

**UNIVERSITÀ DEGLI STUDI DI MILANO**

FACOLTA' DI SCIENZE E TECNOLOGIE

Dipartimento di Fisica

Corso di Laurea Magistrale in Fisica

**Primordial Black Holes and Large Scale Structure**

Relatore: Prof. Antonio Riotto

Correlatore: Prof. Emanuele Castorina

Tesi di Laurea di:  
Emanuele Francesco  
Besana  
Matr. 08737A

Anno Accademico 2022 / 2023

# Contents

<b>Acknowledgements</b>	<b>1</b>
<b>Riassunto</b>	<b>2</b>
<b>Summary</b>	<b>3</b>
<b>1 Introduction</b>	<b>4</b>
<b>2 Elements of Cosmology</b>	<b>6</b>
2.1 Our Universe . . . . .	7
2.1.1 The Geometry . . . . .	8
2.1.2 The Energy Content . . . . .	10
2.1.3 The Dynamics . . . . .	11
2.1.4 The Parameters . . . . .	14
2.1.5 The problems . . . . .	16
2.2 Cosmological Perturbation Theory . . . . .	17
2.2.1 The Gauge Philosophy . . . . .	17
2.2.2 Perturbations in the energy-momentum . . . . .	22
2.2.3 Spacetime perturbations and the Einstein equations . . . . .	26
2.2.4 The evolution of matter perturbations . . . . .	27
2.2.5 Statistical Properties . . . . .	29
2.2.6 Particle Sampling . . . . .	31
2.3 Numerical Results and Fits For The Linear Theory . . . . .	32
<b>3 Nonlinear Clustering</b>	<b>34</b>
3.1 The Spherical Collapse Model . . . . .	34
3.2 The Press-Schechter Formalism . . . . .	37
3.3 Excursion Set Theory . . . . .	40
3.4 Halo Formation In the Excursion Set Theory . . . . .	42
3.4.1 The Conditional Mass Function . . . . .	43
3.4.2 Halo Accretion Rates . . . . .	44
3.5 Quasi-Linear and Non-Linear scaling relations . . . . .	46

<b>4 Primordial Black Holes</b>	<b>51</b>
4.1 Formation of Primordial Black Holes . . . . .	52
4.1.1 Collapse of Inflationary Perturbations . . . . .	52
4.1.2 Threshold Values . . . . .	54
4.1.3 Near Critical Collapse . . . . .	54
4.1.4 PBH Formation During Matter Domination . . . . .	54
4.2 Clustering of PBHs . . . . .	54
4.2.1 The Evolution of Clustering . . . . .	55
4.3 PBHs in structure formation . . . . .	56
4.4 Abundance of PBHs . . . . .	57
4.5 Generating Primordial Perturbations for PBHs in Inflation . . . . .	58
4.5.1 Inflation . . . . .	58
4.5.2 Inflation Models for PBH production . . . . .	59
4.6 Observational Constraints on Non-Evaporated PBHs . . . . .	61
4.6.1 Evaporation Constraints . . . . .	62
4.6.2 Gravitational Lensing Constraints . . . . .	62
4.6.3 Dynamical Constraints . . . . .	62
4.6.4 Accretion Constraints . . . . .	63
4.6.5 Gravitational Wave Constraints . . . . .	63
<b>5 Testing the <math>\Lambda</math>CDM Model</b>	<b>65</b>
5.1 Accelerating Structure Formation . . . . .	68
5.1.1 The Accretion of Dark Matter Halos . . . . .	71
5.1.2 Accretion by Primordial Black Holes . . . . .	76
5.2 The NANOGrav Gravitational Wave Signal and Some Perspectives . . . . .	78
<b>6 Conclusions and outlook</b>	<b>80</b>
<b>References</b>	<b>81</b>



-Mark Z. Danielewski, *House of Leaves*  
*The finest act of seeing is necessarily the  
act of not seeing something else.*

*The truth knocks on the door and you say,  
'Go away, I'm looking for the truth',  
and so it goes away. Puzzling.*

- Robert M. Pirsig, *Zen and the Art of Motorcycle Maintenance*



## Acknowledgements

This work was carried out in Geneva under the supervision of prof. Toni Riotto, whom I have to thank for giving me this great opportunity, and for showing me what real research is. Grazie.

Thank you to all my friends in Geneva. They all have played a role during my stay.

Thank you to Rudi, the only other theoretical physicist I knew, for all the incredible discussions in UniMail, which ranged from elephant meat to spinor bundles. The amount of caffeine we consumed in that building is probably a national record. Your vast computational knowledge has been fundamental and I owe the completion of this thesis to you as well.

A big thank you goes to Mai for the heartfelt discussions, her vitality and the gin tonics (which I did not consume) at Rond Pont, to Ella for her colorfulness and organizational powers and for being a wonderful neighbor. Thank you to the power couple Matyas & Ben, Toby, Reza, Joon, Benny, Luca, Valeria, Sophy, Liza, Lina, Jemima and Tara.

Thank you Соња, for everything.

Grazie anche a coloro che sono stati forgiati nel mio stesso fuoco, la Porcellina, Saddam, la Zebrina, Cappe, Vitö, Ghe. Grazie anche a Riki e Gio, ricordando le due nottate allucinanti a Ginevra.

Infine, grazie alla mia famiglia, mamma, papà e Cri, per esserci stati sempre per me, nonostante la distanza.

## Riassunto

A seguito del lancio del James Webb Space Telescope (JWST), è stata individuata una possibile collezione di galassie sorprendentemente massive e a redshift molto alto.

Questi sistemi forniscono un modo conveniente per verificare una proprietà fondamentale del modello  $\Lambda$ CDM: il contenuto stellare degli aloni di materia oscura non deve eccedere il materiale barionico disponibile in tali aloni. Risulta che la formazione prematura di galassie massive è molto difficile da rendere compatibile con il modello  $\Lambda$ CDM, richiedendo infatti un'elevata efficienza di formazione stellare, purtroppo non osservata nell'Universo a basso redshift [1].

Possibili interpretazioni di questi oggetti potrebbero riguardare semplicemente un'alta efficienza di formazione stellare, oppure il fatto che le proprietà degli oggetti massivi, come redshift fotometrico e massa stellare, sono ancora soggette a diverse fonti di incertezza sperimentale. Nel caso in cui nessuna di queste due spiegazioni reggesse, queste galassie rappresenterebbero una difficile sfida nei confronti del modello cosmologico  $\Lambda$ CDM.

In questo lavoro ci proponiamo di studiare una possibile estensione del  $\Lambda$ CDM, uno scenario in cui una parte della materia oscura è composta da buchi neri primordiali (PBHs) che, se aggiunti alla teoria standard, forniscono un ausilio per la formazione di strutture a larga scala e ad alto redshift.

Per esplorare le eventuali implicazioni di questi buchi neri, prendiamo in considerazione i due modi mediante cui essi influenzano la formazione di strutture, in modo da comprendere le galassie individuate da Labbè *et al.* [2].

In primo luogo, studiamo l'effetto Poissoniano dei PBHs, in cui le fluttuazioni numeriche dei buchi neri provocano un rumore bianco supplementare al power spectrum. Utilizziamo una funzione di massa monocromatica per i PBHs, in modo da specificare completamente il modello solamente con la frazione  $f_{\text{PBH}}$  di materia oscura sotto forma di PBHs e con la loro massa  $m_{\text{PBH}}$ . Trattando queste quantità come parametri, seguiamo [3] e consideriamo perturbazioni di isocurvatura sovrapposte al power spectrum adiabatico. Gli autori, nelle loro analisi, utilizzano il formalismo Press-Schechter, scoprendo che, per spiegare le osservazioni, sono necessari PBHs relativamente massivi, che richiederebbero la generazione di perturbazioni molto grandi da parte dell'inflazione. Ci chiediamo quindi se i vincoli sui parametri trovati in questo articolo possano essere rilassati utilizzando metodi più avanzati per lo studio di formazione delle galassie, come l'extended Press-Schechter (excursion set theory), che ci permette di studiare le genealogie degli aloni in maniera accurata.

Inoltre, in parte motivati dalla recente scoperta di AGN ad alto redshift contenenti buchi neri supermassivi, esaminiamo un modello in cui PBHs molto massivi e rarefatti, con una funzione di massa piccata, funzionano come semi per la formazione di galassie, che in questo caso verrebbero create per accrescimento della materia oscura circostante.

Infine, riflettiamo su come eventuali progressi nella comprensione della fisica dei PBHs, uniti ai dati del JWST e a quelli nuovi sulle onde gravitazionali pubblicati da NANOGrav [4], si riveleranno di fondamentale importanza per risalire alla natura della materia oscura.

## Summary

With the successful launch of the James Webb Space Telescope (JWST), a population of surprisingly massive galaxy candidates has been discovered at high redshifts.

These systems provide a convenient way to test a fundamental property of the  $\Lambda$ CDM model: the stellar content of dark matter halos should not exceed the available baryonic material in those halos. It is found that such early formation of massive galaxies is very difficult to reconcile with the  $\Lambda$ CDM model, demanding a high star formation efficiency, which is never seen in the low-redshift universe. In fact, only if all available baryons in all halos with enough baryons to form those galaxies are converted into stars, rather unrealistically, can the JWST data be explained [1].

Possible explanations for these objects could either be simply a high star formation efficiency, or the fact that the massive objects' properties, like photometric redshift and stellar mass, are still subject to several sources of observational uncertainty. If neither of these explanations hold up, these massive galaxies will pose a serious challenge to the  $\Lambda$ CDM cosmological model.

In this work, we aim to study a possible extension of  $\Lambda$ CDM, a scenario where a fraction of dark matter is composed of a population of primordial black holes (PBHs) that, if added to the standard background, would provide the necessary aid in the formation of high redshift large scale structures.

In order to explore the implications of these black holes we consider the two ways in which they influence structure formation to explain the massive galaxies found by Labbè *et al.* [2].

Firstly, we study the Poisson effect, where the PBHs number count fluctuations create an additional white noise to the power spectrum. We use a monochromatic mass function for PBHs, so that our model is specified completely by the fraction  $f_{\text{PBH}}$  of dark matter in the form of PBHs and their mass  $m_{\text{PBH}}$ . Treating these quantities as parameters, we follow [3] and consider isocurvature perturbations superposed to the standard adiabatic power spectrum. The authors use the Press-Schechter formalism in their analysis, and find that, to explain the observations, relatively massive PBHs are needed, which would require very large perturbations out of inflation. We thus question if the constraints on the parameters found in the said article can be relaxed by utilizing more advanced methods of galaxy formation, such as the extended Press-Schechter model (excursion set theory), through which halo histories can be thoroughly studied.

Furthermore, partly motivated by the recent discovery of very early AGNs containing supermassive black holes, we study a model in which heavy and rarefied PBHs with a peaked mass function can work as seeds for the formation of galaxies, which would be created via accretion of the surrounding dark matter.

Finally, we speculate how advancements in the understanding of PBHs physics, coupled to the JWST and new NANOGrav gravitational waves data [4], will prove to be fundamental in determining the true nature of dark matter.

# 1 Introduction

The synthesis of General Relativity and Quantum Field Theory, coupled to observational measurements, gives rise to the so called  $\Lambda$ CDM model. It details the evolution of our Universe and makes testable predictions [5].

According to it, the Universe starts out with a period of rapid expansion called inflation, to which follows the evolution determined by matter (baryons and dark matter), dark energy and radiation (mostly photons), the only components contained in spacetime. At large scales, the Universe looks rather homogeneous and isotropic, whereas small scales reveal the nonlinear structures (galaxies, clusters etc...) that were supposedly seeded by primordial quantum field fluctuations during inflation.

The onset of structure formation sees baryons falling into the potential wells created by dark matter which, decoupling early on, had a head start in the evolution of its over(under)-densities. Although the subsequent formation of the specific structures themselves is still under way of being understood, their large scale properties (number, mass and luminosity densities etc...) and the approximate redshift range in which they form can be calculated.

Generally, while the  $\Lambda$ CDM model is not free of minor shortcomings [6–13], it could, until now, describe our Universe to a high degree of precision.

The year 2021 saw the launch of the James Webb Space Telescope (JWST). Its temperature, collecting area and infrared wavelength sensitivity allow for the penetration of dust clouds in order to perform very deep observations up to redshifts  $z \gtrsim 15$ , right around the time structures, like active galactic nuclei and galaxies, were just taking form. The infancy and evolution of these objects is now finally possible to explore. Very soon after its launch, the telescope found heavy galaxy and AGN candidates (a lot of which were later confirmed) at high redshifts [2, 14–21], which were quickly pointed out by the community to be in tension with current  $\Lambda$ CDM predictions [1]. In particular, in this work we focus on the candidates found by Labb   et al. [2], consisting of 13 objects with redshifts in the range  $6.5 < z < 9.1$  and masses  $10^9 \lesssim M/M_{\odot} \lesssim 10^{11}$ . The team reports a stellar mass density at redshifts  $z \sim 8 - 9$  that, if confirmed, would result in a stark contrast with predictions based on the parameters of  $\Lambda$ CDM. One immediate resolution of this conundrum requires simply a higher than observed stellar formation efficiency. On the other hand, multiple new models are mostly based on different forms of dark matter, dark energy or on exotic inflationary mechanisms [22–30]. The focus of this work is on the possibility of primordial black holes constituting a fraction of the dark matter [3, 26, 31–33].

Primordial black holes (PBHs) are black holes that are not of astrophysical origin, but could be formed out of different mechanisms, for instance from the collapse of sizeable density perturbations that re-enter the horizon after inflation. They were first proposed by Hawking [34] but did not gain much traction until the LIGO discovery of a black hole binary merger in 2015 [35], which was proposed to be of primordial origin, given the largeness of the masses involved [36–38].

Although still not detected, PBHs are an interesting dark matter candidate, since they have the power of explaining a large variety of cosmological phenomena [39]. In particular, it should be stressed that even their non-existence would prove to be helpful, as it would constrain the possible cosmological inflationary model that gives rise to them.

In this work, in order to study scenarios that give rise to the JWST data points, we leverage the prowess of PBHs in aiding structure formation through two different effects. Firstly, we study the Poisson effect, which in general calls for lighter but denser PBHs that induce a Poisson white noise in the power spectrum, and, using excursion set theory, we follow the history of dark matter haloes in their mergers and accretion over time until  $z \sim 10$ . Moreover, we consider the seed effect, which requires instead heavier and more rarefied black holes that accrete the surrounding matter to form large structures. Interestingly, the latter case is also supported by the high redshift AGNs found by JWST [40–45], that reveal black holes too old and massive to have formed after their host galaxy, possibly then demanding a primordial origin.

Finally, and rather excitingly, in the midst of the writing of this thesis the NANOGrav collaboration (among other collaborations) confirmed a previous result announcing a stochastic gravitational wave background in the nHz region [4, 46–48]. Since the standard interpretation calls for a population of supermassive black hole binaries, we give some perspectives as to whether this could in part be explained by PBHs.

This thesis is organised as follows. In chapter 2 we give an overview of the standard cosmological template that consists of the  $\Lambda$ CDM model and its first order density perturbations. In the third chapter we follow these overdensities to their non-linear regime to develop the theory of excursion sets, which allows to follow closely the history of halos. In chapter 4 the concept of primordial black holes is introduced, along with the various measurements constraining them. Finally, chapter 5 is fully dedicated to our results and considerations, which also look toward the future of PBHs physics.

## 2 Elements of Cosmology

Modern cosmology is based upon the so called *cosmological principle*, the notion that, when coarse grained, the Universe is spatially homogeneous and isotropic, which follows from combining the observational result that galaxies are distributed isotropically on large enough distances, and the supposition that we don't live in a special place in space. All of this has to be supplemented with the understanding that the dynamical evolution of such a Universe is governed by Einstein's theory of General Relativity (GR).

Current popular cosmologies usually consider a Universe with four main components. The first, and the least relevant, is the "ordinary" baryonic matter, the protons, neutrons and electrons, the stuff we can actually see but makes up only 5% of the Universe; another component is the radiation, which is mostly made of photons and neutrinos.

However, the majority of matter is in the form of *dark matter* (DM), for which first evidence was found by Zwicky in 1933 when studying galaxies in the Coma cluster, and was later supported in the 1970s by the measurement of rotation speed of hydrogen gas in the outer legs of galaxies. Today, the strongest evidence of dark matter comes from the gravitational lensing of the *cosmic microwave background* (CMB): as the photons travel through the Universe, the intervening large scale structure curves their paths, giving rise to the cold and hot spots in the CMB. Since the degree of lensing depends on the total amount of matter, we can compare this with the abundance of light elements (baryonic) created via *Big Bang nucleosynthesis*. This is the way we know that the majority of the mass must be non luminiscent. However, despite this knowledge, the true small scale nature of dark matter remains an open question in physics.

The final component consists of what is usually called *dark energy*. By observing distant supernovae explosions, cosmologists were able to deduce that the Universe was decelerating at early times, but on the other hand accelerating at very recent times (from a few billion years ago). Such accelerated expansion is only possible in GR if the Universe is filled with an energy density that doesn't dilute and exerts negative pressure. A candidate for this dark energy is the energy density of empty space, the *cosmological constant*, since naturally it doesn't dilute with the expansion of spacetime. Again, the nature of this component is largely mysterious, despite it contributing to 70% of the overall energy in the Universe today.

Summarizing, the most popular model today is the so called  $\Lambda$ CDM model, a flat Universe in which  $\sim 75\%$  of the energy is due to the cosmological constant (hence the  $\Lambda$  in the name),  $\sim 21\%$  is due to "cold dark matter" (CDM), and the remaining 4% is the luminous matter out of which stars and galaxies are made. The Universe starts with a Hot Big Bang, then expands (but decelerates) at early times, with radiation domination followed by matter domination, and finally starts to accelerate approximately one billion years ago during dark energy domination.

If the cosmological principle held perfectly, if the distribution of matter in the Universe were homogeneous and isotropic, there would be no structures (inhomogeneities) now. We therefore clearly need a mechanism which gives rise to deviations from perfect uniformity, which cannot be found in the standard cosmological model.

It is naturally expected that a general relativistic description of the Universe breaks down at very early times when the Universe is so dense and hot that quantum effects become important; in fact, the standard cosmological model has a number of conceptual problems when applied to the early Universe. Solutions to these problems require extending the  $\Lambda$ CDM paradigm by introducing density perturbations by quantum fluctuations at early times, which are now believed to be responsible for the formation of all the large scale structure that we see. A popular theory that accomplishes the generation of those seeds is the *inflationary theory* in which the Universe is supposed to have gone through a rapid expansion over very little amount of time driven by one (or more) quantum field, the *inflaton*.

Our understanding of the Universe is still very far from complete and, despite the inflationary paradigm, we are still unable to predict the initial conditions for structure formation from first principles, forcing us to rely on a set of parameters taken from observational data.

Once the initial conditions have been specified, one can compute the way in which density perturbations evolve in time in a homogeneous background, as we will do below. When the universe is matter dominated, the perturbations will naturally grow, since a region whose initial density is slightly higher than the mean will attract its surroundings more strongly than average. Therefore overdense regions become more overdense over time, and underdense regions become more rarefied as matter will flow away from them.

The exact rate at which perturbations grow will depend on the cosmological model, and in particular this growth will stop when dark matter comes to dominate the Universe. In any case, once a certain region is overdense enough, it will stop growing and start to collapse, which is the start of the virialization of large objects. In general, as we will see, dark matter starts growing earlier than baryonic matter, and nonlinear, quasi equilibrium dark matter objects are called *dark matter halos*; they are the potential wells into which the baryons will eventually fall and form galaxies. Over time, these galaxies form clusters and superclusters until dark matter starts to dominate the Universe.

## 2.1 Our Universe

The picture we just laid out, that of an expanding Universe, implies that in the far past conditions were such that the environment was very hot and dense. Particles collided frequently and they were all in thermal equilibrium at a temperature  $T$ . A useful relation, which can be derived, relates the temperature of the Universe  $T$  to its age:

$$\frac{T}{1 \text{ MeV}} \approx \left( \frac{t}{1 \text{ s}} \right)^{-1/2} \quad (2.1)$$

This equation tells us that, for instance, the Universe had a temperature of about 1 MeV when it was just one second old. The rates of reactions were extremely high, and a lot of interesting phenomena took place in a short timeframe.

Above 100 GeV, all particles of the Standard Model were in equilibrium and therefore their abundances were roughly equal to one another. This can be taken as the initial condition for our Universe. In just  $10^{-9}$  seconds, the Universe expanded by a factor of  $10^4$  and the temperature dropped rapidly. During this short time the Universe went through successive evolutionary stages.

At around 100 GeV the electroweak (EW) symmetry of the Standard Model was broken in what is called the *EW phase transition*, where the weak and electromagnetic forces decoupled and particles acquired their mass. The detailed dynamics of such a transition are still an open subject of research, even though the basics of it were confirmed by the discovery of the Higgs boson.

As the temperature drops, one by one particles become less energetic, so that the particle-antiparticle annihilation is favored over the reverse process of particle creation. The first particles to disappear in this manner were the most massive quarks, followed shortly after by the massive gauge bosons W and Z, the tau lepton and the Higgs. Finally, at around 150 MeV, the remaining quarks condensed into hadrons, resulting in the *QCD phase transition*.

Particles can fall out of thermal equilibrium when their interaction rate drops below the expansion rate of the Universe. At that "moment" those particles will stop interacting with the environment and a relic abundance will be created. One of the most important decoupling events is the *neutrino decoupling* around one second after the Big Bang, which produced the *cosmic neutrino background* (CνB). During radiation domination, these neutrinos had a large impact on the evolution of the Universe, and their imprints have been recently detected in, among other objects, the clustering of galaxies.

About 1 minute after the Big Bang, the temperature dropped enough to allow the creation of the first light nuclei Helium-4 and Lithium-7. This Big Bang nucleosynthesis produced very few elements heavier than lithium, because there are no stable nuclei with 5 or 8 nucleons that would be required in order to sustain the reaction. Heavier elements instead formed in the insides of stars, via *stellar nucleosynthesis*, where densities allow for the creation of, for instance, carbon out of alpha particles via the  $3\alpha$  channel. Later, with the explosion of stars, these heavier elements, up until iron, spread throughout the Universe; instead, the origin of elements that lie beyond iron is still debated, but they most likely formed via *explosive nucleosynthesis*, either with s-processes during the last phases of stellar evolution, or with r-processes in supernovae explosions or neutron star mergers.

About 370000 years after the Big Bang the first atoms could form. This process, known as *recombination*, where the interaction cross section for the photons dropped enough for the electrons to be captured by the nuclei, produced what we now see as a relic photon background, the cosmic microwave background. The CMB contains temperature fluctuations that, although very small ( $\delta T/T \sim 10^{-5}$ ), contain very important information about the primordial Universe. These fluctuations are the same that produced the nonlinear structures (galaxies, clusters etc...) that we see in our Universe. In fact, a remarkable and stunning

property of the CMB fluctuations is that they span acausal distances. Therefore, either they were created some time before the Big Bang, or they are in contrast with the concept of causality. Indeed, as we've seen above, current theories believe that the Big Bang was not the beginning of time, but the fluctuations were produced during an earlier period of very rapid expansion called *inflation*. The small density fluctuations during inflation finally grew into the structures we see today.

### 2.1.1 The Geometry

A particularly simple description for the geometry of our Universe is obtained from the observation that space, on large scales, is isotropic and homogeneous. These kinds of three-spaces must have constant intrinsic curvature, either positive, negative or null. The metric for each of those cases is the *Friedmann-Lemaître-Robertson-Walker metric* (FLRW), which in polar coordinates reads:

$$ds^2 = -c^2 dt^2 + a^2(t) \left( \frac{dr^2}{1 - kr^2/R_0^2} + r^2 d\Omega^2 \right) \quad (2.2)$$

where  $k = 0, +1, -1$  for a flat, positively and negatively curved Universe respectively.

As we can see, the symmetries of the Universe rendered the metric in a very simple form, in particular the independent components of  $g_{\mu\nu}$  have been reduced only to a scale factor  $a(t)$  and a curvature scale  $R_0$ . We can immediately make some important comments regarding (2.2). First, we notice that it is invariant under the following scalings:

$$a \rightarrow \lambda a, \quad r \rightarrow r/\lambda, \quad R_0 \rightarrow R_0/\lambda \quad (2.3)$$

We can use this freedom to set the scale factor  $a(t)$  equal to one today, i.e.  $a(t_0) = 1$ . The coordinate  $r$ , the *comoving coordinate*, is not physically observable because it can be rescaled by (2.3), contrary to the *physical coordinate*  $r_{\text{phys}} = a(t)r$  on which physical results will depend; in fact, the physical velocity of a galaxy with trajectory  $\vec{r}(t)$  in comoving coordinates is:

$$\vec{v}_{\text{phys}} \equiv \frac{d\vec{r}_{\text{phys}}}{dt} = \frac{da}{dt} \vec{r} + a(t) \frac{d\vec{r}}{dt} \equiv H r_{\text{phys}} + \vec{v}_{\text{pec}} \quad (2.4)$$

where we have introduced the important variable:

$$H(t) \equiv \frac{\dot{a}(t)}{a(t)} \quad (2.5)$$

called the *Hubble parameter*, which controls the *Hubble flow* in the first term of (2.4), the motion of the galaxy due to the expansion of space between the origin and  $r_{\text{phys}} = a(t)r$ . The second term is the *peculiar velocity*, the velocity of the galaxy measured by an observer comoving with the Hubble flow. The local expansion rate of the Universe  $H_0$  is of fundamental importance since, for instance, it determines the age of the Universe; this is one of the reasons why it is important to measure it precisely. Historically, measurements of the Hubble constant used to come with very large uncertainties, so that it is conventional to define:

$$H_0 \equiv 100 h \text{ km s}^{-1} \text{ Mpc}^{-1} \quad (2.6)$$

In general, the  $H_0$  parameter can be estimated from the cosmological model utilizing early Universe measurements or from the local Universe. Interestingly, these two ways give different results. The latest supernovae measurements have given the following result [49]:

$$h = 0.730 \pm 0.010 \quad (2.7)$$

Whereas the measurements from the CMB anisotropy spectrum [5] give:

$$h = 0.674 \pm 0.005 \quad (2.8)$$

This discrepancy is known as the *Hubble tension*, and it is the most serious challenge to the  $\Lambda$ CDM model. The metric (2.2) can actually be rewritten in a more convenient form by redefining the radial coordinate  $d\chi = dr/\sqrt{1 - kr^2/R_0^2}$  such that:

$$ds^2 = -c^2 dt^2 + a^2(t)(d\chi^2 + S_k^2(\chi)d\Omega^2) \quad (2.9)$$

where:

$$S_k(\chi) = R_0 \begin{cases} \sinh \chi / R_0 & k = -1 \\ \chi / R_0 & k = 0 \\ \sin \chi / R_0 & k = +1 \end{cases} \quad (2.10)$$

Note that for a flat Universe ( $k = 0$ ) there is no distinction between the old radial coordinate  $r$  and the new one  $\chi$ . This reformulation is very useful since we can now introduce the concept of *conformal time*  $\eta$  by now changing time coordinate:

$$d\eta = \frac{dt}{a(t)} \quad (2.11)$$

so that we notice that (2.9) is conformally flat:

$$ds^2 = a^2(\eta)(-c^2 d\eta^2 + d\chi^2 + S_k^2(\chi)d\Omega^2) \quad (2.12)$$

Obviously, this form of the metric is particularly convenient for studying the propagation of light, since light cones will be tilted everywhere at 45 degrees, and it will also prove more useful to find the solutions to Einstein's equations.

An important concept in cosmology is the *cosmological redshift*, since what we know of the properties of the Universe is a result of the observation of distant objects. In order to analyze these correctly, we need to know how light propagates as the Universe evolves in time. Consider the scenario in which a galaxy that sends light signals to us positioned at  $(r_e, \theta_e, \phi_e)$ . Since the Universe is isotropic and homogeneous we can freely choose the propagation of light to follow a straight line  $\theta = \phi = \text{const}$  along null geodesics  $ds^2 = 0$ :

$$0 = a(\eta)(-c^2 d\eta^2 + d\chi^2) \rightarrow \Delta\chi(\eta) = \pm c \Delta\eta \quad (2.13)$$

where the  $\pm$  corresponds to outgoing and ingoing photons. If a wave crest is emitted at time  $t_e$  from the galaxy, the time  $t_0$  when it reaches us is given by:

$$c^2(\eta(t_0) - \eta(t_e)) = \chi(r_e) - \chi(0) = \chi(r_e) \quad (2.14)$$

Since the comoving distance of the galaxy does not change with time, a successive wave crest emitted shortly after, at  $t_e + \delta t_e$ , reaches the origin at time  $t_0 + \delta t_0$ :

$$c^2(\eta(t_0 + \delta t_0) - \eta(t_e + \delta t_e)) = \chi(r_e) \quad (2.15)$$

Combining these two equations and noticing that for real applications we can approximate  $\delta t_e \ll t_e$  and  $\delta t_0 \ll t_0$ , we find:

$$\frac{\delta t_0}{a(t_0)} = \frac{\delta t_e}{a(t_e)} \quad (2.16)$$

where we have used the definition of conformal time (2.11). Thus the period of the wave, hence its wavelength, increases proportionally to the scale factor:

$$\frac{\lambda_0}{\lambda_e} = \frac{v_e}{v_0} = \frac{\delta t_0}{\delta t_e} = \frac{1}{a(t_e)} \quad (2.17)$$

where we remember that we have the freedom to choose  $a(t_0) = 1$ . Defining the relative change of wavelength as the *redshift*  $z$ , we have:

$$a(t_e) = \frac{1}{1+z} \quad (2.18)$$

A galaxy at redshift  $z = 1$  thus emitted the observed light when the Universe was half its current size, a galaxy at redshift 2 when the Universe was one third its size, and so on.

### 2.1.2 The Energy Content

The symmetries of the Universe, isotropy and homogeneity, which constrain the form of the metric, also constrain the large scale form of the energy momentum tensor. In fact it takes the structure of a perfect fluid:

$$T_{\mu\nu} = \left( \rho + \frac{P}{c^2} \right) U_\mu U_\nu + P g_{\mu\nu} \quad (2.19)$$

where  $\rho c^2$  and  $P$  are the energy density and pressure in the rest frame of the fluid and  $U_\mu$  is its four-velocity relative to a comoving observer.

To see how these quantities evolve with time, we recall that energy and momentum need to be conserved, a requirement that is packaged in the *continuity equation*:

$$\nabla_\mu T^\mu_\nu = 0 \quad (2.20)$$

Using the Christoffel's symbols for the FLRW metric we find that this becomes:

$$\dot{\rho} + 3\frac{\dot{a}}{a} \left( \rho + \frac{P}{c^2} \right) = 0 \quad (2.21)$$

Naturally, energy conservation in Minkowski space follows from a requirement of time translation, which is not respected in an expanding Universe and therefore the familiar energy conservation is replaced by equation (2.21). Since most cosmological fluids are governed by a constant equation of state  $P/(\rho c^2) = w$ , we can simplify this equation to:

$$\frac{\dot{\rho}}{\rho} = -3(1+w)\frac{\dot{a}}{a} \quad (2.22)$$

so that:

$$\rho \propto a^{-3(1+w)} \quad (2.23)$$

This shows that the evolution of the energy density with time depends on the equation of state of the fluid in question. As we said in the introduction, the Universe is made of matter (dark matter and baryons), radiation and dark energy.

The word "matter" in cosmology in general refers to a fluid whose pressure is negligible compared to its energy density, thus having an equation of state with  $w = 0$ :

$$\rho_m \propto a^{-3} \quad (2.24)$$

which just says that, naturally, energy density in a region of space  $V$  dilutes because the volume itself increases as  $V \propto a^3$  while the energy of that region stays the same. In the Universe we know of two types of matter:

- *Baryons*: In cosmology we refer to "ordinary" matter, like protons and electrons, as baryons, although being technically incorrect.
- *Dark Matter*: Most of the matter in the Universe is in the form of a dark component. The precise (small scale) nature of dark matter is unknown, although at large scales we can treat it as a perfect fluid that weakly interacts with the other species of matter.  
There are many popular dark matter candidates such as WIMPs, axions, MACHOs and, importantly for this work, primordial black holes (PBHs).

The term "radiation", on the other hand, denotes anything for which the equation of state is  $P = \rho c^2/3$  or  $w = 1/3$  equivalently. This gives:

$$\rho_r \propto a^{-4} \quad (2.25)$$

In general there are various examples of radiation in our Universe:

- *Light Particles*: We know that at early times, when the Universe was very hot and dense, all particles of the Standard Model acted as radiation, since at that time their masses were negligible compared to the temperature. As the temperature dropped, particles started to behave like matter as their mass became relevant.

- *Photons*: Clearly, being massless, photons are always a form of radiation. In particular, they were the dominant form of energy during the nucleosynthesis of light elements.
- *Neutrinos*: Neutrinos, we now know, have a very small mass. They have therefore behaved as radiation for most part of the history of the Universe, and only recently have they started behaving like matter.

Finally, at present time the Universe is dominated by a form of dark energy with negative pressure  $P = -\rho c^2$  ( $w = -1$ ) or, in other words, a form of energy that does not dilute with the expansion:

$$\rho \propto a^0 \quad (2.26)$$

Candidates for this type of energy density are:

- *Vacuum energy*: As the Universe expands, more "space" is being created and this energy naturally increases proportionally to the volume, so it doesn't dilute. This energy of empty space, the so called vacuum energy, is actually predicted by quantum field theory as an energy momentum tensor of the form:

$$T_{\mu\nu}^{\text{vac}} = -\rho_{\text{vac}} c^2 g_{\mu\nu} \quad (2.27)$$

Unfortunately, the QFT prediction for the value of this vacuum energy is much greater than that inferred from observations.

- *Cosmological Constant*: A more interesting candidate is the cosmological constant from the Einstein equations:

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu} \quad (2.28)$$

It has now become standard practice to bring the  $\Lambda$  term to the right, treating it as a source of energy:

$$T_{\mu\nu}^{\Lambda} = -\frac{\Lambda c^4}{8\pi G} g_{\mu\nu} \quad (2.29)$$

which has the form of (2.27)

### 2.1.3 The Dynamics

So far, we have used the large scales symmetries of the Universe to determine the form of the metric (2.9), and the form of the energy momentum tensor (2.19). If we want to understand the time evolution, or equivalently find the form of the scale factor  $a(t)$ , we need to solve the Einstein equations for the metric, for the various components of energy density:

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu} \quad (2.30)$$

Setting  $G_0^0$  equal to  $-(8\pi G/c^4)\rho$ , we get the *Friedmann equation*:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2 R_0^2} \quad (2.31)$$

where  $\rho$  is to be understood as a sum of all the components:

$$\rho \equiv \sum_{i=m,r,\Lambda} \rho_i \quad (2.32)$$

where we have renamed (2.26) as  $\rho_{\Lambda}$ . This equation is very important, as it describes the evolution of the Universe in response to the energy density that is contained in it; in order to write it as a differential equation in  $a(t)$ , we need to specify  $\rho(a)$ .

The spatial part of the Einstein equations  $G_j^i$  leads to the second Friedmann equation (also known as the *Raychaudhuri equation*):

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3} \left( \rho + \frac{3P}{c^2} \right) \quad (2.33)$$

which also follows by taking the time derivative of (2.31) and using the continuity equation (2.21) for  $\dot{\rho}$ . Equation (2.31) is mostly written in terms of the Hubble parameter and the *critical density* which, for a flat Universe today, equals:

$$\rho_{\text{crit},0} = \frac{3H_0^2}{8\pi G} \simeq 2.8 \times 10^{11} h^2 M_\odot \text{Mpc}^{-3} \quad (2.34)$$

Rescaling all densities by this quantity:

$$\Omega_i = \frac{\rho_{i,0}}{\rho_{\text{crit},0}} \quad (2.35)$$

the Friedmann equation can be rewritten as:

$$\frac{H^2}{H_0^2} = \Omega_r a^{-4} + \Omega_m^{-3} + \Omega_k a^{-2} + \Omega_\Lambda \quad (2.36)$$

where the "curvature" parameter is  $\Omega_k \equiv -kc^2/(R_0 H_0)^2$ . Evaluating this equation at present time  $t_0$  leads to the constraint:

$$1 = \Omega_r + \Omega_m + \Omega_\Lambda + \Omega_k \quad (2.37)$$

An important task of modern cosmology is to measure as precisely as possible these density parameters, since they govern the whole evolution of the Universe.

Generally, equation (2.36) is a complex differential equation for the function  $a(t)$ . We can find solutions to it for some special cases, which correspond to only certain periods of time in the Universe. For instance, a flat Universe with a single fluid component is a good approximation for the various domination phases (radiation, matter, vacuum energy), because of the different scalings of the densities. In these cases the Friedmann equation becomes:

$$\frac{d \log a}{dt} \approx H_0 \sqrt{\Omega_i} a^{-3/2(1+w_i)} \quad (2.38)$$

Integrating, we find:

$$a(t) \propto \begin{cases} t^{2/3(1+w_i)} & w_i \neq -1 \\ e^{H_0 \sqrt{\Omega_\Lambda} t} & w_i = -1 \end{cases} \quad (2.39)$$

so that we have the following evolutions for radiation domination (RD), matter domination (MD) and dark energy domination (AD):

$$a(t) \propto \begin{cases} t^{1/2} & \text{RD} \\ t^{2/3} & \text{MD} \\ e^{H_0 \sqrt{\Omega_\Lambda} t} & \text{AD} \end{cases} \quad (2.40)$$

In conformal time, the Friedmann equation is:

$$\frac{d \log a}{d\eta} \approx H_0 \sqrt{\Omega_i} a^{-1/2(1+3w_i)} \quad (2.41)$$

and we get:

$$a(\eta) \propto \eta^{2/(1+3w_i)} = \begin{cases} \eta & \text{RD} \\ \eta^2 & \text{MD} \\ -\eta^{-1} & \text{AD} \end{cases} \quad (2.42)$$

It is interesting to explore some of these solutions explicitly:

- *Einstein-de Sitter Universe*: The solution for the pure matter flat Universe  $\Omega_m = 1$ :

$$a(t) = \left( \frac{t}{t_0} \right)^{2/3} \quad (2.43)$$

was first derived by Einstein and de Sitter in an influential paper [50]. This solution is very important because it lets us calculate easily the age of the Universe; in fact, it follows from (2.43) that there is a simple relationship between the Hubble parameter and  $t_0$ :

$$H_0 = \frac{2}{3} \frac{1}{t_0} \quad (2.44)$$

which implies that, given the measurement of  $H_0 \approx 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ , the age of the Universe is:

$$t_0 = \frac{2}{3} \frac{1}{H_0} \approx 9 \text{ billion years} \quad (2.45)$$

This was in fact a big problem in cosmology during the 1980s, since this age implied that the Universe was younger than some of the stars contained in it. (Later a period of accelerated expansion driven by dark energy was introduced, which solved many problems, including this one.)

- *de Sitter space*: This is the solution for a Universe in which the only component is dark energy  $\Omega_\Lambda = 1$ . This solution is very important because it turns out that this is a good approximation for the very early (during inflation) and very late (now) Universe. The time dependence in such a model is very interesting: all physical quantities, like the energy density  $\rho$ , are independent of time. In order for any dynamics to happen, it is necessary to break this time translation invariance by imposing a time-dependent form of the stress energy tensor.
  - *Matter and curvature*: It is interesting to follow the evolution of a Universe with only matter and curvature, and  $\Omega_m \neq 1$ . In fact, before the discovery of dark energy, it was thought that the late Universe's dynamics would be dictated completely by its curvature, given the scaling of its energy density  $\rho_k \propto a^{-2}$  compared to that of matter  $\rho \propto a^{-3}$ .
- To derive the evolution of such a Universe, we first rewrite the two Friedmann equations with the *conformal Hubble parameter*  $\mathcal{H} = a'/a$ :

$$\begin{aligned} \mathcal{H}^2 + \frac{k c^2}{R_0^2} &= \frac{8\pi G}{3} \rho_m a^2 \\ \frac{d\mathcal{H}}{d\eta} &= -\frac{4\pi G}{3} \rho_m a^2 \end{aligned} \quad (2.46)$$

Combining these two we get:

$$2 \frac{d\mathcal{H}}{d\eta} + \mathcal{H}^2 + \frac{k c^2}{R_0^2} = 0 \quad (2.47)$$

Taking out the units out of the equations by rescaling  $\theta \equiv c\eta/R_0$  and  $\tilde{\mathcal{H}} \equiv R_0 \mathcal{H}/c$ , equation (2.47) becomes:

$$\tilde{\mathcal{H}}(\theta) = \frac{1}{a} \frac{da}{d\theta} = \begin{cases} \cot \theta/2 & k = +1 \\ \coth \theta/2 & k = -1 \end{cases} \quad (2.48)$$

which is easily integrated to give:

$$a(\theta) = A \begin{cases} \sin \theta/2^2 & k = +1 \\ \sinh \theta/2^2 & k = -1 \end{cases} \quad (2.49)$$

where  $A$  is an integration constant. Using this solution we can determine  $t(\theta)$ :

$$t(\theta) = \frac{R_0}{c} \int a(\theta) d\theta = \frac{R_0 A}{2c} \begin{cases} \theta - \sin \theta & k = +1 \\ \sinh \theta/2 - \theta & k = -1 \end{cases} \quad (2.50)$$

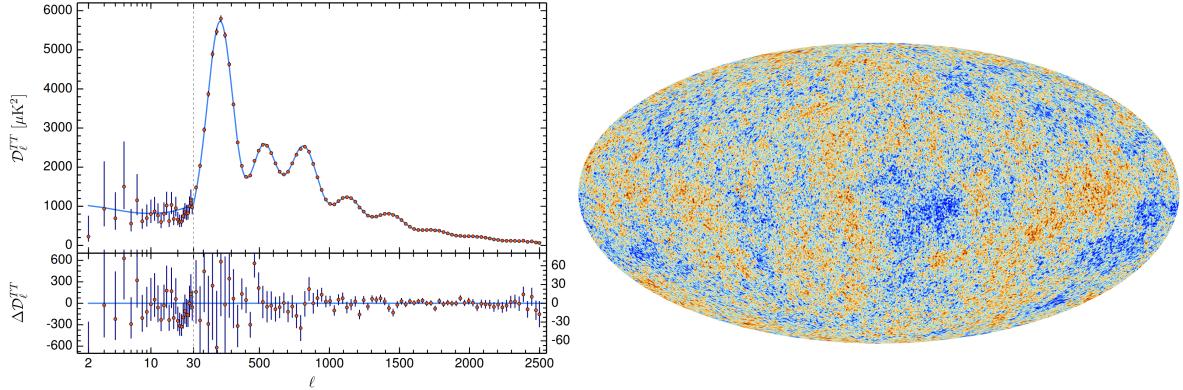


Figure 1: *Left:* CMB temperature power spectrum. Picture adapted from [5]. *Right:* The projection of the CMB's temperature fluctuations. The red spots are hot and the blue ones are cold.

Using equations (2.49) and (2.50) we can derive the function  $a(t)$  in parametric form.

We can clearly see that a Universe with positive curvature recollapses on itself, whereas the one with negative curvature expands forever.

#### 2.1.4 The Parameters

In this section we aim to give an overview of the values of the parameters that define the best model that fits our Universe, the  $\Lambda$ CDM model. All the values and uncertainties are taken from [5].

One quantity that was measured many times and to a great degree of precision is the photon density. The COBE satellite found the CMB temperature to be [51]:

$$T_0 = 2.7255 \pm 0.0006 \text{ K} \quad (2.51)$$

Using basic statistical mechanics for relativistic particles one can find the number density and energy density from the temperature:

$$n_{\gamma,0} = 0.24 \times \left( \frac{k_B T_0}{\hbar c} \right) \approx 410 \text{ photons cm}^{-3} \quad (2.52)$$

and:

$$\rho_{\gamma,0} = 0.66 \times \frac{(k_B T_0)^4}{(\hbar c)^3} \approx 4.6 \times 10^{-34} \text{ g cm}^{-3} \quad (2.53)$$

By rescaling with the critical density we get:

$$\Omega_\gamma \approx 5.4 \times 10^{-5} \quad (2.54)$$

Taking into account neutrinos (as massless particles) we get:

$$\Omega_r = 9 \times 10^{-5} \quad (2.55)$$

The COBE satellite also found that the CMB contains fluctuations, cold and hot spots alternating. These are in fact very small,  $\delta T/T \sim 10^{-5}$ , but contain critical information about the composition of the Universe. The precise shape of the temperature power spectrum depends on the cosmological parameters, thus fitting this curve with data points allows us to measure key cosmological parameters. A representative plot of the CMB power spectrum is found in Figure 1.

These measurements obtain an upper bound on the curvature density fraction:

$$\Omega_k < 0.005 \quad (2.56)$$

We can see that today the energy density in the form of curvature makes up less than 1% of the total energy content.

Turning to baryons, their density is known because it affects the abundances of light elements created during nucleosynthesis and it determines detailed features of the CMB power spectrum. Both of these observations obtain that the baryon density fraction is:

$$\Omega_b \approx 0.05 \quad (2.57)$$

The number density of baryons (mostly protons) is then:

$$n_{b,0} = \frac{\Omega_b \rho_{\text{crit}}}{m_p c^2} \approx 0.3 \times 10^{-6} \text{ cm}^{-3} \quad (2.58)$$

Comparing to the density of photons (2.52) we get the *baryon-to-photon ratio*:

$$\eta \equiv \frac{n_b}{n_\gamma} \approx 6 \times 10^{-10} \quad (2.59)$$

This ratio is of fundamental importance in cosmology, and it is still mysterious why it should be this small. We know that most of the matter in the Universe is in the form of dark matter. The CMB fluctuations actually depend sensitively on the amount of it, so that we can infer its density fraction from them:

$$\Omega_c \approx 0.27 \quad (2.60)$$

where the "c" subscript indicates that we are considering cold dark matter, with an equation of state of the type  $w_c = 0$ . In total, the matter density fraction is thus:

$$\Omega_m \approx 0.32 \quad (2.61)$$

Going back in time, the radiation component becomes more and more important. The scale factor at matter-radiation equality is an important parameter in our model, and it can be estimated from the parameters above to be:

$$a_{eq} = \frac{\Omega_r}{\Omega_m} \approx 2.9 \times 10^{-4} \rightarrow z_{eq} \approx 3400 \quad (2.62)$$

Most of the energy density of the Universe is in the form of dark energy. The way to infer its density fraction is from supernovae observations. In fact, these objects appear fainter than expected in a pure matter universe, therefore the luminosity data can only be fit with a dark energy component, which turns out to be:

$$\Omega_\Lambda \approx 0.68 \quad (2.63)$$

Again, we can estimate the time of matter-dark energy equality:

$$a_{m\Lambda} = \left( \frac{\Omega_m}{\Omega_\Lambda} \right)^{1/3} \approx 0.77 \rightarrow z_{m\Lambda} \approx 0.3 \quad (2.64)$$

Importantly, knowing all parameters, we can estimate the age of the Universe with the Friedmann equation:

$$t_0 = \frac{1}{H_0} \int_0^1 \frac{da}{\sqrt{\Omega_r a^{-2} + \Omega_m a^{-1} + \Omega_\Lambda a^2 + \Omega_k}} \approx 13.8 \times 10^9 \text{ years} \quad (2.65)$$

Evaluating this integral until  $a_{eq}$  or  $a_{m\Lambda}$  gives us instead:

$$\begin{aligned} t_{eq} &\approx 50\,000 \text{ years} \\ t_{m\Lambda} &\approx 10.2 \times 10^9 \text{ years} \end{aligned} \quad (2.66)$$

The following table summarizes the results for the cosmological parameters:

Parameter	Description	Value
$H_0$	Expansion rate	$67.74 \pm 0.46 \text{ km s}^{-1} \text{ Mpc}^{-1}$
$\rho_{\text{crit}}$	Critical Density	$(8.62 \pm 0.12) \times 10^{-27} \text{ kg m}^{-3}$
$\Omega_r$	Radiation density	$(9.02 \pm 0.21) \times 10^{-5}$
$\Omega_\gamma$	Photon density	$(5.38 \pm 0.15) \times 10^{-5}$
$\Omega_\nu$	Neutrino density	$< 0.003$
$\Omega_b$	Baryon density	$0.0493 \pm 0.0006$
$\Omega_m$	Matter density	$0.3153 \pm 0.0073$
$\Omega_\Lambda$	Dark energy density	$0.6847 \pm 0.0073$
$ \Omega_k $	Spatial curvature	$< 0.005$
$t_0$	Age of the Universe	$13.799 \pm 0.021 \text{ billion years}$
$t_{eq}$	Matter-radiation equality	$51100 \pm 0.021 \text{ billion years}$
$z_{eq}$	Redshift of equality	$3402 \pm 26$

Table 1:  $\Lambda$ CDM parameters measured by Planck [5]

### 2.1.5 The problems

The  $\Lambda$ CDM model we just laid out is a well defined, simple and predictive cosmological model. As we have presented it, it is based on a set of assumptions. First of all, our Universe contains radiation, matter (baryons and dark matter) and a cosmological constant  $\Lambda$ , all of it being described on large scales by General Relativity. Therefore, the relevant action is:

$$S = \int d^4x \sqrt{-g} \left( R - 2\Lambda + \frac{1}{4\alpha} F_{\mu\nu} F^{\mu\nu} + \mathcal{L}_m(\psi, A) \right) \quad (2.67)$$

where  $\alpha$  is the fine structure constant,  $G$  is Netwon's constant,  $R$  is the Ricci scalar,  $F_{\mu\nu}$  the electromagnetic field strength tensor and  $\mathcal{L}_m$  is the Lagrangian density for the matter fields  $\psi_m$  coupled to the one form  $A_\mu$ . The cosmological principle, which states that the Universe is spatially homogeneous and isotropic at sufficiently large scales ( $\gtrsim 100 \text{ Mpc}$ ), allows us to write the metric of the Universe in a simple standard form, the FLRW metric (2.9). Solving the Einstein equations brings us finally to the Friedmann equations (2.31) and (2.33). At last, a phase of inflation before the Hot Big Bang is introduced to account for the *horizon problem* and the *flatness problem*, where Gaussian scale invariant fluctuations are produced from a quantum field.

The most immediate generalizations to the  $\Lambda$ CDM model may be produced by modifying the fundamental constants in the action (2.67) or adding new terms. Some of the extensions have to do with promoting Newton's constant to a dynamical degree of freedom, the so called "scalar-tensor theories", by allowing it to depend on a scalar field  $G \rightarrow G(\Phi(x))$  where the dynamics of  $\Phi$  is governed by some terms added to the action [52]; another modification would be again to change the cosmological constant into a dynamical degree of freedom  $\Lambda \rightarrow \Lambda(\Phi(x))$ . In addition, there are some theories called *Maxwell Dilaton theories*, where the fine structure constant is changed  $\alpha \rightarrow \alpha(\Phi(x))$  [53]. Finally, one can think of the addition of new terms into the action, which may be functions of the Ricci scalar, the torsion scalar, the Kretschmann scalar or other invariants.

Despite its success, the  $\Lambda$ CDM model is under intense investigation, motivated by a range of profound inconsistencies found within the theory and the observations.

From a theoretical standpoint, the most important difficulties are the *fine tuning* and the *coincidence* problems. The former is related to the fact that there is a large discrepancy between observations and theoretical expectations on the value of the cosmological constant  $\Lambda$  [6]. The coincidence problem instead has to do with the observed coincidence of the vacuum energy density and the matter density which are approximately equal nowadays despite having very different scaling relations; as a solution to this problem, the anthropic principle has been summoned, stating that these correspondences result from a selection bias towards the existence of human life in the context of a multiverse [54].

From the observational side of things, instead, there are signals in cosmological/astrophysical data that

appear to be in tension ( $\geq 2\sigma$ ) with the  $\Lambda$ CDM model. One of the main challenges is the "Hubble tension" that we introduced above in the context of the measurements of the Hubble constant  $H_0$ . Some other important tensions are the *growth tension* [7], which has to do with an inconsistency between the measure of the growth rate from direct measurements and the one indicated by the Planck/ $\Lambda$ CDM parameter values, the *CMB anisotropy anomalies* [8], which include lack of power on large angular scales, small vs. large scales tension, cold spot anomaly etc., *cosmic dipoles* [9], which indicates the violation of the cosmological principle, *cosmic birefringence* [10], which indicates a hint of a new ingredient beyond the Standard Model, *the age of the Universe problem* [11], where the age of the Universe obtained from local measurements using the ages of the oldest stars in the Milky Way appears to be slightly larger than the age obtained using the CMB, *small scale problems* [12], like the core-cusp problem, the missing satellite problem, too big to fail problem, angular momentum catastrophe etc. and *the Lithium problem* [13], where the measurement of old, metal poor stars in the Milky Way shows that they have 5 times less lithium than is predicted by BBN. For a comprehensive review of all the tensions see [55].

## 2.2 Cosmological Perturbation Theory

In the previous section, making use of large scales observations, we have treated the Universe as perfectly homogeneous and isotropic. However, the presence of structures in the Universe warrants the introduction of inhomogeneities in the metric whose evolution, as long as they stay in the linear regime, can be followed with perturbation theory.

After neutrino decoupling and BBN, the only relevant forces that play a critical role in cosmological evolution are gravity, which is the same for all particles due to the equivalence principle, and electromagnetism. Before recombination, photons, electrons and protons are tightly coupled by their electromagnetic interactions and are thus treated usually as a single fluid. The two main components that drive the evolution of the Universe are the homogeneous densities of the species, which govern the cosmological expansion rate, and their perturbations, which source the local gravitational potential. This physical process is in general highly non-linear, and it can lead to very nontrivial dynamics.

Two important times in the evolution of the Universe are matter-radiation equality, which happens at around  $z_{eq} \sim 3400$ , and photon decoupling  $z_{dec} \sim 1100$  (which basically coincides with recombination). They are reference points because, before equality, the growth of density perturbations is damped by the high pressure exerted by photons, therefore the growth only starts to become significant when the Universe becomes matter dominated. Moreover, before decoupling, the baryons are coupled to photons and their density perturbations oscillate, as opposed to the dark matter ones, which decouples very early on. These oscillations, sound waves in the primordial plasma, arise as a result of the dance between two opposing forces: the clustering effect of gravity battles with the radiation pressure which causes the perturbations to re-expand; the frequency of these oscillations depends on the wavelength of the fluctuations, so that at decoupling modes with different wavelengths are captured at different phases in their evolutions, giving rise to the characteristic CMB patterns, as can be seen in Figure 1.

After decoupling, when the electrons get captured by the protons to make the first atoms, the baryons lose the pressure support of the photons and their perturbations begin to grow, and eventually they fall into the potential wells formed by the already clumped dark matter. This is the start of structure formation.

### 2.2.1 The Gauge Philosophy

In the perturbation theory of General Relativity, one considers a *perturbed spacetime* that is close to being a simple, symmetric spacetime, the *background spacetime*. During our discussion, these are the two objects that constantly play with each other:

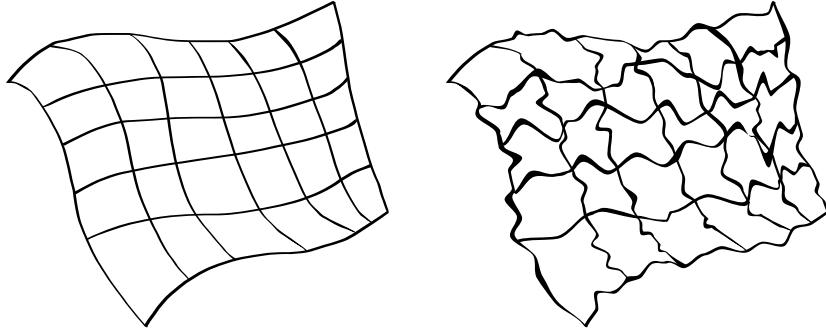


Figure 2: The two spacetimes at play in perturbation theory: the background spacetime on the left and the total, perturbed one on the right

The coordinate system on the perturbed spacetime can be written with the following metric:

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + \delta g_{\mu\nu} \quad (2.68)$$

where  $\bar{g}_{\mu\nu}$  is the metric on the background, unperturbed spacetime and  $\delta g_{\mu\nu}$  is just a small perturbation. We require in general, even though it's not strictly necessary, that the first and second derivatives of the perturbation  $\delta g_{\mu\nu,\rho}$  and  $\delta g_{\mu\nu,\rho\sigma}$  are small. The same idea applies to the Einstein and energy momentum tensors:

$$\begin{aligned} G_{\mu\nu} &= \bar{G}_{\mu\nu} + \delta G_{\mu\nu} \\ T_{\mu\nu} &= \bar{T}_{\mu\nu} + \delta T_{\mu\nu} \end{aligned} \quad (2.69)$$

The background and the whole (perturbed) spacetimes obey the Einstein equations:

$$G_{\mu\nu} = 8\pi G T_{\mu\nu} \quad \text{and} \quad \bar{G}_{\mu\nu} = 8\pi G \bar{T}_{\mu\nu} \quad (2.70)$$

From these we can see, by subtracting them, that the perturbations obey the same equations:

$$\delta G_{\mu\nu} = 8\pi G \delta T_{\mu\nu} \quad (2.71)$$

We need to make a few remarks here. First, the above calculations require a pointwise correspondence in the two spacetimes given by the coordinate system  $\{x^\alpha\}$ , which lets us subtract the two equations, in the following sense: the point  $\bar{P}$  in the background and the mapped one  $P$  in the perturbed spacetime correspond (i.e. are the same physical point) and have the same coordinate. Given a background spacetime though, there exist many coordinate systems, very close to one another, for which (2.68) holds. The choice among all these coordinate systems is called a *gauge choice*, and it is of critical importance in perturbative GR. In this section we develop first order perturbation theory, so that we drop all terms which are "doubly small", i.e. the product of two infinitesimal quantities  $\delta g_{\mu\nu}$ , and the derivatives we cited above. In this way the Einstein equations will become linear in the perturbation  $\delta g_{\mu\nu}$ , whereas in Fourier space the modes don't mix with each other. This is definitely simpler than solving the full nonlinear theory.

Secondly, the fact that the spatial averages  $\bar{G}_{\mu\nu}$  and  $\bar{T}_{\mu\nu}$  satisfy the Einstein equations is only true in linear perturbation theory. In general, we expect this to not be true since GR is a nonlinear theory, so that the perturbations affect the evolution of the mean. This effect is called *backreaction* and the degree to which it influences the background is still a subject of intense research. In fact, since by the time of structure formation these perturbation grow very large, we expect this backreaction to affect non-trivially the evolution of the background universe.

Lastly, we remark that the separation of the quantities into their background and perturbed parts is always performed as to spatially average the perturbations to zero, i.e. the background value at a certain fixed time  $t$  is just the spatial average of the full quantity over the hypersurface  $t = \text{const}$ .

Now that we've seen this separation of quantities, we have to tackle the important problem of *gauge transformations* since, as we said, there are many possible perturbed systems for a given background spacetime,

all close to one another. In this case, a gauge transformation means a coordinate transformation that relates two of these perturbed systems. To develop the formalism, we denote the coordinates of the background by  $x^\alpha$  and the ones in the two different perturbed systems as  $\hat{x}^\alpha$  and  $\tilde{x}^\alpha$ ; the latter are related then by a gauge transformation by:

$$\tilde{x}^\alpha = \hat{x}^\alpha + \xi^\alpha \quad (2.72)$$

where  $\xi$  is a small parameter (we thus drop products of this with itself, and its derivatives).

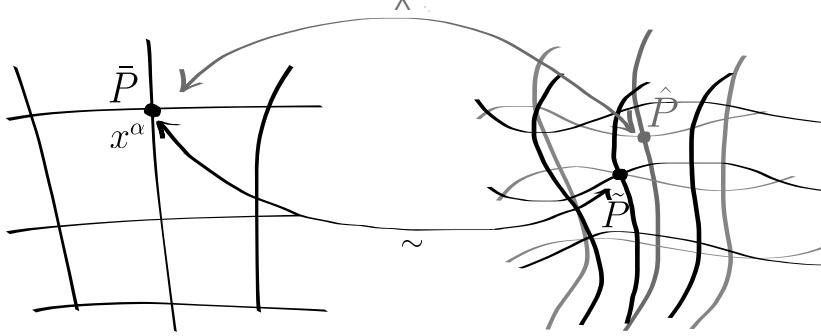


Figure 3: A gauge transformation. On the left, the same background spacetime is mapped on two different perturbed ones, which are related by a coordinate transformation.

In particular, the difference between  $\partial_\beta \xi^\alpha$  and  $\hat{\partial}_\beta \xi^\alpha$  is second order small and thus ignored, so we can actually think of  $\xi$  as living in the background spacetime. We illustrate the situation in Figure 3. We know that the coordinate system  $\{\hat{x}^\alpha\}$  associates point  $\bar{P}$  in the background with  $\hat{P}$ , whereas  $\{\tilde{x}^\alpha\}$  associates it to  $\tilde{P}$ . This is done by:

$$\tilde{x}^\alpha(\tilde{P}) = \hat{x}^\alpha(\hat{P}) = x^\alpha(\bar{P}) \quad (2.73)$$

The gauge transformation relates the coordinates of the same point in the perturbed spacetime, that is:

$$\begin{aligned} \tilde{x}^\alpha(\tilde{P}) &= \hat{x}^\alpha(\tilde{P}) + \xi^\alpha \\ \tilde{x}^\alpha(\hat{P}) &= \hat{x}^\alpha(\hat{P}) + \xi^\alpha \end{aligned} \quad (2.74)$$

where we notice that the difference between  $\xi^\alpha(\tilde{P})$  and  $\xi^\alpha(\hat{P})$  is second order small, and thus we associate  $\xi^\alpha$  to the background:

$$\xi^\alpha = \xi^\alpha(\bar{P}) = \xi^\alpha(x^\beta) \quad (2.75)$$

Putting together equations (2.73) and (2.74) we get the relation between two different points in the same coordinate system:

$$\begin{aligned} \hat{x}^\alpha(\tilde{P}) &= \hat{x}^\alpha(\hat{P}) - \xi^\alpha \\ \tilde{x}^\alpha(\tilde{P}) &= \tilde{x}^\alpha(\hat{P}) - \xi^\alpha \end{aligned} \quad (2.76)$$

We are now ready to understand how various tensors transform under a gauge transformation. We start with the perturbed quantities:

$$\begin{aligned} s &= \bar{s} + \delta s \\ w^\alpha &= \bar{w}^\alpha + \delta w^\alpha \\ B_{\alpha\beta} &= \bar{B}_{\alpha\beta} + \delta B_{\alpha\beta} \end{aligned} \quad (2.77)$$

We are now going to find how the scalar  $s$  changes under a gauge transformation, and just state the results for the other two cases.

Let's first consider the full quantity  $s = \bar{s} + \delta s$ , which lives on the perturbed spacetime. We cannot assign a unique background quantity  $\bar{s}$  to a point in the perturbed spacetime, because different gauges may have the

same values for for  $\bar{s}$ . This implies that there is no unique perturbation  $\delta s$ , but the perturbation is gauge dependent. As we saw above, the perturbations in different gauges are defined by:

$$\begin{aligned}\hat{\delta}s(x^\alpha) &\equiv s(\hat{P}) - \bar{s}(\bar{P}) \\ \tilde{\delta}s(x^\alpha) &\equiv s(\tilde{P}) - \bar{s}(\bar{P})\end{aligned}\quad (2.78)$$

We see that the perturbation  $\delta s$  is just a subtraction between two spacetimes, and therefore we consider it as living on the background spacetime. We can now relate  $\hat{\delta}s$  to  $\tilde{\delta}s$ :

$$s(\tilde{P}) = s(\hat{P}) + \frac{\partial s}{\partial \hat{x}^\alpha}(\hat{P})(\hat{x}^\alpha(\tilde{P}) - \hat{x}^\alpha(\hat{P})) = s(\hat{P}) - \frac{\partial s}{\partial \hat{x}^\alpha}(\hat{P})\xi^\alpha = s(\hat{P}) - \frac{\partial \bar{s}}{\partial x^\alpha}(\bar{P})\xi^\alpha \quad (2.79)$$

where we approximated  $\frac{\partial s}{\partial \hat{x}^\alpha}(\hat{P}) \approx \frac{\partial \bar{s}}{\partial x^\alpha}(\bar{P})$ . However, since we are interested in scenarios where the background is homogeneous and time dependent, we have  $\bar{s} = \bar{s}(\eta, x^i) = \bar{s}(\eta)$  only, thus:

$$\frac{\partial \bar{s}}{\partial x^\alpha}(\bar{P})\xi^\alpha = \frac{\partial \bar{s}}{\partial \eta}(\bar{P})\xi^0 \quad (2.80)$$

so that we get in the end:

$$s(\tilde{P}) = s(\hat{P}) - \bar{s}'\xi^0 \quad (2.81)$$

and the final gauge transformation for scalar perturbations is then:

$$\tilde{\delta}s(x^\alpha) = s(\hat{P}) - \bar{s}'\xi^0 - \bar{s}(\bar{P}) = \hat{\delta}s(x^\alpha) - \bar{s}'\xi^0 \quad (2.82)$$

The other cases, respectively for vectors and tensors in (2.77), are:

$$\begin{aligned}\tilde{\delta}w^\alpha &= \hat{\delta}w^\alpha + \xi_{,\beta}^\alpha \bar{w}^\beta - \bar{w}_{,\beta}^\alpha \xi^\beta \\ \tilde{\delta}B_{\mu\nu} &= \hat{\delta}B_{\mu\nu} - \xi_{,\mu}^\rho \bar{B}_{\rho\nu} - \xi_{,\nu}^\sigma \bar{B}_{\mu\sigma} - \bar{B}_{\mu\nu,\alpha} \xi^\alpha\end{aligned}\quad (2.83)$$

In *cosmological perturbation theory* the background Universe is described by the flat FLRW metric:

$$ds^2 = -dt^2 + a^2(t)d\vec{x}^2 = a^2(\eta)(-d\eta^2 + d\vec{x}^2) \quad (2.84)$$

where  $a(t)$  is the scale factor, and we introduced the conformal time  $d\eta = dt/a$ . The perturbed spacetime can be written generally as:

$$ds^2 = a^2(\eta)[-(1+2A)d\eta^2 + 2B_i dx^i d\eta + (\delta_{ij} + 2E_{ij})dx^i dx^j] \quad (2.85)$$

where all the new parameters  $A$ ,  $B_i$  and  $E_{ij}$  are functions of the spacetime coordinates. It will be extremely useful to perform a *scalar-vector-tensor decomposition* (SVT) of the perturbations. This just means that we split any three vector into the gradient of a scalar and a divergenceless vector:

$$B_i = \partial_i B + \hat{B}_i \quad (2.86)$$

with  $\partial^i \hat{B}_i = 0$ . Importantly, in Fourier space the decomposition becomes  $B_i = i\hat{k}_i B + \hat{B}_i$ , where  $\hat{k} \equiv \vec{k}/|\vec{k}|$ , so that the vector has been split into a piece that is along the same direction of the mode  $\vec{k}$  and a piece that is orthogonal to it. In the same manner we split the tensor perturbations:

$$E_{ij} = C\delta_{ij} + \partial_{(i}\partial_{j)} E + \partial_{(i}\hat{E}_{j)} + \hat{E}_{ij} \quad (2.87)$$

where

$$\begin{aligned}\partial_{(i}\partial_{j)} E &\equiv \left( \partial_i \partial_j - \frac{1}{3} \delta_{ij} \nabla^2 \right) E \\ \partial_{(i}\hat{E}_{j)} &\equiv \frac{1}{2} (\partial_i \hat{E}_j + \partial_j \hat{E}_i)\end{aligned}\quad (2.88)$$

The first term in (2.87) contains the trace of the spatial perturbations  $E_i^i = 3C$ , while the other ones are traceless. As before, the hatted quantities have vanishing divergence  $\partial^i \hat{E}_i = \partial^i \hat{E}_{ij} = 0$ , and therefore are the transverse vector and tensor perturbations. In Fourier space, we write the first of (2.88) as  $-(\hat{k}_i \hat{k}_j - 1/3 \delta_{ij}) E \equiv -\hat{k}_{(i} \hat{k}_{j)} E$ .

What makes the SVT decomposition very powerful and handy is the fact that, at linear order, the Einstein equations for scalars, vectors and tensors don't mix and can therefore be treated completely separately. For the most part we will be interested in scalar perturbation, and at linear order it is consistent to then set vector and tensor perturbations to zero anyway.

Given the transformation of the metric under a gauge transformation  $\tilde{x}^\alpha = x^\alpha + \xi^\alpha$  that we found above:

$$\tilde{\delta} g_{\mu\nu} = \delta g_{\mu\nu} - \xi_{,\mu}^\rho \bar{g}_{\rho\nu} - \xi_{,\nu}^\sigma \bar{g}_{\mu\sigma} - \bar{g}_{\mu\nu,0} \xi^0 \quad (2.89)$$

and the form of the perturbations:

$$\delta g_{\mu\nu} = a^2 \begin{pmatrix} 2A & B_i \\ B_i & 2E_{ij} \end{pmatrix} \quad (2.90)$$

if we decompose the gauge transformation as  $\xi^0 \equiv T$  and  $\xi^i \equiv L^i = \partial^i L + \hat{L}^i$  we get the transformation rules for the perturbations. One can show that:

$$\begin{aligned} \tilde{A} &= A - T' - \mathcal{H}T \\ \tilde{B} &= B + T - L' \\ \tilde{C} &= C - \mathcal{H}T - \frac{1}{3} \nabla^2 L \\ \tilde{E} &= E - L \\ \tilde{\hat{B}}_i &= \hat{B}_i - \hat{L}'_i \\ \tilde{\hat{E}}_i &= \hat{E}_i - \hat{L}_i \\ \tilde{\hat{E}}_{ij} &= \hat{E}_{ij} \end{aligned} \quad (2.91)$$

Before carrying on with the calculations, we would like to comment on the whole philosophy at play here. The perturbations are defined as differences between the perturbed spacetime and the background spacetime, so in which spacetime do they live? What we are doing in perturbation theory is building a description of the perturbed spacetime as a set of fields (the perturbations), that live in the background. Thus the background spacetime is the playing field on which we are creating the perturbed spacetime. As we have said, for the same perturbed spacetime, there exist many equivalent descriptions of this kind and they are all related by gauge transformations. This formalism of gauge theory in the background spacetime is similar to gauge theories used in particle physics.

As we briefly mentioned before, the scalar perturbations are the most important ones, since they couple to pressure and density perturbations and follow the gravitational instability, that is overdense regions become more overdense over time, and viceversa for underdense regions. On the other hand, the vector perturbations couple to rotational velocity perturbations in the cosmic fluid, and therefore are expected to decay in an expanding Universe; in this sense we can neglect them. Finally, it is found that tensor perturbations are gauge invariant. These are gravitational waves, automatically in the traceless and transverse gauge, and are very important since they could have observable effects on CMB anisotropies if strong enough.

We stressed above that the perturbations are not uniquely defined as a consequence of the gauge invariance of GR. The intrinsic problem stemming from this is the possible appearance of fictitious *gauge modes*, sham perturbations that can arise from an inconvenient choice of coordinates. As an example, we can consider the transformation of the homogeneous density of the Universe under a local change of coordinates in the time slicing  $\eta \rightarrow \tilde{\eta} = \eta + \xi^0(\eta, \vec{x})$ , given in (2.82):

$$\rho(\eta) = \rho(\tilde{\eta} - \xi^0(\tilde{\eta}, \vec{x})) = \bar{\rho}(\tilde{\eta}) - \bar{\rho}' \xi^0 \quad (2.92)$$

We see that even in a perfectly homogeneous Universe, density perturbations may arise fictitiously if a certain coordinate chart is used rather than another one. Conversely, taking a perturbed Universe, we can remove the perturbations by declaring that we are using coordinates such that  $\delta\rho = 0$  (choosing the hypersurface of constant time to coincide with the hypersurface of constant energy density); naturally, in this last case, the perturbations will reappear in some form in the metric.

In general there are two ways one can deal with this problem, which involves a compromise between mathematical simplicity and physical "realness", in the following sense.

We can define the following variables, the *Bardeen variables*, which are clever combinations of the metric perturbations that are gauge invariant:

$$\begin{aligned}\Psi &\equiv A + \mathcal{H}(B - E') + (B - E)' \\ \Phi &\equiv -C + \frac{1}{3}\nabla^2 E - \mathcal{H}(B - E')\end{aligned}\tag{2.93}$$

These, being gauge invariant, would be the "real" spacetime perturbations.

An alternative strategy is to fix the gauge and keep track of all the perturbations, both of those in the metric and those in the matter. In this way the math is considerably simpler, and a lot of different gauges are used for different purposes. One that is very important is the *Newtonian gauge*:

$$B = E = 0\tag{2.94}$$

The metric becomes then very simple:

$$ds^2 = a^2(\eta)[-(1+2\Psi)d\eta^2 + (1-2\Phi)\delta_{ij}dx^i dx^j]\tag{2.95}$$

It is evident that this metric is very useful. First, it is diagonal, so the hypersurfaces of constant time are orthogonal to the worldlines of the observers at rest, but also the induced spatial geometry is isotropic. Finally, this metric resembles the one obtained in the weak-field limit of GR, with  $\Psi$  playing the role of gravitational potential. Other useful gauges are the *spatially flat gauge* ( $C = E = 0$ ) and the *synchronous gauge* ( $A = B = 0$ ).

### 2.2.2 Perturbations in the energy-momentum

In this section we consider the perturbations in the energy momentum tensor. As we saw above, the background form of it is completely fixed by the requirements of isotropy and homogeneity (2.19):

$$\bar{T}^{\mu\nu} = (\bar{\rho} + \bar{P})\bar{u}^\mu\bar{u}^\nu + \bar{P}g^{\mu\nu}\tag{2.96}$$

with  $\bar{u}^\mu = 1/a(1, \vec{0})$ . The perturbed energy tensor is (2.69):

$$T^{\mu\nu} = \bar{T}^{\mu\nu} + \delta T^{\mu\nu}\tag{2.97}$$

The perfect fluid degrees of freedom in  $\delta T^{\mu\nu}$  are those that keep  $T^{\mu\nu}$  in the form of a perfect fluid:

$$T^{\mu\nu} = (\rho + P)u^\mu u^\nu + Pg^{\mu\nu}\tag{2.98}$$

and therefore can be taken as the density, pressure and velocity perturbations:

$$\rho = \bar{\rho} + \delta\rho \quad P = \bar{P} + \delta P \quad u^i = \bar{u}^i + \delta u^i \equiv \frac{1}{a}v^i\tag{2.99}$$

Since we have the constraint that  $u_\mu u^\mu = -1$ , the  $\delta u^0$  perturbation is not independent. We now call:

$$v^i \equiv au^i\tag{2.100}$$

the velocity perturbation. We also use the density contrast  $\delta \equiv \delta\rho/\rho$ , introduced above. If we want to express  $u^\mu$  and  $u_\mu$  in terms of  $v_i$ , we write them in the following way:

$$\begin{aligned}u^\mu &= \bar{u}^\mu + \delta u^\mu = (a^{-1} + \delta u^0, a^{-1}\vec{v}) \\ u_\nu &= \bar{u}_\nu + \delta u_\nu = (a^{-1} + \delta u_0, \delta\vec{u})\end{aligned}\tag{2.101}$$

Using the form of the perturbed metric (2.90), we find:

$$u_0 = g_{0\mu} u^\mu = \dots = -a - a^2 \delta u^0 - 2aA \quad (2.102)$$

from which it follows:

$$\delta u_0 = -a^2 \delta u^0 - 2aA \quad (2.103)$$

We get, in the same manner:

$$\delta u_i = u_i = g_{i\mu} u^\mu = -aB_i + av_i \quad (2.104)$$

Finally, by using  $u_\mu u^\mu = -1$  we can solve for  $\delta u^0$ :

$$u_\mu u^\mu = -1 = \dots = -1 - 2a\delta u^0 - 2A \implies \delta u^0 = -\frac{1}{a}A \quad (2.105)$$

So that we have the 4-velocity:

$$\begin{aligned} u^\mu &= \frac{1}{a}(1 - A, v_i) \\ u_\mu &= a(-1 - A, v_i - B_i) \end{aligned} \quad (2.106)$$

By inserting this form into equation (2.98) and neglecting vector perturbations we get:

$$\delta T_v^\mu = \begin{pmatrix} -\delta\rho & (\bar{\rho} + \bar{P})v_i \\ -(\bar{\rho} + \bar{P})v_i & \delta P \delta_j^i \end{pmatrix} \quad (2.107)$$

The only degrees of freedom remaining are those in the spatial part of the energy momentum tensor, which we write as:

$$\delta T_j^i = \delta P + \Pi_j^i \quad (2.108)$$

where  $\Pi_j^i$  is called the *anisotropic stress* and  $\Pi_i^i = 0$ . We will usually write the perturbations in terms of the *momentum density*  $q^i \equiv (\bar{\rho} + \bar{P})v^i$ . Note that if there are several species present, the perturbations add up, i.e.  $T_{\mu\nu} = \sum_a T_{\mu\nu}^{(a)}$ :

$$\delta\rho = \sum_a \delta\rho_a \quad \delta P = \sum_a \delta P_{(a)} \quad q^i = \sum_a q_{(a)}^i \quad \Pi^{ij} = \sum_a \Pi_{(a)}^{ij} \quad (2.109)$$

Notice that the velocities do not add up, but the momentum densities  $q_i$  do.

As for the metric, we can extract scalar, vector and tensor parts out of the perturbations. The  $\delta\rho$  and  $\delta P$  naturally only have scalar parts, whereas we can write the momentum density and velocity as:

$$\begin{aligned} v_i &= \partial_i + \hat{v}_i \\ q_i &= \partial_i + \hat{q}_i \end{aligned} \quad (2.110)$$

and the anisotropic stress as:

$$\Pi_{ij} = \partial_{(i}\partial_{j)}\Pi + \partial_{(i}\hat{\Pi}_{j)} + \hat{\Pi}_{ij} \quad (2.111)$$

All in all, the scalar perturbations are given by the variables  $\delta, \delta P, v$  and  $\Pi$ . These, under a gauge transformation:

$$\tilde{\delta T}_{\mu\nu} = \delta T_{\mu\nu} - \xi_{,\mu}^\rho \tilde{T}_{\rho\nu} - \xi_{,\nu}^\sigma \tilde{T}_{\mu\sigma} - \tilde{T}_{\mu\nu,0} \xi^0 \quad (2.112)$$

transform as:

$$\begin{aligned} \widetilde{\delta\rho} &= \delta\rho - \bar{\rho}' \xi^0 \\ \widetilde{\delta P} &= \delta P - \bar{P}' T \\ \tilde{q}_i &= q_i + (\bar{\rho} + \bar{P})L'_i \\ \tilde{v}_i &= v_i + L'_i \\ \tilde{\Pi}_{ij} &= \Pi_{ij} \end{aligned} \quad (2.113)$$

Again, we can use various gauges to simplify our calculations. Example of useful gauges are the *uniform density gauge* ( $\delta\rho = 0$ ) and the *comoving gauge* ( $v + B = 0$ ). Moreover, just as was the case for the Bardeen potentials (2.93), we can define gauge invariant quantities that follow the "real" matter perturbations. One of these is the *comoving density contrast*  $\Delta$ :

$$\bar{\rho}\Delta \equiv \delta\rho + \bar{\rho}'(v + B) \quad (2.114)$$

which reduces to the density constraint in the comoving gauge, hence the name. Two additional important gauge invariant quantities are:

$$\begin{aligned} \zeta &\equiv -C + \frac{1}{3}\nabla^2 E + \mathcal{H}\frac{\delta\rho}{\bar{\rho}'} \\ \mathcal{R} &\equiv -C + \frac{1}{3}\nabla^2 E - \mathcal{H}(v + B) \end{aligned} \quad (2.115)$$

called the *curvature perturbations*, since they reduce to the intrinsic curvature of constant time hypersurfaces in the uniform density and comoving gauges respectively. These three gauge invariant variables are not independent, but are related by:

$$\zeta = \mathcal{R} + \frac{\mathcal{H}}{\bar{\rho}'}\bar{\rho}\Delta \quad (2.116)$$

We will find that the comoving density contrast vanishes on superhorizon scales, so that in fact:

$$\zeta \rightarrow \mathcal{R} \quad \text{if } k \ll \mathcal{H} \quad (2.117)$$

This means that we can treat  $\zeta$  and  $\mathcal{R}$  as the same variable on superhorizon scales. These curvature perturbations, being constant on these scales, turn out to be a good candidate with which to define the initial conditions of our differential equations.

Once we have defined the perturbations of the energy momentum tensor, we need to find the equations of their evolution in a perturbed spacetime. These equations come from expanding to first order the conservation equation for  $T_{\mu\nu}$ :

$$T_{\nu;\mu}^\mu = 0 \quad (2.118)$$

For the metric, we use the Newtonian gauge (2.95):

$$ds^2 = a^2(\eta)[-(1+2\Psi)d\eta^2 + (1-2\Phi)\delta_{ij}dx^i dx^j] \quad (2.119)$$

Solving the  $\nu = 0$  conservation equation, we get the continuity equation:

$$\delta' = -\left(1 + \frac{\bar{P}}{\bar{\rho}}\right)(\nabla \cdot \vec{v} - 3\Phi') - 3\mathcal{H}\left(\frac{\delta P}{\delta\rho} - \frac{\bar{P}}{\bar{\rho}}\right)\delta \quad (2.120)$$

The last term on the right hand side has to do with the dilution of the energy density due to the expansion of the background Universe, while the first term has to do with the local fluid flow. The second term, on the other hand, is a relativistic effect, and is controlled by the perturbations in the local expansion rate (this arises because we can think of  $(1-\Phi)a$  as a local scale factor in the spatial part of the metric in the Newtonian gauge). For  $P \ll \rho$  we recover the continuity equation in Newtonian fluid dynamics, written in conformal time and with a relativistic correction:

$$\delta' = -\nabla \cdot \vec{v} + 3\Phi' \quad (2.121)$$

If we are working on sub-Hubble scales, where the Newtonian theory is still valid, the relativistic correction is small and we recover:

$$\dot{\delta} = -\frac{1}{a}\nabla \cdot \vec{v} \quad (2.122)$$

The relativistic *Euler equation* is obtained by setting  $\nu = 1$  in (2.118):

$$v'_i = -\left(\mathcal{H} + \frac{\bar{P}'}{\bar{\rho} + \bar{P}}\right)v_i - \frac{1}{\bar{\rho} + \bar{P}}(\partial_i \delta P + \partial^j \Pi_{ij}) - \partial_i \Psi \quad (2.123)$$

It will prove useful to evaluate equations (2.120) and (2.123) for some special cases. For a non-relativistic fluid (matter) with  $P_m = 0$  and  $\Pi_m^{ij} = 0$ , we get:

$$\begin{aligned}\delta'_m &= -\nabla \cdot \vec{v}_m + 3\Phi' \\ \vec{v}'_m &= -\mathcal{H}\vec{v}_m - \nabla\Psi\end{aligned}\tag{2.124}$$

which, combined, give:

$$\delta''_m + \mathcal{H}\delta'_m = \nabla^2\Psi + 3(\Phi'' + \mathcal{H}\Phi')\tag{2.125}$$

Here, we see that the second term on the left hand side is due to friction over an expanding background, whereas the right hand side is completely a relativistic effect.

On the other hand, for a relativistic perfect fluid (radiation) with  $P_r = 1/3\rho_r$  and  $\Pi_r^{ij} = 0$ , we have:

$$\begin{aligned}\delta'_r &= -\frac{4}{3}\nabla \cdot \vec{v}_r + 4\Phi' \\ \vec{v}'_r &= -\frac{1}{4}\nabla\delta_r - \nabla\Psi\end{aligned}\tag{2.126}$$

Combining them yields:

$$\delta''_r - \frac{1}{3}\nabla^2\delta_r = \frac{4}{3}\nabla^2\Psi + 4\Phi''\tag{2.127}$$

The main equations (2.120) and (2.123) apply separately for every component in the Universe, provided that there is no energy and momentum transfer between them, in which case the species interact only gravitationally through the evolution of the Bardeen variables. It is useful to rewrite them in Fourier space with the following conventions for scalar perturbations:

$$\begin{aligned}\partial_i &\rightarrow ik_i \\ v_i &\rightarrow i\hat{k}_i v \\ \Pi_{ij} &\rightarrow -(\bar{\rho} + \bar{P})\hat{k}_{(i}\hat{k}_{k)}\Pi\end{aligned}\tag{2.128}$$

Then:

$$\begin{aligned}\delta' &= \left(1 + \frac{\bar{P}}{\bar{\rho}}\right)(kv + 3\Phi') - 3\mathcal{H}\left(\frac{\delta P}{\delta\rho} - \frac{\bar{P}}{\bar{\rho}}\right)\delta \\ v' &= -\left(\mathcal{H} + \frac{\bar{P}'}{\bar{\rho} + \bar{P}}\right)v - \frac{1}{\bar{\rho} + \bar{P}}k\delta P + \frac{2}{3}k\Pi - k\Psi\end{aligned}\tag{2.129}$$

Some last remarks are needed in order to simplify the equations. Sometimes we may assume a perfect fluid, where the strong interactions between the particles keep the pressure isotropic  $\Pi = 0$ . In addition, for a barotropic fluid we have  $P = P(\rho)$ , so that the pressure perturbations can be written as  $\delta P = c_s^2\delta\rho$ , where:

$$c_s^2 = \left(\frac{\partial P}{\partial\rho}\right)_S = \frac{\bar{P}'}{\bar{\rho}'}\tag{2.130}$$

Moreover, the perturbations will be called *adiabatic* if:

$$\delta P = c_s^2\delta\rho = \frac{\bar{P}'}{\bar{\rho}'}\delta\rho\tag{2.131}$$

In the case of a barotropic fluid, the perturbations are automatically adiabatic, but the reverse need not be true in the most general of cases. If the perturbations are adiabatic, we have reduced the number of independent variables to solve for, only  $\delta$  and  $v$ . It is interesting to note that, although the perfect fluid approximation is not valid for decoupled or weakly interacting species, dark matter is a special case, because it has a negligible velocity dispersion, and therefore behaves like a pressureless perfect fluid.

### 2.2.3 Spacetime perturbations and the Einstein equations

To close out the system of equations, we need to find the evolution of the metric potentials  $\Psi$  and  $\Phi$ . We will thus write the Einstein equations:

$$G_\nu^\mu = 8\pi G T_\nu^\mu \quad (2.132)$$

in the Newtonian gauge. The linearized version of these equations yield:

$$\begin{aligned} \nabla^2 \Phi - 3\mathcal{H}(\Phi' + \mathcal{H}\Psi) &= 4\pi Ga^2 \delta\rho \\ -(\Phi' + \mathcal{H}\Psi) &= 4\pi Ga^2 q \\ \partial_{(i} \partial_{j)} (\Phi - \Psi) &= 8\pi Ga^2 \Pi_{ij} \\ \Phi'' + \mathcal{H}\Psi' + 2\mathcal{H}\Phi' + \frac{1}{3}\nabla^2(\Psi - \Phi) + (2\mathcal{H}' + \mathcal{H}^2)\Psi &= 4\pi Ga^2 \delta P \end{aligned} \quad (2.133)$$

In general, we can neglect the anisotropic stress, so the second equation implies  $\Phi \approx \Psi$ , so the last one reduces to:

$$\Phi'' + 3\mathcal{H}\Phi' + (2\mathcal{H}' + \mathcal{H}^2)\Phi = 4\pi Ga^2 \delta P \quad (2.134)$$

Furthermore, the first two equations can be combined to give:

$$\nabla^2 \Phi = 4\pi Ga^2 \bar{\rho} \Delta \quad (2.135)$$

which is basically the Poisson equation, but now valid on all scales.

In order to find the solutions to these equations, we need initial conditions. Since all the scales, at initial times, were bigger than the horizon, we will give these conditions working in the *superhorizon limit*. Taking in equations (2.124) and (2.126) the limit  $k \ll \mathcal{H}$ , we find:

$$\delta'_m = 3\Phi' \quad \delta'_r = 4\Phi' \quad (2.136)$$

which can be integrated to give:

$$\begin{aligned} \delta_\gamma &= 4\Phi + C_\gamma & \delta_\nu &= \delta_\gamma + S_\nu \\ \delta_c &= \frac{3}{4}\delta_\gamma + S_c & \delta_b &= \frac{3}{4}\delta_\gamma + S_b \end{aligned} \quad (2.137)$$

where the integration constants  $S_a$  are called *isocurvature modes*. A preferred set of initial conditions are given by the *adiabatic mode* with  $S_\nu = S_c = S_b = 0$  and  $C_\gamma \neq 0$ , which are a natural prediction of inflationary models and seem to be favored by observations. To find the constant  $C_\gamma$ , we look at the first Einstein equation (2.133) ignoring the anisotropic stress and taking the superhorizon limit:

$$3\mathcal{H}(\Phi' + \mathcal{H}\Phi) = -4\pi G(\bar{\rho}_\gamma \delta_\gamma + \bar{\rho}_\nu \delta_\nu) \quad (2.138)$$

where we dropped the matter contributions, negligible at early times. Specializing to the adiabatic mode and noticing that  $4\pi Ga^2(\bar{\rho}_\gamma + \bar{\rho}_\nu) = 3/2\mathcal{H}^2$  with  $\mathcal{H} = 1/\eta$ , we find:

$$\eta\Phi' + \Phi = -\frac{1}{2}\delta_\gamma \quad (2.139)$$

which becomes:

$$\eta\Phi'' + 4\Phi' = 0 \quad (2.140)$$

This equation has the growing solution  $\Phi = \text{const}$ , which implies then:

$$\delta_\gamma(\eta_i, \vec{k}) = -2\Phi_i(\vec{k}) \quad (2.141)$$

so that  $C_\gamma = -6\Phi_i$ , where  $\Phi_i$  indicates the value of the potential on superhorizon scales during matter radiation. We finally find that at initial times:

$$\delta_\gamma = \delta_\nu = \frac{4}{3}\delta_c = \frac{4}{3}\delta_b = -2\Phi_i \quad (2.142)$$

so that the initial conditions of the fluctuations are specified completely by the initial value of the potential  $\Phi$  or  $\Psi$  at early times. The value of the potential on superhorizon scales is constant only when the equation of state of the Universe is constant, but it varies between eras. In fact,  $\Phi$  obeys:

$$\Phi'' + 3(1+w)\mathcal{H}\Phi' - w\nabla^2\Phi = 0 \quad (2.143)$$

which shows that, on superhorizon scales it is constant. The value of this constant, however, changes between radiation domination and matter domination. It would be interesting to find a quantity that stays the same on all scales, despite the change of the equation of state. For adiabatic initial conditions, the conserved perturbations are the curvature perturbations introduced above, which in the Newtonian gauge are:

$$\begin{aligned} \zeta &= \Phi - \frac{\delta\rho}{3(\bar{\rho} + \bar{P})} \\ \mathcal{R} &= \Phi - \mathcal{H}\nu \end{aligned} \quad (2.144)$$

It is interesting to notice that, for a Universe dominated by a fluid with a constant equation of motion, on superhorizon scales we get:

$$\mathcal{R} \rightarrow \frac{5+3w}{3+3w}\Phi \quad (2.145)$$

Now, since we have said that  $\mathcal{R}$  remains the same through eras, we see that the value of  $\Phi$  during matter domination (on scales larger than the horizon) drops by a factor of  $9/10$  from its original value  $\Phi(\eta_i)$ . In other words, the density perturbations are frozen on large scales, and only when they enter the horizon do they begin to evolve, as a consequence of causality. For this very reason, the initial value of the perturbations are given at horizon entry, where they are proportional to the initial value of the potential  $\Phi_i$ . The constant of proportionality depends on whether the mode enters the horizon during radiation or matter domination, and whether we are considering matter or radiation perturbations.

From equation (2.145), we see that in radiation domination (superhorizon):

$$\Phi(\eta < \eta_{eq}, \vec{k}) = \frac{2}{3}\mathcal{R}_i(\vec{k}) \quad (2.146)$$

Therefore:

$$\delta_r = \frac{4}{3}\delta_m = -2\Phi = -\frac{4}{3}\mathcal{R}_i \quad (2.147)$$

So we can actually use the value of  $\mathcal{R}_i$  to specify initial conditions. This is relevant since the spectrum of the curvature perturbations can be generated via quantum fluctuations during inflation, which we take to be:

$$P_{\mathcal{R}}(\eta_i, k) = A_s k^{n_s - 4} \quad (2.148)$$

#### 2.2.4 The evolution of matter perturbations

We are now in a position to solve the Einstein equations (2.133) for the evolution of the pressureless matter fluctuations, relevant for the clustering of dark matter and for baryons after decoupling (before decoupling, the baryons are closely tied to the photons, forming a photon-baryon fluid).

First, we have already seen that the evolution of the density fluctuation is driven by the potential  $\Phi$ , therefore it is necessary to find its evolution in time. Analytical solutions for  $\Phi$  can only be found on superhorizon scales, or when either matter or radiation dominates (i.e. we have a constant equation of state  $w$ ); for times close to equality, we have to resort to numerical methods.

Proceeding this way, we know that when the equation of state is approximately constant in time,  $\Phi$  satisfies (2.143):

$$\Phi'' + 3(1+w)\mathcal{H}\Phi' - w\nabla^2\Phi = 0 \quad (2.149)$$

During matter domination we have  $w = 0$ , and the solution is:

$$\Phi(a, \vec{k}) = C_1(\vec{k}) + C_2(\vec{k})a^{-5/2} \quad \text{matter era} \quad (2.150)$$

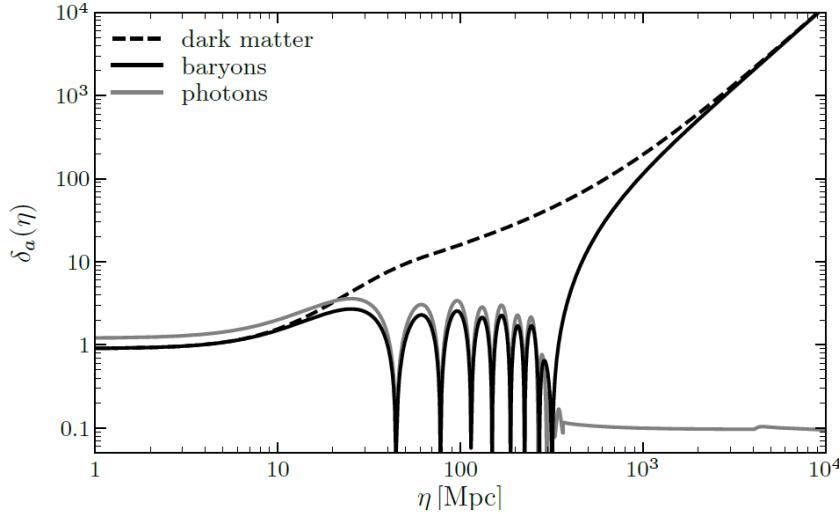


Figure 4: The evolution of the density perturbations for dark matter, baryons and photons, for a mode that entered the horizon in the radiation era  $k = 0.25 h \text{ Mpc}^{-1}$ . Note how dark matter gets a head start, whereas the photons and the baryons oscillate until decoupling. Figure taken from [56]

where we obviously recover the growing constant mode on superhorizon scales. In addition we notice, however, that the solution is valid on all scales, meaning that the potential is a constant on all scales during the matter era.

During radiation domination we have  $w = 1/3$  and the solution is:

$$\Phi(\eta, \vec{k}) = 2\mathcal{R}_i \frac{\sin \varphi - \varphi \cos \varphi}{\varphi^3} \quad \text{radiation era} \quad (2.151)$$

where  $\varphi = k\eta/\sqrt{3}$  and the normalization is fixed by the initial conditions  $\Phi(0, \vec{k}) = 2/3\mathcal{R}_i$ . We see that if a certain mode enters the horizon in the radiation era,  $\Phi$  oscillates with a frequency  $1/\sqrt{3}k$  and with a decaying amplitude  $\eta^{-2} \propto a^{-2}$ .

Having obtained these equations, we can find the evolution of the matter density contrast. During the matter era, since matter is the dominant component, the Poisson equation reads:

$$\nabla^2 \Phi \simeq 4\pi G a^2 \bar{\rho}_m \Delta_m \quad (2.152)$$

Using the evolution of the potential (2.150) and the fact that  $a^2 \bar{\rho}_m \propto a^{-1}$ , we find:

$$\Delta_m(a, \vec{k}) = -\frac{k^2 \Phi}{4\pi G a^2 \bar{\rho}_m} = \tilde{C}_1(\vec{k})a + \tilde{C}_2(\vec{k})a^{-3/2} \quad \text{matter era} \quad (2.153)$$

Notice the discrepancy on superhorizon scales, where the comoving density contrast  $\Delta_m$  grows with the scale factor  $a$ , whereas  $\delta_m$  stays constant. This is resolved by noticing that, inside the horizon, we have  $\delta_m \rightarrow \Delta_m$ , so there is no problem (this is a general feature: the density contrast is not something we can measure on superhorizon scales, therefore the gauge problem disappears).

During the radiation era, it is generally more difficult to find a solution. Since matter is the subdominant component, we cannot use the same trick as before, but instead we have to work with the continuity and Euler equations. We derived above equation (2.125):

$$\delta_m'' + \mathcal{H}\delta_m' = \nabla^2 \Phi + 3(\Phi'' + \mathcal{H}\Phi') \quad (2.154)$$

where in general  $\Phi = \Phi_m + \Phi_r$  is sourced by both radiation and matter, where the  $\Phi_m$  contribution is a constant, whereas the  $\Phi_r$  is rapidly oscillating inside the horizon. It turns out that the radiation mode is

suppressed compared to the constant mode, and therefore  $\delta_m$  is sourced by  $\Phi_m$  even deep in the radiation era. This reasoning brings us to the so-called *Meszaros equation*:

$$\frac{d^2\delta_m}{dy^2} + \frac{2+3y}{2y(1+y)} \frac{d\delta_m}{dy} - \frac{3}{2y(1+y)} \delta_m = 0 \quad (2.155)$$

where  $y = a/a_{eq}$ . The solutions are:

$$\delta_m \propto \begin{cases} 1 + \frac{3}{2}y & \\ (1 + \frac{3}{2}y) \log\left(\frac{\sqrt{1+y+1}}{\sqrt{1+y-1}}\right) - 3\sqrt{1+y} & \end{cases} \quad (2.156)$$

We see that in the limit  $y \ll 1$  (in radiation domination), the growing mode is logarithmic in  $a$ , i.e.  $\delta_m \propto \log y \propto \log a$ , whereas in the matter era  $y \gg 1$ , we have the majority of the growth, since  $\delta_m \propto y \propto a$ .

While we will not treat the evolution of the radiation perturbations, we give here a brief summary of the physics, represented in Figure 4. We can see that, initially, the fluctuations in all components are approximately equal. Before decoupling, the photons and the baryons act as a single fluid with pressure provided by the photons and mass density by the baryons, and they oscillate together. In the meantime, dark matter experiences a (slow) logarithmic growth during radiation domination, which gives it a head start for matter clustering, and then a power law growth after decoupling. After decoupling, the baryons lose the pressure support of the photons and their overdensities begin to grow, finally to match those of dark matter at late times. For the formation of galaxies, it is crucial that dark matter gets a head start, which aids the growth of baryon perturbations, otherwise the latter wouldn't cluster fast enough.

### 2.2.5 Statistical Properties

In this chapter, thus far, we have looked at the linear evolution of the individual Fourier modes of the density perturbation field  $\delta(\vec{x}, t) \equiv \rho(\vec{x}, t)/\bar{\rho}(t) - 1$ .

Here our goal is to describe the statistical properties of these perturbations, which are needed if one wants to relate the theory to observations.

In general, if our objective is to completely specify the density perturbation field  $\delta(\vec{x})$ , we would need to find the value of it at every point in space (equivalently find  $\delta(\vec{k})$  for every mode  $\vec{k}$ ). This process is, firstly, largely impractical, since we would need to specify an infinite number of values, and secondly, it is also unnecessary because the observed large structure of the Universe is a single realization of a random process, and therefore we need a description of the (over)density field only from a statistical standpoint. In this case, we should concern ourselves to identify the random process that generates the cosmic density field, rather than a certain realization of the field itself.

Following this philosophy, the statistical properties of a random perturbation field (at a fixed time), are specified by giving the probability for any particular realization of  $\delta(\vec{x})$ . Let's divide the Universe into  $n$  cells which are centered at  $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n$ , then the random perturbation field  $\delta(\vec{x})$  is characterized by the probability distribution function:

$$\mathbb{P}(\delta_1, \delta_2, \dots, \delta_n) d\delta_1 d\delta_2 \dots d\delta_n \quad (2.157)$$

where  $\delta_i \equiv \delta(\vec{x}_i)$ . This specifies the probability that the field  $\delta$  has values in the range  $\delta_i$  to  $\delta_i + d\delta_i$  at positions  $\vec{x}_i$  where  $i = 1, 2, \dots, n$ . This distribution function is completely determined if we know all its moments:

$$\langle \delta_1^{\ell_1} \delta_2^{\ell_2} \dots \delta_n^{\ell_n} \rangle = \int \delta_1^{\ell_1} \delta_2^{\ell_2} \dots \delta_n^{\ell_n} \mathbb{P}(\delta_1, \delta_2, \dots, \delta_n) d\delta_1 d\delta_2 \dots d\delta_n \quad (2.158)$$

where  $\ell_i$  are non-negative integers. Physically, the cosmological principle constrains the form of these moments, since it requires that the cosmological density field be statistically homogeneous and isotropic.

The moment of order 1 is, by definition, zero, because we are defining the overdensity with respect to the mean itself. The most important moment is the *2 point correlation function*:

$$\xi(|\vec{x}_1 - \vec{x}_2|) \equiv \langle \delta_1 \delta_2 \rangle \quad (2.159)$$

Note that it depends only on the distance between the two points, because of the requirement of isotropy and homogeneity of the underlying space, which is what we remarked above.

A natural comment on (2.158) is the following. The expectation value  $\langle \dots \rangle$  denotes an ensemble average of the stochastic process that created the random field  $\delta$ , which is the natural object to compute, but it is not what we actually observe. Instead, having only one Universe to work with, we are peering at just one of the infinite realizations of a random process. To relate the observations to the theoretical predictions, we need to assume *ergodicity*, which states that ensemble averages are equal to spatial averages if the averaging volume is large enough (in fact, it needs to be infinite for the equality to hold); in this case, different part of the Universe can be viewed as different realizations of the same random process, and the volume average approximates the average over different members of the ensemble.

The shortcoming of this philosophy is that in fact the Universe we observe is not infinite, so any survey of the sky will introduce some statistical fluctuations called *sample variance*.

Switching to Fourier space, we can see that:

$$\begin{aligned} \langle \delta(\vec{k}) \delta(\vec{k}') \rangle &= \int d^3x d^3x' e^{-i\vec{k} \cdot \vec{x}} e^{-i\vec{k}' \cdot \vec{x}'} \langle \delta(\vec{x}) \delta(\vec{x}') \rangle = \int d^3r d^3x' e^{-i\vec{k} \cdot \vec{r}} e^{-i(\vec{k} - \vec{k}') \cdot \vec{x}'} \xi(r) \\ &= (2\pi)^3 \delta_D(\vec{k} - \vec{k}') \int d^3r e^{-i\vec{k} \cdot \vec{r}} \xi(r) \equiv (2\pi)^3 \delta_D(\vec{k} - \vec{k}') P(k) \end{aligned} \quad (2.160)$$

where we introduced the separation  $\vec{r} \equiv \vec{x} - \vec{x}'$ . The delta function implies that Fourier modes with different wavelength are independent of each other. The important piece here is the so called *power spectrum*  $P(k)$ , which is just the three dimensional Fourier transform of the two point correlation function  $\xi(x)$ :

$$\begin{aligned} P(k) &= \int d^3r e^{-i\vec{k} \cdot \vec{r}} \xi(r) = \int_0^{2\pi} d\phi \int_{-1}^{+1} d(\cos \theta) \int_0^\infty dr r^2 e^{-ikr \cos \theta} \xi(r) \\ 2\pi \int_0^\infty dr \frac{r^2}{ikr} (e^{ikr} - e^{-ikr}) \xi(r) &= \frac{4\pi}{k} \int_0^\infty dr r \sin(kr) \xi(r) \end{aligned} \quad (2.161)$$

and, using the above:

$$\begin{aligned} \xi(r) &= \frac{1}{(2\pi)^3} \int d^3x P(k) e^{i\vec{k} \cdot \vec{r}} = \int \frac{dk}{k} \frac{k^3}{2\pi^2} P(k) \frac{\sin(kr)}{kr} \\ &= \int \frac{dk}{k} \Delta^2(k) \frac{\sin(kr)}{kr} \end{aligned} \quad (2.162)$$

where we introduced the *dimensionless power spectrum*:

$$\Delta^2(k) = \frac{k^3}{2\pi^2} P(k) \quad (2.163)$$

While the correlation function is the natural object to consider in observations, the power spectrum is easier to deal with and predict. However these two concepts are related, as we just saw.

The volume average of  $\xi(x)$  inside a sphere of radius  $x$  is defined as:

$$\bar{\xi}(x) \equiv \frac{\int d^3x' \xi(x')}{\int d^3x'} = \frac{3}{4\pi x^2} \int_0^\infty dx' \xi(x') 4\pi x'^2 \quad (2.164)$$

which is related to the power spectrum by:

$$\bar{\xi}(x) = \int_0^\infty \Delta^2(k) \frac{3(\sin kx - kx \cos kx)}{(kx)^3} \frac{dk}{k} \quad (2.165)$$

The window function in (2.165) dies off faster with increasing  $k$  compared to the one in (2.162), so that  $\tilde{\xi}$  gives a cleaner measure of the power spectrum at scales  $k \sim 1/x$  than does  $\xi$ :

$$\Delta^2(k) \simeq \tilde{\xi}(1/k) \quad (2.166)$$

Also,  $\tilde{\xi}$  can be interpreted as the characteristic squared overdensity within the radius  $x$ .

It is clear from this discussion that it is somewhat difficult to completely specify a general random field, since we would need to specify an infinite number of moments.

Luckily, inflation predicts that the initial conditions of the universe were, to a good degree of approximation, described by a *Gaussian random field*. This means that the probability distribution function (2.157) is:

$$\mathbb{P}(\delta_1, \delta_2, \dots, \delta_n) = \frac{1}{((2\pi)^n \det \xi)^{1/2}} \exp\left(-\frac{1}{2} \sum_{i,j} \delta_i \langle \delta_i \delta_j \rangle \delta_j\right) \quad (2.167)$$

We see therefore that all  $n$ -point functions are determined by functional integrals over  $\mathbb{P}$ , and thus are specified in terms of the two point function  $\langle \delta_i \delta_j \rangle$ . Basically, in the end we just need to know the two point function in order to compute everything else. In particular, the one point distribution function is:

$$\mathbb{P}(\delta) d\delta = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left(-\frac{\delta^2}{2\sigma^2}\right) d\delta \quad (2.168)$$

where  $\sigma^2 = \xi(0)$  is the variance of the density perturbation field.

Furthermore, as long as the evolution is linear, the statistics remain Gaussian.

### 2.2.6 Particle Sampling

In many applications, and in the present work, it is useful to represent the cosmic density field by a set of particles. This is useful, for instance, when one is considering the evolution of the 2 point function beyond the linear regime, where one can imagine that objects start to form clusters. In this section we consider *Poisson sampling*, a simple model to sample a continuous mass distribution with particles.

The simplest way to proceed is to divide the space into infinitesimal cells, and sample the perturbation field such that the number of particles in each cell has a Poisson distribution with a mean proportional to the mean density of particles. We also suppose that the dimension of each cell  $\Delta V$  is small enough that the probability for it to contain more than one particle is zero. In this situation, we need only specify  $p^{(1)}(\vec{x})$ , the probability that a certain cell located at position  $\vec{x}$  contains one single particle, which is given by the average occupancy of a cell:

$$p^{(1)}(\vec{x}) = (1 + \delta(\vec{x})) \bar{n} \Delta V \quad (2.169)$$

where  $\bar{n} = \bar{\rho}/m$  is the average number density of particles. In this way we can calculate:

$$\langle p^{(1)}(\vec{x}_i) p^{(1)}(\vec{x}_i + \vec{x}) \rangle = (\bar{n} \Delta V)^2 (1 + \xi(x)) \quad (2.170)$$

where the average is spatial, since we (always) assume ergodicity. This is the probability of finding a pair of particles in two arbitrary cells at comoving separation  $x$ . The right hand side, if the particles were distributed randomly, would be  $(\bar{n} \Delta V)^2$ , so (2.170) tells us that the two point correlation function  $\xi(x)$  is a measure of the excess in the number of pairs of particles at separation  $x$  over a random distribution. The conditional probability for finding two particles at comoving separation  $x$ , given that one cell contains a particle is:

$$\frac{\langle p^{(1)}(\vec{x}_i) p^{(1)}(\vec{x}_i + \vec{x}) \rangle}{\langle p^{(1)}(\vec{x}_i) \rangle} = (\bar{n} \Delta V)(1 + \xi(x)) \quad (2.171)$$

This is the mean number of neighbors each particle has when looking at a comoving distance  $x$ . Instead, the mean number of neighbors within a spherical shell that is centered on a particle and has proper radius  $r$  and thickness  $dr$  is:

$$dN(r) = 4\pi r^2 \bar{n} (1 + \xi(r)) dr \quad (2.172)$$

which, when integrated, gives the mean number of particles within a proper distance  $r$  from a specific particle:

$$N(x) = \bar{n}a^3 \int_0^x ds 4\pi s^2 (1 + \xi(s)) \quad (2.173)$$

where we switched from the proper radius to the comoving one  $r = ax$ . This formula (2.172) suggests the interpretation of  $\bar{\rho}(1 + \xi(r))$  as the mean density profile around each particle.

The concept we just introduced is important in this work, since it turns out that it will let us follow the evolution of the statistics of primordial black holes on smaller scale, where they can cluster.

## 2.3 Numerical Results and Fits For The Linear Theory

Up until now, we have seen that the density contrast (comoving or not) evolves in the following way:

$$\delta(\vec{k}, t) = C_1(\vec{k})D_+(t) + C_2(\vec{k})D_-(t) \quad (2.174)$$

where the factors  $D_{\pm}(t)$  describe the linear evolution. We refer to  $D_+(t)$  and  $D_-(t)$  as the growing and decaying modes, respectively. The growing mode is usually called the *linear growth function*, and it is normalized so that  $D(t_0) = 1$ .

We saw, in addition, that the matter perturbations inside the horizon grow as  $a$  during the matter era, while they experience a logarithmic growth during the radiation era; instead, outside the horizon they grow as  $a$  again during matter domination and as  $a^2$  during radiation domination. Since the moment of horizon entry depends on the wavenumber of the mode,  $k = aH$ , this leads to a  $k$  dependent growth. A convenient way to relate the initial conditions and the density perturbations that we observe in the post-recombination Universe, is through a *linear transfer function*,  $T(k)$ , defined as:

$$T(k) = \frac{D_+(t_i)}{D_+(t)} \frac{\delta(\vec{k}, t)}{\delta(\vec{k}, t_i)} \quad (2.175)$$

where  $t_i$  is some initial time after inflation (when all modes of interest were superhorizon), and it depends only on the magnitude of the wavevector as a consequence of the isotropic evolution. The transfer function is well fitted by [57]:

$$T(k) = \frac{\log(1 + 2.34q)}{2.34q} (1 + 3.89q + (16.1q)^2 + (5.46q)^3 + (6.71q)^4)^{-1/4} \quad (2.176)$$

where:

$$q \equiv \frac{1}{\Gamma} \left( \frac{k}{h \text{ Mpc}^{-1}} \right) \quad \text{and} \quad \Gamma = \Omega_m h \exp(-\Omega_b(1 + \sqrt{2h}/\Omega_m)) \quad (2.177)$$

and the growth factor in a Universe with matter and dark energy is:

$$D_{\text{ad}}(a) = \frac{5\Omega_m}{2} \frac{H(a)}{H_0} \int_0^a \frac{da'}{(a'H(a)/H_0)^3} \quad (2.178)$$

By defining:

$$\begin{aligned} \Omega_m(z) &= \frac{\Omega_m}{\Omega_{\Lambda,0}a^3 + \Omega_m} \\ \Omega_{\Lambda}(z) &= \frac{\Omega_{\Lambda}}{\Omega_{\Lambda} + \Omega_m/a^3} \end{aligned} \quad (2.179)$$

we get the following approximation for the growth factor [58]:

$$D_{\text{ad}}(z) = \frac{5\Omega_m(z)}{2} \left[ \Omega_m(z)^{-4/7} - \Omega_{\Lambda}(z) + \left( 1 + \frac{\Omega_m(z)}{2} \right) \left( 1 + \frac{\Omega_{\Lambda}(z)}{70} \right) \right]^{-1} \quad (2.180)$$

By using these formulas we can see that the power spectrum as defined by (2.160) is:

$$P(k, t) = T^2(k) \frac{D_+^2(t)}{D_+^2(t_i)} P(k, t_i) \quad (2.181)$$

where the transfer function satisfies  $T(k=0)=1$ , and the primordial matter power spectrum is recovered from (2.148):

$$P(k, t_i) = A_s k^{n_s} \quad (2.182)$$

The adiabatic power spectrum for the linear perturbations is therefore the following:

$$P_{\text{ad}}(k, a) = 2\pi^2 A_s \frac{k^{n_s} c^{n_s+3}}{H_0^{n_s+3}} T^2(k) \left( \frac{D_{\text{ad}}(a)}{D_{\text{ad}}(a=1)} \right)^2 \quad (2.183)$$

\*\*\*

## Chapter references

This chapter was written mainly following two textbooks, namely inspiration was taken from Baumann [56] and Mo, Van den Bosch and White [59]. The section about the problems of the  $\Lambda$ CDM model is a summary of [55], and finally the discussion regarding gauge transformations was taken from a series of lecture notes on perturbation theory by Kurki-Suonio [60].

### 3 Nonlinear Clustering

So far, our treatment has been purely concerned with the linear regime, where  $\delta \ll 1$ , however, many objects in the present-day Universe, such as galaxies and clusters of galaxies, have densities that are much higher than the average density of the Universe; it follows that if we want to understand the formation of these structures, we must go beyond the linear perturbation theory and explore the nonlinear territory, where  $\delta \gg 1$ .

In this section we thus address the problem of gravitational collapse of overdensities in the nonlinear regime. As a general rule, the nonlinear gravitational dynamics is difficult to deal with from an analytical standpoint, and as a consequence the last word has to be left to detailed computer  $N$ -body simulations; the problem with numerical simulations, however, is that they tend to obscure the basic physics contained in the equations, so that they act essentially as a "black box". Luckily, if certain assumptions about the symmetry of the problem are made, simple analytical models can be studied, which give ultimately some insights into the complex physics of gravitational clustering.

In the following we discuss a simple model, the spherical collapse model, which serves both as an example where we can follow analytically the nonlinear evolution, and as a bridge to the so called *Press-Schechter theory*, which utilizes the results of the spherical collapse model, and follow up with *excursion set theory*, which formalizes Press-Schechter and allows us to be even more precise.

Furthermore, we present a fairly sophisticated approach involving nonlinear scaling relations, which will allow us to track the two point correlation function into the quasi-linear and fully non-linear regimes of the clustering process of a system of particles.

#### 3.1 The Spherical Collapse Model

Consider a flat, matter dominated universe. The average density is equal to the critical density and evolves in time as:

$$\bar{\rho}(t) = \frac{1}{6\pi G} \frac{1}{t^2} \quad (3.1)$$

where we used the results for the Einstein-de Sitter Universe obtained in equation (2.43).

At some initial time  $t_i$  we create a spherically symmetric matter overdensity by compressing a region of radius  $R_i$  to one with a smaller radius  $r_i < R_i$ . Using the conservation of mass, we find that the initial density of the perturbation is:

$$\rho_i = \frac{\bar{\rho}_i R_i^3}{r_i^3} \equiv \bar{\rho}_i(1 + \delta_i) \quad (3.2)$$

Note that we just introduced the initial density contrast  $\delta_i$ , but it will not be restricted to values  $\delta_i \ll 1$ , like it was in the previous sections.

The important point in this problem is that we are requiring spherical symmetry, from which follow some simplifications. First, the evolution of the background is decoupled from that of the perturbation, so that the former is not influenced by the latter, and secondly we can think of the overdensity as composed of infinite thin shells. Our analysis will remain valid unless the shells cross with one another, in which case the mass contained in the shell is no longer constant, making our assumption (3.2) invalid. The shortcomings of this model will become clear below.

Let us now study the evolution of a single mass shell of radius  $r(t)$ , which evolves according to the Newtonian equation:

$$\frac{d^2 r}{dt^2} = -\frac{GM(r)}{r^2} \quad (3.3)$$

where  $M(r)$  is the mass enclosed in the shell. If we integrate this equation we get the conservation of energy:

$$\frac{1}{2} \left( \frac{dr}{dt} \right)^2 - \frac{GM}{r} = E \quad (3.4)$$

where  $E$  is the specific energy of the mass shell. Both  $E$  and  $M(r)$  are constant throughout the collapse. We will study here the case  $E < 0$ , so that the overdensity acts like a closed universe with positive curvature.

The solutions of these equations were given in parametric form in the matter and curvature solution in the first section:

$$\begin{aligned} r(\theta) &= A(1 - \cos \theta) \\ t(\theta) &= B(\theta - \sin \theta) \end{aligned} \quad (3.5)$$

with  $A$  and  $B$  constants to be determined by the initial conditions: take the solutions (3.5) at early times where  $\theta \ll 1$  and expand them to lowest order:

$$\begin{aligned} r(\theta) &= A(\theta^2/2 - \theta^4/24) \\ t(\theta) &= B(\theta^3/6 - \theta^5/120) \end{aligned} \quad (3.6)$$

If we insert them into equations (3.3) and (3.4) we get the following relations:

$$A^3 = GM B^2 \quad A = \frac{GM}{-2E} \quad (3.7)$$

Now, assume that at the initial time  $t_i \ll 1$  the radius and the velocity of the mass shell are  $r_i$  and  $v_i$ , so that the energy is  $E = v_i^2/2 - GM/r_i$ . If we insert this into the second expression of (3.7) we get:

$$\frac{r_i}{2A} = 1 - \frac{(v_i/H_i r_i)^2}{1 + \delta_i} \quad (3.8)$$

where  $H_i$  is the initial Hubble constant, and we used the fact that the average initial  $\delta_i$  is related to  $M$  and  $r_i$  by:

$$M = \rho_i \frac{4\pi}{3} r_i^3 = \bar{\rho}_i (1 + \delta_i) \frac{4\pi}{3} r_i^3 \quad (3.9)$$

where we have used (3.2). Since  $M$  is constant, we can write  $v_i$  as:

$$v_i = \frac{dr_i}{dt_i} = H_i r_i \left( 1 - \frac{1}{3H_i t_i} \frac{\delta_i}{1 + \delta_i} \frac{d \log \delta_i}{d \log t_i} \right) \quad (3.10)$$

At sufficiently early times  $t_i \ll 1$  and  $\delta_i \ll 1$ , we have the following scalings:

$$\begin{aligned} H_i t_i &\approx 2/3 \\ \delta_i &\approx t_i^{2/3} \\ \frac{v_i}{H_i r_i} &\approx 1 - \frac{\delta_i}{3} \end{aligned} \quad (3.11)$$

If we insert these into (3.7) and (3.8), we finally get the equations for  $A$  and  $B$  as a function of initial conditions:

$$A = \frac{3}{10} \frac{r_i}{\delta_i} \quad B = \frac{3}{4} \left( \frac{3}{5} \right)^{3/2} \frac{1}{\delta_i} \quad (3.12)$$

The motion of the mass shell is therefore specified, in a given cosmology, via the initial conditions on the radius  $r_i$  of the mass shell and the mean overdensity  $\delta_i$ .

The mass shell reaches the maximum expansion at  $\theta = \pi$  (the maximum of the function  $r(\theta)$ ) and then turns around. This means that the radius and time at maximum expansion are:

$$r_{\max} = 2A = \frac{6}{10} \frac{r_i}{\delta_i} \quad t_{\max} = B\pi = \frac{3\pi}{4} \left( \frac{3}{5} \right)^{3/2} \frac{1}{\delta_i} \quad (3.13)$$

To get an idea of how the shell evolves, let's calculate the overdensity and follow its evolution. The densities evolve as:

$$\begin{aligned} \rho(\theta) &= \frac{M}{(4\pi/3)r^3(\theta)} = \frac{3M}{4\pi A^3} \frac{1}{(1 - \cos \theta)^3} \\ \bar{\rho}(\theta) &= \frac{1}{6\pi G} \frac{1}{t^2(\theta)} = \frac{1}{6\pi G B^2} \frac{1}{(\theta - \sin \theta)^2} \end{aligned} \quad (3.14)$$

The ratio of the (3.14) determines the density contrast:

$$1 + \delta = \frac{\rho(\theta)}{\bar{\rho}(\theta)} = \frac{9}{2} \frac{(\theta - \sin \theta)^2}{(1 - \cos \theta)^3} \quad (3.15)$$

where we have used the first of (3.7). We can now consider three different stages of the evolution. Consider the *linear regime* at early times, where  $\delta$  is still really small. Expanding (3.15) to lowest order we get:

$$\begin{aligned} \delta &= \frac{9}{2} \frac{(\theta^3/6 - \theta^5/120)^2}{(\theta^2/2 - \theta^4/24)^3} - 1 \simeq \frac{9}{2} \frac{\theta^6}{36} \frac{8}{\theta^6} \frac{(1 - \theta^2/10)}{1 - \theta^2/4} - 1 \\ &\simeq 1 - \frac{\theta^2}{10} + \frac{\theta^2}{4} + \mathcal{O}(\theta^4) = \frac{3}{20} \theta^2 = \frac{3}{20} \left(\frac{6}{B}\right)^{2/3} t^{2/3} \equiv \delta_{\text{lin}}(t) \end{aligned} \quad (3.16)$$

which is just the linear growth function in a matter dominated universe, a result that we already found in the first chapter.

The *turn around point*,  $\theta = \pi$ , is characterized by:

$$\delta(\theta = \pi) = \frac{9\pi^2}{16} - 1 \approx 4.55 \quad (3.17)$$

It is also useful to ask what the extrapolated linear solution gives at the turn around point, albeit being just an artificial concept:

$$\delta_{\text{lin}}(t) = \frac{3}{20} (6\pi)^{2/3} (t/t_{\max})^{2/3} \implies \delta_{\text{lin}}(t_{\max}) \approx 1.06 \quad (3.18)$$

where we used the second equation of (3.13). This test, whether  $\delta_{\text{lin}}$  exceeds its limit of validity  $\delta_{\text{lin}} \ll 1$ , is a good test to judge by extrapolation if the perturbation is decoupled from the Hubble flow.

Finally, in the *collapse time*, the solution diverges:

$$\delta(\theta_{\text{coll}} = 2\pi) = \infty \quad (3.19)$$

Before discussing the divergence, let's again calculate the result for the extrapolated density contrast in the linear regime, when the matter has collapsed. Using  $t_{\text{coll}} = 2t_{\max} = 2B\pi$ , we find:

$$\delta_{\text{lin}}(t_{\text{coll}}) = \frac{3}{20} (12\pi)^{2/3} \approx 1.69 \quad (3.20)$$

We can therefore say that when the extrapolated linear density contrast reaches 1.69, the region will have collapsed.

The divergence we found in (3.19) is a symptom of having assumed perfect spherical symmetry for the initial perturbation. In fact, our treatment cannot be extended to arbitrarily small values of  $r$ . As the mass shell turns around and begins to collapse, particles that are in the mass shell in question can cross the mass shells that were originally inside it and therefore, as was said above, the mass conservation of each shell is not a valid assumption anymore. Indeed, by the time  $t = 2t_{\max}$ , all the mass shells have crossed path so many times that the object formed is considered a virialized halo.

In our cosmological model, the halos that form are the dark matter halos, which are the locations where the baryons finally collapse into and form galaxies. In order to estimate the real value of the density of these halos, we can use the virial theorem, which states that, for virialized objects, there is a relation between the average kinetic and potential energy:

$$T = -\frac{1}{2}V \quad (3.21)$$

If we pair this with the conservation of energy, we get that at the turn around point  $E = V_{\text{turn}}$ , so that after virialization:

$$T_{\text{vir}} + V_{\text{vir}} = \frac{1}{2}V_{\text{vir}} = V_{\text{turn}} \quad (3.22)$$

This means that the radius of the virialized object is halved from the turning point  $r_{\text{vir}} = r_{\max}/2$ , and the density is  $\rho_{\text{vir}} = 8\rho_{\text{turn}}$ . We can compare this to the background density at the time of virialization, which we take to be  $t_{\text{vir}} \approx t_{\text{col}} = 2t_{\max}$  for simplicity. Since  $\bar{\rho} \propto t^{-2}$ , this means that  $\bar{\rho}_{\text{vir}} = \bar{\rho}_{\text{turn}}/4$ . Finally, the desity of the object is:

$$1 + \delta_{\text{vir}} = \frac{\rho_{\text{vir}}}{\bar{\rho}_{\text{vir}}} = \frac{8\rho_{\text{turn}}}{\bar{\rho}_{\text{turn}}/4} = 32(1 + \delta_{\text{turn}}) = 18\pi^2 \approx 178 \quad (3.23)$$

where we have used (3.17). We found that the density of dark matter halos is therefore around 200 times greater than the average density of the Universe, a result that agrees with numerical simulations.

### 3.2 The Press-Schechter Formalism

In the standard  $\Lambda$ CDM paradigm of cosmological structure formation, the formation of galaxies begins with the gravitational collapse of the overdense regions into virialized halos of dark matter. Hereafter, the baryons, bound in the potential wells of the said halos, proceed to cool, condense and form galaxies (remember that, as seen in Table 1, the density of dark matter outweighs that of baryons by about six to one).

It is of fundamental importance to understand the formation, the properties and the abundance of these DM halos, if our goal is to ultimately grasp the workings of galaxies.

Press and Schechter (PS) [61] introduced a formalism which made use of the result we found above: whenever a density perturbation in linear theory exceeds the threshold  $\delta_c \simeq 1.69$ , a virialized halo will have formed.

The first step in this theory is to introduce a linearly evolved *smoothed density field* by averaging out the contributions below a certain scale  $R$  to the density contrast field, so that we obtain a new field, which we shall call  $\delta_R(t, \vec{x}) \equiv \delta(t, \vec{x}; R)$ , and is given by a convolution with a specified *window function* or *filter*  $W(\vec{x}, R)$  (which is assumed to be spherically symmetric):

$$\delta_R(t, \vec{x}) = \int d^3x' W(|\vec{x} - \vec{x}'|; R) \delta(t, \vec{x}') \quad (3.24)$$

so that in Fourier space it becomes a product:  $\delta_R(t, \vec{k}) = W(\vec{k}; R) \delta(t, \vec{k})$ .

The purpose of the window function is therefore to weigh in a specific manner (depending on the analytic form) the density field. The most common type of window function is a sphere in real space, the so called top-hat function:

$$W(r; R) = \begin{cases} \frac{3}{4\pi R^3} & r < R \\ 0 & r > R \end{cases} \quad (3.25)$$

with a Fourier transform:

$$W(k; R) = \frac{3}{(kR)^3} (\sin(kR) - kR \cos(kR)) \quad (3.26)$$

The smoothed field  $\delta_R(t, \vec{x})$  is then, in this case, the average density in spheres of radius  $R$  around the point  $\vec{x}$ . The problem with this choice in this window function, however, is that the jump in real space introduces power on all scales in Fourier space, so that it is useful to smoothen the boundary of  $W$ , which we do by, for instance, introducing a Gaussian window function:

$$W(x; R) = \frac{1}{(2\pi)^{3/2} R^3} \exp\left(-\frac{x^2}{2R^2}\right) \quad (3.27)$$

with a Gaussian Fourier transform:

$$W(k; R) = \exp\left(-\frac{k^2 R^2}{2}\right) \quad (3.28)$$

Any given window can be labeled by the mean mass contained in it  $M(R)$ , instead of the size  $R$ , where the relation between the two is given by:

$$M(R) \equiv \bar{\rho} V(R) \quad (3.29)$$

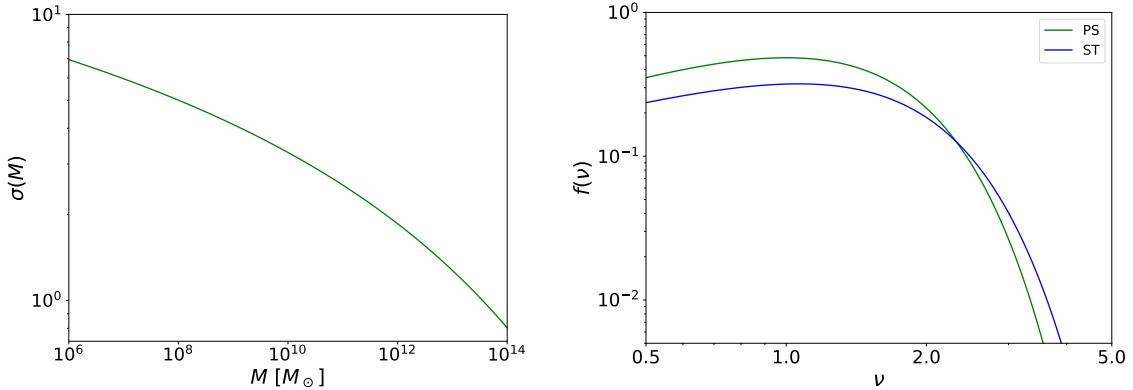


Figure 5: *Left:* The variance of the (smoothed) overdensity field as a function of mass  $M$  contained in the smoothing scale, introduced in (3.31). *Right:* Comparison between the Press-Schechter and the Sheth-Tormen halo multiplicities of eq.s (3.37) and (3.38).

where  $\bar{\rho}$  is the mean mass density, by convention evaluated at present time, since the mass of the collapsed object is conserved, and  $V(R)$  is the volume associated to the convolution window:

$$V(R) = \begin{cases} \frac{4\pi R^3}{3} & \text{Top-hat filter} \\ (2\pi)^{3/2} R^3 & \text{Gaussian filter} \end{cases} \quad (3.30)$$

Since the smoothed density fluctuation field  $\delta_R(t, \vec{x})$  itself a Gaussian random variable, being a sum of Gaussian random variables, in order to find its probability distribution we need only to know its mean (which is zero) and its variance:

$$\sigma^2(R, t) = \langle \delta_R(t, \vec{x})^2 \rangle = \int d \log k \Delta(t, k)^2 |W(k; R)|^2 \quad (3.31)$$

where  $\Delta(t, k)^2$  is the dimensionless matter power spectrum. Evaluating this variance for the specific scale  $R = 8h^{-1}$  Mpc, gives the parameter  $\sigma_8$  which is often used as a measure of the amplitude of the (linear) power spectrum. Its value measured today is roughly  $\sigma_8 \approx 0.811(6)$ . Figure 5 shows the variance of the smoothed density field as a function of the mass contained in the averaging volume, for the  $\Lambda$ CDM parameters given in Table 1, and for a Gaussian window function.

As we said above, the Press-Schechter formalism of galaxy formation assumes that the mass inside a region in which the smoothed density fluctuation is greater or equal than a critical value  $\delta_c$  (which generally depends on the redshift  $z$ ), corresponds to a virialized object with mass given by  $M(R)$ , which is given by (3.29) and (3.30). Using this fact we can thus calculate important statistical quantities, like the number of halos in a given mass range. Let  $n_h(t, \vec{x}, M)$  be the number of halos of mass  $M$  at position  $\vec{x}$  and at time  $t$ , and  $\bar{n}_h(t, M) \equiv \langle n_h(t, \vec{x}, M) \rangle$  be its mean value.

Since we said that the smoothed density field is a Gaussian random variable, the probability that a region of space with mass  $M$  (or inside a scale  $R$ ) has an overdensity  $\delta_M$  is therefore given by (2.168):

$$\mathbb{P}(\delta_M) = \frac{1}{\sqrt{2\pi}\sigma(M)^2} \exp\left(-\frac{1}{2} \frac{\delta_M^2}{\sigma(M)^2}\right) \quad (3.32)$$

where  $\sigma(M)^2$  was just introduced in (3.31), and the time dependence is suppressed. Now the probability for a region with overdensity  $\delta_M$  to exceed the critical value  $\delta_c$  is:

$$\mathbb{P}(\delta_M > \delta_c) = \int_{\delta_c}^{\infty} d\delta_M \mathbb{P}(\delta_M) = \int_{\nu}^{\infty} dx \exp\left(-\frac{x^2}{2}\right) = \frac{1}{2} \operatorname{erfc}\left(\frac{\nu}{2}\right) \quad (3.33)$$

where  $\text{erfc}(x)$  is the complementary error function, and  $\nu(M) \equiv \delta_c/\sigma(M)$  is the *peak height*, the height of the threshold in units of standard deviations of the smoothed density distribution. Note that in this model, the collapse of a region of mass  $M$  is defined so that it occurs when the smoothed density  $\delta_R(t, \vec{x})$  equals  $\delta_c$  on the appropriate (mass) scale.

Note that in the standard  $\Lambda\text{CDM}$  model, the variance  $\sigma(M)$  is a monotonically decreasing function of the scale  $M$ , which implies that small scale fluctuations are the first to collapse. This type of structure formation is called *bottom up*, because large-scale structure are formed when smaller and smaller objects coalesce into larger ones.

The Press-Schechter approach has a problem however, which is clear when looking at the result (3.33). In fact, in the limit  $R \rightarrow 0$ ,  $\tilde{\mathbb{P}}$  should give the fraction of all mass in virialized objects, however, since  $\sigma(M)$  becomes arbitrarily large when  $R \rightarrow 0$  and  $\text{erfc}(0) = 1$ , in this limit we get that only half of the mass collapses into halos, which is not consistent with numerical simulations. This issue arises because this approach does not account for the fact that while at a particular smoothing scale  $R$  the smoothed density field  $\delta_R(t, \vec{x})$  could be smaller than  $\delta_c$ , it may well be possible that it is bigger than the critical value at a bigger smoothing scale  $R' > R$ ; then it would seem natural that this larger volume should collapse to form a virialized object, and in the meanwhile swallow up the smaller parts within it. This effect clearly has the consequence of increasing our result in (3.33).

Press and Schechter "solved" this problem by multiplying (3.33) by a factor of 2,  $\tilde{\mathbb{P}} = 2\mathbb{P}$ , even though the reason why it should have been exactly a factor of 2 was far from convincing. In the literature, this issue of the existence of regions below a certain threshold on a particular scale, which are above threshold on a bigger scale, is called the *cloud-in-cloud* problem.

Some time later, Bond, Cole and Efstathiou and Kaiser [62] introduced an extension of the PS theory, using *excursion set theory*, that explained the factor of 2.

Multiplying by the factor of 2, one finds that the probability that a halo formed in the mass range  $[M, M+dM]$  is:

$$P([M, M+dM]) = |\tilde{\mathbb{P}}(\delta_{M+dM} > \delta_c) - \tilde{\mathbb{P}}(\delta_M > \delta_c)| \approx -\frac{d\tilde{\mathbb{P}}}{dM} \quad (3.34)$$

The abundance of halos of mass  $M$ , called the *halo mass function*, is then obtained by multiplying (3.34) by the maximum number of such halos in a region of mean density  $\bar{\rho}$ , which is given by  $\bar{\rho}/M$ :

$$\frac{d\bar{n}_h}{dM}(z, M) = -\frac{\bar{\rho}}{M} \frac{d\tilde{\mathbb{P}}}{dM} = -\sqrt{\frac{2}{\pi}} \nu(M) \exp\left(-\frac{\nu(M)^2}{2}\right) \frac{\bar{\rho}}{M^2} \frac{d \log \sigma(M)}{d \log M} \quad (3.35)$$

Note that this equation contains a time (redshift) dependence through  $\sigma(M)$ , where it is found in the power spectrum, as is written in (3.31). Using the  $\Lambda\text{CDM}$  linear matter power spectrum, we find that the mass function is a power law for small masses (high  $\nu$ ), while it has an exponential fall-off at large mass scales (low  $\nu$ ). Note that through this formalism, we can understand the way nonlinear structure develops: looking at (3.35), we see that halos with mass  $M$  can only form in significant number only when the argument in the exponential is not too large, so that  $\sigma(M) \gtrsim \delta_c$ ; in fact, by defining a characteristic mass  $M^*$  by:

$$\sigma(M^*, t) = \delta_c \quad (3.36)$$

only halos with mass  $M \lesssim M^*$  can have formed in large number before time  $t$ .

The function of the peak height  $\nu$  appearing in the mass function is called the *halo multiplicity*:

$$f_{PS}(\nu) = \sqrt{\frac{2}{\pi}} \nu \exp\left(-\frac{\nu^2}{2}\right) \quad (3.37)$$

The PS mass function captures all the essential features of structure formation, but it disagrees with the results of numerical  $N$ -body simulations, in particular it underpredicts the abundance of rare high-mass halos by about a factor of 10, and overpredicts that of low-mass halos by a factor of 2. One way to obtain a better fit with the simulations would be to realize that the PS formalism assumes that the collapse of density perturbations is described by the spherical collapse model. However, it can be shown that the collapse of overdensities in a Gaussian density field is in general ellipsoidal, rather than spherical.

Sheth and Tormen [63], considering these effects, proposed a new fitting function for the halo multiplicity:

$$f_{ST}(\nu) = A \sqrt{\frac{2}{\pi}} [1 + (a\nu^2)^{-p}] \sqrt{a\nu^2} \exp\left(-\frac{a\nu^2}{2}\right) \quad (3.38)$$

where  $A = 0.32$ ,  $a = 0.75$  and  $p = 0.3$ . Figure 5 shows a comparison between the halo multiplicities of Press-Schechter and Sheth-Tormen.

### 3.3 Excursion Set Theory

As was discussed in the previous section, the Press-Schechter derivation of the halo mass function (3.35) is plagued by a problem: it does not allow for the possibility that  $\delta_R(\vec{x})$  can be smaller than the critical density  $\delta_c$  at that smoothing scale  $R$ , but could be larger at a bigger scale  $R' > R$ . In this case it is only natural to assume that the larger volume  $R'$  would swallow up the smaller scale and form a virialized object. This problem was therefore solved by increasing, by a factor of 2, the probability in (3.33).

To really solve the "cloud-in-cloud" problem, as was obtained in [62] using what's called the *excursion set formalism*, we would need to compute the largest value of the smoothing scale  $R$  for which the  $\delta_R(\vec{x})$  exceeds the threshold  $\delta_c$ .

We first consider the smoothed density field  $\delta_R(\vec{x})$  now not as a function of the comoving coordinate  $\vec{x}$ , but as a function of the scale  $R$ , so that we can just write  $\delta(R)$ . For very large scales, we know that  $\delta_R(\vec{x}) \ll \delta_c$ , so the probability that the region lies above the boundary is vanishingly small. As we decrease  $R$ , the variance becomes larger and therefore the threshold will be crossed at some point, which is the first up-crossing of the barrier. The problem is therefore to compute the equation of motion of  $\delta(S)$ , where we denoted  $\sigma^2(R) \equiv S$ , and used the fact that the variance is a bijective function of the scale, so that we can interchange them freely. Note that each location  $\vec{x}$  now corresponds to a trajectory  $\delta(S)$ , which reflects the value of the smoothed density field at that point in space, when smoothed over a scale  $S$ .

Throughout the discussion it is important to remember that the variance is a (monotonically) decreasing function of the scale  $R$ , and as a consequence we have that bigger  $S$  means smaller  $M$  and viceversa. Therefore the problem is to calculate the first up-crossing between the values  $S$  and  $S + dS$ .

Consider a certain value  $S = S_1$  as the starting point, with  $\delta(S_1) \equiv \delta_1 < \delta_c$ . For a given change in  $S$  space, we may have a certain probability of reaching the value  $\delta_2$  after an increment  $\Delta S = S_2 - S_1 > 0$  (decreasing the scale  $R$  or  $M$ ), which in general depends both on the size of the step  $\Delta S$  and the value of  $\delta(S)$  on other scales. If the probability distribution that we are looking for depends on the value of the density field on other scales, the equations governing its motion are non-trivial, therefore we solve this problem by assuming a *sharp-k* smoothing window function, used to define (3.31):

$$W(k, R) = \Theta(1 - kR) \quad (3.39)$$

The problem with this sharp transition in Fourier space is that the integration of  $W(\vec{x}, R)$  over all space diverges, so it is not straightforward to associate to it a well defined volume; formally, we assign a volume as if we were working with a tophat filter in real space:

$$V(R) = \frac{4}{3} \pi R^3 \quad (3.40)$$

In any case, using this filter, we can see that decreasing the filter scale just corresponds to adding a set of independent Fourier modes to the smoothed density. These modes, of course, have not played a role in determining  $\delta(S)$  at other smoothing scales, thus the probability of a jump in density  $\Delta\delta$  associated with the jump in variance  $\Delta S$  is a Gaussian with zero mean and variance equal to  $\Delta S$ , independently of the starting point  $\delta_1$  in  $S$  space.

What one finds is the following: the trajectory  $\delta(S)$  as we change the smoothing scale executes a Brownian random walk, of which examples are shown below:

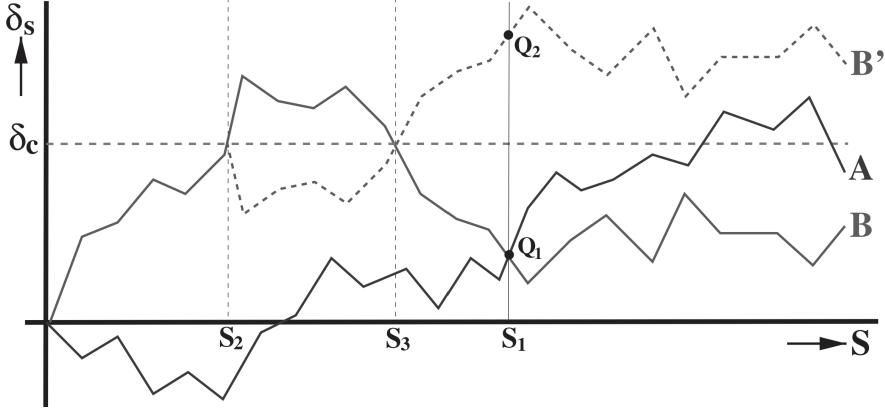


Figure 6: Gaussian random walks of the smoothed density field  $\delta_s$  as the mass scale  $S$  is varied, governed by the diffusion equation (3.45). Three relevant trajectories A, B and B' are show, where B' is the mirror of B from the point  $(\delta_c, S_2)$  onwards. Since the transition probability is a Gaussian, B and B' are equally probable.

In light of this, the goal of excursion set then becomes to calculate the largest smoothing scale  $R$ , or equivalently  $M$ , at which the random walk pierces the barrier  $\delta_c$ . This line of reasoning introduces naturally the missing factor of 2 from the Press-Schechter theory in the following way.

Consider the two random walks A and B in Figure 6 and a mass scale  $M_1$  with an associated variance  $S_1$ , where the continuous vertical line is positioned. The Press-Schechter ansatz assumes that the fraction of trajectories that have  $\delta_1 > \delta_c$  is the fraction of mass elements in collapsed objects with a mass  $M > M_1$ . The problem is that, for instance, the trajectory B is above the critical density in between the scales  $S_2$  and  $S_3$ , so that it should be found as part of a halo with mass  $M > M_3 > M_1$ . The Press-Schechter theory does not account for these type of trajectories, like A and B, and therefore some mass is missing, i.e. the factor of 2. A correction for this can be made with a heuristic approach, if one realizes that the trajectory from  $(\delta_2, S_2)$  to  $(\delta_1, S_1)$  is equally as likely as the trajectory B', the mirror of B for  $S \geq S_2$ , since the transition probability is a Gaussian which depends only on the size of the step, as we remarked above. This implies that each trajectory that is missed by the PS theory, corresponds to a mirrored trajectory whose smoothed density field is above the threshold at the desired mass scale, and therefore the fraction of mass in halos with mass greater than the said scale is given by twice the fraction of actual trajectories above threshold. This reasoning gives a natural explanation for the missing factor of 2.

More mathematically, we know that in the case of a sharp-k filter, the transition probability from  $\delta_1$  to  $\delta_2 = \delta_1 + \Delta\delta$  is:

$$\Pi(\delta_2, S_2)d\delta_2 = \Psi(\Delta\delta; \Delta S)d(\Delta\delta) \quad (3.41)$$

where the  $\Psi$  only depends on the size of the jump  $\Delta S$ , as remarked above:

$$\Psi(\Delta\delta; \Delta S)d(\Delta\delta) = \frac{1}{\sqrt{2\pi\Delta S}} \exp\left(-\frac{(\Delta\delta)^2}{2\Delta S}\right)d(\Delta\delta) \quad (3.42)$$

Since we now want to find the evolution equation of  $\Pi(\delta, S)$ , we start with the relationship between it and its value at a subsequent step  $\Pi(\delta, S + \Delta S)$ :

$$\Pi(\delta, S + \Delta S) = \int d(\Delta\delta)\Psi(\Delta\delta; \Delta S)\Pi(\delta - \Delta\delta, S) \quad (3.43)$$

Now we Taylor expand (3.43), keeping terms only up to  $(\Delta\delta)^2$  and integrating to find:

$$\frac{\partial\Pi}{\partial S} = \lim_{\Delta S \rightarrow 0} \left( \frac{\langle (\Delta\delta)^2 \rangle}{2\Delta S} \frac{\partial^2\Pi}{\partial\delta^2} - \frac{\langle \Delta\delta \rangle}{\Delta S} \frac{\partial\Pi}{\partial\delta} \right) \quad (3.44)$$

Using the fact that the transition probability is Gaussian with zero mean  $\langle \Delta\delta \rangle = 0$  and variance  $\langle (\Delta\delta)^2 \rangle = \Delta S$ , we simplify things further:

$$\frac{\partial \Pi}{\partial S} = \frac{1}{2} \frac{\partial^2 \Pi}{\partial \delta^2} \quad (3.45)$$

which is precisely the diffusion equation. In order to calculate the first up-crossing of  $\delta(S)$  at  $\delta_c$ , we wish to solve (3.45) for trajectories that have a starting point  $(\delta(S_0), S_0)$  and reach a point  $(\delta(S), S)$  without exceeding a critical value  $\delta_c$  at smaller  $S$ . This is equivalent to the diffusion equation with an absorbing barrier at  $\delta(S) = \delta_c$ . This problem has the following solution:

$$\Pi(\delta, S) = \frac{1}{\sqrt{2\pi\Delta S}} \left( \exp\left(-\frac{(\Delta\delta)^2}{2\Delta S}\right) - \exp\left(-\frac{(2(\delta_c - \delta_0) - \Delta\delta)^2}{2\Delta S}\right) \right) \quad (3.46)$$

where  $\Delta S = S - S_0$  and  $\Delta\delta = \delta - \delta_0$ . We see that the first term in (3.46) is the Gaussian probability distribution that represents the points above threshold at  $S$ , while the second term accounts for the trajectories that have been removed because they crossed above threshold at a bigger scale  $S' < S$ , but would have crossed back below by  $S$ , which are the trajectories that the Press-Schechter formalism fails to take into account.

At this point, from the discussion it is clear that the fraction of trajectories that have crossed the threshold at or prior to a certain scale  $S(M)$  is the complement of the sum of (3.46) until  $\delta_c$ :

$$\mathbb{P}(S) = 1 - \int_{-\infty}^{\delta_c} \Pi(\delta, S) d\delta = \text{erfc}\left(\frac{\delta_c - \delta_0}{\sqrt{2\Delta S}}\right) \quad (3.47)$$

Taking now the starting point to be  $(\delta_0, S_0) = (0, 0)$  at some large value of the smoothing scale, the above equation yields precisely the Press-Schechter mass function in the following way. The differential probability for a first piercing of the threshold is:

$$\begin{aligned} f(S) dS &\equiv \frac{d\mathbb{P}(S)}{dS} dS = \left( \int_{-\infty}^{\delta_c} \frac{\partial \Pi}{\partial S} d\delta \right) dS = \frac{1}{2} \frac{\partial \Pi}{\partial \delta} \Big|_{-\infty}^{\delta_c} dS \\ &= \frac{\delta_c}{\sqrt{2\pi S^{3/2}}} \exp\left(-\frac{\delta_c^2}{2S}\right) dS \end{aligned} \quad (3.48)$$

Now the fraction of mass in collapsed objects in a narrow interval of masses is found by changing the variable to  $M$ :

$$\frac{d\mathbb{P}(M)}{dM} = \frac{1}{\sqrt{2\pi S}} \frac{\delta_c}{S} \left| \frac{dS}{dM} \right| \exp\left(-\frac{\delta_c^2}{2S}\right) \quad (3.49)$$

The mass function is given by multiplying (3.49) by  $\bar{\rho}/M$ , yielding the Press-Schechter result, without the ad-hoc factor of two.

Note that in the excursion set formalism, the sharp-k filter Fourier window function only serves to simplify the calculations, but in principle other filters can be used at the cost of increasing mathematical complexity. As we noted above, this is because the steps in  $S$  space are not independent anymore, so it is necessary to compute the trajectory all at once to account for the correlations. It is important however to underline that the lack of correlations is not a prediction of excursion set theory, but only a simplifying assumption.

### 3.4 Halo Formation In the Excursion Set Theory

Excursion set theory is a powerful tool to calculate the properties of the progenitors which give rise to a given class of collapsed objects. For example, one can calculate the mass function at  $z = 5$  of these halos which by  $z = 0$  end up in a massive cluster-sized halo of mass  $10^{15} M_\odot$ . Basically, excursion sets allow us to track how a dark matter halo assembled its mass via mergers of smaller mass halos.

Since we need to make explicit the time dependence, we will write the evolution of the smoothed density field as  $\delta(S, a) = D(a)\delta(S, a = 1)$  where  $D(a)$  is the linear growth factor (2.178) (normalized so that  $D(a = 1) = 1$ ).

As we've seen in the last section, the formulas depend only on the ratio  $\delta_c/\sigma(M)$ , so that it is often easier to extract the growth factor from the rms variance and envision the height of the critical density threshold as a function of time. In that way, collapse at a scale factor  $a' \neq 1$  corresponds to the  $a = 1$  density fluctuation penetrating a barrier of height  $\delta_c(a') = \delta_c/D(a')$ .

While at first the possibility of using excursion sets to understand the formation histories of halos was discussed [62], only later was the problem studied more formally [64]; following the latter, we will write:

$$\omega(a) \equiv \delta_c/D(a) \quad (3.50)$$

Consider now a spherical region of mass  $M_1$  corresponding to a variance  $S_1 = \sigma^2(M_1)$ , and indicate  $\omega_1 = \omega(t_1) = \delta_c/D(t_1)$  so that this object collapses at time  $t_1$ . We are now interested in the fraction of  $M_1$  that was already in collapsed objects of a certain mass at an earlier time  $t_2 < t_1$ . To calculate this, we calculate the probability of a random walk originating at  $(\delta_1, S_1)$  to execute a first up-crossing of the barrier  $\omega(t_2)$  at  $S = S_2 > S_1$ , corresponding to a mass scale  $M_2 < M_1$ . This is exactly the same problem as above, although with a shift in the origin, which we now take to be  $(\delta_1, S_1)$ , hence the probability we are looking for is the following:

$$f(S_2, \omega_2 | S_1, \omega_1) dS_2 = \frac{1}{\sqrt{2\pi}} \frac{\Delta\omega}{\Delta S^{3/2}} \exp\left(-\frac{(\Delta\omega)^2}{2\Delta S}\right) dS_2 \quad (3.51)$$

where  $\Delta\omega = \omega_2 - \omega_1$  is the difference in barrier heights,  $\Delta S = S_2 - S_1$ . According to the excursion set theory, (3.51) is the conditional probability that a trajectory pierces the barrier  $\omega_2$  in an interval of width  $dS_2$  about  $S_2$  on the condition that the trajectory first pierces  $\omega_1$  at  $S_1$ , or in other words, it gives the fraction of mass elements in  $(\delta_1, M_1)$  that were in collapsed objects of mass  $M_2$  at the earlier time  $t_1$ .

Note that in this section, the interpretation of (3.51) is different from the one we gave above, where the initial condition was fixed at a large scale  $S_1 \ll S_2$ . Here, the two barriers represent the critical density at two different times, so that the change in barrier height  $\Delta\omega$  represents a shift in time.

### 3.4.1 The Conditional Mass Function

Using (3.51), a quantity of interest that descends from it is the *conditional mass function*. Given a halo of mass  $M_1$ , one can obtain the average number of progenitors at  $t_2$  in the mass range  $(M_2, M_2 + dM_2)$  that have merged with it by the time  $t_1$ . This is done by converting (3.51) from mass weighting to number weighting:

$$\frac{dn(M_2, \omega_2 | M_1, \omega_1)}{dM_2} dM_2 = \frac{M_2}{M_1} f(S_2, \omega_2 | S_1, \omega_1) \Big| \frac{dS_2}{dM_2} \Big| dM_2 \quad (3.52)$$

The following figures show the conditional mass function for standard  $\Lambda$ CDM power spectra, for four halo masses at three different redshifts  $z_2 = 1, z_2 = 1/4, z_2 = 3$ :

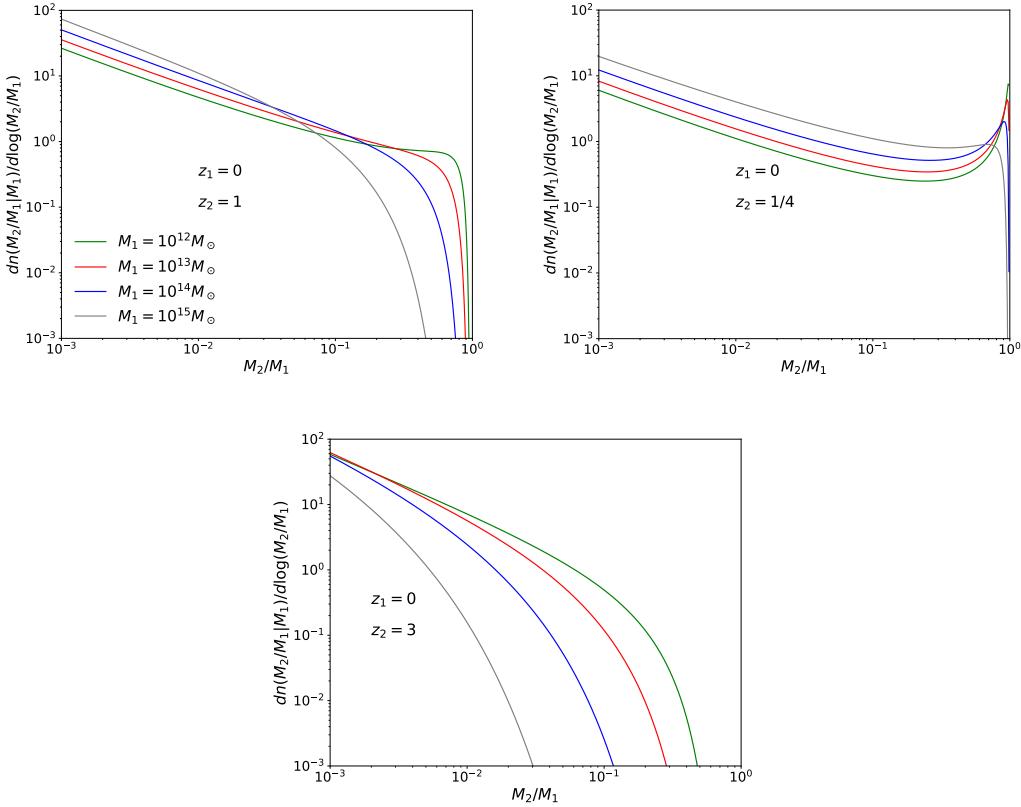


Figure 7: The conditional mass function of equation (3.52) at three different redshifts, for four different halo masses

The evolution towards more fragmentation as the redshift  $z_2$  of the smaller objects increases is evident in Figure 7. Also notice that more massive halos tend to fragment into smaller sub-units more quickly. The progenitor mass functions (3.52) have been tested by numerical simulations, which show that although the agreement is good when the time difference  $\Delta\omega$  between the epochs is small, the conditional mass function significantly underestimates the mass fraction in high mass progenitors for relatively large  $\Delta\omega$  (see for instance [65]).

### 3.4.2 Halo Accretion Rates

Another possibility is that of following the growth of a dark matter halo in time through mergers/accretion from lower mass objects. In order to do this, we ask what is the probability that some object with a certain mass will be incorporated into a bigger halo at some later time.

Therefore, we clearly need to ask the opposite question compared to the previous section, so that we start with an already existing object of mass  $M_2$  at scale  $S_2$  (and formed at time  $\omega_2$ ), and look for the posterior probability:

$$\begin{aligned} f(S_1, \omega_1 | S_2, \omega_2) dS_1 &= \frac{f(S_2, \omega_2 | S_1, \omega_1) f(S_1, \omega_1)}{f(S_2, \omega_2)} dS_1 \\ &= \frac{1}{\sqrt{2\pi}} \frac{\omega_1(\omega_2 - \omega_1)}{\omega_2} \left( \frac{S_2}{(S_1(S_2 - S_1))} \right)^{3/2} \exp\left(-\frac{(\omega_1 S_2 - \omega_2 S_1)}{2S_1 S_2 (S_2 - S_1)}\right) dS_1 \end{aligned} \quad (3.53)$$

where  $f(S, \omega)$  is given by (3.48). According to the formalism we developed, this formula is the conditional probability that a halo of mass  $M_2$  at time  $t_2$  is incorporated into a halo with a mass between  $M_1 > M_2$  and

$M_1 + dM_1$  at a later time  $t_2 > t_1$ . If we then set  $M_1 \equiv M_2 + \Delta M$  and  $t_1 = t_2 + \Delta t$ , equation (3.53) gives the probability for the halo to gain a mass  $\Delta M$  by merging/accretion in the time interval  $\Delta t$ .

Finally, the rate at which a halo with initial mass  $M_2$  transits to a halo with mass between  $M_2$  and  $M_1 = M_2 + \Delta M$  per unit  $\Delta M$  and per unit Hubble time is:

$$\frac{d^2R}{d \log \Delta M d \log a} = \sqrt{\frac{2}{\pi}} \frac{\Delta M}{M_1} \frac{\omega/\sigma(M_1)}{(1-S_1/S_2)^{3/2}} \exp\left(-\frac{\omega^2(S_2-S_1)}{2S_1S_2}\right) \left| \frac{d \log \omega}{d \log a} \right| \left| \frac{d \log \sigma}{d \log M_1} \right| \quad (3.54)$$

where we took the limit  $\omega_2 - \omega_1 \rightarrow 0$  to give a quantity per unit time, i.e. the rate. Note that a change in mass  $\Delta M$  can be, in principle, due to a cumulative effect of multiple mergers. However, the accretion rate is only interested in infinitesimal shifts in time, so that the change in mass is most likely due to a single event.

Thus the formula (3.54), the *halo accretion rate*, gives the number of mergers per unit of Hubble time of a halo with mass  $M_2$  with a smaller object with mass  $M_1$ . A quantity of interest is also (3.54), but weighted by the shift in mass we are considering, the *mass-weighted halo accretion rate*, which is basically the rate of mass increase.

The accretion rates and the mass accretion rates for the standard  $\Lambda$ CDM cosmology are shown in the following figures:

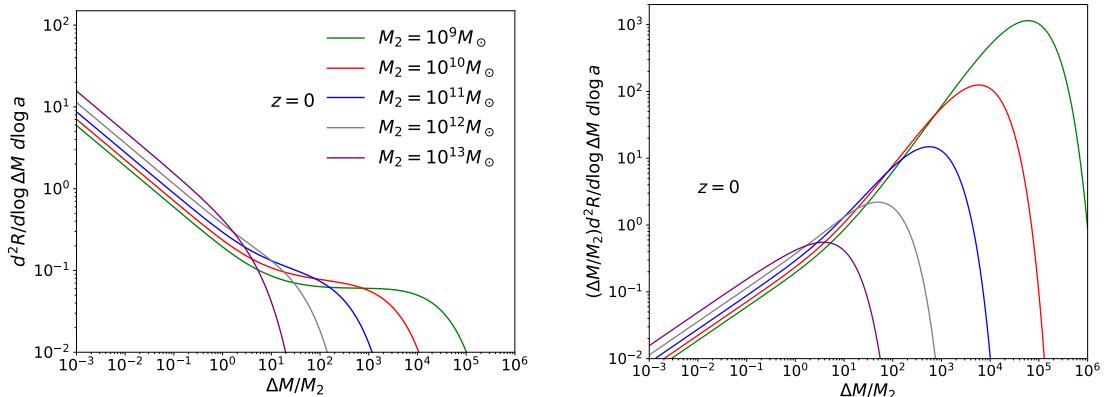


Figure 8: Halo accretion rates per logarithmic interval of mass change and Hubble time for five different initial masses at  $z = 0$ . *Left*: the probability for a change in  $\Delta M$  as a function of the mass shift itself as given in (3.54). *Right*: the fractional mass accretion rate given by (3.54) times  $\Delta M/M_2$ .

Note the following points. The halo mass accretion rate on the left of Figure 8 diverges for small increases, which means that the smaller mergers are most frequent in general. However, the average mass increase rate converges at  $\Delta M \rightarrow 0$  and is dominated by the more infrequent high mass mergers, i.e. the halos get most of their mass increase by these larger mergers compared to accreting smaller objects, albeit in very large quantities.

Finally, the *total mass accretion rate* is defined by integrating (3.54) over  $\log \Delta M$ , which gives the total mass increase per unit of Hubble time of an object with initial mass  $M_2$ . The following figure computes this integral and shows it in unit of halo mass and at three different redshifts  $z = 0$ ,  $z = 0.5$  and  $z = 1$ :

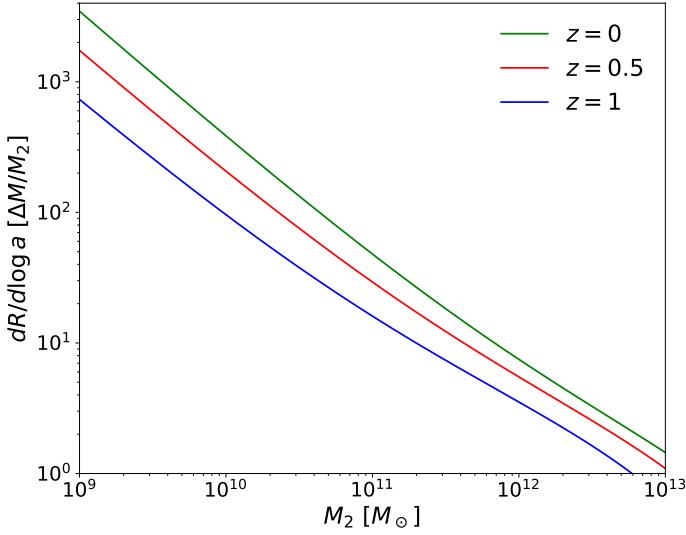


Figure 9: Total mass increase rate, defined as the integral of (3.54) over  $\log \Delta M$ , in units of halo mass as a function of initial masses for three different redshifts  $z = 0, z = 0.5, z = 1$

Figure 9 shows clearly that smaller halos increase their mass many times over the span of a Hubble radius, but this increase gets attenuated as we get farther from the present epoch (increasing redshift). An inclusive interpretation of Figures 8 and Figure 9 is that halos residing in the smaller mass scales tend to increase their mass more rapidly, mostly through mergers with larger objects, whereas bigger halos accrete more slowly and predominantly objects that are smaller in size.

### 3.5 Quasi-Linear and Non-Linear scaling relations

So far the evolution of the perturbations has been concerned with the density contrast  $\delta(\vec{x})$  or its Fourier transform, and from it we derived the statistical measures we wanted, like  $\xi(x, z)$  and the power spectrum  $P(z, k)$ . Alternatively, we can also study the evolution of the statistics themselves and work with individual particles, which we do in this section.

We know that given an initial density contrast, we can easily obtain the evolved one in linear theory via a growth function. If there is some way to relate the nonlinear and linear density contrasts, then we can start to understand how phenomena like clustering work and evolve. We can make an ansatz in the following fashion.

Let  $v_{\text{rel}}(a, x)$  be the relative pair velocity of particles separated by a comoving distance  $x$  at a specific epoch  $a$ , averaged over the entire universe, which we take as a measure of gravitational clustering at that scale and time. Let:

$$h(a, x) \equiv -\frac{v_{\text{rel}}(a, x)}{\dot{a}x} \quad (3.55)$$

This is the ratio between the relative pair velocity of two particles and the Hubble flow velocity. We can distinguish two limits immediately to find a solution for (3.55) in terms of  $\xi$ .

In the extreme nonlinear regime ( $\xi \gg 1$ ), the structures do not expand with the Universe and they form a closed system, which is true if the relative pair velocity of particles mostly balances the Hubble velocity  $v_{\text{rel}}(a, x) \simeq -\dot{a}x$ , so  $h(a, x) \simeq 1$ .

The behavior of  $h(a, x)$  in the linear regime ( $\xi \ll 1$ ) is a bit more complicated. Let's first define the peculiar

velocity field  $\vec{v}(\vec{x})$ , so that the relative pair velocity, at comoving distance  $\vec{r} = \vec{x} - \vec{y}$ , is:

$$\begin{aligned}\vec{v}_{\text{rel}} &\equiv \langle (\vec{v}(\vec{x}) - \vec{v}(\vec{y}))(1 + \delta(\vec{x}))(1 + \delta(\vec{y})) \rangle \\ &\simeq \langle (\vec{v}(\vec{x}) - \vec{v}(\vec{y}))\delta(\vec{x}) \rangle + \langle (\vec{v}(\vec{x}) - \vec{v}(\vec{y}))\delta(\vec{y}) \rangle\end{aligned}\quad (3.56)$$

to first order, and  $\langle \vec{v}(\vec{x}) - \vec{v}(\vec{y}) \rangle = 0$ . Now denote  $\vec{v}_{xy} \equiv \vec{v}(\vec{x}) - \vec{v}(\vec{y})$  and  $\vec{x} \equiv \vec{y} + \vec{r}$ . The radial component of the relative velocity is:

$$\vec{v}_{xy} \cdot \vec{r} = \int \vec{v}(\vec{k}) \cdot \vec{r} (e^{i\vec{k} \cdot (\vec{r} + \vec{y})} - e^{i\vec{k} \cdot \vec{y}}) \frac{d^3 k}{(2\pi)^3} \quad (3.57)$$

From (2.122) we know that this quantity is related to  $\delta(\vec{k})$  by:

$$\vec{v}(\vec{k}) = iHa \left( \frac{\delta(\vec{k})}{k^2} \right) \vec{k} \quad (3.58)$$

Putting this in (3.57) and writing the density contrasts in Fourier space we have:

$$\begin{aligned}\vec{v}_{xy} \cdot \vec{r}(\delta(\vec{x}) + \delta(\vec{y})) &= \\ iHa \int \frac{d^3 k}{(2\pi)^3} \int \frac{d^3 p}{(2\pi)^3} \left( \frac{\vec{k} \cdot \vec{r}}{k^2} \right) \delta(\vec{k}) \delta^*(\vec{p}) e^{i(\vec{k}-\vec{p}) \cdot \vec{y}} (e^{i\vec{k} \cdot \vec{r}} - 1) (e^{-i\vec{p} \cdot \vec{y}} + 1) &\end{aligned}\quad (3.59)$$

Taking the average of this expression, and remembering the definition of the power spectrum  $P(k)$  (2.160):

$$\begin{aligned}\vec{v}_{\text{rel}} \cdot \vec{r} &\equiv \langle \vec{v}_{xy} \cdot \vec{r}(\delta(\vec{x}) + \delta(\vec{y})) \rangle \\ &= iHa \int \frac{d^3 k}{(2\pi)^3} \frac{P(k)}{k^2} (\vec{k} \cdot \vec{r}) (e^{i\vec{k} \cdot \vec{r}} - e^{-i\vec{k} \cdot \vec{r}}) \\ &= -2Ha \int \frac{d^3 k}{(2\pi)^3} \frac{P(k)}{k^2} (\vec{k} \cdot \vec{r}) \sin(\vec{k} \cdot \vec{r})\end{aligned}\quad (3.60)$$

The symmetries of the problem in fact require that  $\vec{v}_{\text{rel}}$  be in the direction of  $\vec{r}$ , so  $\vec{v}_{\text{rel}} \cdot \vec{r} = v_{\text{rel}} r$ . Integrating (3.60) in the angular variables we obtain:

$$v_{\text{rel}} r = \frac{Ha}{r\pi^2} \int_0^\infty \frac{dk}{k} P(k) (kr \cos kr - \sin kr) \quad (3.61)$$

but we have an expression for the averaged two point function that closely resembles this (2.165), and in fact the two can be related:

$$v_{\text{rel}} r = -\frac{2}{3} (Har^2) \bar{\xi} \quad (3.62)$$

Now we can remember that in the definition (3.55), where the denominator can be rewritten in terms of the comoving separation  $r$  by  $Hr_{\text{prop}} = Har$  (we switch momentarily notation from  $x$  to  $r_{\text{prop}}$  to denote the proper separation between the particles). Then:

$$v_{\text{rel}}(r)r = -\frac{v_{\text{rel}}(r)}{Hr_{\text{prop}}} = -\frac{v_{\text{rel}}(r)}{Har} = \frac{2}{3} \bar{\xi} \quad (3.63)$$

This calculation was performed in order to show that our function  $h(a, x)$  depends on the pair  $(a, x)$  only through the averaged two point function  $\bar{\xi}(a, x)$  in the linear limit, while  $h \approx -1$  in the nonlinear limit. This suggests the following ansatz: the function  $h$  depends on the epoch  $a$  and the separation  $x$  only through a measure of the density contrast, which we shall take as  $\bar{\xi}$ , since we just found the illuminating result (3.63). In plain words, we are assuming  $h(a, x) = h[\bar{\xi}(a, x)]$ . By finding this functional relation between the two functions, it turns out that we can connect the linear regime  $\xi_L$  to the quasilinear  $\xi_{QL}$  and nonlinear  $\xi_{NL}$  regimes.

The derivation starts with the mean number of neighbors (2.173), from which we can say that the average

particle flux through a spherical shell of radius  $x$  centered on a given particle is  $4\pi\bar{n}a^2x^2(1+\xi(x,a))v_{\text{rel}}(x)$  (the  $\bar{n}(t)$  is the average particle number density at the epoch  $t$ , so that it is dependent on time. The combination  $\bar{n}(t)a^3$  is therefore time independent). The conservation law for the number of particles is therefore naturally:

$$\frac{\partial N}{\partial t} + 4\pi\bar{n}a^2x^2(1+\xi(x,a))v_{\text{rel}}(x) = 0 \quad (3.64)$$

Taking the partial derivative with respect to  $x$  of this equation, we find:

$$\frac{\partial \xi}{\partial t} + \frac{1}{ax^2}\frac{\partial}{\partial x}(x^2(1+\xi)v) = 0 \quad (3.65)$$

where it is understood that  $v$  is  $v_{\text{rel}}$ . This equation is based on the simple assumption of pair conservation and applies both in the linear and nonlinear regimes. From the definition (2.164), we can derive the following relation:

$$\begin{aligned} \frac{1}{3x^2}\frac{\partial}{\partial x}(x^3(1+\bar{\xi})) &= \\ &= \frac{1}{3x^2}(3x^2(1+\bar{\xi}) + x^3\frac{\partial}{\partial x}\left(\frac{3}{4\pi x^3}\int_0^x dy 4\pi y^2\xi(y)\right)) \\ &= 1 + \bar{\xi} + \frac{1}{3x^2}(x^3\left(-\frac{9}{4\pi x^4}\int_0^x dy 4\pi y^2\xi(y)\right) + 3x^2\xi(x)) \\ &= 1 + \bar{\xi} + \frac{1}{3x^2}\left(\left(-\frac{9}{4\pi x}\int_0^x dy 4\pi y^2\xi(y)\right) + 3x^2\xi(x)\right) \\ &= 1 + \bar{\xi} - \underbrace{\frac{3}{4\pi x^3}\int_0^x dy 4\pi y^2\xi(y)}_{=\bar{\xi}(x)} + \xi(x) = 1 + \xi(x) \end{aligned} \quad (3.66)$$

So we found:

$$\frac{1}{3x^2}\frac{\partial}{\partial x}(x^3(1+\bar{\xi})) = 1 + \xi(x) \quad (3.67)$$

If we put this in (3.65), we get:

$$\frac{1}{3x^2}\frac{\partial}{\partial x}(x^3\frac{\partial}{\partial t}(1+\bar{\xi})) = -\frac{1}{ax^2}\frac{\partial}{\partial x}\left(\frac{v}{3}\frac{\partial}{\partial x}(x^3(1+\bar{\xi}))\right) \quad (3.68)$$

By integrating over  $x$  (and by setting the arbitrary function of  $t$  to zero that would appear, in order to reproduce correct limiting behavior), we get:

$$x^3\frac{\partial}{\partial x}(1+\bar{\xi}) = -\frac{v}{a}\frac{\partial}{\partial x}(x^3(1+\bar{\xi})) \quad (3.69)$$

and switching variables  $t \rightarrow a$ , as in  $\partial/\partial t = \dot{a}\partial/\partial a$  brings this to:

$$x^3\dot{a}\frac{\partial}{\partial a}(1+\bar{\xi}) = -\frac{v}{a}\frac{\partial}{\partial x}(x^3(1+\bar{\xi})) \quad (3.70)$$

so in the end:

$$a\frac{\partial}{\partial a}(1+\bar{\xi}(a,x)) = -\left(\frac{v}{\dot{a}x}\right)\frac{1}{X^2}\frac{\partial}{\partial x}(x^3(1+\bar{\xi}(a,x))) \quad (3.71)$$

or, using the definition of  $h$  (3.55):

$$\left(\frac{\partial}{\partial \log a} - h\frac{\partial}{\partial \log x}\right)(1+\bar{\xi}) = 3h(1+\bar{\xi}) \quad (3.72)$$

This equation (3.72) shows clearly that the behavior of the average correlation function  $\bar{\xi}(a, x)$  is essentially controlled by  $h(a, x)$ . As we said above, we are now going to make the following assumption:

$$h(a, x) = h[\bar{\xi}(x, a)] \quad (3.73)$$

which allows us to actually find a solution to (3.72) with the following procedure. Let  $A = \log a$ ,  $X = \log x$  and  $D(X, A) = (1 + \bar{\xi})$ . We define curves in this space of new variables which satisfy:

$$\frac{dX}{dA} \Big|_c = -h[D(X, A)] \quad (3.74)$$

which means that the tangent to the curve at any point  $X, A, D$  is constrained by the value of  $h$  at that point. In this new space the l.h.s of (3.72) is a total derivative, so we can rewrite the differential equation as:

$$\left( \frac{\partial D}{\partial A} - h(D) \frac{\partial D}{\partial X} \right)_c = \left( \frac{\partial D}{\partial A} + \frac{\partial D}{\partial X} \frac{dX}{dA} \right)_c \equiv \frac{dD}{dA} \Big|_c = 3hD \quad (3.75)$$

By integrating:

$$\exp\left(\frac{1}{3} \int \frac{dD}{Dh(D)}\right) = \exp A + c \propto a \quad (3.76)$$

by squaring and determining the constant  $c$  from initial conditions at  $a_0$ , we get:

$$\exp\left(\frac{2}{3} \int_{\bar{\xi}(a_0, l)}^{\bar{\xi}(a, x)} \frac{d\bar{\xi}}{h(\bar{\xi})(1 + \bar{\xi})}\right) = \frac{a^2}{a_0^2} = \frac{\bar{\xi}_L(a, l)}{\bar{\xi}_L(a_0, l)} \quad (3.77)$$

Here  $l$  is the scale at the initial time, where the regime is still linear. Now we need to relate the scales  $l$  and  $x$ . Equation (3.74) can be written, using (3.75) in the following way:

$$\frac{dX}{dA} = -h = \frac{1}{3D} \frac{dD}{dA} \quad (3.78)$$

giving:

$$3X + \log D = \log(x^3(1 + \bar{\xi})) = \text{constant} \quad (3.79)$$

but using the initial conditions in the linear regime:

$$x^3(1 + \bar{\xi}) = l^3 \quad (3.80)$$

This shows that  $\bar{\xi}_L$  should be evaluated at  $l = x(1 + \bar{\xi})^{1/3}$ , therefore, as long as the evolution is linear  $\bar{\xi} \ll 1$  and  $R \sim x$ , but when clustering develops  $\bar{\xi} \gg 1$  the scale  $x$  is smaller than the initial one  $l$ . The end result is thus:

$$\bar{\xi}_L(a, l) = \exp\left(\frac{2}{3} \int_{\bar{\xi}_L(a_0, l)}^{\bar{\xi}(a, x)} \frac{d\mu}{h(\mu)(1 + \mu)}\right) \quad l = x(1 + \bar{\xi}(a, x))^{1/3} \quad (3.81)$$

So, given the function  $h(\bar{\xi})$ , we can find a relationship between the linear and the nonlinear regimes, i.e. the mapping  $\bar{\xi}(a, x) = U[\bar{\xi}_L(a, l)]$ . The lower limit in the integral in (3.81) is chosen in order to give the result  $\log \bar{\xi}$  for small values of  $\bar{\xi}$ .

By solving (3.81) for an arbitrary scale  $x$ , one can obtain the mapping between  $\bar{\xi}_L$  and  $\bar{\xi}$  in the nonlinear regime. However, there is an easier way to obtain the form of this mapping, along the following line of thought. Obviously in the linear regime we have  $U[\bar{\xi}_L] \approx \bar{\xi}_L$ . To determine  $U$  in the quasi-linear regime, consider a region surrounding a density peak in the linear stage, around which clustering takes place. We argued above that from (2.172) it follows that the density profile around this peak is proportional to  $1 + \xi$ :

$$\rho(x) \approx \bar{\rho}(1 + \xi(x)) \quad (3.82)$$

which says that the initial mean density contrast scales with the initial shell radius  $l$  as  $\bar{\delta}_i(l) \propto \bar{\xi}_L(l)$ . Using the results for the collapsing spherical shell model (3.13), the shell will expand to a maximum radius of

$x_{\max} \propto l/\bar{\delta}_i \propto l/\bar{\xi}_L(l)$ . Now taking the final effective radius as proportional to  $x_{\max}$ , the final average correlation function will be:

$$\bar{\xi}_{QL}(x) \propto \rho \propto \frac{M}{x_{\max}} \propto \frac{l^3}{(l^3/\bar{\xi}_L(l))^3} \propto \bar{\xi}_L(l)^3 \quad (3.83)$$

That is, the final average correlation function in the quasilinear regime is the cube of the initial correlation function at  $l$ , where  $l^3 \propto x^3 \bar{\xi}_L^3 \propto x^3 \bar{\xi}_{QL}(x)$ . Note that the result we obtained is independent of the initial power spectrum of the objects that cluster. If the initial power spectrum is a power law, with  $\bar{\xi}_L \propto x^{-(n+3)}$ , we get the scaling:

$$\bar{\xi}_{QL} \propto x^{-3(n+3)/(n+4)} \quad (3.84)$$

For the completely nonlinear regime, we can use the *stable clustering hypothesis*, which roughly says that, although separation between clusters is altered by the expansion of the Universe, their internal structure is constant in time. This would require the correlation function to have the form  $\bar{\xi}_{NL}(a, x) = a^3 F(ax)$ , where  $F$  is some unknown function of its argument. From previous considerations, we know that  $\bar{\xi}_{NL}$  is a function of  $\bar{\xi}_L$ :

$$\bar{\xi}_{NL}(a, x) = a^3 F(ax) = U[\bar{\xi}_L(a, x)] \quad (3.85)$$

We know that the linear correlation function evolves as  $a^2$ , therefore we can write  $\bar{\xi}_L(a, l) \equiv a^2 Q(l^3)$  (the exponent 3 in the argument is there for convenience). In our case this means  $l^3 = x^3 \bar{\xi}_{NL}(a, x) = (ax)^3 F(ax) = r^3 F(r)$ , where  $r = ax$ . With this in mind, (3.85) reads:

$$a^3 F(r) = U[\bar{\xi}_L(a, x)] = U[a^2 Q(l^3)] = U[a^2 Q(r^3 F(r))] \quad (3.86)$$

Now consider this equality at a fixed  $r$ : the unknown function  $U$  needs to satisfy  $U[c_1 a^2] = c_2 a^3$ , thus it must be that:

$$U[z] \propto z^{3/2} \quad (3.87)$$

Finally the mapping from the linear to the nonlinear regimes is:

$$\bar{\xi}_{NL}(x) \propto \bar{\xi}_L^{3/2}(l) \quad (3.88)$$

Again, if the initial power spectrum happens to be a power law, we get:

$$\bar{\xi}_{NL}(x) \propto x^{-\gamma} \quad \gamma = \frac{3(n+3)}{n+5} \quad (3.89)$$

Putting all our results together:

$$\bar{\xi}(x) \propto \begin{cases} \bar{\xi}_L(l) & (\bar{\xi}_L < 1, \bar{\xi} < 1) \\ \bar{\xi}_L^3(l) & (1 < \bar{\xi}_L < 5.85, 1 < \bar{\xi} < 200) \\ \bar{\xi}_L^{3/2}(l) & (\bar{\xi}_L > 5.85, \bar{\xi} > 200) \end{cases} \quad (3.90)$$

where we have assumed that the linear result is valid up to  $\bar{\xi} \approx 1$  and the virialisation to occur at  $\bar{\xi} \approx 200$ , which are the results suggested by the spherical model. Also, we must remember the relationship between the initial scale and the clustering scale (3.80).

\*\*\*

## Chapter references

The main body of the chapter was taken from Mo, Van den Bosch & White [59] and Baumann [56]. A lot of details in the spherical collapse model were taken from Padmanabhan (1999) [66]. For the excursion set theory, the inspiration came from a review by Zentner (2006) [67]. Finally, the last section regarding quasi-linear and non-linear scaling relations is mostly a summary of the corresponding section in Padmanabhan (1999) [66].

## 4 Primordial Black Holes

The history of primordial black holes (PBHs) begins in the sixties when Zeldovich and Novikov pointed at the possibility that, if black holes formed in the early Universe, they could have grown in size by accreting the surrounding radiation. It was only in 1971 that Hawking proposed the scenario [34] in which overdense regions in the early Universe could in principle undergo gravitational collapse and form black holes. This was the seed for the idea of the modern mechanism of PBHs formation.

We know that in astrophysical settings only objects which are heavier than about 3 solar masses can collapse to form black holes (otherwise they form neutron stars), however extremely strong gravitational forces in the early Universe allow formation of not only massive BHs, but also of BHs that could be as small as the Planck mass scale  $\sim 10^{-5} g$ . In fact, after the advent of inflationary cosmology, numerous models have been proposed as to the formation of these primordial black holes that rely mostly on the collapse of initial evolved perturbations, leading to the study of PBH properties such as mass and abundance. On the other hand, observational information about the (non) existence of these objects can provide important constraint on the inflationary models.

Depending on their mass, PBHs leave traces that are found by different observational signals. Firstly, [68] finds that, based on particle emission (Hawking radiation) from black holes, the lifetime of the hole is:

$$t_0 \approx \frac{G^2}{\hbar c^4} \frac{M_0^3}{3\alpha_0} \quad (4.1)$$

where  $\alpha(m)$  is a numerical coefficient that depends on which particle species can be emitted at a significant rate. Since most of the decay time of the hole is spent near its original mass  $M_0$ , the coefficient can be taken to be its value  $\alpha_0$  at that mass (provided that  $\alpha(M)$  does not change rapidly with mass near  $M_0$ ). As a consequence, black holes with masses smaller than the following characteristic scale have already evaporated (by time  $t_0$ ):

$$M_0 \approx \left( \frac{3\hbar c^4 \alpha_0}{G^2} t_0 \right)^{1/3} \approx 10^{15} g \left( \frac{\alpha_0}{4 \times 10^{-4}} \right)^{1/3} \left( \frac{t_0}{13.8 \text{ Gyr}} \right)^{1/3} \quad (4.2)$$

We can see clearly that PBHs with mass smaller than  $\approx 10^{15} g$  cannot exist in the present universe. These objects, though, leave some traces that are useful to determine their abundance in the early Universe, e.g. through changes in light elements abundances produced by Big Bang Nucleosynthesis due to high energy particles emitted by the holes. Observations for this purpose can place constraint on the abundance of these low mass PBHs; on the other hand, primordial black holes with mass greater than  $10^{15} g$  have not evaporated yet. We specify that in this work we focus on non-evaporated PBHs, since they are the most relevant ones.

One interesting question regarding non-evaporating black holes is whether they comprise all of the dark matter content of the Universe. Stringent upper limits on this fact have been placed for a vast PBH mass range (see below). Currently, it appears that PBHs do not explain all the dark matter, but could constitute only a fraction of it.

The LIGO discovery of the merger event (GW150914) of a black hole binary [35] triggered a renewed interest in primordial black holes in the stellar mass range, given the unexpected largeness of the detected ones (around  $30 M_\odot$  each). In particular, after the LIGO announcement, several research groups have independently indicated that the inferred merger rate could be explained by the merger of PBHs [36–38].

Following in the same footsteps, PBHs have been the object of several publications in recent years, one of the main reason being that PBHs could explain the DM found in galactic halos. In addition, as we have said, PBHs could provide at least part of the GWs signal from coalescing binary black holes. Finally, a very interesting possibility is that of PBHs providing the seeds for the supermassive black holes (SMBH) residing in galactic nuclei.

However, since the scenario in which PBHs constitute all of the DM was very quickly excluded, models where PBHs only constitute a fraction of the DM are now studied.

As for the GWs signal, the study of gravitational waves is a potent and useful tool to probe new regions of the PBH parameter space that could not be accessed only by electromagnetic signals, and the next generation advanced gravitational wave interferometers such as the Einstein Telescope [69], Cosmic Explorer [7],

LISA [70] and DECIGO [71] covering different frequency bands will follow. It is very probable that more merger events will be detected and the statistical information of BHs and BH binaries such as mass, spin, eccentricity, redshift etc... will become available. With these data in our hands, we will be able to assess if there is a PBH model that explains it, or if the real scenario is a mixture of multiple possibilities (PBH and astrophysical BHs).

Although these arguments are the ones attracting the most attention, several other cosmological phenomena might be explained by PBHs. For instance, the effect we will consider in this work is the accelerated structure formation via the Poisson or seed effect (see below). A review of other important phenomena in which PBHs play a role are found in [39].

We should note that at the time of writing, there is still no definitive evidence for primordial black holes. However, despite this lack of evidence, PBHs are an interesting concept since they have been invoked to explain numerous cosmological observations, as we have seen, and studying these allows us to place constraints on the fraction of dark matter composed of PBHs and in turn on the cosmological model that generates them. These constraints (non evaporating PBHs) are analyzed in more detail below, and they involve the use of PBHs to explain lensing phenomena, the heating of stars in our Galactic disc, the origin of MACHOs, the generation of large-scale structure, the LIGO observations etc...

## 4.1 Formation of Primordial Black Holes

Several mechanics of PBH formation in the early Universe have been proposed. Examples are the possibility of formation of PBH by domain walls or the via the nucleation of vacuum bubbles which nucleate during the inflation period. However, the most commonly studied scenario is the collapse of superhorizon fluctuations, which we care to give an analysis of in this section.

### 4.1.1 Collapse of Inflationary Perturbations

As pointed out, primordial black holes could have formed in the early Universe, for instance during radiation domination, via the collapse of overdense regions generated by inflation.

This scenario comes with the assumption of an enhancement of the primordial fluctuations at small scales, of which the amplitude is way above the value required to match the CMB observations at larger scales. From the theory of Gaussian random fields [57], those large and rare peaks of the density fluctuations are nearly spherical, thus we can make the approximation that those which gravitationally collapse to black holes are spherical.

After being stretched out of the Hubble horizon at the end of inflation, the fluctuations remain frozen (as we have seen in the first chapter, this means that the gauge-invariant curvature perturbations remain constant) until a later era in which they re-enter the horizon again. If those perturbations exceed a certain threshold, they will start to collapse and form a black hole; otherwise, they will disperse because of pressure gradients which prevent the full collapse.

Mathematically speaking, the early Universe after inflation is well described by the spatially flat FLRW metric:

$$ds^2 = -dt^2 + a(t)^2 \delta_{ij} dx^i dx^j \quad (4.3)$$

We recall that solving the Einstein equations with a energy source given by a perfect fluid yields the background Friedmann equation:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3} \bar{\rho}(t) \quad (4.4)$$

On this background we then consider a locally perturbed region that could eventually collapse to a black hole. Since, as we remarked above, we consider this region to be spherically symmetric and at first much bigger than the Hubble horizon size, we may apply the *separate universe approach*, i.e. we assume a metric of the form:

$$ds^2 = -dt^2 + a(t)^2 e^{2\psi(r)} \delta_{ij} dx^i dx^j \quad (4.5)$$

where  $\psi > 0$  and is assumed to go to zero as  $r$  goes to infinity. Equation (4.5) is equivalent to a FLRW metric of a closed Universe:

$$ds^2 = -dt^2 + a(t)^2 \left( \frac{dR^2}{1-K(R)R^2} + R^2(d\theta^2 + \sin^2 \theta d\phi^2) \right) \quad (4.6)$$

where we made the following coordinate change:

$$\begin{aligned} K &= -\frac{\psi'(r)}{r} \frac{2+r\psi'(r)}{e^{2\psi(r)}} \\ R &= r e^{\psi(r)} \end{aligned} \quad (4.7)$$

We note that the spatial ( $t = 0$  hypersurface) Ricci tensor is given by:

$$R^{(3)} = \frac{K}{a^2} \left( 1 + \frac{d \log K(R)}{3d \log R} \right) \quad (4.8)$$

Ignoring the spatial derivative of  $K$ , the time-time component of the Einstein equations gives:

$$H^2 + \frac{K(r)}{a^2} = \frac{8\pi G}{3}\rho \quad (4.9)$$

which, as was obvious from the form of the metric (4.6), is the Friedmann equation with a small inhomogeneity induced by the  $K(r)$  term.

We use equations (4.4) and (4.9) in order to define the density contrast:

$$\Delta \equiv \frac{\rho - \bar{\rho}}{\bar{\rho}} = \frac{3K}{8\pi G \bar{\rho} a^2} = \frac{K}{H^2 a^2} \quad (4.10)$$

Since  $H^2 \propto a^{-4}$  during radiation domination,  $\Delta$  is vanishingly small initially, but it grows to be of order unity as the Universe evolves. As we know from the study of the Friedmann equations, if we ignore the spatial dependence on  $K$ , the Universe with  $K > 0$  would stop expanding at the point where  $\dot{a} = 0$  and recollapse. This point of inversion is, from (4.9),  $3K/a^2 = 8\pi G\rho$ , which means that the comoving scale of this positively curved region becomes the order of the Hubble size, at which the separate Universe approach fails. We nevertheless assume that (4.9) can still be used as a perturbed metric for black hole formation, which was actually shown to be valid in non-linear numerical studies.

Putting  $3K/a^2 = 8\pi G\rho$  into (4.10) we see that  $\Delta = 1$  is the time when the Universe starts collapsing. We also assume this to be the time of black hole formation  $t = t_c$ . Since a perturbation on scales smaller than the Jeans scale cannot collapse, we set the collapse to happen at  $c_s^2 k^2/a^2 = H^2$  or  $k^2/a^2 = 3H^2$  for  $c_s^2 = 1/3$  in radiation domination. We thus have:

$$1 = \Delta(t_c) = \frac{K}{k^2 H^2 a^2} = \frac{K}{c_s^2 k^2} \quad (4.11)$$

This implies that we should identify  $K$  with  $c_s^2 k^2$ . The condition for black hole formation is then that the density contrast at the time when the scale of interest re-enters the Hubble horizon is greater than  $\Delta_c = c_s^2$ :

$$\Delta(t_k) = \frac{K}{H^2(t_k)a^2(t_k)} = \frac{c_s^2 k^2}{H^2(t_k)a^2(t_k)} \geq c_s^2 = \frac{1}{3} \quad (4.12)$$

where  $t_k$  is the time when the scale is equal to the Hubble horizon  $k/a = H$ . Roughly speaking the mass of the black hole that is formed in the radiation dominated era is equal to the horizon mass  $M_H = (4\pi/3)\rho H^{-3}$  at the time of formation:

$$M_{\text{PBH}} = \gamma M_H \Big|_{\text{formation}} = \gamma \frac{4\pi}{3} \rho_{\text{form}} H_{\text{form}}^{-3} = \gamma \frac{4\pi}{3} \frac{3H_{\text{form}}^2}{8\pi G} H_{\text{form}}^{-3} = \gamma \frac{1}{2G} H_{\text{form}}^{-1} \quad (4.13)$$

where the correction factor  $\gamma$  can be evaluated to be around  $\gamma \approx 0.2$  by a simple analytic calculation.

### 4.1.2 Threshold Values

In equation (4.12) we used the threshold condition that the density contrast become comparable to 1/3. A thorough investigation for this type of PBH formation mechanism has been studied in the literature (for a review of thresholds see [72]), and in general numerical simulations demonstrate that there is no unique value for the threshold, but different density/curvature perturbation profiles collapse above different thresholds, which is only a natural conclusion from a physical point of view. In terms of the comoving density contrast, the spread of the threshold was found to be 0.3 to 0.66, a range where our value of 1/3 lies, despite being obtained from a rather crude reasoning.

### 4.1.3 Near Critical Collapse

We have said that the PBH mass is roughly of the order of the Hubble horizon at the time of formation (when the scale enters the horizon). In reality, the final PBH mass, after the completion of accretion processes from the FLRW background, depends of the specific profile of the fluctuation, its initial overdensity amplitude and the equation of state parameter  $w$  of the background fluid. Numerical simulations have although shown that the black hole mass follows a scaling law [73] when the perturbation  $\delta$  is close to the critical value  $\delta - \delta_c \lesssim 10^{-2}$ :

$$M_{\text{PBH}} = KM_H(t_H)(\delta - \delta_c)^\gamma \quad (4.14)$$

where  $\gamma$  is a universal exponent which only depends on the background fluid's equation of state  $\gamma = \gamma(\omega)$ , whereas the constant  $K$  has a numerical value of  $\mathcal{O}(1)$  and depends both on the shape of the curvature fluctuation and  $\omega$ .

### 4.1.4 PBH Formation During Matter Domination

Until now we have considered the collapse of overdensities during the radiation domination era. Of course, PBHs may also form in the matter dominated era, and in the literature it was emphasized that taking into account deviations from our spherical collapse configuration is essential. Another effect that may turn out to be important is the fact that matter cannot be exactly a non relativistic fluid with  $P = 0$ , but may have a small pressure term. Interestingly, a massive scalar field is known to behave like non relativistic matter when it oscillates around the potential's minimum. Therefore PBH formation in the matter era may become important for models where the scalar field oscillates after inflation and dominates the universe.

## 4.2 Clustering of PBHs

As we have said above, the LIGO/Virgo collaboration's detection of gravitational waves signals coming from a coalescence of massive black holes has renewed interest in the possibility that these might be of primordial origin and formed in the early Universe. In particular, the research is focused on the likelihood that these PBHs comprise a fraction  $f_{\text{PBH}}$  of the total dark matter content of the Universe.

Since the evolution of these black holes through cosmic history until their merger is of fundamental importance to making predictions for these types of GWs signals, it could be of use to consider the clustering aspects of the problem since, for instance, PBH binaries could be disrupted by encounters with a third PBH if they live in clusters, and therefore the merger rate would be modified.

Interestingly, the clustering of primordial black holes could help relax some stringent bounds on the mass distributions of PBHs arising from microlensing experiments and from the CMB (see below).

From these examples it is clear that the clustering of these primordial black holes presents an important and interesting question.

To help describe mathematically the spatial distribution of the PBHs, we need to characterize the two point function as a function of separation  $x \equiv |\vec{x}|$ . We can use the peak approach to large scale structure [74] to write the overdensity of discrete tracer objects (PBHs in our case) at positions  $\vec{x}_i = \vec{x}_i(z)$ :

$$\delta_{\text{PBH}} = \frac{n(\vec{x})}{\bar{n}_{\text{PBH}}} = \frac{1}{\bar{n}_{\text{PBH}}} \sum_i \delta_D(\vec{x} - \vec{x}_i) - 1 \quad (4.15)$$

where  $\delta_D(\vec{x})$  is the Dirac delta function, and we introduced the average number density of PBH per comoving volume:

$$\bar{n}_{\text{PBH}} = \frac{\bar{\rho}_{\text{PBH}}}{m_{\text{PBH}}} = \frac{\Omega_{\text{PBH}} \rho_{\text{crit}}}{m_{\text{PBH}}} = \frac{f_{\text{PBH}} \Omega_c \rho_{\text{crit}}}{m_{\text{PBH}}} \quad (4.16)$$

where we assumed a monochromatic mass function for PBHs, such that a general PBH model is uniquely defined by the BH mass  $m_{\text{PBH}}$  and the fraction  $f_{\text{PBH}}$  of dark matter in the form of PBHs.

The two point correlation function of this fluctuation field is, as we defined it in (2.159):

$$\begin{aligned} \langle \delta_{\text{PBH}}(\vec{x}) \delta_{\text{PBH}}(\vec{0}) \rangle &= \frac{1}{\bar{n}_{\text{PBH}}^2} \left\langle \sum_{i,j} \delta_D(\vec{x} - \vec{x}_i) \delta_D(\vec{x}_j) \right\rangle - \frac{1}{\bar{n}_{\text{PBH}}} \left\langle \sum_i \delta_D(\vec{x} - \vec{x}_i) \right\rangle - \frac{1}{\bar{n}_{\text{PBH}}} \left\langle \sum_j \delta_D(\vec{x}_j) \right\rangle + 1 \\ &= \frac{1}{\bar{n}_{\text{PBH}}^2} \delta_D(\vec{x}) \left\langle \sum_i \delta_D(\vec{x} - \vec{x}_i) \right\rangle + \frac{1}{\bar{n}_{\text{PBH}}^2} \left\langle \sum_{i \neq j} \delta_D(\vec{x} - \vec{x}_i) \delta_D(\vec{x}_j) \right\rangle - 1 \\ &= \frac{1}{\bar{n}_{\text{PBH}}} \delta_D(\vec{x}) + \frac{1}{\bar{n}_{\text{PBH}}^2} + \left\langle \sum_{i \neq j} \delta_D(\vec{x} - \vec{x}_i) \delta_D(\vec{x}_j) \right\rangle - 1 \\ &\equiv \frac{1}{\bar{n}_{\text{PBH}}} \delta_D(\vec{x}) + \xi_{\text{PBH}}(x) \end{aligned} \quad (4.17)$$

where we split the sum into an  $i = j$  and an  $i \neq j$  part, which correspond respectively to the correlation of the PBHs with themselves and with different PBHs, and we introduced the *reduced correlation function* for PBHs  $\xi_{\text{PBH}}(x)$ . The first term appears due to "self-pairs", which are usually ignored in the calculation of real space correlations. Taking the Fourier transform we get the power spectrum:

$$\Delta^2(k, z) = \frac{k^3}{2\pi^2} \int d^3x e^{i\vec{k}\cdot\vec{x}} \langle \delta_{\text{PBH}}(\vec{x}) \delta_{\text{PBH}}(\vec{0}) \rangle = \frac{k^3}{2\pi^2} \left( \frac{1}{\bar{n}_{\text{PBH}}} + \int d^3x e^{i\vec{k}\cdot\vec{x}} \xi_{\text{PBH}}(x, z) \right) \quad (4.18)$$

We see that the self-pair term contributes to the Poisson noise  $\frac{1}{\bar{n}_{\text{PBH}}}$ , which is related to the discrete nature of the PBHs and is thus present in any distribution of discrete objects, regardless of their clustering (which is controlled by the reduced correlation function).

The only requirement is that the power spectrum be positive definite, which implies that the Fourier transform of the reduced two point function is greater or equal to  $-\frac{1}{\bar{n}_{\text{PBH}}}$ .

Note also that another constraint arises from the intrinsic exclusion volume associated to black holes, which cannot form at an arbitrary close distance. As a result, the conditional probability to find a PBH at a comoving distance  $x$  from another, which is proportional to  $1 + \xi_{\text{PBH}}(x)$ , must vanish at a certain small scale  $x \leq x_{\text{exc}}$ , which for most scenarios is approximately the Hubble radius  $x_H$  at formation time:

$$\xi_{\text{PBH}}(x) \approx -1 \quad \text{for } x \leq x_{\text{exc}} \quad (4.19)$$

so that PBHs are anti-correlated at short distances, which is expected since there can be at most one PBH per horizon volume.

#### 4.2.1 The Evolution of Clustering

Hierarchical clustering implies that, below a certain characteristic PBH *clustering length*, the fluctuations in the PBH number counts are dominated by the correlation function term in (4.18), while the Poisson shot noise dominates on larger scales.

One interesting conclusion found in [75] is that, in the PBH formation mechanism of collapsing density perturbations, while the value of the initial clustering length is sensitive to the shape of the primordial curvature power spectrum, for a narrow spectral feature it is in fact significantly smaller than the mean comoving PBH separation, rendering initial clustering not relevant. In other words, the PBHs are Poisson

distributed at the time of formation, rather irrespectively of the details of the power spectrum. In light of this, the initial PBH power spectrum can be approximated letting  $\xi(x, 0) \approx 0$  for all scales in (4.18):

$$\Delta_i^2(k, 0) \equiv \Delta_i^2(k) = \frac{k^3}{2\pi^2} \int d^3x e^{i\vec{k}\cdot\vec{x}} \langle \delta_{\text{PBH}}(\vec{x}) \delta_{\text{PBH}}(\vec{0}) \rangle \approx \left( \frac{k}{k_*} \right)^3 \quad (4.20)$$

where we introduced a characteristic wavenumber:

$$k_* \equiv (2\pi^2 \bar{n}_{\text{PBH}})^{1/3} \quad (4.21)$$

inversely proportional to the mean separation between PBHs.

In the linear regime, the density contrast is frozen until matter radiation equality, and subsequently grows as [76]:

$$\Delta_{\text{PBH},L}^2(k, z) = (1 + f_{\text{PBH}}(D_+(z) - 1))^2 \Delta_i^2(k) \quad (4.22)$$

where  $D_+(a)$  is the linear growth function for the isocurvature perturbations induced by the black holes, and can be approximated by:

$$D_+(z) = \left( 1 + \frac{3\gamma}{2\alpha_-} \frac{1+z_{\text{eq}}}{1+z} \right)^{\alpha_-} \quad (4.23)$$

where:

$$\gamma = \frac{\Omega_c}{\Omega_m} \quad \alpha_- = \frac{1}{4} (\sqrt{1+24\gamma} - 1) \quad (4.24)$$

As was argued above (3.90), the linear perturbations in the PBH number density enter the quasi-linear stage when the power spectrum (or the correlation function) is  $\sim 1$ . This transition happens at a different redshift for every scale, which can be found through:

$$\Delta_{\text{PBH},L}^2(k_{\text{L-QL}}(z), z) \approx 1 \quad (4.25)$$

This means:

$$k_{\text{L-QL}}(z) \approx k_* (1 + f_{\text{PBH}}(D_+(z) - 1))^{-3/2} \quad (4.26)$$

We can write the power spectrum in the quasi-linear regime using (3.84) and (2.166) with a scale invariant initial power spectrum  $n \approx 1$ :

$$\Delta_{\text{PBH},\text{QL}}^2(k, z) \approx \left( \frac{k}{k_{\text{L-QL}}(z)} \right)^{9/4} \quad (4.27)$$

For the fully nonlinear regime, again we can use (3.90), so that the transition happens when  $\Delta_{\text{PBH},\text{QL}}^2 \sim 200$ . We define, as before, the scale at which the transition happens:

$$\Delta_{\text{PBH},\text{QL}}^2(k_{\text{QL-NL}}(z), z) \approx 200 \quad (4.28)$$

which means:

$$k_{\text{QL-NL}}(z) \approx 10.5 k_{\text{L-QL}}(z) \quad (4.29)$$

For the nonlinear power spectrum we can use again (3.89) with (2.166):

$$\Delta_{\text{PBH},\text{NL}}^2 \approx 200 \left( \frac{k}{k_{\text{QL-NL}}(z)} \right)^{9/5} \quad (4.30)$$

### 4.3 PBHs in structure formation

Assuming a monochromatic mass function, where the PBH population is characterized by  $m_{\text{PBH}}$  and  $f_{\text{PBH}}$ , at scales where the discrete nature of these objects becomes relevant, there are two different effects that can influence structure formation, and they will dominate in different parts of the parameter space.

In general, PBHs of mass  $m_{\text{PBH}}$  provide a source of fluctuations for objects of mass  $M$  either via the *seed effect*, where the effect of a single black hole generates an initial fluctuation  $m_{\text{PBH}}/M$ , or via the *Poisson*

*effect*, in which the  $\sqrt{N}$  fluctuation in the number density of PBHs will generate a perturbation of order  $(f_{\text{PBH}} m_{\text{PBH}} / M)^{1/2}$ . Both types of perturbations will grow to bind objects of mass  $M$ . The relationship between these effects is nontrivial, and one needs to determine the dominant one at different scales. We will investigate the case of a monochromatic mass function, which is relevant for the present work.

We have just said that for PBHs of mass  $m_{\text{PBH}}$ , the initial fluctuations on a mass scale  $M$  are:

$$\delta_i \simeq \begin{cases} m_{\text{PBH}}/M & \text{seed} \\ (f_{\text{PBH}} m_{\text{PBH}}/M)^{1/2} & \text{Poisson} \end{cases} \quad (4.31)$$

Clearly, if PBHs provide all the dark matter  $f_{\text{PBH}} \sim 1$ , then the Poisson effect dominates on all scales  $M$ , whereas in scenarios where  $f_{\text{PBH}} \ll 1$ , the seed effect starts to dominate at a scale  $M < m_{\text{PBH}}/f_{\text{PBH}}$ , and the Poisson effect at larger scales. In fact, the subdivision (4.31) only holds if the seed effect is considered for a dark matter fraction  $f_{\text{PBH}} < m_{\text{PBH}}/M$ , since otherwise a region of mass  $M$  would contain more than one black hole, causing conflict between different seeds. In other words, a population of PBHs will generate a fluctuation in the density field that, at large scales, is a Poissonian noise; whenever we look at relatively small scales (provided that  $f_{\text{PBH}}$  is not 1 or close), i.e. near a single black hole, we start to notice a fluctuation in the density field that's different in nature from Poissonian white noise, and in general can only be studied via numerical simulations.

The effect of these fluctuations in time can be evaluated as follows. The  $\sqrt{N}$  fluctuation is frozen when radiation dominates the energy density, and starts to grow from matter-radiation equality as  $\sim (1+z)^{-1}$ . On the other hand, since  $z_{\text{eq}} \sim 4000$  and an overdense region binds when  $\delta \sim 1$ , the mass binding at redshift  $z$  is:

$$M \simeq \begin{cases} 4000 m_{\text{PBH}} z^{-1} & \text{seed} \\ 10^7 f_{\text{PBH}} m_{\text{PBH}} z^{-2} & \text{Poisson} \end{cases} \quad (4.32)$$

For a given value of  $f_{\text{PBH}}$  then, (4.32) and the condition  $M < m_{\text{PBH}}/f_{\text{PBH}}$  imply that the seed effects dominates at redshifts  $z > z_{\text{eq}} f_{\text{PBH}} \sim 4000 f_{\text{PBH}}$ . This condition has a simple interpretation. Since the fluctuations grow as  $(1+z)^{-1}$  after equality, the fraction of bound regions in the Universe at a given redshift is  $f_{\text{PBH}} z_{\text{eq}}/z$ , which exceeds one for  $f_{\text{PBH}} > z/z_{\text{eq}}$ . In this case the competition between different seeds (black holes) will reduce the mass of the bound region to at most  $M \sim m_{\text{PBH}}/f_{\text{PBH}}$ , precisely the value of  $M$  above which the Poisson noise dominates. In particular, [76] finds numerically that, given a redshift  $z$ , the cross-over value for  $f_{\text{PBH}}$  is:

$$f_{\text{PBH}} \simeq \frac{0.02(1+z)}{100} \quad (4.33)$$

Therefore, despite there being no apparent dependence on  $f_{\text{PBH}}$  in the seed effect (4.32), there is an upper limit that bounds the parameter region in which the result is valid.

The considerations were carried out for a monochromatic PBH mass function. In general, if PBHs have an extended mass function, both the Poisson and seed effect could operate at the same time on different scales. For instance, if one imagines a PBHs population with two different masses (one big and one small), one could provide the dark matter and generate the Poisson effect, while the other may provide a low density of supermassive black holes which act as seeds.

The case of the extended mass function is more subtle and difficult to follow analytically, but in general it is expected to be more physically relevant than the monochromatic counterpart.

#### 4.4 Abundance of PBHs

In order to investigate the abundance of PBHs, we introduce a parameter that represents the mass fraction of PBHs in the Universe at formation:

$$\beta \equiv \left. \frac{\rho_{\text{PBH}}}{\rho_{\text{tot}}} \right|_{\text{at formation}} = \left( \frac{H_0}{H_{\text{form}}} \right)^2 \left( \frac{a_{\text{form}}}{a_0} \right)^{-3} \Omega_c f_{\text{PBH}} \quad (4.34)$$

By using the relation between the mass of the PBH and the horizon at formation (4.13), the mass fraction can be written as:

$$\beta \simeq 3.7 \times 10^{-9} \left( \frac{\gamma}{0.2} \right)^{-1/2} \left( \frac{g_{*,\text{form}}}{10.75} \right)^{1/4} \left( \frac{M_{\text{PBH}}}{M_{\odot}} \right)^{1/2} f_{\text{PBH}} \quad (4.35)$$

where  $g_{*,\text{form}}$  is the number of relativistic degrees of freedom. It is clear that an observational constraint on  $f_{\text{PBH}}$  can be reinterpreted as one on  $\beta$ .

Technically, the abundance of primordial black holes is usually estimated via a statistical scheme. Since the collapse of sufficiently overdense regions at horizon re-entry is the formation mechanism of PBH we are considering, once the probability distribution for these density fluctuations is given,  $\beta$  can be regarded as the probability that the density contrast is larger than threshold for PBH formation. We can evaluate  $\beta$  therefore as:

$$\beta = \gamma \int_{\delta_{\text{th}}}^1 P(\delta) d\delta \quad (4.36)$$

where  $\delta_{\text{th}}$  is the threshold for PBH formation. For a Gaussian distribution we find:

$$\beta \simeq \frac{\gamma}{\sqrt{2\pi}\nu_{\text{th}}} \exp\left(-\frac{\nu_{\text{th}}^2}{2}\right) \quad (4.37)$$

where  $\nu_{\text{th}} = \delta_{\text{th}}/\sigma_{\text{PBH}}$  and the variance is:

$$\sigma_{\text{PBH}}^2 = \int d \log k k P_{\delta}(k) W^2(k, R) = \int d \log k W^2(k, R) \frac{16}{81} (kR)^4 P_{\mathcal{R}}(k) \quad (4.38)$$

where  $P_{\delta}$  and  $P_{\mathcal{R}}$  are the primordial density fluctuations and curvature perturbations power spectra respectively, and  $W(k, R)$  is a smoothing function at a comoving scale  $R \approx 1/(a_{\text{form}} H_{\text{form}})$ .

The equation we just found is very important, since we can interpret a constraint on  $\beta$  as one on the peak height  $\nu$ , which would give us a hint for constructing physically relevant inflationary models where PBHs indeed form. For instance, if the constraint for  $M_{\text{PBH}}$  is  $f_{\text{PBH}} < 10^{-3}$ , it would mean  $\nu_{\text{th}} > 6.27$ , that is  $\sigma_{\text{PBH}} < 0.08$  with  $\delta_{\text{th}} = 0.5$ .

Finally, note that in these calculations we have used a Gaussian distribution function for the primordial density fluctuations  $P(\delta)$ . However, as we will see below, in order to have efficient PBH formation we need to consider non-standard inflationary paradigms which violate the slow roll condition. In this case, therefore, the effect of eventual non-Gaussianity can be relevant for some inflationary models.

## 4.5 Generating Primordial Perturbations for PBHs in Inflation

### 4.5.1 Inflation

As we have seen briefly in the first section, it is widely accepted that before the Hot Big Bang the Universe underwent a period of rapid expansion called inflation where  $\ddot{a} > 0$ . The remarkable feature of inflation is that it provides a natural mechanism with which to create primordial density fluctuations that in the end grow into structures in the late Universe by gravitational instability (e.g. via spherical collapse as we've explained above), and therefore provide observability. In fact, inflation can make the Universe homogeneous and isotropic, while generating these local primordial density fluctuations from quantum vacuum fluctuations. Below we briefly review this formalism in order to give a general idea of the conditions needed to seed the formation of primordial black holes via the collapse of overdense regions during radiation domination.

Inflation can only be realised only if the Universe is initially filled with some sort of "dark energy" component. Furthermore, this should be dynamic such that inflation ends successfully and continuously connects to the Hot Big Bang. The most simple and famous realisation of such a component is accomplished through a scalar field  $\phi$  called *inflaton*. To start with, we decompose the field into an homogeneous mode and its fluctuations:

$$\phi(\vec{x}, t) = \phi_0(t) + \delta\phi(\vec{x}, t) \quad (4.39)$$

The energy density and pressure of the homogeneous mode are given by:

$$\begin{aligned}\rho &= \frac{1}{2}\dot{\phi}_0^2 + V(\phi_0) \\ P &= \frac{1}{2}\dot{\phi}_0^2 - V(\phi_0)\end{aligned}\tag{4.40}$$

where  $V(\phi)$  is the potential of the field. Now we take into account the second Friedmann equation (2.33):

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\rho(1+3\omega)\tag{4.41}$$

We see that  $\ddot{a} > 0$  would require  $\omega < 1/3$ . We see that such a constraint is realised by the inflaton if in (4.40) we consider the velocity to be sufficiently small  $\dot{\phi}_0^2 \ll V(\phi_0)$ . Such inflation is naturally known as slow roll inflation, the success of which is determined by the flatness of the potential  $V(\phi_0)$  as we show below. The dynamics during inflation are determined by a combination of the inflaton potential and the Klein Gordon equation in curved spacetime for the homogeneous mode:

$$\begin{aligned}H^2 &= \frac{1}{3M_{Pl}^2} \left( \frac{1}{2}\dot{\phi}_0^2 + V(\phi_0) \right) \\ \ddot{\phi}_0 + 3H\dot{\phi}_0 + V'(\phi_0) &= 0\end{aligned}\tag{4.42}$$

where  $V' = \partial_\phi V$ . As we said, the characteristic of inflation is that all quantities are slowly varying, and thus we require that both the velocity  $\dot{\phi}_0$  and the acceleration  $\ddot{\phi}_0$  (we want the slow roll behavior to persist) of the field are small. To this end we introduce two *slow roll parameters*, the first of which is:

$$\epsilon \equiv -\frac{\dot{H}}{H^2} = \frac{3/2\dot{\phi}_0^2}{1/2\dot{\phi}_0^2 + V(\phi_0)} \approx \frac{M_{Pl}^2}{2} \left( \frac{V'}{V} \right)^2\tag{4.43}$$

where the first equality follows from combining (4.42) and the second one from using the limits  $\dot{\phi}_0 \ll 1$  and  $\ddot{\phi}_0 \ll 1$  in both equations (4.42). The second slow roll parameter is:

$$\delta \equiv -\frac{\ddot{\phi}_0}{H\dot{\phi}_0}\tag{4.44}$$

If both the speed and the acceleration of the inflaton are small  $\epsilon, \delta \ll 1$ , then the inflationary expansion will last for a long time.

We can see that when  $\delta$  is small the inflaton's field is determined by the slope of the potential, from which we get:

$$\eta \equiv \delta + \epsilon \approx M_{Pl}^2 \frac{V''}{V}\tag{4.45}$$

Therefore  $\epsilon, \delta \ll 1$  implies  $\epsilon, \eta \ll 1$ , so that a convenient way to judge whether a potential  $V(\phi)$  can lead to slow roll inflation is to ask if  $\epsilon, \eta \ll 1$ , which are related to the slope of the potential.

#### 4.5.2 Inflation Models for PBH production

In order to understand what type of models are needed to form PBHs from the collapse of overdensities, we introduce at first the primordial curvature power spectrum induced by the primordial fluctuations in the inflaton field:

$$P_{\mathcal{R}}(k) = \frac{1}{2\epsilon M_{Pl}^2} \left( \frac{H}{2\pi} \right)^2 \Big|_{aH=k}\tag{4.46}$$

During slow roll inflation both  $H$  and  $\epsilon$  vary only slowly and the power spectrum is therefore close to scale invariant. In fact we can extract the  $k$  dependence of  $P_{\mathcal{R}}$ , the *spectral index*  $n_s$ , in the following way:

$$n_s \equiv \frac{d \log P_{\mathcal{R}}}{d \log k} + 1 = -6\epsilon + 2\eta + 1\tag{4.47}$$

Current CMB observations have indicated that  $P_{\mathcal{R}} \sim 10^{-9}$  (often known as COBE normalization) over CMB observables scales, whereas the PBH formation is effective enough to be observationally interesting for  $P_{\mathcal{R}} \sim 10^{-2} - 10^{-1}$ . As a consequence, in order to obtain large enough perturbation on small scales (that are consistent with observations), we would need to relax some of the assumptions we made in the inflationary paradigm, since this implies  $\epsilon, \eta \ll 1$ .

For instance, to generate a  $\sim 30M_{\odot}$  PBH, i.e. at a scale  $k_{30M_{\odot} \text{ PBH}} \sim 2.9 \times 10^5 \text{ Mpc}^{-1}$  (interesting for LIGO observations), we would need:

$$n_s - 1 = \frac{\log(P_{\mathcal{R}}(k_{30M_{\odot} \text{ PBH}})/P_{\mathcal{R}}(k_{\text{CMB}}))}{\log(k_{30M_{\odot} \text{ PBH}}/k_{\text{CMB}})} \approx 0.85 \quad (4.48)$$

where we use  $k_{\text{CMB}}$  as a typical observable scale by CMB observations. Clearly this is in contrast with the results [5] which indicate  $n_s \sim 0.965$ .

In other words, for PBH formation, we would need an amplified blue-tilted power spectrum. We now review some models which accomplish this.

- *Inflection Inflation Model*: [77] proposed a model which can produce a peak in the primordial power spectrum by considering a potential with an inflection point where the inflaton temporarily slows down during inflation. The reason this model can produce a peak in the power spectrum is the following. If we look at the power spectrum in (4.46), one can see that when the inflaton temporarily slows down the parameter  $\epsilon$  becomes suppressed and thus a peak is formed at the scales which exit the Horizon during this plateau phase.

This model in reality needs a near-inflection point in order to avoid the eternal inflation that would be produced by a full inflection point. It was found later that the standard slow roll conditions might be violated during the plateau phase and in fact the amplification of the fluctuations can be realized by this model, but  $P_{\mathcal{R}} \approx 10^{-4}$  at most.

- *Running Mass Inflation Model*: In this model the time dependence of the parameter  $\eta$  during inflation could be large, so that the scale-dependence of  $n_s$  could become relevant [78]. Therefore  $\eta$  could be small so as to be consistent with the Planck results on CMB scales, but takes a larger value on smaller scales that are relevant for PBH formation. For these kinds of models, the primordial power spectrum is not a simple power law anymore but includes in a perturbative fashion the scale dependence of  $n_s$ :

$$P_{\mathcal{R}}(k) = A_s \left( \frac{k}{k_*} \right)^{n(k)} \quad (4.49)$$

where:

$$n(k) = n_s - 1 + \frac{1}{2!} \alpha_s \log(k/k_*) + \frac{1}{3!} \beta_s \log^2(k/k_*) + \dots \quad (4.50)$$

where  $\alpha_s$  and  $\beta_s$  are called, respectively, the "running" and "running of running" of the spectral index. Tweaking these parameters would give us the results required to explain CMB observations and produce PBHs. However the Planck results [5] find (at the pivot scale  $k_* = 0.05 \text{ Mpc}^{-1}$ ):

$$\begin{aligned} \beta_s &= 0.009 \pm 0.012 \\ \alpha_s &= 0.0011 \pm 0.0099 \end{aligned} \quad (4.51)$$

so that there seems to be no evidence for any significant deviation from a simple power law on all scales.

- *Chaotic New Inflation Model*: In this scenario [79] the inflaton's potential is a double well. The inflaton would start its slow roll in the vicinity of the origin (an unstable maximum) but its initial amplitude would largely exceed the global minimum, which gives rise to chaotic inflation. By tuning the parameters of the model, two inflationary phases could be realized. After a first phase of standard chaotic inflation (with a large field value), the inflaton oscillates around the origin; however, in this model the potential has a feature which can realize a new second phase of inflation around the

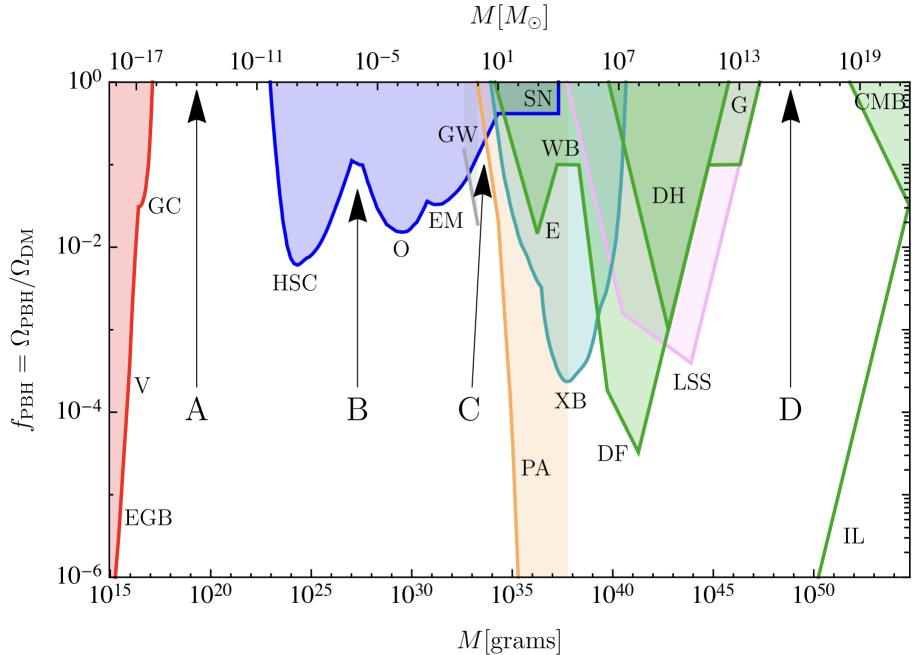


Figure 10: Constraint on the PBH dark matter fraction  $f_{\text{PBH}}$  for a monochromatic mass function. The individual bounds are from evaporation (red), lensing (blue), gravitational waves (GW) (grey), dynamical effects (green), accretion (light blue), CMB distortions (orange) and large-scale structure (purple). The evaporation limits come from the extragalactic  $\gamma$ -ray background (EGB), the Voyager positron flux (V) and annihilation-line radiation from the Galactic centre (GC). The lensing constraints derive from microlensing of supernovae (SN) and of stars in M31 by Subaru (HSC), the Magellanic Clouds by the Experience pour la Recherche d'Objets Sombres (EROS) and Massive Compact Halo Object (MACHO) collaborations (EM), and the Galactic bulge by the Optical Gravitational Lensing Experiment (OGLE) (O). The dynamical bounds are from wide binaries (WB), star clusters in Eridanus II (E), halo dynamical friction (DF), galaxy tidal distortions (G), heating of stars in the Galactic disk (DH) and the cosmic microwave background dipole (CMB). The large-scale structure (LSS) limits are due to the requirement that various cosmological structures do not form earlier than observed. The accretion constraints derive from X-ray binaries (XB) and Planck measurements of cosmic microwave background distortions (PA). The four mass windows (A, B, C, D) indicate regions in which PBHs could have an appreciable density, assuming the validity of the mentioned constraints. Figure adapted from [72]

origin.

In this scenario, modes which exit the Hubble horizon at the transition era from the first to the second inflationary phases (where the slow roll conditions are temporarily violated) would be highly amplified in the power spectrum.

We note that in this section we have only worked with and reviewed the case of single field inflation, however in the literature some interesting cases of multi field inflation have been considered, such as the curvaton model and double inflation model.

#### 4.6 Observational Constraints on Non-Evaporated PBHs

As we discussed above PBHs are used to explain multiple cosmological phenomena, which in turn help us to place constraints on  $f_{\text{PBH}}$  through their observation. In this section we will be reviewing these constraints, which are summarized in Figure 10. We can immediately see that there exist only a few mass windows where PBHs could provide all the dark matter ( $f_{\text{PBH}}=1$ ), and this was clear already more than a decade ago [80]. We note that in this section we assume the PBHs to have a monochromatic mass function.

The main constraints on  $f_{\text{PBH}}$  are "direct", in the sense that they are derived by investigating the effects that PBHs directly trigger by their gravitational potential and are thus independent of the mechanism involved in their formation. These effects are classified roughly in the following manner: evaporation, gravitational lensing, numerous dynamical effects (among which the growth of large-scale structure), accretion and gravitational waves.

Before giving a brief review of each one be, it must be stressed that most constraints have a certain (varying) degree of uncertainty and some come with a caveat. For instance, for some the observations are well understood (like CMB and lensing data) but there are uncertainties in the physics of the black holes, whereas for others, it is the observations themselves that are not understood (or depend on some assumptions). Furthermore, many constraints depend on astrophysical parameters that are not indicated explicitly. In other words, some of these limits might be significantly relaxed in the near future, or even disappear entirely. Finally, some constraints have been omitted for brevity, and a more detailed review can be found in [81], which is the most comprehensive one on this topic to date.

#### 4.6.1 Evaporation Constraints

One of the strongest constraints is on primordial black holes with masses  $M \geq 10^{15} M_{\odot}$  using the non-observation of  $\gamma$ -rays which originate from their evaporation [82]; in other words PBHs emit Hawking radiation that should not exceed the observed background. Further constraints from evaporation use positron data from Voyager 1 [83], measurements of the 511 keV annihilation-line radiation from the Galactic center [84, 85] and  $\gamma$ -ray and radio observations of the Galactic center [86]. Having summarized the most important contraints from evaporation, we note that there exist other ones that are not mentioned in this work, see [81] for a thorough review. Some of the constraints discussed are plotted in red on the left of Figure 10. We note here that the mass region  $10^{-15} \leq M_{\text{PBH}}/M_{\odot} \leq 10^{-11}$  is free of observational bounds, since the constraints coming from the possible capture of PBHs from neutron stars or white dwarfs were recently discredited.

#### 4.6.2 Gravitational Lensing Constraints

Constraints on the PBH abundance via gravitational lensing are retrieved using observations of Andromeda with the *Subaru Hyper Suprime-Camera* (HSC) [87], observations of stars in the Large and Small Magellanic clouds which probe the fraction of the Galactic halo in PBHs via EROS [88], the MACHO project [89] and observation of the *Optical Gravitational Lensing Experiment* (OGLE) in the Galactic bulge [90]. Lensing constraints are shown in blue in Figure 10.

#### 4.6.3 Dynamical Constraints

To a certain degree, PBHs affect any astrophysical system via gravitational interaction. As a consequence, by evaluating the impact of this encounter and later comparing with observation, it is possible to place new constraints on PBH abundance for large masses.

This line of reasoning was applied to: wide binaries in the Milky Way since they are particularly vulnerable to disruption by PBHs [91], globular clusters [92] and a star cluster near the center of the dwarf galaxy Eridanus 2 that has not been disrupted by halo objects [93].

Furthermore, PBHs moving randomly in the Galactic halo have a chance to pass through the galactic disk and the stars, being pulled by gravitational attraction, acquire velocity. Since the direction of the velocity gain of a star is random, the time evolution of the disk stars is described by a random walk. Therefore, the variance of the stars' velocities increases over time, so that the disk becomes hotter over time. By requiring that the velocity increased by the PBHs does not exceed the observed velocity, it is possible to put bounds on the PBH abundance. This was shown in [94].

A further limit in this mass range arises because halo objects will be dragged into the nucleus of the Milky Way by the dynamical friction of the spheroid stars and halo objects themselves [92].

The last dynamical effect we mention is acquired by requiring that various types of structures do not form too early through their seed or Poisson effect [95]. Analogous constraints can be derived for dwarf galaxies

and galaxy clusters. We also note that the Lyman- $\alpha$  forest is influenced by the Poisson effect [96]. The dynamical constraints are shown in green in Figure 10 and the large scale structures effect is in pink.

#### 4.6.4 Accretion Constraints

Numerous works in the 1980s and 1990s pointed out that the accretion of background gas by PBHs could have a large luminosity, which consequently imposes limits on the PBH abundance. It is possible to place a constraint coming from PBH interactions with the interstellar medium which would signify a non-neglectable X-ray flux, thus contributing to the number density of compact X-ray objects in galaxies [97]. Planck measurements can also be used as a probe of PBH abundance since the process of accretion is accompanied by the emission of radiation [98]. The results from accretion constraints are shown in light blue in Figure 10.

#### 4.6.5 Gravitational Wave Constraints

Like stellar black holes, PBHs can undergo mergers and as a consequence emit gravitational waves, in which case they would generate a stochastic background if constituting a high enough fraction of the dark matter content. Figure 10 contains the constraints in [99] (in grey) where strong limits on the PBH abundance are derived by considering the first LIGO/Virgo observational run to the observable merger rate of PBHs. Note that this constraint does not extend up to  $f_{\text{PBH}} = 1$  since for sufficiently high PBH densities, tidal disruption decreases the number of completed merger processes.

An interesting figure is the following, which compares the redshifts at which the various discussed observations happen:

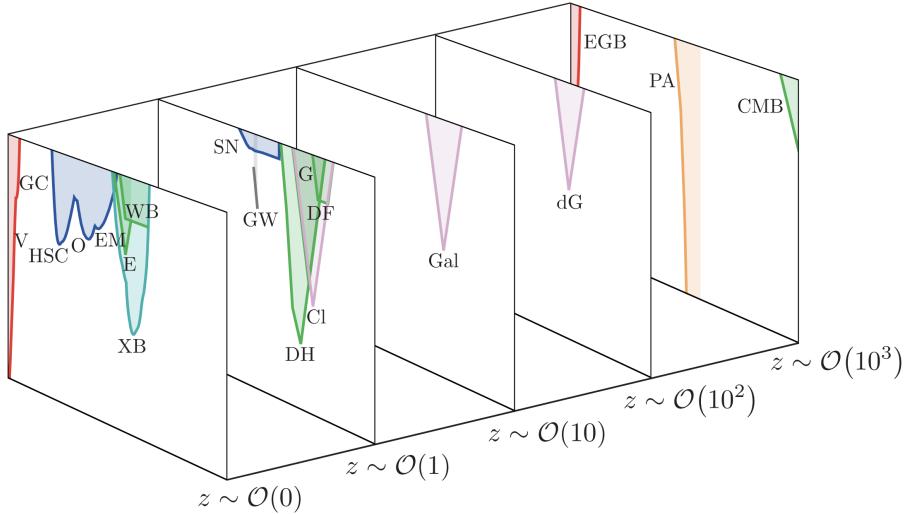


Figure 11: Constraints on the  $f_{\text{PBH}}$  parameter grouped in various redshift bins. Figure adapted from [72]

As was remarked above, however, these PBH constraints hold only when supported by initial assumptions. In fact, as [100] points out, a monochromatic PBH mass function is not very realistic, since gas accretion and merger histories will both tend to broaden and shift the mass function over cosmic time, let alone the fact that PBHs may form in a broad spectrum of masses to begin with. In fact, it has been found that a PBH mass function spanning from  $1 M_\odot$  to  $1000 M_\odot$ , where the DM fraction does not exceed 40% at any mass, is still viable. In general, several works have pointed out that a recalculation of these constraints for wide mass distributions tends to have an effect of relaxing a few of these exclusion zones.

In addition, Figure 11 shows that the constraints act at different epochs. For instance, most lensing and dynamical constraints are important near present time  $z \simeq 0$ , constraints from type IA supernovae at  $z \simeq 1$ , those arising from X-rays and the cosmic infrared background at  $z \simeq 10$  and finally the CMB and accretion

constraints only play a role at  $z \simeq 10^3 - 10^4$ . Overall, one can think of PBHs model where these constraints are evaded by way of a time evolving mass function through merging and accretion, enough not to be excluded by accretion constraints at early times, as well as those arising from gravitational lensing at late times, i.e. PBHs with mean masses growing from  $0.01 - 10 M_\odot$  early on, to  $10 - 1000 M_\odot$  in the late Universe.

Moreover, the constraints we laid out were derived assuming little PBHs clustering. However, it was found that clustered PBHs in the stellar mass range could in principle comprise the totality of DM. In general, it is expected that taking into account clustering is bound to alleviate most of the constraints, and shift them to lower masses.

\*\*\*

## Chapter references

The introduction to PBHs was taken from the reviews by Carr *et al.* (2023) [39] and the one by Escrivà *et al.* (2022) [101]. The sections detailing the formation mechanism for PBHs, their abundance and the models of inflation are all a mix of two reviews, namely the one from Sasaki *et al.* (2018) [102] and the one by Escrivà *et al.* (2022) [101]. The part regarding the clustering is a mixture of three papers, respectively Baldauf *et al.* (2013) [103], De Luca *et al.* (2020) [104] and Desjacques & Riotto (2019) [105]. Finally, the part describing the role of PBHs in structure formation was taken from Carr & Silk (2018) [106] and the list of constraints is a summary of Carr *et al.* (2021) [81].

## 5 Testing the $\Lambda$ CDM Model

A big thank you goes to Rudi. Our countless hours in Uni-Mail and his fundamental contributions to this work have made it possible for me to finish my thesis.

*"Dude, you're trying to understand what Python is doing! No one does that! In Python you just DO"*

In the standard  $\Lambda$ CDM paradigm, the dark matter overdensity field gets a head start in its growth compared to the baryons and photons, meaning that after recombination the baryons can collapse in already formed dark matter halos and give way to galaxy formation. How, then, are dark matter halos related to galaxies? One simple approximation is to assume that the maximum amount of baryonic material that can be found in a halo is restricted by the cosmic baryon fraction  $f_b = \Omega_b/\Omega_m$ :

$$M_b = f_b M_{\text{halo}} \quad (5.1)$$

This means that the stellar mass in a halo will have an upper bound given by  $M_*(M_{\text{halo}}) \leq M_b(M_{\text{halo}})$ . It turns out that from these simple considerations we can impose stringent constraints on the cosmological model without actually having to lay out the more complicated phenomena involved in galaxy formation. For instance, we can calculate the cumulative comoving star mass density contained in dark matter halos of a specified mass at a given redshift. Using the mass function  $\frac{dn_h}{dM}(z, M)$  we introduced in (3.35), the comoving mass density of halos above a given halo mass threshold is given by:

$$\rho_m(> M_{\text{halo}}, z) = \int_{M_{\text{halo}}}^{\infty} dM M \frac{dn}{dM}(z, M) \quad (5.2)$$

By making the rather straightforward assumption that the largest stellar content of a halo, given the baryon-DM mass fraction, is  $M_* = \epsilon f_b M_{\text{halo}}$ , where  $\epsilon \leq 1$  is the conversion efficiency of baryons into stars, the *stellar formation efficiency* (SFE), we can relate the comoving mass density of stars contained in halos more massive than a certain threshold to the density of matter:

$$\rho_*(> M_{\text{halo}}, z) = \epsilon f_b \rho_m(> M_{\text{halo}}, z) \quad (5.3)$$

Note that this simple approximation fails miserably. Galaxy masses do not scale linearly with halo masses, but the relationship appears to be more complicated (the goal of modeling galaxy formation from known physics within  $\Lambda$ CDM is an entire field of its own).

In turn (5.3) can be rewritten by specifying the stellar mass of the galaxies  $M_*$ :

$$\rho_*(> M_*, z) = \epsilon f_b \rho_m(> \frac{M_*}{f_b \epsilon} = M_{\text{halo}}, z) = \epsilon f_b \int_{M_{\text{halo}}}^{\infty} dM M \frac{dn}{dM}(z, M) \quad (5.4)$$

The question of the consistency of stellar mass functions and the underlying cosmological dark matter halo mass functions has become considerably more important and pressing with the successful launch of the James Webb Space Telescope in 2021 [107], to which followed a swarm of observations of high-redshift galaxy candidates [2, 14–21].

These objects are surprisingly massive galaxy candidates at  $z \gtrsim 10$  with inferred stellar masses  $\gtrsim 10^7 M_\odot$ , among which the current redshift record is held by [17] at  $z \gtrsim 16.7$ . Importantly, these findings are based on photometry only, making their redshifts and spectral energy distribution fits uncertain.

Still, a subset of them, with lower masses, have been recently spectroscopically confirmed [108, 109]; these are galaxies in the redshift range  $10 \lesssim z \lesssim 13$  and have masses of order  $10^7 - 10^8 M_\odot$ . Very rapidly, studies have shown that large cosmological hydrodynamical simulations (in  $\Lambda$ CDM) are compatible with these confirmed high redshift galaxies [110, 111].

However, the more massive objects, yet to be confirmed, situated so early in cosmic history, would be difficult to reconcile with the theory built upon  $\Lambda$ CDM [1].

To illustrate the problem, we work with candidates found in [2]<sup>1</sup> (L22 from now on) shown in the following figure:

---

<sup>1</sup>All the data in L22 can be found in <https://github.com/ivolabbe/red-massive-candidates>

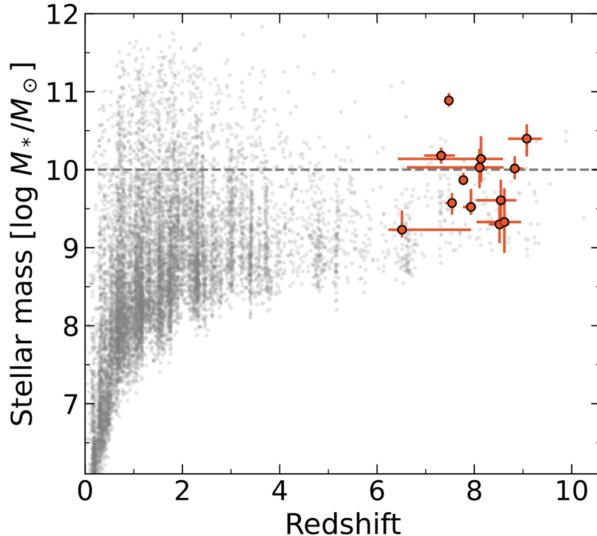


Figure 12: Data points shown in red found by L22. All galaxies have photometric redshifts in the range  $6.5 < z < 9.1$  and six galaxy candidates have masses  $M_* > 10^{10} M_\odot$ . Picture taken from [2].

As shown in Figure 12, the data consists of 13 objects (in red) with photometric redshifts  $6.5 < z < 9.1$ , of which 6 have fiducial masses  $> 10^{10} M_\odot$  with redshifts  $7.4 < z < 9.1$ . The brightest galaxy in the sample is at  $z = 7.5$  and may have a mass that's as high as  $M_* \approx 10^{11} M_\odot$ , more massive than present-day Milky Way.

The masses derived for the samples are intriguing if placed in the context of previous studies, since no candidate galaxies with  $M_* > 10^{10.5} M_\odot$  had been found beyond  $z \sim 7$ , and no candidates with  $M_* > 10^{10} M_\odot$  beyond  $z \sim 8$ .

This, in fact, is in contrast with the  $\Lambda$ CDM mass function fits, as the following discussion will make clear. First we look at the following figures:

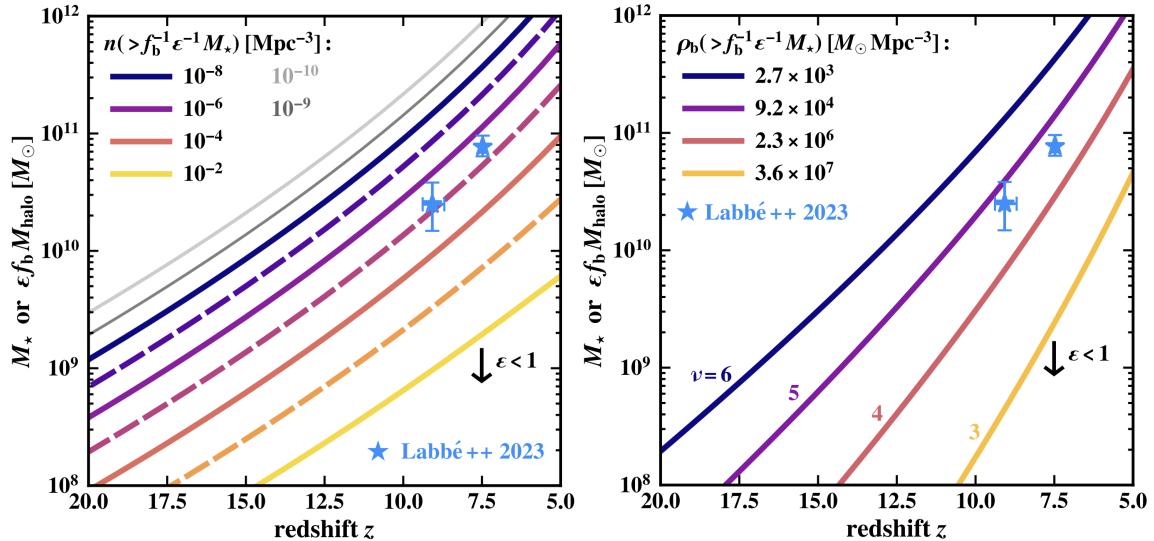


Figure 13: *Left:* lines of constant comoving halo number densities as functions of redshift and stellar mass. *Right:* lines of fixed comoving baryon density as functions of redshift and stellar mass/peak height. In both pictures the SFE is set to 1, and the lines shift downwards as  $\epsilon$  decreases from its ideal value, as indicated by the arrows in the bottom right. The L22 galaxies are shown as blue stars and are represented with uncertainties indicating  $1\sigma$  intervals. Both figures are taken from [1]

On the left, Figure 13 shows contour lines of fixed cumulative halo number densities (mass function integrated over all masses) ranging from  $10^{-10} Mpc^{-3}$  to  $10^{-2} Mpc^{-3}$ , as functions of redshift and stellar masses, related to the halo masses through  $M_* = \epsilon f_b M_{\text{halo}}$ , and taking  $\epsilon = 1$  for the moment, albeit it being an unreasonable value. The two most massive high-redshift galaxy candidates from L22, at  $z \approx 7.5$  ( $M_* \approx 10^{11}$ ) and  $z \approx 9.1$  ( $M_* \approx 10^{10.5}$ ) are shown in Figure 13 as blue stars. These objects, even with ideal SFE conditions, are unexpectedly massive, requiring a comoving number density at that redshift no higher than  $\approx 10^{-5.2} Mpc^{-3}$  ( $\epsilon = 1$ ); in fact we get  $\approx 10^{-7}$  and  $\approx 10^{-9.3}$  for  $\epsilon = 0.32$  and  $\epsilon = 0.1$ . The problem here is that by comparison, the candidates were found in a survey of  $38 \text{ arcmin}^2$ , a volume of  $V \approx 10^5 Mpc^3$  (at each of the redshifts), far too small to contain both of these objects.

The right panel of Figure 13 instead views the fundamental discrepancy at play from a different standpoint, by looking at the comoving baryon density for fixed peak height, as a function of redshift and stellar mass once again. We know that typical halos at  $z$  have  $v \sim 1$ , which corresponds to 24% of the mass in the Universe residing in halos at least that massive. Larger values of the peak height correspond to increasingly rare objects at that epoch. In fact, as we can see in the picture, the L22 galaxies would require at least  $v \approx 4.5$  for  $\epsilon = 1$  meaning that, by multiplying the  $\rho_b$  by the volume of the survey, at most a fraction  $6.2 \times 10^{-5}$  of the baryons are contained in halos massive enough to host these galaxies, which is in contrast with the measured galaxy candidates. Relaxing the SFE results in even worse predictions, as we get  $v \approx 5.4$  and  $v \approx 6.4$  for  $\epsilon = 0.32$  and  $\epsilon = 0.1$ .

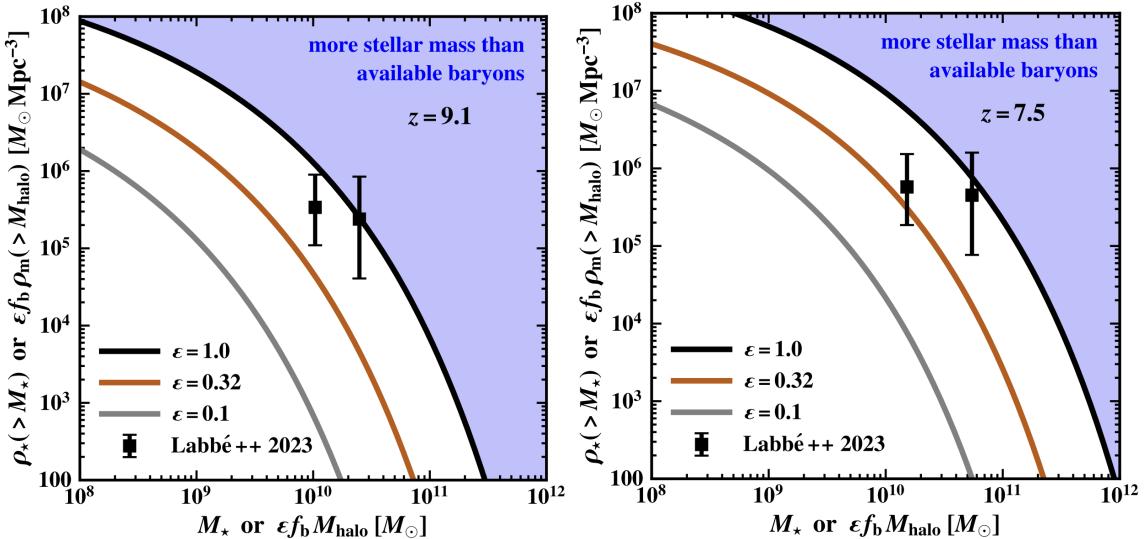


Figure 14: Comoving cumulative stellar mass densities (5.4) as a function of  $M_*$  and plotted for various SFEs, at the two representative redshifts. It is clear from both figures that the data (black markers) in L22 are compatible with  $\Lambda$ CDM predictions only if close to all baryons in halos are turned into stars, which is an unrealistic scenario. This means that the required cosmic allotment of baryons  $f_b$  is way higher than the maximum allowed in our model. Both figures are taken from [1]

Figure 14 shows instead the cumulative stellar mass density  $\rho_*(>M_*)$  as introduced in (5.4) as a function of stellar mass  $M_*$  for the two representative redshifts  $z \approx 9.1$  and  $z \approx 7.5$ . We can see that the L22 objects lie at the extreme of the  $\Lambda$ CDM expectations, implying implausible values of stellar formation efficiency  $\epsilon(z = 9.1) \approx 0.99$  and  $\epsilon(z = 7.5) \approx 0.84$ . Even when considering the  $1\sigma$  error, the data become consistent with an already high SFE  $\epsilon(z \approx 9) \geq 0.57$ . Assuming more realistic values like  $\epsilon = 0.1$  or  $\epsilon = 0.32$  yields far more obvious discrepancies at these redshifts.

Summarizing the problem, it is found that, if confirmed, the massive early galaxies found by L22 are in serious tension with the  $\Lambda$ CDM expectations, since they are only consistent with the model only through unrealistic levels of SFE  $\epsilon \geq 0.57$ . These efficiencies are not observed in the low-redshift Universe, with

$\epsilon \lesssim 0.2$  for all galaxies. Even though we expect theoretically an increase in SFE with redshift,  $\epsilon \geq 0.57$  is still highly improbable [112]. Even then, pushing the SFE to its limit would result in a high abundance of ionizing photons that could contradict the measurements of the cosmic reionization history.

A high SFE would imply that most of the stellar mass at  $z \approx 8, 9$  resides in the most massive galaxies which, although an extreme conclusion, would be in line with the notion that the central regions of present-day massive elliptical galaxies hold the oldest stars in the Universe, and with the finding that already at  $z \gtrsim 2$  the stars in the central regions of massive galaxies make up around 10% to 20% of the total stellar mass density at that redshift.

This discrepancy is robust to uncertainties in the cosmological parameters in  $\Lambda$ CDM, since the precision of the relevant parameter is at  $\lesssim 1\%$  level [5], implying that potential new physics might need to be invoked if the explanation doesn't reside in an error in the derivation of the stellar densities, or in postulating that basically all the stellar mass density in the Universe at those redshifts resides in massive galaxies (high SFE).

## 5.1 Accelerating Structure Formation

Possible solutions to the problem involve extending the base  $\Lambda$ CDM model by accelerating structure formation, and multiple attempts have been made in order to achieve this.

Firstly, we hinted at the possibility of trivially accepting a high SFE, which however carries problems of its own. A solution was proposed invoking "fuzzy dark matter" models that suppress the abundance of small-scale structures [22]. Instead, a recent study has looked at the effect of an "early dark matter" component, which allows for earlier structure formation and a higher abundance of halos at high redshifts, yielding a lower SFE, from  $\epsilon = 0.72$  to  $\epsilon = 0.99$  [23, 24].

Another philosophy involves modifying the small-scale power spectrum. The power spectrum is, in fact, strictly constrained by HST measurements of galaxy clustering, Ly- $\alpha$  forests data and the UV luminosity function at scales  $k \lesssim 3 h Mpc^{-1}$ ; all of these measurements are fitted very well by the  $\Lambda$ CDM model. At smaller scales, however, data points are still scarce and deviations from the model could affect the abundance of  $M \lesssim 10^{11} M_\odot$  objects. There have recently been various proposals that hinge on this observational shortcoming, ranging from axion-like particle dark matter models [25, 26], to the introduction of non-Gaussian tails in the primordial power spectrum [27, 28], of which the existence and formation is still subject to intense work.

Other works have looked at the possibility of "supermassive dark stars", early stars made almost entirely of hydrogen and helium but powered by dark matter heating rather than fusion, that grew from  $1 M_\odot$  at their inception to masses  $> 10^6 M_\odot$ , making them visible by JWST [29].

Finally, studies have shown that cosmic strings, providing a highly non-Gaussian contribution to the density field, could help with the generation of early massive galaxies [30].

We consider instead the possibility of enhancing the matter power spectrum with the introduction of a population of primordial black holes, a scenario which has been studied extensively in the context of the JWST results [3, 26, 31–33].

We follow at first the authors in [3], who employ a monochromatic PBH mass spectrum and consider the scenario in which PBHs produce a Poissonian noise on top of the adiabatic perturbations. The idea is again to take advantage of the lack of small-scale data, using a linear power spectrum for the isocurvature perturbations [76]:

$$P(k, z) = P_{ad}(k, z) + P_{iso}(k, z) \quad (5.5)$$

where  $P_{ad}(k, z)$  is the linear matter power spectrum and:

$$P_{iso}(k, z) = \frac{(f_{PBH} D_+(z))^2}{\bar{n}_{PBH}} \quad (5.6)$$

with linear growth factor  $D_+(z)$  given by:

$$D_+(z) = \left( 1 + \frac{3\gamma}{2a_-} \frac{a}{a_{eq}} \right)^{a_-} - 1 \quad (5.7)$$

The authors expect linear perturbation theory to break down at a certain small (mass) scale, overproducing structures below it, therefore a conservative cut off is put at  $M \sim m_{\text{PBH}}$ :

$$P_{\text{iso}}(k, z) = 0 \quad k > (2\pi^2 \bar{n}_{\text{PBH}} / f_{\text{PBH}})^{1/3} \quad (5.8)$$

Using the full power spectrum (5.5), the Press-Schechter formalism is used to calculate the comoving stellar density, with the Sheth-Tormen halo multiplicity and a Gaussian window function:

$$\rho_*(> M_*, z) = \epsilon f_b \int_{M_{\text{halo}}}^{\infty} M \frac{dn(M, z, f_{\text{PBH}}, m_{\text{PBH}})}{dM} dM \quad (5.9)$$

They then go on to compare the results of this calculation with the following data of stellar masses  $M_*$  and redshifts in L22:  $(M_* \sim 10^{10} M_\odot, z \sim 8)$  and  $(M_* \sim 10^{10.5} M_\odot, z \sim 10)$ . In particular, L22 reports  $\rho_*(\gtrsim 10^{10} M_\odot) \approx 1.3_{-0.6}^{+1.1} \times 10^6 M_\odot \text{ Mpc}^{-3}$  and  $\rho_*(\gtrsim 10^{10.5} M_\odot) \approx 9_{-6}^{+11} \times 10^5 M_\odot \text{ Mpc}^{-3}$ .

The  $(f_{\text{PBH}}, m_{\text{PBH}})$  parameters that result in massive galaxies are found to be:

$M_*$ [ $M_\odot$ ]	$\epsilon$	$z$	$\rho_*(> M_*, z)$ [ $M_\odot \text{ Mpc}^{-3}$ ]	$f_{\text{PBH}} \cdot m_{\text{PBH}}$
$10^{10}$	1	8	$\sim 10^6$	$2.4 \times 10^4$
$10^{10}$	0.1	8	$\sim 10^6$	$1.8 \times 10^6$
$10^{10.5}$	1	10	$\sim 9 \times 10^5$	$1.8 \times 10^5$
$10^{10.5}$	0.1	10	$\sim 9 \times 10^5$	$6.1 \times 10^6$

Table 2: PBH parameters space results in [3]

To illustrate their results, they use the three following models:

Model	$m_{\text{PBH}}$ [ $M_\odot$ ]	$f_{\text{PBH}}$	$m_{\text{PBH}} f_{\text{PBH}}$ [ $M_\odot$ ]
M1	$3 \times 10^5$	0.0003	90
M2	$10^9$	$10^{-5}$	$10^4$
M3	$10^{10}$	$10^{-4}$	$10^6$

Table 3: Representative PBH models used in [2]

For which the power spectra are, at  $z = 0$ :

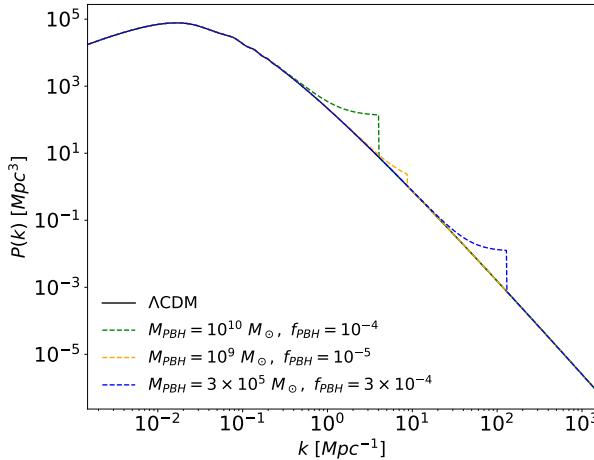


Figure 15: Enhanced power spectra for the three models that are represented in Table 3

We now question whether it is possible to invoke lighter PBHs masses, or in other words relax the constraints in Table 2. In fact, we know that the range in which (5.5) holds is  $f_{\text{PBH}} \lesssim 1$  and  $f_{\text{PBH}} \gtrsim 0.002$

from (4.33) at  $z = 10$ , so Table 2 implies  $m_{\text{PBH}} \gtrsim 10^6 M_\odot$ , heavy black holes that are very hard to form in the standard collapse scenario.

Firstly we investigate whether for the above models of PBHs, the linear approximation is appropriate for the values of  $M_{\text{halo}}$  that are important for the JWST galaxies. For this purpose, we calculate the quasilinear and nonlinear transitions, and ask if the corresponding halo mass at these scales is of the order of  $M_{\text{halo}} \approx 10^{11}$ , using (4.26) and (4.29), and for the mass taking (3.29) with a radius  $R = 2\pi/k$ . The results are the following:

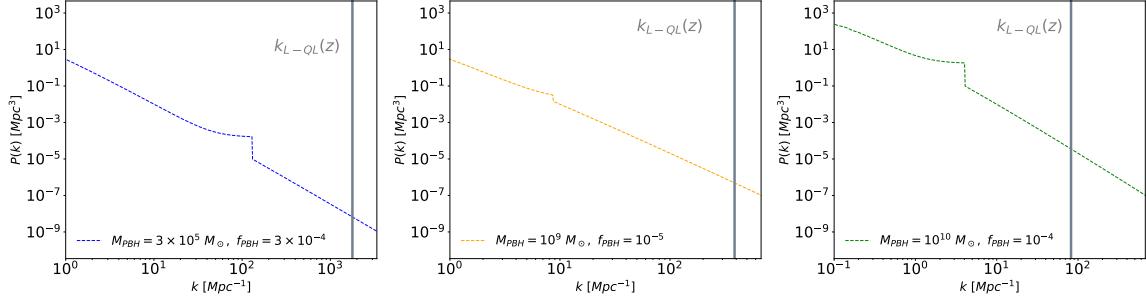


Figure 16: Power spectra (5.5) for the different representative models in Table 3 at  $z \sim 10$  and the quasi linear transition scale (4.26) plotted as the grey vertical line. As we can see the quasi linear transition happens beyond the relevant mass scales.

where the line is the result for the quasi-linear transition scale. Numerically:

Model	$k_{\text{L-QL}} [\text{Mpc}^{-1}]$	$M_{k_{\text{L-QL}}} [M_\odot]$
M1	$1.77 \times 10^3$	$7 \times 10^2$
M2	$3.8 \times 10^2$	$7 \times 10^4$
M3	81	$7 \times 10^6$

Table 4: Results for the linear to quasi linear transitions scales for the models used in Table 3. It is clear that the scales are way beyond the range of interest for the enhancement of halo masses  $M_{\text{halo}} \sim 10^{11} M_\odot$

As we can see from both the pictures and the numerical results, the quasi linear scale sits to the right of the power spectrum amplification at the mass scale we are interested in, so that a further amplification by primordial black holes' clustering is in fact not relevant in this case.

Next, we notice that the authors use a Gaussian window function to calculate the star density, so it might be useful to consider the same problem with different filters. This is motivated by the following figure:

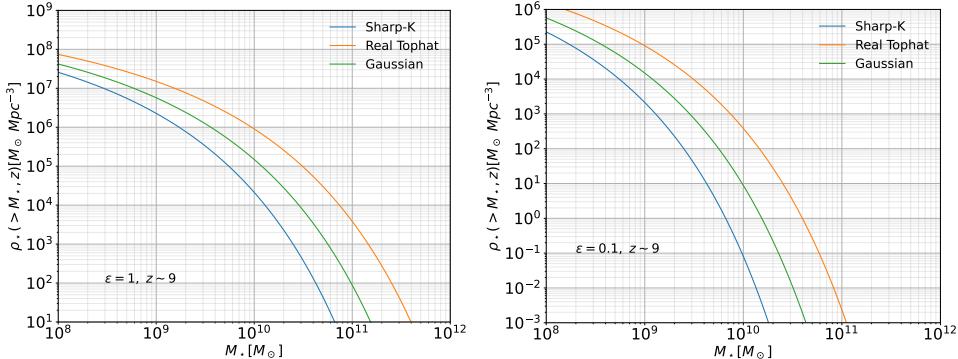


Figure 17: Cumulative stellar mass density at a representative redshift ( $z \sim 9$ ) with  $\epsilon = 1$  (left) and  $\epsilon = 0.1$  (right) with three different filters. It is clear that the stellar density calculated with a real space tophat window function is considerably bigger compared to the other ones.

Figure 17 is the  $\Lambda$ CDM  $\rho_*$  calculated with three different filters at  $z \sim 10$ , for the extreme cases of the SFE. We can see clearly that the real tophat window stellar density, at the masses of interests  $M_* \sim 10^{10} M_\odot$ , is at least an order of magnitude bigger than the next one. In addition, the difference is accentuated when we lower the stellar formation efficiency to  $\epsilon = 0.1$ , which we know to be closer to the real value compared to the extreme  $\epsilon = 1$ .

Using the real space tophat smoothing we thus proceed and calculate the parameter space  $(f_{\text{PBH}}, m_{\text{PBH}})$  required to explain the JWST results. The following table shows the changes that we get by just switching the filter:

$M_*$ [ $M_\odot$ ]	$\epsilon$	$z$	$\rho_*(> M_*, z)$ [ $M_\odot \text{ Mpc}^{-3}$ ]	$f_{\text{PBH}} \cdot m_{\text{PBH}}$
$10^{10}$	1	8	$\sim 10^6$	?
$10^{10}$	0.1	8	$\sim 10^6$	$10^4$
$10^{10.5}$	1	10	$\sim 9 \times 10^5$	$3 \times 10^4$
$10^{10.5}$	0.1	10	$\sim 9 \times 10^5$	$10^5$

Table 5: PBH parameters space results with a real space tophat window function

As we can see, if we compare the results of Table 5 with Table 2, the constraints on the PBH parameters are not as stringent, since we would need to invoke the existence of  $m_{\text{PBH}} \sim 10^4 M_\odot$  instead of  $m_{\text{PBH}} \sim 10^6$  as the results in [3] suggest. (Note that we place a question mark at the top since the stellar density doesn't require in principle PBHs.)

### 5.1.1 The Accretion of Dark Matter Halos

In this section we try to explore a different path, namely the issue regarding the origin of the L22 galaxies. These, in fact, might have also come from younger and smaller sized halos that accreted matter over a period of time.

As a way of studying this problem we use excursion sets and consider the accretion rates (3.54) in the framework where the power spectrum is enhanced (5.5). In this case the variance  $\sigma^2$  is, using the Fourier top-hat window (3.39):

$$\begin{aligned} \sigma^2(M, z) &= \int_0^\infty \frac{dk}{k} \frac{k^3}{2\pi^2} (P_{\text{ad}}(k, z) + P_{\text{iso}}(k, z)) |W(kR)|^2 = \sigma_{\text{ad}}^2(z, M) + \int_0^{1/R} dk \frac{k^2}{2\pi^2} \frac{f_{\text{PBH}}^2}{\bar{n}_{\text{PBH}}} D_+^2(z) \\ &= \sigma_{\text{ad}}^2(z, M) + \frac{1}{2\pi^2} \frac{f_{\text{PBH}}^2}{\bar{n}_{\text{PBH}}} D_+^2(z) \frac{1}{3R^3} = \sigma_{\text{ad}}^2(z, M) + \frac{f_{\text{PBH}}^2}{\bar{n}_{\text{PBH}}} D_+^2(z) \frac{2}{9\pi} \frac{\bar{\rho}}{M} \\ &\equiv \sigma_{\text{ad}}^2(0, M) \left( \frac{D_{\text{ad}}(z)}{D_{\text{ad}}(0)} \right)^2 + \sigma_{\text{iso}}^2(0, M) D_+^2(z) \end{aligned} \quad (5.10)$$

If our goal is to use the excursion set formalism, we need to shift the time dependence from the variance  $S$  to the barrier  $\delta_c$ . Since at first glance the time dependence of  $\sigma^2$  doesn't look separable like in the adiabatic case, we avoid the problem by taking a closer look at (5.10) and in particular at its growth factors (2.180), (5.7), for which we find the following approximations at our redshifts of interests  $a \sim 10^{-1}$ :

$$\begin{aligned} D_{\text{ad}}(a) &\approx a \\ D_+(a) &\approx \left( \frac{3\gamma}{2a_-} \frac{a}{a_{\text{eq}}} \right)^{a_-} \approx \frac{1}{2} \frac{3\gamma}{2a_-} \frac{a}{a_{\text{eq}}} \end{aligned} \quad (5.11)$$

where the approximation in the first equation is nearly perfect (better than a 0.5% error in the range  $8 < z < 50$ ) and the 1/2 in the second equation fits the l.h.s. to better than  $\approx 5\%$  in the range  $8 < z < 20$ .

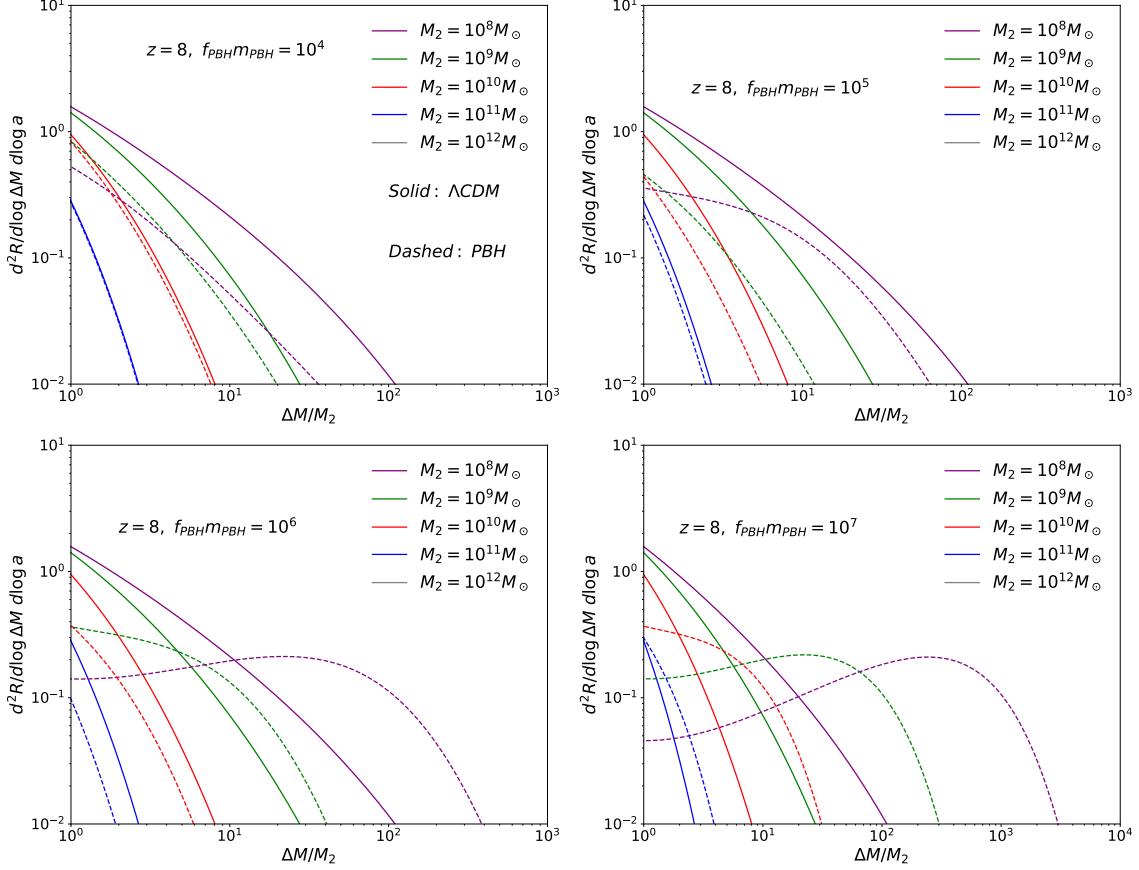
We see that in both of the above equations the  $a$  can be factorized, bringing (5.10) to:

$$\begin{aligned}\sigma^2(z, M, f_{\text{PBH}}, m_{\text{PBH}}) &\approx \left[ \sigma_{\text{ad}}^2(0, M) \left( \frac{1}{D_{\text{ad}}(z=0)} \right)^2 + \sigma_{\text{iso}}^2(0, M) \left( \frac{1}{2} \frac{3\gamma}{2a_-} \frac{1}{a_{\text{eq}}} \right)^2 \right] a^2 \\ &\equiv \left[ \sigma_{\text{ad}}^2(0, M) \left( \frac{1}{D_{\text{ad}}(z=0)} \right)^2 + \sigma_{\text{iso}}^2(0, M) \left( \frac{1}{2} \frac{3\gamma}{2a_-} \frac{1}{a_{\text{eq}}} \right)^2 \right] d(a)^2\end{aligned}\quad (5.12)$$

where  $d(a) \equiv a$  is our new growth factor. We are now in a position to use the excursion set formalism with PBHs, and in particular equation (5.15), with the new definition:

$$\omega(a) = \frac{\delta_c}{d(a)} = \frac{\delta_c}{a} \quad (5.13)$$

The following figures represent the halo accretion rates with the modified variance (5.10) for various models of PBHs, which depend on the parameters only in the combination  $f_{\text{PBH}}m_{\text{PBH}}$ :



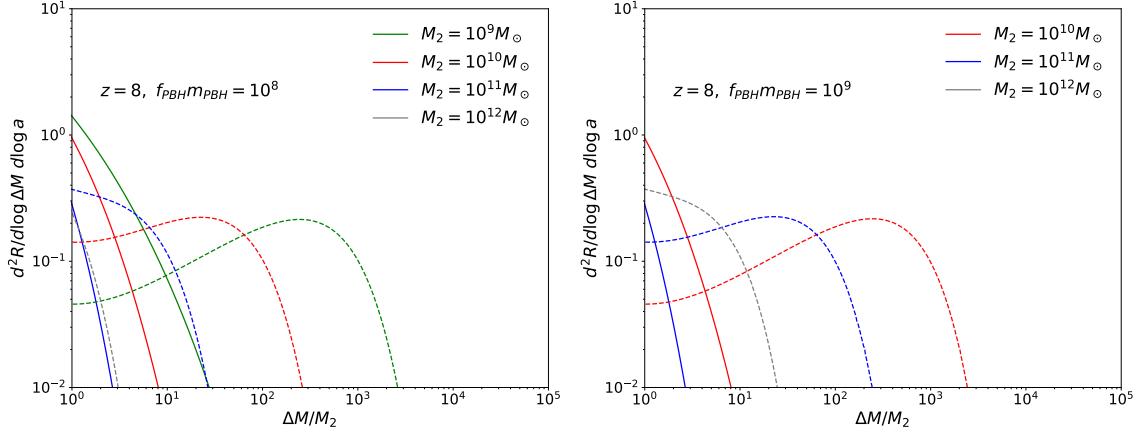


Figure 19: Accretion rates for the  $\Lambda$ CDM model (solid) and the enhanced PBH models for various values of the product  $f_{\text{PBH}} m_{\text{PBH}}$ , since the rate (3.54) only depends on the parameters in that combination.

It is also possible to integrate these plots, finding the amount of mass accreted by an initial halo of mass  $M_2$  per logarithmic interval of time:

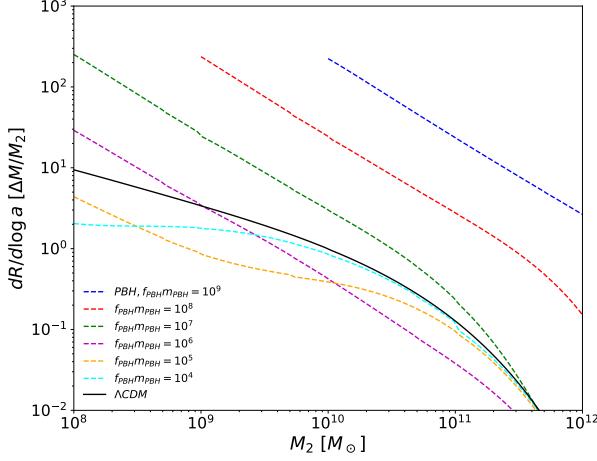


Figure 20: Total mass accretion rates for every PBH model considered in Figure 19.

The above figures suggest that the halos found by JWST may also be a product of the accretion mechanism of smaller halos at a higher redshift. In the Press-Schechter calculation of the stellar mass density, we are in fact only counting the contribution that comes from halos formed at  $z \sim 10$ . Therefore, we now set to find the contribution of the accreting halos to the  $\rho_*$ .

First, in this analysis we are using the sharp-k window function, so it is necessary to compute the allowed parameter space in this case beforehand, not taking into considerations the transition probabilities. We use the JWST results again in L22(v3) to find:

$M_*$ [ $M_\odot$ ]	$\epsilon$	$z$	$\rho_*(> M_*, z)$ [ $M_\odot \text{ Mpc}^{-3}$ ]	$f_{\text{PBH}} \cdot m_{\text{PBH}}$
$10^{10}$	1	9	$\sim 5 \times 10^5$	$4 \times 10^5$
$10^{10}$	0.1	9	$\sim 5 \times 10^5$	$2 \times 10^7$
$10^{10.5}$	1	8	$\sim 5 \times 10^5$	$10^6$
$10^{10.5}$	0.1	8	$\sim 5 \times 10^5$	$4.5 \times 10^7$

Table 6: PBH parameters space results with a sharp-k window function

We redefine the halo mass function that enters (5.4) to account for the probability of a halo of mass  $M_i$  at redshift  $z_i$  to transition (via accretion) to a final JWST halo of mass  $M_f$  at redshift  $z_f$ , that is, the cumulative stellar mass density becomes:

$$\rho_*(> M_{*,f}, z_f) = \epsilon f_b \int_{M_{\text{halo},f}}^{\infty} \left| \frac{dS_f}{dM_f} \right| dM_f \int_0^{M_f/10} dM_i M_i \frac{dn(M_i, z_i)}{dM_i} P(M_i \rightarrow M_f, z_i \rightarrow z_f) \quad (5.14)$$

where we imply the usual relationship between halo mass and its stellar content's mass  $M_{\text{halo},f} = M_{*,f}/\epsilon f_b$ , taking the transition probability from (3.53):

$$P(M_i \rightarrow M_f, z_i \rightarrow z_f) = \frac{1}{\sqrt{2\pi}} \frac{\omega_f(\omega_i - \omega_f)}{\omega_i} \left( \frac{S_i}{(S_f(S_i - S_f))} \right)^{3/2} \exp \left( -\frac{(\omega_f S_i - \omega_i S_f)^2}{2S_f S_i (S_i - S_f)} \right) \quad (5.15)$$

This in fact gives the probability for the halo  $M_i$  to gain a mass  $M_f - M_i$  by accretion in the time interval  $\Delta z$ , where we have put  $z_i = z_f + \Delta z$ . The final stellar mass density will be the sum of the Press-Schechter term and (5.14).

Before showing the results we first note that, in order to speed up the computation, we fit both the variance and its derivative w.r.t. the mass to a 15th and 8th degree polynomial respectively, which are both precise to 0.03% in the mass range  $2 < M_*/M_\odot < 17$ ; note that the degree of the polynomial for the derivative is lower because we find that the results are not much sensitive to it, thus allowing us to save even more computation time.

In order to proceed with the calculation, we need to first study the dependence of the integral and the integrand in (5.14) on the parameters and on the upper/lower bound of the integration. We therefore look at the integrand at a fixed  $z_f$  and  $M_f$  for various values of the  $\Delta z$  parameter, as a function of  $M_i$ . Since we are working with  $M_* \approx 10^{10} M_\odot$ , the following plots are with  $M_f = 10^{11} M_\odot, M_f = 10^{12} M_\odot, M_f = 10^{13} M_\odot$ ; we also take  $f_{\text{PBH}} m_{\text{PBH}} \approx 5 \times 10^5$  as a reference (the plots for the other PBH parameters are the same, but with a shift in the y direction):

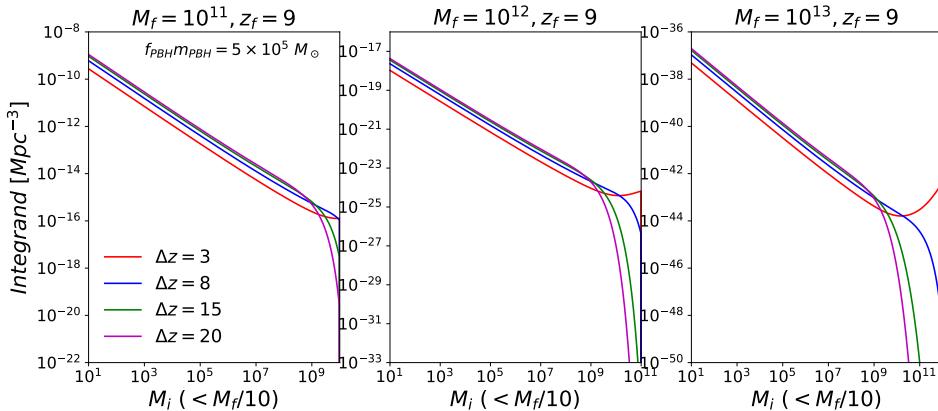


Figure 21: The plot represents the integrand of (5.14) for fixed  $M_f$  (outer variable) and as a function of  $M_i$  (inner variable) with  $10M_\odot < M_i < M_f/10$ ; we also show how the integrand changes as  $\Delta z$  is varied ( $z_f = 9$ ). The PBH parameters are such that  $f_{\text{PBH}} m_{\text{PBH}} \approx 5 \times 10^5$ .

Although these pictures are just representative for some parameters, there are some important observations to be made from them. In fact we see immediately that the most important contributions to the integral will come from the smaller masses, even though they are more distant to the  $M_f$ : since the mass function diverges for low halo masses, this behavior tells us that it is the dominant term in the integrand. In addition we find that the inner integral is insensitive to the upper bound of the integration since, although the transition probability increases for  $M_i \rightarrow M_f$ , it cannot make up for the mass function that decreases too rapidly. In the following plot we change the both the upper and lower bound of the inner integral:

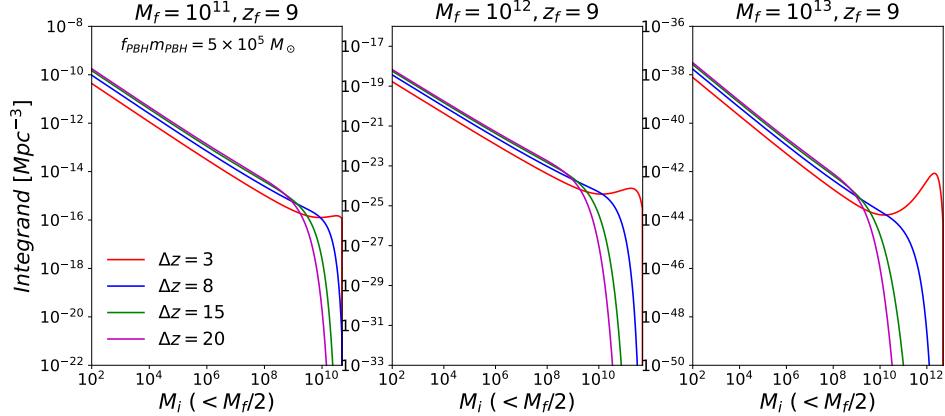


Figure 22: The plot represents the integrand of (5.14) for fixed  $M_f$  (outer variable) and as a function of  $M_i$  (inner variable) capped at  $100M_\odot < M_i < M_f/10$ ; we also show how the integrand changes as  $\Delta z$  is varied ( $z_f = 9$ ). The PBH parameters are such that  $f_{\text{PBH}} m_{\text{PBH}} \approx 5 \times 10^5$ .

As we just said, we can see that the function is not sensitive to the inner integral's upper bound, and it is only slightly sensitive to the lower bound (as is clear by comparing the value of the function at smaller masses between Figure 21 and Figure 22). Lastly it is obvious that the integral will not change significantly when  $M_f \gtrsim 10^{13} M_\odot$ .

In this sense the integral will be sensitive mostly to the  $\Delta z$  parameter, therefore we plot, for every set of parameters in Table 6, the stellar mass density as a function of  $\Delta z$ , divided by its value when taking out the transition probability which we call  $\tilde{\rho}_* \approx 5 \times 10^5 M_\odot \text{ Mpc}^{-3}$  (the Press-Schechter term):

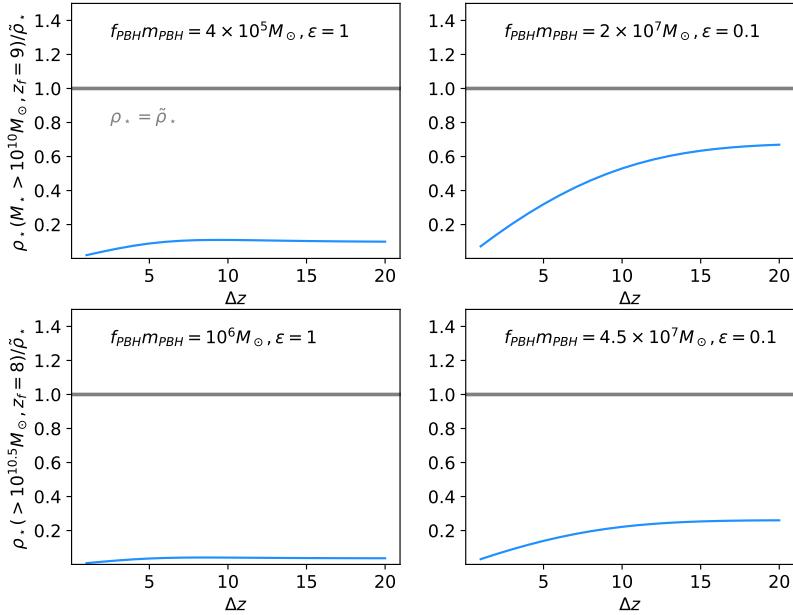


Figure 23: The cumulative stellar mass density (5.14) divided by  $\tilde{\rho}_* \approx 5 \times 10^5 M_\odot \text{ Mpc}^{-3}$  for the parameters in Table 6. We can see that the accretion model cannot relax the PBHs parameter constraint found in the simpler model without transition probabilities.

These figures show that the new model is unable to relax the PBHs parameters found with Press-Schechter, since the same parameters yield a comoving stellar mass density with excursion sets that's around 10% smaller on average. In this sense the  $\rho_*$  gets its most important contribution from the Press-

Schechter term, whereas the terms coming from halos that accreted are negligible to the calculation.

### 5.1.2 Accretion by Primordial Black Holes

Primordial black holes, as explained, can also form structures via the so called seed mechanism. Although the mainstream view is that SMBHs residing in galactic nuclei form after the galaxy itself, recent findings of high redshifts AGNs by JWST with black holes masses between  $10^6 - 10^8 M_\odot$  have challenged this notion [40–45], since cosmological simulations find it challenging to reproduce the growth of such massive black holes given the limited amount of time since the Big Bang. In particular, these black holes at redshifts  $4 < z < 11$  are over-massive relative to their host galaxies, when compared to the local relation  $M_{\text{BH}} - M_{\text{star}}$ , which may indicate that early epochs BHs may form and grow faster than the stellar populations in the galaxies that host them.

Although many models have been proposed in the literature to explain early AGNs (among other phenomena), there is a possibility of seeding galaxies with heavy PBHs. Partly motivated by this reasoning, in this section we offer an explanation of the L22 galaxies using PBHs with the following peaked mass function:

$$\left( \frac{dn}{dm} \right)_{\text{PBH}} = A \frac{\bar{\rho}_{\text{dm}} f_{\text{PBH}}}{m^2} \exp\left( -\frac{(1-m/m_B)^2}{B} \right) \quad (5.16)$$

where  $m_B$  is the black hole mass around which the function is peaked and  $B$  is a parameter of the model. The proportionality constant is fixed by the physical requirement that:

$$f_{\text{PBH}} = \frac{\int_0^\infty dm m \left( \frac{dn}{dm} \right)_{\text{PBH}}}{\bar{\rho}_{\text{dm}}} \quad (5.17)$$

Note that the assumption we are making here is that the PBH mass function is peaked enough for us to consider only a constant  $f_{\text{PBH}}$ , and not a dark matter fraction of the type  $f = f(m)$ .

In this context, the mechanism by which halos form is the accretion of dark matter around isolated PBHs. In fact, if we take a monochromatic PBH mass function around  $M_{\text{PBH}}$ , the mass fraction in halos of mass  $M_N = NM_{\text{PBH}}$  at redshift  $z$  is found to be [113]:

$$f_N(z) = \operatorname{erf}\left(\sqrt{\frac{N}{N_*(z)}}\right) - \operatorname{erf}\left(\sqrt{\frac{N-1}{N_*(z)}}\right) \quad (5.18)$$

where:

$$N_*(z) \simeq \left( \frac{2600}{1+z} \right)^2 f_{\text{PBH}}^2 \quad (5.19)$$

For our purposes, since the PBH mass function is very peaked, we can use this method to determine the fraction of isolated PBHs by calculating  $f_1(z)$ . The number we get will be the fraction of mass in halos of mass  $M_{\text{halo}} = M_{\text{PBH}} \simeq m_B$ , which are just isolated PBHs. For  $m_B \simeq 10^9 M_\odot$  we have the observational upper constraint  $f_{\text{PBH}} \sim 10^{-4}$ , so we find that for  $f_{\text{PBH}} \lesssim 10^{-4}$ :

$$f_1(z \simeq 10) = \operatorname{erf}\left(\sqrt{\frac{1}{N_*(10)}}\right) \simeq 1 \quad (5.20)$$

In this sense, in our model, all the PBHs we have in the Universe are isolated and accrete dark matter independently, generating a halo mass function of similar form of that of the PBHs. This is in line with the seed effect dominating for  $f_{\text{PBH}} \gtrsim 0.002$  at  $z \simeq 10$ . Note that (5.18) is based on a Poissonian PBHs power spectrum, which is not what we are using, making this just an approximation; a similar analytical result for the seed effect does not yet exist in the literature.

The relationship between the mass of the halos and that of its central black hole is given by [114]:

$$m_H(z) = 3 \left( \frac{1000}{1+z} \right) m_B \quad (5.21)$$

Therefore, the halo mass function resulting from (5.16) is:

$$\left(\frac{dn}{dm}\right)_{\text{halo}} = C \frac{\bar{\rho}_{dm} f_{\text{PBH}}}{m^2} \exp\left(-\frac{(1-m/m_H)^2}{B}\right) \quad (5.22)$$

The constant C is fixed by asking that the numerical densities of halos and PBHs are the same (from (5.20) we see that to every halo corresponds only one black hole), or in other words:

$$\int_0^\infty dm \left(\frac{dn}{dm}\right)_{\text{PBH}} = \int_0^\infty dm \left(\frac{dn}{dm}\right)_{\text{halo}} \implies C = A \cdot \frac{\int_0^\infty dm \frac{1}{m^2} e^{-(1-m/m_B)^2/B}}{\int_0^\infty dm \frac{1}{m^2} e^{-(1-m/m_H)^2/B}} = A \frac{m_H}{m_B} \quad (5.23)$$

We calculate these constants A and C for every case we do. In the following Figure we plot a representative case of the unnormalized mass functions we are using, just to show the shapes, in the case where  $m_B = 4 \times 10^8 M_\odot$  and  $B = 10^{-1.9}$  (the latter we use in every case):

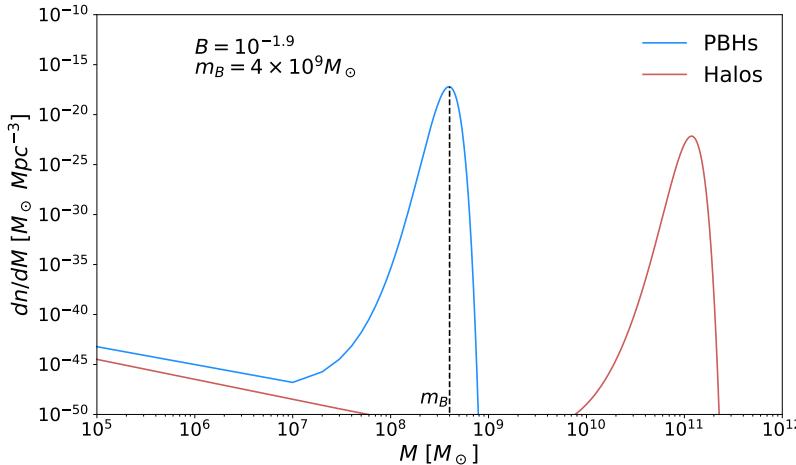


Figure 24: PBH mass function peaked at  $m_B$  in blue and the resulting halo mass function from dark matter accretion.

Once the halo mass function is known, we can calculate the cumulative stellar mass function:

$$\rho_*(> M_*, z) = \epsilon f_b \int_{M_{\text{halo}}}^\infty dm \left(\frac{dn}{dm}\right)_{\text{halo}} = (\epsilon f_b)(C \bar{\rho}_{dm} f_{\text{PBH}}) \int_{M_{\text{halo}}}^\infty dm \frac{1}{m^2} \exp\left(-\frac{(1-m/m_H)^2}{B}\right) \quad (5.24)$$

We find the following results:

$$\rho_*(\gtrsim M_* = 10^{10} M_\odot, z \approx 9) = \begin{cases} (1.5 \times 10^{12} f_{\text{PBH}}) M_\odot Mpc^{-3} & m_B = 4 \times 10^8 M_\odot, \epsilon = 1 \\ (2.5 \times 10^{11} f_{\text{PBH}}) M_\odot Mpc^{-3} & m_B = 2 \times 10^8 M_\odot, \epsilon = 1 \\ (1 \times 10^{10} f_{\text{PBH}}) M_\odot Mpc^{-3} & m_B = 1.85 \times 10^8 M_\odot, \epsilon = 1 \end{cases} \quad (5.25)$$

Reducing the stellar formation efficiency:

$$\rho_*(\gtrsim M_* = 10^{10} M_\odot, z \approx 9) = \begin{cases} (1.5 \times 10^{11} f_{\text{PBH}}) M_\odot Mpc^{-3} & m_B = 4 \times 10^9 M_\odot, \epsilon = 0.1 \\ (1 \times 10^{10} f_{\text{PBH}}) M_\odot Mpc^{-3} & m_B = 2 \times 10^9 M_\odot, \epsilon = 0.1 \end{cases} \quad (5.26)$$

We can see from these results that the stellar density depends very sensitively on the position of the peak of the PBH mass function (and therefore that of the halo mass function). This is because of the lower bound of the integral, that either comes before the peak, or cuts it in half, or misses it completely.

Instead, for the heavier masses we get:

$$\rho_*(\gtrsim M_* = 10^{10.5} M_\odot, z \approx 8) = \begin{cases} (1.3 \times 10^{11} f_{\text{PBH}}) M_\odot Mpc^{-3} & m_B = 7 \times 10^8 M_\odot, \epsilon = 1 \\ (1 \times 10^{10} f_{\text{PBH}}) M_\odot Mpc^{-3} & m_B = 5 \times 10^8 M_\odot, \epsilon = 1 \end{cases} \quad (5.27)$$

and:

$$\rho_*(\gtrsim M_* = 10^{10.5} M_\odot, z \approx 8) = \begin{cases} (1 \times 10^{11} f_{\text{PBH}}) M_\odot \text{Mpc}^{-3} & m_B = 7 \times 10^9 M_\odot, \epsilon = 0.1 \\ (1 \times 10^{10} f_{\text{PBH}}) M_\odot \text{Mpc}^{-3} & m_B = 6 \times 10^9 M_\odot, \epsilon = 0.1 \end{cases} \quad (5.28)$$

Since the observational upper bound is  $f_{\text{PBH}} \sim 10^{-4}$  for the mass  $m_B \sim 10^8 M_\odot$ , we find that this model can explain the JWST results with relatively heavy PBHs,  $m_B \sim 10^8 M_\odot$  for  $\epsilon = 1$  and  $m_B \sim 10^9 M_\odot$  for  $\epsilon = 0.1$  in both cases  $M_* \approx 10^{10} M_\odot$  and  $M_* \approx 10^{10.5} M_\odot$ .

We care to make a few comments on the upper bound given by effects due to dynamical friction. These, in fact, were derived with the requirement of at least two PBHs per structure, which is not realised in our model. This then begs the question: how many PBHs per structure does a realistic model have? How do we analytically calculate this? Although our model was approximating a single PBHs per structure, the answers to these questions are still speculative, since they are proving to be a very hard challenge. The bounds given by dynamical effects will need to be revised when new results, both analytical or numerical, come out.

## 5.2 The NANOGrav Gravitational Wave Signal and Some Perspectives

Depending on the astrophysical or cosmological source, gravitational waves (GW) signals could in principle exist over a wide range of frequencies, which experiments, both present and future, aim to probe.

The detection of GWs in the nHz region follows the original template discussed in [115], where the propagation of light is timed to measure modifications in the distance between freely falling reference masses. The interest in the measurement of low-frequency GWs detection was sparked when the authors in [116] calculated the GW response of electromagnetic signals traveling between Earth and distant spacecrafcts. Later on it was suggested that nHz GWs could be detected using galactic pulsars and the solar system baricenter as references, relying on the regularity of planetary motions and the pulsars' signals [117, 118]. Note that the following properties are crucial. Firstly, the fact that pulsars are such accurate clocks allows for very precise measurements of the pulsars' parameters from the time-of-arrival of their pulses. Secondly, but more importantly, is that even in the case of a measurement of this signal, merely the consistency of amplitude and spectral shape cannot attribute it to GWs; in fact, the signal could arise from intrinsic pulsar processes or from a common systematic noise. Definitive proof of a *stochastic* GW origin is confirmed by establishing the presence of phase-coherent inter-pulsar correlations with a characteristic spatial pattern derived for the first time in [119], the *Hellings-Downs correlation*: for an isotropic gravitational wave background (GWB), the correlation between the timing delays for any pair of pulsars is a universal, quasi-quadrupolar function of their angular separation.

These kind of measurements are carried with pulsar timing arrays (PTAs), since they are sensitive to frequencies in the nHz regime. Excitingly, in 2020 the NANOGrav collaboration announced evidence of a stochastic GWB [120], with similar results obtained by other PTAs [121–123]. Very recent data from NANOGrav [4], EPTA [46], PPTA [47] and CPTA [48] not only strengthen the signal, but also show evidence for the crucial Hellings-Downs correlation.

It is important, at this stage, to speculate as to the origin of this background. The most natural candidate, the one cited in the original NANOGrav article, is a population of supermassive black hole binaries (SMBHB). Indeed, we know that galaxy mergers are the main drivers of hierarchical structure formation over cosmic time and lead to the formation of close SMBHB, the black holes being those at the center of the respective galaxies. The most massive of these SMBHBs, with masses of  $10^9 - 10^{10} M_\odot$ , emit GWs with slowly evolving frequencies, contributing possibly to the nHz signal.

Although the measurements of the GWB are still at an early stage, it was found that the frequency dependence of the observed GWB may slightly differ from that predicted by the simplest SMBHB model, according to which these binaries evolve by losing energy through gravitational waves emission [4, 124].

There are two classes of interpretation for this discrepancy. The first is that these SMBHB may be losing energy due to non-gravitational effect as well, for instance through interactions with their environment. The more radical explanation instead sustains that the GWB may be originating from cosmological sources,

among which some candidates are cosmic strings, a first order phase transition in the early Universe, domain walls, primordial fluctuations and other exotic scenarios. Most of these other scenarios were found to fit the NANOGrav data at least as well as the SMBHB alone [125].

For the purposes of this thesis, some of the literature accounts for this gravitational wave signal invoking PBHs.

Very straightforwardly, the authors in [126] fit a monochromatic population of PBHs to the NANOGrav data, using both models with and without black hole clustering. They find that it is not possible to consistently explain the data with homogeneously distributed PBHs, whereas the clustering helps in the formation of binary systems, thus opening a possibility for a description of the observed signal. This conclusion, however, hinges on the evasion of the constraints on the PBH parameters, given that some were formulated assuming no PBH clustering and using a specific formation mechanism.

On the other hand, the authors in [33] have tried to fit both the NANOGrav and JWST data, assuming no PBHs clustering. They find that the allowed parameter spaces necessary to fit the data only slightly overlap in excluded regions.

There exist in principle other ways in which PBHs could be related to the GWs signal. In particular, the effect of large curvature fluctuations generated during inflation, where the GWs are produced by a second-order effect resulting from scalar perturbations re-entering the horizon, was extensively discussed in the literature [4, 127, 128]. In practice, on top of GWs, these large perturbations could lead to the production of PBHs.

Lastly, it was recently proposed that ultra-low mass PBHs with  $m_{\text{PBH}} \sim 10^8$  g could be the source of the signal [129]. This effect would be due to the nontrivial evolution of first-order scalar perturbations during early matter domination.

The NANOGrav stochastic gravitational wave signal has lately caused a big stir. However, only future PTA datasets may be able to distinguish between the proposed cosmological and astrophysical scenarios. Moreover, observations at higher frequencies will be crucial, since some cosmological scenarios predict a turndown of the frequency spectrum that precludes LISA measurements, while on the other hand the astrophysical models hint at BH mergers detectable by LISA.

Whether primordial black holes have anything to do with the JWST and PTA measurements is still up for debate, which will only be (partially) resolved with the help of richer datasets in the near future.

It is important to underline the work that still needs to be done in the context of PBHs physics as well. Their formation mechanisms could in principle radically change the induced cosmological effects, from the formation of large scale structures to the generation of gravitational wave signals; in addition, the constraints on the PBHs parameters, depending on various assumptions, will need to be revised once the latter become more stringent or, perhaps, more relaxed.

For now, analytical methods have proved to be very challenging, forcing the resolution to numerical simulations; these however are still in their infancy, with the only all-encompassing study halting at  $z \sim 100$  [76].

\*\*\*

## Chapter references

The start and end of this chapter are mostly summaries of sections of various articles, all of which have been already cited when necessary; in particular, the first subsection is essentially taken from [1]. Otherwise, all the results and discussions in between are completely original.

## 6 Conclusions and outlook

New high redshift and high mass galaxies have been discovered by the James Webb Space Telescope. If confirmed, they could pose a serious threat to the  $\Lambda$ CDM model, which underpredicts the stellar mass density.

The physics community has since invoked multiple mechanisms that could in principle accelerate structure formation; they mostly involve modified theories of dark matter/dark energy or exotic inflationary mechanisms. In this work, we focus on the interesting possibility that part of the dark matter content in the Universe is made up of primordial black holes, black holes which didn't form via the collapse of stars, but rather at earlier epochs through, for instance, the collapse of large density perturbations.

We study firstly the Poisson mechanism, which was treated (among other articles) in [3], where it was found that relatively heavy PBHs  $m_{\text{PBH}} \gtrsim 10^5 M_\odot$  were needed. Looking to relax this result, we involve halo accretion and mergers over time through excursion set theory to calculate the stellar mass density. We find that the PBHs masses cannot be decreased, obtaining on average only 10% of the mass density of a simple Press-Schechter calculation with the same parameters.

On the other hand, we also investigate the seed effect. The model we consider sees relatively rarefied PBHs evolving in isolation and accreting dark matter to form galaxies. We find that heavy PBHs with  $m_{\text{PBH}} \sim 10^8 M_\odot$  can explain the JWST galaxies given a high SFE  $\epsilon = 1$ , whereas we would need  $m_{\text{PBH}} \sim 10^9 M_\odot$  if  $\epsilon = 0.1$ .

These results were obtained assuming little to no initial PBHs clustering and monochromatic/very peaked mass functions. More realistic models in the future will have to account for more complicated effects, while at the same time study the relaxations of the PBHs constraints that they possibly imply. These more complex studies, combining JWST and NANOGrav results, will better our understanding regarding PBHs as a dark matter candidate.

Unfortunately, at the moment analytical calculations are proving to be very difficult to carry out, while numerical simulations only reach redshifts of order  $z \sim 100$ .

## References

- [1] M. Boylan-Kolchin, “Stress Testing  $\Lambda$ CDM with High-redshift Galaxy Candidates,” 8 2022.
- [2] I. Labb  , P. van Dokkum, E. Nelson, R. Bezanson, K. A. Suess, J. Leja, G. Brammer, K. Whitaker, E. Mathews, M. Stefanon, and B. Wang, “A population of red candidate massive galaxies  $\sim$ 600 myr after the big bang,” *Nature*, vol. 616, pp. 266–269, feb 2023.
- [3] B. Liu and V. Bromm, “Accelerating Early Massive Galaxy Formation with Primordial Black Holes,” *Astrophys. J. Lett.*, vol. 937, no. 2, p. L30, 2022.
- [4] G. Agazie *et al.*, “The NANOGrav 15 yr Data Set: Evidence for a Gravitational-wave Background,” *Astrophys. J. Lett.*, vol. 951, no. 1, p. L8, 2023.
- [5] N. Aghanim *et al.*, “Planck 2018 results. VI. Cosmological parameters,” *Astron. Astrophys.*, vol. 641, p. A6, 2020. [Erratum: *Astron. Astrophys.* 652, C4 (2021)].
- [6] S. Weinberg, “The cosmological constant problem,” *Rev. Mod. Phys.*, vol. 61, pp. 1–23, Jan 1989.
- [7] C. A. *et al.*, “Dark energy survey year 1 results: Cosmological constraints from galaxy clustering and weak lensing,” *Physical Review D*, vol. 98, aug 2018.
- [8] Y. A. *et al.*, “Planck 2018 results. VII. Isotropy and statistics of the CMB,” *Astronomy & Astrophysics*, vol. 641, p. A7, sep 2020.
- [9] C. Dalang and C. Bonvin, “On the kinematic cosmic dipole tension,” *Monthly Notices of the Royal Astronomical Society*, vol. 512, pp. 3895–3905, mar 2022.
- [10] Y. Minami and E. Komatsu, “New extraction of the cosmic birefringence from the planck 2018 polarization data,” *Physical Review Letters*, vol. 125, nov 2020.
- [11] L. Verde, P. Protopapas, and R. Jimenez, “Planck and the local universe: quantifying the tension,” 2013.
- [12] J. S. Bullock and M. Boylan-Kolchin, “Small-Scale Challenges to the  $\Lambda$ CDM Paradigm,” *Annual Review of Astronomy and Astrophysics*, vol. 55, pp. 343–387, aug 2017.
- [13] B. D. Fields, “The primordial lithium problem,” *Annual Review of Nuclear and Particle Science*, vol. 61, pp. 47–68, nov 2011.
- [14] N. J. Adams, C. J. Conselice, L. Ferreira, D. Austin, J. A. A. Trussler, I. Juod   balis, S. M. Wilkins, J. Caruana, P. Dayal, A. Verma, and A. P. Vijayan, “Discovery and properties of ultra-high redshift galaxies (9  $\leq$  z  $\leq$  12) in the iWST/i ERO SMACS 0723 field,” *Monthly Notices of the Royal Astronomical Society*, vol. 518, pp. 4755–4766, nov 2022.
- [15] H. Atek, M. Shuntov, L. J. Furtak, J. Richard, J.-P. Kneib, G. Mahler, A. Zitrin, H. J. McCracken, S. Charlot, J. Chevallard, and I. Chemerynska, “Revealing galaxy candidates out to z 16 with JWST observations of the lensing cluster SMACS0723,” , vol. 519, pp. 1201–1220, Feb. 2023.
- [16] M. Castellano, A. Fontana, T. Treu, P. Santini, E. Merlin, N. Leethochawalit, M. Trenti, E. Vanzella, U. Mestric, A. Bonchi, D. Belfiori, M. Nonino, D. Paris, G. Polenta, G. Roberts-Borsani, K. Boyett, M. Brada  , A. Calabro  , K. Glazebrook, C. Grillo, S. Mascia, C. Mason, A. Mercurio, T. Morishita, T. Nanayakkara, L. Pentericci, P. Rosati, B. Vulcani, X. Wang, and L. Yang, “Early results from glass-jwst. iii. galaxy candidates at z 9–15\*,” *The Astrophysical Journal Letters*, vol. 938, p. L15, oct 2022.
- [17] C. T. Donnan, D. J. McLeod, J. S. Dunlop, R. J. McLure, A. C. Carnall, R. Begley, F. Cullen, M. L. Hamadouche, R. A. A. Bowler, D. Magee, H. J. McCracken, B. Milvang-Jensen, A. Moneti, and T. Targett, “The evolution of the galaxy UV luminosity function at redshifts  $z \simeq 8 – 15$  from deep JWST and ground-based near-infrared imaging,” *Monthly Notices of the Royal Astronomical Society*, vol. 518, pp. 6011–6040, nov 2022.

- [18] S. L. F. et al., "A long time ago in a galaxy far, far away: A candidate  $z \approx 12$  galaxy in early jwst ceers imaging," *The Astrophysical Journal Letters*, vol. 940, p. L55, dec 2022.
- [19] T. Morishita and M. Stiavelli, "Physical Characterization of Early Galaxies in the Webb's First Deep Field SMACS J0723.3-7327," *The Astrophysical Journal Letters*, vol. 946, p. L35, mar 2023.
- [20] R. P. Naidu, P. A. Oesch, P. van Dokkum, E. J. Nelson, K. A. Suess, G. Brammer, K. E. Whitaker, G. Illingworth, R. Bouwens, S. Tacchella, J. Matthee, N. Allen, R. Bezanson, C. Conroy, I. Labbe, J. Leja, E. Leonova, D. Magee, S. H. Price, D. J. Setton, V. Strait, M. Stefanon, S. Toft, J. R. Weaver, and A. Weibel, "Two remarkably luminous galaxy candidates at  $z \approx 10-12$  revealed by jwst," *The Astrophysical Journal Letters*, vol. 940, p. L14, nov 2022.
- [21] H. Yan, Z. Ma, C. Ling, C. Cheng, and J.-S. Huang, "First Batch of  $z \approx 11-20$  Candidate Objects Revealed by the James Webb Space Telescope Early Release Observations on SMACS 0723-73," *The Astrophysical Journal Letters*, vol. 942, p. L9, dec 2022.
- [22] Y. Gong, B. Yue, Y. Cao, and X. Chen, "Fuzzy Dark Matter as a Solution to Reconcile the Stellar Mass Density of High- $z$  Massive Galaxies and Reionization History," *Astrophys. J.*, vol. 947, no. 1, p. 28, 2023.
- [23] A. Klypin, V. Poulin, F. Prada, J. Primack, M. Kamionkowski, V. Avila-Reese, A. Rodriguez-Puebla, P. Behroozi, D. Hellinger, and T. L. Smith, "Clustering and halo abundances in early dark energy cosmological models," *Monthly Notices of the Royal Astronomical Society*, vol. 504, pp. 769–781, mar 2021.
- [24] T. L. Smith, M. Lucca, V. Poulin, G. F. Abellán, L. Balkenhol, K. Benabed, S. Galli, and R. Murgia, "Hints of early dark energy in Planck, SPT, and ACT data: New physics or systematics?," *Physical Review D*, vol. 106, aug 2022.
- [25] G. Hütsi, M. Raidal, J. Urrutia, V. Vaskonen, and H. Veermäe, "Did JWST observe imprints of axion miniclusters or primordial black holes?," *Phys. Rev. D*, vol. 107, no. 4, p. 043502, 2023.
- [26] S.-Y. Guo, M. Khlopov, X. Liu, L. Wu, Y. Wu, and B. Zhu, "Footprints of Axion-Like Particle in Pulsar Timing Array Data and JWST Observations," 6 2023.
- [27] J. M. Ezquiaga, J. García-Bellido, and V. Vennin, "Massive galaxy clusters like "el gordo" hint at primordial quantum diffusion," 2023.
- [28] M. Biagetti, G. Franciolini, and A. Riotto, "High-redshift JWST Observations and Primordial Non-Gaussianity," *Astrophys. J.*, vol. 944, no. 2, p. 113, 2023.
- [29] C. Ilie, J. Paulin, and K. Freese, "Supermassive Dark Star candidates seen by JWST?," 4 2023.
- [30] H. Jiao, R. Brandenberger, and A. Refregier, "Early structure formation from cosmic string loops in light of early JWST observations," *Phys. Rev. D*, vol. 108, no. 4, p. 043510, 2023.
- [31] G.-W. Yuan, L. Lei, Y.-Z. Wang, B. Wang, Y.-Y. Wang, C. Chen, Z.-Q. Shen, Y.-F. Cai, and Y.-Z. Fan, "Rapidly growing primordial black holes as seeds of the massive high-redshift JWST Galaxies," 3 2023.
- [32] B.-Y. Su, N. Li, and L. Feng, "An inflation model for massive primordial black holes to interpret the JWST observations," 6 2023.
- [33] Y. Gouttēnoire, S. Trifinopoulos, G. Valogiannis, and M. Vanvlasselaer, "Scrutinizing the Primordial Black Holes Interpretation of PTA Gravitational Waves and JWST Early Galaxies," 7 2023.
- [34] S. Hawking, "Gravitationally collapsed objects of very low mass," *Mon. Not. Roy. Astron. Soc.*, vol. 152, p. 75, 1971.

- [35] A. et al., “Observation of gravitational waves from a binary black hole merger,” *Physical Review Letters*, vol. 116, feb 2016.
- [36] S. Bird, I. Cholis, J. B. Muñoz, Y. Ali-Haïmoud, M. Kamionkowski, E. D. Kovetz, A. Raccanelli, and A. G. Riess, “Did LIGO detect dark matter?,” *Phys. Rev. Lett.*, vol. 116, no. 20, p. 201301, 2016.
- [37] S. Clesse and J. García-Bellido, “The clustering of massive Primordial Black Holes as Dark Matter: measuring their mass distribution with Advanced LIGO,” *Phys. Dark Univ.*, vol. 15, pp. 142–147, 2017.
- [38] M. Sasaki, T. Suyama, T. Tanaka, and S. Yokoyama, “Primordial Black Hole Scenario for the Gravitational-Wave Event GW150914,” *Phys. Rev. Lett.*, vol. 117, no. 6, p. 061101, 2016. [Erratum: Phys.Rev.Lett. 121, 059901 (2018)].
- [39] B. Carr, S. Clesse, J. Garcia-Bellido, M. Hawkins, and F. Kuhnel, “Observational Evidence for Primordial Black Holes: A Positivist Perspective,” 6 2023.
- [40] R. Maiolino *et al.*, “JADES. The diverse population of infant Black Holes at  $4 < z < 11$ : merging, tiny, poor, but mighty,” 8 2023.
- [41] R. Maiolino *et al.*, “A small and vigorous black hole in the early Universe,” 5 2023.
- [42] J. Matthee *et al.*, “Little Red Dots: an abundant population of faint AGN at  $z \sim 5$  revealed by the EIGER and FRESCO JWST surveys,” 6 2023.
- [43] Y. H. et al., “JWST/NIRSpec First Census of Broad-Line AGNs at  $z = 4 - 7$ : Detection of 10 Faint AGNs with  $M_{\text{BH}} \sim 10^6 - 10^7 M_{\odot}$  and Their Host Galaxy Properties,” 2023.
- [44] H. Übler; Roberto Maiolino ; Emma Curtis-Lake et al., “GA-NIFS: A massive black hole in a low-metallicity AGN at  $z \sim 5.55$  revealed by JWST/NIRSpec IFS,” 2023.
- [45] D. D. K. M. O. K. I. et al., “Hidden little monsters: Spectroscopic identification of low-mass, broad-line agn at  $z > 5$  with ceers,” 2023.
- [46] J. Antoniadis *et al.*, “The second data release from the European Pulsar Timing Array III. Search for gravitational wave signals,” 6 2023.
- [47] D. J. Reardon *et al.*, “Search for an Isotropic Gravitational-wave Background with the Parkes Pulsar Timing Array,” *Astrophys. J. Lett.*, vol. 951, no. 1, p. L6, 2023.
- [48] H. Xu *et al.*, “Searching for the Nano-Hertz Stochastic Gravitational Wave Background with the Chinese Pulsar Timing Array Data Release I,” *Res. Astron. Astrophys.*, vol. 23, no. 7, p. 075024, 2023.
- [49] A. G. Riess, S. Casertano, W. Yuan, L. M. Macri, and D. Scolnic, “Large Magellanic Cloud Cepheid Standards Provide a 1% Foundation for the Determination of the Hubble Constant and Stronger Evidence for Physics beyond  $\Lambda$ CDM ,” *The Astrophysical Journal*, vol. 876, p. 85, may 2019.
- [50] A. Einstein and W. de Sitter, “On the relation between the expansion and the mean density of the universe,” *Proceedings of the National Academy of Sciences*, vol. 18, no. 3, pp. 213–214, 1932.
- [51] D. J. Fixsen, E. S. Cheng, J. M. Gales, J. C. Mather, R. A. Shafer, and E. L. Wright, “The Cosmic Microwave Background Spectrum from the Full COBE/FIRAS Data Set,” *The Astrophysical Journal*, vol. 473, pp. 576–587, dec 1996.
- [52] C. Deffayet, X. Gao, D. A. Steer, and G. Zahariade, “From k-essence to generalized galileons,” *Physical Review D*, vol. 84, sep 2011.
- [53] J. D. Barrow and A. A. H. Graham, “General dynamics of varying-alpha universes,” *Physical Review D*, vol. 88, nov 2013.

- [54] L. Susskind, “The anthropic landscape of string theory,” 2003.
- [55] L. Perivolaropoulos and F. Skara, “Challenges for  $\Lambda$ CDM: An update,” *New Astronomy Reviews*, vol. 95, p. 101659, dec 2022.
- [56] D. Baumann, *Cosmology*. Cambridge University Press, 7 2022.
- [57] J. M. Bardeen, J. R. Bond, N. Kaiser, and A. S. Szalay, “The Statistics of Peaks of Gaussian Random Fields,” *Astrophys. J.*, vol. 304, pp. 15–61, 1986.
- [58] S. M. Carroll, W. H. Press, and E. L. Turner, “The cosmological constant.,,” , vol. 30, pp. 499–542, Jan. 1992.
- [59] H. Mo, F. C. van den Bosch, and S. White, *Galaxy Formation and Evolution*. 2010.
- [60] H. Kurki-Suonio, “Cosmological perturbation theory I,” 6 2022.
- [61] W. H. Press and P. Schechter, “Formation of galaxies and clusters of galaxies by selfsimilar gravitational condensation,” *Astrophys. J.*, vol. 187, pp. 425–438, 1974.
- [62] J. R. Bond, S. Cole, G. Efstathiou, and N. Kaiser, “Excursion set mass functions for hierarchical Gaussian fluctuations,” *Astrophys. J.*, vol. 379, p. 440, 1991.
- [63] R. K. Sheth and G. Tormen, “Large scale bias and the peak background split,” *Mon. Not. Roy. Astron. Soc.*, vol. 308, p. 119, 1999.
- [64] C. G. Lacey and S. Cole, “Merger rates in hierarchical models of galaxy formation,” *Mon. Not. Roy. Astron. Soc.*, vol. 262, pp. 627–649, 1993.
- [65] S. Cole, J. Helly, C. S. Frenk, and H. Parkinson, “The statistical properties of cold dark matter halo formation,” *Monthly Notices of the Royal Astronomical Society*, vol. 383, pp. 546–556, 12 2007.
- [66] T. Padmanabhan, “Aspects of gravitational clustering,” 1999.
- [67] A. R. Zentner, “The excursion set theory of halo mass functions, halo clustering, and halo growth,” *International Journal of Modern Physics D*, vol. 16, pp. 763–815, may 2007.
- [68] D. N. Page, “Particle emission rates from a black hole: Massless particles from an uncharged, nonrotating hole,” *Phys. Rev. D*, vol. 13, pp. 198–206, Jan 1976.
- [69] M. P *et al.*, “The einstein telescope: a third-generation gravitational wave observatory,” *Classical and Quantum Gravity*, vol. 27, p. 194002, sep 2010.
- [70] P. A. Seoane *et al.*, “The Gravitational Universe,” 5 2013.
- [71] M. S. Turner, “Dark energy and the new cosmology,” 8 2001.
- [72] A. Escrivà, F. Kuhnel, and Y. Tada, “Primordial Black Holes,” 11 2022.
- [73] M. W. Choptuik, “Universality and scaling in gravitational collapse of a massless scalar field,” *Phys. Rev. Lett.*, vol. 70, pp. 9–12, Jan 1993.
- [74] N. Kaiser, “On the Spatial correlations of Abell clusters,” *Astrophys. J. Lett.*, vol. 284, pp. L9–L12, 1984.
- [75] V. Desjacques and A. Riotto, “Spatial clustering of primordial black holes,” *Phys. Rev. D*, vol. 98, no. 12, p. 123533, 2018.
- [76] D. Inman and Y. Ali-Haïmoud, “Early structure formation in primordial black hole cosmologies,” *Phys. Rev. D*, vol. 100, no. 8, p. 083528, 2019.

- [77] J. García-Bellido and E. R. Morales, “Primordial black holes from single field models of inflation,” *Physics of the Dark Universe*, vol. 18, pp. 47–54, dec 2017.
- [78] H. Di and Y. Gong, “Primordial black holes and second order gravitational waves from ultra-slow-roll inflation,” *Journal of Cosmology and Astroparticle Physics*, vol. 2018, pp. 007–007, jul 2018.
- [79] J. Yokoyama, “Chaotic new inflation and formation of primordial black holes,” *Phys. Rev. D*, vol. 58, p. 083510, 1998.
- [80] A. Barrau, D. Blais, G. Boudoul, and D. Polarski, “Peculiar relics from primordial black holes in the inflationary paradigm,” *Annalen Phys.*, vol. 13, pp. 115–123, 2004.
- [81] B. Carr, K. Kohri, Y. Sendouda, and J. Yokoyama, “Constraints on primordial black holes,” *Rept. Prog. Phys.*, vol. 84, no. 11, p. 116902, 2021.
- [82] A. Arbey, J. Auffinger, and J. Silk, “Constraining primordial black hole masses with the isotropic gamma ray background,” *Physical Review D*, vol. 101, jan 2020.
- [83] M. Boudaud and M. Cirelli, “Voyager 1  $e^\pm$  further constrain primordial black holes as dark matter,” *Physical Review Letters*, vol. 122, jan 2019.
- [84] W. DeRocco and P. W. Graham, “Constraining primordial black hole abundance with the galactic 511 keV line,” *Physical Review Letters*, vol. 123, dec 2019.
- [85] R. Laha, “Primordial Black Holes as a Dark Matter Candidate Are Severely Constrained by the Galactic Center 511 keV  $\gamma$ -Ray Line,” *Phys. Rev. Lett.*, vol. 123, no. 25, p. 251101, 2019.
- [86] R. Laha, J. B. Muñoz, and T. R. Slatyer, “INTEGRAL constraints on primordial black holes and particle dark matter,” *Phys. Rev. D*, vol. 101, no. 12, p. 123514, 2020.
- [87] D. Croon, D. McKeen, N. Raj, and Z. Wang, “Subaru-HSC through a different lens: Microlensing by extended dark matter structures,” *Phys. Rev. D*, vol. 102, no. 8, p. 083021, 2020.
- [88] P. Tisserand *et al.*, “Limits on the Macho Content of the Galactic Halo from the EROS-2 Survey of the Magellanic Clouds,” *Astron. Astrophys.*, vol. 469, pp. 387–404, 2007.
- [89] C. Hamadache *et al.*, “Galactic bulge microlensing optical depth from eros-2,” *Astron. Astrophys.*, vol. 454, pp. 185–199, 2006.
- [90] H. Niikura, M. Takada, S. Yokoyama, T. Sumi, and S. Masaki, “Constraints on Earth-mass primordial black holes from OGLE 5-year microlensing events,” *Phys. Rev. D*, vol. 99, no. 8, p. 083503, 2019.
- [91] M. A. Monroy-Rodríguez and C. Allen, “The end of the MACHO era- revisited: new limits on MACHO masses from halo wide binaries,” *Astrophys. J.*, vol. 790, no. 2, p. 159, 2014.
- [92] B. J. Carr and M. Sakellariadou, “Dynamical constraints on dark compact objects,” *Astrophys. J.*, vol. 516, pp. 195–220, 1999.
- [93] T. D. Brandt, “Constraints on MACHO Dark Matter from Compact Stellar Systems in Ultra-Faint Dwarf Galaxies,” *Astrophys. J. Lett.*, vol. 824, no. 2, p. L31, 2016.
- [94] C. G. Lacey and J. P. Ostriker, “Massive black holes in galactic halos ?,” , vol. 299, pp. 633–652, Dec. 1985.
- [95] B. Carr and J. Silk, “Primordial Black Holes as Generators of Cosmic Structures,” *Mon. Not. Roy. Astron. Soc.*, vol. 478, no. 3, pp. 3756–3775, 2018.
- [96] R. Murgia, G. Scelfo, M. Viel, and A. Raccanelli, “Lyman- $\alpha$  Forest Constraints on Primordial Black Holes as Dark Matter,” *Phys. Rev. Lett.*, vol. 123, no. 7, p. 071102, 2019.

- [97] J. Manshanden, D. Gaggero, G. Bertone, R. M. T. Connors, and M. Ricotti, “Multi-wavelength astronomical searches for primordial black holes,” *JCAP*, vol. 06, p. 026, 2019.
- [98] P. D. Serpico, V. Poulin, D. Inman, and K. Kohri, “Cosmic microwave background bounds on primordial black holes including dark matter halo accretion,” *Phys. Rev. Res.*, vol. 2, no. 2, p. 023204, 2020.
- [99] M. Raidal, V. Vaskonen, and H. Veermäe, “Gravitational Waves from Primordial Black Hole Mergers,” *JCAP*, vol. 09, p. 037, 2017.
- [100] M. Trashorras, J. García-Bellido, and S. Nesseris, “The clustering dynamics of primordial black holes in n-body simulations,” *Universe*, vol. 7, p. 18, jan 2021.
- [101] A. Escrivà, F. Kuhnel, and Y. Tada, “Primordial black holes,” 2023.
- [102] M. Sasaki, T. Suyama, T. Tanaka, and S. Yokoyama, “Primordial black holes - perspectives in gravitational wave astronomy,” *Classical and Quantum Gravity*, vol. 35, p. 063001, feb 2018.
- [103] T. Baldauf, U. Seljak, R. E. Smith, N. Hamaus, and V. Desjacques, “Halo stochasticity from exclusion and nonlinear clustering,” *Physical Review D*, vol. 88, oct 2013.
- [104] V. D. Luca, V. Desjacques, G. Franciolini, and A. Riotto, “The clustering evolution of primordial black holes,” *Journal of Cosmology and Astroparticle Physics*, vol. 2020, pp. 028–028, nov 2020.
- [105] V. Desjacques and A. Riotto, “Spatial clustering of primordial black holes,” *Physical Review D*, vol. 98, dec 2018.
- [106] B. Carr and J. Silk, “Primordial black holes as generators of cosmic structures,” *Monthly Notices of the Royal Astronomical Society*, vol. 478, pp. 3756–3775, may 2018.
- [107] Gardner, Jonathan P.; Mather, John C.; Abbott, Randy et al., “The James Webb Space Telescope Mission,” *Publications of the Astronomical Society of the Pacific*, vol. 135, no. 1048, 2023.
- [108] E. C.-L. et al., “Spectroscopic confirmation of four metal-poor galaxies at  $z=10.3\text{--}13.2$ ,” 2023.
- [109] B. E. R. et al., “Identification and properties of intense star-forming galaxies at redshifts  $z>10$ ,” *Nature Astronomy*, vol. 7, pp. 611–621, apr 2023.
- [110] J. McCaffrey, S. Hardin, J. Wise, and J. Regan, “No tension: Jwst galaxies at  $z > 10$  consistent with cosmological simulations,” 2023.
- [111] B. W. Keller, F. Munshi, M. Trebitsch, and M. Tremmel, “Can cosmological simulations reproduce the spectroscopically confirmed galaxies seen at  $z \geq 10$ ?,” *The Astrophysical Journal Letters*, vol. 943, p. L28, feb 2023.
- [112] S. Tacchella, S. Bose, C. Conroy, D. J. Eisenstein, and B. D. Johnson, “A redshift-independent efficiency model: Star formation and stellar masses in dark matter halos at  $z < 4$ ,” *The Astrophysical Journal*, vol. 868, p. 92, nov 2018.
- [113] G. Hütsi, M. Raidal, and H. Veermäe, “Small-scale structure of primordial black hole dark matter and its implications for accretion,” *Physical Review D*, vol. 100, oct 2019.
- [114] K. J. Mack, J. P. Ostriker, and M. Ricotti, “Growth of structure seeded by primordial black holes,” *The Astrophysical Journal*, vol. 665, pp. 1277–1287, aug 2007.
- [115] F. A. E. Pirani, “On the physical significance of the Riemann tensor,” *Acta Physica Polonica*, vol. 15, pp. 389–405, Jan. 1956.
- [116] F. Estabrook and H. Wahlquist, “Response of Doppler spacecraft tracking to gravitational radiation.,” *General Relativity and Gravitation*, vol. 6, pp. 439–447, Oct. 1975.

- [117] M. Sazhin, “Opportunities for detecting ultralong gravitational waves,” , vol. 22, pp. 36–38, Feb. 1978.
- [118] S. Detweiler, “Pulsar timing measurements and the search for gravitational waves,” , vol. 234, pp. 1100–1104, Dec. 1979.
- [119] R. Hellings and G. Downs, “Upper limits on the isotropic gravitational radiation background from pulsar timing analysis.,” , vol. 265, pp. L39–L42, Feb. 1983.
- [120] Z. A. et al., “The NANOGrav 12.5 yr data set: Search for an isotropic stochastic gravitational-wave background,” *The Astrophysical Journal Letters*, vol. 905, p. L34, dec 2020.
- [121] S. C. et al., “Common-red-signal analysis with 24-yr high-precision timing of the european pulsar timing array: inferences in the stochastic gravitational-wave background search,” *Monthly Notices of the Royal Astronomical Society*, vol. 508, pp. 4970–4993, oct 2021.
- [122] J. A. et al., “The international pulsar timing array second data release: Search for an isotropic gravitational wave background,” *Monthly Notices of the Royal Astronomical Society*, vol. 510, pp. 4873–4887, jan 2022.
- [123] B. G. et al., “On the evidence for a common-spectrum process in the search for the nanohertz gravitational-wave background with the parkes pulsar timing array,” *The Astrophysical Journal Letters*, vol. 917, p. L19, aug 2021.
- [124] J. Antoniadis *et al.*, “The second data release from the European Pulsar Timing Array: V. Implications for massive black holes, dark matter and the early Universe,” 6 2023.
- [125] J. e. a. Ellis, “What is the source of the PTA GW signal?,” 8 2023.
- [126] P. F. Depta, K. Schmidt-Hoberg, P. Schwaller, and C. Tasillo, “Do pulsar timing arrays observe merging primordial black holes?,” 6 2023.
- [127] G. Franciolini, A. Iovino, Junior., V. Vaskonen, and H. Veermae, “The recent gravitational wave observation by pulsar timing arrays and primordial black holes: the importance of non-gaussianities,” 6 2023.
- [128] K. Inomata, K. Kohri, and T. Terada, “The Detected Stochastic Gravitational Waves and Subsolar-Mass Primordial Black Holes,” 6 2023.
- [129] N. Bhaumik, R. K. Jain, and M. Lewicki, “Ultra-low mass PBHs in the early universe can explain the PTA signal,” 8 2023.