indicates the Hubble scale. In terms of time, we can rewrite Eq. (4.45) with Eq. (4.37) and Eq. (4.40) such that

$$t_H = -\frac{2\Omega_w}{3\bar{H}_0 k_a^2}. \tag{4.46}$$

Finally, we must ensure that

$$t_f \leq t_H \text{ or}$$
$$\delta_{ini} \geq \frac{5\Omega_w^{\frac{3}{2}} x_{ini}^{\frac{3}{2}}}{k_a^3}, \tag{4.47}$$

where we used Eq. (4.41) and Eq. (B.7). From the above, we can see that the critical threshold depends both on the scale and time and it can be set from the saturation of the inequality as

$$\delta_c = \frac{5\Omega_w^{\frac{3}{2}} x_{ini}^{\frac{3}{2}}}{k_a^3}. \tag{4.48}$$

Let us now evaluate two cases: one considering the dark matter as a pressureless fluid and another considering really small but finite pressure.

4.6 $\bar{p} = 0$

A pressureless fluid representing dark matter is a good starting point due to its simplicity. In this case, Eq. (4.47) tells us that, for every scale k, we can always find a sufficient time in the past such that $x_{ini} \ll 1$, leading to the collapse of all density perturbations. In other words, if the universe is old enough, all perturbations will eventually collapse into a black hole until the universe reaches the bounce. If there was no bounce, the universe would completely collapse into black holes and our model would be highly unstable.

It is indeed expected that a fluid with no pressure leads to a total collapse since no forces are opposing the collapse and dust only interacts via gravity. Thus all dark matter that we see today would indeed be contained in primordial black holes. However, this hypothesis was already refuted by some works by constraining the PBH abundance in dark matter based on observational effects (see Refs. [25, 114]). Hence we must consider a dust fluid with a small but nonvanishing equation of state.

4.7 $\bar{p} = w\bar{\rho}$

In the presence of pressure, it is well known that the Jeans' length determines the smaller scale sufficient for a black hole to be formed. The pressure forces oppose the collapse and only for scales above this limit a black hole can be formed. Thus, we are interested in the physical radius in Super-Jeans/Sub-Hubble scales, that is, $r_j < r < r_H$. The Jeans comoving scale k_j is given by [48]

$$k_J = \sqrt{\frac{3}{2}}\frac{1}{xc}\frac{|\bar{H}|}{c_s}. \tag{4.49}$$

Thus, no structure smaller than the Jeans length can collapse and we have a lower momenta cutoff $k < k_j$. Plugging this requirement plus our upper limit in Eq. (4.49) leads to

$$\begin{gathered} k_H < k \leq k_j \text{ or} \\ \left(\frac{1}{\sqrt{\Omega_w}}\right)^{\frac{3}{2}} > \frac{x^{\frac{2}{3}}}{k^3} \geq \left(\frac{2c_s^2}{3\sqrt{\Omega_w}}\right)^{\frac{3}{2}} \end{gathered} \tag{4.50}$$

where we have used Eq. (4.37) and Eq. (4.40). The above relation indicates that for scales/times that do not satisfy the inequality, the perturbation modes do not contribute to PBH formation. Hence we must evaluate the collapse starting between the Jeans-Hubble time and always ending when the perturbations reach the Hubble length. We can see in Fig. 4 the respective Jeans and Hubble time for each different scale. Note that the interval between both times is always constant for all modes, as it depends only on the dark matter equation of state w. Thus, for smaller w, the gap between the times increases and the PBH formation is enhanced as there is more time for them to be formed.

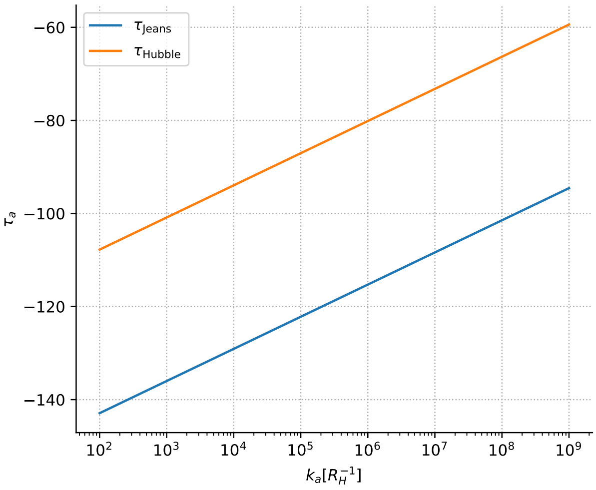


Figura 4 – Plot of the Jeans and Hubble time computed with the NumCosmo library for $w = 10^{-10}$ and $x_b = 10^{35}$. These times correspond to when the matter density perturbations reach either the Jeans or the Hubble length respectively for each mode k. The time parameter in the y-axis is given by Eq. (4.22).

Finally, we have the critical threshold in Eq. (4.48) with Eq. (4.50) corresponding to the allowed scales. Hence, not all perturbations will collapse into primordial black holes. Note that for smaller values of c_s^2, there is more time between the Jeans and the Hubble scale, which allows for more perturbations to collapse. Thus the abundance of PBH in this model is directly related to the equation of state of dark matter, as we will see in the next section.

In this context, let us now analyze Eq. (4.48) for the super-Jeans/Sub-Hubble scales, depicted in Fig. 5. Each plot begins for δ_c being computed for $t_i = t_j$ and goes until $t_i = t_H$. The first case corresponds to the collapse starting when the modes just achieved the Jeans length and thus have the maximum amount of time to collapse until reaching the Hubble length. The end of the time interval corresponds to a collapse taking place right before the perturbation reaches the Hubble length, which leads to a maximum value of the critical threshold since there is a minimum amount of time for the collapse of the perturbations. We can see that all plots go from $\delta_c \approx 10^{-14}$ at Jean crossing time to $\delta_c \approx 5.0$ at the Hubble scale. However, the closer we get to the Hubble time as the initial time, the worse the approximation $\delta_c \ll 1$ gets and thus this should be disregarded.

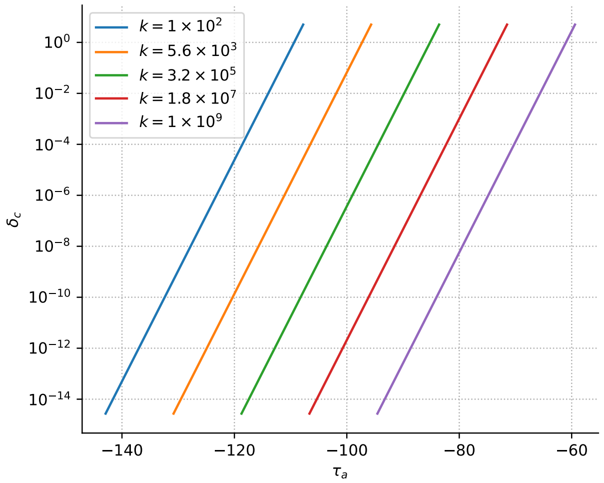


Figura 5 – Plot of the critical threshold vs time computed with the NumCosmo library for $w = 10^{-10}$ and $x_b = 10^{35}$. Each plot refers to a different scale k, while points on the same plot represents a different initial time for the collapse. The time interval for each mode starts at the Jeans scale, leading to the smallest threshold values, and ends for initial times when the perturbations reach the Hubble scale, which leads to higher threshold values. The time parameter in the x-axis is given by Eq. (4.22).

If $t_i = t_H$, we have that $\delta_c \sim \infty$, since there is no time for the collapse to happen. In Fig. 3 we have solved Eq. (4.48) analytically for larger values of δ. Nevertheless, such values of density contrast will never be achieved during the contracting phase.

4.8 Filtered Variance

Let us now compute the variance which will determine the distribution for PBH. Since sub-Jeans and super-Hubble scales do not contribute to the PBH formation that we are interested in, we want to compute the integral in Eq. (4.27) between k_H and k_j. To do so, we first need to evaluate $P_{\tilde{\delta}}(k)$ using Eqs. (4.10)-(B.13) and the field expansions in Eqs. (B.25) and (B.26). Explicitly,

$$\hat{\Psi}(\eta, \mathbf{x}) = \int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^{\frac{3}{2}}} \frac{\bar{H}}{2c\bar{a}k^2} \left[\Pi_{\zeta_k} e^{i\mathbf{kx}} a_k + \Pi^*_{\zeta_k} e^{-i\mathbf{kx}} a^\dagger_k\right], \tag{4.51}$$

$$\bar{D}^2\hat{\Psi}(\eta, \mathbf{x}) = -\int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^{\frac{3}{2}}} \frac{\bar{H}}{2ca^3} \left[\Pi_{\zeta_k} e^{i\mathbf{kx}} a_k + \Pi^*_{\zeta_k} e^{-i\mathbf{kx}} a^\dagger_k\right] \tag{4.52}$$

and

$$\begin{aligned} \delta\hat{\tilde{\rho}}(\mathbf{x}) = \int_{-\infty}^{\infty} \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^{\frac{3}{2}}} \left\{ \left(\frac{\bar{H}}{c\kappa\bar{a}^3}\right) \Pi_{\zeta_k} e^{i\mathbf{kx}} a_k + \right. \\ \left. \left(\frac{\bar{H}}{c\kappa a^3}\right) \Pi^*_{\zeta_k} e^{-i\mathbf{kx}} a^\dagger_k \right\}. \end{aligned} \tag{4.53}$$

Thus, the two-point function and the variance are given by

$$\left\langle \hat{\tilde{\delta}}(\mathbf{x}) \hat{\tilde{\delta}}(\mathbf{y}) \right\rangle = \frac{1}{2\pi^2} \int_0^\infty \frac{\mathrm{d}k}{k} \left[P_{\tilde{\delta}}(k) \frac{\sin kR}{kR}\right] \tag{4.54}$$

and

$$P_{\tilde{\delta}}(k) = \left|\Pi_{\zeta_k}\right|^2 \left(\frac{c^2}{9\bar{a}^6 \bar{H}^2 (1+w)^2}\right), \tag{4.55}$$

where we used the flat Friedmann equation

$$\bar{H}^2 = \frac{c^2 \kappa \bar{\rho}}{3}. \tag{4.56}$$

We are now able to compute the variance modes numerically using the above relation together with Eq. (A.66) in the fluid gauge and our numerical code.

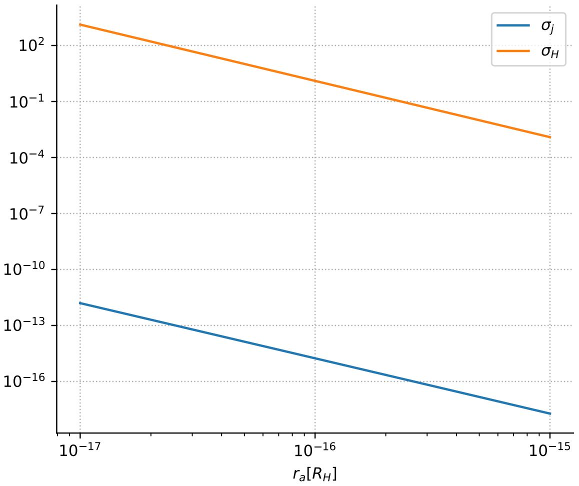


Figura 6 – Plot of the filtered variance computed with the NumCosmo library for $w = 10^{-10}$ and $x_b = 10^{35}$. Each point corresponds to the variance computed when the perturbations become either the Jeans (blue plot) or the Hubble (orange plot) size for each scale.

.

The values for the filtered variance computed with Eq. (4.27) from k_H to k_j are displayed in Fig. 6. We have plotted the values of σ_r for two different times versus radius. The sub-indexes H and j indicate that we are computing the variance at a time when the perturbations are the size of the Jeans length or the Hubble horizon respectively. We can see a linear relation between the filtered variance and the scale radius, such that the former decays with the latter. Larger values of variance indicate a broader Gaussian distribution, which enhances the PBH formation. Thus, from this figure, we see that the formation of PBHs on smaller scales is prioritized rather than on larger scales. We now shall compute the PBH abundance in this model.

4.9 Mass Function

To analyze the abundance of collapsed objects, we use the Press-Schechter (PS) formalism, first proposed in [130]. This ansatz presumes that the objects are formed through a nonlinear collapse and have the property of being universal regarding different cosmological models. Other works have proposed specific mass functions for the primordial black hole production for specific mass ranges, for example [131, 132].

However, in this work, we restrain ourselves to the universal PS formalism to analyze all possible scales. Following this approach, the density of collapsed objects per mass scale is given by the mass function

$$\frac{dn(z,M)}{dM} = -\frac{\bar{\rho}_m(z)}{M}\frac{d\beta(z,M)}{dM}, \tag{4.57}$$

where $\frac{dn(M,z)}{dM}$ is the mass function, z is the object's redshift, M is the object's mass, and β is the fraction of collapsed objects in the mass range of $M + dM$.

Assuming that the objects will collapse when $\delta > \delta_c$, we can measure the mass fraction inside spheres of radius r for a time t that is constituted of collapsed objects as

$$\begin{aligned} \beta(t,r) \equiv \frac{\rho_{PBH}}{\rho} &= \int_{\delta_c}^{\infty} \mathrm{d}\delta \, P(\delta_r) \\ &= \operatorname{erfc}\left(\frac{\delta_c}{\sqrt{2}\sigma_r(t)}\right). \end{aligned} \tag{4.58}$$

However, this function is a probability density of how likely are perturbations to collapse after a time t. For instance, if we compute it at the time when the perturbations reach the Hubble size, $\delta_c = \infty$, this integral will vanish, denoting that there are no more PBHs to be formed after this time. In our context, we are interested in analyzing how many PBHs were formed in the whole contracting phase so we can compute its abundance today. This function shall be given by

$$F(t,r)_{PBH} \equiv \max(\beta(t_i,r)), \text{ for } t_i \leq t. \tag{4.59}$$

The maximum of the beta function indicates for which time there is a higher probability of PBHs to be formed, which will be given by the Jeans time t_j. Thus $F(t,r)$ with $t_i = t_j$ denotes the actual mass fraction of PBHs formed during the entire bouncing phase. Hence, as an approximation, we will compute the mass fraction for the maximum probability of forming PBHs, i.e., always beginning at t_j, to find the abundance of PBHs today.

Another crucial quantity is the abundance of collapsed objects in the universe Ω_{PBH}, defined in Ref. [133] as the ratio between the density of PBHs and the background universe's density today. This parameter will serve as a guide to compare our theoretical predictions with observational data [25]. Since both the background and the PBHs are formed by dust, both grow with the same power of the scale factor and thus the PBH abundance today will be given by

$$\Omega_{PBH} = F(t,r)\Omega_{DM} = \operatorname{erfc}\left(\frac{\delta_c}{\sqrt{2}\sigma_r(t_j)}\right)\Omega_{DM}. \tag{4.60}$$

In Fig. 7 we can see the values of the density function in Eq. (4.59) for different mass scales. Keep in mind that everything is being computed for the collapse starting at the Jeans time for every scale.

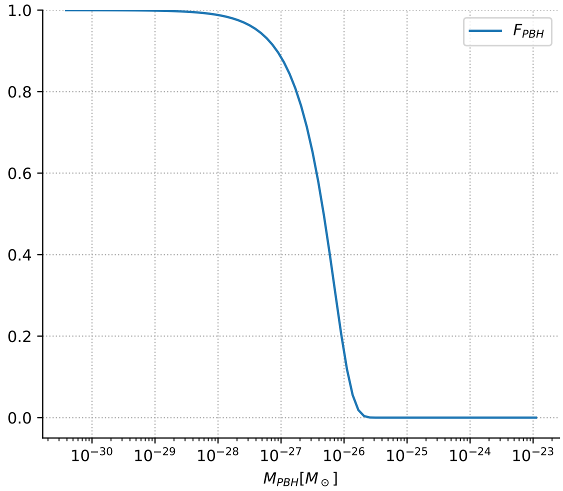


Figura 7 – Plot of the mass fraction of PBHs forming during bounce versus their formation mass in solar masses for $w = 10^{-10}$ and $x_b = 10^{35}$. The computations were done using the NumCosmo library.

The complementary error function in Eq. (4.58) has higher values when the argument is closer to one, i.e., when the filtered variance and the critical density contrast are in the same scale. We can see in Fig. 7 that this is only true for small mass scales, where the density function is equal to one. However, at these mass scales, every formed PBH black hole would eventually evaporate. From Ref. [25], the evaporation constraint determines that only black holes with masses $M > 10^{-18}M_\odot$ would have not completely evaporated today. Thus, for larger scales, the density fraction equals zero, and there is no relevant PBH formation in this model. Nonetheless, since we are dealing only with dust, which has a small and not completely determined equation of state, let us analyze the same function for different values of w, depicted in Fig. 8.

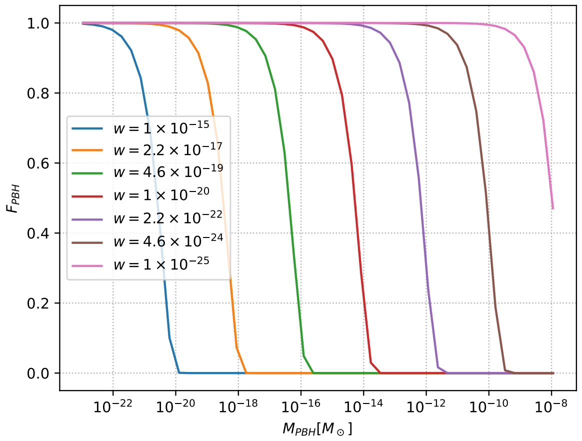


Figura 8 – Plot of the mass fraction of PBHs forming during bounce versus their formation mass in solar masses for different value of w and $x_b = 10^{35}$. The computations were done using the NumCosmo library.

For smaller values for the equation of state of cold dark matter, we can see an enhancement in the density function. This case reflects our previous discussion of a pressureless fluid. As we go further into the smaller values, almost every scale will collapse into a black hole and thus the density function equals one.

4.10 Discussion and Conclusions

In this study, we investigated the formation of PBHs in a flat quantum bouncing model containing only dark matter with a small equation of state parameter w. Despite the spectra from this model exhibiting a slight blue tilt, which deviates from observations of the Cosmic Microwave Background (CMB), this discrepancy should not preclude the model, as the inclusion of radiation is expected to induce a different spectrum that may be compatible with observations [38]. The dark matter-only model serves as an initial attempt to quantify PBH formation during the contracting phase.

Constraints for PBH formation were established, requiring their lengths to fall between the Jeans and Hubble scales. These upper and lower bounds enabled the computation of the critical density contrast, detailed in Sec. 4.5 and App. C. We obtained a time/scale-dependent critical threshold different from the usual constant due to the contracting dynamics. This behavior was expected as there is not only one characteristic scale for the collapse in the bounce model since the perturbations may collapse for any scale between the Jeans and the Hubble scale, which affects the threshold calculation. For the case of a pressureless field, perturbations for all scales may collapse and the bounce/contracting phase duration would be the only constraint for PBH formation.

For a small but non-zero pressure, we computed the mass fraction of primordial black holes in the universe today, given by Eq. (4.59). In Fig. 7, for a fixed value of w, we can see an enhancement of the density function for smaller mass scales of PBHs. However, for $w = 10^{-10}$, the formed primordial black holes would have such small masses at smaller scales that they would have evaporated completely today. Figure 1 was generated in Ref. [25] by imposing observational constraints on the mass fraction of PBHs. They considered a range of effects such as black hole evaporation, microlensing, gravitational wave measurements, and others. The white regions of the graph correspond to accepted values of the mass fraction and the colored represent physical observational effects that exclude the possibility of PBHs constituting DM in that mass range. We can see that PBHs under the mass range $M < 10^{-18} M_{\odot}$ are disregarded due to evaporation, which implies that the formed PBHs depicted in Fig. 7 would have completely evaporated today.

Still on Fig. 7, for larger scales, there is no significant formation as the time interval for the collapse is not sufficient nor is the variance large enough. Also, the variance decreases with scale, as seen in Fig. 6, while the critical threshold remains constant as it is always computed for the Jeans time at the given scale. Hence the complementary error function decreases since the argument grows and the PBH fraction becomes insignificant.

We have seen in Eq. (4.50) that the time interval for which the perturbations become super-Jeans/sub-Hubble is proportional to $c_s^3 = w^{\frac{3}{2}}$. Thus smaller values for the equation of state allow for more time for the perturbations to collapse and therefore enhance the primordial black hole formation. In Fig. 8, we see that only models featuring sufficiently small values for the equation of state of dark matter ($w < 10^{-17}$) may lead to a non-vanishing mass fraction of primordial black hole at relevant scales ($M > 10^{-18} M_{\odot}$). If we go even further on small values of w, more scales start to collapse as we have $F \sim 1$ for all scales as we approach the pressureless fluid scenario.

Another approach for dark matter would be to consider a non-zero temperature. In Ref [134] dark matter is treated as a non-relativistic gas and, even for cases where $w \ll 1$, the non-vanishing temperature would erase structure formation for scales smaller than the corresponding free-streaming length, which once again would contribute to decrease PBH formation in this model. If we treat DM as a relativistic gas, its temperature grows with $\bar{a}^{-2}$ [116]. Thus, its temperature and pressure would diverge close to the bounce and prevent the collapse and PBH formation. We conclude that the formation of PBHs in the flat-dust bounce in observable scales is improbable as the analysis is robust against the formation on larger scales, leading to an insignificant fraction of PBHs as DM today.

The next phase of this research would be to apply the same methodology to a universe populated with radiation and dark matter, which more closely resembles the physical universe and the interaction between both fluids may affect PBH production. If such a model proves reasonable, it would also be interesting to improve the computation of the LTB model in App. C for larger values of the equation of state. Despite the Jeans scale being a good approximation, the introduction of a non-pressureless fluid would also require the study of the pressure impact in our characteristic scales. Furthermore, in a radiation-dominated quantum bouncing model, one expects a greater blue tilt on the spectral index, which may affect the density contrast modes for smaller scales. Said effect could lead to a larger formation of PBHs. This analysis is left for future work.

Parte II

Observational Cosmology

5 CLUSTER NUMBER COUNTS AND MEAN MASS ANALYSIS IN LSST-DESC

Galaxy clusters are the largest gravitational bounded objects and are formed through the collapse of matter over-densities. There are several properties of such objects that can be used to probe cosmology and thus it is a great source of interest within the cosmological community. The clusters are mostly constituted by dark matter (~ 90%), intergalactic gas (~ 9%) and galaxies (~ 1%). Due to their size and matter content, plus the fact that they are seeded by primordial density fluctuations in the matter density field, galaxy clusters are used to study the large-scale structure formation in the universe [52, 53]. Among the various applications, one may measure the gravitational redshift generated by clusters [135, 136] and compare it to models that predict light emitted from the center of the cluster to have larger redshifts. We can also study gravitational lensing effects through light emitted from sources and deflected by galaxy clusters [137, 138], amid other properties. In this work, we shall focus on the cluster abundance in the Universe (or cluster number counts) and their respective mass.

In order to conduct cluster analysis, data is required, which will be provided by the survey called Legacy Survey of Space and Time (LSST) [56], conducted by the Rubin observatory during a ten-year interval. The survey will provide a gigantic dataset on a variety of topics, ranging from solar systems to cosmological scales. Within this framework, the work presented in this chapter was developed as a part of the Dark Energy Science Collaboration (DESC), which is dedicated to the cosmological analysis of the Rubin LSST data. We start the chapter with a brief overview of the LSST-DESC collaboration and then move to the cluster analysis that was developed in the collaboration. We present the theory behind the analysis and discuss the code implementation.

5.1 LSST-DESC

5.1.1 Telescope and Survey

The Legacy Survey of Space and Time will be administered by the Vera Rubin Observatory (named for the American astronomer Vera Florence Cooper Rubin). The observatory is a large, wide-field imaging telescope, currently under construction on Mount Cerro Pachon in Chile. The main facility is the Simonyi Survey Telescope, complemented by the Rubin Auxiliary Telescope (AuxTel) to provide important additional information such as atmospheric transmission to correct the Rubin LSST catalogs. In Fig. 9 we can see the Simonyi telescope in the center and the Auxiliary telescope on the right. The upper part of the observatory contains the telescope mount, where we can see the LSST camera located in the center of the four arms in Fig. 10.

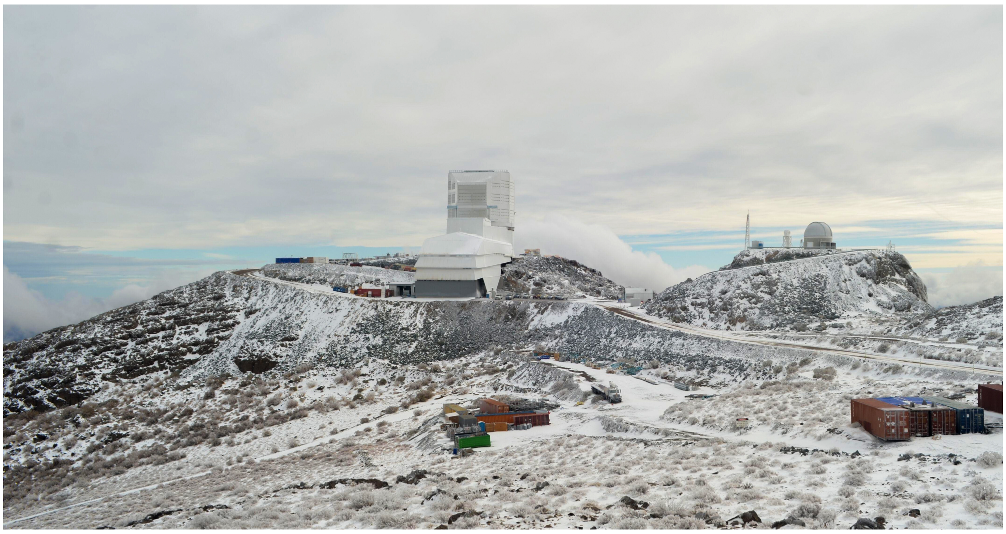


Figura 9 – Image from Ref. [56] of the Vera Rubin observatory under construction.

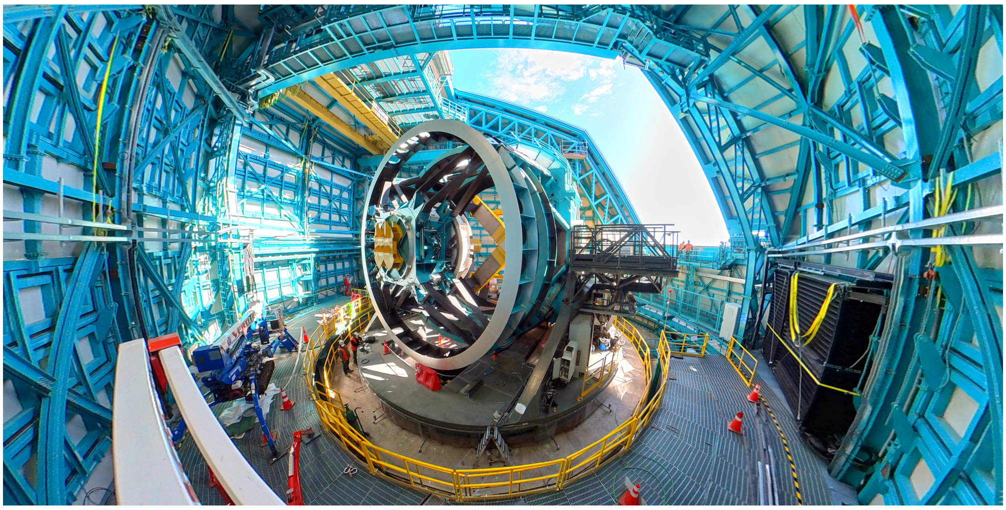


Figura 10 – Image from Ref. [56] of the inside of the Vera Rubin observatory dome.

The LSST survey will provide valuable insights into cosmology, dark matter, and dark energy [139]. By analyzing the positions, shapes, and redshifts of galaxies, the LSST will enhance our understanding of the clustering of matter at different scales and probe weak gravitational lensing effects. The survey will also focus on cluster abundance, with the detection of hundreds of thousands of galaxy clusters up to a redshift of 1.2. The LSST's galaxy catalogs will play a crucial role in calibrating the mass of galaxy clusters through weak lensing effects. Additionally, the LSST will contribute to the study of Type Ia supernovae, expanding the dataset for measuring the expansion rate of the universe with unprecedented precision.

5.1.2 DESC and the Data Challenge

The Dark Energy Science Collaboration was created in June 2012. The collaboration works closely with the LSST survey. Each of the primarily LSST probes (galaxy weak lensing, galaxy clustering, galaxy clusters, and Supernovae Time Domain) is assigned to a dedicated Working Group (WG) in DESC. Also, the modeling and combined probes are developed within DESC to correlate and analyze the data with statistical methods. To all the analysis WGs, the developed codes and sky simulations are crucial to the success of the LSST project, as the simulations are used to forecast the performance of the DESC pipeline[1]. Therefore, we will now discuss the Data Challenge 2 simulation.

The Data Challenge 2 (DC2) is a vast simulated astronomical dataset covering 440 deg2 of the sky, designed to help developing and testing the pipeline and analysis tools of DESC for interpreting the LSST data [140]. The data production has several outcomes and in this section, we want to focus on the cluster catalog generation, used to test the cluster counts code implementation.

The data production is based on an N-body simulation. By modeling the galaxy population related to the simulation, we can generate the cosmoDC2 extra-galactic catalog. Thus the CosmoDC2 is just a galaxy catalog that imitates the measurements expected by the LSST. Next, we run the redMaPPer cluster finder [141] on cosmoDC2, which identifies galaxy clusters through the presence of over-densities of red sequence galaxies. The resulting cluster catalog provides the cluster positions, redshifts, and richnesses

---

[1] All contributions to the development of the DESC pipeline are pushed in a GitHub repository at <https://github.com/LSSTDESC>. For more information about the collaboration, check <https://lsstdesc.org/>.

(calculated as the sum of membership probabilities of galaxies around the cluster) along with the list of potential member galaxies.

The galaxy cluster catalog can also be compared with the dark matter simulation, which we assume to be the "true"data, so we can evaluate the effectiveness of the cluster finder algorithm. In the N-body simulation, we identify dark matter halos by a cluster finder algorithm, for example the friend-of-friend (FoF) halo finder, which defines groups containing all particles separated by a distance less than a given separation and linking length. Next, each halo is assigned a mass M_{FoF} (the sum of the individual dark matter particles associated between them) and a spherical overdensity mass M_{200c}, which represents the required overdensity necessary for a massive gravitational bounded object collapse into a cluster. Assuming that each dark matter halo is related to a galaxy cluster, we can analyze if there is fake detection by the galaxy cluster finder or if the resulting catalog is incomplete. We will now describe the cluster abundance analysis.

5.2 Cluster Abundance

The abundance of galaxy clusters is strongly related to the universe's matter density Ω_m and the rms mass fluctuation σ_8 [54, 55], which is a measure of the amplitude of the linear power spectrum in a scale of $8h^{-1}Mpc$. In Ref. [142] it was found that the total number of clusters in the universe is proportional to a power law depending on σ_8 and Ω_m for clusters detected in Sunyaev Zeldovich surveys, i.e., $N_{tot} \propto \sigma_8^{9.8}\Omega_m^{2.9}$. Cluster abundance has already been used in several analyses as a probe to cosmology (see Refs. [143, 144, 145]). In the following sections, we will briefly review the concepts of cluster formation and how to compute its abundance. Considering that we want to confront observational data with the theoretical predictions in Sec.5.3.1, we will continuously make some comments about the observational implications of the theoretical prediction.

5.2.1 Cluster Formation

Cluster formation is believed to be a rare event generated by high peaks in the matter density [15], characterized by the density contrast $\delta(\mathbf{x})$. We consider the same mechanisms of critical collapse and the Press-Schechter formalism already discussed in Sec. 4.4. In this section we will only discuss additional information regarding cluster formation.

We will use Eq. (4.57) to predict the abundance of clusters in desired ranges rather than in the entire universe as done for primordial black holes. To make it easier to see the connection between the mass function and cosmology, we will split this relation into a collapse-dependent part and a cosmology-dependent part. In this context, we rewrite Eq. (4.57) as

$$\frac{dn(M,z)}{dM} = \frac{\bar{\rho}_m(z)}{M} f(\sigma_R) \left(\frac{1}{\sigma_R(z)} \frac{d\sigma_R(z)}{dM} \right) \tag{5.1}$$

such that

$$f(\sigma_R) = \frac{1}{\sqrt{2\pi}} \frac{\delta_c(z)}{\sigma_R(z)} \exp\left(-\frac{\delta_c^2(z)}{2\delta_R^2(z)} \right). \tag{5.2}$$

The function $f(\sigma_R)$ is called the multiplicity function. An important property is that for an Einstein-de Sitter universe, it is shown in Ref. [146] that

$$\frac{\delta_c(z)}{\sigma_R(z)} = \frac{\delta_c^0 D(z)}{\sigma_R^0 D(z)} \tag{5.3}$$

where $D(z)$ is called the growth function. Thus this fraction will weakly depend on the cosmology and we can decouple the multiplicity function from the cosmological information in $\bar{\rho}_m$ and σ_R in the mass function. Finally, $f(\sigma_R)$ will mostly only depend on the type of collapse[2,3].

The clusters are formed through a nonlinear collapse and can be quantified with Eq. (5.1) using the Press-Schechter formalism. However, the question of the right mass function remains. With the introduction of N-body simulations, different authors tried to refine the PS formalism for clusters by adjusting the multiplicity function. The simulations can be used to reproduce the formation of dark matter halos, which we associate each to a galaxy cluster[4], making it possible to check whether a mass function can properly describe the formation of these objects. Following this approach, different multiplicity

---

[2] Although the cosmological dependency is low, it is not completely negligible, which seeded the efforts to find more accurate multiplicity functions.

[3] (5.3) is the original PS proposal for the spherical collapse

[4] Note that since galaxy clusters are mainly composed of dark matter, the effective mass of galaxies inside the cluster is negligible and we can approximate its mass to the halo mass.

functions were proposed such as the Jenkins mass function [147] based on an ellipsoidal collapse, with

$$f(\sigma_R) = a \exp\left[-|b - \ln \sigma_R|^c\right], \tag{5.4}$$

where $a \approx 0.3$, $b \approx 0.6$ and $c \approx 0.38$, and the Tinker mass function [148] based on spherical overdensity, with

$$f(\sigma_R) = A\left[\left(\frac{\sigma_R}{b}\right)^{-a} + 1\right] \exp\left[-c/\sigma_R^2\right], \tag{5.5}$$

where the parameters A, a, b and c are fitted in the simulation depending on the spherical overdensity factor. While these proposals performed well for the simulations and can be a good approximation for real data, we still do not have a universal mass function that has been proved to properly describe cluster formation in the real universe. With the tools for cluster formation developed in this section, we can now analyze its abundance in the universe.

5.2.2 Binned Cluster Abundance

The mass function is not an observable itself, it is a density per mass range and thus we need to adjust our theoretical predictions that will be used to perform cluster analysis with real data. We can use Eq. (5.1) to count the number of clusters inside a comoving volume V for a given range of mass, which is independent of the cosmology and therefore a good observational candidate. This section is devoted to studying how to compute the theoretical prediction for this new quantity.

We can integrate the mass function from a minimum to an infinite mass to obtain the number of clusters in this range for a desired redshift[5]. This calculation is given by

$$\frac{dN(M > M_{min}, z)}{dz} = \frac{dV}{dz d\Omega} \int_{M_{min}}^{\infty} \mathrm{d}M \, \frac{dn}{dM} \tag{5.6}$$

where $\frac{dV}{dzd\Omega}$ is the comoving volume element and $d\Omega$ is the solid angle. Considering a flat universe, the volume element is given by

$$\frac{dV}{dz d\Omega} = \frac{c}{H} r^2(z) d\Omega \;\text{ with }\; r(z) = \int_0^z \mathrm{d}z' \frac{c}{H}. \tag{5.7}$$

---

[5] Usually a minimal mass M_{min} is chosen as the cluster mass is bounded from below

Finally, if we want the theoretical prediction of the total number of galaxy clusters in a redshift range (z_{min}, z_{max}) with masses bigger than M_{min}, this is provided by

$$N^{tot}_{pred} = \int_{z_{min}}^{z_{max}} \frac{dV}{dzd\Omega} \int_{M_{min}}^{\infty} \mathrm{d}M \, \frac{dn}{dM}. \tag{5.8}$$

There are two ways that we can use Eq. (5.8) depending on the observations. The first is to choose a range of redshift and mass to count the number of clusters inside it (the binned approach), which involves performing the above integrals for different ranges of mass and redshift inside each bin. The second is to use the total number of clusters and multiply by the probability of each cluster being at a redshift z_i with mass M_i. The second option requires data from each cluster and while it can be useful to propagate individual errors, the observational requirements are more complex. Thus, we will focus on the binned approach which only requires that we observe the number of clusters inside bins of mass and redshift.

In the binned analysis, we will separate the data interval into smaller bins. Given a cluster catalog where the cluster masses and redshift range between (M_{min}, M_{max}) and (z_{min}, z_{max}) respectively, we will count the number of clusters in each interval $M^{bin}_{\beta} = (M_{\beta}, M_{\beta+1})$ and $z^{bin}_{\alpha} = (z_{\alpha}, z_{\alpha+1})$ such that the sum of all intervals recovers the initial ranges. Hence, the number of clusters inside the α-redshift bin and β-mass bin is given by

$$N^{\beta}_{\alpha} = \int_{z_{\alpha}}^{z_{\alpha+1}} \mathrm{d}z \, \frac{dV}{dzd\Omega} \int_{M_{\beta}}^{M_{\beta+1}} \mathrm{d}M \, \frac{dn}{dM}. \tag{5.9}$$

Having the prediction inside the given interval, we want to compare it to the data, which is not that easily obtained.

5.2.3 Cluster Data and Richness Proxy

There is one problem when comparing the above theoretical prediction to real data: we cannot measure cluster masses directly. We intend to compare our predictions to actual data to fit some cosmological parameters, which is not possible in the above scenario. Thus, we need to use a proxy, a direct measurement that can be used to represent the original observable. In this work, we shall focus on the richness property of clusters as an indirect observation of mass.

The richness of a cluster is a common mass proxy. The definition of this attribute may differ between distinct analyses, as discussed in Ref. [149], and in this work, we define richness as the number of galaxies inside a galaxy cluster. This is a quantity that can be directly obtained from optical surveys and thus we adopt it as an observational mass.

We follow Refs. [150, 151], which model the relation between the observed richness λ with the cluster mass M at a redshift z by a Gaussian distribution, that is,

$$P\left(\ln\lambda|M,z\right)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{\chi^2\left(M,z,\ln\lambda\right)}{2\sigma^2\left(M,z\right)}\right\} \tag{5.10}$$

with

$$\begin{aligned}\chi(N,M,z)\equiv\ln\lambda-\Bigg[&A+B\ln\left(\frac{M}{M_{\text{pivot}}}\right)\\&+B_z\ln\left(\frac{1+z}{1+z_{\text{pivot}}}\right)+C_z\left[\ln\left(\frac{1+z}{1+z_{\text{pivot}}}\right)\right]^2\Bigg].\end{aligned} \tag{5.11}$$

and

$$\sigma(M,z)=\sigma_0+q\ln\left(\frac{M}{M_{\text{pivot}}}\right)+q_z\ln\left(\frac{1+z}{1+z_{\text{pivot}}}\right)+p_z\left[\ln\left(\frac{1+z}{1+z_{\text{pivot}}}\right)\right]^2. \tag{5.12}$$

In these relations, M_{pivot} and z_{pivot} are constants usually set as the mean mass and redshift obtained from the data respectively. These parameters provide a scale for the analysis related to the used cluster catalog. The Gaussian depends on the mass-proxy parameters $A, B, B_z, C_z, \sigma_0, q, q_z$, and p_z, which may be fitted for a given cluster dataset. Since the clusters true mass and its richness are related through a Gaussian probability, there is a probabilistic relation between these quantities, such that there is not a bijective correspondence among them. Thus it is troublesome to find the best fit for all the parameters needed for the analysis. Let us analyze the necessity of the 6 parameters.

The presence of A and σ_0 is intuitive as they represent a constant mean value for richness and the standard deviation. The parameters B and q are also expected if we assume the number of galaxies inside a cluster to increase with its mass. Regarding the redshift dependency in the remaining parameters, it was shown in Ref. [151] that the presence of such allows the model to include non-monotonic behaviors which led to better results when comparing to Planck data [152]. In the paper, they attribute the non-monotonic dependency on the redshift to projection effects at different cluster redshifts.

Considering the richness proxy discussed above instead of the true mass, our prediction is now given by

$$N_{\alpha}^{\beta} = \int_{z_{\alpha}}^{z_{\alpha+1}} \mathrm{d}z \, \frac{dV}{dz d\Omega} \int_{\ln \lambda_{\beta}}^{\ln \lambda_{\beta+1}} \mathrm{d} \ln \lambda \int_{0}^{\infty} \mathrm{d}M \, \frac{dn}{dM} P \left(\ln \lambda | M, z\right). \tag{5.13}$$

Notice that in this case, our binning is in the richness-redshift parameter space, as we do not have actual measurements of cluster mass to define the intervals. We have seen in Sec. 5.2.1 that cluster formation in the universe is related to the cosmological parameters via the mean density of the universe, the power spectrum, and the density contrast. Consequently, the cluster abundance in Eq. (5.13) is also sensible to cosmology and thus it is a viable probe to constrain such parameters.

Regarding the newly 6 added richness-mass parameters in Eq. (5.13) via Eq. (5.10), the number of clusters inside a bin will not be sufficient to determine the proxy-mass relation and thus we will need to add more information. In this sense, we describe in the next section how to compute the mean mass of the clusters inside a bin, which will later be added as new information in our analysis. This new quantity has a clear relation to the cluster masses and thus is well suited to fit the richness-mass parameters. We will briefly discuss how this type of data can be obtained.

5.2.4 Mean Mass Prediction

Following the same path developed for Eq. (5.13), we just have to multiply the predicted density function for a cluster with mass M at redshift z by its mass M. Thus, the mean mass of clusters inside a redshift and richness bin is given by

$$\langle M \rangle_{\alpha}^{\beta} = \frac{1}{N_{\alpha}^{\beta}} \int_{z_{\alpha}}^{z_{\alpha+1}} \mathrm{d}z \, \frac{dV}{dz d\Omega} \int_{\ln \lambda_{\beta}}^{\ln \lambda_{\beta+1}} \mathrm{d} \ln \lambda \int_{0}^{\infty} \mathrm{d}M \, \frac{dn}{dM} P \left(\ln \lambda | M, z\right) M. \tag{5.14}$$

However, this takes us back to the discussion of how to obtain cluster masses. Although the mean mass inside a redshift-richness bin is more feasible than having data of singular cluster masses, it is still an impossible direct measure. After computing the above prediction, we need to estimate the cluster mass directly from the data to construct our likelihood. One option is to use gravitational lensing to infer the mean mass of clusters inside a bin (check [153] for the description of this method.). This approach is currently being implemented in the DESC cluster pipeline. Let us now discuss the likelihood of the analysis.

5.2.5 Likelihood

There are two important statistical properties to define the cluster abundance likelihood. First, the number of clusters is distributed by a Poisson distribution inside each bin. However, regarding the second property, there is a correlation between the abundance in different bins, which we will call sample covariance. Thus, we relate both characteristics in a Gauss-Poisson Compound (GPC) likelihood, proposed by Hu in Ref. [154].

In Ref. [155] it was stated that if the shot noise contribution is negligible compared to the sample covariance, the GPC likelihood would reduce to a Gaussian given by

$$L = \frac{1}{\sqrt{2\pi}^{c}\,|\Sigma|} \exp\left(-\frac{1}{2}\left[\mathbf{N}_{obs} - \mathbf{N}_{pred}\right]^{T} \Sigma \left[\mathbf{N}_{obs} - \mathbf{N}_{pred}\right]\right), \tag{5.15}$$

where c is the total number of bins, $\mathbf{N} = \{N^{\alpha}_{\beta}\}$ is the vector containing the counts for all the bins, and $\Sigma = \Sigma_{SC} + \Sigma_{SN}$ is the covariance that takes into account the shot noise and the sample covariance contribution. The accuracy of such approximation was tested in Ref. [156], which concluded that the Gaussian approximation with both contributions in the covariance had the same accuracy as the compounded likelihood. Thus, in this work, we use the likelihood in Eq. (5.15). We assume the same likelihood for the mean mass analysis. We will now discuss how the cluster abundance analysis was implemented in Firecrown, a library developed inside the DESC collaboration.

5.3 Firecrown

Firecrown is a Python package that offers the DESC framework for implementing likelihoods, including dedicated implementations of specific likelihoods[6]. The code is being developed to prepare the DESC pipeline for the upcoming LSST data. It will be utilized by multiple working groups at the end of the analysis, providing the necessary likelihood and statistical tools.

---

[6] The code can be found on the GitHub page https://github.com/LSSTDESC/firecrown

The implementation has an extensibility design that allows users to extend or modify the code with new functionalities without requiring extensive changes to the core codebase. This design approach promotes flexibility and maintainability by allowing developers to add new features or components without disrupting the existing code structure. The code also has automatic features such as the update of the required parameters when new objects are added to the structure, addition of new parameters to be sampled for new likelihood implementations, automatic name presetting, and others.

Firecrown is also connected to the CosmoSIS [157] package, the Numerical Cosmology library (NumCosmo) [117], and the Cobaya [158] package, which allows the user to perform Markov Chain Monte Carlo (MCMC) analysis using several samplers and cosmology-oriented codes. Among the different samplers, the user can use the APES sampler included in NumCosmo, which we will discuss in the next chapter. In this context, our goal was to implement the likelihood described in Eq. (5.15) in Firecrown, among the required cluster functions and information needed to perform an MCMC run. We will describe the code functionalities in the next section.

5.3.1 Cluster Code Implementation

We can split the code into two main parts: the modeling and the likelihood. To perform the cluster number counts analysis, we first needed to implement the modeling for the count prediction given by Eq. (5.13). We developed a new cluster module inside Firecrown to compute the integrals and the necessary quantities, where we used the Core Cosmology Library (CCL) [159] to set the cosmology and the cluster mass functions. Inside the module, we designed the *ClusterAbundance*, *ClusterMass*, and *ClusterRedshift* objects to define the observable-proxy relation and its parameters, which are now all contained in the Firecrown[7] package.

There is extra information that is required for the cluster abundance prediction. As discussed in Sec. 5.1.2, we use cluster catalogs to validate our implementations. Specifically, we used the catalog generated by the RedMaPPer algorithm. Thus, we have to take into account that the cluster finder will not provide a true catalog. It may fail to find all the clusters or it may identify clusters that do not exist. These two features are named completeness and purity respectively (check Ref. [160] for their definition), which

---

[7] https://github.com/LSSTDESC/firecrown

we have to take into account when implementing the cluster prediction otherwise it will not match the data. Considering these errors, the new prediction for a cluster catalog becomes

$$N_\alpha^\beta = \int_{z_\alpha}^{z_{\alpha+1}} \mathrm{d}z \, \frac{dV}{dzd\Omega} \int_{\lambda_\beta}^{\lambda_{\beta+1}} \mathrm{d}\lambda \, \frac{1}{\mathbf{p}(\lambda, z)} \int_0^\infty \mathrm{d}M \, C(M, z) \frac{dn}{dM} P\left(\ln \lambda | M, z\right), \quad (5.16)$$

where $C(M, z)$ is the completeness and $\mathbf{p}(\lambda, z)$ is the purity, known as selection functions. This function may be fitted for each cluster finder. In Firecrown, we have fitted this relation for the RedMaPPer algorithm and added it to the prediction code.

With the modeling, we then proceed to the likelihood implementation. In Firecrown, the *Statiscs* module is responsible to organize the data and call the prediction module, so we have all the ingredients needed to build the likelihood in Eq. (5.15). It is also responsible for instantiating the parameters that will be fitted in the analysis. Therefore, we created the *ClusterNumberCounts* object inside the *Statistics* module to read cluster data and build the cluster abundance likelihood for different types of analysis. The object handles all the combinations from true-mass and true-redshift to any other given proxies that may be implemented in the prediction module by the user.

We have also built a connection inside the repository to read data from the *SACC* data format[8], another DESC tool to pass information inside the pipeline, where we created new data types in the library that are now used in our analysis. With all the statistics computed by the *ClusterNumberCounts* object, we were then able to build the likelihood using the *Likelihood* class in Firecrown. The final code can be used for individual computations or with MCMC samplers. In the next section, we will provide some sanity checks and examples of the algorithm used with CosmoSIS.

5.3.2 Code Testing

To check the effectiveness of the code, we first ran the analysis for controlled data. The goal was to run the Firecrown cluster abundance likelihood in CosmoSIS for a completely known dataset and fit two cosmological parameters (σ_8 and Ω_m) to compare with their fiducial values as well as to fit the proxy parameters. To do so, we generated cluster catalogs using the Numerical Cosmology Library (NumCosmo) [117]. In this

---

[8] https://github.com/LSSTDESC/sacc

manner, we have the true information to be passed for Firecrown and we expect to find the true values of the fitted parameters as well. For the catalog generation, we used the Tinker mean mass function [148] and some preset parameters for the mass-richness relation. The catalogs had information about cluster mass, richness, and redshift. Note that throughout some of our tests, the contours have narrow widths since we are using our simulated catalogs. Real data are much more noisy and we expect to have less restrictive constraints on the parameters.

The first test consisted of analyzing the catalog for the true mass and true redshift of the clusters. The result of the analysis is displayed in Fig. 11. Using the same mass function as the one to generate the catalog, we wanted to analyze if the generated MCMC chain using the Firecrown likelihood would be able to fit the true cosmological parameters. We binned the cluster counts data in redshift and mass bins and used the *ClusterTrueMass* module to predict the number of clusters inside each bin. We can see the fit of the final chain in Fig. 11. In this controlled environment, we were able to find the true cosmological parameters denoted by the dotted line within the 2-σ region.

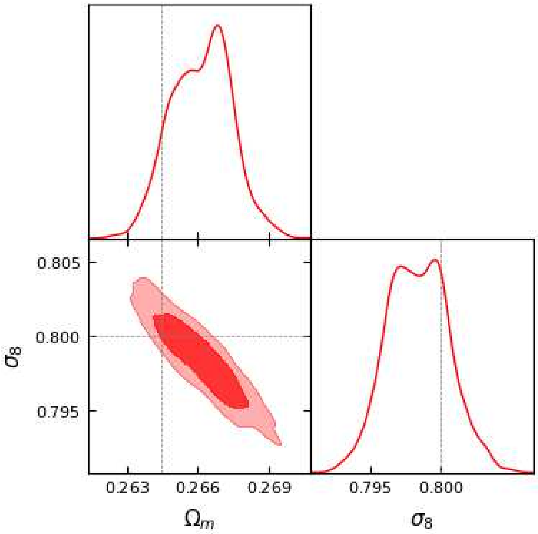


Figura 11 – Contours generated with GetDist from the CosmoSIS samples using the Firecrown likelihood for Cluster Number Counts. In this case, we utilized data of true mass and true redshift of the clusters. The red contours represent the 1 and 2-σ limits. The dotted line represents the original cosmology used to generate the NumCosmo catalog.

We then performed the same test using the richness of information from the catalog. In this test, we used the same mass proxy, including the true parameters and mass function as the ones used to generate the data in the catalog. The results are depicted in Fig. 12. We once again expected to obtain the original cosmology used to generate the catalog. However, since there is a scatter between the richness and the cluster masses, we chose to repeat the test for three datasets, i.e., three simulated catalogs, with different random seeds to properly evaluate the capabilities of the code and simulate different realizations of the universe (check Ref. [156] for a discussion on the subject). Analyzing the contours, we can see that in this ideal scenario, we were able to recover the cosmology up to the $2 - \sigma$ limit.

We proceeded to generalize the previous test by freeing the mass-proxy parameters, which we fitted within the same analysis as described above. The results of this analysis are shown in Fig. 13. Once again we were able to find the original values for both the cosmological and proxy parameters from the catalog within the $2 - \sigma$ region.

Next, we moved to a less ideal test where we do not generate our own catalog and

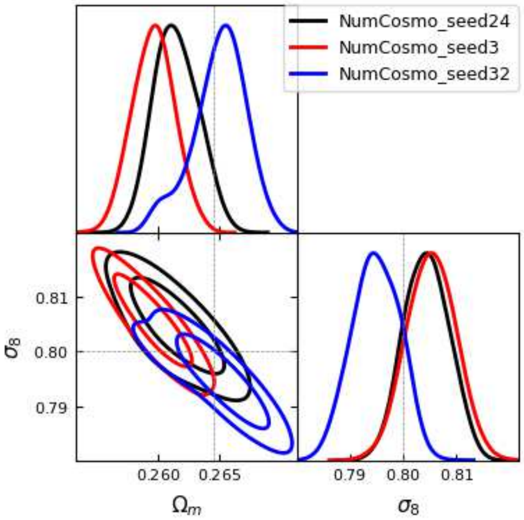


Figura 12 – Contours generated with GetDist from the CosmoSIS samples using the Firecrown likelihood for Cluster Number Counts. In this case, we utilized data of richness and true redshift of the clusters. The richness-mass parameters used were the same as the ones to generate the data. The three sets of data were generated using the NumCosmo library with different random seeds. The dotted line represents the original cosmology used to generate the catalogs.

use one from an N-body simulation where we do not know all true parameters of the analysis. As discussed in Sec. 4.9, the used mass functions implemented in PyCCL were either deduced from first principles or calibrated with N-body simulations. To test their accuracy, we performed a cluster number counts analysis for different mass functions using the SkySim catalog, which is a dark-matter halo catalog[9]. In this case, we perform the analysis in the mass-redshift space since the catalog has the true mass of the dark matter halos obtained from the N-body simulation. The resulting samples can be seen in Fig. 14. We can see that the different mass functions did not properly describe the catalog, which was expected since we do not know the real mass function of the simulation.

We did a last validation of the code using the RedMaPPer cluster catalog, which is closer to a real analysis since we have to fit all the required parameters. To prepare the data, the galaxy cluster catalog has been matched to its equivalent halo catalog from the N-body simulation so we can assign the mean mass for the galaxy clusters in each redshift-richness bin. We also have information about the real proxy parameters.

---

[9] The catalog can be found in the Git repository https://github.com/LSSTDESC/gcr-catalogs.

Therefore we were able to make a joint analysis with the cluster counts and mean mass data. We modeled both the counts prediction and their mean mass with the Boqcuet mass function and ran the analysis using the Firecrown likelihood to fix jointly the proxy parameters and the cosmology. In this case, we used Eq. (5.16) for the counts prediction and its equivalent for the mean mass[10], where both the purity and completeness functions have been fitted for the RedMaPPer cluster finder. The final contours are displayed in Fig. 15 and once again we can see that we found the true values for the catalog.

After the tests, we conclude that the implemented code in Firecrown behaves as expected. Thus, our code implementation in Firecrown to predict cluster abundance was successful and the code may be used by the LSST-DESC collaboration. In the next chapter, we discuss the APES sampler developed by me and Sandro Vitenti.

---

[10] It is worth noticing that for the mean mass, we only account for the completeness since we are using the matched catalog and we assume the catalog to be pure. Bear in mind that this has to be taken into account when computing the counts denominator in Eq. (5.14).

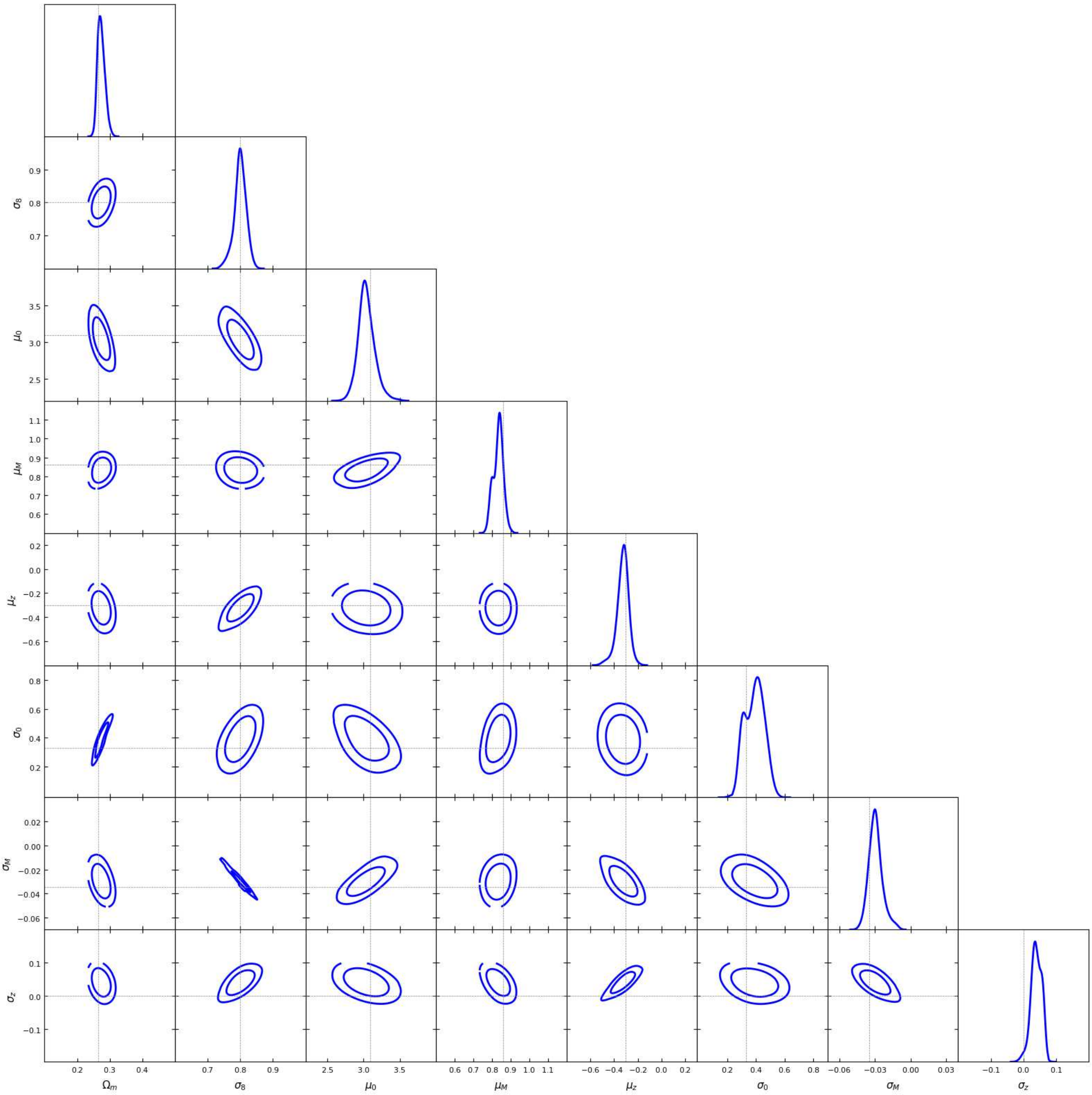


Figura 13 – Contours generated with GetDist from the CosmoSIS samples using the Firecrown likelihood for Cluster Number Counts. In this case, we utilized data of richness and true redshift of the clusters, to simultaneously fit the mass-proxy parameters and the cosmology. The dotted line represents the original values used to generate the NumCosmo catalog.

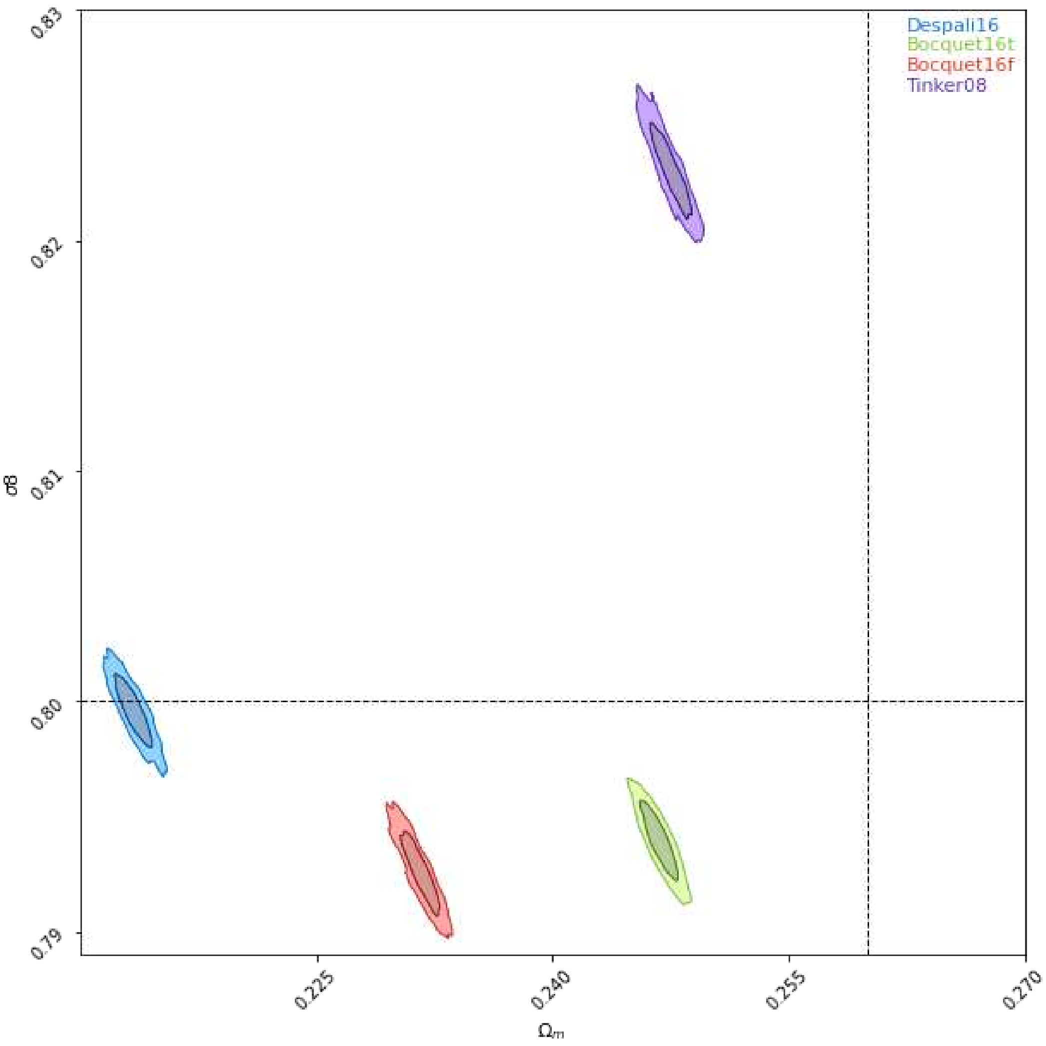


Figura 14 – Contours generated with GetDist from the CosmoSIS samples using the Firecrown likelihood for Cluster Number Counts. In this case, we utilized data of true mas and true redshift of the clusters from the SkySim catalog. The dotted line represents the original cosmology values from the catalog. The names of the mass functions are related to their function calls in the CCL library.

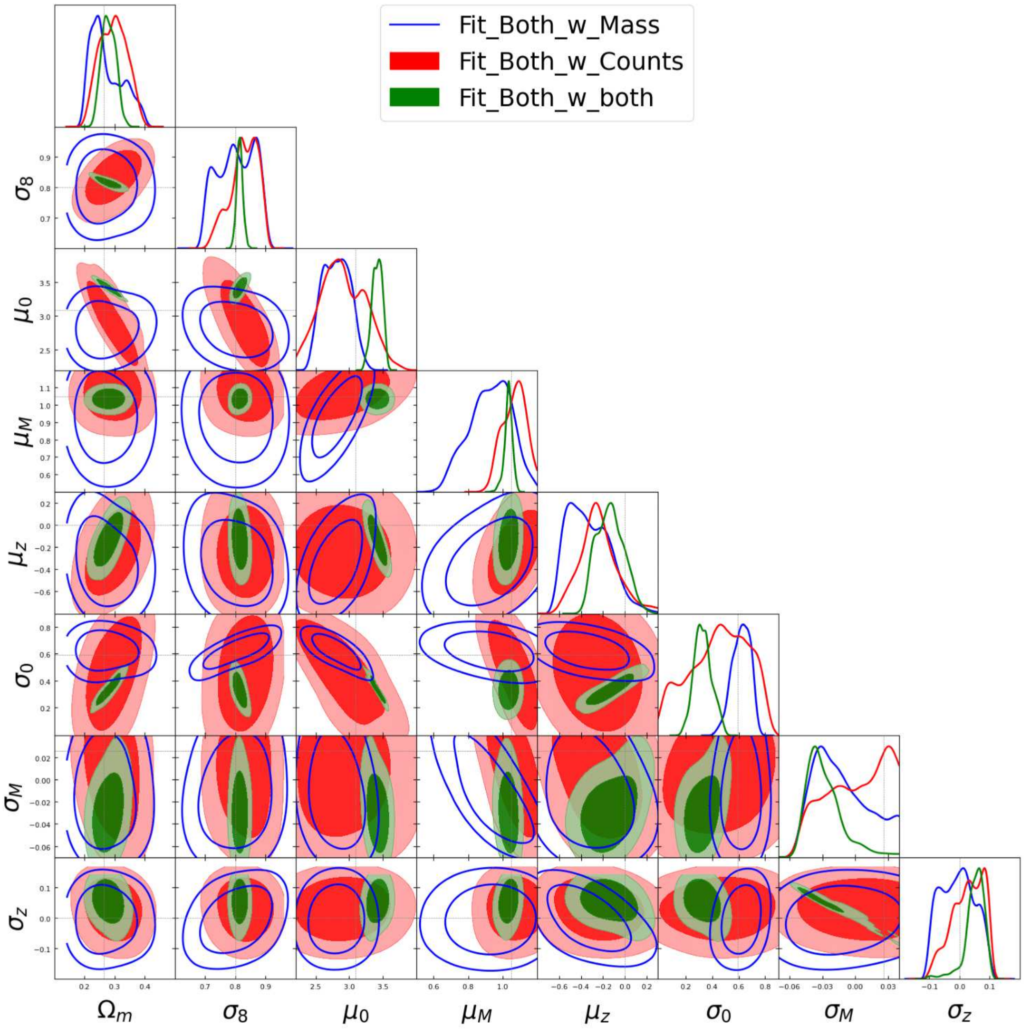


Figura 15 – Contours generated with GetDist from the CosmoSIS samples using the Firecrown likelihood for Cluster Number Counts and Cluster Mean Mass with selection function errors. We utilized data of richness and true redshift for the counts analysis and true mean mass and true redshift of the clusters for the mean mass analysis. Both analyses used the RedMaPPer catalog and its corresponding match with the N-body simulation for the cluster masses. The dotted line represents the original cosmology values from the catalog. We repeated the analysis using just the count’s likelihood, the mean mass likelihood, and a joint likelihood.

6 APES: APPROXIMATE POSTERIOR ENSEMBLE SAMPLER

To fit some likelihood parameters for the cluster number counts analysis, we needed to use Markov Chain Monte Carlo (MCMC). This technique is often used in the area of cosmology and thus knowledge of such a method has become essential among scientists of this area. During the PhD, my advisor Sandro Dias Pinto Vitenti and I have studied ways to develop a more efficient MCMC technique when compared to other popular methods in the community, such as the *Emcee* [161]. This led to the development of the approximate posterior ensemble sampler, denominated APES. We have published a paper related to this work in Ref. [60], which we present in a compact version below.

6.1 Motivation

Markov Chain Monte Carlo methods have become a prominent tool in Bayesian inference since their introduction by Metropolis in 1953 [57]. It consists of a versatile and effective method for generating samples from the posterior distribution without having to sample directly from the distribution, which is usually computationally demanding. In particular, MCMC methods have found widespread application in cosmology and astrophysics for estimating model parameters from observational data, as evidenced by several studies. They were initially deployed in the analysis of the cosmic microwave data using Metropolis-Hastings algorithms [162, 164]. In recent years, MCMC methods have been the subject of intense research in cosmology, with numerous studies exploring their use in various contexts such as the 21-cm signal [165], the cosmic UV background [166], and others.

The development of ensemble samplers, an MCMC algorithm that deals with proposals from multiple points at each time step, led to the development of advanced algorithms that outperform traditional single-particle methods (see [167, 161, 168]). Nevertheless, ensemble samplers face significant challenges when applied to high-dimensional problems. In such scenarios, algorithms designed for low-to-moderate dimensions often resemble random walk samplers, leading to large self-correlation between points and high autocorrelation of the posterior sample. These issues compromise the efficiency of the algorithm and are commonly referred to as the curse of dimensionality [169]. Further discussions about this problem can be found in [170].

There have been several proposals to address the curse of dimensionality in high-dimensional problems. In this paper, we present a new method called the Approximate Posterior Ensemble Sampler (APES). APES is a Metropolis-Hastings ensemble sampler that employs the APES move to generate high-dimensional samples with low autocorrelation and rapid convergence. This proposal employs kernel density estimation and radial basis function interpolation to construct an approximation of the posterior, which is then used to propose points for the Metropolis-Hastings ensemble step. The result is an asymptotically high acceptance ratio and an effective solution to high-dimensional problems. The algorithms presented in this work are implemented in the NumCosmo library [117], which fully integrates the sampling algorithms with cosmological and astrophysical models.[1].

6.2 Ensemble Sampler

Regarding MCMC algorithms, one limitation of the method is the lack of information about the sample when proposing new points. The transitional kernel, used to propose points and update the chain at each time iteration, only depends on the current position and thus cannot use information about the sample. Explicitly, given the target probability distribution $\pi(x)$ for $x \in \mathbb{V} \subset \mathbb{R}^n$ we define a Metropolis-Hastings transition kernel as

$$\begin{aligned} T(y|x) = {} & \alpha(x,y)q(y|x) \\ & + \delta^n(x-y)\int_{\mathbb{V}} \left[1-\alpha(x,z)\right] q(z|x)\mathrm{d}^n z, \end{aligned} \tag{6.1}$$

where $y \in \mathbb{V}$, $q(y|x)$ is an arbitrary conditional probability distribution, which we call proposal distribution on $\mathbb{V}$, $\mathrm{d}^n z$ a measure on $\mathbb{V}$, and the acceptance probability $\alpha(x,y)$ is defined by

$$\alpha(x,y) \equiv \mathrm{MIN}\left[1, \frac{q(x|y)\pi(y)}{q(y|x)\pi(x)}\right]. \tag{6.2}$$

The transitional kernel in Eq. (6.1) satisfies the detailed balance condition $T(y|x)\pi(x) = T(x|y)\pi(y)$. This condition assures that the desired distribution $\pi(x)$ is

---

[1] The project repository can be found here: NumCosmo github

left invariant by the transition kernel, that is,

$$\int_{\mathbb{V}} T(y|x)\pi(x)\mathrm{d}^n x = \pi(y), \tag{6.3}$$

which results in the the accepted points y of the chain being distributed by $\pi(y)$. This is true for any conditional distribution $q(y|x)$. However, conditionals that are easy to sample from are preferable and one usually should opt for distributions that are well adapted to the desired distribution $\pi(x)$ in the sense that the proposed points y should lead to higher values of $\pi(y)$. This choice results in high acceptance ratios in the sample. Also, proposals that lead to a high auto-correlation in the sample (points such that given the current position x and the new point y, we have $\pi(x) \approx \pi(y)$) are not recommended for the algorithm due to its inefficiency.

As seen in Eq. (6.1), the transition kernel has no information about other sample points besides the current position x. In this work, we use the ensemble sampler to circumvent this problem by changing the probability space $\mathbb{V}$ and $\pi(x)$. Instead of dealing with one n-dimensional point in the parametric space, the ensemble sampler considers a set of L independent and identically distributed (iid) realizations of $\pi(x)$

$$\mathbf{x} \equiv (x_1, x_2, \ldots, x_L)\,, \tag{6.4}$$

where $x_i \in \mathbb{V}$ for $i = 1, 2, \ldots, L$. Each component x_i is referred to as a walker and an ensemble point as an $L \times n$ dimensional point $\mathbf{x} \in \mathbb{V}^L$. The target joint distribution for the iid realizations of $\pi(x)$ is

$$\Pi(\mathbf{x}) = \prod_{i=1}^{L} \pi(x_i). \tag{6.5}$$

The ensemble sampler has the same goal to generate samples from a desired distribution without having to sample directly from it and works as other MCMC algorithms. To generate the sample, the ensemble point $\mathbf{x}$ must be iterated for each step (or time), represented with a superscript ^{j}, starting at the $j = 0$, such that each update, or iteration, leads to a new time $j + 1$. The proposal distribution is conditioned on the previous ensemble point, that is, $Q(\mathbf{y}|\mathbf{x})$, which is one of the advantages of this method. If $\mathbf{x}$ provides an approximated sample of $\Pi(x)$ then its information can be used to build a good proposal for the next step. In the case of this work, we choose to update each walker individually such that the point update for each walker is given by the transition kernel

$$\begin{aligned} T\left(y_i\middle|x_i,\mathbf{x}_{[i]}\right) &= \alpha_i(\mathbf{x},y_i)q\left(y_i\middle|x_i,\mathbf{x}_{[i]}\right) \\ &\quad+\delta^n(x_i-y_i)\int_{\mathbb{V}}\left[1-\alpha_i(\mathbf{x},z)\right]q(z|\mathbf{x})\mathrm{d}^n z, \end{aligned} \tag{6.6}$$

and the acceptance probability $\alpha_i(\mathbf{x},y_i)$ for the i-th walker is defined by

$$\alpha_i(\mathbf{x},y_i) \equiv \mathrm{MIN}\left[1,\frac{q\left(x_i\middle|y_i,\mathbf{x}_{[i]}\right)\pi(y_i)}{q(y_i|\mathbf{x})\pi(x_i)}\right]. \tag{6.7}$$

In this context, an ensemble sampler iteration consists of updating all the walkers at a time j. In the above equations, we have defined a partition called the complementary ensemble, which is given as $\mathbf{x}=\left(x_i,\mathbf{x}_{[i]}\right)$ where $x_i\in\mathbb{V}$ is the i-th component of $\mathbf{x}$ to be updated and $\mathbf{x}_{[i]}$ is an element of $\mathbb{V}^{L-1}$ obtained by removing the i-th component of $\mathbf{x}$, called the complementary ensemble. This technique that allows us to update each walker individually is called *partial resampling*, see for example [171, 172].

Any transition kernel $T\left(y_i\middle|x_i,\mathbf{x}_{[i]}\right)$ that leaves invariant the conditional

$$\Pi\left(x_i\middle|\mathbf{x}_{[i]}\right)=\frac{\Pi(\mathbf{x})}{\int_{\mathbb{V}}\Pi\left(\mathbf{x}\right)\mathrm{d}^n x_i}=\pi(x_i), \tag{6.8}$$

also leaves $\Pi(\mathbf{x})$ invariant, which is proven in Ref. [60]. Thus our n-dimensional point proposal can be used such that the ensemble points will be distributed by the joint distribution in Eq. (6.5). This proposal also satisfies the detailed balance and thus assures that the chain is reversible, see for example [173].

The ensemble sampler with *partial resampling* has already been used in the literature. For example [167, 161] used the complementary set to develop an affine invariant proposal, which assures that the sampler is efficient when dealing with non-isotropic distributions [167, 174]. The stretch move, for example, proved itself to be efficient for generic target distributions. However, the presence of a high-dimensional term in the acceptance probability of this algorithm led to its increase/decrease too fast in high-dimensional problems. In high-dimensional problems, the majority of the proposed points end up in spaces with low probability since they consist of a large volume of the parametric space, diminishing the acceptance ratio and increasing the auto-correlation, which can slow down the convergence.

In this work, we introduce a new proposal for the ensemble sampler algorithm. The method uses kernel density estimation and radial basis interpolation to generate proposal points for a Metropolis-Hastings algorithm. These strategies are particularly well-suited for addressing challenges posed by high-dimensional problems and offer a robust algorithm tailored to such settings.

6.3 The APES Proposal

In our proposal, we use the information of the complementary ensemble $\mathbf{x}_{[i]}$ to create an approximate distribution, $\tilde{\pi}(x|\mathbf{x}_{[i]})$, of the target $\pi(x)$. For each step, we need to update all components of $\mathbf{x}$ serially and thus we need to create a split in the ensemble point so the approximation is constant for each iteration and can be used to update all the points in a parallelized manner. We propose splitting the ensemble point into two parts, a technique previously utilized in the `emcee` [161] and other similar algorithms, such that $\mathbf{x} = (\mathbf{x}_{L_1}, \mathbf{x}_{L_2})$, where $\mathbf{x}_{L_1} = (x_1, x_2, \dots, x_{L/2})$ and $\mathbf{x}_{L_2} = (x_{L/2}, x_{L/2+1}, \dots, x_L)$, and we assume L is even for simplicity. We define two index sets as $L_1 = [1, 2, \dots, L/2]$ and $L_2 = [L/2, L/2 + 1, \dots, L]$, and the function

$$s(i) = \begin{cases} L_1 & \text{if} \quad i \in L_2 \\ L_2 & \text{if} \quad i \in L_1 \end{cases} \tag{6.9}$$

Since $\mathbf{x}_{L_2}$ is a subset of $\mathbf{x}_{[i]}$ for any $i \in L_1$, we can build an approximation using $\mathbf{x}_{L_2}$ and use it to update all elements of $\mathbf{x}_{L_1}$ in parallel. We use the same approach to update $\mathbf{x}_{L_2}$ based on the approximation built from $\mathbf{x}_{L_1}$. Let us now discuss the method to construct the approximation using $\mathbf{x}_{L_i}$ for $i = 1, 2$.

Given a sample $\boldsymbol{S}$ of m points in $\mathbb{R}^n$, the approximation is given by

$$\tilde{\pi}(x|\boldsymbol{S}) = \sum_{k=1}^{m} \frac{w_k}{h^n} K_k \left[\frac{D_k(x, x_k)}{h} \right], \qquad \sum_{k=1}^{m} w_k = 1, \tag{6.10}$$

where D_k is the Mahalanobis distance between $x \in \mathbb{V}$ and a point $x_k \in \boldsymbol{S}$ $(k = 1, 2, \dots, m)$ that depends on a covariance matrix $\mathbf{C}_k$, h is the bandwidth parameter, $w_k \geq 0$ the weights and K_k is a normalized multivariate distribution on x.

To generate a sample from $\tilde{\pi}(x|\boldsymbol{S})$ we randomly select a kernel based on the weights and generate one random realization from that kernel. To use the approximation in Eq. (6.10), we need to define the matrices $\mathbf{C}_k$ and the distances $D_k^2(x, x_k)$, kernel function K, weights w_k, and bandwidth h, which we briefly discuss in 6.3.2 (see [60] for a detailed description).

Having built approximation, we set the proposal distribution to be the approximation itself, i.e., $q\left(y_i \middle| x_i, \mathbf{x}_{[i]}\right) = \tilde{\pi}\left(x \middle| \mathbf{x}_{s(i)}\right)$. Consequently, the acceptance probability defined in Eq. (6.7) is then given by:

$$\alpha_i\left(\mathbf{x}_{s(i)}, y_i\right) = \text{MIN}\left[1, \frac{\tilde{\pi}\left(x_i \middle| \mathbf{x}_{s(i)}\right)\pi(y_i)}{\tilde{\pi}\left(y_i \middle| \mathbf{x}_{s(i)}\right)\pi(x_i)}\right]. \tag{6.11}$$

When the approximation $\tilde{\pi}(x|\mathbf{x}s(i))$ effectively represents $\pi(x)$, such that $\tilde{\pi}(x|\mathbf{x}s(i)) = \pi(x)\ [1 + \delta(x)]$ with $|\delta(x)| \ll 1$, the acceptance probability takes the form:

$$\alpha_i\left(\mathbf{x}_{s(i)}, y_i\right) = \text{MIN}\left[1, \frac{1 + \delta(x_i)}{1 + \delta(y_i)}\right] \approx 1 + \text{MIN}\left[0, \delta(x_i) - \delta(y_i)\right]. \tag{6.12}$$

As a consequence, even in cases where errors are relatively large, reaching up to $\delta(x) \sim 20\%$, the acceptance probability can remain higher than that of many other MCMC samplers, approaching unity when errors are small. The subsequent section will offer a comprehensive explanation of the step proposal and the APES algorithm.

6.3.1 The APES Move

The algorithm starts with the determination of an initial point $\mathbf{x}^0 \in \mathbb{V}^L$ either sampled from prior distributions or the best fit. The approximation $\tilde{\pi}\left(x \middle| \mathbf{x}_{L_2}^0\right)$ is then calculated, from which $L/2$ samples are drawn, resulting in $\mathbf{x}_{L_1}^{\star 1}$. The acceptance probability for each sample $x_k^{\star 1}$ is determined using Eq. (6.11), which is given by:

$$\alpha_k\left(\mathbf{x}_{L_2}^0, x_k^{\star 1}\right) = \text{MIN}\left[1, \frac{\tilde{\pi}\left(x_k^0 \middle| \mathbf{x}_{L_2}^0\right)\pi\left(x_k^{\star 1}\right)}{\tilde{\pi}\left(x_k^{\star 1} \middle| \mathbf{x}_{L_2}^0\right)\pi\left(x_k^0\right)}\right], \tag{6.13}$$

forming the tuple ($t = 1$):

$$\boldsymbol{\alpha}_{L_1}^t = \left[\alpha_k\left(\mathbf{x}_{L_2}^{t-1}, x_k^{\star t}\right)\right]_{k \in L_1}. \tag{6.14}$$

If the sample point $x_k^{\star 1}$ is accepted, $x_k^1 = x_k^{\star 1}$, otherwise $x_k^1 = x_k^0$. The algorithm then proceeds to calculate the next set of points, $\mathbf{x}_{L_2}^1$, which is done by computing the approximation $\tilde{\pi}\left(x \middle| \mathbf{x}^1_{L_1}\right)$ and using it to draw $L/2$ samples, referred to as $\mathbf{x}^{\star 1}_{L_2}$. These samples are accepted with a probability

$$\alpha_k\left(\mathbf{x}^1_{L_1}, x_k^{\star 1}\right) = \mathrm{MIN}\left[1, \frac{\tilde{\pi}\left(x_k^0 \middle| \mathbf{x}^1_{L_1}\right)\pi\left(x_k^{\star 1}\right)}{\tilde{\pi}\left(x_k^{\star 1} \middle| \mathbf{x}^1_{L_1}\right)\pi\left(x_k^0\right)}\right], \tag{6.15}$$

which compose the second tuple ($t = 1$):

$$\boldsymbol{\alpha}^t_{L_2} = \left[\alpha_k\left(\mathbf{x}^t_{L_1}, x_k^{\star t}\right)\right]_{k \in L_2}.$$

The new value of x_k^1 is set to $x_k^{\star 1}$ when the probability condition is satisfied; otherwise, it retains its previous value, x_k^0. Following these two steps, the subsequent sample point, denoted as $\mathbf{x}^1$, is determined. This process is iterated until the target time t is achieved. A comprehensive illustration of this process is presented in Algorithm 1.

Algorithm 1 APES Algorithm

```
Separate the $L$ walkers into two blocks $L_1$ and $L_2$
Compute the initial sample with $L$ points $\mathbf{x}^0$ and set $t = 0$
while $t \leq t_{\max}$ do
    Compute $\tilde{\pi}\left(x \middle| \mathbf{x}^t_{L_2}\right)$ and draw $\mathbf{x}^{\star t+1}_{L_1}$
    for $k \in L_1$ do (parallelized loop)
        Compute $\pi\left(x_k^{\star t+1}\right)$ and $\alpha_k\left(\mathbf{x}^t_{L_2}, x_k^{\star t+1}\right)$
    end for
    Update $\mathbf{x}^{t+1}_{L_1}$ using $\boldsymbol{\alpha}^{t+1}_{L_1}$ computed above
    Compute $\tilde{\pi}\left(x \middle| \mathbf{x}^{t+1}_{L_1}\right)$ and draw $\mathbf{x}^{\star t+1}_{L_2}$
    for $k \in L_2$ do (parallelized loop)
        Compute $\pi\left(x_k^{\star t+1}\right)$ and $\alpha_k\left(\mathbf{x}^{t+1}_{L_1}, x_k^{\star t+1}\right)$
    end for
    Update $\mathbf{x}^{t+1}_{L_2}$ using $\boldsymbol{\alpha}^{t+1}_{L_2}$ computed above
    Set $\mathbf{x}^{t+1} = \left(\mathbf{x}^{t+1}_{L_1},\ \mathbf{x}^{t+1}_{L_2}\right)$
    $t = t + 1$
end while
```

6.3.2 APES Options

The APES proposal's success relies on the accuracy of approximations made using two sub-samples, L_i, at each iteration. This accuracy is connected to the choice of distance functions and their associated covariance, the kernel function K_k, the bandwidth parameters, and the weights. In this section, we will briefly discuss the possible options to optimize the algorithm for different problems.

Regarding the choice of distance functions, there are two different options: the "same kernel approach", where a single covariance matrix, $\widehat{\mathbf{C}}\left(\mathbf{x}_{L_i}\right)$, is used for all distance computations, meaning $\mathbf{C}_k = \widehat{\mathbf{C}}\left(\mathbf{x}_{L_i}\right)$; and the variable kernel approach, a different covariance matrix $\mathbf{C}_k$ is determined for each x_k with $k \in L_i$ using a defined number of the nearest points to x_k from $\mathbf{x}_{L_i}$. In the same kernel approach, we recognized the presence of outliers in the sample $\mathbf{x}_L$, which we mitigated by incorporating the Orthogonalized Gnanadesikan-Kettenring (OGK) algorithm [175], a robust estimation of the covariance. In the variable kernel approach, we define the number of nearest neighbors m_k of x_k as a fraction p of $L/2$ and use the fast k-Nearest Neighbor (kNN) search algorithm implementation [176] to find the m_k closest points to x_k which are then used to compute the covariance matrix $\mathbf{C}_k$ of each of these sub-samples $\mathbf{x}_{L_i}[\mathbf{x}_k, p]$ for use in the kernel K_k. The variable kernel approach exhibits adaptability in diverse situations by adjusting to variances at each point and yielding precise approximations. However, it's important to note that a potential limitation lies in its requirement for a greater number of walkers, which scales proportionally with the dimension of $\mathbb{V}$.

To efficiently compute the approximation in Eq. (6.10) and facilitate easy sampling, this study proposes two options: Gaussian (Gauss) and Student's t (ST) kernels. Both of these kernels are widely recognized and have readily accessible efficient implementations. Their mathematical expressions are respectively:

$$K_k^{\mathrm{G}}(d) = \frac{\exp\left[-\frac{d^2}{2}\right]}{\sqrt{(2\pi)^n |\mathbf{C}_k|}}, \tag{6.16}$$

$$K_k^{\mathrm{ST}}(d) = \frac{\Gamma[(\nu+n)/2]}{\Gamma(\nu/2)\sqrt{(\nu\pi)^d\,|\mathbf{C}_k|}}\left[1 + \frac{d^2}{\nu}\right]^{-\frac{\nu+n}{2}}, \tag{6.17}$$

where $|\mathbf{C}_k|$ is the determinant of $\mathbf{C}_k$. The approximation is constructed using a sample

$\mathbf{x}_{L_i}$ of $\pi(x)$. It's essential to note that the majority of points in $\mathbf{x}$ primarily originate from the high-probability region of $\mathbb{V}$, which may result in a neglect of the distribution's tails. The Gaussian kernel is suitable when $\pi(x)$ experiences rapid decay outside its primary region of probability. Conversely, if $\pi(x)$ features prolonged tails or plateaus, the Cauchy kernel is more appropriate.

There are two options to compute the weights in Eq. (6.10): kernel density estimation [177, 178] and radial basis interpolation. The first one involves the use of kernel functions to estimate the target probability distribution, such that the multivariate density estimation is defined by Eq. (6.10) with a fixed weight parameter, $w_k = 2/L$ [179]. This approach is characterized by its simplicity and computational efficiency. However, the accuracy of this method relies on the adequacy of sample points $\mathbf{x}^t_{L_i}$ at time t, as it solely relies on these points for information and does not consider additional knowledge about the distribution $\pi(x)$.

The second approach involves the use of Radial Basis Interpolation (RBI), which uses the approximation (6.10) to approximate the actual distribution function $\pi(x)$ by fitting the weights w_k by solving the following system at each MCMC ensemble iteration:

$$\tilde{\pi}(x_i) = \pi(x_i), \quad \text{for} \quad i = 1, 2, ..., L/2. \tag{6.18}$$

This requires solving the linear system

$$\mathbf{K}\boldsymbol{w} = \boldsymbol{\pi}, \tag{6.19}$$

$$\mathbf{K}_{ij} = \frac{1}{h^n} K_j \left[\frac{D_j(x_i, x_j)}{h} \right], \qquad \boldsymbol{w}_j = w_j, \qquad \boldsymbol{\pi}_i = \pi(x_i), \tag{6.20}$$

where the second line defines the components of the matrix $\mathbf{K}$ and vectors $\boldsymbol{w}$ and $\boldsymbol{\pi}$. The challenge of solving Eq. (6.19) is to find a solution that lets $w_i \geq 0$ since the w_i values represent probabilities in (6.10). This is a Non-Negative Least Squares (NNLS) problem, which we addressed using a C-based algorithm derived from the algorithm described in [180] and the Lapack Library for linear algebra computations.

Lastly, the bandwidth h is computed with the Rule-of-Thumb (RoT) for density estimation interpolation described in [179]. In the case of Gauss and ST kernels, it is given by:

$$h_{\mathrm{RoT}}^{\mathrm{G}} = \left[\frac{4}{L/2(n+2)}\right]^{\frac{1}{n+4}},$$

$$h_{\mathrm{RoT}}^{\mathrm{ST}} = \left[\frac{16(\nu-2)^2(1+n+\nu)(3+n+\nu)}{(2+n)(n+\nu)(2+n+\nu)(n+2\nu)(2+n+2\nu)L/2}\right]^{\frac{1}{n+4}}.$$

These values serve as a preliminary estimate for the bandwidth, such that the final value for both the same and variable kernel approaches is calculated by setting $h = oh_{RoT}$ and $h = oh_{RoT}/p$ respectively, where p is the sample fraction and o is the over-smoothing parameter. The value of o can be freely chosen by the user or fitted by a cross-validation procedure available with the algorithm.

6.4 Computational Tests

This section evaluates the APES algorithm by comparing it with the Stretch-Move algorithm (Stretch) as implemented in `Emcee` [167, 161]. We use autocorrelation time (τ), the measure of the number of iterations to sample a new independent point, and the acceptance ratio of the sample points as metrics. NumCosmo library's diagnostics are applied and we verify the results and diagnostics using GetDist [181]. Different kernel functions, interpolation, and robustness options are available for density estimation, which shall be fine-tuned by the user. The `VKDE` is suitable for non-centralized distributions, `KDE` for centralized ones. Interpolation is useful in high-probability areas but computationally intensive. The "Robust"option addresses outliers. All kernels can be combined with the interpolation option. A quick reference guide with labels for each option is provided in Tables 1 and 2.

Finally, we have introduced a simplified interface for APES that offers compatibility with `Emcee` and `Zeus`, allowing users to utilize APES without the need for the complete NumCosmo statistical framework[2]. Now we analyze the Rosenbrock distribution and a cosmological model to show the algorithm's effectiveness (see [60] for more examples).

---

[2] See notebooks/apes_tests/rosenbrock_simple.ipynb for an example.

Tabela 1 – NumCosmo approximation options. The Robust column relates to the option to use the robust covariance estimation based on the OGK algorithm. The Interpolation column relates to the option to use RBI interpolation to approximate the desired distribution.

Approximation Options	Kernel	Robust	Interpolation
`KDE`	Same	-	-
`VKDE`	Variable	-	-
`Robust-KDE`	Same	X	-
`Robust-VKDE`	Variable	X	-
`Interp-KDE`	Same	-	X
`Interp-VKDE`	Variable	-	X
`Interp-Robust-KDE`	Same	X	X
`Interp-Robust-VKDE`	Variable	X	X

Tabela 2 – NumCosmo kernel types for density estimation.

Kernel Type	Density Function	Properties
Gauss	Eq. (6.16)	Light-tailed
ST3	Eq. (6.17) $\nu = 3$	Heavier tails than Gaussian
Cauchy	Eq. (6.17) $\nu = 1$	Heaviest tails

6.4.1 The 2-dimensional Rosenbrock Distribution

We tested the sampler by generating samples from the 2-dimensional Rosenbrock distribution (see [182] for a discussion on the use of n-dimensional Rosenbrock distributions to test MCMC algorithms). The distribution, denoted as $\pi(x_1, x_2)$, is challenging to sample from due to its thin tails and is expressed as

$$\pi(x_1, x_2) \propto \exp\left[-100\left(x_2 - x_1^2\right)^2 + (1 - x_1)^2\right], \; x_1, \, x_2 \in \mathbb{R}. \tag{6.21}$$

where we did not include the normalization. We first study the interpolation of this distribution using the NumCosmo library generating perfect samples from it and computing the approximation using various kernels and options to find the best approach.

In Fig. 16, we present different approximations of (6.21), all of which were generated using the same sample of 160 draws from (6.21). The VKDE method outperforms the KDE method in describing the distribution due to its local covariance adaptation, which effectively captures position-dependent correlations between x_1 and x_2.

Tabela 3 – Relative difference, acceptance ratio, and over-smooth for each interpolation applied to the same sample of 160 draws from the Rosenbrock distribution.

Interpolation	$\lvert\tilde{\pi}/\pi - 1\rvert < 0.2$	$\overline{\alpha}$	o
`Interp-KDE:Cauchy`	9%	64%	0.05
`Interp-KDE:ST3`	23%	65%	0.07
`Interp-KDE:Gauss`	32%	66%	0.05
`Interp-VKDE:Cauchy`	9%	69%	0.08
`Interp-VKDE:ST3`	33%	72%	0.13
`Interp-VKDE:Gauss`	51%	76%	0.14

To assess the approximation, we generated a test sample $\boldsymbol{T}$ of 100,000 points from Eq. (6.21) and calculated the percentage of points with relative errors less than 20%, as shown in Fig. 3. We also computed the mean acceptance probability. The results indicate that Gaussian kernels offer better fits, especially in regions of higher probability.

Fig. 17 shows the relationship between approximation precision and proposal effectiveness. Heavier-tailed kernels (Cauchy kernels) demonstrated higher acceptance probabilities in the tails, making them more suitable for sampling in these regions. Let us now analyze the MCMC-generated sample with the chosen options.

Next, we executed both algorithms, generating $5,000,000$ points with 15625 iterations and 320 walkers. We discarded the initial 5000 iterations to allow chains to reach a stationary distribution before collecting samples for analysis. Although a smaller burn-in for APES might suffice, we removed the same initial iterations to ensure a fair comparison between the algorithms. Examining the results in Table 4 and Fig. 18, we observe that the APES-generated Rosenbrock sample converged to the true variance value of 4 for the Rosenbrock distribution. In contrast, the Stretch sample shows signs of not converging yet, as indicated in Fig. 18. The autocorrelation time for the APES sample is considerably smaller, with a τ around 10, whereas the Stretch method has an autocorrelation around $\tau = 1400$. This independence advantage, coupled with APES's higher acceptance ratio (47% vs. 23% for Stretch), makes APES significantly more efficient, with 140 times less autocorrelation.

It is essential to note that APES comes with a higher computational cost. It requires around 3 minutes for 5,000,000 points when using 320 walkers, while the Stretch move has virtually no computational cost. Despite this drawback, APES proves more

Figura 16 – We generated plots of approximated distributions using 160 points distributed according to Eq. (6.21). For each approximation, we also plotted the 4σ ellipse in red, which represents the covariance of each kernel. The first line shows plots of the approximated distribution using `Interp-KDE` with all supported kernels. In the second line, we used `VKDE` with $p = 0.05$ of the nearest sample points to compute the kernel covariance for each point. All samples were generated with an over-smooth factor o obtained by applying split cross-validation with 60% of the sample. Finally, the third line shows a plot of the original Rosenbrock function for comparison. We renormalized all plots so that the maximum probability is equal to one, and we displayed the probability decay until a threshold of 10^{-8}. The interpolation with heavier tails tends to overestimate the probability away from the peaks, which is a desirable feature to enable our MCMC algorithm to explore the entire parametric space. Note that the fixed kernel approach is unable to adapt to the local features of the distribution and therefore fails to describe the tails properly. This highlights the advantages of using the VKDE approach to capture the intricate features of the Rosenbrock distribution.

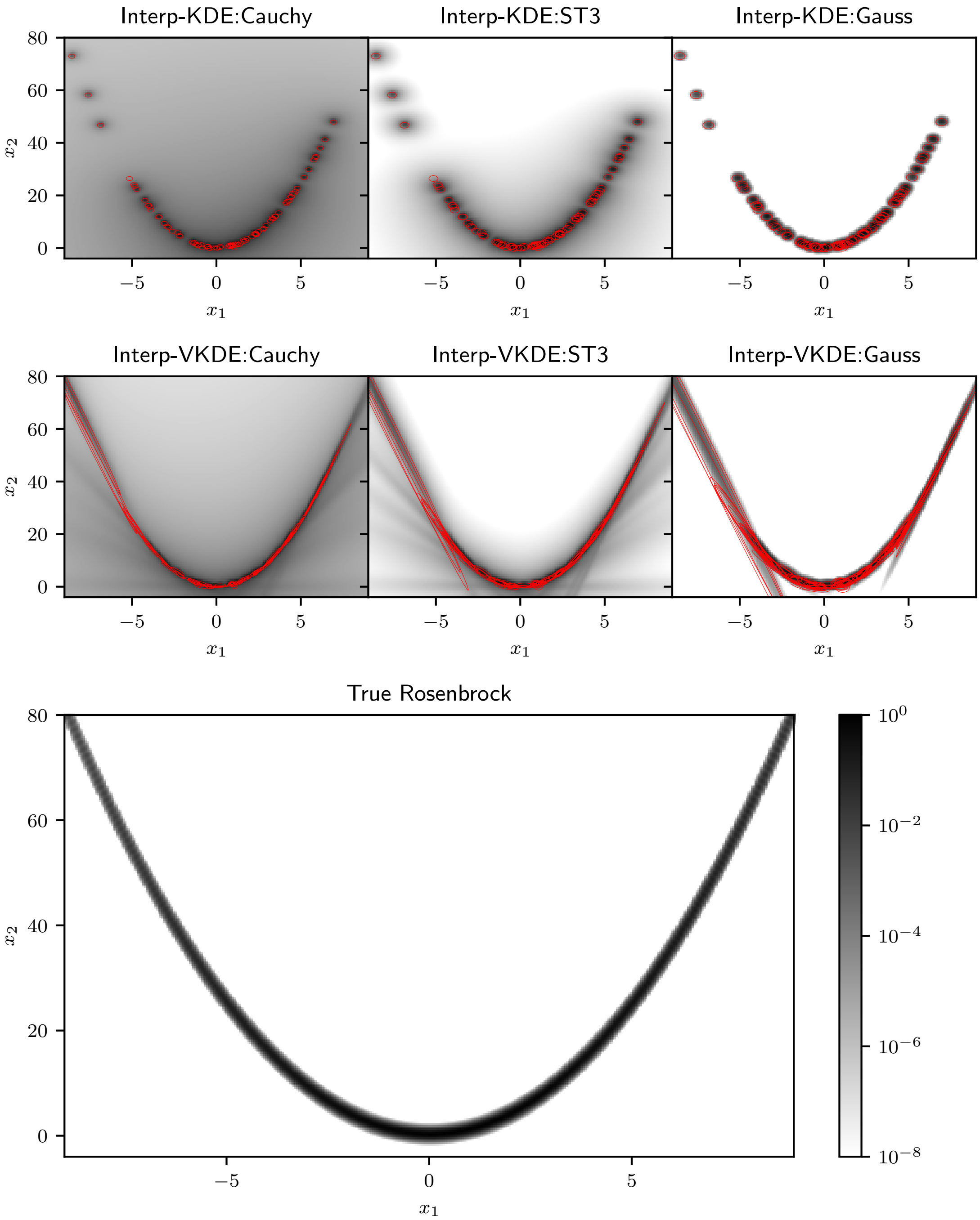
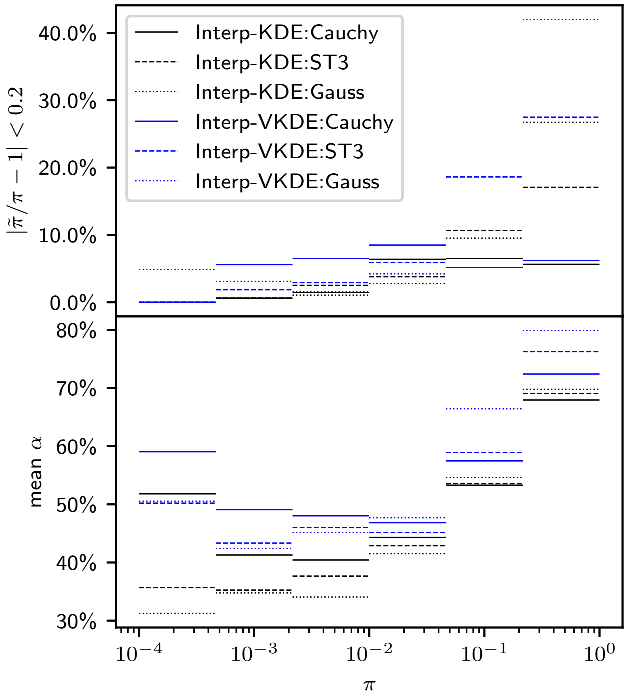


Figura 17 – The upper panel of the graph displays the percentage of test points for which the relative error was smaller than 20% in bins of $\pi(x)$. It is important to note that all approximations experience a loss of precision as they approach the tails, with lighter-tailed kernels performing better near the peak but suffering from decreased precision in the tails. In contrast, heavier-tailed kernels demonstrate greater stability across the entire range, providing a better overall approximation. To explore the relationship between approximation precision and proposal effectiveness, we have included a plot of the mean acceptance probability in the figure below (see equation (6.13)). The acceptance probability is shown as a function of the true distribution on the proposed point. Notably, heavier-tailed kernels have the highest acceptance probability in the tails, indicating that these kernels provide a more effective proposal and increase the likelihood of sampling in the tails.



efficient, producing samples with lower autocorrelation and better tail exploration. Future work could focus on optimizing APES to reduce its computational burden.

You can find the notebook exploring the approximations of the Rosenbrock distribution,[3] and the notebook with the MCMC runs and their analysis at NumCosmo github project.[4]

---

[3] notebooks/stats_dist_tests/stats_dist_rosenbrock.ipynb

[4] notebooks/apes_tests/rosenbrock_mcmc.ipynb

Tabela 4 – Results of an ensemble sampler run using APES and the Stretch move to generate a 2-dimensional Rosenbrock sample. The APES configuration is `Interp-VKDE:Cauchy` with an over-smooth parameter of 0.2 and a local fraction of walkers of 0.05. Both samples were generated using a burn-in of 5000 iterations and a total of 1.6 million points (5000 iterations multiplied by 320 chains). The mean variance and other mean results were computed using all chains. Analytical results were also included for comparison.

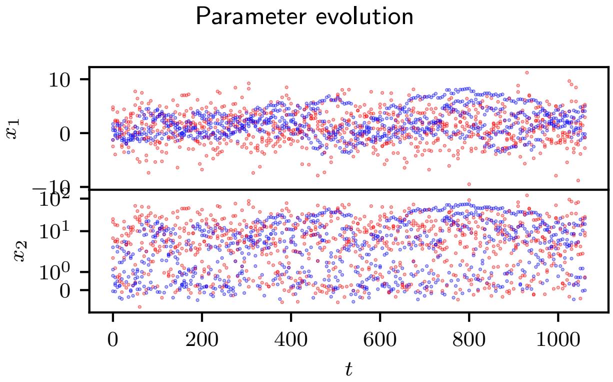
2-dimensional Rosenbrock Run			
	APES	Stretch	Analytical (6.21)
Number of walkers	320	320	–
Number of points	5×10^6	5×10^6	–
τ_{x_1}	6.3	1437.8	–
τ_{x_2}	10.7	1421.6	–
$\bar{x}_1$	1.0	0.9	1
$\bar{x}_2$	11.0	10.2	11
$\overline{-2\ln\pi}$	2.0	2.0	2
$\mathrm{Var}\,(x_1)$	10.0	9.5	10
$\mathrm{Var}\,(x_2)$	236.7	173.3	$2401/10$
$\mathrm{Cor}\,(x_1,x_2)$	0.41	0.35	$20/49 \approx 0.41$
$\mathrm{Var}\,(-2\ln\pi)$	4.0	3.6	4
Acceptance ratio	47%	23%	–

6.4.2 Cosmological Model

This section explores the standard cosmological model's parametric space, incorporating a single massive neutrino. The analysis involves fitting six parameters: namely the Hubble parameter H_0, the cold dark matter density parameter Ω_{c0}, the curvature parameter Ω_{k0}, the dark energy equation of state w, the supernovae type Ia absolute magnitude $\mathcal{M}_1$, and a massive neutrino $m_{\nu 0}$ (in electron-volt eV), as described in detail in [183]. The six-dimensional parametric space is represented as x. We use flat priors for these parameters, except for the cases where $m\nu 0 > 0$ and $\Omega_{k0} > -0.3$, where the priors are non-flat.

The likelihood function used in our analysis comprises the Pantheon+ likelihood [184, 185], the Baryon Acoustic Oscillation (BAO) derived cosmological distances obtained from Data Release 16 (DR16) of the SDSS-III [186, 187], and the cosmological chronometers collated in [188]. This analysis focuses on poorly constrained parametric spaces to study the sampler's behavior.

Figura 18 – Plot showing the iterations (x_1^t, x_2^t) of the MCMC samples generated from the Rosenbrock distribution using APES and Stretch proposals. The plot includes three chains (out of 320 possible) from each algorithm, where the red points correspond to the APES sample, and the blue points correspond to the Stretch sample. The iterations shown in the plot were obtained after discarding the burn-in period. The visual inspection reveals that the APES algorithm explores the tails more efficiently, with more points distributed away from the mean and exhibiting only small autocorrelation. On the other hand, the Stretch algorithm shows a stronger autocorrelation, with the generated points tending to stay closer to their previous positions. The statistical properties for these two runs are summarized in Tabela 4.



In NumCosmo APES, we can set a best-fit before running the MCMC, which accelerates chain convergence. However, in this analysis, we chose not to use this option to study APES's burn-in phases. We initialized all runs with points sampled from Gaussian distributions for the following parameters

$$\begin{aligned} H_0 &= 70 \pm 1, & \Omega_{c0} &= 0.25 \pm 0.01, & \Omega_{k0} &= 0.0 \pm 0.01, \\ w &= -1 \pm 0.01, & m_{\nu 0} &= 10^{-5} \pm 0.01, & \mathcal{M}_1 &= -19.25 \pm 1. \end{aligned}$$

In our analysis, two key points are noted. First, using more walkers extends the burn-in phase but improves the quality of approximations and reduces autocorrelation. This applies to the Stretch move as well. We conducted experiments with five different configurations: APES with 600 walkers, APES with 2000 walkers, and Stretch with 64, 600, and 2000 walkers, running approximately 2×10^6 points in each case. Posterior parameters and covariances were computed after removing the burn-in phase, defined as the iteration maximizing the effective sample size. The results are depicted in Tabela 5. We noted that the burn-in phase increases with the number of walkers, but both APES and Stretch with 600 walkers exhibit larger effective sample sizes, indicating better convergence with smaller autocorrelations.

Despite differing burn-in phases, all configurations eventually converge to the same mean and covariance for the parameters of interest. Stretch with 64 walkers has a notably shorter burn-in phase, suggesting a possible strategy where fewer walkers are initially used to pass through the burn-in phase quickly, followed by restarting the analysis with more walkers. Investigations into this approach and the mixing of APES with other methods are left for future work.

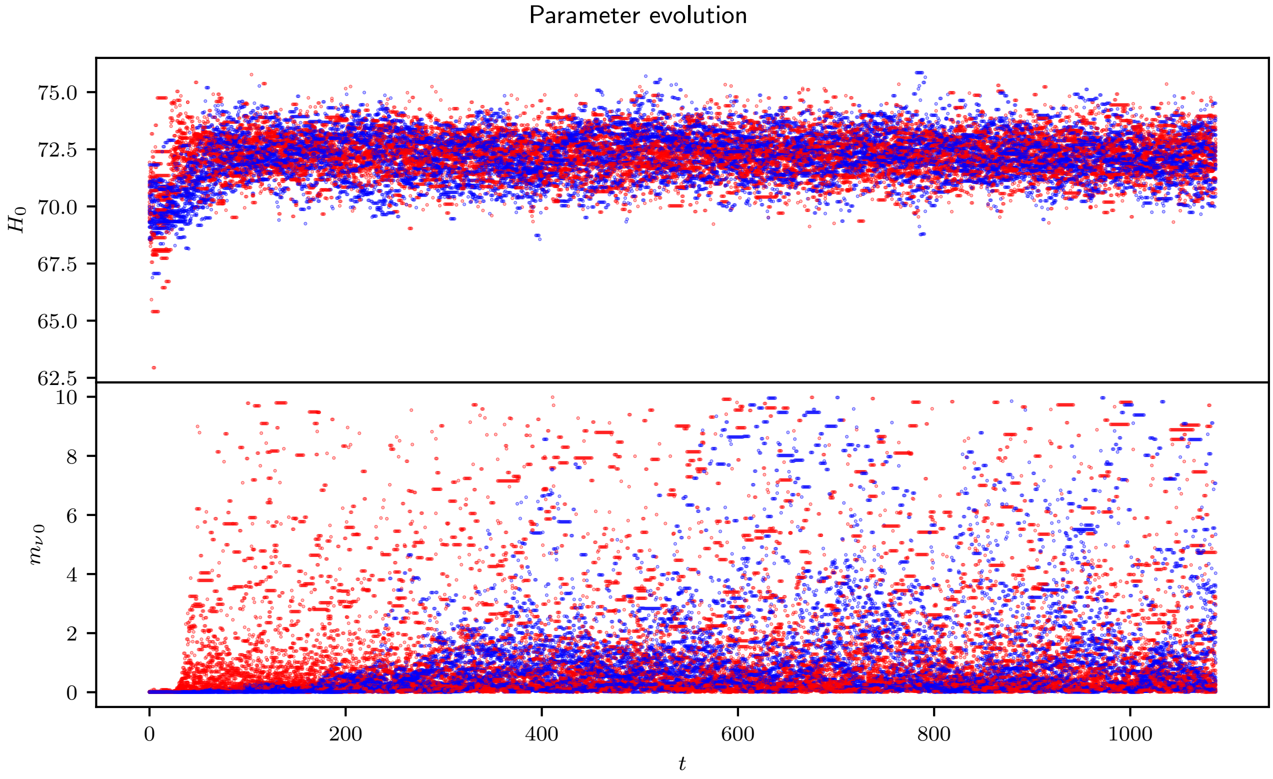
In the most efficient configurations (both using 600 walkers), APES has about 20 times less autocorrelation than Stretch, which is visually apparent inFigura 19 that illustrates the evolution of 20 different chains. Notably, both methods exhibit similar behavior when chains are far from the mean. They often get stuck at certain values for numerous iterations, and in the case of Stretch, points tend to cluster together, further increasing autocorrelation.

High autocorrelation is observed with Stretch with 2000 walkers, where the burn-in phase constitutes a substantial portion of the total iterations. In such cases, caution is required in the analysis process. APES proves advantageous in numerous scenarios due to its typically low autocorrelation. By minimizing autocorrelation, APES enhances the reliability and robustness of analyses, reducing the risk of false convergence and improving the estimation of the target distribution.

Tabela 5 – Results from the cosmological analysis are presented. The APES data were generated using the `Interp-VKDE:Cauchy` option with a local fraction of $p = 0.05$ and an over-smooth factor of 0.2. We generated two samples using 2000 walkers using both APES and Stretch. Additionally, we included a run using Stretch and just 64 walkers. Although the APES chains had converged since the 200th iteration, the larger autocorrelation resulting from the Stretch move necessitated increasing the number of steps to obtain a better approximation of the posterior using this sampler. In this study, we intentionally included the slow-convergence phase of the Stretch move to show its influence and bias on the parameters. Increasing the burn-in to 500 iterations removes the bias and reduces the autocorrelation to about 40 for all parameters. All tests converged to the same means and variances: $\overline{H}_0 = 72.3 \pm 0.9$, $\overline{\Omega}_{c0} = 0.27 \pm 0.04$, $\overline{\Omega}_{k0} = -0.21 \pm 0.05$, $\overline{w} = -0.9 \pm 0.05$, $\overline{m}_{\nu 0} = 1.4 \pm 2.0$, and $\overline{\mathcal{M}}_1 = -19.3 \pm 0.03$.

Pantheon BAO sample					
	APES	APES	Stretch	Stretch	Stretch
Number of walkers	600	2000	64	600	2000
Total number of points	2×10^6	2×10^6	2×10^6	2×10^6	2×10^6
Burn-in: steps	68	130	64	544	670
Burn-in: points	$\approx 4.1 \times 10^4$	$\approx 2.6 \times 10^5$	$\approx 6.0 \times 10^4$	$\approx 3.3 \times 10^5$	$\approx 1.3 \times 10^6$
Effective sample size	$\approx 2.9 \times 10^5$	$\approx 4.7 \times 10^5$	$\approx 9.9 \times 10^3$	$\approx 1.9 \times 10^4$	$\approx 2.6 \times 10^3$
τ_{H_0}	3.9	2.6	103	48	23
$\tau_{\Omega_{c0}}$	6.4	3.7	196	71	21
$\tau_{\Omega_{k0}}$	4.9	3.2	80	83	25
τ_w	5.2	2.5	87	79	24
$\tau_{m_{\nu 0}}$	6.6	3.7	196	86	23
$\tau_{\mathcal{M}_1}$	3.9	2.6	102	57	24
Acceptance ratio	40%	49%	43%	43%	44%

Figura 19 – The figure illustrates the parameter evolution as a function of the iteration t, focusing on the burn-in phase and the initial mixing for MCMC runs using the APES and Stretch methods in a cosmological analysis. We specifically examine the most efficient configuration of 600 walkers for both methods. To maintain clarity in the visualization, we included only 20 chains in the plot. The parameter evolution shown pertains to two specific parameters: H_0, which exhibits smaller autocorrelation, and $m_{\nu 0}$, which exhibits higher autocorrelation. It is worth noting that the burn-in for the Stretch method occurs around $t = 544$, as indicated in Tabela 5. This can be observed in the figure as it takes this duration for the Stretch method to begin exploring the region with $m_{\nu 0} > 0.1$.


Parameter evolution

6.5 Discussion

In summary, the radial basis interpolation technique used in the APES algorithm has proven highly efficient in generating approximate posterior distributions and corresponding sample points. It exhibits smaller autocorrelation times and better acceptance ratios when compared to other MCMC algorithms. Our assessment, which compared APES and the Stretch sampler on the same posterior distributions, highlighted APES's superiority, especially in scenarios involving adaptive sampling like the Rosenbrock distribution.

However, it's crucial to note that in cases with high autocorrelation, chains can falsely appear to have converged prematurely during the analysis. Slow convergence can lead to false positive diagnoses, as evident in our experiments. For the Rosenbrock distribution, the tails aren't well-represented during the slow phase, as chains continue exploring pertinent areas of the parametric space, indicated by the significant autocorrelation. Additionally, the slow convergence phase can bias the mean and report lower variance due to the over-representation of the distribution center.

We also assessed the algorithm's performance for cosmology likelihood, revealing that the APES algorithm led to a significantly smaller autocorrelation time than the Stretch move sampler. For wider samples, the APES algorithm is expected to have a smaller autocorrelation time, a finding supported by our experiments. We explored the impact of the number of walkers on the algorithm's efficiency and noted that a large number of walkers can extend the burn-in phase. However, the number of walkers must be adjusted according to the problem's dimension, as shown in [60] for the truncated multivariate Gaussian distribution test. We're currently investigating alternative methods to minimize extended burn-in phases.

We rigorously tested the robust method in various scenarios and found that while it didn't significantly improve the overall sampling process, it came with substantial computational costs. In practical settings prioritizing computational efficiency, the standard approach remains preferable. However, further research into more computationally efficient robust methods could be valuable for future applications. Robust methods may be beneficial during the burn-in phase or in situations with many outliers where accurate covariance estimation is vital, although the added computational cost should be considered when making this choice.

The versatility of the APES algorithm allows integration with various approximation techniques, making it a powerful tool for diverse applications. Its adaptable framework and efficient target distribution sampling lead to faster convergence rates and less autocorrelation compared to other MCMC algorithms. While future enhancements, such as additional kernels, cross-validation, and fitting techniques, could further improve its capabilities, the current framework is highly effective for solving a wide range of problems.

7 CONCLUSION

In this work, we have developed three different projects: the formation of primordial black holes in a contracting model; the cluster pipeline development in LSST; and the APES: Approximate Posterior Ensemble Sampler. The first two projects are directly connected as they involve the formation of large-scale structures in the universe. Both PBH and galaxy clusters are formed through critical collapse and the study of their abundance in the universe is made similarly. Regarding APES, MCMC algorithms are essential for cosmological analysis and we have seen its use on CosmoDC2, the simulated catalog from DESC. Let us discuss our main results for each project separately.

To compute the abundance of PBH, we needed two quantities: the variance related to density contrast modes and the critical threshold δ_c in a bouncing scenario. The variance was obtained through a numerical code that uses the adiabatic approximation to impose initial conditions for the problem. During the contracting phase of a dust-bouncing model, the perturbations' power spectra grow as we get closer to the bounce. This leads to an increase in the variance related to power spectra and consequently to an enhancement of the probability for PBH production.

Regarding the critical threshold, we developed an analytical solution to compute such a quantity. In the bouncing scenario, this calculation was not fully understood as some approximations used in the inflationary scenario do not hold. We used the LTB metric to describe a local collapse of the perturbations that would form PBHs. We computed the Einstein field equations for such metric and obtained a parametric solution for both the metric's dynamical degree of freedom and the necessary time for the collapse. After imposing some constraints on the initial/final formation time of PBH, we were able to obtain an expression for the critical threshold when the perturbations are small.

With the threshold and variance described above, we computed the probability of PBH to be formed during the contracting phase and finally the abundance of DM today. We concluded that although the production of PBH in certain mass scales may be enhanced during the dust bounce, the formed black holes have such small masses that they would have completely evaporated today. Consequently their abundance in DM today would be irrelevant. If we consider cold dark matter with null pressure, we note that the PBH formation is enhanced. In this case of a pressureless field, all PBH could constitute dark matter, since the only restriction for formation is the bounce time.

The next natural step of this research is to add radiation into the contracting model, a description closer to the physical universe. The addition of a second fluid will lead to a new interaction term into the Hamiltonian and may enhance the perturbations' power spectra. In this case, we also need to analyze the LTB set of solutions as the pressure will not be irrelevant in the presence of radiation.

Regarding the cluster pipeline inside DESC-LSST, we tackled several different problems that needed development in the code. We contributed to the creation of new data files to store and share data inside the collaboration, the SACC format. We have also implemented different likelihoods in Firecrown for galaxy cluster analysis. This package has likelihood implementation for cluster number counts, weak lensing, different proxies such as richness, and others. We have mostly contributed to the cluster number counts analysis, where we implemented theoretical predictions into the code for the case of real masses/redshifts and the case where we use cluster richness as a proxy.

After the code was implemented, we performed several tests with simulated catalogs (CosmoDC2). The code led to the same cosmological parameters as the fiducial ones from the simulations. Our tests indicated that the code is behaving as expected and can be used to perform real data analysis for the upcoming LSST data in 2025. So far, the cluster working group inside DESC has a preliminary pipeline that can do either cluster number counts or stacked lensing profile analysis separately. The next goal of the collaboration is to perform a joint analysis for both probes that involves a covariance between the different data types. This is currently being developed in the collaboration.

Lastly, let us discuss the APES. The curse of dimensionality is a well-known problem among ensemble samplers. For complex multidimensional target distributions, usual MCMC algorithms can get stuck in poorly distributed areas. Using kernel density estimation and radial basis interpolation, the APES algorithm can easily be adapted for different areas of the distribution and avoid such problems. Our algorithm can better describe the posterior by the use of those techniques and we have tested the code for different distributions.

We have compared the efficiency of the APES mostly with the stretch move, adopted by popular MCMC algorithms such as the emcee. We have seen that for complex distributions such as the Rosenbrock distribution, the APES samples not only had a smaller autocorrelation time but also provided a better description of the distribution at its tails, where common algorithms are not as efficient. In the cosmological model example, which had a well-behaved distribution, we once again saw a bigger acceptance ratio of the APES when compared to the stretch move, while maintaining a smaller autocorrelation time. The APES algorithm is adaptive and thus it is well-suited for complex distributions.

REFERÊNCIAS

[1] Zel'dovich, Y. B.; Novikov, I. D. The Hypothesis of Cores Retarded during Expansion and the Hot Cosmological Model. *Soviet Astronomy*, v. 10, p. 602, fev. 1967.

[2] HAWKING, S. Gravitationally collapsed objects of very low mass. *Monthly Notices of the Royal Astronomical Society*, Oxford University Press, v. 152, n. 1, p. 75–78, 1971.

[3] HAWKING, S. W. Black hole explosions? *Nature*, Nature Publishing Group UK London, v. 248, n. 5443, p. 30–31, 1974.

[4] CARR, B. J.; HAWKING, S. W. Black holes in the early universe. *Monthly Notices of the Royal Astronomical Society*, Oxford University Press Oxford, UK, v. 168, n. 2, p. 399–415, 1974.

[5] CARR, B. J. The primordial black hole mass spectrum. v. 201, p. 1–19, out. 1975.

[6] KIM, I. H.; LEE, C. H. Constraints on the spectral index from primordial black holes. *Physical Review D*, APS, v. 54, n. 10, p. 6001, 1996.

[7] RICOTTI, M.; OSTRIKER, J. P.; MACK, K. J. Effect of primordial black holes on the cosmic microwave background and cosmological parameter estimates. *The Astrophysical Journal*, IOP Publishing, v. 680, n. 2, p. 829, 2008.

[8] DOMÈNECH, G.; TAKHISTOV, V.; SASAKI, M. Exploring evaporating primordial black holes with gravitational waves. *Physics Letters B*, Elsevier BV, v. 823, p. 136722, dec 2021. Disponível em: <https://doi.org/10.1016%2Fj.physletb.2021.136722>.

[9] WANG, S.; ZHAO, Z.-C. GW200105 and GW200115 are compatible with a scenario of primordial black hole binary coalescences. *The European Physical Journal C*, Springer Science and Business Media LLC, v. 82, n. 1, jan 2022. Disponível em: <https://doi.org/10.1140%2Fepjc%2Fs10052-021-09981-1>.

[10] Chapline, G. F. Cosmological effects of primordial black holes. *Nature*, v. 253, n. 5489, p. 251–252, jan. 1975.

[11] ABBOTT, B. P. et al. Binary black hole mergers in the first advanced ligo observing run. *Physical Review X*, APS, v. 6, n. 4, p. 041015, 2016.

[12] ABBOTT, B. et al. Binary black hole population properties inferred from the first and second observing runs of advanced ligo and advanced virgo. *The Astrophysical Journal Letters*, IOP Publishing, v. 882, n. 2, p. L24, 2019.

[13] STAROBINSKII, A. A. Spectrum of relict gravitational radiation and the early state of the universe. *ZhETF Pis ma Redaktsiiu*, v. 30, p. 719–723, 1979.

[14] GUTH, A. H. The inflationary universe: A possible solution to the horizon and flatness problems. *Phys. Rev. D*, v. 23, p. 347–356, 1981.

[15] BARDEEN, J. M.; STEINHARDT, P. J.; TURNER, M. S. Spontaneous creation of almost scale-free density perturbations in an inflationary universe. *Physical Review D*, APS, v. 28, n. 4, p. 679, 1983.

[16] LINDE, A. D. A new inflationary universe scenario: a possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems. *Physics Letters B*, Elsevier, v. 108, n. 6, p. 389–393, 1982.

[17] BULLOCK, J. S.; PRIMACK, J. R. Non-gaussian fluctuations and primordial black holes from inflation. *Physical Review D*, APS, v. 55, n. 12, p. 7423, 1997.

[18] YOKOYAMA, J. Chaotic new inflation and formation of primordial black holes. *Physical Review D*, APS, v. 58, n. 8, p. 083510, 1998.

[19] JOSAN, A. S.; GREEN, A. M. Constraints from primordial black hole formation at the end of inflation. *Physical Review D*, APS, v. 82, n. 4, p. 047303, 2010.

[20] BALLESTEROS, G.; TAOSO, M. Primordial black hole dark matter from single field inflation. *Physical Review D*, APS, v. 97, n. 2, p. 023501, 2018.

[21] WANG, X.; ZHANG, Y.-l.; SASAKI, M. Enhanced curvature perturbation and primordial black hole formation in two-stage inflation with a break. *arXiv preprint arXiv:2404.02492*, 2024.

[22] CARR, B. et al. Primordial black hole formation during slow reheating after inflation. *Physical Review D*, APS, v. 97, n. 12, p. 123535, 2018.

[23] MARTIN, J.; PAPANIKOLAOU, T.; VENNIN, V. Primordial black holes from the preheating instability in single-field inflation. *Journal of Cosmology and Astroparticle Physics*, IOP Publishing, v. 2020, n. 01, p. 024, 2020.

[24] CARR, B.; KÜHNEL, F. Primordial black holes as dark matter candidates. *SciPost Physics Lecture Notes*, p. 048, 2022.

[25] VILLANUEVA-DOMINGO, P.; MENA, O.; PALOMARES-RUIZ, S. A brief review on primordial black holes as dark matter. *Frontiers in Astronomy and Space Sciences*, Frontiers Media SA, v. 8, p. 681084, 2021.

[26] GARCÍA-BELLIDO, J. Massive primordial black holes as dark matter and their detection with gravitational waves. In: IOP PUBLISHING. *Journal of Physics: Conference Series*. [S.l.], 2017. v. 840, n. 1, p. 012032.

[27] PINTO-NETO, N. Bouncing quantum cosmology. *Universe*, MDPI, v. 7, n. 4, p. 110, 2021.

[28] BRANDENBERGER, R.; PETER, P. Bouncing cosmologies: Progress and problems. *Foundations of Physics*, Springer Science and Business Media LLC, v. 47, n. 6, p. 797–850, feb 2017. Disponível em: <https://doi.org/10.1007%2Fs10701-016-0057-0>.

[29] VITENTI, S. D. P.; FALCIANO, F. T.; PINTO-NETO, N. Covariant bardeen perturbation formalism. 2013.

[30] GASPERINI, M.; VENEZIANO, G. Inflation, deflation, and frame-independence in string cosmology. *Mod. Phys. Lett. A*, v. 8, p. 3701–3713, 1993.

[31] GASPERINI, M.; VENEZIANO, G. Dilaton production in string cosmology. *Phys. Rev. D*, v. 50, p. 2519–2540, 8 1994.

[32] LYTH, D. H. The primordial curvature perturbation in the ekpyrotic universe. *Physics Letters B*, Elsevier, v. 524, n. 1-2, p. 1–4, 2002.

[33] FINELLI, F.; BRANDENBERGER, R. Generation of a scale-invariant spectrum of adiabatic fluctuations in cosmological models with a contracting phase. *Physical Review D*, APS, v. 65, n. 10, p. 103522, 2002.

[34] WANDS, D. Duality invariance of cosmological perturbation spectra. *Phys. Rev. D*, v. 60, n. 2, p. 023507, 7 1999.

[35] BRANDENBERGER, R.; FINELLI, F. On the spectrum of fluctuations in an effective field theory of the ekpyrotic universe. *J. High Energy Phys.*, v. 11, p. 56, 11 2001.

[36] PETER, P.; PINTO-NETO, N. Primordial perturbations in a nonsingular bouncing universe model. *Phys. Rev. D*, v. 66, n. 6, p. 063509, sep 2002. ISSN 0556-2821.

[37] HWANG, J.; NOH, H. Non-singular big-bounces and evolution of linear fluctuations. *Phys. Rev. D*, v. 65, p. 124010, 2002.

[38] VITENTI, S. D. P.; PINTO-NETO, N. Primordial power spectrum from a contracting pre-bounce phase. In: DAMOUR, T.; JANTZEN, R.; RUFFINI, R. (Ed.). *The Twelfth Marcel Grossmann Meeting: On Recent Developments in Theoretical and Experimental General Relativity, Astrophysics and Relativistic Field Theories: Proceedings of the MG12 Meeting on General Relativity*. [S.l.]: World Scientific Publishing Company, 2012. p. 1424–1426.

[39] VITENTI, S.; FALCIANO, F.; PINTO-NETO, N. Quantum cosmological perturbations of generic fluids in quantum universes. *Physical Review D*, APS, v. 87, n. 10, p. 103503, 2013.

[40] BATTEFELD, D.; PETER, P. A critical review of classical bouncing cosmologies. *Physics Reports*, Elsevier BV, v. 571, p. 1–66, apr 2015. Disponível em: <https://doi.org/10.1016%2Fj.physrep.2014.12.004>.

[41] PETER, P.; PINHO, E. J.; PINTO-NETO, N. Noninflationary model with scale invariant cosmological perturbations. *Physical Review D*, APS, v. 75, n. 2, p. 023516, 2007.

[42] AGULLO, I.; OLMEDO, J.; SREENATH, V. Observational consequences of bianchi i spacetimes in loop quantum cosmology. *Physical Review D*, APS, v. 102, n. 4, p. 043523, 2020.

[43] SAKELLARIADOU, M. Phenomenology of loop quantum cosmology. In: IOP PUBLISHING. *Journal of Physics: Conference Series*. [S.l.], 2010. v. 222, n. 1, p. 012027.

[44] CARR, B.; COLEY, A. Persistence of black holes through a cosmological bounce. *International Journal of Modern Physics D*, World Scientific, v. 20, n. 14, p. 2733–2738, 2011.

[45] CORMAN, M.; EAST, W. E.; RIPLEY, J. L. Evolution of black holes through a nonsingular cosmological bounce. *Journal of Cosmology and Astroparticle Physics*, IOP Publishing, v. 2022, n. 09, p. 063, 2022.

[46] CHEN, J.-W. et al. Tracing primordial black holes in nonsingular bouncing cosmology. *Physics Letters B*, Elsevier, v. 769, p. 561–568, 2017.

[47] CHEN, J.-W. et al. Enhance primordial black hole abundance through the non-linear processes around bounce point. *Journal of Cosmology and Astroparticle Physics*, IOP Publishing, v. 2023, n. 01, p. 015, 2023.

[48] QUINTIN, J.; BRANDENBERGER, R. H. Black hole formation in a contracting universe. *Journal of Cosmology and Astroparticle Physics*, IOP Publishing,

v. 2016, n. 11, p. 029–029, nov. 2016. ISSN 1475-7516. Disponível em: <http://dx.doi.org/10.1088/1475-7516/2016/11/029>.

[49] BANERJEE, S.; PAPANIKOLAOU, T.; SARIDAKIS, E. N. Constraining f (r) bouncing cosmologies through primordial black holes. *Physical Review D*, APS, v. 106, n. 12, p. 124012, 2022.

[50] PAPANIKOLAOU, T. et al. Primordial black holes and induced gravitational waves in non-singular matter bouncing cosmology. *arXiv preprint arXiv:2404.03779*, 2024.

[51] VITENTI, S. D. P.; PINTO-NETO, N. Large adiabatic scalar perturbations in a regular bouncing universe. *Physical Review D*, APS, v. 85, n. 2, p. 023524, 2012.

[52] KRAVTSOV, A. V.; BORGANI, S. Formation of galaxy clusters. *Annual Review of Astronomy and Astrophysics*, Annual Reviews, v. 50, n. 1, p. 353–409, sep 2012. Disponível em: <https://doi.org/10.1146%2Fannurev-astro-081811-125502>.

[53] FAN, X.; BAHCALL, N. A.; CEN, R. Determining the amplitude of mass fluctuations in the universe. *The Astrophysical Journal*, IOP Publishing, v. 490, n. 2, p. L123, 1997.

[54] WANG, L.; STEINHARDT, P. J. Cluster abundance constraints for cosmological models with a time-varying, spatially inhomogeneous energy component with negative pressure. *The Astrophysical Journal*, IOP Publishing, v. 508, n. 2, p. 483, 1998.

[55] ABDULLAH, M. H.; KLYPIN, A.; WILSON, G. Cosmological constraints on ωm and σ8 from cluster abundances using the galwcat19 optical-spectroscopic sdss catalog. *The Astrophysical Journal*, IOP Publishing, v. 901, n. 2, p. 90, 2020.

[56] CORPORATION, L. *Legacy Survey of Space and Time (LSST)*. 2023 (accessed June 15, 2023). <https://www.lsst.org/>.

[57] METROPOLIS, N. et al. Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, v. 21, p. 1087–1092, 1953.

[58] TROTTA, R. Bayes in the sky: Bayesian inference and model selection in cosmology. *Contemporary Physics*, Taylor & Francis, v. 49, n. 2, 2008.

[59] BARROSO, E. et al. Primordial black hole formation in a dust bouncing model. *arXiv preprint arXiv:2405.00207*, 2024.

[60] VITENTI, S. D. P.; BARROSO, E. J. Apes: approximate posterior ensemble sampler. *Monthly Notices of the Royal Astronomical Society*, Oxford University Press, v. 525, n. 1, p. 73–90, 2023.

[61] FRIEDMAN, A. Über die krümmung des raumes. *Zeitschrift für Physik*, v. 10, p. 377–386, 1922.

[62] LEMAÎTRE, G. Un univers homogène de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extra-galactiques. *Annales de la Société Scientifique de Bruxelles, A47, p. 49-59*, v. 47, p. 49–59, 1927.

[63] ROBERTSON, H. P. Kinematics and world-structure. *The Astrophysical Journal*, v. 82, p. 284, 1935.

[64] WALKER, A. G. On milne's theory of world-structure. *Proceedings of the London Mathematical Society*, Wiley Online Library, v. 2, n. 1, p. 90–127, 1937.

[65] PEEBLES, P. J. E.; RATRA, B. The cosmological constant and dark energy. *Reviews of modern physics*, APS, v. 75, n. 2, p. 559, 2003.

[66] KNOBEL, C. An introduction into the theory of cosmological structure formation. *arXiv preprint arXiv:1208.5931*, 2012.

[67] MAROTO, A. L.; RAMIREZ, J. A conceptual tour about the standard cosmological model. *arXiv preprint astro-ph/0409280*, 2004.

[68] KINNEY, W. H. Cosmology, inflation, and the physics of nothing. *NATO Sci. Ser. II*, v. 123, p. 189, 2003.

[69] WALD, R. M. *General relativity*. [S.l.]: University of Chicago press, 2010.

[70] CARROLL, S. M. *Spacetime and geometry*. [S.l.]: Cambridge University Press, 2019.

[71] ISHAM, C. J. *Modern differential geometry for physicists*. [S.l.]: World Scientific Publishing Company, 1999. v. 61.

[72] PERLMUTTER, S. et al. Measurements of ω and λ from 42 high-redshift supernovae. *The Astrophysical Journal*, IOP Publishing, v. 517, n. 2, p. 565, 1999.

[73] SPERGEL, D. N. et al. First year wilkinson microwave anisotropy probe (wmap) observations: Determination of cosmological parameters. *The Astrophysical Journal Supplement Series*, American Astronomical Society, v. 148, n. 1, p. 175–194, sep 2003. Disponível em: <https://doi.org/10.1086%2F377226>.

[74] AL., N. A. et. Planck 2018 results. vi. cosmological parameters. *Astronomy & Astrophysics*, EDP Sciences, v. 641, p. A6, sep 2020. Disponível em: <https://doi.org/10.1051%2F0004-6361%2F201833910>.

[75] YORK, D. G. et al. The sloan digital sky survey: Technical summary. *The Astronomical Journal*, IOP Publishing, v. 120, n. 3, p. 1579, 2000.

[76] NORBERG, P. et al. The 2df galaxy redshift survey: the dependence of galaxy clustering on luminosity and spectral type. *Monthly Notices of the Royal Astronomical Society*, Blackwell Science Ltd, v. 332, n. 4, p. 827–838, 2002.

[77] WEINBERG, S. Gravitation and cosmology: principles and applications of the general theory of relativity. 1972.

[78] LINDE, A. D. A new inflationary universe scenario: a possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems. *Physics Letters B*, Elsevier, v. 108, n. 6, p. 389–393, 1982.

[79] RATRA, B.; VOGELEY, M. S. The beginning and evolution of the universe. *Publications of the Astronomical Society of the Pacific*, IOP Publishing, v. 120, n. 865, p. 235–265, mar 2008. Disponível em: <https://doi.org/10.1086%2F529495>.

[80] BRANDENBERGER, R.; MUKHANOV, V.; SORNBORGER, A. Cosmological theory without singularities. *Physical Review D*, APS, v. 48, n. 4, p. 1629, 1993.

[81] NOVELLO, M.; BERGLIAFFA, S. P. Bouncing cosmologies. *Physics reports*, Elsevier, v. 463, n. 4, p. 127–213, 2008.

[82] PETER, P. Observing alternatives to inflation. arXiv, 2009. Disponível em: <https://arxiv.org/abs/0912.1293>.

[83] MATHEW, J. *Bounce inflation driven by Higgs field*. arXiv, 2018. Disponível em: <https://arxiv.org/abs/1811.06001>.

[84] SENATORE, L. Lectures on inflation. In: *New Frontiers in Fields and Strings*. WORLD SCIENTIFIC, 2016. Disponível em: <https://doi.org/10.1142%2F9789813149441_0008>.

[85] GONZALEZ, J. A. V.; PADILLA, L. E.; MATOS, T. Inflationary cosmology: from theory to observations. *Revista Mexicana de Fí sica E*, Sociedad Mexicana de Fisica A C, v. 17, n. 1 Jan-Jun, p. 73–91, jan 2020. Disponível em: <https://doi.org/10.31349%2Frevmexfise.17.73>.

[86] PEEBLES, P. Recombination of the primeval plasma. *The Astrophysical Journal*, v. 153, p. 1, 1968.

[87] SPERGEL, D. N. et al. First year wilkinson microwave anisotropy probe (wmap) observations: Determination of cosmological parameters. *The Astrophysical Journal Supplement Series*, American Astronomical Society, v. 148, n. 1, p. 175–194, sep 2003. Disponível em: <https://doi.org/10.1086%2F377226>.

[88] COIL, A. L. The large-scale structure of the universe. In: *Planets, Stars and Stellar Systems*. Springer Netherlands, 2013. p. 387–421. Disponível em: <https://doi.org/10.1007%2F978-94-007-5609-0_8>.

[89] PEREIRA, T. d. S. *Teoria inflacionária em universos anisotrópicos*. Tese (Doutorado) — Universidade de São Paulo, 2008.

[90] VITENTI, S. D. P. *Estudo das perturbaçoes em universos com ricochete*. Tese (Doutorado) — Ph. D. thesis, Centro Brasileiro de Pesquisas Fısicas, Rio de Janeiro, Brazil, 2011.

[91] STAROBINSKY, A. A. A new type of isotropic cosmological models without singularity. *Physics Letters B*, Elsevier, v. 91, n. 1, p. 99–102, 1980.

[92] GUTH, A. H. Inflationary universe: A possible solution to the horizon and flatness problems. *Physical Review D*, APS, v. 23, n. 2, p. 347, 1981.

[93] LINDE, A. Phys lett b 108: 389. *Albrecht A, Steinhardt PJ (1982) Phys Rev Lett*, v. 48, p. 1220, 1982.

[94] MARTIN, J. *The Theory of Inflation*. arXiv, 2018. Disponível em: <https://arxiv.org/abs/1807.11075>.

[95] TSUJIKAWA, S. *Introductory review of cosmic inflation*. arXiv, 2003. Disponível em: <https://arxiv.org/abs/hep-ph/0304257>.

[96] NETO, N. P. *Hamiltonian formulation of General Relativity and applications*. [S.l.]: PPGCosmo, 2020. (Cadernos de Astrofísica, Cosmologia e Gravitação).

[97] PINTO-NETO, N. Quantum cosmology: how to interpret and obtain results. *Brazilian Journal of Physics*, SciELO Brasil, v. 30, p. 330–345, 2000.

[98] HALLIWELL, J. J. Introductory lectures on quantum cosmology. *Introductory lectures on quantum cosmology*, 1990.

[99] PINTO-NETO, N.; FABRIS, J. C. Quantum cosmology from the de broglie-bohm perspective. 2013.

[100] DEWITT, B. S. Quantum theory of gravity. i. the canonical theory. *Physical Review*, APS, v. 160, n. 5, p. 1113, 1967.

[101] KIEFER, C.; PETER, P. Time in quantum cosmology. *Universe*, MDPI, v. 8, n. 1, p. 36, 2022.

[102] MAŁKIEWICZ, P.; PETER, P.; VITENTI, S. Quantum empty bianchi i spacetime with internal time. *Physical Review D*, APS, v. 101, n. 4, p. 046012, 2020.

[103] MUKHANOV, V. *Physical foundations of cosmology*. [S.l.]: Cambridge university press, 2005.

[104] TOLMAN, R. C. *Relativity, Thermodynamics and Cosmology*. [S.l.]: Dover, 1934.

[105] LEMAîTRE, A. G. l'Univers en expansion. *Annales de la Société Scientifique de Bruxelles*, A 53, p. 51, 1933.

[106] TOLMAN, R. C. Static solutions of einstein's field equations for spheres of fluid. *Physical Review*, APS, v. 55, n. 4, p. 364, 1939.

[107] OPPENHEIMER, J. R.; VOLKOFF, G. M. On massive neutron cores. *Physical Review*, APS, v. 55, n. 4, p. 374, 1939.

[108] CARR, B. J. Some cosmological consequences of primordial black-hole evaporations. *Astrophysical Journal, vol. 206, May 15, 1976, pt. 1, p. 8-25. Research supported by the Science Research Council of England*, v. 206, p. 8–25, 1976.

[109] CARR, B. et al. New cosmological constraints on primordial black holes. *Physical Review D*, APS, v. 81, n. 10, p. 104019, 2010.

[110] ADE, P. A. et al. Planck 2015 results-xiii. cosmological parameters. *Astronomy & Astrophysics*, EDP sciences, v. 594, p. A13, 2016.

[111] BERGSTRöM, L. Dark matter candidates. *New Journal of Physics*, IOP Publishing, v. 11, n. 10, p. 105006, oct 2009. Disponível em: <https://doi.org/10.1088%2F1367-2630%2F11%2F10%2F105006>.

[112] CYBURT, R. H.; FIELDS, B. D.; OLIVE, K. A. Primordial nucleosynthesis in light of wmap. *Physics Letters B*, Elsevier, v. 567, n. 3-4, p. 227–234, 2003.

[113] ALCOCK, C. et al. The macho project: limits on planetary mass dark matter in the galactic halo from gravitational microlensing. *The Astrophysical Journal*, IOP Publishing, v. 471, n. 2, p. 774, 1996.

[114] CARR, B. et al. Constraints on primordial black holes. *Reports on Progress in Physics*, IOP Publishing, v. 84, n. 11, p. 116902, 2021.

[115] BARDEEN, J. M. Gauge-invariant cosmological perturbations. *Phys. Rev. D*, v. 22, p. 1882–1905, 10 1980.

[116] MUKHANOV, V. F.; FELDMAN, H. A.; BRANDENBERGER, R. H. Theory of cosmological perturbations. *Physics reports*, Elsevier, v. 215, n. 5-6, p. 203–333, 1992.

[117] VITENTI, S. D. P.; Penna-Lima, M. *NumCosmo: Numerical Cosmology*. 2014. Astrophysics Source Code Library ascl:1408.013.

[118] MUKHANOV, V.; WINITZKI, S. *Introduction to Quantum Effects in Gravity*. Cambridge University Press, 2007. ISBN 9780521868341. Disponível em: <https://books.google.com.br/books?id=vmwHoxf2958C>.

[119] COLLABORATION, P. Planck 2018 results. x. constraints on inflation. 2018.

[120] WANG, Y.; XUE, W. Inflation and alternatives with blue tensor spectra. *Journal of Cosmology and Astroparticle Physics*, IOP Publishing, v. 2014, n. 10, p. 075–075, out. 2014. ISSN 1475-7516. Disponível em: <http://dx.doi.org/10.1088/1475-7516/2014/10/075>.

[121] CAI, Y.-F.; WILSON-EWING, E. A Λcdm bounce scenario. *J. Cosmol. Astropart. Phys.*, v. 3, p. 006, mar. 2015. Disponível em: <http://adsabs.harvard.edu/abs/2015JCAP...03..006C>.

[122] KUROYANAGI, S.; TAKAHASHI, T.; YOKOYAMA, S. Blue-tilted inflationary tensor spectrum and reheating in the light of nanograv results. *Journal of Cosmology and Astroparticle Physics*, IOP Publishing, v. 2021, n. 01, p. 071, 2021.

[123] WU, Y.-M.; CHEN, Z.-C.; HUANG, Q.-G. Search for stochastic gravitational-wave background from massive gravity in the nanograv 12.5-year dataset. *Physical Review D*, APS, v. 107, n. 4, p. 042003, 2023.

[124] NIEMEYER, J. C.; JEDAMZIK, K. Near-critical gravitational collapse and the initial mass function of primordial black holes. *Physical Review Letters*, APS, v. 80, n. 25, p. 5481, 1998.

[125] MUSCO, I. Threshold for primordial black holes: Dependence on the shape of the cosmological perturbations. *Physical Review D*, American Physical Society (APS), v. 100, n. 12, dec 2019. Disponível em: <https://doi.org/10.1103%2Fphysrevd.100.123524>.

[126] BARDEEN, J. M. et al. The statistics of peaks of gaussian random fields. *Astrophysical Journal, Part 1 (ISSN 0004-637X), vol. 304, May 1, 1986, p. 15-61. SERC-supported research.*, v. 304, p. 15–61, 1986.

[127] MARTIN, J. et al. The Best Inflationary Models After Planck. *J. Cosmol. Astropart. Phys.*, v. 1403, p. 039, 2014.

[128] INC., W. R. *Mathematica, Version 14.0*. Champaign, IL, 2024. Disponível em: <https://www.wolfram.com/mathematica>.

[129] DEY, D. et al. The equilibrium condition in gravitational collapse and its application to a cosmological scenario. *arXiv preprint arXiv:2303.16789*, 2023.

[130] PRESS, W. H.; SCHECHTER, P. Formation of galaxies and clusters of galaxies by self-similar gravitational condensation. *The Astrophysical Journal*, v. 187, p. 425–438, 1974.

[131] CAI, Y.-F. et al. Primordial black hole mass functions as a probe of cosmic origin. *arXiv preprint arXiv:2301.09403*, 2023.

[132] BI, X.-M.; CHEN, L.; WANG, K. Primordial black hole mass function with mass gap. *Monthly Notices of the Royal Astronomical Society*, Oxford University Press, v. 527, n. 2, p. 3962–3967, 2024.

[133] CARR, B.; KüHNEL, F.; SANDSTAD, M. Primordial black holes as dark matter. *Physical Review D*, American Physical Society (APS), v. 94, n. 8, oct 2016. Disponível em: <https://doi.org/10.1103%2Fphysrevd.94.083504>.

[134] ARMENDARIZ-PICON, C.; NEELAKANTA, J. T. How cold is cold dark matter? *Journal of Cosmology and Astroparticle Physics*, IOP Publishing, v. 2014, n. 03, p. 049, 2014.

[135] POUND, R. V.; JR, G. A. R. Gravitational red-shift in nuclear resonance. *Physical Review Letters*, APS, v. 3, n. 9, p. 439, 1959.

[136] POUND, R. V.; SNIDER, J. L. Effect of gravity on gamma radiation. *Physical Review*, APS, v. 140, n. 3B, p. B788, 1965.

[137] SCHNEIDER, P. et al. Gravitational lenses as astrophysical tools. *Gravitational Lenses*, Springer, p. 467–515, 1992.

[138] KUIJKEN, K. et al. Gravitational lensing analysis of the kilo-degree survey. *Monthly Notices of the Royal Astronomical Society*, The Royal Astronomical Society, v. 454, n. 4, p. 3500–3532, 2015.

[139] Ivezić, Ž. et al. LSST: From Science Drivers to Reference Design and Anticipated Data Products. *Astrophysical Journal*, v. 873, p. 111, mar. 2019.

[140] ABOLFATHI, B. et al. The lsst desc dc2 simulated sky survey. *The Astrophysical Journal Supplement Series*, IOP Publishing, v. 253, n. 1, p. 31, 2021.

[141] RYKOFF, E. et al. redmapper. i. algorithm and sdss dr8 catalog. *The Astrophysical Journal*, IOP Publishing, v. 785, n. 2, p. 104, 2014.

[142] BOLLIET, B. et al. Including massive neutrinos in thermal sunyaev zeldovich power spectrum and cluster counts analyses. *Monthly Notices of the Royal Astronomical Society*, Oxford University Press, v. 497, n. 2, p. 1332–1347, 2020.

[143] ROZO, E.; WECHSLER et al. Cosmological constraints from the sloan digital sky survey maxbcg cluster catalog. *The Astrophysical Journal*, IOP Publishing, v. 708, n. 1, p. 645, 2009.

[144] ABBOTT, M. A. e. a. T. Dark energy survey year 1 results: Cosmological constraints from cluster abundances and weak lensing. *Physical Review D*, American Physical Society (APS), v. 102, n. 2, jul 2020. Disponível em: <https://doi.org/10.1103%2Fphysrevd.102.023509>.

[145] KIRBY, M. et al. *Improved Cosmological Constraints from SDSS redMaPPer Clusters via X-ray Follow-up of a Complete Subsample of Systems*. 2019.

[146] LIMA, M. P. *Abundância de Aglomerados de Galáxias como Observável Cosmológico: Aplicações aos Levantamentos Fotométricos DES e SDSS*. Tese (Doutorado) — Centro Brasileiro de Pesquisas Físicas, 2010.

[147] JENKINS, A. et al. The mass function of dark matter haloes. *Monthly Notices of the Royal Astronomical Society*, Blackwell Science Ltd Oxford, UK, v. 321, n. 2, p. 372–384, 2001.

[148] TINKER, J. et al. Toward a halo mass function for precision cosmology: The limits of universality. *Astrophys. J.*, v. 688, p. 709–728, 2008.

[149] YEE, H.; LÓPEZ-CRUZ, O. A quantitative measure of the richness of galaxy clusters. *The Astronomical Journal*, IOP Publishing, v. 117, n. 5, p. 1985, 1999.

[150] LIMA, M.; HU, W. Self-calibration of cluster dark energy studies: Observable-mass distribution. *Physical review D*, APS, v. 72, n. 4, p. 043006, 2005.

[151] MURATA, R. et al. The mass–richness relation of optically selected clusters from weak gravitational lensing and abundance with subaru hsc first-year data. *Publications of the Astronomical Society of Japan*, Oxford University Press, v. 71, n. 5, p. 107, 2019.

[152] COLLABORATION, P. et al. Planck intermediate results. xlvi. reduction of large-scale systematic effects in hfi polarization maps and estimation of the reionization optical depth. *A&A*, v. 596, p. A107, 2016.

[153] MANDELBAUM, R. et al. Weak lensing shear calibration with simulations of the hsc survey. *Monthly Notices of the Royal Astronomical Society*, Oxford University Press, v. 481, n. 3, p. 3170–3195, 2018.

[154] HU, W.; KRAVTSOV, A. V. Sample variance considerations for cluster surveys. *The Astrophysical Journal*, IOP Publishing, v. 584, n. 2, p. 702, 2003.

[155] LIMA, M.; HU, W. Self-calibration of cluster dark energy studies: Counts in cells. *Physical review D*, APS, v. 70, n. 4, p. 043504, 2004.

[156] PAYERNE, C. et al. Testing the accuracy of likelihoods for cluster abundance cosmology. *Monthly Notices of the Royal Astronomical Society*, Oxford University Press, v. 520, n. 4, p. 6223–6236, 2023.

[157] ZUNTZ, J. et al. Cosmosis: Modular cosmological parameter estimation. *Astronomy and Computing*, Elsevier, v. 12, p. 45–59, 2015.

[158] TORRADO, J.; LEWIS, A. Cobaya: Code for bayesian analysis of hierarchical physical models. *Journal of Cosmology and Astroparticle Physics*, IOP Publishing, v. 2021, n. 05, p. 057, 2021.

[159] CHISARI, N. E. et al. Core cosmology library: Precision cosmological predictions for lsst. *The Astrophysical Journal Supplement Series*, IOP Publishing, v. 242, n. 1, p. 2, 2019.

[160] AGUENA, M.; LIMA, M. Effects of completeness and purity on cluster dark energy constraints. *Physical Review D*, American Physical Society (APS), v. 98, n. 12, dec 2018. Disponível em: <https://doi.org/10.1103%2Fphysrevd.98.123529>.

[161] FOREMAN-MACKEY, D. et al. emcee : The MCMC hammer. *Publications of the Astronomical Society of the Pacific*, University of Chicago Press, v. 125, n. 925, p. 306–312, Mar 2013.

[162] CHRISTENSEN, N. et al. Ii: Bayesian methods for cosmological parameter estimation from cosmic microwave background measurements. *Class. Quant. Grav.*, v. 18, p. 2677, 2001.

[163] CHRISTENSEN, N. et al. Bayesian methods for cosmological parameter estimation from cosmic microwave background measurements. *Classical and Quantum Gravity*, IOP Publishing, v. 18, n. 14, p. 2677, 2001.

[164] HASTINGS, W. Monte carlo samping methods using markov chains and their applications. *Biometrika*, v. 57, p. 97–109, 1970.

[165] HARKER, G. J. et al. An mcmc approach to extracting the global 21-cm signal during the cosmic dawn from sky-averaged radio observations. *Monthly Notices of the Royal Astronomical Society*, Blackwell Publishing Ltd Oxford, UK, v. 419, n. 2, p. 1070–1084, 2012.

[166]CARUSO, D. et al. Mcmc determination of the cosmic uv background at z? 0 from h α fluorescence. *Monthly Notices of the Royal Astronomical Society*, Oxford University Press, v. 482, n. 2, p. 2833–2837, 2019.

[167]GOODMAN, J.; WEARE, J. Ensemble samplers with affine invariance. *CAMCoS*, Mathematical Sciences Publishers, v. 5, n. 1, p. 65–80, 2010.

[168]AKERET, J. et al. Cosmohammer: Cosmological parameter estimation with the mcmc hammer. *Astronomy and Computing*, Elsevier, v. 2, p. 27–39, 2013.

[169]JEFFREY, N.; WANDELT, B. D. Solving high-dimensional parameter inference: marginal posterior densities & moment networks. *arXiv preprint arXiv:2011.05991*, 2020.

[170]MORZFELD, M.; TONG, X. T.; MARZOUK, Y. M. Localization for mcmc: sampling high-dimensional posterior distributions with local structure. *Journal of Computational Physics*, Elsevier, v. 380, p. 1–28, 2019.

[171]LIU, J. S.; SABATTI, C. Generalised gibbs sampler and multigrid monte carlo for bayesian computation. *Biometrika*, Oxford University Press, v. 87, n. 2, p. 353–369, 2000.

[172]LIU, J. S. *Monte Carlo strategies in scientific computing*. [S.l.]: Springer, 2001. ISBN 0387952306.

[173]ROBERT, C.; CASELLA, G. *Monte Carlo statistical methods*. [S.l.]: Springer Science & Business Media, 2013.

[174]COULLON, J.; WEBBER, R. J. *Ensemble sampler for infinite-dimensional inverse problems*. 2020.

[175]MARONNA, R. A.; ZAMAR, R. H. Robust estimates of location and dispersion for high-dimensional datasets. *Technometrics*, Taylor & Francis, v. 44, n. 4, p. 307–317, 2002.

[176]MA, L. *Kdtree*. [S.l.]: GitHub, 2017. <https://github.com/begeekmyfriend/kdtree>.

[177]ROSENBLATT, M. Remarks on Some Nonparametric Estimates of a Density Function. *The Annals of Mathematical Statistics*, Institute of Mathematical Statistics, v. 27, n. 3, p. 832 – 837, 1956. Disponível em: <https://doi.org/10.1214/aoms/1177728190>.

[178]PARZEN, E. On Estimation of a Probability Density Function and Mode. *The Annals of Mathematical Statistics*, Institute of Mathematical Statistics, v. 33, n. 3, p. 1065 – 1076, 1962. Disponível em: <https://doi.org/10.1214/aoms/1177704472>.

[179] SILVERMAN, B. W. *Density estimation for statistics and data analysis*. [S.l.]: Routledge, 2018.

[180] BRO, R.; JONG, S. D. A fast non-negativity-constrained least squares algorithm. *Journal of Chemometrics*, v. 11, n. 5, p. 393–401, 1997. Disponível em: <https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/abs/10.1002/%28SICI%291099-128X%28199709/10%2911%3A5%3C393%3A%3AAID-CEM483%3E3.0.CO%3B2-L>.

[181] LEWIS, A. GetDist: a Python package for analysing Monte Carlo samples. *ArXiv e-prints*, 2019. Disponível em: <https://getdist.readthedocs.io>.

[182] PAGANI, F.; WIEGAND, M.; NADARAJAH, S. *An n-dimensional Rosenbrock Distribution for MCMC Testing*. 2020.

[183] DOUX, C. et al. Cosmological constraints from a joint analysis of cosmic microwave background and spectroscopic tracers of the large-scale structure. *Mon. Not. R. Astron. Soc.*, v. 480, p. 5386–5411, nov. 2018. Disponível em: <http://adsabs.harvard.edu/abs/2018MNRAS.480.5386D>.

[184] BROUT, D. et al. The Pantheon+ Analysis: Cosmological Constraints. *Astrophys. J.*, v. 938, n. 2, p. 110, 2022.

[185] SCOLNIC, D. et al. The Pantheon+ Analysis: The Full Data Set and Light-curve Release. *Astrophys. J.*, v. 938, n. 2, p. 113, 2022.

[186] ALAM, S. et al. The clustering of galaxies in the completed sdss-iii baryon oscillation spectroscopic survey: cosmological analysis of the dr12 galaxy sample. *Mon. Not. R. Astron. Soc.*, v. 470, p. 2617–2652, jul. 2017.

[187] ALAM, S. et al. Completed sdss-iv extended baryon oscillation spectroscopic survey: Cosmological implications from two decades of spectroscopic surveys at the apache point observatory. *Phys. Rev. D*, v. 103, n. 8, p. 083533, 2021.

[188] GÓMEZ-VALENT, A.; AMENDOLA, L. h_0 from cosmic chronometers and type ia supernovae, with gaussian processes and the weighted polynomial regression method. In: *15th Marcel Grossmann Meeting on Recent Developments in Theoretical and Experimental General Relativity, Astrophysics, and Relativistic Field Theories*. [S.l.: s.n.], 2019.

[189] VITENTI, S. D. P.; FALCIANO, F.; PINTO-NETO, N. Covariant bardeen perturbation formalism. *Physical Review D*, APS, v. 89, n. 10, p. 103538, 2014.

[190] PINTO-NETO, N. et al. Vector perturbations in bouncing cosmology. *Physical Review D*, APS, v. 101, n. 12, p. 123519, 2020.

[191] VITENTI, S. D. P.; PINTO-NETO, N. Large adiabatic scalar perturbations in a regular bouncing universe. *Physical Review D*, APS, v. 85, n. 2, p. 023524, 2012.

[192] BARDEEN, J. M. Gauge-invariant cosmological perturbations. *Physical Review D*, APS, v. 22, n. 8, p. 1882, 1980.

[193] VITENTI, S.; FALCIANO, F.; PINTO-NETO, N. Covariant bardeen perturbation formalism. *Physical Review D*, APS, v. 89, n. 10, p. 103538, 2014.

[194] MUKHANOV, V.; WINITZKI, S. *Introduction to quantum effects in gravity*. [S.l.]: Cambridge university press, 2007.

[195] SUMMERS, S. On the stone-von neumann uniqueness theorem and its ramifications. 12 1998.

[196] REED, M. *Methods of modern mathematical physics: Functional analysis*. [S.l.]: Elsevier, 2012.

[197] MUKHANOV, V. F.; CHIBISOV, G. Quantum fluctuations and a nonsingular universe. *ZhETF Pisma Redaktsiiu*, v. 33, p. 549–553, 1981.

[198] HAWKING, S. W. The development of irregularities in a single bubble inflationary universe. *Physics Letters B*, Elsevier, v. 115, n. 4, p. 295–297, 1982.

[199] GREINER, W.; REINHARDT, J. *Field quantization*. [S.l.]: Springer Science & Business Media, 2013.

[200] STEPANIAN, A.; KOHANDEL, M. *Unitary Inequivalent Representations and Quantum Physics*. 2017.

[201] PENNA-LIMA, M.; PINTO-NETO, N.; VITENTI, S. D. New formalism to define vacuum states for scalar fields in curved spacetimes. *Physical Review D*, APS, v. 107, n. 6, p. 065019, 2023.

[202] BIRRELL, N. D.; DAVIES, P. C. W. Quantum fields in curved space. Cambridge university press, 1984.

[203] WALD, R. M. *Quantum field theory in curved spacetime and black hole thermodynamics*. [S.l.]: University of Chicago press, 1994.

[204] BIRRELL, N. D.; DAVIES, P. C. W. Quantum fields in curved space. Cambridge university press, 1984.

[205] PENNA-LIMA, M.; PINTO-NETO, N.; VITENTI, S. D. New formalism to define vacuum states for scalar fields in curved space-times. *arXiv preprint arXiv:2207.08270*, 2022.

[206] BAUGH, C.; MURDIN, P. Correlation function and power spectra in cosmology. *Encyclopedia of Astronomy and Astrophysics,(IOP, London, UK, 2006)*, 2006.

[207] GONÇALVES, S. M. Black hole formation from massive scalar field collapse in the einstein–de sitter universe. *Physical Review D*, APS, v. 62, n. 12, p. 124006, 2000.

Apêndices

APÊNDICE A – PERTURBATION THEORY

In this appendix, we review perturbation theory in general relativity. We consider a Friedmann-Lemaître-Robertson-Walker (FLRW) background Universe, which is spatially homogeneous and isotropic. Even though this model is a good approximation of the real Universe globally, we want to consider a physical Universe that slightly differ from this background locally, which can be modeled with perturbation theory.

We represent the quantities defined in the background universe and mapped in the physical Universe with an over-bar sign, such as in the background metric tensor $\bar{g}^{\mu\nu}$. The physical, or perturbed, quantities are labeled without the bar sign. In this section, we study the perturbation theory for a general background curvature $\bar{K}$ for learning purposes, although we assume that the universe is flat in the rest of this work, i.e., $\tilde{K} = 0$.

A.1 Definitions and Gauge Dependency

Even though the FLRW Universe is a good first approximation for the real Universe, it does not completely describes the physical Universe. It is expected that the physical, which we will also call the perturbed, Universe differs minimally from the Friedman background version [69, 116]. The background model is defined in a manifold M^{bg} that is assumed to be spatially homogeneous and isotropic. To analyze how the perturbed manifold M representing the physical Universe differs from this ideal model, we first need to analyze how to connect both manifolds trough a diffeomorphism. We will analyze this map and then compare the mapped quantities to the perturbed ones.

Following Ref. [189], considering a quantity Q^{bg} defined on a background $3+1$ manifold M^{bg}, we introduce this quantity in the physical $3+1$ manifold M through a one parameter diffeomorphism map f_λ, such that we label these mapped quantities $\bar{Q} \in M$. This diffeomorphism map can be seen in Fig. 20.

Knowing how to define the background quantities in the physical manifold, we are able to compute the difference between them and their perturbed versions, that is,

$$\delta Q = Q - \bar{Q}, \tag{A.1}$$

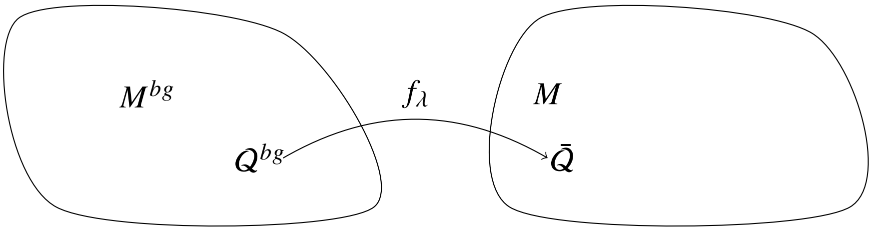


Figura 20 – Map between the background and the physical manifolds.

where Q is the quantity defined in the physical manifold and we take δQ to be small. From now on, we refer to quantities brought from the background to the physical manifold as background quantities, labeled with an over-bar sign.

Since background quantities depends on the map used to connect the manifolds, we must analyze the dependency given by the parameter λ. Also, we have to take into account that the quantities definitions depend on the coordinate system. Thus, we shall call the dependency on the diffeomorphism map plus the coordinate system as a Gauge dependency. Let us consider transformations generated by an infinitesimal change in the coordinates, characterized by a vector A, that is, $x_\mu \longrightarrow y_\mu = x_\mu + A_\mu$. Under this transformation, a general tensor changes as

$$\begin{aligned}\bar{T}_{\sigma\alpha\lambda\ldots}(x) &\longrightarrow \bar{T}_{\sigma\alpha\lambda\ldots}(y) \\ &= \bar{T}_{\sigma'\alpha'\lambda'\ldots}(x)\frac{\partial y_\nu}{\partial x^{\nu'}}\frac{\partial y_\alpha}{\partial x^{\alpha'}}\ldots. \end{aligned} \tag{A.2}$$

Expanding $\bar{T}_{\sigma\alpha\lambda\ldots}(x = y - A)$ in a Taylor series, we get

$$\bar{T}_{\sigma\alpha\lambda\ldots}(y) = \bar{T}_{\sigma\alpha\lambda\ldots}(x) + \pounds_A \bar{T}_{\sigma\alpha\lambda\ldots}(x). \tag{A.3}$$

However, the coordinate choice also affects the physical quantities. Explicitly,

$$T_{\sigma\alpha\lambda\ldots}(y) = \bar{T}_{\sigma\alpha\lambda\ldots}(x) + \pounds_A \bar{T}_{\sigma\alpha\lambda\ldots}(x), \tag{A.4}$$

such that perturbations as in Eq. (A.1) remain unchanged.

On the other hand, we assume that for a small λ change, the new background quantities vary with the same form, that is,

$$\bar{T}_{\sigma\alpha\lambda\ldots}(y) = \bar{T}_{\sigma\alpha\lambda\ldots}(x) + \pounds_B \bar{T}_{\sigma\alpha\lambda\ldots}(x), \tag{A.5}$$

where B is a vector that characterizes the change in the diffeomorphism map. Since the physical quantities are not affected by this map, the resulted perturbations will be

$$\delta T_{\sigma\alpha\lambda\ldots}(x) \longrightarrow \bar{T}_{\sigma\alpha\lambda\ldots}(x) - \pounds_B \bar{T}_{\sigma\alpha\lambda\ldots}(\bar{x}). \tag{A.6}$$

Thus, we refer to a gauge transformation the total change[1]

$$\bar{T} \longrightarrow \bar{T} + \pounds_{A+B}\bar{T} \tag{A.7}$$

$$T \longrightarrow T + \pounds_A \bar{T}. \tag{A.8}$$

In this work, we want to make a coordinate choice such that the background quantities are invariant under a Gauge choice, that is, $B = -A$, which leads to

$$\bar{T} \longrightarrow \bar{T}, \tag{A.9}$$

and

$$\delta T \longrightarrow \delta T - \pounds_B T. \tag{A.10}$$

So from now one, we assume the a gauge choice will only affect the perturbed quantities and left background variables unchanged.

We now apply this same formalism to the metric tensor. In perturbation theory, we assume that the metric defined in the background and mapped in the perturbed manifold is a FLRW metric, that is,

$$\begin{aligned} \bar{g}_{\mu\nu} &= f_{\lambda=0} g^{bg}_{\mu\nu}, \quad \mu, \nu = 0, 1, 2, 3. \\ &= -c^2 dt^2 + \bar{\gamma}_{ij} dx^i dx^j, \quad ij = 1, 2, 3, \end{aligned} \tag{A.11}$$

being $\bar{g}_{\mu\nu}$ the FLRW metric, t the cosmic time and $\bar{\gamma}_{ij} = \bar{a}^2 \delta_{ij}$ the spatial metric. We assume that the physical metric $g^{\mu\nu}$ has a small deviation from this metric, that is,

$$\delta g_{\mu\nu} = g_{\mu\nu} - \bar{g}_{\mu\nu}. \tag{A.12}$$

Finally, under a gauge transformation represented by a vector ξ, the metric perturbation changes as

$$\begin{aligned} \delta g^{\lambda}_{\mu\nu} &= \delta g_{\mu\nu} - \pounds_\xi \bar{g}_{\mu\nu} \\ &= \delta g_{\mu\nu} - \bar{\nabla}_\mu \xi_\nu - \bar{\nabla}_\nu \xi_\mu. \end{aligned} \tag{A.13}$$

The next step is to decompose and analyze the metric and matter perturbations.

---

[1] We are omitting the tensor indexes for reading simplification purposes.

A.2 Metric and Matter Scalar Perturbations

A.2.1 Metric Perturbations

In this work, we will only focus on scalar perturbations. For vector perturbations, check Ref. [190]. Following Refs. [189, 191], we may write the physical metric as

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + \delta g_{\mu\nu}, \tag{A.14}$$

such that $\delta g_{\mu\nu}$ represents the scalar perturbations given by

$$\delta g_{00} = 2\phi \tag{A.15}$$

$$\delta g_{0i} = -\bar{D}_i \mathcal{B} \tag{A.16}$$

$$\delta g_{ij} = 2\left(\psi \bar{\gamma}_{ij} - \bar{D}_i \bar{D}_j \mathcal{E}\right). \tag{A.17}$$

In the above, $\mathcal{B}, \phi, \psi$ and $\mathcal{E}$ are scalar perturbations of the metric. These equations represents the possible projections of the metric in time and spatial background directions and the scalar components are derived trough the Scalar-Vector-Tensor (SVT) decomposition.

Let us now establish a foliation for the physical universe, determining how the spatial hypersurfaces with constant times are selected. This will allow us to study the projections for different tensors and analyze their components. In this study, we adopt a mixed foliation approach for the perturbations. This approach involves using a geodesic foliation defined by a vector $\bar{n}^\mu$ for the background manifold, as discussed in Sec. 2.1.1, while parameterizing the perturbed manifold with $n^\mu = (1-\phi)\bar{n}^\mu$ [90]. By choosing this foliation we achieve more physical comparisons between the perturbed and background quantities, since the foliation for the perturbed manifold aligns with the background's up to a small deviation and we can use the same notion of hypersurfaces for both manifolds. Thus, given a generic tensor $W^{\mu\nu}$, we will denote perturbations projected on the background foliation, that is

$$\delta W^{\mu\nu} n_\nu n_\mu = \delta W_{00} = W^{\mu\nu} \bar{n}_\mu \bar{n}_\nu - \bar{W}^{\mu\nu} \bar{n}_\mu \bar{n}_\nu. \tag{A.18}$$

Let us now define the matter perturbations that will be needed for the Einstein's equations.

A.2.2 Energy-Momentum Tensor Perturbations

The physical matter content should not differ much from a perfect fluid approximation that leads to the Friedmann equations. The background and the perturbed energy-momentum tensors are respectively given by

$$\bar{T}_{\mu}{}^{\nu} = \bar{\rho}\bar{n}_{\mu}\bar{n}^{\nu} + \bar{p}\bar{\gamma}_{\mu}{}^{\nu} \tag{A.19}$$

and

$$T_{\mu}{}^{\nu} = \bar{T}_{\mu}{}^{\nu} + \delta T_{\mu}{}^{\nu} = \delta\rho\bar{n}_{\mu}\bar{n}^{\nu} + (\bar{\rho} + \bar{p})(\bar{D}_{\mu}\mathcal{V}\bar{n}^{\nu} + \bar{n}_{\mu}(\bar{D}^{\nu}\mathcal{V} + \bar{D}^{\nu}\mathcal{B})) + \delta\bar{p}\bar{\gamma}_{\mu}{}^{\nu}, \tag{A.20}$$

where $\bar{\rho}$ is the background energy density, $\bar{p}$ is the background pressure and we considered that there is no anisotropic pressure nor anisotropic pressure perturbation. In the above, $\mathcal{V}$ is the perturbation on the fluid velocity

$$u_{\mu} = (1 - \phi)\bar{n}_{\mu} + \bar{D}_{\mu}\mathcal{V}. \tag{A.21}$$

In this approach, we are incorporating the cosmological constant contribution from Eq. (2.12) into the energy-momentum tensor. We now have all the ingredients to compute the movement equations through the Einstein equations, which will provide relations between the geometry and the matter content in perturbation theory.

A.3 Equations of Motion

Our objective is to use the Einstein equations formulated for the physical manifold in order to establish connections among the perturbed quantities. These equations are represented by

$$G_{\mu\nu} = \kappa T_{\mu\nu}, \tag{A.22}$$

where the cosmological constant term is already included in the energy-momentum tensor, given by equation (A.20). After multiplying Eq. (A.22) by $g^{\nu\alpha}$ and subtracting the result by $\bar{G}_{\mu}{}^{\alpha}$ from both sides, we get the relation for the perturbations

$$\delta G_{\mu}{}^{\alpha} = \kappa\delta T_{\mu}{}^{\alpha}. \tag{A.23}$$

We can now compute Eq. (A.23) for all components, which is explicitly done in Ref. [191]. In this context, the scalar equations of motion for a Universe with a FLRW background metric results in[2]

$$\delta G_0{}^0 = \frac{2\bar{\Theta}\delta\Theta}{3} + \frac{\delta\mathcal{R}}{2} = \kappa\delta\rho = \kappa\delta T_0{}^0 \tag{A.24}$$

$$\begin{aligned}\delta G_i{}^0 &= \frac{2}{3}\bar{D}_i\left(\delta\Theta - 3\bar{K}\delta\sigma - \bar{D}^2\delta\sigma\right)\\ &= \bar{D}_i[\kappa(\bar{\rho}+\bar{p})\mathcal{V}] = \kappa\delta T_i{}^0\end{aligned} \tag{A.25}$$

$$\begin{aligned}\delta G_i{}^i &= -2\left(\delta\dot{\Theta} + \bar{\Theta}\delta\Theta + \phi\dot{\bar{\Theta}} - \bar{D}_\mu\hat{a}^\mu + \frac{\delta\mathcal{R}}{4}\right)\\ &= 3\kappa\delta p = \kappa\delta T_i{}^i\end{aligned} \tag{A.26}$$

$$\begin{aligned}\delta G_j{}^i &= -\bar{D}^i\bar{D}_j\left(\psi - \phi - \delta\dot{\sigma} - \frac{\bar{\Theta}}{3}\delta\sigma\right)\\ &= 0 = \kappa\delta T_j{}^i, \quad \text{for } i \neq j.\end{aligned} \tag{A.27}$$

In the equations above, $\delta\Theta$ is the perturbed expansion rate, $\delta\mathcal{R}$ is the perturbed curvature scalar, $\delta\sigma$ is the perturbed shear and a^i is the worldline acceleration with respect to the constant cosmic time hyper-surfaces. These perturbation quantities can be written in terms of the scalar metric perturbations as

$$\delta\Theta = \bar{D}^2\delta\sigma + 3\left(\frac{\bar{\Theta}}{3}\phi + \dot{\psi}\right) \tag{A.28}$$

$$\delta\mathcal{R} = -4(\bar{D}^2 + 3\bar{K})\psi \tag{A.29}$$

$$\delta\sigma = -(\dot{\mathcal{E}} - \mathcal{B}) + \frac{2}{3}\bar{\Theta}\mathcal{E} \tag{A.30}$$

$$a^\mu = -\bar{D}^\mu\phi. \tag{A.31}$$

The relations in Eqs. (A.24)-(A.27) depend on the scalar metric perturbations, which defines the physical quantities that are gauge dependent. We are generally not interested in a theory that depends on the gauge since the physical meaning of the quantities may change. Therefore, our aim for the remainder of this appendix is to reframe these variables in a manner that renders them gauge invariant to express our future results. We will also briefly discuss some possible gauge choices.

---

[2] Had we defined the metric perturbations with the vector perturbations as well, we could get these same scalar equations (A.24)-(A.27) respectively by contracting equation (A.23) with the temporal vectors $\bar{n}^\mu\bar{n}_\alpha$, taking the divergent of the mixed projection, the trace of the spatial projection and the double derivative of the spatial projection.

A.4 Gauge Transformations and Gauge Invariant Quantities

To address the problem of gauge dependence, our first step is to examine the behavior of perturbations under a gauge transformation. By understanding these changes, we can determine the appropriate modifications required to express the perturbations in a gauge invariant form. Let us start with the scalar metric perturbations.

A.4.1 Scalar Metric Variables Transformation

Under a gauge transformation, the metric perturbation transforms as Eq. (A.13), being ξ the generator of the transformation with the general form[3]

$$\xi_\mu \equiv \xi^{\parallel}\bar{n}_\mu + \bar{D}_\mu \xi^{\perp}, \tag{A.32}$$

where $\xi^{\parallel}$ and $\xi^{\perp}$ are the temporal and spatial components respectively. We can analyze the change in the temporal and spatial projections of the metric perturbation to analyze how each scalar perturbation transforms under this gauge transformation. Explicitly, from Eq (A.13),

$$\delta g^{\lambda}_{nn} = \delta g_{\bar{n}\bar{n}} + 2\bar{\nabla}_{\bar{n}}\xi_{\bar{n}} = 2\phi - 2\dot{\xi}^{\parallel}, \tag{A.33}$$

$$\begin{aligned} \gamma[\delta g_{\bar{n}\nu}{}^{\lambda}] &= \gamma[\delta g_{\bar{n}\nu} - \bar{\nabla}_{\bar{n}}\xi_\nu - \bar{\nabla}_\nu \xi_{\bar{n}}] \\ &= -\bar{D}_\nu \mathcal{B} + \bar{D}^\nu \xi^{\parallel} - \partial_{ct}(\bar{D}^\nu \xi^{\perp}) \end{aligned} \tag{A.34}$$

and

$$\begin{aligned} \gamma[\delta \bar{g}^{\lambda}_{\mu\nu}] &= \gamma[\delta \bar{g}_{\mu\nu} - \bar{\nabla}_\mu \xi_\nu - \bar{\nabla}_\nu \xi_\mu] \\ &= 2\left(\psi\bar{\gamma}_{\mu\nu} - \bar{D}_\mu \bar{D}_\nu \mathcal{E}\right) - 2\bar{\Theta}_{\mu\nu}\xi^{\parallel} - 2\bar{D}_\mu \bar{D}_\nu \xi^{\perp}. \end{aligned} \tag{A.35}$$

---

[3] We are considering that ξ_μ is the same order as the perturbations. Also, using the SVT decomposition, we are only taking into account the scalar components of this vector.

Knowing that $\bar{\Theta}_{\mu\nu} = \frac{\bar{\Theta}}{3}\bar{\gamma}_{\mu\nu}$ and $[\partial_{ct}, \bar{\nabla}^{\mu}] = -\frac{2}{3}\bar{\Theta}\bar{D}^{\mu}$ [90], we deduce that the scalar perturbations evolve under gauge transformation as

$$\phi \longrightarrow \phi + \dot{\xi}^{\parallel} \tag{A.36}$$

$$\mathcal{B} \longrightarrow \mathcal{B} - \xi^{\parallel} - \frac{2\bar{\Theta}}{3}\xi^{\perp} + \dot{\xi}^{\perp} \tag{A.37}$$

$$\mathcal{E} \longrightarrow \mathcal{E} + \xi^{\perp} \tag{A.38}$$

$$\psi \longrightarrow \psi - \frac{\bar{\Theta}}{3}\xi^{\parallel}. \tag{A.39}$$

Using equations (A.37) and (A.38), we can also compute the gauge transformation for the shear perturbation as

$$\delta\sigma \longrightarrow \delta\sigma - \xi^{\parallel}. \tag{A.40}$$

With the transformations of the metric variables we can form two new gauge invariant variables, known as the Bardeen variables [192], defined as

$$\Phi \equiv \phi + \dot{\delta\sigma}, \quad \Psi \equiv \psi - \frac{\bar{\Theta}}{3}\delta\sigma. \tag{A.41}$$

We must be careful when interpretation these variables. They have no direct physical meaning since they are just a combination of variables that form gauge invariant quantities. Nevertheless, they are commonly used, and we will see that in the Newtonian gauge ($\delta\sigma = 0$) the first variable is equivalent to the Newtonian potential. Thus, these variables are convenient if the goal is to work in this gauge.

We could have also formed different gauge invariant variables by combining Eqs. (A.36)-(A.39). When working in certain gauges, one should always analyze which variables fit best the problem in question to simplify the analysis. We will briefly discuss this point in the end of this section. We shall now compute the variation of the energy-momentum tensor under a gauge transformation and combine gauge dependent variables to form invariant ones.

A.4.2 Energy-Momentum Tensor Transformations

Following the same approach using Eq. (A.13), the energy-momentum tensor transforms under a gauge transformation as

$$\delta T^{\lambda}_{\mu\nu} \longrightarrow \delta T_{\mu\nu} - \pounds_{\xi}\delta T_{\mu\nu}. \tag{A.42}$$

To analyze all the components, we start with the temporal projection of Eq. (A.42), such that

$$\delta T^{\lambda}_{\bar{n}\bar{n}} \rightarrow \delta T_{\bar{n}\bar{n}} - \pounds_{\xi}\delta T_{\bar{n}\bar{n}}. \tag{A.43}$$

Comparing this relation to Eq. (A.20), we deduce that the energy density transforms as

$$\delta\rho \rightarrow \delta\rho - \xi^{\|}\dot{\bar{\rho}}. \tag{A.44}$$

For the fluid velocity[4] and pressure perturbation, we use the temporal-spatial projection and the spatial-spatial projection of Eq. (A.42). In this context, these variables change under a gauge transformation as

$$\mathcal{V} \rightarrow \mathcal{V} + \xi^{\|}, \tag{A.45}$$

and

$$\delta p \rightarrow \delta p + \xi^{\|}\dot{\bar{p}}. \tag{A.46}$$

We want to combine these variables with the metric components to generate gauge invariant variables. To have a similar relations to the Bardeen variables, we will define all the gauge invariant variables using the shear perturbation given in equation (A.29) [90]. Therefore, we can combine these variables to generate their gauge invariant forms, which are

$$\tilde{\delta\rho} \equiv \delta\rho - \delta\sigma\dot{\bar{\rho}} \tag{A.47}$$

$$\tilde{\mathcal{V}} \equiv \mathcal{V} + \delta\sigma \tag{A.48}$$

$$\tilde{\delta p} \equiv \delta p - \delta\sigma\dot{\bar{p}}. \tag{A.49}$$

It is interesting to use only one variable to define the gauge-invariant forms of the other variables (which in our case is $\delta\sigma$), so when we do make a gauge choice, there is a clear interpretation.

Next, we want to rewrite some of the movement equations in terms of these invariant variables, which will lead to the definition of the curvature perturbation variable that connects both the geometric and matter components.

---

[4] It is important to notice that, had we considered the vector perturbations as well, the perturbation of the fluid velocity would be given by $V_{\mu} = \bar{D}_{\mu}\mathcal{V} + \mathbb{V}_{\mu}$. Therefore the gauge transformation would be on the vector V_{μ} and that is why the scalar term changes as in equation (A.45).

A.4.3 Curvature Perturbation

First, we see that the relation in Eq. (A.27) becomes trivial in terms of the Bardeen variables, that is,

$$\Psi = \Phi. \tag{A.50}$$

Next, we want to combine Eqs. (A.25) and (A.26). Starting with Eq. (A.25), we integrate both sides of the equation in the space variables and multiply the result by $\bar{\Theta}$. Then, we subtract the result by Eq. (A.24), such that

$$-(3\bar{K} + \bar{D}^2)\Psi = \frac{\kappa}{2}\left[\delta\rho - \bar{\Theta}(\bar{\rho} + \bar{p})\mathcal{V}\right]. \tag{A.51}$$

In the above, we also used the relation for $\delta\mathcal{R}$ given in Eq. (A.29).

Using Eqs. (A.47) and (A.48), plus the continuity background equation

$$\dot{\bar{\rho}} + \bar{\Theta}(\bar{\rho} + \bar{p}) = 0, \tag{A.52}$$

we can rewrite Eq. (A.51) with the gauge-invariant quantities as

$$-(3\bar{K} + \bar{D}^2)\Psi = \frac{\kappa}{2}\left[\tilde{\delta\rho} - 3\bar{\Theta}(\bar{\rho} + \bar{p})\tilde{\mathcal{V}}\right]. \tag{A.53}$$

We can rearrange Eq. (A.53) so

$$-\bar{D}^2\Psi + \frac{3\kappa(\bar{\rho} + \bar{p})}{2}\left(\frac{\bar{\Theta}}{3}\tilde{\mathcal{V}} - \frac{2\bar{K}\Psi}{(\bar{\rho} + \bar{p})}\right) = \frac{\kappa}{2}\tilde{\delta\rho}. \tag{A.54}$$

To simplify this relation, we define the curvature perturbation

$$\zeta \equiv \Psi - \frac{2\bar{K}\Psi}{\kappa(\bar{\rho} + \bar{p})} + \frac{\bar{\Theta}}{3}\tilde{\mathcal{V}}, \tag{A.55}$$

which is gauge invariant by constructionso. Replacing this relation in Eq. (A.54),

$$\zeta - \Psi - \frac{2\bar{D}^2\Psi}{3\kappa(\bar{\rho} + \bar{p})} = \frac{\tilde{\delta\rho}}{3(\bar{\rho} + \bar{p})}. \tag{A.56}$$

With Eq. (A.56), we have a relation for the density contrast that is gauge-invariant and is related with the Bardeen variables, which connect both the matter and the geometry content of the universe. We now have a well defined perturbation theory and all the needed gauge invariant variables. Let us see some possible gauge choices.

A.5 Gauge Choices

As discussed, each gauge-invariant combination of variables may be chosen depending on what gauge we intend to work with. In this section, we will analyze the Newtonian, Synchronous, Constant Curvature gauge, and Fluid's Gauge and define the best suited variables to work in each gauge.

A.5.1 Newtonian Gauge

In the Newtonian gauge, we fix the gauge in Eq. (A.32) to eliminate the scalar part of the shear, ensuring that $\delta\sigma = 0$, which fully determines the gauge. This implies that spatial and orthogonal quantities in the background remain spatial and orthogonal when inserted in the physical manifold. For this choice, we can construct gauge-invariant variables by combining all perturbation components with $\delta\sigma$. This allows us to obtain a collection of quantities that remain invariant under gauge transformations and, in the case of choosing the Newtonian gauge, reduce to the familiar Newtonian variables.

In this gauge, the chosen gauge-invariant variables are

$$\begin{aligned} \tilde{\delta\rho} &\equiv \delta\rho - \delta\sigma\dot{\bar{\rho}}, & \tilde{\mathcal{V}} &\equiv \mathcal{V} + \delta\sigma^{\mathrm{s}} \\ \Psi &\equiv \psi - \frac{\bar{\Theta}}{3}\delta\sigma^{\mathrm{s}}, & \Phi &\equiv \phi + \delta\dot{\sigma}^{\mathrm{s}}, \\ \zeta &\equiv \Psi - \frac{2\bar{K}\Psi}{\kappa(\bar{\rho}+\bar{p})} + \frac{\bar{\Theta}}{3}\tilde{\mathcal{V}}, \end{aligned} \tag{A.57}$$

which were already discussed in this Appendix.

A.5.2 Synchronous Gauge

The Synchronous gauge is determined by selecting the background hypersurfaces in such a way that their normal vector fields follow geodesics. Thus, given the acceleration in Eq. (A.31), by choosing $\phi = 0$, we achieve geodesic normal fields. However, analyzing Eq. (A.36), this choice does not completely fix the gauge since ϕ only changes with the derivative of the temporal component under gauge transformations. Rather than addressing this caveat from a gauge-fixing perspective, we adopt a similar approach as in the Newtonian gauge, seeking variables that reduce to those in the Synchronous gauge when $\phi = 0$. To do so, we introduce the integral

$$I_\phi \equiv \int_{t_a}^{t} \mathrm{d}t_1 \phi\left(t_1\right). \tag{A.58}$$

In the context above, we then define the variables

$$\begin{aligned} \tilde{\delta}^s_\rho &\equiv \delta_\rho - \bar{\Theta} I_\phi \quad & \tilde{\mathcal{V}}^s &\equiv \mathcal{V} - I_\phi \\ \Sigma &\equiv \delta\sigma^{\mathrm{s}} + I_\phi, \quad & \Psi^s &\equiv \psi + \tfrac{\bar{\Theta}}{3} I_\phi, \end{aligned} \tag{A.59}$$

where δ_ρ is the density contrast. Although these combinations are not gauge-invariant, they follow a simple rule under gauge transformations,

$$\begin{aligned} \tilde{\delta}^s_\rho &\to \delta^s_\rho - \bar{\Theta}\xi^{\parallel} \quad & \tilde{\mathcal{V}}^s &\to \mathcal{V} - \xi^{\parallel} \\ \Sigma &\to \delta\sigma^{\mathrm{s}} + \xi^{\parallel}, \quad & \Psi^s &\to \psi + \tfrac{\bar{\Theta}}{3}\xi^{\parallel} \end{aligned}\,. \tag{A.60}$$

Consequently, we can freely choose the value of one of these functions at a specific time t to completely fix the gauge. As a common practice, the Cold Dark Matter (CDM) velocity potential $\tilde{V}^s_c$ is initially set to zero, remaining constant throughout the evolution due to its equation of motion. When this additional step is taken, the mentioned variables above become gauge-invariant.

A.5.3 Constant Curvature Gauge

The intrinsic curvature perturbation is governed by the metric potential ψ as given by Eq. (A.29). Hence, selecting $\psi = 0$ establishes a gauge where the spatial hypersurfaces possess zero intrinsic curvature. Similar to the Newtonian gauge, this choice entirely fixes the gauge freedom. In this context, we introduce the gauge-invariant variables

$$\begin{aligned} \tilde{\delta}^c_\rho &\equiv \delta_\rho + 3\psi, \quad & \tilde{\mathcal{V}}^c &\equiv \mathcal{V} + \frac{3\psi}{\bar{\Theta}} \\ \Sigma^c &\equiv \delta\sigma^{\mathrm{s}} - \frac{3\psi}{\bar{\Theta}}, \quad & \Phi^c &\equiv \phi + {}_{\bar{v}}\left(\frac{3\psi}{\bar{\Theta}}\right)' \end{aligned} \tag{A.61}$$

which have direct physical interpretation in this gauge.

In the next appendix, we want to quantize the perturbations defined in Sec. A.4. We begin with the classical Mukhanov-Sasaki action, which we will show in App. B.1.2 that is related to the curvature variable. Then, we will use this action to dictate the dynamics of our variables and proceed to the quantization approach.

A.5.4 Fluid's Gauge

In App.C, we compute the solution for the Einstein equations for an inhomogeneous dust-perfect fluid with a spherically symmetric local metric. To use this solution and compare it with our perturbations defined in this appendix, we must ensure that measurements in the local metric are in the same Gauge as our perturbations. Since we projected the energy-momentum tensor in the direction of the velocity of the fluid in Eq. (C.8), we consider the Gauge where the fluid is at rest and we shall now see how to define our perturbation theory in this Gauge choice.

Essentially, following [189], we want to make sure that the local tensor in (C.6) can be seen as the perturbed energy-momentum tensor $\delta T_{\mu\nu}$. Explicitly, we want

$$T_{\mu\nu} \propto \delta T_{\mu\nu} = (\delta\rho - 2\phi) n_\mu n_\nu + 2(\rho + p) n_{(\mu} \bar{D}_{\nu)} \mathcal{V} + (\delta p)\ \mathrm{h}_{\mu\nu} + 2p \left(n_{(\mu} \bar{D}_{\nu)} \mathcal{B} + \bar{D}_\mu \bar{D}_\nu \mathcal{E} \right), \tag{A.62}$$

where we consider no anisotropic pressure. Thus, the fluid's Gauge is obtained by setting

$$\mathcal{V} = \mathcal{E} = 0 \ \text{ and} \tag{A.63}$$

$$\mathcal{B}|_{t_1} = 0. \tag{A.64}$$

With these Gauge choices,

$$\delta T_{\mu\nu} = (\delta\rho - 2\phi) n_\mu n_\nu + (\delta p \psi)\ \mathrm{h}_{\mu\nu} + 2p \left(\bar{D}_\mu \bar{D}_\nu \mathcal{E} \right). \tag{A.65}$$

The first choice for $\mathcal{V}$ in Eq. (A.63) assures that the perturbed fluid is at rest with the background universe. The other two conditions guarantee that, at least for an initial time t_1, there are no off-diagonal terms and we have an isotropic perturbed fluid. For different times, it is not possible to set a Gauge such that $\mathcal{E} = 0$ for $t \neq t_1$ [51]. Nonetheless, we can argue that this term is proportional to the pressure of the fluid which has an almost vanishing value for cold dark matter and thus can be discarded. Also, since our main goal is to make proper measurements of the density contrast that depends on Eq. (4.7), this term is not relevant to our computations. Finally, in the fluid's Gauge, the gauge-invariant density contrast may be interpreted as its physical equivalent, that is,

$$\tilde{\delta\rho}^{F} \equiv \delta\rho + \mathcal{V} \dot{\bar{\rho}} = \delta\rho \tag{A.66}$$

where we used Eq. (4.7).

APÊNDICE B – QUANTIZING THE PERTURBATIONS

In Section 3.3.2, we delved into the quantization of the background theory. In Chapter 4, our focus shifts to examining the growth of perturbations within this quantum background, which will seed the primordial black hole formation. Consequently, the next natural step is to quantize the perturbations. Fortunately, the cosmological perturbations are represented through a gauge theory that can be seen independently from the background. This independence permits us to separate the quantization of both spaces.

To quantize our perturbation theory, we start with the classical Mukhanov-Sasaki action to apply the canonical quantization formalism [116, 194]. We study the dynamics given by the variation principle and show that up to the second order, this action is related to our perturbation theory. We then proceed with the canonical quantization formalism. However, we will see that in dynamical/curved spacetimes the solution to our equations of motion is not straightforward since we have a time-dependent mass. Moreover, there is an ambiguity of the vacuum in curved spacetimes. This ambiguity stems from the fact that the Stone-Von-Neuman theorem, which typically guarantees a unitary equivalence between different Schrödinger representations, doesn't hold in curved spacetimes [195, 196]. To face these issues, we employ the adiabatic vacuum approach, which involves using the WKB (Wentzel-Kramers-Brillouin) ansatz to propose solutions for the equation of motion. Finally, we will use these field solutions to compute the spectrum of the theory. From now on we are considering $\tilde{K} = 0$.[1]

B.1 Classical Mukhanov-Sasaki action

Our starting point involves revisiting the classical Einstein-Hilbert action for a perfect fluid

$$S = -\frac{1}{2\kappa}\int_M d^4x\sqrt{-\bar{g}}\bar{\mathcal{R}} + \int_M d^4x\sqrt{-\bar{g}}\bar{p}. \tag{B.1}$$

---

[1] Had we considered $\tilde{K} \neq 0$, the Mukhanov-Sasaki action would have extra terms and all of our calculations in the following sections would have to be adapted.

To derive the dynamic equations for our perturbation theory, we need the second-order variation of the Einstein-Hilbert action. However, directly deriving the perturbed action from this action is not simple. Instead, we will adopt the perturbed action previously derived in Ref. [39] in the case of a flat Universe and analyze its dynamics in the next section. We will also demonstrate that the curvature perturbation defined in Eq. (A.55) can also be deduced from the Mukhanov-Sasaki action, establishing the equivalence between these two theories.

B.1.1 Second Order Variation and Equation of Motion

The perturbations are described by the Einstein-Hilbert action expanded up to second order, resulting in the Mukhanov-Sasaki Lagrangian

$$L_{MS} = \int d^3\mathbf{x} \frac{1}{\kappa c} \left(\dot{\zeta}^2 z^2 + c_s^2 z^2 \zeta \Delta \zeta \right), \tag{B.2}$$

where $c_s^2 = \left(\frac{\partial \bar{\rho}}{\partial \bar{p}}\right)_S = w$ is the speed of sound, $\Delta = \bar{D}^2$ is the spatial Laplacian operator and

$$z^2 = \frac{c^2 \kappa \bar{a}^3 (\bar{\rho} + \bar{p})}{2 \bar{H}^2 c_s^2}. \tag{B.3}$$

To compute the density contrast in Eq. (4.10), we will need the curvature modes related to the Mukhanov-Sasaki Lagrangian. The extremization of Eq. (B.2) in Fourier space yields the equation of motion

$$\frac{\ddot{\zeta}_k}{\kappa c} + 2 \frac{\dot{z}}{z} \frac{\dot{\zeta}_k}{\kappa c} + \frac{k^2}{\bar{a}^2} \frac{\zeta_k}{\kappa c} = 0, \tag{B.4}$$

where k is the comoving wave number of the modes. Since obtaining analytical solutions to (B.4) is non-trivial, we will employ a numerical code to compute the modes, which requires the use of dimensionless variables. In dimensionless units, denoted with a subscript A, we redefine our variables as[2]

$$\begin{aligned} \zeta_{k_A,A} &\equiv \frac{\zeta_k}{\sqrt{8\pi l_p^2 R_H}}, && \text{(B.5)} \\ \Pi_{\zeta_k,A} &\equiv \frac{\Pi_{\zeta_k}\sqrt{R_H}}{\sqrt{8\pi l_p^2}}, && \text{(B.6)} \\ k_A &\equiv kR_H, && \text{(B.7)} \\ t_A &\equiv \frac{ct}{R_H}, && \text{(B.8)} \end{aligned}$$

where l_p is the Planck length, $R_H = \frac{c}{H_0}$ and

$$\Pi_\zeta \equiv \frac{\partial L_{MS}}{\partial(\dot{\zeta})} = \frac{2z^2\dot{\zeta}}{\kappa c}. \tag{B.9}$$

The equation of motion becomes

$$\begin{aligned} \dot{\Pi}_{k_A,A} + \frac{2c_s^2 z^2 k_A^2}{a^2}\zeta_{k_A,A} &= 0 \;\; or \\ \dot{\Pi}_{k_A,A} + 2\omega_{k_a}^2 z^2 \zeta_{k_A,A} &= 0. \end{aligned} \tag{B.10}$$

Here, $\omega_A^2 \equiv \frac{c_s^2 k_a^2}{\bar{a}^2}$ is the dimensionless frequency and the dimensionless conjugated momenta are given by

$$\Pi_{k_A,A} = 2z^2\dot{\zeta}_{k_A,A}. \tag{B.11}$$

Eq. (B.10) can be interpreted as a harmonic oscillator with a time-dependent frequency $\omega_{k_a}(t_a)$ and mass $m(t_a) = 2z^2$. From here on, we are going to use the dimensionless variables with the same notation as described above.

We expect that the classical and the quantum modes have the same equations of motion, and thus Eq. (B.4) is valid for both cases. Before we continue with the canonical quantization procedure, we prove that the definition in Eq. (A.55) is consistent with Eq. (B.12) to show that the Mukhanov-Sasaki action represents the same perturbation

---

[2] A derivative of a dimensionless variable will also be dimensionless. However, we use the same dot notation.

theory developed in App. A.4. Note that the Mukhanov-Sasaki variable is related to the curvature perturbation by

$$v \equiv -\zeta z \sqrt{\frac{2}{\kappa c}}. \tag{B.12}$$

We will prove the equivalence by rewriting Eq. (B.12) in terms of Bardeen variable Φ using Eq. (B.4) and the relation from [39]

$$\bar{D}^2 \Phi = -\frac{z^2 \bar{\Theta}}{3\bar{a}^3} \dot{\zeta} = -\frac{\bar{\Theta}}{6\bar{a}^3} \Pi_\zeta. \tag{B.13}$$

We will later show that the same relation can be achieved within the framework of App. A.

We start by writing the left side of Eq. (B.13) in the Fourier space so the spatial Laplacian operator can be replaced by $\bar{D}^2 = -\frac{\mathbf{k}^2}{\bar{a}^2}$. Explicitly,

$$-\int_{-\infty}^{\infty} \frac{\mathrm{d}^3 k}{(2\pi)^{\frac{3}{2}}} \bar{a} \mathbf{k}^2 \Phi_{\mathbf{k}}(\tau) e^{-i\mathbf{k}.\mathbf{x}} = \frac{z^2 \bar{\Theta}}{3} \int_{-\infty}^{\infty} \frac{\mathrm{d}^3 k}{(2\pi)^{\frac{3}{2}}} \left(\dot{\zeta}_{\mathbf{k}} e^{-i\mathbf{k}.\mathbf{x}} \right). \tag{B.14}$$

Dividing both sides by $\bar{\Theta}$ and taking the time derivative from both sides leads to

$$-\int_{-\infty}^{\infty} \frac{\mathrm{d}^3 k}{(2\pi)^{\frac{3}{2}}} \bar{a} \mathbf{k}^2 \left[\frac{\Phi_{\mathbf{k}}}{3} + \frac{\partial}{c\partial t} \left(\frac{\Phi_{\mathbf{k}}(\tau)}{\bar{\Theta}} \right) \right] e^{-i\mathbf{k}.\mathbf{x}} = \int_{-\infty}^{\infty} \frac{\mathrm{d}^3 k}{(2\pi)^{\frac{3}{2}}} \frac{2z\dot{z}}{3} \dot{\zeta}_{\mathbf{k}}(\tau) e^{-i\mathbf{k}.\mathbf{x}} +$$
$$\frac{cz^2}{3} \int_{-\infty}^{\infty} \frac{\mathrm{d}^3 k}{(2\pi)^{\frac{3}{2}}} \left(\ddot{\zeta}_{\mathbf{k}} e^{-i\mathbf{k}.\mathbf{x}} \right), \tag{B.15}$$

where we used $\dot{\bar{a}} = \frac{c\bar{a}\Theta}{3}$ in the first term. We can now use Eq. (B.4) and Eq. (B.12) such that

$$-\int_{-\infty}^{\infty} \frac{\mathrm{d}^3 k}{(2\pi)^{\frac{3}{2}}} \bar{a} \mathbf{k}^2 \left[c\frac{\Phi_{\mathbf{k}}}{3} + \frac{\partial}{c\partial t} \left(\frac{\Phi_{\mathbf{k}}(\tau)}{\bar{\Theta}} \right) \right] e^{-i\mathbf{k}.\mathbf{x}} = -\frac{c^2 z^2 c_s^2}{3\bar{a}^2} \int_{-\infty}^{\infty} \frac{\mathrm{d}^3 k}{(2\pi)^{\frac{3}{2}}} \left(\zeta_{\mathbf{k}} e^{-i\mathbf{k}.\mathbf{x}} \right). \tag{B.16}$$

Furthermore,if we go back to the coordinate space we have

$$\bar{D}^2 \left[\bar{a}^3 \frac{c\Phi}{3} + \frac{\bar{a}^3 \partial}{c\partial t} \left(\frac{\Phi}{\bar{\Theta}} \right) \right] = \frac{c^2}{3} z^2 c_s^2 \bar{D}^2 \zeta. \tag{B.17}$$

Finally,

$$\zeta = \frac{3\bar{a}^3}{c^2 z^2 c_s^2} \left[\frac{\partial}{c\partial t} \left(\frac{\Phi}{\bar{\Theta}} \right) + \frac{c}{3} \Phi \right]. \tag{B.18}$$

In the next section, we will show that this relation can be obtained in the perturbation theory context developed in App. A, and thus connect the perturbation theory to the Mukhanov-Sasaki Lagrangian.

B.1.2 Connection with Perturbation Theory

We need to prove that the definition in Eq. (A.55) is consistent with Eq. (B.18). We include the curvature factor $\bar{K}$ in this calculus to prove that the equivalence is general, even though we consider $\bar{K} = 0$. Starting with the relations in Eq. (A.41), we can compute

$$\begin{aligned}\dot{\Psi} + \frac{\bar{\Theta}}{3}\Phi &= \frac{\bar{\Theta}}{3}\phi + \dot{\psi} - \frac{\dot{\bar{\Theta}}\delta\sigma}{3} \\ &= \frac{\bar{\Theta}}{3}\phi + \dot{\psi} + \delta\sigma\left[\frac{\kappa}{2}(\bar{p}+\bar{\rho}) - \bar{K}\right], \end{aligned} \tag{B.19}$$

where we used the Friedmann equations in Eqs. (2.20) and (2.21). Combining Eq. (B.19) with Eq. (A.28), we get

$$\dot{\Psi} + \frac{\bar{\Theta}}{3}\Phi = \frac{\delta\bar{\Theta} - \bar{D}^2\delta\sigma}{3} + \delta\sigma\left[\frac{\kappa}{2}(\bar{p}+\bar{\rho}) - \bar{K}\right]. \tag{B.20}$$

Plugging this result into Eq. (A.25),

$$\frac{3\dot{\Psi} + \bar{\Theta}\Phi}{3} = \frac{\kappa}{2}(\bar{\rho}+\bar{p})\tilde{\mathcal{V}}. \tag{B.21}$$

In the above, we used the gauge invariant form in Eq. (A.48). Replacing $\tilde{\mathcal{V}}$ by the curvature perturbation variable in Eq. (A.55) gives

$$\frac{3\dot{\Psi} + \bar{\Theta}\Phi}{9} = \frac{\kappa}{2}(\bar{\rho}+\bar{p})\frac{1}{\bar{\Theta}}\left[\zeta - \Psi + \frac{2\bar{K}\Psi}{\kappa(\bar{\rho}+\bar{p})}\right]. \tag{B.22}$$

We can use Eq. (2.21) such that

$$\frac{3\dot{\Psi} + \bar{\Theta}\Phi}{9} = \frac{\kappa}{2}(\bar{\rho}+\bar{p})\frac{1}{\bar{\Theta}}\left[\zeta + \frac{2\dot{\bar{\Theta}}\Psi}{3\kappa(\bar{\rho}+\bar{p})}\right] \tag{B.23}$$

and finally

$$\zeta = \frac{3\bar{a}^3}{c^2 z^2 c_s^2}\left[\frac{\partial}{c\partial t}\left(\frac{\Phi}{\bar{\Theta}}\right) + \frac{c}{3}\Phi\right], \tag{B.24}$$

where we have used Eq. (B.3) in the last passage and also the fact that $\Phi = \Psi$ for a perfect fluid with no anisotropic pressure perturbation.

By Eqs. (B.18) and (B.24), we can see that the perturbed theory developed in App. A is in agreement with the second-order variation of the Einstein-Hilbert action which leads to the Mukhanov-Sasaki Lagrangian. Thus our previous results are still valid and we now start the quantization formalism by computing the conjugated momentum of the classical theory and then defining the commutation relations for the quantum operators from the Poisson brackets.

B.2 Commutation Relations

In Chapter 4, the background is characterized by a quantum contracting phase where an effective quantum fluid dominates near the bounce. Consequently, we seek quantized perturbations for consistency, although both theories can be viewed independently. In particular, the use of quantized perturbations evolving on classical backgrounds has been widely used since the results of Mukhanov and Chibisov [197] and Hawking [198] to derive the power spectrum of primordial perturbations, which are in turn used to describe the formation of structure on our universe.

The use of quantum fields to describe primordial perturbations means that said fields will have statistical properties. We may then partition the universe into local spatial regions and consider each one as a realization of a random process and compare its statistical properties with our theoretical predictions.[3]

We proceed to quantize our fields by promoting them to Hermitian operators that act on a Fock space. In terms of the usual Fourier mode expansion, our quantum operators, denoted with a hat superscript$\hat{}$from now on, are given by

$$\hat{\zeta}(\mathbf{x},t) = \frac{1}{(2\pi)^{\frac{3}{2}}} \int \mathrm{d}^3\mathbf{k} \left(e^{i\mathbf{k}.\mathbf{x}} \zeta_k^*(t) a_k + e^{-i\mathbf{k}.\mathbf{x}} \zeta_k(t) a_k^\dagger \right) \tag{B.25}$$

and

$$\hat{\Pi}_\zeta(\mathbf{x},t) = \frac{1}{(2\pi)^{\frac{3}{2}}} \int \mathrm{d}^3\mathbf{k} \left(e^{i\mathbf{k}.\mathbf{x}} \Pi_{\zeta_k}^*(t) a_k + e^{-i\mathbf{k}.\mathbf{x}} \Pi_{\zeta_k}(t) a_k^\dagger \right), \tag{B.26}$$

where $\mathbf{k}$ is the momentum vector with modulus k and a_k and $a_k^\dagger$ are the annihilation and creation operators respectively. Demanding that the quantum fields satisfy the canonical commutation relations

$$\left[\hat{\zeta}(\mathbf{x},t), \hat{\Pi}_\zeta(\mathbf{y},t)\right] = i\hbar\delta(\mathbf{x}-\mathbf{y})\,, \tag{B.27}$$

---

[3] Here, an important remark must be made: in inflationary models, one usually postulates this quantum to classical statistical connection, but the specific mechanism that converts a quantum universe to a classical one is still an open problem [27, 103]. In our model, even if the perturbative level predictions do not depend strongly on the bouncing mechanism if one assumes the De Broglie-Bohm interpretation applied to Canonical Quantum Gravity, this problem is automatically solved [99].

and that the complex modes satisfy

$$\dot{\zeta}_k \zeta_k^* - \zeta_k \dot{\zeta}_k^* = i\hbar\,, \tag{B.28}$$

$$\dot{\Pi}_k \Pi_k^* - \Pi_k \dot{\Pi}_k^* = i\hbar\,, \tag{B.29}$$

imply that the creation and annihilation operators $a_k, a_k^\dagger$ satisfy the usual creation and annihilation algebra

$$\left[a_{k_1}, a_{k_2}\right] = \left[a_{k_1}^\dagger, a_{k_2}^\dagger\right] = 0 \tag{B.30}$$

and

$$\left[a_{k_1}, a_{k_2}^\dagger\right] = \delta(k_1 - k_2). \tag{B.31}$$

Equation (B.28) represents the only constraint for choosing the basis of our problem and it is known as the vacuum normalization. To fully specify our operators and solve Eq. (B.10), we need to set initial conditions that will physically determine the annihilation operator and, consequently, the vacuum state of the theory.

B.3 Vacuum Choice Problem

In the conventional canonical quantization formalism, see Ref. [199] for an example, it is common to use the 4-dimensional Fourier basis to expand the field operators and find the solutions for the equation of motion in terms of annihilation and creation operators defined in this expansion. However, this approach involves making a specific choice for the vacuum state. By defining the annihilation operator as the one that annihilates the vacuum state, we implicitly select a particular vacuum state for the theory, which may not be the same had we chosen another base expansion. Let us first analyze how this vacuum choice is defined.

In usual quantum mechanics, the Stone-Von-Neuman theorem assures that for every Weyl representation of the canonical commutation relation for one degree of freedom, there is a unitary equivalence to the Schrödinger representation [195, 196]. Explicitly, for a separable Hilbert space $\mathbb{H}$, let $U(t)$ and $V(s)$ with $t, s \in \mathbb{R}$ be two continuous unitary groups on $\mathbb{H}$ satisfying the Weyl relations. Consider that there are closed sub-spaces $\mathbb{H}_l$ satisfying

$$\mathbb{H} = \sum_{l=1}^{\mathcal{N}} \mathbb{H}_l, \ \mathcal{N} \in \mathbb{Z}^+ \tag{B.32}$$

and

$$U(t) : \mathbb{H}_l \to \mathbb{H}_l, V(s) : \mathbb{H}_l \to \mathbb{H}_l \text{ for all } s, t \in \mathbb{R}. \tag{B.33}$$

The theorem states that for each l, there is an operator $T_l : \mathbb{H}_l \to L^2(\mathbb{R})$ such that $T_l U(t) T^{-1}$ represents a translation to the left by (t) and $T_l V(s) T_l^{-1}$ is a multiplication by e^{isx}.

The Stone-Von Neumann theorem implies that there is a Hilbert space isomorphism for operators U and V defined in different sub-spaces [200]. Thus there is an equivalence between all the base representations and therefore all the choices lead to the same vacuum of the theory in quantum mechanics (since we can always find the transformation T that connects the spaces). When it comes to quantum field theory, we have infinity degrees of freedom ($\mathcal{N} \to \infty$), as there is particle creation and annihilation for infinity energy numbers, and thus this affirmation is not true. Each base choice can lead to a different vacuum of the theory and we can at most define a set of equivalent solutions that leads to the same vacuum, but we cannot connect all the base choices.

The problem can be seen from the local nature of quantum field theory. While it is not possible to define a field at infinity, we can approximate its behavior locally and introduce a cutoff for our calculations. This local characteristic enables us to perform computations in quantum field theory and demonstrate that effects beyond the cutoff have negligible influence on local physics. However, this approach poses a challenge when considering global effects, as there is no universally applicable method for defining them across all local representations [200]. Consequently, the vacuum choice problem emerges. This is also closely connected to the challenge of defining a theory of Quantum Gravity since there is no preferred approach for defining the vacuum state in such a theory.

In non-curved spacetimes, the most used approach to solve this problem is to use the Minkowski vacuum. The choice of vacuum for a curved quantum field theory is extended analyzed in Ref. [201]. The authors propose criteria to define the vacuum such that it is constant in time and such that its basis expansion only differs by a phase from the solutions from the equations of motion. This criterion is used to define the temporal part of the basis that does not necessarily follow the equations of motion of the problem. In this work, we will use the adiabatic vacuum approach [204, 194], which we explain next.

B.4 Adiabatic Vacuum

In quantum field theory, due to the non-applicability of the Stone-Von Neumann Theorem, different choices of Hilbert Space that are consistent with the commutation relations in Eq. (B.27) are not unitary equivalent, which means that they lead to different physical predictions [203]. Therefore, one also needs a prescription to construct the associated Hilbert Space of a quantum field theory.

Although in usual Minkowski space-time one may use its symmetries to define the Hilbert space, in curved spacetimes one needs other techniques to construct said space [202]. This problem can be mapped to a choice of operators $\hat{a}_k$ that annihilate the vacuum state $|0\rangle$, which in turn depends on the choice of mode functions $\zeta_k(t)$ that satisfy the normalization condition in Eq. (B.28) [118]. Since such a condition is preserved by the time evolution, it suffices to choose a set of initial conditions $\{\zeta_k(t_0), \Pi_k(t_0)\}$ at an initial time t_0 [201].

A widely used prescription to define such a vacuum state is known as the adiabatic vacuum prescription [204, 118]. Its main idea is to set initial conditions by demanding that the mode functions $\zeta_k(t)$ coincide with their $\mathcal{N}$ order adiabatic approximation, ${}^{(\mathcal{N})}\zeta_k(t)$. This is implemented if one chooses the initial conditions [204]

$$ {}^{(\mathcal{N})}\zeta_k(t_0) = \frac{1}{\sqrt{\omega_k(t_0)}} e^{i\alpha_k(t_0)} \text{ and } {}^{(\mathcal{N})}\dot{\zeta}_k(t_0) = i\omega_k(t_0)\zeta_k(t_0), \tag{B.34} $$

where $\omega_k^2 = \frac{c_s^2 k^2}{\bar{a}^2}$ in our case and α_k is defined by the relation

$$ {}^{(\mathcal{N})}\zeta_k(t_0)\dot{\alpha}_k(t_0) = 1. \tag{B.35} $$

Using these initial conditions for the modes, a well-defined vacuum state is obtained, which is in turn used to construct the Fock Space by successive applications of the creation operator $\hat{a}_k^\dagger$ and their linear combinations [118]. Now that we have defined the appropriate initial conditions for quantization, we may see how to compute the perturbations' power spectrum.

B.5 Spectrum

The power spectrum of a theory plays a crucial role in understanding the formation of large-scale structures. The power spectrum also allows one to completely describe the statistics of the problem if the density field is described by Gaussian fluctuations [206]. To compute the power spectrum in our cosmological model, we first calculate the two-point correlation function of the field variable $\zeta(\mathbf{x}, t_i)$ at an initial time t_i. This function measures the spatial correlation between fluctuations at different points in space. The correlation function is expressed as an integral over Fourier modes, yielding the desired spatial correlation, and takes the form

$$\begin{aligned} \left\langle \hat{\zeta}(\mathbf{x}, t_i)\hat{\zeta}(\mathbf{y}, t_i) \right\rangle &= \frac{1}{(2\pi)^{\frac{2}{3}}} \int \mathrm{d}^3\mathbf{k} \left[|\zeta_k(t_i)|^2 e^{-i\mathbf{k}(\mathbf{x}-\mathbf{y})} \right] \\ &= \int \frac{\mathrm{d}k}{k} \left[P_\zeta(k) \frac{\sin kR}{kR} \right], \end{aligned} \tag{B.36}$$

where $R = |\mathbf{x} - \mathbf{y}|$ and we performed an integral over solid angles in the last equality. In the above, the power spectrum P_ζ is defined as[4]

$$P_\zeta(k) \equiv \frac{k^3 |\zeta_k(t_i)|^2}{2\pi^2}. \tag{B.37}$$

---

[4] Note that the power spectrum is already dimensionless, and thus we can use the same definition with our dimensionless variables as well.

APÊNDICE C – EINSTEIN EQUATIONS FOR SPHERICAL COLLAPSE

In this Appendix, we are interested in solving the Einstein equations for a metric representing the spherical collapse. We intend to use these solutions to compute the critical threshold in Sec. 4.5. We adapted the calculations initially done in Ref. [207] and corrected by Ref. [23]. In these references, a scalar field is used as the source for the energy-momentum tensor. We shall now perform the same calculations for a pressureless barotropic fluid.

An inhomogeneous but spherically symmetric space can be represented by the metric[1]

$$\mathrm{d}s^2 = -c^2\mathrm{d}t^2 + e^{-2\Lambda(t,r)}\mathrm{d}r^2 + R^2(t,r)\left(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\varphi^2\right), \tag{C.1}$$

where $\Lambda(r,t)$ and $R(t,r)$ are functions of the local coordinates to be defined and we are considering spherical coordinates. Since none of these functions do not depend on the angular variables, this metric may represent any type of spherical collapse. To find the local functions that describe our problem, we need to solve Einstein's equations for an energy-momentum tensor that corresponds to our symmetries. We shall represent $\partial_r \equiv {}'$.

Following [23], the Einstein's tensor components are

$$G_{tt} = \frac{1}{R^2}\left[1 + \dot{R}^2 - 2\dot{\Lambda}\dot{R}R - Re^{2\Lambda}\left(2\Lambda' R' + 2R'' + \frac{R'^2}{R}\right)\right] \tag{C.2}$$

$$G_{tr} = -\frac{2}{R}\left(\dot{R}' + \dot{\Lambda}R'\right) \tag{C.3}$$

$$G_{rr} = \frac{1}{R^2}\left[R'^2 - e^{-2\Lambda}\left(\dot{R}^2 + 2R\ddot{R} + 1\right)\right] \tag{C.4}$$

$$G_{\theta\theta} = \sin^{-2}\theta G_{\varphi\varphi} = R\left(\dot{R}\dot{\Lambda} + \Lambda' R' e^{2\Lambda} + R'' e^{2\Lambda} - \ddot{R} + \ddot{\Lambda}R - R\dot{\Lambda}^2\right). \tag{C.5}$$

We want to consider an isotropic and inhomogeneous barotropic fluid. Thus our energy-momentum tensor will have the form of a perfect fluid

$$T_{\mu\nu} = (\rho(r,t) + p(r,t))u_\mu u_\nu + p(r,t)g_{\mu\nu} \tag{C.6}$$

---

[1] This metric can also be written with a lapse function that shifts the time component. However, one can always redefine the coordinate to incorporate this quantity.

where $u_\mu u^\mu = -1$ is the velocity vector of the fluid, which we assume to be orthogonal to the spatial hypersurfaces. It is worth mentioning that in Eq. (C.6), both the pressure and the energy density depend on the spatial radius since we want to consider a physical metric modeled as a perturbation around the background quantities. Also, the equation of state parameter of the fluid is

$$w = \frac{p}{\rho} \tag{C.7}$$

and we are considering $w \ll 1$ for cold dark matter. In this context, the energy-momentum tensor components are

$$T_{uu} = \rho(t,r) \tag{C.8}$$

$$T^r{}_r = T^\theta{}_\theta = T^\varphi{}_\varphi = p(t,r) \tag{C.9}$$

$$T^r{}_u = 0, \tag{C.10}$$

$$T_{ru} = 0, \tag{C.11}$$

where the indexes u and r indicate both the fluid's velocity direction and the radial direction respectively. Let us now move to the Einstein's equations (EE).

We can redefine our variables as

$$k(t,r) = 1 - R'^2 e^{2\Lambda}, \quad m(t,r) = \frac{R}{2}\left(\dot{R}^2 + k\right), \tag{C.12}$$

such that the EE will be given by

$$k' = \frac{8\pi G}{c^2} R R' \left(T_{uu} + T^r{}_r\right) + 2R'\left(\ddot{R} + \dot{\Lambda}\dot{R}\right), \tag{C.13}$$

$$\dot{k} = \frac{8\pi G}{c^2} R R' T^r{}_u, \tag{C.14}$$

$$m' = \frac{4\pi G}{c^2} R^2 R' T_{uu} - \frac{4\pi G}{c^4} R^2 \dot{R} T_{ru}, \tag{C.15}$$

$$\dot{m} = \frac{4\pi G}{c^2} R^2 R' T^r{}_u - \frac{4\pi G}{c^2} \dot{R} R^2 T^r{}_r. \tag{C.16}$$

Additionally, the energy conservation of $T^{\mu\nu}$ gives the constraint

$$\nabla_\mu T^{\mu\nu} = 0, \text{ or}$$

$$\nabla_\mu\left(\rho u^\mu u^\nu + p\gamma^{\mu\nu}\right) = \left(\dot{\rho} + \frac{3H}{c}(\rho + p)\right) u^\nu = 0. \tag{C.17}$$

Thus, our system is given by

$$\mathrm{d}s^2 = -c^2\mathrm{d}t^2 + \frac{(R')^2}{1-k}\mathrm{d}r^2 + R^2(t,r)\left(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\varphi^2\right) \tag{C.18}$$

with

$$k' = \frac{8\pi G}{c^2}RR'\left(\rho + p\right) + 2R'\left(\ddot{R} + \dot{\Lambda}\dot{R}\right), \tag{C.19}$$

$$\dot{k} = 0, \tag{C.20}$$

$$m' = \frac{4\pi G}{c^2}R^2R'\rho \tag{C.21}$$

$$\dot{m} = -\frac{4\pi G}{c^2}\dot{R}R^2p = -(4\pi G)\dot{R}R^2w\rho = 0. \tag{C.22}$$

plus Eqs. (C.7) and (C.17). One can replace Eqs. (C.22) and (C.20) into (C.19) using the definitions in (C.12) yielding the final equations

$$m' = \frac{4\pi G}{c^2}R^2R'\rho \tag{C.23}$$

$$(\dot{R})^2 = \frac{2m}{R} - k. \tag{C.24}$$

We can recognize in Eqs. (C.23) and (C.24) the Lemaitre-Tolman-Bondi (LTB) class of solutions [105]. One may check that the metric in Eq. (C.18) describes the Schwarshild metric if m is constant, the Einstein-de Sitter universe if $R = a(t)r$ and $k = 0$ and the closed Friedmann universe if $R = a(t)r$ and $k = r^2$. We want to find analytical solutions for the system without considering a specific type of these local functions. Luckily, the LTB class of solutions is one of the few cases where there is an analytical parametric solution to Einstein's field equations, given by

$$R(\theta, r) = \frac{2m}{k}\sin^2\left(\frac{\theta}{2}\right) \tag{C.25}$$

$$t(\theta, r) = t_1(r) + \frac{m}{k^{\frac{3}{2}}}\left(\theta - \pi - \sin\theta\right). \tag{C.26}$$

In the above, θ is a parameter in the range $[-2\pi, 0]$ and t_1 is an integration constant[2]. We still have to find m and k, which will require initial conditions. Let us first analyze m.

---

[2] Our definition of the parameter θ differs from Ref. [23] by $-\pi$ for simplification purposes.

From the right side of Eq. (C.23), we see that m is related to the mass density of a spherical volume. To analyze this quantity, we start by rewriting the energy density from the spherical collapse metric as

$$\rho(t,r) = \bar{\rho}(t)(1+\delta(t,r)), \tag{C.27}$$

such that δ is the density contrast $\delta = \frac{\rho(t,r)-\bar{\rho}(t)}{\bar{\rho}(t)}$ and $\bar{\rho}$ the background energy density. We consider that there is a small overdensity in our spherical system when compared to the background density. We want to study a top-hat spherical collapse so that outside the over-dense spherical shell characterized by a critical radius r_c, the energy density is approximately the background density, i.e., $(r > r_c) \longrightarrow \rho \approx \bar{\rho}$. We can add this information to the density contrast by making the redefinition

$$\delta(t,r) \rightarrow \delta(t,r)\Theta(r-r_c) \tag{C.28}$$

where Θ is the Heaviside function. In this context, Eq. (C.23) becomes

$$m(r) = \int_0^r dr_1 \, \frac{4\pi G R^2 R'}{c^2} \left(\bar{\rho} + \delta\Theta(r-r_c)\right). \tag{C.29}$$

We still have freedom in our metric to define the local function R, which can be fixed with initial conditions. The simplest choice would be that at initial collapse time t_{ini}, $R(t_{ini}, r) = r$, which implies that the 3-dimensional spherical shell is initially at rest[3]. This leads to

$$m(r) = \frac{4\pi r^3 G\bar{\rho}(t_{ini})}{3c^2}\left(1 + \frac{3}{r^3}\int_0^r dr_1 (r_1)^2 \delta(t_{ini}, r_1)\Theta(r-r_c)\right). \tag{C.30}$$

The integral represents the expected value of the density contrast inside a spherical region with radius r and m gives the total mass inside this region. To evaluate it, we need to assume that at the initial time, our spherical metric possesses the same symmetries as the background and thus it is homogeneous. This means that the density contrast has a uniform distribution and does not depend on position at t_{ini}[4]. Consequently,

$$\delta(t_{ini}, r) = \delta(t_{ini}) \equiv \delta_{ini}. \tag{C.31}$$

---

[3] Note that any other initial condition is valid and sufficient to completely define m.

[4] This is only true at t_{ini}. At later times, the perturbation evolves and depends on the radial position.

We are representing any variable V at initial time t_{ini} as $V(t_{ini}) \equiv V_{ini}$ to simplify the notation. Plugging Eq. (C.31) in (C.30) leads to

$$m(r) = \begin{cases} M\left(\frac{r^3}{r_c^3}\right) & r \leq r_c \\ M + \frac{M}{1+\delta_{ini}}\left(\frac{r^3}{r_\mathrm{c}^3} - 1\right) & r > r_c \end{cases}, \tag{C.32}$$

such that we rewrote the terms as a function of

$$M = m(r_c) = \frac{4\pi G\bar{\rho}_{ini} r_c^3}{3c^2}\left(1 + \delta_{ini}\right). \tag{C.33}$$

We may now compute k using the second relation in Eq. (C.12), which requires us to evaluate $\dot{R}(t,r)$. To do so, we define the inhomogenous Hubble function

$$H \equiv \frac{\dot{R}(t,r)}{R(t,r)}. \tag{C.34}$$

Since we have the vanishing relations in Eqs. (C.22) and (C.20), R is the only local degree of freedom that evolves in time and provides the dynamics of the universe according to Eq. (C.25). Hence it is natural that we use this quantity to define the Hubble function. Assuming that the local metric is homogeneous at t_{ini},

$$H^2(t_{ini}, r) = \bar{H}^2(t_{ini}) = \frac{\kappa\bar{\rho}c^2}{3} = \frac{2M}{1+\delta_{ini}}\frac{1}{r_c^3}. \tag{C.35}$$

Thus, we can use this relation to compute k for $t = t_{ini}$ since this variable does not evolve in time, which leads to

$$k(r) = \frac{2m}{r} - r^2 H_{ini}^2 = \begin{cases} 2M\frac{\delta_{ini}}{1+\delta_{ini}}\frac{r^2}{r_\mathrm{c}^3} & r \leq r_c \\ 2M\frac{\delta_\mathrm{ini}}{1+\delta_{ini}}\frac{1}{r} & r > r_c \end{cases}. \tag{C.36}$$

Now that we have properly computed our local quantities, we can rewrite our solutions in Eq. (C.25) and (C.26) for $r \leq r_c$ as

$$R(\theta, r) = \frac{r(1+\delta_{ini})}{\delta_{ini}}\sin^2\left(\frac{\theta}{2}\right), \tag{C.37}$$

$$\Delta t(\theta, r) = \frac{1+\delta_{ini}}{2H_{ini}\delta^{\frac{3}{2}}(t_{ini})}\left(\theta - \pi - \sin\theta\right), \tag{C.38}$$

such that $\Delta t = t_i - t_j$ for arbitrary times and θ_{ini} is obtained from Eq. (C.25) at t_{ini}, i.e,

$$\sin^2\left(\frac{\theta_{ini}}{2}\right) = \frac{\delta_{ini}}{1+\delta_{ini}}. \tag{C.39}$$

We are now able to use these solutions to compute the critical value for the density contrast.