UNIVERSITY
OF OSLO

Maksym Brilenkov

From theory to practice: Reproducibility and Open Science in the CMB field

Thesis submitted for the degree of Philosophiae Doctor

Institute of Theoretical Astrophysics
Faculty of Mathematics and Natural Sciences

2023

Series of dissertations submitted to the
Faculty of Mathematics and Natural Sciences, University of Oslo
No. 2606

ISSN 1501-7710

Cover: UiO.
Print production: Graphics Center, University of Oslo.

To my family

Abstract

The study of the Cosmic Microwave Background (CMB) radiation is a pillar of modern physical cosmology that enables us to understand the history, the evolution of structures, and the global geometrical properties of the Universe. Furthermore, being one of the oldest signals in the Universe, the CMB brings us unique information about the period of inflation — when the Universe was in its infancy. It is theorized that inflation produced primordial gravitational waves, and these may have left an imprint on the polarization of CMB light in the form of B-modes. Thus, detecting these patterns is a vital stress test for the current theory and our understanding of the Universe in general.

CMB research is a broad theme that requires a complex approach. This thesis touches upon theoretical, numerical, and observational aspects of cosmology that all revolve around CMB radiation. The first part of the manuscript describes the conventional ΛCDM model and my work on its extension by the inclusion of relativistic particles with the subsequent derivation of analytical perturbation theory for the case of a flat Universe in the weak gravitational field regime. I also present my work on spatial topology, where I solved the Helmholtz equation and derived the expressions for the gravitational potential in trigonometric representation for the cases of $T \times T \times R$ and $T \times R \times R$ topologies.

The latter parts are concerned with CMB experiments and data analysis. In particular, I describe a few major missions that shaped the field into what we know today, provide an overview of the joint analysis approach (based on the Gibbs sampling technique) that formed the foundations of the `Commander3` pipeline, and present some of the main results of the BeyondPlanck collaboration that I have been a member of during these years. I then expand on the ideas of BeyondPlanck and describe the Cosmoglobe project as its main successor — a community-driven effort with a strong emphasis on Open Science and reproducibility. These efforts include easy distribution of raw data and user-friendly code compilation procedures, along with documentation, data products, etc. Lastly, I discuss simulation-based validation of the `Commander3` pipeline.

Studiet av den kosmiske bakgrunnsstrålingen (Cosmic Microwave Background; CMB) er en hjørnestein i studiet av det tidlige univers, og den gjør oss i stand til å forstå historien, utviklingen av strukturer og de globale geometriske egenskapene til universet. Siden den kosmiske bakgrunnsstrålingen er blant de eldste signalene i universet, så tror man idag at dette signalet inneholder informasjon om selve skapelsesøyeblikket, i form av primordiale gravitasjonsbølgene fra kosmisk inflasjon. Å oppdage disse mønstrene er en avgjørende stresstest for den nåværende teorien og vår forståelse av universet generelt.

Studiet av bakgrunnsstrålingen er et bredt forskningstema som krever en kompleks tilnærming til oppgaven. Denne oppgaven berører teoretiske, numeriske og observasjonsgrener av kosmologi, som alle dreier seg om CMB-stråling. Den første delen av manuskriptet omhandler den konvensjonelle ΛCDM-modellen og mitt arbeid med dens utvidelse ved å inkludere relativistiske partikler med påfølgende utledning av analytisk forstyrrelsesteori for tilfellet med et flatt univers i det svake gravitasjonsfeltregimet. Dernest presenterer jeg mitt bidrag til arbeidet med romlig topologi, hvor jeg løste Helmholtz-ligningen og utledet uttrykkene for gravitasjonspotensialet i trigonometrisk representasjon for tilfellene av $T \times T \times R$ og $T \times R \times R$ topologier.

De siste delene omhandler noen viktige CMB-eksperimenter og dataanalysen av disse. Spesielt beskriver jeg store bidrag som formet feltet til det vi kjenner i dag; jeg gir en oversikt over den felles analysetilnærmingen (basert på Gibbs prøvetakingsteknikk) som dannet grunnlaget for `Commander3`-pipelinen; og jeg presenterer noen av hovedresultatene av BeyondPlanck-samarbeidet som jeg var medlem av i disse årene. Deretter utvider jeg ideene til BeyondPlanck og beskriver Cosmoglobe-prosjektet som dets viktigste etterfølger — en fellesskapsdrevet innsats med sterk vekt på åpen vitenskap og reproduserbarhet. Disse inkluderer enkel distribusjon av rådata og brukervennlige kodekompileringsprosedyrer, sammen med dokumentasjon, dataprodukter osv. Til slutt diskuterer jeg simuleringsbasert validering av `Commander3`-pipelinen.

Acknowledgements

This thesis culminates more than a decade-long scientific journey I undertook by enrolling in the undergraduate physics program at Odesa National University, Ukraine, in 2010. Since then, I have studied in a variety of European universities, been part of several international research projects, and contributed to a number of scientific papers. But, most importantly, I met many great people along the way who helped me to become who I am today, and without whom, undoubtedly, this manuscript would not see the light of day.

First and foremost, I am very grateful to my Norwegian supervisors, **Hans Kristian Eriksen** and **Ingunn Kathrine Wehus**, for granting me this opportunity to become a fully-fledged researcher as well as for their patience, knowledge, and all the help they provided (especially during these turbulent times). Likewise, I would like to thank my Ukrainian supervisors, **Maxim Eingorn** and **Alexander Zhuk**, for introducing me to the world of cosmology and accompanying me on this long journey from the beginning while providing guidance along the way.

I would also like to thank my institute colleagues **Eirik Gjerløw**, **Ranajoy Banerji**, **Kristian Joten Andersen**, **Mikołaj Szydlarski**, and many others for all the discussions, comments, feedback, and chattings about `Commander`, `CMake`, and other challenges both during and after working hours. Special thanks go to **Duncan Watts**, who was willing to sacrifice some of his precious time to clarify and explain things I would otherwise have struggled to understand on my own.

Last but not least, my family has my utmost respect and gratitude for enabling and supporting me throughout this journey. Especially my parents, to whom I am indebted not only for my education but also for my very existence. And, of course, my wife and partner, **Liudmyla**, who fills my life with happiness and makes me a better version of myself. Without you, it would be close to impossible to make it through my master, not to mention the PhD program.

Maksym Brilenkov
Oslo, March 2023

List of Papers

Paper I

M. Brilenkov, E. Canay, and M. Eingorn. “Cosmological Perturbations Engendered by Discrete Relativistic Species”. Submitted to the *Physics of the Dark Universe*, arXiv:2206.13495v1.

Paper II

M. Eingorn, A. McLaughlin II, E. Canay, M. Brilenkov, and A. Zhuk “Gravitational Interaction in the Chimney Lattice Universe”. *MDPI Universe* 2021, 7(4), 101; DOI: 10.3390/universe7040101

Paper III

M. Eingorn, N. O’Briant, K. Arzu, M. Brilenkov, and A. Zhuk “Gravitational potentials and forces in the Lattice Universe: a slab”. *The European Physical Journal Plus* 136, 205 (2021); DOI: 10.1140/epjp/s13360-021-01139-y

Paper IV

The BeyondPlanck Collaboration. “BeyondPlanck I. Global Bayesian analysis of the *Planck* Low Frequency Instrument data”. 2023, *A&A*, in press, arXiv:2011.05609v2.

Paper V

S. Gerakakis, M. Brilenkov, M. Ieronymaki, M. San, D. J. Watts, and the BeyondPlanck Collaboration. “From BeyondPlanck to Cosmoglobe: Open Science, Reproducibility, and Data Longevity.” 2023, *The Open Journal of Astrophysics*, Vol. 6 arXiv:2205.11262v2.

Paper VI

M. Brilenkov, K. S. F. Fornazier, L. T. Hergt, G. A. Hoerning, A. Marins, T. Murokoshi, F. Rahman, N.-O. Stutzer, Y. Zhou, F. .B. Abdalla, and the BeyondPlanck Collaboration. “BeyondPlanck IV. On end-to-end simulations in CMB analysis — Bayesian versus frequentist statistics”. 2023, *A&A*, submitted arXiv:2209.04437v1.

Additional papers published during the PhD studies, but not included in the thesis

1. *Constraints on the spectral index of polarized synchrotron emission from WMAP and Faraday-corrected S-PASS data*, Fuskeland, U., et al. 2021, A&A, 646, A69
2. *A Monte Carlo comparison between template-based and Wiener-filter CMB dipole estimators*, Thommesen, H., et al. 2020, A&A, 643, A179
3. *BeyondPlanck II. CMB map-making through Gibbs sampling*, Keihänen, E., et al., A&A, in press, [arxiv:2011.06024]
4. *BeyondPlanck III. Commander3*, Galloway, M., et al., A&A, in press, [arxiv:2201.03509]
5. *BeyondPlanck VI. Noise characterization and modelling*, Ihle, H. T., et al., A&A, in press, [arxiv:2011.06650]
6. *BeyondPlanck V. Minimal ADC Corrections for Planck LFI*, Herman, D., et al., A&A, in press, [arxiv:2203.13635]
7. *BeyondPlanck VII. Bayesian estimation of gain and absolute calibration for CMB experiments*, Gjerløw, E., et al., A&A, in press, [arxiv:2011.08082]
8. *BeyondPlanck VIII. Efficient Sidelobe Convolution and Correction through Spin Harmonics*, Galloway, M., et al., A&A, in press, [arxiv:2201.03478]
9. *BeyondPlanck X. Bandpass and beam leakage corrections*, Svalheim, T. L., et al., A&A, in press, [arxiv:2201.03417]
10. *BeyondPlanck X. Planck LFI frequency maps with sample-based error propagation*, Basyrov, A., et al., A&A, in press, [arxiv:2208.14293]
11. *BeyondPlanck XI. Bayesian CMB analysis with sample-based end-to-end error propagation*, Colombo, L., et al., A&A, in press, [arxiv:2208.14276]
12. *BeyondPlanck XII. Cosmological parameter constraints with end-to-end error propagation*, Paradiso, S., et al., A&A, in press, [arxiv:2205.10104]
13. *BeyondPlanck XIII. Intensity foreground sampling, degeneracies, and priors*, Andersen, K. J., et al., A&A, in press, [arxiv:2201.08188]
14. *BeyondPlanck XVI. Limits on Large-Scale Polarized Anomalous Microwave Emission from Planck LFI and WMAP*, Herman, D., et al., A&A, in press, [arxiv:2201.03530]
15. *From BeyondPlanck to Cosmoglobe: Preliminary WMAP Q-band analysis*, Watts, D., et al., A&A, in press, [arxiv:2202.11979]

16. *Probing Cosmic Inflation with the LiteBIRD Cosmic Microwave Background Polarization Survey*, LiteBIRD collaboration, et al., PTEP, doi:10.48550/arXiv.2202.02773

17. *Tensor-to-scalar ratio forecasts for extended LiteBIRD frequency configurations*, Fuskeland, U., et al., A&A, submitted, [arxiv:2302.05228]

18. *Cosmoglobe DR1 results. I. Improved Wilkinson Microwave Anisotropy Probe maps through Bayesian end-to-end analysis*, Watts, D., et al., A&A, submitted, [arxiv:2303.08095]

Contents

Introduction

Our place in the Universe is not unique. This simple idea is a pillar of modern cosmology and is often interpreted as *the Universe does not have a preferred position or direction in space.* This statement — known as the **cosmological principle** — while being both simple and elegant, unfortunately, is not exact. Rather, everything depends on the scale at which one considers the Universe we live in. Specifically, the Universe is homogeneous and isotropic on very large scales [137].

How do we know that? Because light carries information. When looking at an object, we do not see the object itself but the light either emitted or reflected by it and captured by our eyes — the detectors. Our eyes are quite limited detectors since they can distinguish only *visible light* — one of the many parts of the *electromagnetic spectrum*[1]. There are so many things we can observe using visible light alone. For instance, *stars* [128] — the spherical gaseous objects bound by self-gravity that radiate light particles, *photons*, due to the nuclear fusion in their cores — are the primary source of visible light in the Universe. Together with planets, comets, asteroids, interstellar gas, and dust, millions of stars can form *galaxies* [88] — gravitationally bound conglomerates of mass, which typically rotate around a supermassive black hole in its center.

Of course, there are additional crucial components to a galaxy [17, 29], and one particularly important is *Dark Matter* (DM) — an unknown substance that does not interact with light directly but only through gravity. Together with *Dark Energy*[2] (DE), DM constitutes one of the main puzzles in modern cosmology, and an incredible research effort is dedicated to understanding its nature by the scientific community. Together, DE and DM represent 95 % of the matter and energy content of the Universe, and the mathematical framework that describes this is known as the Λ *Cold Dark Matter* (ΛCDM) model. Key to this model is the idea that the Universe started its evolution from a highly dense state and then rapidly expanded in size — the Hot Big Bang — with its subsequent cooling. This model also relies heavily on the theory of *General Relativity* (GR) to describe gravity and is extended by the theory of inflation [72, 138, 144, 145] to solve some important cosmological problems such as the flatness and the horizon problems [3, 63, 86].

Overall, the ΛCDM model is an incredible success and is today regarded as the standard concordance model of modern physical cosmology. Among other things, it gives a good explanation for the existence of the *Cosmic Microwave*

---

[1] Others include *radio waves*, *microwaves*, *infrared light*, *ultraviolet light*, *X-rays* and γ*-rays*, each corresponding to their specific frequency.

[2] The term Dark Energy was first mentioned by Michael Turner in 1999 paper [69] in full analogy with Fritz Zwicky's [164] Dark Matter.

Background (CMB) radiation — the primordial light formed roughly 370 000 years after the Big Bang [134, 148]. The CMB is surprisingly homogeneous, with slight variations at the level of only tens or a few hundreds of μK across the entire sky compared to its background value of ~ 2.7K. Being released during the early ages of the Universe after having been in thermal equilibrium with matter during that period, it carries important information about how the Universe came to be, and is today a valuable tool to test cosmological models. That is why a massive amount of theoretical and observational cosmology research is concentrated around this light.

Despite its success, ΛCDM suffers from some serious issues among which the most prominent ones are the unknown nature(s) of DE and DM, the *cosmological constant problem* [21, 93], and perhaps even the existence of CMB anomalies [112, 118, 125]. It is therefore of interest to consider various extensions and generalizations of this model, and constrain these modifications by observations. Only a few examples are modified gravity models [24, 74] or dynamical DE models [50, 84, 92, 105, 155]. An aspect that is considered in this thesis is the extension of ΛCDM by the inclusion of relativistic particles such as neutrinos or *Warm Dark Matter* (WDM) (see [37, 82, 142, 156], and references therein), since such species may have an important observable effect on the CMB. Furthermore, the analysis of the corresponding model potentially allows us to estimate, e.g., neutrino masses, predict physical properties of WDM, or, perhaps, even advance in solving the missing satellites problem (see [95] and references therein). In this regard, it makes sense to generalize the analytical perturbation theory formulated for a case of the standard ΛCDM Universe. This is done in Paper I.

Research during the last two decades suggests that various CMB anomalies on large angular scales, such as the suppression of the quadrupole moment or the quadrupole and octopole alignment could possibly be imprints of non-trivial spatial topology [15, 16], i.e., the global shape of the Universe — a fundamental topic related to the questions of boundedness[3], curvature[4], and connectivity[5] of the Universe. GR does not dictate topology; thus, potentially, a wide range of topologies are possible, except for some "exotic" non-simply connected spacetimes with singularities [56]. Papers II and III expand on these ideas further and consider two distinct topologies — chimney and slab.

Historically, the existence of the CMB was predicted by Alpher and Gamow in 1948 [4], and it was later discovered in 1964 by astronomers Penzias and Wilson [106]. It quickly became a main topic of study for a plethora of satellite, balloon-bourne, and ground-based experiments, and perhaps the most widely known are *COBE*, which constrained the CMB temperature value as $T_0 = (2.725 \pm 0.002)$K [96]; *WMAP* [13] that measured the CMB angular power spectrum down to scales of about a quarter of a degree [12]; and, of course, *Planck* [152, 157] — the European Space Agency (ESA) led satellite mission that mapped primary temperature anisotropies of the CMB to the cosmic variance limit on all angular

---

[3] Is the Universe finite or infinite?

[4] Is the curvature of the Universe flat, hyperbolic or spherical?

[5] Is the Universe simply (in agreement with concordance cosmology) or multiply connected?

scales larger than about 7' and produced the highest signal-to-noise ratio maps of the entire sky between 30 GHz and 857 GHz published to date [109, 114, 121].

The next milestone is the study of CMB polarization, which, at least in theory, carries an imprint of the period of inflation. The key idea here is that *gravitational waves* created during the period of exponential expansion emerged as ripples in spacetime, and these ripples should have left a specific, curly imprint on large angular scales in the CMB polarization pattern, called *B modes*. While no other mechanism is known to produce such an effect, the signal is supposed to be very weak, making it difficult to detect in practice. Therefore, a new generation of experiments has been proposed explicitly for this task. One of them is *LiteBIRD* [27] — the Japan Aerospace Exploration Agency (JAXA) led mission expected to be launched in late 2020's.

By capturing the electromagnetic radiation and transforming it into electrical signals, CMB experiments generate *terabytes* (TB) of computer data which are analyzed using various statistical techniques and numerical algorithms on *high-performance computing* (HPC) clusters. The entire data analysis process is quite involved, and an overview of a typical procedure is described in the second part of Chapter 2. This procedure can roughly be split into six distinct but connected stages, namely: (1) data acquisition; (2) data cleaning; (3) map making; (4) component separation; (5) power spectrum estimation; and (6) cosmological parameter estimation. A combination of several or all of these steps in one single entity is called a *pipeline*.

Usually, each step mentioned above is handled by a specific group of dedicated professionals that specialize in their respective chunk of work. However, the recent BeyondPlanck project challenged that paradigm by developing `Commander3` — the `Fortran`-based code which can jointly analyze CMB data via a Bayesian Gibbs sampling technique. The beauty and power of `Commander3` lie in the fact that it does everything starting from raw *time-ordered data* (TOD) and finishing with cosmological parameter estimation. Even more so, since we in BeyondPlanck recognized the importance of Open Science and reproducibility, I have spent a substantial amount of time developing a `CMake`-based infrastructure that enables users to install `Commander3` and all its dependencies automatically, with a handful of simple commands. In addition, I have implemented a stand-alone simulation module that effectively transforms `Commander3` from a data analysis code into a CMB data simulator, allowing much easier debugging and validation. These endeavors are discussed in Papers V and VI, while an overview of `Commander3` products is given in Paper IV. The machinery developed in these papers forms an important computational foundation for the entire BeyondPlanck pipeline, and will hopefully continue to serve the CMB community in the future.

Overall, CMB cosmology is an ever-expanding and fascinating but incredibly complex research field that allows both creation and testing of various cosmological theories. My work lies on the intersection of theory, observation, and numerical practices, and this thesis constitutes a modest attempt to bring these components together. The manuscript is submitted as partial fulfillment of the requirements for the degree of *Philosophiae Doctor* at the Institute of Theoretical

Astrophysics, University of Oslo. It was performed as part of the CMB&CO research group under the supervision of professor Hans Kristian Eriksen, and professor Ingunn Kathrine Wehus. It was supported by grant agreements №776282 (COMPET-4;BeyondPLanck) and №772253 (ERC;bits2cosmology).

The thesis comprises this introduction, three distinct chapters, and a list of six papers in total. Chapter 1 gives a brief overview of the standard cosmological model. Chapter 2 provides a brief overview of major CMB experiments and then dives into the CMB data analysis concentrating mainly on `Commander3` theory. Lastly, Chapter 3 describes our approaches to reproducibility and Open Science in BeyondPlanck and Cosmoglobe projects. In addition to the six papers included in the thesis, I have also contributed to 18 other papers during my PhD studies, both individually and as part of the two international *LiteBIRD* and Cosmoglobe collaborations.

Chapter 1

Cosmology

According to the concordance model of cosmology (for an in-depth overview, see, for instance, [35, 85, 97, 158]), all matter in the visible Universe was initially condensed into a tiny volume smaller than the size of a hydrogen atom[1]. When the Universe became 10^{-36} s old [137], it started to expand at an explosive speed. Nothing is certain of what was before, and it is believed that the physical laws we know today were not working in the same way at the time. What we do know, however, is the fact that the Universe continued to enlarge ever since, although the speed of expansion was and is not homogeneous in time, with some periods marked as exponential, usual, and accelerated expansions. The purpose of this chapter is to give a brief overview of the Universe's timeline. In addition, it provides a small tour of the ΛCDM model, touches upon the topic of the topology of the Universe, and describes the CMB.

1.1 Brief History of the Universe

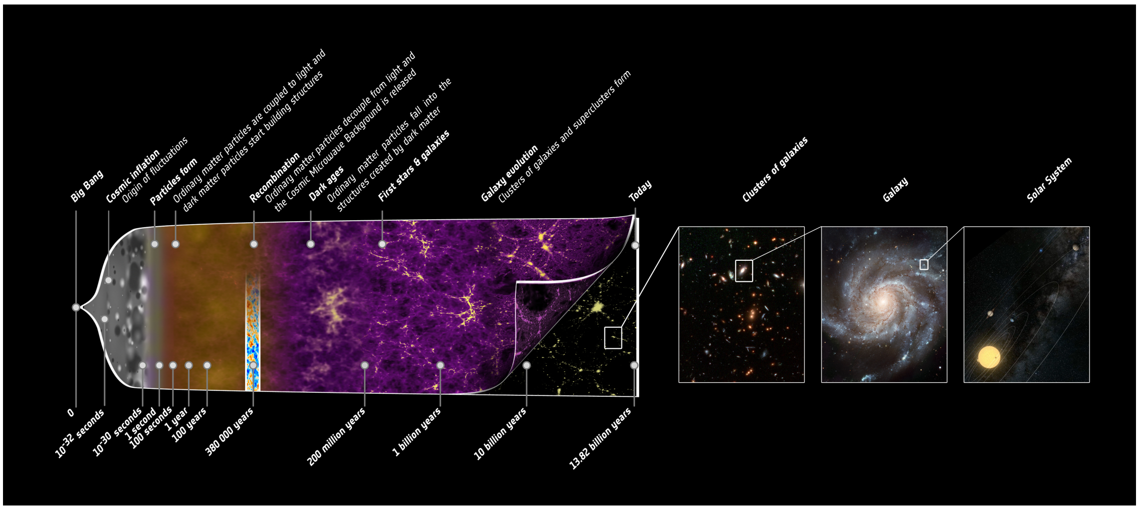

Figure 1.1: Schematic depiction of the Universe's timeline. Image Credit: NASA.

The Universe's evolutionary timeline can be roughly divided into several distinct stages (see figure 1.1), namely the very early Universe, early Universe, dark ages and the formation of large-scale structure, present Universe, and future Universe.

The Very Early Universe. The matter density and temperature was so extreme in the newborn Universe that it is not possible to recreate similar conditions in a

---

[1] The first person who came up with the term "primeval atom" and "cosmic egg" was Lemaître in his 1931 papers [81, 89].

laboratory. We can therefore only make theoretical models of what kind of place our Universe was at that time, and these may or may not be true. However, it is believed that during this era, not only was it impossible for atomic and subatomic particles to form, but even the four fundamental forces, namely gravity, electromagnetism, and weak and strong nuclear forces, were combined into one single entity (see, e.g., [134] and references therein). At this time, the Universe is theorized to have undergone a rapid exponential expansion called *inflation* [72, 138, 144, 145] that, although lasted only for a tiny fraction of a second, allowed the space to stretch by a factor of $O\left(10^{28}\right)$ [35]. Driven by the vacuum energy of one (or several) scalar field(s), it was accompanied by quantum fluctuations whose size increased together with space. These irregularities became the future seeds from which all the structures (such as galaxies, clusters of galaxies, etc.) grew later on. Inflation finished with a period called *reheating*[2] that signified the beginning of the *radiation dominated era* [35, 158]. From here on, we talk about the "traditional" Big Bang.

The Early Universe. Reheating left the Universe in an incredibly hot and dense state filled with a *quark-gluon plasma* (see, e.g., [104, 162] and references therein). During this period, the unified force broke into the familiar forces of nature, signifying the beginning of the *quark epoch* during which the fundamental particles called *quarks* and *gluons* travelled freely through space. At $t \simeq 10^{-5} - 10^{-4}$ s after the Big Bang [162], *quantum chromodynamics* (QCD) phase transitions started to occur, and quarks and gluons formed "colorless particles", such as neutrons and protons, collectively known as *hadrons*. The conversion was completed when the overall plasma temperature decreased below that corresponding to the mass of the *pion* ($m_\pi \simeq 140$ MeV) and *muon* ($m_\mu \simeq 106$ MeV) particles. Most of these hadrons were short-lived, and they quickly annihilated with their respective antihadron pairs [51]. Around one second after the Big Bang [88], the *hadron epoch* was over, and the *Cosmic Neutrino Background* (CνB) was formed since *neutrinos*[3] decoupled from the rest of the fluid and began to freely traverse the space.

At the same time as the CνB was released, the formation of various isotopes of hydrogen (^{2}H), helium (^{3}He, ^{4}He), and lithium (^{7}Li) started to occur. This process — known as *Big Bang Nucleosynthesis* (BBN) [100] — ended about 20 minutes after the Big Bang [25]. The majority of the mass-energy at this time was in the form of relativistic particles. Photons frequently interacted with electrons, protons, and light nuclei, all being part of a single entity, a hot ionized plasma. The Universe was opaque at this time, and the mean free path of the photons was quite short. However, as the Universe became 18 000 years old

---

[2] The process of energy transfer between the inflationary field(s) and the conventional matter [134].

[3] These neutrinos are believed to have survived until the present day. As their energies are very low, it is challenging, if not impossible, to observe them directly. However, there is a number of indirect evidences from the CMB and BBN (see, e.g., [5] and references therein) that CνB had taken place.

[148], the opportunity arose for free electrons to combine with He nuclei into neutral He atoms. The lengthy process of *helium recombination*[4] (HeIII → HeII) went on for tens of thousands of years while the Universe was expanding and cooling further. Relativistic particles and baryons were still effectively a single fluid, but photons were losing energy as their wavelengths were increasing with the expansion of space. Eventually, 47 000 years [88, 137] after the Big Bang, the energy density of baryons started to dominate that of photons, marking the beginning of the *matter dominated era* [35, 158] — an extremely long period of time during which the energy density of matter dictated the evolution of structures in the Universe.

An important milestone in structure formation occurred when the Universe was 130 000 years old, caused by another phase of atomic formation — the *second helium recombination* (HeII → HeI). It was succeeded by the *hydrogen recombination* with the peak of atomic production at roughly 260 000 years (see [148], figure 5). Once bound by the respective nuclei, highly energetic electrons were releasing one or several photons during the transition to the low energy states. This was a *photon decoupling*, and it was continuously going on until about 370 000 years after the Big Bang [134, 148]. By that time, most of the protons were bound into atoms, and although some of the decoupled photons were absorbed by these newly formed atoms, the remaining ones were untouched. Since the Universe became electroneutral, these photons could now freely propagate through space, thus making the Universe transparent to electromagnetic radiation for the first time in history. Initially, decoupled photons had a temperature of 0.26 eV or 3000 K [134], but during billions of years of cosmic expansion, their wavelengths stretched. Eventually, instead of visible light, they became redshifted into the microwave range, which earned them the name *Cosmic Microwave Background radiation*[5].

Dark Ages and Large-Scale Structure. The photons' wavelengths stretched together with space, and the CMB lost energy while propagating freely through space after recombination. At some point, the CMB redshifted from visible light,

---

[4] In itself, the term "recombination" is inaccurate since atoms started to form for the very first time. In addition, it is common in the literature to address the period of *hydrogen recombination* (with $z \sim 1000$ [148]) as *the recombination* since it defines the *surface of last scattering* (with $z \simeq 1063$ and a width of $\Delta z \approx 80$ [158]), i.e., the shell at which most of the CMB photons started to free-stream towards us. Here, z is the *redshift* that quantifies the stretching of the wavelength of observed light due to the systematic expansion of the Universe. Mathematically, it is defined as

$$z \equiv \frac{\lambda_{\text{obs}} - \lambda_{\text{em}}}{\lambda_{\text{em}}} = \frac{1}{a} - 1 \simeq \frac{v}{c}\,, \tag{1.1}$$

where λ_{em} is the emission wavelength and λ_{obs} is the wavelength we observe on Earth; v is the speed of nearby receding object, and c is the speed of light in a vacuum. Eq.(1.1) effectively shows that an abstract quantity such as the normalized scale factor a (discussed in section 1.2) is directly tied to physical ones.

[5] The "background" in the name means that these photons are coming to us from all the directions (like background noise) unlike the ones from the distant stars or galaxies.

and the Universe became “dark”; it remained so until the first stars and galaxies emerged under the influence of DM. Indeed, generated anywhere between inflation and the QCD phase transitions, DM was undergoing the period of “stagspansion” [129] during the radiation dominated era. After its end, clusters of DM started to gradually accumulate mass by gathering into bigger and bigger structures. However, these were unable to undergo gravitational collapse since DM cannot lose energy by emitting light. The situation changed after recombination since baryonic matter, decoupled from radiation, was attracted by DM and so formed (faster than it would in the absence of DM) huge gas clouds composed mainly of hydrogen. Structures then formed from these clouds via *hierarchical clustering* (see [88] for a thorough discussion of the topic), meaning that smaller objects (such as stars and clusters of stars) formed first and these assembled into larger ones (such as galaxies and clusters of galaxies). The stars then reionized the intergalactic hydrogen marking the beginning of the *reionization epoch* somewhere between 100 and 950 million years[6] or $30 > z > 6$ [88]. The Universe “lighted up again” due to the presence of visible light sources, so the Dark Ages gradually ended. After the *Large-Scale Structure* (LSS) was formed, the Universe was largely unchanged for billions of years to come. However, roughly 9.8×10^9 years after the Big Bang [137], the matter dominated era ended, and we start talking about the present Universe.

Present Universe. In the matter domination epoch, the Universe’s expansion slowed down due to the gravitational force. However, at some point the Universe started its accelerated expansion yet again. To explain such behavior, the DE concept was introduced. Like DM, this entity is also unable to directly interact with light [50, 84, 105]. It is believed to comprise $\sim 68\%$ of the total energy density of the Universe, which makes it the most dominant component, followed by DM. The effect of the DE (which may be represented by the cosmological constant Λ or a dynamical field) is believed to remain unchanged through time, and this is unlike gravity which becomes weaker and weaker the further the objects are. After about 10 billion years of cosmic time, the Universe started a period of acceleration expansion.

Future Universe. The Universe dominated by Λ will continue its accelerated expansion until infinity resulting in a “Big Freeze” — the state of the Universe where everything has cooled down and effectively stopped.

1.2 Quantitative Description: General Relativity Treatment

Before we start any cosmological calculations, we need to decide on a coordinate system. Since space is constantly stretching, it would have proven difficult to divide the Universe into an equidistant grid if it were not for the assumption of homogeneity and isotropy of space. We will call the length element of the

---

[6] Calculated from the redshift using cosmology calculator [161]: https://astro.ucla.edu/~wright/CosmoCalc.html

grid the *scale factor*, $\mathcal{R}$, and it is a time-dependent entity. To not be bothered with the issue of dimensionality, we can simplify the picture by introducing the *normalized scale factor* [158]

$$a(t) \equiv \frac{\mathcal{R}(t)}{\mathcal{R}_0}, \quad \mathcal{R}_0 = \mathcal{R}(t_0)\,, \tag{1.2}$$

which is equal to unity at the present time t_0. This approach allows us to change the usual *physical coordinates* into the so-called *comoving coordinates* via a simple transformation law,

$$\mathbf{r} = a\mathbf{x}\,. \tag{1.3}$$

Here $\mathbf{r}$ and $\mathbf{x}$ stand for the physical and comoving distances, respectively.

The scale factor governs the universal expansion rate. This implies that an equation or a set of equations should exist that describe how the Universe, together with all its matter content, evolves with time. And indeed, they do exist, and Alexander Friedmann was the first to derive them in 1922 [48].

In its most basic form, the Friedmann equations can be obtained from classical Newtonian views alone (see [85] and references therein), employing only the cosmological principle, the notion of comoving coordinates, and the energy conservation laws. Although elegant, this approach is too simplistic since it does not account for relativistic effects. This is done in the theory of GR — the main topic of this section.

1.2.1 Metric

Every theory of relativity has to account for two key aspects — invariants[7] and changes of reference frames. While developing his theory of relativity [34, 42], Einstein postulated that instead of time, the universal unchanged quantity is the speed of light in a vacuum, c. Secondly, he postulated that the laws of nature are the same in inertial reference frames.

These two postulates allow us to obtain a mathematically justifiable theory of gravity. As always, we start small initially. Let us consider the line distance between two objects. From usual calculus, we know that in order to calculate the distance, we should: (1) define a basis, (2) split this line into a set of infinitesimally small elements of equal length, (3) calculate these lengths via Pythagoras' theorem, and (4) sum up (or integrate) across all the line elements with boundary conditions being the beginning and the end of the line.

In GR, space and time form one single entity — *spacetime* — and that implies we should work in a 4-dimensional space. The invariant interval can then be written as[8]

$$ds^2 = \left(dx^0\right)^2 - \left(dx^1\right)^2 - \left(dx^2\right)^2 - \left(dx^3\right)^2\,, \tag{1.4}$$

where the first element, dx^0, is the time coordinate, and the others are spatial coordinates. The line element — called the *metric* — should be the same

---

[7] These are quantities that stay the same under basis transformations.

[8] I am using the $(+,-,-,-)$ metric signature.

regardless of the basis we choose. Hence, we can write the expression (1.4) in a more compact form

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu \,, \tag{1.5}$$

where $g_{\mu\nu}$ is called the *metric tensor* and the summation symbol is ommitted using the *Einstein summation convention.*

1.2.2 Field Equations

The *principle of least action* is the fundamental ingredient that allows us to infer the dynamics of the system. To utilize it in GR, we consider a gravitational field interacting with matter in a given volume element of spacetime. The total *action*, in this case, can be written as [134]

$$\begin{aligned} S_{\text{tot}} &= S_{\text{EH}} + S_\Lambda + S_{\text{M}}, \quad \text{where} \qquad (1.6)\\ S_{\text{EH}} &= \frac{1}{2\kappa}\int R\sqrt{-g}d^4x,\\ S_\Lambda &= -\frac{\Lambda}{\kappa}\int \sqrt{-g}d^4x,\\ S_{\text{M}} &= \int \mathcal{L}_{\text{M}}\sqrt{-g}d^4x\,. \end{aligned}$$

Here S_{EH} is called the *Einstein-Hilbert action* and it is dependent on the *Ricci scalar*, R, that describes the local curvature of spacetime; S_Λ describes the contribution of the cosmological constant Λ; S_{M} models the contribution of both baryonic and dark matter through the Lagrangian density $\mathcal{L}_{\text{M}}$; $\kappa = 8\pi G_N/c^4$ is called *Einstein constant* [34] expressed through the Newtonian constant, G_N, and the speed of light in vacuum, c. The integrals in (1.6) are taken via the local volume of spacetime, $\sqrt{-g}d^4x$, where g is the determinant of the metric tensor, $g_{\mu\nu}$.

By varying the total action in (1.6) with respect to the metric, it is possible to derive the *Einstein field equations* [158]

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R - \Lambda g_{\mu\nu} = \kappa T_{\mu\nu}, \quad \text{where} \quad T_{\mu\nu} = \frac{2}{\sqrt{-g}}\frac{\delta\left(\sqrt{-g}\mathcal{L}_{\text{M}}\right)}{\delta g^{\mu\nu}} \tag{1.7}$$

is called the *stress-energy* tensor. Eqs.(1.7) have ten independent components and fully describe the gravitational field in an arbitrary coordinate system.

The expanding homogeneous and isotropic Universe is described by the so-called *Friedmann-Lemaître-Robertson-Walker* (FLRW) metric, which represents a particular solution of eqs.(1.7). It takes the form

$$ds^2 = dt^2 - a(t)\left(\frac{dr^2}{1-Kr^2} + r^2 d\theta^2 + r^2\sin^2\theta d\phi^2\right)\,, \tag{1.8}$$

where K describes the curvature of spacetime and (r, ϕ, θ) are the spherical coordinates. Given this metric, Einstein's equations (1.7) are naturally reduced to the so-called *Friedmann equations* [48] (for derivation see, e.g., [158] and references therein)

$$H^2 + \frac{Kc^2}{a^2} = \frac{8\pi G_N}{3}\left(\rho + \rho_\Lambda\right), \quad \rho_\Lambda \equiv \frac{\Lambda c^2}{8\pi G_N}\,, \tag{1.9}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G_N}{3}\left(\rho + \frac{3p}{c^2}\right) - \frac{4\pi G_N}{3}\left(\rho_\Lambda + \frac{3p_\Lambda}{c^2}\right), \quad p_\Lambda = -\rho_\Lambda c^2\,, \tag{1.10}$$

where $H \equiv \dot{a}/a$ is the *Hubble expansion rate*[9], a is the scale factor; p and ρ are the pressure and energy density of matter, respectively, K/a^2 is the spatial curvature of spacetime with possible values for K being 0 (for the spatially flat Universe), -1 (for the hyperbolic, i.e., open Universe), and 1 (for the spherical, i.e., closed Universe); and $p_\Lambda/\rho_\Lambda = -1$ corresponds to the equation of state for Λ describing DE [98].

The Friedmann equations (1.9)-(1.10) describe the evolution of the scale factor with time; thus, they allow us to mathematically deduce the past, the present, and the future of the Universe. For instance, we see that eq.(1.10) implies that the expansion rate depends on the two competing terms on the right-hand side — the matter-radiation density of a perfect fluid and the cosmological constant. Given that the density and pressure of the matter-radiation fluid cannot be negative, the minus sign implies that the expansion rate will be negative if it were not for Λ. Hence, everything depends on the sign of the cosmological constant: if it is negative, then the Universe decelerates in its expansion; if it is positive and dominant over matter and radiation, then the universal expansion is accelerated. Since we know from observations that the latter statement is correct, we can conclude that Λ is positive and dominant at late times.

1.2.3 ΛCDM Model Parameters

When solving the first Friedmann equation (1.9), it is convenient to introduce the fractional energy densities at the present time in the following form [158]:

$$\Omega_i \equiv \frac{\rho_i(t_0)}{\rho_{\text{crit}}}, \quad \rho_{\text{crit}} \equiv \frac{3H_0^2}{8\pi G_N}, \quad \sum_i \Omega_i = 1\,, \tag{1.12}$$

where ρ_{crit} is the *critical density*[10] and ρ_i encompasses the energy densities of CDM, baryons, radiation, curvature and Λ. The quantities Ω_i effectively

---

[9] By observing Cepheid variable stars in distant but nearby galaxies, Edwin Hubble was the first to experimentally confirm the systematic expansion of the Universe [67]. He observed that galaxies are moving away from us, with the recession velocity $\mathbf{v}$ being proportional to the distance $\mathbf{r}$,

$$\mathbf{v} = H_0\mathbf{r}\,, \tag{1.11}$$

where the proportionality constant, H_0, is called the *Hubble constant* and eq.(1.11) is known as the *Hubble law*.

[10] The critical density is defined to be the matter energy density of present time for a spatially flat ($k = 0$) universe with $\Lambda = 0$ [158].

transform the Friedmann equation (1.9) into

$$\frac{\dot{a}}{a} = H_0 \sqrt{\Omega_{\rm rad} a^{-4} + \left(\Omega_{\rm CDM} + \Omega_b\right) a^{-3} + \Omega_K a^{-2} + \Omega_\Lambda}\,, \tag{1.13}$$

which mathematically depicts the picture of the Universe's timeline described in the beginning of this section. Indeed, the very early Universe corresponds to the scale factor $a \to 0$, which allows the radiation term to dominate the right-hand side. On the other hand, in the limit of $a \to \infty$, the right-hand side is dominated by Λ, which implies that the future evolution of the Universe is dictated by DE. The present Universe corresponds to $a(t_0) = 1$ and so its evolution right now is affected by all of these terms. However, the *Planck* 2018 [124] CMB data analysis results favor a spatially flat Universe with $|\Omega_K| \lesssim 0.005$. Furthermore, assuming a negligible radiation density ($\Omega_{\rm rad} = 10^{-4}$) and remembering that by their definition the sum of Ω_i's should add up to 1, we can express one of the parameters Ω_Λ, $\Omega_{\rm CDM}$, and Ω_b via the other two. Usually, the Ω_b and $\Omega_{\rm CDM}$ are chosen to be free; thus, I will take these as the *first two parameters* of the ΛCDM model.

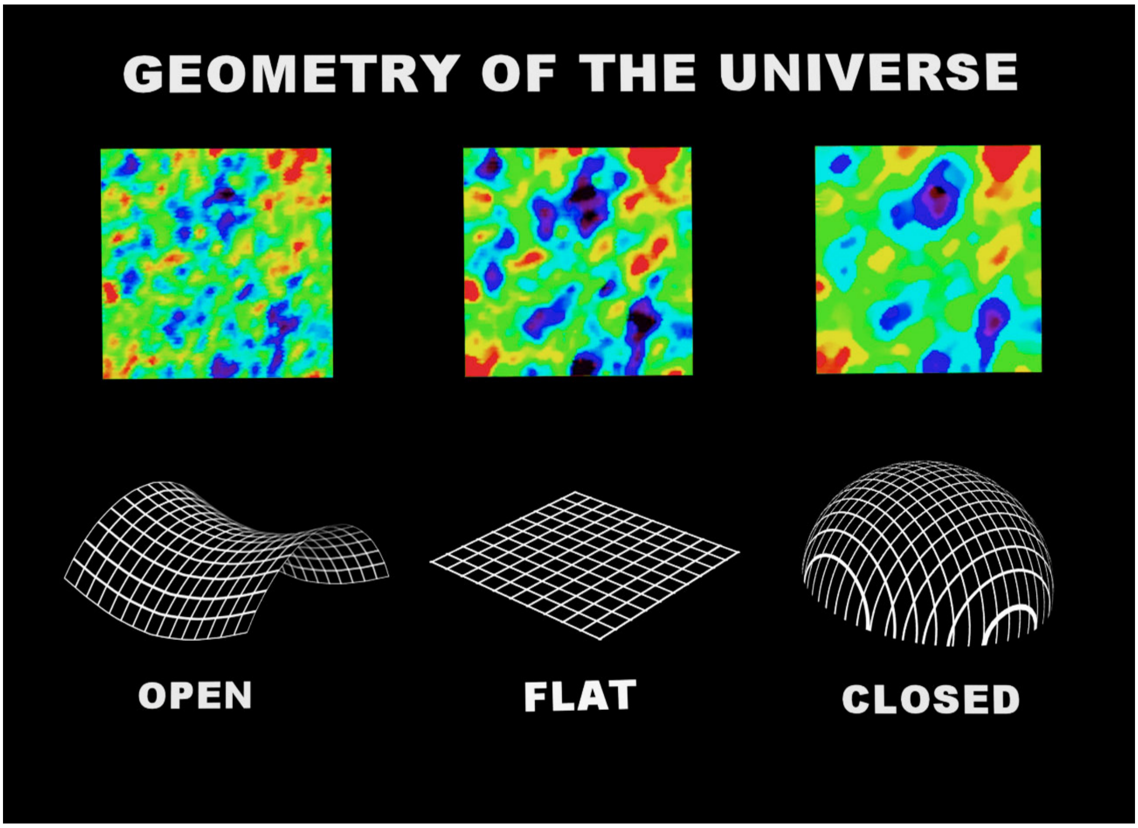

Figure 1.2: Schematic illustration of the CMB anisotropies for the case of open, flat, and closed Universe. Image Credit: NASA/WMAP science team.

Another set of parameters may be defined by assuming that the power spectrum of initial scalar perturbations is described in terms of a power law function. That is,

$$\mathcal{P}_s(k) = \mathcal{A}_s \left(\frac{k}{k_0}\right)^{n_s - 1}, \tag{1.14}$$

where n_s is called the scalar spectral index and is a parametrization of the power spectrum tilt, and $\mathcal{A}_s$ is the power spectrum amplitude characterized by the pivot wave number k_0.

One more parameter that has its imprint on the CMB comes from the primordial plasma of baryons and photons. Since the plasma was a very hot and dense substance, there were constantly two forces battling each other for the upper hand — the outward pressure and gravity. The result of these encounters was the creation of *cosmic sound* — the spherical waves that move away from the overdense regions. These are called *baryon acoustic oscillations* (BAO).

As already mentioned, the older the Universe, the colder the plasma; thus, the tight coupling between photons and baryons weakens with time. At some point of recombination, called the *drag epoch*, the Compton scattering rate of photons off baryons became equal to the expansion rate of the Universe. That was the moment when baryons and photons decoupled.

This epoch is characterized by the redshift z_{drag}, and this allows us to calculate the value of the *sound horizon* — the comoving distance traveled by cosmic sound waves from the Big Bang till the drag epoch — as [158]

$$L_{\mathrm{SH}} = \int_0^{t_{\mathrm{drag}}} \frac{c_s(t)}{H(t)} \frac{da}{a^2(t)} \,, \tag{1.15}$$

which is used to infer another free cosmological parameter, namely the *angular size of the acoustic oscillations*, via the formula [124]

$$\theta_{\mathrm{BAO}} = L_{\mathrm{SH}}/D_{\mathrm{M}} \,, \tag{1.16}$$

where D_{M} is the comoving angular diameter distance[11].

The last parameter is the *optical depth to reionization* τ, which is defined as [35]

$$\tau \equiv \int_{\eta}^{\eta_0} n_e \sigma_{\mathrm{T}} a d\eta', \quad \dot{\tau} \equiv \frac{d\tau}{d\eta} = -n_e \sigma_{\mathrm{T}} a \,, \tag{1.17}$$

where n_e is the electron number density, η is the conformal time[12], and σ_{T} is the Thomson cross section[13].

All in all, I consider the following free parameters to comprise the standard[14] ΛCDM model: the baryon density at the present time, Ω_b; the dark matter

---

[11] $D_{\mathrm{M}} = (1+z)\, D_{\mathrm{A}}$, where D_{A} is the angular diameter distance i.e. the ratio of an object's proper (transverse) size to the angle subtended by the object on the sky (angular size) [158].

[12] *Conformal time* is the total comoving distance the light has traveled since the beginning of time, $t = 0$. Mathematically,

$$\eta \equiv \int_0^t \frac{dt'}{a(t')} = \int_0^{a_0} \frac{da}{aH} \,. \tag{1.18}$$

Since nothing can travel faster than light in a vacuum, the regions with a distance greater than η will not be causally connected. Therefore, we can think of η as the *comoving horizon.*

[13] The process of Thomson scattering and its significance to the polarization of the CMB is discussed in section 1.2.6.2

[14] Here, the important remark should be made. Generally, it is up to an individual to choose which parameters to treat as free since others can be derived directly or indirectly from them.

Table 1.1: Planck 2018 [124] constraints on base-ΛCDM parameters (estimated with 68% confidence intervals) derived from CMB power spectra, CMB lensing reconstruction, and BAO (not from the CMB, but the large-scale structure). The first six are free parameters, while others are derived quantities.

Parameter	TT, TE, EE + lowE	TT, TE, EE + lowE + Lensing	TT, TE, EE + lowE Lensing + BAO
$\Omega_{\rm b}h^2$	0.02236 ± 0.00015	0.02237 ± 0.00015	0.02242 ± 0.00014
$\Omega_{\rm c}h^2$	0.1202 ± 0.0014	0.1200 ± 0.0012	0.11933 ± 0.00091
$100\theta_{\rm BAO}$	1.04090 ± 0.00031	1.04092 ± 0.00031	1.04101 ± 0.00029
τ	$0.0544^{+0.0070}_{-0.0081}$	0.0544 ± 0.0073	0.0561 ± 0.0071
$\ln(10^{10}\mathcal{A}_{\rm s})$	3.045 ± 0.016	3.044 ± 0.014	3.047 ± 0.014
$n_{\rm s}$	0.9649 ± 0.0044	0.9649 ± 0.0042	0.9665 ± 0.0038
Ω_Λ	0.6834 ± 0.0084	0.6847 ± 0.0073	0.6889 ± 0.0056
t_0	13.800 ± 0.024	13.797 ± 0.023	13.787 ± 0.020
$\Omega_{\rm m}$	0.3166 ± 0.0084	0.3153 ± 0.0073	0.3111 ± 0.0056
σ_8	0.8120 ± 0.0073	0.8111 ± 0.0060	0.8102 ± 0.0060
$z_{\rm re}$	7.68 ± 0.79	7.67 ± 0.73	7.82 ± 0.71
H_0	67.27 ± 0.60	67.36 ± 0.54	67.66 ± 0.42
$10^9\mathcal{A}_{\rm s}e^{-2\tau}$	1.884 ± 0.012	1.883 ± 0.011	1.881 ± 0.010

density at the present time, $\Omega_{\rm DM}$; the angular size of BAO on the CMB sky (or at recombination), $\theta_{\rm BAO}$; the optical depth to reionization, τ; the scalar spectral index, n_s; and the scalar perturbation amplitude, $\mathcal{A}_s$. The current values of these parameters estimated by *Planck* are listed in table 1.1.

1.2.4 Problems and Extensions of ΛCDM Model

Even though the concordance model is a tremendous success due to its ability to account for most cosmological observations today, it nevertheless suffers from major drawbacks and unresolved problems. Most striking of these include (but are not limited to) the exact nature(s) of DE and DM, the so-called *cosmological constant* and *coincidence problems*[15] [71], and the existence of the CMB anomalies [112, 118, 125]. All this encourages the search for more general solutions and models, considering various extensions to ΛCDM and the standard theory of

---

The *minimal* number of free parameters in the standard ΛCDM model is seven, which includes Ω_K. This list can be extended even further by adding, for instance, the CMB temperature $T_{\rm CMB}$, or neutrino mass. The *typical* setup is to consider six, which implies "forcing" Ω_K to be 0, because the CMB data suggests the case of (almost) spatially flat Universe (the schematic illustration of the CMB anisotropies in the same sky patch in the case of different K is depicted in figure 1.2). For further discussion of the topic as to why we should not set $\Omega_K = 0$ I refer the reader to [8] and references therein.

[15] The *cosmological constant problem* stems from the fact that Λ has an extremely small theoretical value in Planck units [21, 93] with the observed value being by 120 (!) orders of magnitude lower than the Planck scale. At the same time, coincidentally, its energy density is on par with that of the present matter density.

gravity in the hope of discovering new exotic physics. Examples of this include signatures of modified gravity [24, 74] such as $f(R)$ gravity[16] [30, 99, 143] or unimodular gravity[17] [7, 45, 101]; DE in the form of "quintessence" (or dynamical DE) [50, 84, 92, 105, 155], etc. For a comprehensive recent review of these topics, I refer the reader to [20] and references therein.

Another example of extensions of the ΛCDM model is the introduction of an extra relativistic component into the ΛCDM setting, which could leave some observable traces in the CMB and the large-scale structure [37, 82, 142, 156], and may potentially, allow us to solve another cosmological challenge — the so-called *missing satellites problem* (see, specifically, [95] and references therein). In this regard, the main goal of Paper I was to reformulate the analytical perturbation theory for the case of the ΛCDM model plus relativistic particles. This was done by considering the system of point-like particles with comoving radii $\mathbf{r}_n$, momenta $\mathbf{q}_n$, and masses m_n with a given form of the energy-momentum tensor [1, 77] and the perturbed FLRW metric in the Poisson gauge form,

$$ds^2 = a^2\left[(1+2\Phi)d\eta^2 + 2B_i d\eta dx^i + (-\delta_{ij} + 2\Psi\delta_{ij})\, dx^i dx^j\right] . \tag{1.19}$$

As a result, the analytic expressions for the first-order scalar (Φ, Ψ) and vector (B_α) perturbations[18] together with the equations of motion for a given particle in the system were derived with a subsequent study of their asymptotic behavior. These expressions are valid for arbitrary cosmological (i.e., sub and superhorizon) scales and constitute the generalization of their counterparts in [38]. Even more so, they are of Yukawa-type with a characteristic finite and time-dependent screening length (i.e., the radius of interaction) that was also confirmed for various extensions of ΛCDM plus extra perfect fluids with linear and non-linear equations of state [22, 39, 40, 41]. These results can be used for the future high-precision cosmological simulations of structure formation, which involves the research of the roles played by neutrinos and WDM in it.

1.2.5 A Few Words on the Topology of the Universe

The CMB anomalies (such as the lack of large-angle CMB temperature anisotropies, the suppression of the quadrupole moment, and the "axis of evil"[19] [15, 16]) may suggest that our Universe is of non-trivial topology. Indeed, topology governs the global properties of the Universe and effectively describes its shape. Not only is it a fundamental topic related to the finiteness or the infiniteness of the Universe, but it also may lead to exciting observable features such as multiple images [91, 108, 133] of faraway objects (stars, galaxies, etc.). The search for these repeated patterns falls into the domain of cosmic crystallography [52, 79], which didn't discover such images within well-separated areas of the sky [53].

---

[16] In $f(R)$ gravities the Ricci scalar is replaced by some functions of the Ricci scalar in the Einstein-Hilbert action.

[17] The determinant of the metric is fixed to be -1.

[18] Described in more detail in section 1.2.6.

[19] The preferable alignment of the CMB quadrupole and octopole.

Still, *multiply connected*[20] spaces such as $T \times T \times R$ (an equal-sided chimney) or $T \times R \times R$ (a slab) could have provided solutions for the CMB anomalies mentioned above. This will not break the current theory of gravity since GR allows universes that are multiply connected while having finite volume and zero (or non-zero) curvature [113].

Historically, the term *cosmic topology* was introduced in the 1990s [76], but first discussions of the topic can be traced way back to works of Schwarzschild [139, 146], Friedmann [49] and others (see [90] for a brief historical overview). However, studies were primarily abandoned until the observational CMB data arrived from the *COBE*, *WMAP*, and *Planck* experiments. In particular, a search for matched circles in the 7-year *WMAP* data constrained the lower bound on the comoving distance to our nearest topological clone for a spatially flat universe to be $\simeq 27.9$Gpc [15]. *Planck* 2013 results [113] constrained the radius R_i of the largest sphere inscribed into the topological domain for $\Omega_K = 0$ to be $R_i > 0.71 d_{\mathrm{rec}}$ and $R_i > 0.50 d_{\mathrm{rec}}$ for chimney and slab, respectively (where $d_{\mathrm{rec}} \sim 14$Gpc is the distance to the recombination surface). *Planck* 2015 results [119] made these constraints even stronger for a slab topology resulting in $R_i > 0.56 d_{\mathrm{rec}}$.

By definition, the slab topology has only one finite dimension, which was tempting to interpret as the “axis of evil” [9, 47]. Therefore, it was interesting to see what kind of gravitational potential would be induced by the point-like massive particles in this case. This was the idea behind Paper II.

It had been already shown that there is no way to obtain any physically meaningful result for the gravitational potential sourced by the matter density fluctuations in the Newtonian limit [18]. The situation is changed, however, if one considers the relativistic effects. In this case, the resulting equation obtained from the perturbed Einstein equations is then not of Poisson but of Helmholtz-type [38, 40, 41]. Consequently, we can obtain physically reasonable solutions in two distinct forms: (1) the expansion of the delta function in terms of Fourier series (due to the periodical property along the toroidal dimension) and (2) the direct summation of Yukawa-type potentials which correspond to the source particle and all its images. Lastly, it was demonstrated that the latter solution is preferable for numerical computations since it requires fewer terms in the series to achieve the required precision.

Paper II had a similar reasoning. However, in this case, we assumed the chimney topology instead. The first two solutions were also based on Fourier series and Yukawa-type potentials. There was also an alternative solution for the potential, which was formulated using the Ewald sums applied to Yukawa-type potentials. Still, in this case, the plain summation of Yukawa-type potentials was preferable for the numerical computations.

---

[20] The space is called *simply connected* if any path between two random points can be continuously transformed into any other path between them while the points are unchanged. In addition, every closed loop in the domain can be continuously shrunk into a point [132]. Simply put, this is a space that does not have holes spanning through it. Otherwise, the topological space is *multiply connected*. An example of simply connected space is a sphere, while a doughnut or coffee cup with a handle is multiply connected.

That being said, the latest results based on the matched circles searches [2, 107] place even tighter constraints than *Planck*, and put limits on topological length scales for a number of Euclidean topologies, among which are slab and chimney. It does not mean, however, that the search has stopped since there are many cases of non-trivial topologies for flat (and non-flat) geometries.

1.2.6 Cosmic Microwave Background

Since matter is distributed unevenly throughout spacetime, the baryonic and DM densities vary from region to region. Together with photons and neutrinos, the denser regions perturb the curvature of spacetime and generate gravitational potential variations. To account for these, we can generalize the metric into[21]

$$ds^2 = \left[1 + 2\Psi\left(\mathbf{x}, t\right)\right] dt^2 - a^2 \left[1 - 2\Phi\left(\mathbf{x}, t\right)\right] \delta_{ij} dx^i dx^j \ , \tag{1.20}$$

where Φ describes the Newtonian potential and Ψ characterizes perturbations of curvature. Equation (1.20) includes only *scalar* perturbations that couples matter/radiation with gravitational potentials. However, these are not the only ones. Others include *vector* and *tensor* perturbations. While the former is connected with topological defects, the latter is supposed to describe primordial gravitational waves (GWs) — the ripples of spacetime generated during the inflation period, which are particularly important for the CMB. Indeed, these ripples distort spacetime by inducing a local quadrupole moment, meaning that GWs should have an imprint on the CMB polarization — an extremely important topic in observational cosmology, and the main theme of this section.

1.2.6.1 Temperature Anisotropies

Given the metric (1.20), it is possible to derive (for the complete treatment, see [35] and references therein) the *Boltzmann equations* that simultaneously account for the coupling between baryons, photons, neutrinos, and dark matter, and effectively describe the distribution of matter in the Universe. In its unintegrated form, it can be written as

$$\frac{df}{dt} = C\left[f\right] \ , \tag{1.21}$$

where $f = f\left(\mathbf{p}(t), \mathbf{x}(t), t\right)$ corresponds to the one-particle distribution function of the position $\mathbf{x}$, momentum $\mathbf{p}$, and time t; and $C\left[f\right]$ is the functional of f that accounts for all the collisions between all the particles in the combined fluid. We are particularly interested in the photon distribution function, which can be expanded around its zero-order Bose-Einstein value as

$$f\left(\mathbf{x}, p, \hat{\mathbf{p}}, t\right) = \left[\exp\left\{\frac{p}{T(t)\left[1 + \Theta(\mathbf{x}, \hat{\mathbf{p}}, t)\right]}\right\} - 1\right]^{-1} \ , \tag{1.22}$$

---

[21] The metric in eq.(1.20) is written in *Newtonian gauge*, which is not the only one — in principle, it is possible to use others as well.

where $T(t)$ characterizes the temperature of the homogeneous and isotropic Universe, and Θ accounts for perturbations to the distribution and describes the inhomogeneous and anisotropic Universe. Apart from the observer's position $\mathbf{x}$ and the time t, Θ is a function of only the direction of photon propagation represented by the unit vector $\hat{\mathbf{p}}$ and not by the complete photon momentum $\mathbf{p}$. That is because, as part of the single fluid, the CMB photons were continuously scattered off the electrons until the Universe became large enough that the mean free path of the primordial photons became infinite. That was the moment when the CMB photons scattered for the last time, decoupled from the fluid and free-streamed into space. This fact is precisely what makes the CMB so valuable — since they have been traveling ever since without interactions, the photon energy, albeit lesser due to the frequency shift, carries the imprint of how the Universe was during its early times.

This is supported by observations, which show that the Universe was very smooth to first order, with the CMB being almost a perfect black body having the temperature of about $T = 2.725K$ [46]. However, there are tiny deviations in temperature across the sky, which are believed to be the imprints of primordial perturbations. These anisotropies $\delta T/T$ are described in terms of Θ, which allows us to write the expression for the CMB temperature as

$$T\left(\mathbf{x}, \hat{\mathbf{p}}, \eta\right) = T(\eta)\left[1 + \Theta\left(\mathbf{x}, \hat{\mathbf{p}}, \eta\right)\right] , \tag{1.23}$$

where Θ can be expanded in terms of spherical harmonics as [35]

$$\Theta\left(\mathbf{x}, \hat{\mathbf{p}}, \eta\right) = \sum_{l=1}^{\infty} \sum_{m=-l}^{l=1} a_{lm}\left(\mathbf{x}, \eta\right) Y_{lm}(\hat{\mathbf{p}}) , \tag{1.24}$$

where l and m correspond to the conjugate of $\hat{\mathbf{p}}$; and Y_{lm}'s are spherical harmonics. All the information about the temperature field is contained in the spherical harmonic coefficients or the amplitudes a_{lm}. The properties of the amplitudes stem from the assumption of isotropy and homogeneity of spacetime, i.e., the mean value is zero, and the amplitudes' variances are written as

$$\langle a_{lm} \rangle = 0, \quad \langle a_{lm} a^*_{l'm'} \rangle = \delta_{ll'} \delta_{mm'} C_l \, . \tag{1.25}$$

Here, C_l is the angular power spectrum that essentially connects the variance with the angular scale and is used to compare the observed sky with theoretical models. Indeed, the angular power spectrum and the best-fit ΛCDM model are shown in figure 1.3. While it is outside of the scope of this work to describe what physical processes result in the locations and positions of the peaks of the power spectrum, it is crucial to note that the power spectrum perfectly fits the ΛCDM model discussed in this chapter.

Another point to mention is that the large error bars are due to the so-called *cosmic variance*, which is essentially an uncertainty in the estimate of C_l from one CMB map,

$$\hat{C}_l = \frac{1}{2l+1} \sum_{m=-l}^{l} \left|a_{lm}\right|^2 , \quad \frac{\Delta \hat{C}_l}{\hat{C}_l} = \sqrt{\frac{2}{2l+1}} , \tag{1.26}$$

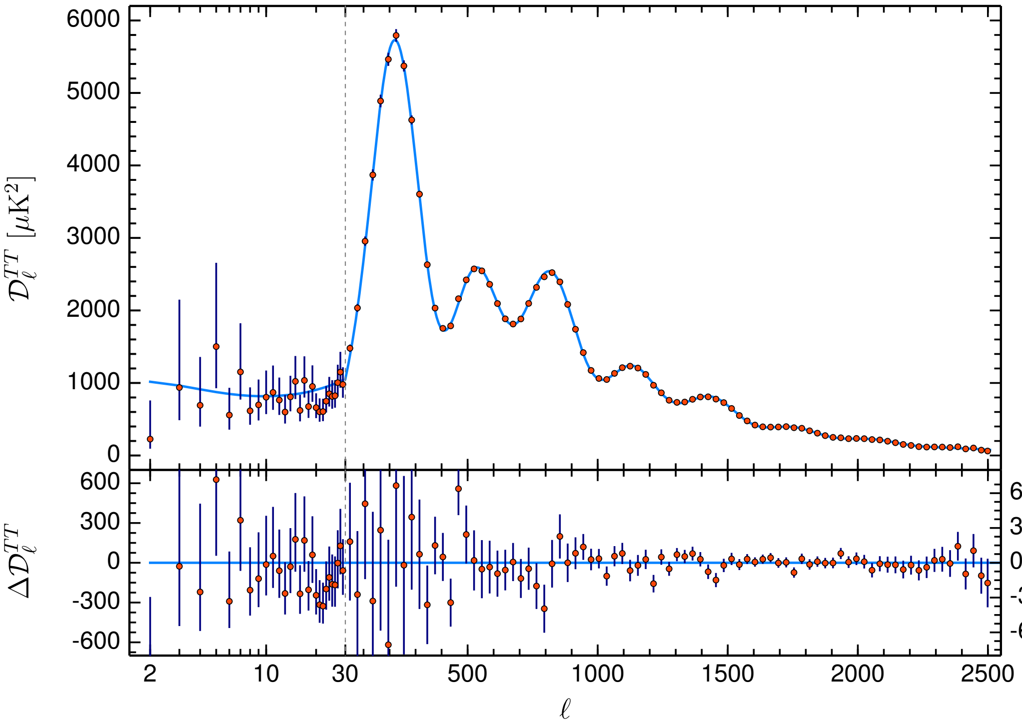

Figure 1.3: Planck 2018 [124] temperature angular power spectrum, $D_l \equiv l\,(l+1)\,C_l/2\pi$. Image credit: Planck Collaboration.

where $\hat{C}_l$ indicates an estimator.

1.2.6.2 Polarization of the CMB Light

The light wave is composed of three parts: electric, **E**, magnetic, **B**, and propagation, **k**, vectors, respectively. The *polarization* of light effectively describes the spatial orientation of magnetic and electric fields. Since the CMB was emitted from an anisotropic process, it is slightly ($\sim 10\%$) polarized. The anisotropic process here is the *Thomson scattering* — the elastic scattering of photons off the electrons that did not change their wavelengths but only their directions.

Generally, there are several different types of polarization, but the Thomson scattering results in the CMB being only *linearly polarized.* No known cosmological process produces *circular polarization* of the CMB light, although there are some theories (for instance, see [75] and references therein) that predict circular polarization. Mathematically, it is common to describe the polarization in terms of four Stokes parameters

$$\begin{aligned} I &\equiv |E_x|^2 + |E_y|^2 \,, \quad & Q &\equiv |E_x|^2 - |E_y|^2 \,, \\ U &\equiv 2\mathrm{Re}\left(E_x E_y^*\right) , \quad & V &\equiv 2\mathrm{Im}\left(E_x E_y^*\right) , \end{aligned} \tag{1.27}$$

where I describes the intensity of light, V describes circular polarization, and Q and U describe linear polarization in two systems relatively rotated by 45°.

Eq.(1.27) employs the so-called *Q-U decomposition*, which is very convenient to describe the linearly polarized light. However, due to the lack of rotational invariance, it is not the best way to represent the CMB.

Indeed, the CMB field can be described as a spin-2 field, i.e., the vector field with a difference of 180°. By employing spin-2 spherical harmonics, we write the Stokes Q and U parameters in the so-called *E-B decomposition* as

$$\left(Q \pm iU\right)(\hat{\mathbf{n}}) = \sum_{\ell=2}^{\ell_{\max}} \sum_{m=-\ell}^{\ell} {}_{\pm 2}a_{\ell m}\, {}_{\pm 2}Y_{\ell m}(\hat{\mathbf{n}})\,, \tag{1.28}$$

where the spin-k spherical harmonic coefficients, ${}_k a_{\ell m}$, construct the E-B mode power spectra

$$a^E_{\ell m} = -\frac{1}{2}\left({}_2a_{\ell m} + {}_{-2}a_{\ell m}\right), \quad \delta_{ll'}\delta_{mm'}C_l^{EE} = \left\langle a^E_{lm}(a^E_{l'm'})^*\right\rangle, \tag{1.29}$$

$$a^B_{\ell m} = -\frac{1}{2\mathrm{i}}\left({}_2a_{\ell m} - {}_{-2}a_{\ell m}\right), \quad \delta_{ll'}\delta_{mm'}C_l^{BB} = \left\langle a^B_{lm}(a^B_{l'm'})^*\right\rangle, \tag{1.30}$$

and cross-correlations C^{XY}, where (X, Y) are taken from pairs (T, E, B), can be constructed similarly (see [163] for more details).

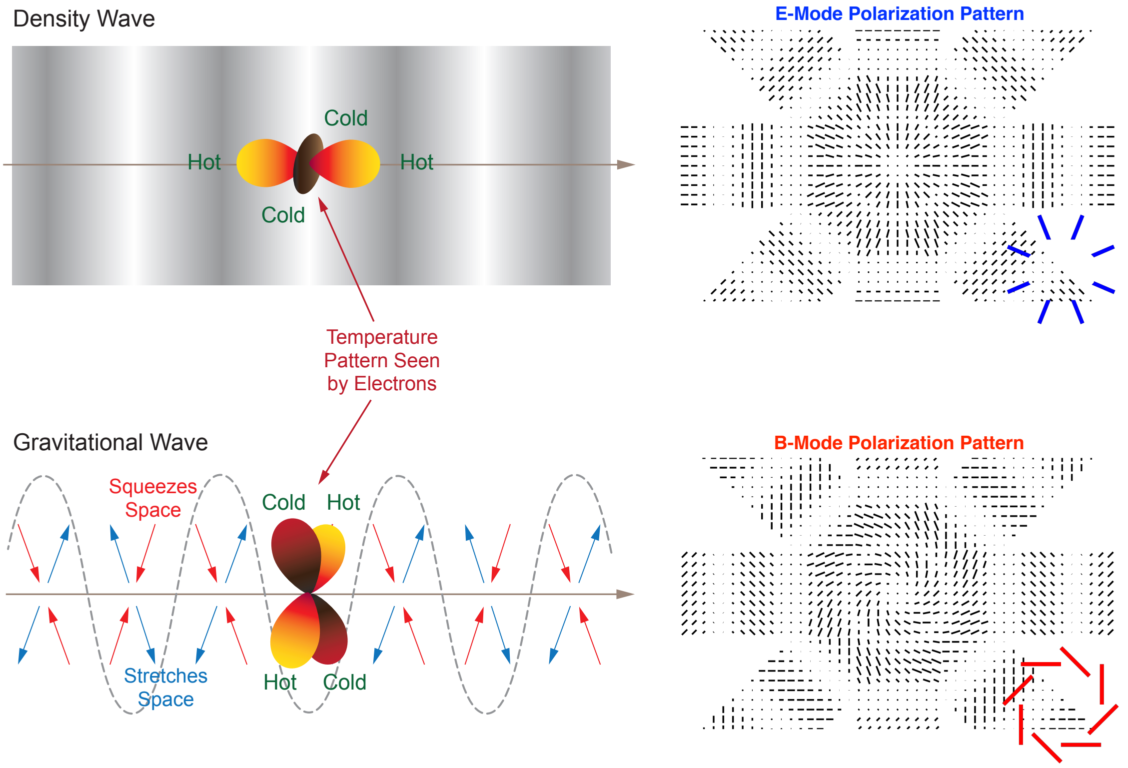

Figure 1.4: The polarization pattern of E-B decomposition induced by scalar (top) and tensor (bottom) perturbations. Image Credit: BICEP/Keck collaboration [26].

To understand where the names E- and B-modes come from, it is helpful to look at their polarization pattern depicted in figure 1.4. It is clear that E-

and B-modes are curl- and divergence-free in a full analogy with electric and magnetic fields; hence, the names. This decomposition is rotationally invariant, with their power spectra specified by the conditions at the last-scattering surface.

As was already mentioned before, tensor perturbations have an imprint on the CMB polarization. In particular, they are responsible for the B-mode production [163] — scalar perturbations can only produce E-modes. Therefore, by detecting large-scale B-modes one can directly test the theory of inflation that predicts primordial gravitational waves. In addition, the B-mode amplitude is associated with the inflation energy scale, allowing for a distinction between different inflation scenarios. Indeed, by introducing the *tensor-to-scalar* ratio

$$r \equiv \frac{P_k^T}{P_k^S}\Big|_{k=0.02\mathrm{Mpc}/h} \approx \frac{C_l^{BB}}{C_l^{EE}}\Big|_{l\approx 100}\,, \tag{1.31}$$

we can probe the potential energy of inflation since the latter is calculated via the so-called *slow-roll parameter*, ϵ, proportional to the ratio of primordial power spectra P_k^X (where $X = (S, T)$ for scalar and tensor perturbations, respectively) calculated at the time of recombination.

The parameter r is of particular importance since, by estimating its value, we are able to rule out classes of incorrect inflationary models. Hence, it is not surprising that many scientists around the globe contribute to newer and newer experiments to further tighten our constraints on r and other cosmological parameters. The most influential of these experiments are described in the next chapter.

Chapter 2

CMB Data Analysis

The never-ending quest to catch the CMB photons started in 1965 by the American astronomers Arno Penzias and Robert Wilson [106], which made them the Nobel Prize winners 14 years later. Incredibly homogeneous and uniformly distributed light on the sky with an almost perfect black body spectrum has captivated the hearts and minds of numerous researchers ever since. Hence, it is no surprise that many CMB missions were designed to squeeze as much information as possible from this signal. This chapter is dedicated to these experiments and the CMB data analysis techniques used by the ones I was involved with during my PhD at the University of Oslo.

2.1 CMB Experiments Overview

The multitude of CMB experiments conducted throughout the years can roughly be divided into *ground-based*, *balloon-bourne*, and *satellite-based*. Although it is beyond the scope of this thesis to go through each one that ever existed (the topic no less fascinating!), it is important to mention several that undoubtedly shaped how the entire field operates today. And the first in line should be the *Cosmic Background Explorer* (*COBE*) — the NASA satellite operated from 1989 until 1993. It is precisely due to *COBE* we now have the first scientific evidence that CMB is a black body and that there are inherent anisotropies within its distribution on the sky.

The incredible success of *COBE* resulted in a Nobel Prize award in 2006 to John C. Mather and George F. Smoot, and led to the development of another major satellite-based experiment sponsored by NASA — the *Wilkinson Microwave Anisotropy Probe* (*WMAP*) — operated from 2001 until 2010, and targeting five frequencies ranging from 23 to 94 GHz [13]. It was a mission designed specifically to determine the geometry and the content of the Universe by mapping the CMB sky to a resolution of $14'$ *Full Width Half Maximum* (FWHM). It is not an exaggeration to say that *WMAP* provided a piece of key evidence to establish ΛCDM as the standard model of cosmology.

WMAP was succeeded by *Planck* – the European Space Agency (ESA) led satellite mission that mapped the temperature anisotropies of the sky in nine distinct frequencies with its two unprecedentedly sensitive[1] instruments, namely the *Low-Frequency Instrument* (LFI) and *High-Frequency Instrument* (HFI). *Planck* data will remain an invaluable tool for any temperature and polarisation analysis in the foreseeable future.

---

[1] *Planck* had higher resolution and was even more sensitive than *WMAP*, which was 45 times more sensitive than *COBE*. Figure 2.1 shows the visual comparison in resolution from *COBE*, *WMAP* and *Planck* of the same patch of the sky to further clarify this point.

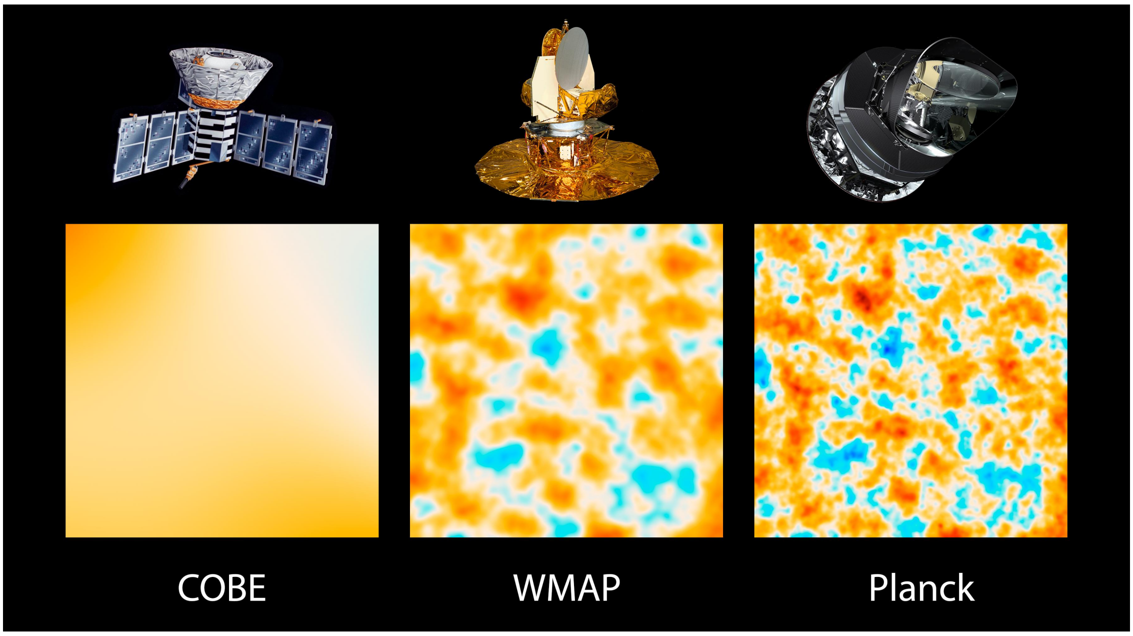

Figure 2.1: The same ten square degree patch of the sky of CMB anisotropies as an example of resolution improvement from *COBE* until *Planck*. Image credit: NASA/JPL-Caltech/ESA.

2.2 Data Analysis

Any CMB mission continuously scans (either a chunk of or the entire) sky via detectors that cover various frequency channels to measure temperature and polarization signals. During this process, both the orientation of the telescope and the set of data points, d_1, ..., d_N, called *time-ordered data* (TOD), are recorded together with other auxiliary data such as *half-wave plate* (HWP) position or the instrument's thermal/electric monitor. Overall, the measured signal often comprises TB of data and consists not only of CMB (*background signal*) but of many other components (*foregrounds*) that can be roughly divided into several types:

Diffuse Galactic foregrounds originate from the Milky Way galaxy and include:

- **Synchrotron** emission that dominates the radio sky in both temperature and polarization, and is generated by relativistic electrons ejected from supernovae and subsequently gyrating in the Milky Way magnetic field. The polarization fraction of synchrotron radiation varies across the sky, with a theoretical maximum being 75%. However, in practice, typical values rarely exceed 15% and reach 30% - 50% only in "Galactic Spurs" [120].
- **Free-free emission** (or *bremsstrahlung*) are photons generated by thermal electrons that scatter on protons in the ionized *interstellar medium* (ISM). It is generated when ISM temperature is comparable[2] to the binding energy

---

[2] About $10^3 - 10^4$K.

of hydrogen atoms; thus, the component correlates with the HII emission line, which makes it an indicator of the star formation regions. Even though the free-free component is unpolarized due to being generated from an isotropic process [136], it is nevertheless crucial for CMB experiments since it is the only component present at every frequency between 1 and 1000 GHz.

- **Anomalous microwave emission** (AME) is emission of an unknown origin observed around 20GHz, and was first identified by Leitch et al. [80] in 1997. Currently, the most accepted mechanism for generating this radiation is believed to be the rotation of the dust grains, whose kinetic energy is comparable to the surrounding gas. The component is at most weakly polarized with a *polarization fraction*, $p < 1\%$ [66], and is correlated with thermal dust, although it exhibits different spectral behavior.
- **Thermal dust emission** is generated via the vibration of tiny dust grains[3] in the ISM [36], typically composed of at least silicates and carbon. The component dominates the microwave sky at high frequencies. Furthermore, since they are keen to align with the local magnetic field, the set of dust grains produces a net polarization of $p \lesssim 22\%$ [126].

Extragalactic sources such as galaxies or clusters of galaxies emit light that also obstructs our vision of CMB distribution. Two components to be cautious about are the *Cosmic Infrared Background* (CIB) [64] — the infrared light from far away galaxies that was redshifted into the microwave frequency range — and the *Thermal Sunyaev-Zel'dovich* effect [[149] — the effect of CMB light distortion due to it being inverse Compton scattered with the hot electrons while the photons were passing through galaxies. Although the latter is only relevant on small angular scales, the former has similar *spectral energy distribution* (SED) to that of thermal dust, which makes it hard to separate.

Zodiacal Emission is the radiation coming from the dust particles thermalized by the Sun and is essentially the time-varied component due to the satellite's motion around the Sun. It is vital at high frequencies [111].

The end goal of any CMB data analyst is to calculate the CMB power spectra and subsequently constrain the cosmological parameters described in the previous chapter. This is done in a sequence of several distinct steps that together comprise what is called the (traditional) *pipeline*:

Data Preprocessing and Calibration. To get meaningful results, the raw TODs should be calibrated by known sources; flagged; and cleaned by removing any gross time-domain systematics (via, e.g., filtering or template subtraction).

---

[3] The sizes ranges from 1nm to 1μm in diameter.

Map Making is a process that significantly reduces the data set size without losing important information. The idea is to use knowledge of an instrument to estimate the sky signal given as a set of Stokes parameters at each observational frequency. As a result of this operation, a set of sky (frequency) maps are produced that ease visual observation of spatial structures of the sky.

Component Separation is a process of separating CMB from foregrounds described above, given the sufficient number of frequency maps. As a result of this process, the ensemble of component maps is obtained. If the number of frequency maps is insufficient to obtain a clean map, then masking is used together with foreground templates from other sources.

Power Spectra Estimation. Using the CMB in combination with frequency maps, the power spectra (two-point correlation functions) of T, E, and B polarisation modes are calculated.

Cosmological Parameters Estimation. The last step is to derive the best-fit cosmological parameters by comparing the theoretical power spectra with the ones obtained from data analysis.

Each of the steps above have traditionally been handled by a specific group of dedicated professionals all over the globe. Usually, due to their effort, a pipeline was developed to analyze a specific dataset. Instead, the BeyondPlanck project proposes establishing a common framework to handle many different datasets jointly. To give a general overview of this approach, we should first define a data model.

2.2.1 Sky Parametrisation and Ideal Instrumental Model

Mathematically speaking, the single data point, d_t^i, produced by the detector i at time t can be written as

$$d_t^i = g(t) \left[I_{\hat{\mathrm{p}}(t)} + Q_{\hat{\mathrm{p}}(t)} \cos\left(2\psi_t^i\right) + U_{\hat{\mathrm{p}}(t)} \sin\left(2\psi_t^i\right) \right] + n_t^i \,, \tag{2.1}$$

where (I, Q, U) are the Stokes parameters observed at position $\hat{\mathbf{p}}(t)$ and time t; the detector is sensitive to the polarization with angle ψ along the direction $\hat{p}(t)$ on the sky; n_t^i corresponds to *noise* — the artificial signal produced by various components in the electronic chain; $g(t)$ is called *gain*, and its purpose is to convert *Volts* (the digitized units) into K_{CMB} (physical units).

Under the assumption that all detectors have the same narrow spectral bandpass and angular resolution, the expression (2.1) can be written in a more compact form by combining the signal from all the detectors into a single vector, $\mathbf{d}$, such that

$$\mathbf{d}_t = \mathbf{G}\mathbf{P}_{tp(t)}\mathbf{s}_{p(t)} + \mathbf{n}_t \,, \tag{2.2}$$

where $\mathbf{P}$ is called the *pointing matrix*, $\mathbf{s}_{p(t)}$ and $\mathbf{n}_t$ are vectors of sky and noise signals, respectively, and $p(t)$ is the pixel observed by the position vector $\hat{\mathbf{p}}(t)$ at time t. The elements in eq.(2.2) have the following properties

- $\mathbf{G}$ is a diagonal $N_{\mathrm{TOD}} \times N_{\mathrm{TOD}}$ matrix with values g_t on its main diagonal.
- $\mathbf{P}$ is a sparse matrix that translates between time- and pixel-valued vectors with one non-zero entry per raw and are given by

$$\mathbf{P}_{tp(t)} = \begin{bmatrix} 1 & \cos(2\psi_t) & \sin(2\psi_t) \end{bmatrix} . \tag{2.3}$$

- $\mathbf{s}$ is a vector of pixel-ordered Stokes parameters with each element is given by

$$\mathbf{s}_{p(t)} = \begin{bmatrix} \mathbf{I}_p & \mathbf{Q}_p & \mathbf{U}_p \end{bmatrix} . \tag{2.4}$$

- $\mathbf{n}$ is a N_{samp} element vector of time-ordered noise in the time-stream.

The data model in (2.2) can be generalized by noting that (1) no detector measures a single frequency signal, but instead, a range of frequencies, and (2) the data sets are usually provided in temperature instead of brightness units, and so the unit conversion is necessary. By defining the bandpass profile, $\tau(\nu)$, and the unit conversion factors, U, as described in Paper IV, the data model can be written as

$$\mathbf{d} = U\mathbf{G}\mathbf{P} \int \mathbf{s}(\nu)\tau(\nu)d\nu + \mathbf{n} , \tag{2.5}$$

where conversion from units X_i into units X_j is given by

$$U_{ij} = \left(\int \tau(\nu) \frac{dI_\nu}{dX_i} d\nu \right) \bigg/ \left(\int \tau(\nu) \frac{dI_\nu}{dX_j} d\nu \right) , \tag{2.6}$$

and the sky model of ith component is defined in the form of $s_i(\nu) = a_i f_i\left(\nu, \nu_{0,i}, \beta\right)$ with a_i being linear amplitude of the foreground or CMB component with respect to the reference frequency $\nu_{0,i}$, and $f\left(\nu, \beta\right)$ is some frequency scaling function dependent on arbitrary frequency ν and ensemble of spectral parameters β. The BeyondPlanck sky model in brightness temperature units is given by

$$
\begin{aligned}
\mathbf{s}_{\mathrm{RJ}} = {} & \left[a_{\mathrm{CMB}} + a_{\mathrm{quad}}(\nu)\right] \frac{x^2 e^x}{\left(e^x - 1\right)^2} \\
& + a_{\mathrm{s}} \left(\frac{\nu}{\nu_{0,\mathrm{s}}}\right)^{\beta_{\mathrm{s}}} + a_{\mathrm{ff}} \left(\frac{\nu_{0,\mathrm{ff}}}{\nu}\right)^2 \frac{g_{\mathrm{ff}}\left(\nu, T_e\right)}{g_{\mathrm{ff}}\left(\nu_{0,\mathrm{ff},T_e}\right)} \\
& + a_{\mathrm{ame}} \left(\frac{\nu_{0,\mathrm{ame}}}{\nu}\right)^2 \frac{f_{\mathrm{ame}}\left(\nu \cdot \frac{30\mathrm{GHz}}{\nu_p}\right)}{f_{\mathrm{ame}}\left(\nu_{0,\mathrm{ame}} \cdot \frac{30\mathrm{GHz}}{\nu_p}\right)} \\
& + a_{\mathrm{d}} \left(\frac{\nu}{\nu_{0,\mathrm{d}}}\right)^{\beta_{\mathrm{d}}+1} \frac{\exp\left(\frac{h\nu_{0,\mathrm{d}}}{k_{\mathrm{B}} T_{\mathrm{B}}}\right) - 1}{\exp\left(\frac{h\nu}{k_{\mathrm{B}} T_{\mathrm{B}}}\right) - 1} \\
& + U_{\mathrm{mJy}} \sum_{j=1}^{N_{\mathrm{src}}} a_{j,\mathrm{src}} \left(\frac{\nu}{\nu_{0,\mathrm{src}}}\right)^{\alpha_{j,\mathrm{src}}-2}, \quad x = \frac{h\nu}{kT_0}\,,
\end{aligned}
\tag{2.7}
$$

where the only polarised components are the amplitudes a_{CMB}, a_{s} and a_{d}.

Since we want to constrain the signal measured by the detector j given the data, it is useful to introduce the *mixing matrix*, $\mathbf{M}_i^j$, as

$$
\mathbf{s}^j = \sum_{i=1}^{N_{\mathrm{comp}}} a_i \left(U_j \int f_i(\nu, \beta) \tau_j(\nu) d\nu\right) \equiv \sum_{j=1}^{N_{\mathrm{comp}}} \mathbf{M}_i^j \mathbf{a}_i = \mathbf{M}^j \mathbf{a}\,, \tag{2.8}
$$

that results in the *idealised data model*

$$
\mathbf{d} = \mathbf{GPMa} + \mathbf{n}\,. \tag{2.9}
$$

2.2.2 Systematics and Real Instrumental Model

The instrument produces data by scanning the same sky areas multiple times during the observational periods. If we lived in an ideal world, the observations of the same sky area but different instrumental orientations would all be described by the data model above. However, in the real world, all sorts of things can occur (such as random spikes in electronics, weather conditions, bad instrument calibration, etc.) that may impact, modify and/or corrupt the dataset. The effects originating from instrumental non-idealities are collectively known as *systematics* and are a topic of great concern for any CMB scientist.

2.2.2.1 Bandpass

Several factors describe the sensitivity of the detectors. The first one is, of course, the bandpass profile introduced above. The issue here is that τ is usually measured in the laboratory prior to the experiment, making it very hard to estimate precisely in the case of multi-frequency detectors. Therefore, the measurements of τ are often inaccurate and this can introduce a shift in its central frequency. To model this effect, we introduce a linear shift, Δ_{bp}, in the bandpass profile

$$\tau(\nu) = \tau_0 \left(\nu + \Delta_{\mathrm{bp}}\right) , \tag{2.10}$$

with τ_0 being the laboratory measurement. The exact choices of Δ_{bp} are treated in [150] and are beyond the scope of this work. What is important to note, however, is the fact that since two detectors measure the sky signal within the same frequency channel differently, this may generate a so-called *bandpass mismatch error* in the frequency maps.

2.2.2.2 Beam Profile

The next factor related to the detector's spatial sensitivity is called the *beam profile*, and it is effectively the point spread function that characterizes how the real instrument scans the portion of the sky instead of a single point. The signal, s_t^{beam}, as seen by the instrument through the beam, $\mathbf{b}$, is then given in terms of convolution operation

$$s_t^{\mathrm{beam}} = \int_{4\pi} \mathbf{b}_t(\hat{n})\mathbf{s}(\hat{n})d\Omega, \quad \mathbf{b}_t(\hat{n}) = \mathbf{R}_t\left(\hat{n}, \hat{n}'\right)\mathbf{b}(\hat{n}') , \tag{2.11}$$

where the integral is taken over the solid angle, Ω, and the beam rotations (due to scanning strategy) are accounted for in terms of a time-dependent rotation matrix, $\mathbf{R}_t$. Operation in eq.(2.11) is linear; thus, it makes sense to define a new matrix, $\mathbf{B}$, so it will act as an operator on a signal, $\mathbf{s}^{\mathrm{beam}} = \mathbf{B}\mathbf{s}$, which will then modify eq.(2.9) as

$$\mathbf{d} = \mathbf{GPBMa} + \mathbf{n} . \tag{2.12}$$

Several optimizations can be introduced to calculate the convolution integral in eq.(2.11). For instance, instead of directly computing it in pixel space, we can transform the convolution into multiplication in harmonic space[4]. The integral is then evaluated via the help of reduced Wigner matrices (see [159] and [127] for the algorithmic details and its optimization as well as its formulation [55] for BeyondPlanck analysis) that substantially reduces the operational cost by a factor of $O\left(\sqrt{N_{\mathrm{pix}}}\right)$[5] for each beam calculation step. In addition, we may also assume that the beam is azimuthally symmetric, allowing us to expand the beam in terms of Legendre transforms (or *beam transfer function*) b_l. The full

---

[4] This stems from the so-called *convolution theorem* that allows the convolution of two functions in one domain to be expressed as a pointwise product in the other domain.

[5] ***H****ierarchical* ***E****qual* ***A****rea iso-****L****atitude* ***Pix****elization* (HEALPix) scheme is the algorithm to pixelise the sphere in the form of tessellation by partitioning its surface into twelve base pixels which are then recursively divided into four sub-pixels to reach the desired resolution (see Górski at el [61] for more details). The resolution parameter, N_{side}, controlls the total number of pixels of the sky map via the formula

$$N_{\mathrm{pix}} = 12 \times N_{\mathrm{side}}^2 . \tag{2.13}$$

convolution for, e.g., temperature signal, is then written in terms of spherical harmonics s_{lm} as

$$s^{\text{beam}}(\hat{n}) = \sum_{l=0}^{l_{\max}} \sum_{m=-l}^{l} b_l s_{lm} Y_{lm}(\hat{n}) \,, \tag{2.14}$$

where

$$b_l = \exp\left[-\frac{1}{2} l\,(l+1)\,\frac{\sigma^2_{\text{FWHM}}}{8 \ln 2}\right] , \tag{2.15}$$

for a Gaussian beam.

The assumptions of azimuthal symmetry should be treated with caution since the real instrumental beam is usually not symmetric. Luckily, we can distinguish between the symmetric (*main beam*) and asymmetric (*sidelobes*) parts and so compute them separately (due to the convolution being a linear operation). In the case of *Planck*, the beam imperfections were caused by diffraction around primary and secondary mirrors (see figure 2.2), thus, making beam structures highly frequency dependent. The sidelobes, however, have relatively small bandlimits, $l_{\max}$, that allow for their direct computation without too much computational cost.

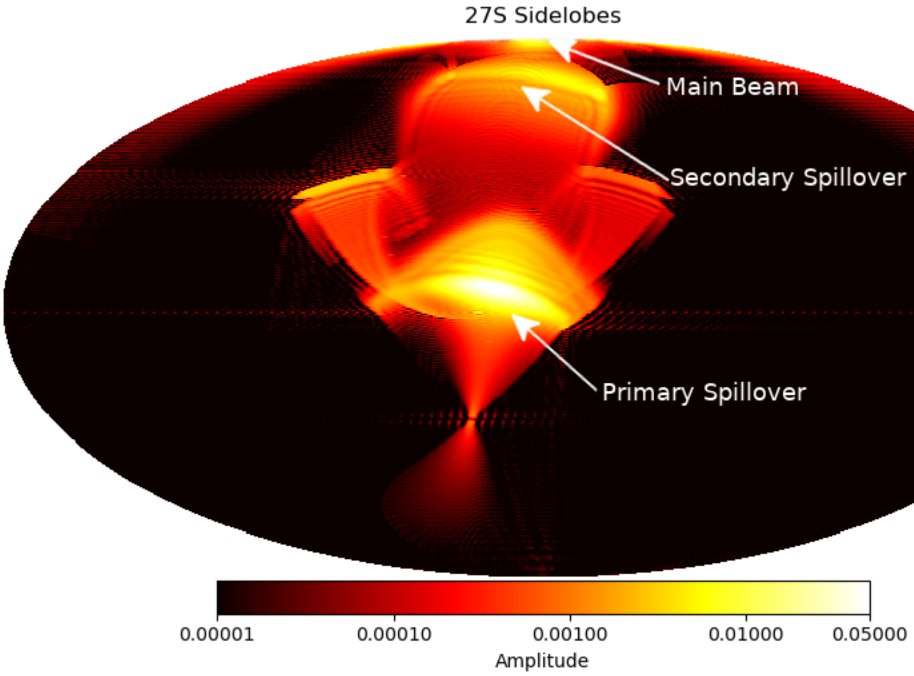

Figure 2.2: Illustration of the beam imperfections of the real instrument detector (30GHz in this case). The sidelobe spillover is caused by diffraction around *Planck* mirrors. Image credit: BEYONDPLANCK collaboration.

Like in the case of bandpass profiles, any given detector possesses a slightly different beam. If we were to combine bandpasses and beams assuming that the underlying signal is the same from any detector, we would get what is called *temperature-to-polarization leakage*, which is intensity signal being incorrectly attributed to the Q- and U-parameters due to instrumental non-idealities. Since I is much stronger than other Stokes parameters, this effect tend to pose a serious challenge unless properly accounted for.

2.2.2.3 Gain

Let us return to the gain, g, in eq.(2.12). *Gain calibration* is a process that requires knowledge of a strong astrophysical signal to calibrate against. For microwave frequencies, the best such signal is the *CMB dipole* that is a result of the Doppler effect due to the Sun's motion (*solar dipole*) with respect to the CMB rest-frame as well as the Earth's orbiting around the Sun (*orbital dipole*). The solar dipole is much stronger but, unfortunately, its parameters are inherently unknown; thus, we will use the orbital dipole (which we can measure precisely because we know Earth's orbital velocity around the Sun very accurately) to calibrate an entire system. These observations, coupled with the Gibbs sampling technique introduced in the next section, allow us to iteratively (re)estimate the Solar dipole within the `Commander` pipeline and use it to measure the relative calibration of various detectors.

2.2.2.4 Noise

A review of instrumental model (2.9) would not be complete without a description of instrumental noise. This can effectively be split into two parts, namely *correlated noise* and *white noise*

$$\mathbf{n} = \mathbf{n}_{\mathrm{corr}} + \mathbf{n}_{\mathrm{wn}}\,. \tag{2.16}$$

which we assume to be Gaussian distributed but with covariance matrices $\mathbf{N}_{\mathrm{corr}} \equiv \left\langle \mathbf{n}_{\mathrm{corr}}\mathbf{n}_{\mathrm{corr}}^{\mathrm{T}} \right\rangle$ and $\mathbf{N}_{\mathrm{wn}} \equiv \left\langle \mathbf{n}_{\mathrm{wn}}\mathbf{n}_{\mathrm{wn}}^{\mathrm{T}} \right\rangle$. The white noise is due to the thermal motions of electrons in the radiometer's electric circuits, while the correlated noise is due to the gain fluctuations which modulate the system's temperature, T_{sys}.

Given the receiver bandwidth, $\Delta\nu$, and the root mean square gain variation, Δg, over the sample of duration, Δt, we can use the *radiometer equation* to write down the expression for a standard deviation of the total noise as

$$\sigma_0 = T_{\mathrm{sys}}\sqrt{\frac{1}{\Delta\nu\Delta t} + \left(\frac{\Delta g}{g}\right)^2}\,. \tag{2.17}$$

The total noise is best described in terms of *power spectral density* (PSD)

$$P(f) = \mathbf{N}_{\mathrm{corr}} + \mathbf{N}_{\mathrm{wn}} = \sigma_0^2\left[1 + \left(\frac{f}{f_k}\right)^{\alpha}\right]\,, \tag{2.18}$$

where f_k is the knee frequency and α defines the slope of the spectrum at low frequencies. It turns out that eq.(2.18) needs one more term to fully describe the noise profiles for 30 and 44GHz detectors. As described in [70], there is a

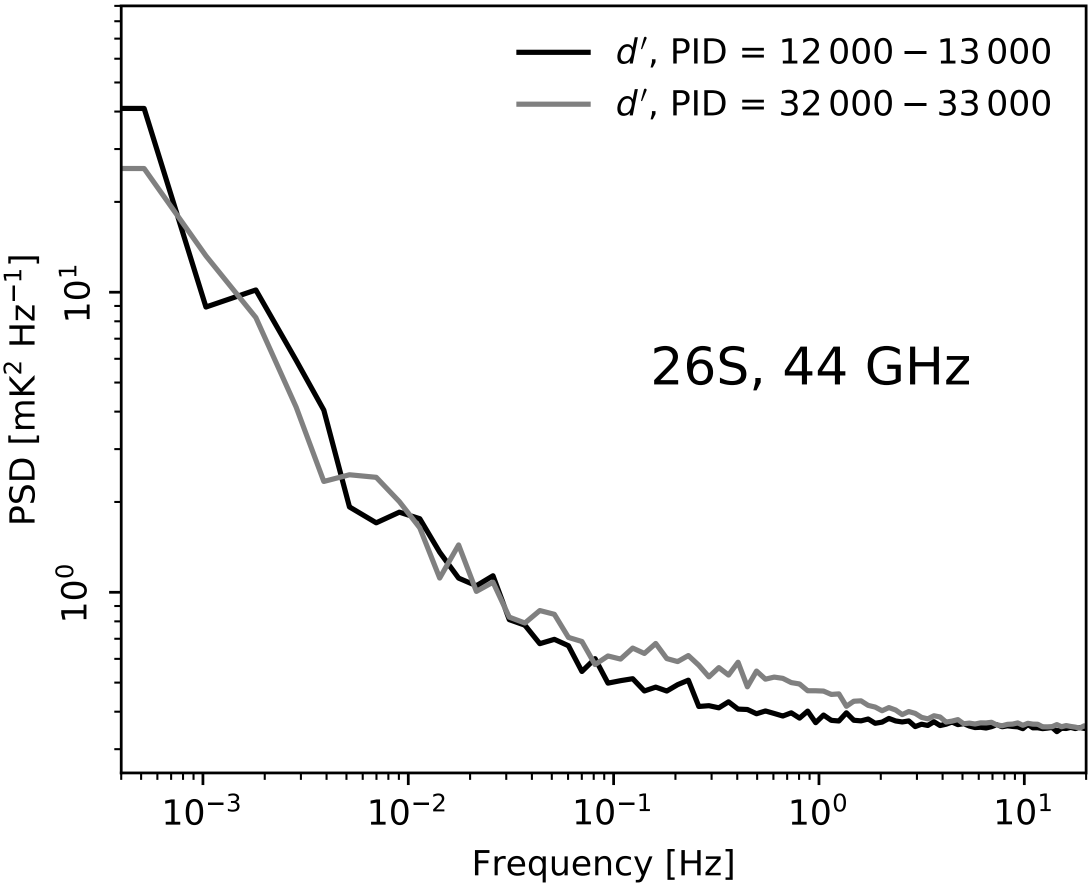

Figure 2.3: Illustration of noise excess in the PSD for 26S radiometer in the frequency range 0.1–10 Hz. The averaging was done over ten PIDs (with 100 PID intervals) for two ranges, namely 12 000–13 000 (black curve) and 32 000–33 000 (grey curve). Image credit: BeyondPlanck collaboration.

significant noise excess for a majority of detectors between frequencies 0.01 and 1Hz, which we model as

$$P(f) = A_p \exp\left[-\frac{1}{2}\left(\frac{\log_{10} f - \log_{10} f_p}{\sigma_{\mathrm{dex}}}\right)^2\right], \tag{2.19}$$

where f_p and A_p are the freely fitted amplitude and peak frequency of this term, respectively. This point is illustrated in figure 2.3.

The combination of eq.(2.18) and eq.(2.19) results in much better χ^2 compared to standard $1/f$ noise profile presented in eq.(2.18). For convenience, we group the spectral noise parameters into one variable $\xi_n = \{\sigma_0, \alpha, f_k\}$, which we are going to sample using Gibbs sampling later on.

2.2.2.5 1Hz Spikes Corrections

A last correction in BeyondPlanck is the 1Hz spike[6] corrections of the time-streams. These are present in the output from all detectors affecting the 44 GHz channel the most. Although the LFI differencing scheme suppressed most of this contaminant, residuals remain in the differenced data.

2.2.2.6 Global Parametric Data Model

In light of everything above, a complete parametric data model (2.12) for a given radiometer, j, signal component, c, time sample, t, and pixel index, p, takes the following form

$$\begin{aligned} d_{j,t} = g_{j,t}\mathbf{P}_{tp,j} &\left[\mathbf{B}^{\mathrm{symm}}_{pp',j}\sum_c \mathbf{M}_{cj}\left(\beta_{p'},\Delta^j_{\mathrm{bp}}\right)a^c_{p'} + \mathbf{B}^{4\pi}_{j,t}\mathbf{s}^{\mathrm{orb}}_j + \mathbf{B}^{\mathrm{asymm}}_{j,t}\mathbf{s}^{\mathrm{fsl}}_t\right] \\ &+ a^{\mathrm{1Hz}}\mathbf{s}^{\mathrm{1Hz}}_j + \mathbf{n}^{\mathrm{corr}}_{j,t} + \mathbf{n}^{\mathrm{wn}}_{j,t}\,. \end{aligned} \tag{2.20}$$

Here, the majority of components are already encountered, i.e., g stands for instrumental gain; $\mathbf{P}$ is the pointing matrix; $\mathbf{a}$ are the amplitudes of astrophysical components defined at the same reference frequencies as mixing matrix $\mathbf{M}$, which is dependent on spectral parameters, β, and the bandpass corrections, Δ_{bp}; $\mathbf{B}^{\mathrm{symm}}$, $\mathbf{B}^{\mathrm{asymm}}$, $\mathbf{B}^{4\pi}$ are the beam matrices for symmetric, asymmetric and the 4π beams, respectively; $\mathbf{s}^{\mathrm{orb}}$, $\mathbf{s}^{\mathrm{fsl}}$, $\mathbf{s}^{\mathrm{1Hz}}$ are the orbital dipole signal, the far sidelobes corrections, and the 1Hz spikes; and, finally, $\mathbf{n}^{\mathrm{corr}}$ together with $\mathbf{n}^{\mathrm{wn}}$ comprise noise contribution described by the combined PSD of eqs.(2.18) and (2.19). Together with white noise (essentially, a stochastic variable of the model), the set of free parameters is comprised of $\{\mathbf{g},\Delta_{\mathrm{bp}},\mathbf{n}_{\mathrm{corr}},\mathbf{a},\beta\}$, while others are either constructed as the deterministic functions of these or constitute an internal part of the original dataset.

2.2.3 Complete Gibbs Chain

In the heart of the analysis lies the Gibbs sampling — a variation of the Metropolis algorithm (see [57] and references therein) that allows subdividing the total joint posterior distribution into subvolumes that are easier to probe, by iteratively sampling each parameter from its respective conditional distribution. Mathematically, if we have a parameter vector ω which we subdivide into N components

$$\omega = (\omega_1, ..., \omega_N)\ , \tag{2.21}$$

---

[6] 1Hz spikes are caused by electronic interference between housekeeping and scientific data.

then each iteration of the Gibbs sampler will cycle through the components of ω and draw the given vector component from the probability distribution

$$\begin{aligned}
\omega_1 &\leftarrow \mathcal{P}\left(\omega_1|\omega_2, ..., \omega_N, d\right) \\
\omega_2 &\leftarrow \mathcal{P}\left(\omega_2|\omega_1, ..., \omega_N, d\right) \\
&\vdots \\
\omega_N &\leftarrow \mathcal{P}\left(\omega_N|\omega_1, ..., \omega_{N-1}, d\right)
\end{aligned} \tag{2.22}$$

conditional on the data, d, while other components remain fixed. If we are on the ith iteration, then the jth component will be updated according to the rule

$$\omega_j^i \leftarrow \mathcal{P}\left(\omega_j|\omega_1^i, ..., \omega_{j-1}^i, \omega_{j+1}^{i-1}, ..., \omega_N^{i-1}, d\right) \tag{2.23}$$

where ω_N^{i-1} implies the Nth component has the value calculated during previous iteration.

Given the joint posterior distribution $\mathcal{P}\left(\mathbf{g}, \mathbf{n}_{\mathrm{corr}}, \xi_n, \Delta_{\mathrm{bp}}, \mathbf{a}, \beta, C_l|\mathbf{d}\right)$ for the data model (2.20) we can rewrite the Gibbs chain (2.22) as

$$\begin{align}
\boldsymbol{g} &\leftarrow \mathcal{P}(\boldsymbol{g} \quad\; |\, \boldsymbol{d}, \qquad\qquad \xi_n, \mathbf{a}^{1\mathrm{Hz}}, \Delta_{\mathrm{bp}}, \mathbf{a}, \beta, C_\ell) \tag{2.24} \\
\boldsymbol{n}_{\mathrm{corr}} &\leftarrow \mathcal{P}(\boldsymbol{n}_{\mathrm{corr}} \,|\, \boldsymbol{d}, \boldsymbol{g}, \qquad\; \xi_n, \mathbf{a}^{1\mathrm{Hz}}, \Delta_{\mathrm{bp}}, \mathbf{a}, \beta, C_\ell) \tag{2.25} \\
\xi_n &\leftarrow \mathcal{P}(\xi_n \quad\; |\, \boldsymbol{d}, \boldsymbol{g}, \boldsymbol{n}_{\mathrm{corr}}, \quad \mathbf{a}^{1\mathrm{Hz}}, \Delta_{\mathrm{bp}}, \mathbf{a}, \beta, C_\ell) \tag{2.26} \\
\mathbf{a}^{1\mathrm{Hz}} &\leftarrow \mathcal{P}(\mathbf{a}^{1\mathrm{Hz}} \,|\, \boldsymbol{d}, \boldsymbol{g}, \boldsymbol{n}_{\mathrm{corr}}, \xi_n, \qquad \Delta_{\mathrm{bp}}, \mathbf{a}, \beta, C_\ell) \tag{2.27} \\
\Delta_{\mathrm{bp}} &\leftarrow \mathcal{P}(\Delta_{\mathrm{bp}} \;|\, \boldsymbol{d}, \boldsymbol{g}, \boldsymbol{n}_{\mathrm{corr}}, \xi_n, \mathbf{a}^{1\mathrm{Hz}}, \qquad \mathbf{a}, \beta, C_\ell) \tag{2.28} \\
\beta &\leftarrow \mathcal{P}(\beta \quad\; |\, \boldsymbol{d}, \boldsymbol{g}, \boldsymbol{n}_{\mathrm{corr}}, \xi_n, \mathbf{a}^{1\mathrm{Hz}}, \Delta_{\mathrm{bp}}, \qquad C_\ell) \tag{2.29} \\
\mathbf{a} &\leftarrow \mathcal{P}(\mathbf{a} \quad\; |\, \boldsymbol{d}, \boldsymbol{g}, \boldsymbol{n}_{\mathrm{corr}}, \xi_n, \mathbf{a}^{1\mathrm{Hz}}, \Delta_{\mathrm{bp}}, \quad \beta, C_\ell) \tag{2.30} \\
C_\ell &\leftarrow \mathcal{P}(C_\ell \quad |\, \boldsymbol{d}, \boldsymbol{g}, \boldsymbol{n}_{\mathrm{corr}}, \xi_n, \mathbf{a}^{1\mathrm{Hz}}, \Delta_{\mathrm{bp}}, \mathbf{a}, \;\beta \quad\;), \tag{2.31}
\end{align}$$

Eqs.(2.24)-(2.31) is a schematic representation of the BEYONDPLANCK approach. It is worth noting that this is not a pure Gibbs sampler. For instance, the spectral parameters β are not conditional on amplitudes $\mathbf{a}$ (see lines five and six), but these are rather drawn from their marginal distribution, and then the amplitudes are drawn conditionally with respect to the spectral parameters. The advantage of this step lies in the shorter Markov correlation lengths compared to standard Gibbs sampling, which means we get convergence faster.

2.2.4 Map Making

Essentially, the *map making* is the linear operation of constructing the unbiased estimator, $\hat{\mathbf{s}}$, over the statistical ensemble of instrumental noise realizations by applying some operator, $\mathbf{W}$, to transform the input time-streams, $\mathbf{d}$, into a map, $\mathbf{m}$, defined in the pixel domain. There are various approaches to the problem (see, e.g., [153] and references therein), but most revolve around two classes, namely *optimal* and *destriping*. All of them assume that the noise is Gaussian

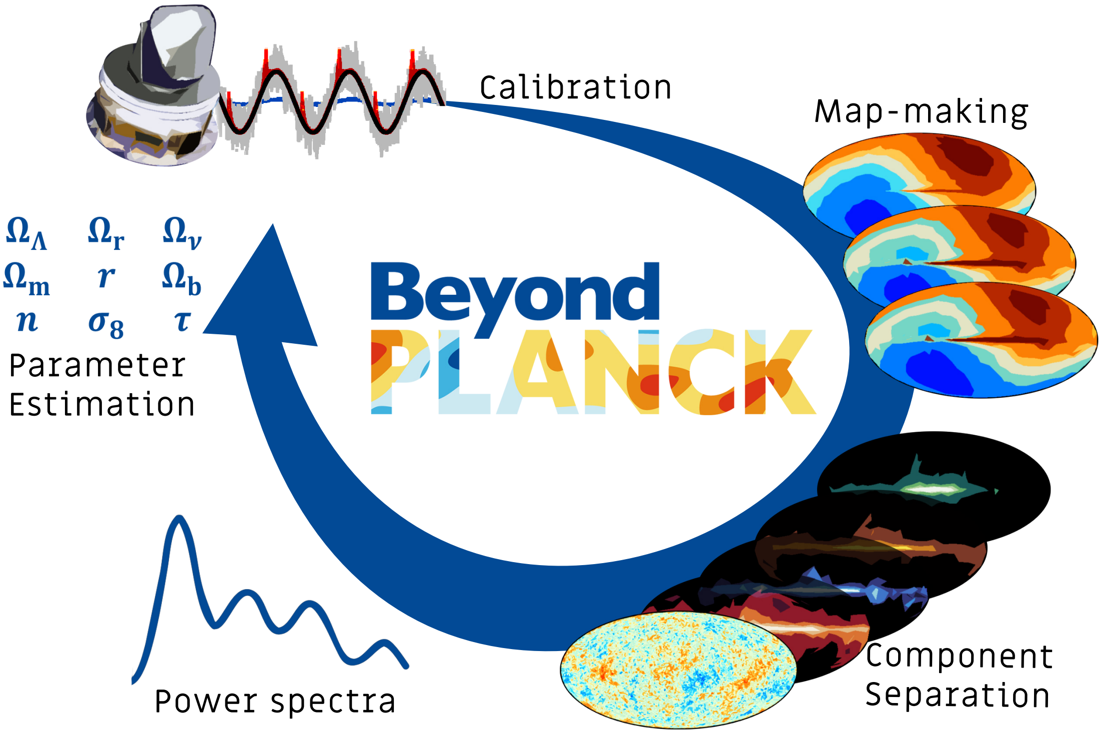

Figure 2.4: Schematic illustration of BeyondPlanck pipeline. Image credit: BeyondPlanck collaboration.

distributed, which typically requires artifacts to be removed from the TOD through preprocessing.

The *Maximum Likelihood* mapmaking (or the *generalized least squares* (GLS)) is based on the idea of applying the maximum likelihood principle to the noise distribution, which results in the set of minimum-variance and unbiased maps. Indeed, let us consider the simplified data model of eq.(2.12). Ignoring the gain, we can rewrite it with respect to the noise vector as

$$\mathbf{n} = \mathbf{d} - \mathbf{P}\mathbf{s}\,. \tag{2.32}$$

Eq.(2.32) is used to construct the *signal likelihood* from the multivariate Gaussian distribution

$$\mathcal{L} = \prod_i \mathcal{P}\left(d_t - P_{tp} s_p | m_p\right) \propto \exp\left[-\frac{1}{2}\left(\mathbf{d} - \mathbf{P}\mathbf{s}\right)^{\mathrm{T}} \mathbf{C}^{-1} \left(\mathbf{d} - \mathbf{P}\mathbf{s}\right)\right], \tag{2.33}$$

which is then minimized with respect to the map

$$\frac{\partial \ln \mathcal{L}}{\partial \mathbf{m}} = 0\,, \tag{2.34}$$

to obtain the *map maker equation*

$$\mathbf{m} = \left(\mathbf{P}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{P}\right)^{-1} \mathbf{P}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{d}\,, \tag{2.35}$$

where $\mathbf{C} = \left\langle \mathbf{n}\mathbf{n}^{\mathrm{T}} \right\rangle$ is called the *noise covariance matrix* and is a dense matrix that may be hard to compute directly for large datasets. In this case, the iterative methods are employed, such as the *Conjugate Gradient Method* (see [141] which makes an initial guess of a solution and then reiterates it until the necessary accuracy is reached.

In the case of classical map making algorithms, eq.(2.35) is solved by Conjugate Gradients which is expensive. However, this burden is alleviated if we use Gibbs sampling, which will break it into two separate steps: (1) the removal of noise and (2) the construction of the sky signal. To understand how it works, let us return to the eq.(2.24) where we will consider the noise component as a sum of two parts, namely the correlated $1/f$ noise, n_{corr}, and the white noise, n_{wn}. The correlated noise is often modeled as a sequence of baselines with a fixed length, N_a, and the amplitudes, $\mathbf{a}$. Therefore, the noise term becomes

$$\mathbf{n} = \mathbf{F}\mathbf{a} + \mathbf{n}_{\mathrm{wn}} \,, \tag{2.36}$$

where $\mathbf{F}$ is the matrix that projects the amplitudes into a full data stream and is equal to the unity matrix since we choose $N_a = 1$ sample for simplicity.

The essence of Gibbs sampling is constructing the joint posterior probability distribution and then splitting it into several parts to ease the computation procedure. Assuming the pixelized map, $\mathbf{m}$, to be statistically independent of $\mathbf{a}$, and that the prior of $\mathbf{m}$ to be uniform, $\mathcal{P}(\mathbf{m}) = 1$, we write the joint posterior distribution as [73]

$$\mathcal{P}\left(\mathbf{m}, \mathbf{a} | \mathbf{d}, \mathbf{C}_{\mathrm{w}}, \mathbf{C}_a\right) \propto \mathcal{P}\left(\mathbf{d} | \mathbf{m}, \mathbf{a}, \mathbf{C}_{\mathrm{w}}\right) \mathcal{P}\left(\mathbf{a} | \mathbf{C}_a\right) \,, \tag{2.37}$$

where $\mathbf{C}_{\mathrm{w}}$ and $\mathbf{C}_a$ are the covariance matrices of $\mathbf{n}_{\mathrm{w}}$ and $\mathbf{a}$, which we assume to know apriori, while the likelihood and prior probabilities are assumed to be drawn from Multivariate Gaussian, i.e.,

$$\begin{aligned} \mathcal{P}\left(\mathbf{d} | \mathbf{m}, \mathbf{a}, \mathbf{C}_{\mathrm{w}}\right) &= \frac{\exp\left[-\frac{1}{2}\left(\mathbf{d} - \mathbf{F}\mathbf{a} - \mathbf{P}\mathbf{m}\right)^{\mathrm{T}} \mathbf{C}_{\mathrm{w}}^{-1} \left(\mathbf{d} - \mathbf{F}\mathbf{a} - \mathbf{P}\mathbf{m}\right)\right]}{\sqrt{|2\pi\mathbf{C}_{\mathrm{w}}|}} \,, \\ \mathcal{P}\left(\mathbf{a} | \mathbf{C}_a\right) &= \frac{\exp\left(-\mathbf{a}^{\mathrm{T}}\mathbf{C}_a^{-1}\mathbf{a}\right)}{\sqrt{|2\pi\mathbf{C}_a|}} \,. \end{aligned} \tag{2.38}$$

Therefore, one full Gibbs cycle would consist of two sampling steps, namely

$$\mathbf{m}' \leftarrow \mathcal{P}\left(\mathbf{m} | \mathbf{a}; \mathbf{d}, \mathbf{C}_{\mathrm{w}}\right), \tag{2.39}$$

$$\mathbf{a}' \leftarrow \mathcal{P}\left(\mathbf{a} | \mathbf{m}; \mathbf{d}, \mathbf{C}_{\mathrm{w}}, \mathbf{C}_a\right) \,, \tag{2.40}$$

where the symbol "←" indicates that the sample is drawn from the distribution on the right-hand side, and the symbol ";" separates parameters (with fixed values) that will not be sampled during the current sampling step. The process depicted in eq.(2.39) is summarized as follows (the actual justification can be found in [73] and appendix of Paper IV):

1. Drawing the sample $\mathbf{m}'$ of sky map, $\mathbf{m}$, from $\mathcal{P}\left(\mathbf{d}|\mathbf{m},\mathbf{a},\mathbf{C}_{\mathrm{w}}\right)$ described in eq.(2.39) is akin to solving

$$\mathbf{m}' = \left(\mathbf{P}^{\mathrm{T}}\mathbf{C}_{\mathrm{w}}^{-1}\mathbf{P}\right)^{-1}\left(\mathbf{P}^{\mathrm{T}}\mathbf{C}_{\mathrm{w}}^{-1}\mathbf{d}' + \mathbf{C}_{\mathrm{w}}^{-1/2}\omega_1\right) , \tag{2.41}$$

   where ω_1 is a random vector obtained from normalised Gaussian distribution and $\mathbf{d}' = \mathbf{d} - \mathbf{Fa}$ represents the current estimate of noise-free TOD.

2. Drawing the sample $\mathbf{a}'$ from joint distribution of eq.(2.40) is equivalent to solving

$$\begin{aligned}\mathbf{a}' &= \left(\mathbf{F}^{\mathrm{T}}\mathbf{C}_{\mathrm{w}}^{-1}\mathbf{F} + \mathbf{C}_{a}^{-1}\right)^{-1}\mathbf{b},\\ \mathbf{b} &= \mathbf{F}^{\mathrm{T}}\mathbf{C}_{\mathrm{w}}^{-1}\mathbf{d} + \mathbf{F}^{\mathrm{T}}\mathbf{C}_{\mathrm{w}}^{-1/2}\omega_2 + \mathbf{C}_{a}^{-1/2}\omega_3 ,\end{aligned} \tag{2.42}$$

   where $(\omega_2, \omega_3) \sim \mathcal{N}\left(\mathbf{0},\mathbf{1}\right)$ (similarly to ω_1) are random vectors whose sizes are equal to the number of elements of their respective covariance matrices.

The steps above with eqs.(2.41)-(2.42) represent the essence of the *Gibbs map maker*, and one of its main strengths lies in error propagation. It can be used both as a traditional GLS estimator that calculates the maximum likelihood function (as explained in [73]) and as a sampling algorithm that draws samples from the underlying posterior distribution, whose mean values are the best estimate of the true sky. In the latter scenario, the residual noise in the map estimate is the samples' scatter. Figure 2.5 illustrates the results of the algorithm from the BeyondPlanck analysis (see [10] for details).

2.2.5 Component Separation

The goal of *component separation* is to estimate a set of physical and/or statistical parameters which provide the best description of a given astrophysical component. This set of parameters is model dependent and, in the most basic scenario, comprises an emission in pixels or a map of the component of interest. Often, this map is obtained via the minimization of total error variance [31]

$$\chi^2 = \sum_p \left|s_p - \hat{s}_p\right|^2 , \tag{2.43}$$

where $\hat{s}$ is the estimated value of a true component emission, s, defined on some space of interest, p, which can be anything, i.e., the set of pixels, the spherical harmonic modes, the Fourier modes, etc. Although the idea behind eq.(2.43) seems trivial, it is quite challenging in practice. Fortunately, throughout the years of data analysis, many algorithms have emerged which can roughly be divided into two groups depending on the final objective.

The first group is concerned only with extracting the CMB signal while treating everything else as unnecessary. The most basic example in this group is called the *Internal Linear Combination* (ILC), and its main idea is the following.

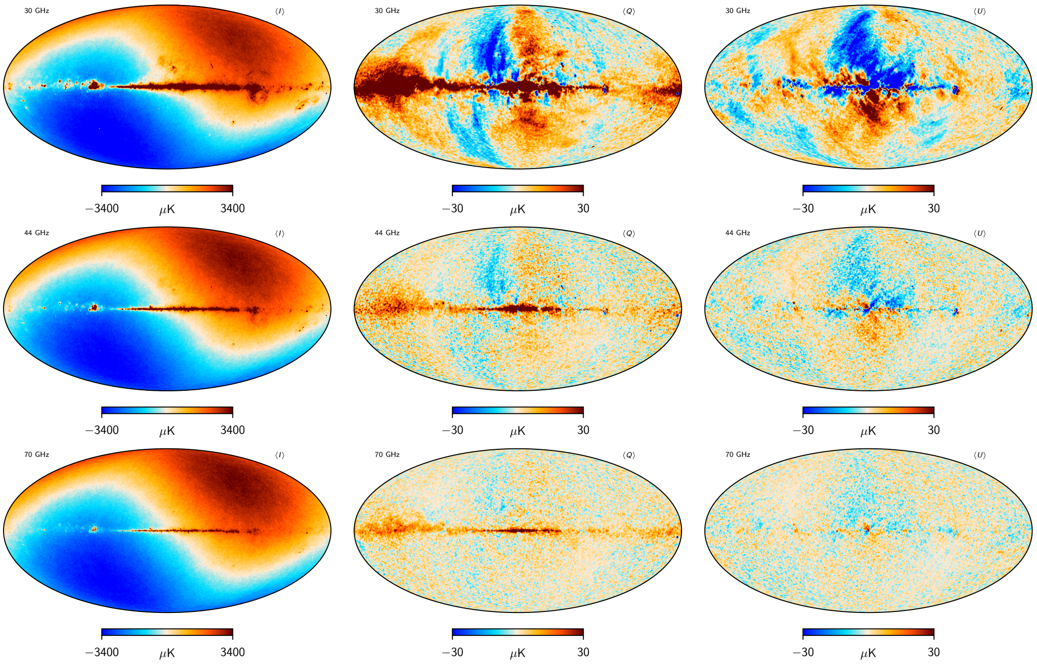

Figure 2.5: Posterior mean maps for the 30 (top row), 44 (middle row), and 70 (bottom row) GHz frequency channels. The columns correspond to Stokes I (left), Q (middle), and U (right). Both 30 and 44 GHz maps have a resolution of $N_{\text{side}} = 512$, while 70GHz is at the $N_{\text{side}} = 1024$. All maps are in Galactic coordinates with polarisation maps smoothed to 1° FWHM angular resolution. Image credit: BeyondPlanck collaboration.

Assuming the template of the component's emission is equal at all frequencies (with observations being calibrated with respect to this component), the TOD vector is modeled as follows [31]

$$d_p^i = s_p + f_p^i + n_p^i \,, \tag{2.44}$$

where for each frequency channel i, we have a foreground, f_p^i, and noise, n_p^i, contributions, respectively, in addition to the CMB component, s_p. The form of eq.(2.44) essentially tells us that any data point is the measurement of signal with some error $f_p^i + n_p^i$; thus, we aim to construct $\hat{s}$ as the averaged over all measurements, each weighted by some weights, w^i, that maximize some criterion of $\hat{s}$ while keeping the component unchanged. This allows us to write

$$\hat{s}_p = \sum_i w_p^i d_p^i, \quad \text{where} \quad \sum_i w_p^i = 1 \,. \tag{2.45}$$

In its simplest possible form, ILC assumes that (1) the weights are independent of p, (2) the CMB signal s is uncorrelated with noise and other components. Hence, eq.(2.45) becomes

$$\hat{s}_p = s_p + \sum_i w^i f_p^i + \sum_i w^i n_p^i \,, \tag{2.46}$$

and the task is to minimise the variance, σ^2, of $\hat{s}$ given by

$$\sigma^2 = \sum_{l\geq 1} \sum_{m=-l}^{l} \mathbf{w}^{\dagger}\mathbf{C}\left(l,m\right)\mathbf{w} = \mathbf{w}^{\dagger}\mathbf{C}\mathbf{w}\,, \tag{2.47}$$

where $\mathbf{C}$ as usual is the signal covariance matrix ($\mathbf{C}\left(l,m\right) = \left\langle \mathbf{d}\mathbf{d}^{\dagger}\right\rangle$, $\mathbf{d} = \mathbf{d}(l,m)$) summed over all modes (l,m). The minimization is done with the help of the Lagrangian multiplier method, which implies solving the following system of equations[7]

$$\frac{\partial}{\partial w^i}\left[\sigma^2 + \lambda\left(1 - \sum_i w^i\right)\right] = 0, \quad \text{where} \quad \sum_i w^i = 0\,, \tag{2.48}$$

with the solution of the form

$$w^i = \frac{\sum_j \left[C^{-1}\right]_{ij}}{\sum_{ij}\left[C^{-1}\right]_{ij}}\,. \tag{2.49}$$

Although ILC is simple and elegant, it has a notable disadvantage associated with the total variance of the resulting ILC map. Indeed, since it is determined by the weights, w^i, which are strongly constrained by the foreground components, the ILC method will work well for the center of the map (the galactic plane). But the high galactic latitudes (low-noise) regions, which are of interest for CMB studies, may be ignored.

Still, ILC proved quite popular among different groups; thus, several variations of this method appeared over the years. For example, the *WMAP* team applied the pixel-based version (instead of the harmonic one) of the ILC algorithm to its data under the constraints that the cosmological component remains the same [11]. Another example applied to WMAP data is the *Needlet ILC* (NILC) [33] variance, which, as the name suggests, operates in the wavelet domain where the input maps are decomposed into needlets in various angular scales. The variance is then minimized on these scales to estimate a set of weights that can vary across the sky. These weights are later combined to produce the CMB map. Other variations of ILC include (but are not limited to) the *Modified ILC Algorithm* (MILCA) [68] and the *Spin-SILC* [131].

The ILC and its variants are not the only algorithms in the first group. Indeed, the template removal methods such as *Spectral Estimation Via Expectation Maximisation* (SEVEM) [94] or *Spectral Matching Independent Component Analysis* (SMICA) [32] were also applied extensively in *Planck* (see, e.g., [78] and [110]). The exact treatment of these algorithms is beyond the scope of this work, and I strongly encourage the reader to look for appropriate reviews such as [31] and others. What I want to concentrate on from now, however, is the second — the *parametric* – group, to which belongs the *Gibbs sampling component*

---

[7] Please note that the notation where i appears as superscript has nothing to do with tensors. I write it in this way simply out of personal convenience.

separation method first described in [43] and [44]. It was used extensively by *Planck* [115], and it now forms a pivotal point in the BeyondPlanck approach. Below I present the general ideas behind the algorithm. For a complete treatment (including all justifications and algorithm improvements), I refer the interested reader to [6] and Paper IV.

Given the pixelized map, m_ν, where all time-dependent components (such as far sidelobe, orbital dipole, correlated noise, etc.) were explicitly subtracted from the TODs before the map making procedure, we can write the data model of eq.(2.20) as

$$m_{\nu,p} = g_\nu B^{\mathrm{symm}}_{pp',\nu} \sum_c M_{c\nu}\left(\beta_{p'}, \Delta^j_{\mathrm{bp}}\right) a^c_{p'} + n^{\mathrm{w}}_{\nu,p}\,, \tag{2.50}$$

which can be rewritten more compactly as

$$\mathbf{m}_\nu = \mathbf{A}_\nu(\beta)\mathbf{a} + \mathbf{n}^{\mathrm{w}}_\nu\,, \tag{2.51}$$

Instead of using the complete Gibbs chain summarized in eqs.(2.24)-(2.31), we are going to jointly sample only the component amplitudes $\mathbf{a}$ and spectral parameters β via the conditional distribution

$$\mathcal{P}\left(\mathbf{a}, \beta|\mathbf{m}_\nu\right) = \mathcal{P}\left(\beta|\mathbf{m}_\nu\right)\mathcal{P}\left(\mathbf{a}|\mathbf{m}_\nu, \beta\right)\,, \tag{2.52}$$

This will lead us to a shorter correlation length than the complete Gibbs sampler. The eqs.(2.29)-(2.30) for β and $\mathbf{a}$ are then transformed into

$$\beta \leftarrow \mathcal{P}(\beta\,|\,\mathbf{m}_\nu)\,, \tag{2.53}$$
$$\mathbf{a} \leftarrow \mathcal{P}(\mathbf{a}\,|\,\mathbf{m}_\nu, \beta)\,. \tag{2.54}$$

Such two-step sampling approach is described in detail in [147] and results in the log-posterior for β of the form (written in our notation)

$$-2\ln\mathcal{P}\left(\beta|\mathbf{m}_\nu\right) = \sum_\nu \left(\mathbf{A}^{\mathrm{T}}_\nu\mathbf{N}^{-1}_\nu\mathbf{m}_\nu\right)^{\mathrm{T}}\left(\mathbf{A}^{\mathrm{T}}_\nu\mathbf{N}^{-1}_\nu\mathbf{A}_\nu\right)^{-1}\left(\mathbf{A}^{\mathrm{T}}_\nu\mathbf{N}^{-1}_\nu\mathbf{m}_\nu\right)\,. \tag{2.55}$$

Given this equation, the sampling for spectral parameters can be done via a tuned Metropolis algorithm, as described in [151] and [6]. However, it is important to note that the expression in eq.(2.55) can only be used for intensity sampling since its derivation relies heavily on the assumption that beam responses are identical for any given frequency channel. For polarization, the likelihood is defined in terms of map residual from eq.(2.51), i.e.,

$$-2\ln\mathcal{P}\left(\beta|\mathbf{m}_\nu, \mathbf{a}\right) = \sum_\nu \left(\frac{\mathbf{m}_\nu - \mathbf{A}_\nu(\beta)\mathbf{a}}{\sigma_\nu(p)}\right)^2\,, \tag{2.56}$$

here $\sigma_\nu(p)$ is the *std* map for detector frequency ν.

After finding β, it is time to turn to the second component in eq.(2.53). Fortunately, due to the assumptions of noise being Gaussian white noise in map eq.(2.51) and that $\sum_\nu \mathbf{A}_\nu(\beta)$ is the deterministic linear operator, we write the sampling equation for amplitudes as a Multivariate Gaussian (see Paper IV, Appendix 1, eq.(A.11))

$$\left(\mathbf{S}^{-1}+\sum_\nu \mathbf{A}_\nu^{\mathrm{T}}\mathbf{N}_\nu^{-1}\mathbf{A}_\nu\right)=\mathbf{S}^{-1}\mu+\mathbf{S}^{-1/2}\eta_0+\sum_\nu \mathbf{A}_\nu^{\mathrm{T}}\mathbf{N}_\nu^{-1}\mathbf{m}_\nu+\sum_\nu \mathbf{A}_\nu^{\mathrm{T}}\mathbf{N}_\nu^{-1/2}\eta_\nu\,, \tag{2.57}$$

where the prior is simply the signal covariance matrix, $\mathbf{S}=\mathbf{S}(C_l)$, whose inverse can be put to zero if needed. The elements η_ν, η_0, and μ are (once again) the random vectors drawn from normalized Gaussian distribution.

Eq.(2.57) is one of the most challenging steps to handle from a numerical point of view. A novel algorithm was developed and described in [140], whose main idea was based on the usage of the CG method together with a pseudo-inverse preconditioner which allowed the use of high-resolution instrumental beams. This technique formed a foundation of the `Commander2` code that was later transformed into `Commander3` and used by the BeyondPlanck collaboration to process *Planck* LFI data. Results from the BeyondPlanck component separation are depicted in figure 2.6 (for more details, see [6] and Paper IV).

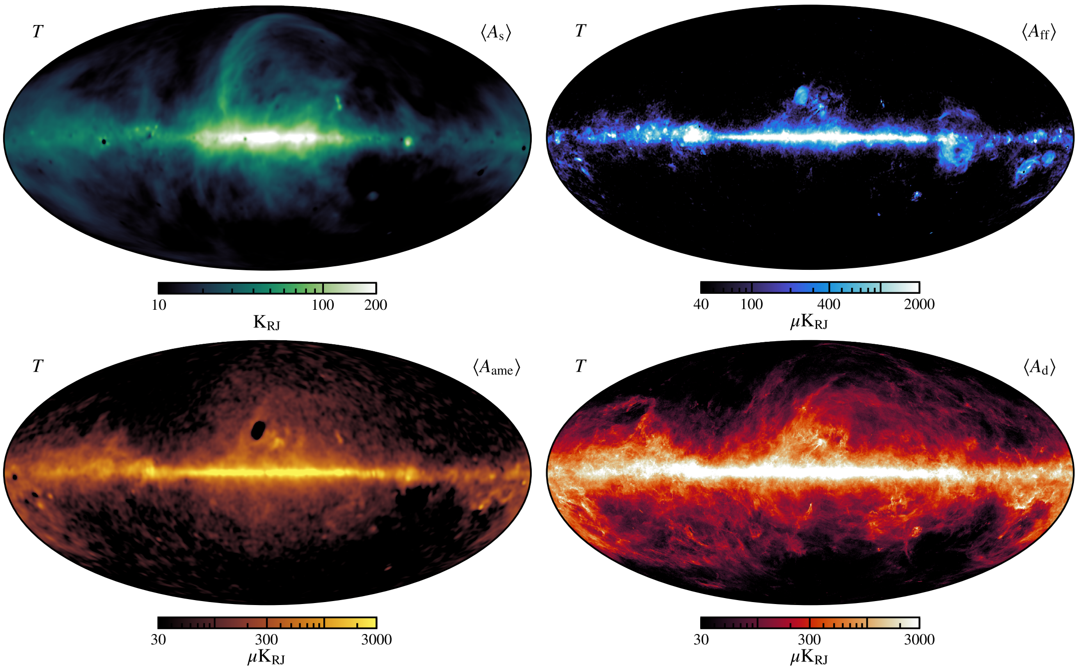

Figure 2.6: Synchrotron (top left), free-free (top right), AME (bottom left), and thermal dust (bottom right) posterior mean amplitude maps at 120, 30, 120, and 10′ FWHM angular resolutions. Image credit: BeyondPlanck collaboration.

2.2.6 Power Spectrum and Parameter Estimation

The two last steps in any CMB processing pipeline are the angular power spectra estimation and the calculation of cosmological parameters from it. In order to obtain the former, BeyondPlanck employed the *Blackwell-Rao* (BR) estimation

[23, 116, 135] for temperature (intermediate angular scales) and the standard *brute-force likelihood evaluation* [102, 123] for polarization (large angular scales). There is, however, a substantial improvement based on the *Karhunen-Loéve* (or *signal-to-noise eigenmode*) *compression* [59, 154] that allows trimming the number of Gibbs samples needed for convergence by scaling down the covariance matrix dimensionality. In addition, we are now sampling the CMB sky maps from the frequency maps (foreground-subtracted) with a well-defined noise term, which allows us to use the BR estimator for l up to 600 (for comparison, *Planck* was able to do this for $l \leq 200$ [122]. Indeed, let us start by writing the posterior in the form of a Multivariate Gaussian

$$\mathcal{P}\left(C_l|\hat{\mathbf{s}}_{\mathrm{CMB}}\right) \propto \frac{\exp\left[-\frac{1}{2}\hat{\mathbf{s}}_{\mathrm{CMB}}^{\mathrm{T}}\left(\mathbf{S}(C_l)+\mathbf{N}\right)^{-1}\hat{\mathbf{s}}_{\mathrm{CMB}}\right]}{\sqrt{|\mathbf{S}(C_l)+\mathbf{N}|}}\,, \tag{2.58}$$

with $\hat{\mathbf{s}}_{\mathrm{CMB}}$ being the CMB+noise map while $\mathbf{S}(C_l)$ and $\mathbf{N}$ are the covariance matrices of the total-effective signal and noise, respectively. Eq.(2.48) introduces nothing novel to the field since it had already been used at the time of *COBE*-DMR [62]. The originality of our approach, however, stems from the fact that we evaluate both $\hat{\mathbf{s}}_{\mathrm{CMB}}$ and $\mathbf{N}$ by averaging over all Gibbs samples, i.e.,

$$\hat{\mathbf{s}}_{\mathrm{CMB}} = \left\langle \hat{\mathbf{s}}_{\mathrm{CMB}}^{i} \right\rangle, \quad \mathbf{N} = \left\langle \left(\mathbf{s}_{\mathrm{CMB}}^{i} - \mathbf{s}_{\mathrm{CMB}}\right)\left(\mathbf{s}_{\mathrm{CMB}}^{i} - \mathbf{s}_{\mathrm{CMB}}\right)^{\mathrm{T}} \right\rangle, \tag{2.59}$$

where i is the sample number. Since averaging is done via all Monte Carlo samples, there is no need to independently account for each source of systematics in the covariance matrix. As an example, figure 2.7 (reproduced from Colombo at el. [28] shows the resulting noise matrices.

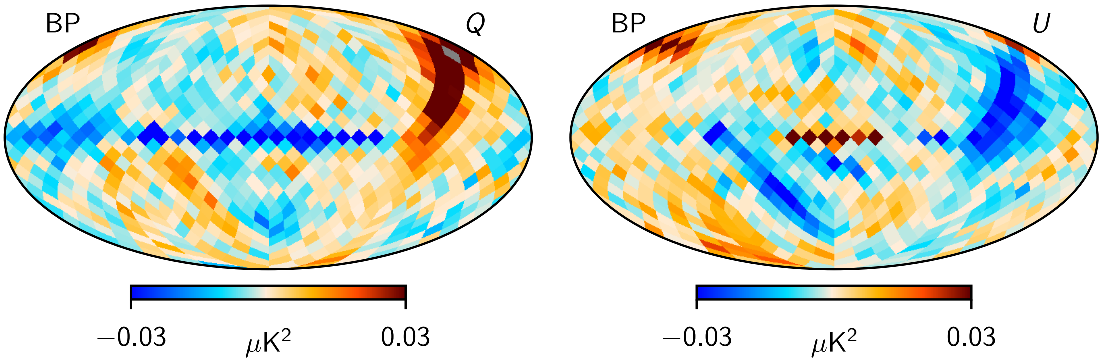

Figure 2.7: The low resolution ($N_{\mathrm{nside}} = 8$) noise covariance matrix of CMB Q and U Stokes parameters constructed by BEYONDPLANCK without any filters. Image credit: BEYONDPLANCK collaboration.

Since the estimator above strongly depends on the size of the covariance matrix, it cannot be used for high angular resolution analysis. Therefore, we employ the *Gaussianized Blackwell-Rao* estimator [135] to achieve faster convergence for $l \lesssim 700$. This requires first evaluating the univariate likelihood

Table 2.1: Constraints on the 6 ΛCDM base parameters with confidence intervals at 68% from CMB data alone and adding lensing + BAO.

Parameter	BeyondPlanck	BeyondPlanck + *Planck*	BeyondPlanck + *Planck* + Lensing + BAO
$\Omega_{\mathrm{b}}h^2$	$0.0228^{+0.0011}_{-0.0012}$	0.02224 ± 0.00022	0.02239 ± 0.00020
$\Omega_{\mathrm{c}}h^2$	$0.130^{+0.019}_{-0.028}$	0.1218 ± 0.0021	0.1189 ± 0.0011
$100\theta_{\mathrm{MC}}$	$1.043^{+0.006}_{-0.008}$	1.0406 ± 0.0005	1.0410 ± 0.0004
τ	0.065 ± 0.012	0.070 ± 0.012	0.070 ± 0.010
$\ln(10^{10}A_{\mathrm{s}})$	$3.10^{+0.10}_{-0.11}$	3.078 ± 0.022	3.071 ± 0.018
n_{s}	$0.973^{+0.021}_{-0.029}$	0.961 ± 0.006	0.967 ± 0.004
Ω_{Λ}	$0.63^{+0.14}_{-0.08}$	0.673 ± 0.014	0.691 ± 0.007
t_0	$13.7^{+0.3}_{-0.2}$	13.83 ± 0.04	13.79 ± 0.03
Ω_{m}	$0.37^{+0.08}_{-0.14}$	0.327 ± 0.014	0.309 ± 0.007
σ_8	$0.87^{+0.12}_{-0.14}$	0.830 ± 0.010	0.819 ± 0.007
z_{re}	8.8 ± 1.2	9.2 ± 1.1	9.2 ± 0.9
H_0	$65.9^{+4.3}_{-5.6}$	66.6 ± 0.9	67.8 ± 0.5
$10^9 A_{\mathrm{s}}e^{-2\tau}$	$1.96^{+0.17}_{-0.22}$	1.888 ± 0.010	1.875 ± 0.006

separately for any given multipole

$$\mathcal{P}\left(C_l|\mathbf{s}^{\mathrm{CMB}}\right) = \sum_{i=1}^{n_{\mathrm{samp}}} \frac{\exp\left(-\frac{2l+1}{2}\frac{\sigma_l^i}{C_l}\right)}{|C_l|^{(2l+1)/2}}, \quad \sigma_l^i \equiv \sum_m \frac{\left|s_{lm}^i\right|^2}{2l+1}\,, \tag{2.60}$$

with the subsequent evaluation of the final likelihood

$$\mathcal{P}\left(C_l|\mathbf{d}\right) \approx \left(\Pi_l \frac{\partial C_l}{\partial x_l}\right)^{-1} \exp\left[-\frac{1}{2}\left(x-\mu\right)^{\mathrm{T}} \mathbf{C}^{-1}\left(x-\mu\right)\right]\,, \tag{2.61}$$

where the covariance matrix $\mathbf{C}_{ll'} = \langle (x_l - \mu_l)(x_{l'} - \mu_{l'}) \rangle$ is given in terms of the transformed input power spectrum, $\mathbf{x} = \{x_l(C_l)\}$ and estimated from the averaging over Gibbs samples. The $\partial C_l/\partial x_l$ is the Jacobian. It is important to note that the approach outlined here was used only for the temperature case, and the complete treatment (which includes polarization) can be found in Paper IV of this thesis.

The likelihoods described above are used as inputs to *Cosmological Monte Carlo* (CosmoMC) [83] code to sample cosmological parameters, and results are presented in table 2.1. To be consistent with the one reproduced in chapter 1 from *Planck* analysis, both free and derived parameters are depicted with the inclusion of constraints from BAO and CMB lensing.

Chapter 3

Open Science in the CMB Field

Usually, reproducibility in science (generally) and in the CMB field (particularly) is defined and conducted as a competition between different groups, each trying to measure the same sky signal using various statistical techniques and numerical codes. Although this approach works for classical experiments, there are some issues[1] it is unable to resolve. Furthermore, due to the intrinsic relationship between astrophysical components and instrument calibration, one simply cannot confine oneself to a specific experiment but must employ the joint analysis ideas described in the previous chapter. Indeed, as demonstrated in Paper IV and the suite of companion papers (summarized in table 3.1), joint Bayesian analysis is a potent tool to explore posterior distributions, quantify uncertainties in the data and, consequently, constrain the sky. Therefore, it makes sense to expand the scope of BeyondPlanck to a broader range of experiments such as *WMAP*, *SPIDER*, *QUIET*, *LiteBIRD*, etc. This is a primary motivation for the Cosmoglobe project, whose key goals can be summarized as follows:

- Integrate the best data from a range of CMB experiments starting with radio and finishing with sub-mm wavelengths using a single model global analysis.
- Employ the flexible and well-tested Bayesian parameter estimation technique developed by *Planck* and refined in BeyondPlanck.
- Have the ability to use data in terms of sky maps in addition to raw TODs.
- Build a community around it with a strong emphasis on Open Science and reproducibility.

Clearly, such ambitious goals require strong support from the community; thus, reproducibility and Open Science will play a pivotal role in this endeavor. The purpose of this chapter is to describe our approach to these topics in BeyondPlanck and, subsequently, Cosmoglobe.

3.1 Reproducibility in BeyondPlanck

In the center of BeyondPlanck was `Commander3` — the `Fortran`-based code of about 60 000 lines, designed to run specifically on HPC systems. It supports

---

[1] For example, incorrect gain measurements can lead to fake polarization signals when differencing two detectors. This is prominent in *Planck* where the incorrect gain led to large-scale polarization signals — so-called *poorly measured modes*. A similar effect can be observed in *WMAP* where incorrectly characterized pickup between the two horns created a large-scale polarization signal.

Table 3.1: Overview of BeyondPlanck and preliminary Cosmoglobe papers.

Reference	Title
Pipeline	
[14]	I. Global Bayesian analysis of the *Planck* Low Frequency Instrument data
[73]	II. CMB mapmaking through Gibbs sampling
[54]	III. `Commander3`
[19]	IV. On end-to-end simulations in CMB analysis — Bayesian versus frequentist statistics
Instrument characterization	
[65]	V. Minimal ADC Corrections for Planck LFI
[70]	VI. Noise characterization and modelling
[60]	VII. Bayesian estimation of gain and absolute calibration for CMB experiments
[55]	VIII. Efficient Sidelobe Convolution and Correction through Spin Harmonics
[150] . . .	IX. Bandpass and beam leakage corrections
Cosmological and astrophysical results	
[10]	X. *Planck* LFI frequency maps with sample-based error propagation
[28]	XI. Bayesian CMB analysis with sample-based end-to-end error propagation
[103] . . .	XII. Cosmological parameter estimation with end-to-end error propagation
[6]	XIII. Intensity foregrounds, degeneracies and priors
[151] . . .	XIV. Polarized foreground emission between 30 and 70 GHz
[66]	XV. Limits on Large-Scale Polarized Anomalous Microwave Emission from *Planck* LFI and *WMAP*
Cosmoglobe	
[58]	From BeyondPlanck to Cosmoglobe: Open Science, reproducibility, and data longevity
[160] . . .	From BeyondPlanck to Cosmoglobe: Preliminary WMAP Q-band analysis

MPI parallelization, in-memory TOD compression, efficient and fast I/O, FFT, and Spherical Harmonic operations. It is beyond the scope of this work to go through the exact implementation details thoroughly described in [54]. However, to give an idea of the actual speed we were able to achieve, I reproduced table 3.2 from [54] with CPU hours spent on each step of the BeyondPlanck LFI analysis.

If we are going to proceed with the goals of Cosmoglobe, `Commander3` will be a vital part of the analysis; thus, the community should be able to install and run it without investing too much of their time learning or dealing with annoying bugs (which, naturally, will occur due to the speed of `Commander` development). These considerations allow us to quantify the requirements for reproducibility in Cosmoglobe. Eventually, the end user should be able to: (1) fetch and install the code with ease; (2) acquire all data streams and other auxiliary files; (3) have access to documentation; (4) contribute to Cosmoglobe via, e.g., extending `Commander` by introducing new processing module for complementary data set.

Every point is associated with a set of unique challenges. For instance, to install the code, one should be proficient in understanding different `Fortran` compilers, know how to compile third-party libraries `Commander3` is dependent on, etc. Furthermore, the lack of a common industry standard in addressing each of these points adds up to the complexity of the problem. Therefore, it is important to understand that there is no "right" way of dealing with such issues.

Table 3.2: Computational resources required for end-to-end BEYONDPLANCK processing. All times correspond to CPU hours. All reported times are averaged over more than 100 samples, and vary by $\lesssim 5\,\%$ from sample to sample. Reproduced from [54].

ITEM	30 GHz	44 GHz	70 GHz	SUM	REFERENCE
Data volume					
Uncompressed TOD volume	761 GB	1 633 GB	5 522 GB	7 915 GB	
Compressed TOD volume	86 GB	178 GB	597 GB	861 GB	
Non-TOD-related RAM usage				659 GB	
Total RAM requirements				**1 520 GB**	
Processing time (cost per run)					
TOD initialization/IO time	3.8 h	4.3 h	12.5 h	20.6 h	
Other initialization				43.4 h	
Total initialization				**64.0 h**	
Gibbs sampling steps (cost per sample)					
Huffman decompression	1.1 h	1.8 h	7.1 h	10.0 h	[54]
TOD projection (P operation)	0.3 h	0.7 h	3.1 h	4.1 h	[14]
Sidelobe evaluation ($\boldsymbol{s}_{\mathrm{sl}}$)	1.1 h	2.1 h	6.5 h	9.7 h	[55]
Orbital dipole ($\boldsymbol{s}_{\mathrm{orb}}$)	0.5 h	1.1 h	4.6 h	6.2 h	[60]
Gain sampling (g)	0.6 h	0.7 h	4.7 h	6.0 h	[60]
1 Hz spike sampling ($s_{1\mathrm{hz}}$)	0.2 h	0.3 h	1.9 h	2.4 h	[14]
Correlated noise sampling ($\boldsymbol{n}_{\mathrm{corr}}$)	1.7 h	3.6 h	24.8 h	30.1 h	[70]
Correlated noise PSD sampling (ξ_{n})	3.3 h	4.0 h	1.1 h	8.4 h	[70]
TOD binning (P^t operation)	0.2 h	0.5 h	4.1 h	4.8 h	[10]
Sum of other TOD processing	1.3 h	2.5 h	10.9 h	14.7 h	[54]
TOD processing cost per sample	**10.4 h**	**17.4 h**	**69.1 h**	**96.9 h**	
Amplitude sampling, $P(\mathbf{a} \mid \boldsymbol{d}, \omega\backslash\mathbf{a})$				23.9 h	[6]
Spectral index sampling, $P(\beta \mid \boldsymbol{d}, \omega\backslash\beta)$				40.3 h	[151]
Other steps				0.6 h	[14]
Total cost per sample				**163.9 h**	

So, even though based on the facts and logical considerations, every statement below may still be subjective and, perhaps, will be. Since the majority of HPC systems are run on Linux[2], the starting point of our discussion should be Linux kernel.

3.1.1 Containers

Started as an innocent hobby project, the Linux kernel and the GNU/Linux Operating System has grown into something extremely big and indispensable. Almost all OS used in HPC systems uses Linux kernel — the brain of the system that allocates and deallocates the available physical resources to processes that run on the system. Obviously, the entire system design and operation are beyond the scope of this work. Instead, I want to concentrate on the Linux kernel's particular features, namely *namespaces* and *control groups* (`cgroups`).

---

[2] The top500.org gives an overview of the most powerful HPC systems on the planet. The share of Linux OS variants is more than 70% as of November 2022.

CGroups are the part of the Linux kernel that effectively manages, tracks, and accounts for system resources (such as CPU, memory, I/O, network, etc.) used for a specific application via hierarchical partitioning of groups of processes. Essentially, `cgroups` allow users (instead of the kernel) to limit the system's resource consumption by a specific process, its child, or a group of processes.

Namespaces encapsulate important global system resources to hide them from unwanted processes. All processes in Linux kernel have their roots in `init`[3]; some of them are more "privileged" than others, meaning that they have rights to locate and "kill" other processes. Sometimes such behavior is undesirable, so the namespaces were invented to split the common processes' tree into multiple subtrees, each having its root. This provides a necessary layer of abstraction to hide the system processes from its children. In other words, by belonging to a new namespace, the child process becomes the root, making its parent invisible to the child (since the parent is part of another namespace). However, the parent can still see the child's process, even though it belongs to a different namespace. In order to track it, the child gets *local* Process IDentifier (`PID`) within its new namespace, in addition to the unique *global* one (assigned to track the child in a global namespace). `PID` is one of the system namespaces available. Others include[4] mount (`mnt`), UNIX Time-sharing System (`uts`), Interprocess Communication (`ipc`), Process IDentifier (`PID`), Network (`net`), User ID (`user`), and Control Group[5] (`cgroup`). Together with `cgroups` briefly summarized above, namespaces form a core functionality of what is known as containers.

Containers use an underlying (typically, Linux) kernel to provide isolated environments to run a specific (group of the) process(es). In this way, the virtual environments are run on top of the host kernel (instead of the hypervisor) and include only the application together with the necessary libraries (instead of the entire OS), which effectively reduces resource consumption. The isolation is provided by `namespaces` (meaning that each container is run in and has access to only its dedicated `namespace`), while actual resource allocation is performed via `cgroups`.

3.1.1.1 Docker and Singularity

Even though it is possible to run containers yourself, the task is quite tedious; thus, users typically use various container management systems to build, deploy

---

[3] The `init` process is the first one to start (so-called `PID1`) by the kernel and is de facto the progenitor of all other processes on the system. Once the "On" button is pushed, and the system starts loading, the kernel is looking for `init` and, if not found, will execute the `panic()` method. For those who are interested, the source code of the method is inside `init/main.c:kernel_init`.

[4] Since the work on Linux kernel is ongoing, the actual number is a subject to change.

[5] It is important to distinguish between the `cgroups` (a kernel feature that limits system resource allocation) and the `cgroup namespace` (which isolates the root directory of `cgroups` so it becomes invisible to the unwanted processes).

and run containers easily. One of the most popular ones is called *Docker* – the tool to manage containers efficiently. The technology is not new[6], and throughout the years, it has evolved into an entire *Platform as a Service* (PaaS) instance.

Due to its simplicity of use, it was tempting to ship `Commander` in the form of a Docker container, which would effectively make it the main system in terms of reproducibility. However, this approach had several downsides, the first and foremost being computation speed. Even though Docker took fewer resources than any traditional hypervisor, which made it much faster, the run speed of the codes compiled for a specific CPU was still far superior to the Docker containers. It would be fine to use containers for test runs which took minutes or hours at most. However, the time scale for an entire LFI `Commander` run amounted to more than a month, which made even 5% loss in computation speed a significant flaw.

The second issue was the availability and root access. Indeed, by the time of our considerations, the Docker infrastructure required root privileges[7] in order to be installed and run. Clearly, this was unfeasible for the average HPC user. Luckily, the `Singularity`[8] software existed — the container platform designed specifically to run on HPC systems. It still required administrative access to be installed, but once done, the user could run containers without root access. This limited the availability of the software on many HPC systems worldwide, so we decided to use something else as our primary installation/distribution tool for reproducibility purposes. Although eventually, we put up the `Commander`3 Docker container as described in Paper V.

3.1.2 Tools of the Trade: CMake

Initially, `Commander` relied on Make as its main compilation system. While it has many advantages and desired features, it is not the best tool for compiling complicated codes that depend on many additional third-party libraries. Thus, we needed something better which would comply with the following requirements: (1) free and Open Source; (2) cross-platform; (3) automation; (4) long-term support; (5) multi-language support; (6) minimal dependencies; and (7) Integrated Development Environment (IDE) support.

The considered tools are summarized in Paper V and can roughly be divided into low-level, high-level systems, and package managers. Indeed, Make is a low-level system, while, e.g., `Meson` is a high-level one, which essentially employs Make under the hood to install the code. It is a `Python`-based tool that complies with all requirements above. However, at the time of our consideration, it was not as popular as its counterpart — `CMake` — which made it unclear whether it would survive in the long run.

---

[6] Docker was first released as an Open Source software in March 2013.

[7] The situation has changed with the introduction of Docker Engine v20.10 that introduced so-called *rootless mode* of operation. Read more at https://docs.docker.com/engine/security/rootless/

[8] The name was changed to `Apptainer`. See the official website for more details: http://apptainer.org/

All in all, we landed on `CMake`[9], which is an Open Source family of tools designed for testing, compiling, and packaging software. It has been around since 1999 and is used extensively in astrophysics[10] and beyond[11], making it less likely to be abandoned in the future. In addition, `CMake` supports nested project structures, cross-compilation, multiple platforms, and languages and can incorporate third-party libraries into the build. Among its notable weaknesses are a lack of a common standard of file organization, a steep learning curve for its syntax, and non-trivial documentation that newcomers find challenging to navigate. However, its strengths outweigh its weaknesses, which made us choose it as our primary, high-level compilation system, with Make being the low-level one.

The `CMake-Commander` code organization is depicted in figure 3.1. Instead of mixing `Fortran` and `CMake` code-base, we separated them completely, making the structure more transparent and easier to maintain. Moreover, to avoid the intermixing between source code and `CMake`-produced Make configuration files, we are using the so-called “out-of-source” build approach, which forbids running `CMake` inside the `root` directory and instead requires running it from a dedicated folder (usually named `build`). Thus, to compile `Commander`, one needs to perform the following steps:

```
> git clone https://github.com/Cosmoglobe/Commander.git
> mkdir Commander/build && cd Commander/build
> cmake -DCMAKE_INSTALL_PREFIX=$HOME/local \
        -DCMAKE_C_COMPILER=icc \
        -DCMAKE_CXX_COMPILER=icpc \
        -DCMAKE_Fortran_COMPILER=ifort \
        -DMPI_C_COMPILER=mpiicc \
        -DMPI_CXX_COMPILER=mpiicpc \
        -DMPI_Fortran_COMPILER=mpiifort \
        ..
> cmake --build . --target install -j 8
```

The first two lines download the `Commander3` source code and create the `build` directory to store the configuration files, and the last two are analogous to the infamous `configure`, `make`, and `make install`. In principle, the steps depicted above are the only ones the user will need to type in their keyboard to install `Commander`. Everything else is done automatically, i.e., `CMake` searches for the required dependencies and, if they are missing, it will download, check hashes of the sources, compile and install all libraries and then compile and link them to `Commander` itself. Of course, the entire installation process is system dependent, but it usually takes no more than ten minutes.

In reality, various issues arise that require a certain level of human intervention. In such cases, we started to put up the FAQ section in `Commander` online documentation hosted on GitHub pages[12]. Although it is in the pre-alfa state

---

[9] https://cmake.org/

[10] Examples include OpenBLAS, HDF5, FFTW3, and Time Ordered Astrophysics Scalable Tools (TOAST).

[11] Perhaps, the most notable ones from this category will be Netflix and KDE.

[12] https://cosmoglobe.github.io/Commander/#/

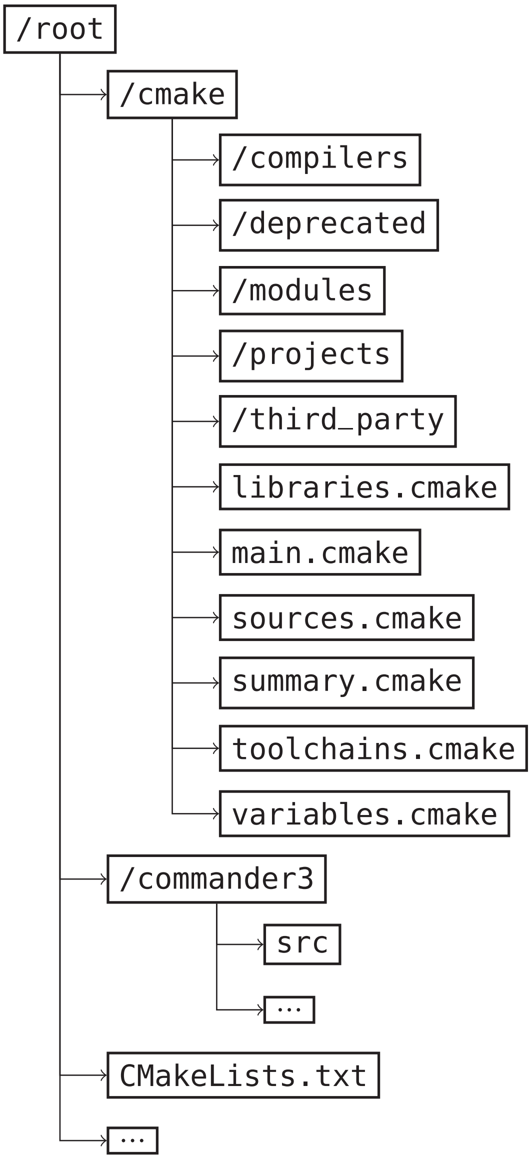

Figure 3.1: Overview of `CMake` infrastructure inside `Commander3` directory (dir). As per convention, the root contains the top `CMakeLists.txt` file with general `CMake-Commander` configuration. But, the main bulk of code is inside `cmake` dir, where the `main.cmake` mimics the logic of a typical compiled language such as, e.g., `C++`, and serves as the "entry point" to the program. Other scripts contain information such as the `Commander3` dependencies' web credentials (`sources.cmake`); custom-defined configuration `CMake`-variables for `Commander3` installation (`variables.cmake`); available compiler toolchains (`toolchains.cmake`), and the particular configuration for it (`compilers` dir); custom and third-party `Find<Name>.cmake` modules (`modules` and `third_party` dirs); general and specific configurations for each subproject (`libraries.cmake` and `projects` dir); the module that outputs installation summary in the terminal screen (`summary.cmake`); and the deprecated codes (`deprecated`).

as of the time of writing this thesis, the documentation contains a lot of information regarding `Commander3` compilation, the parameter file, as well as some information of code extension. In addition, `CMake` also produces the log files of each compilation step, located inside `logs` directory[13].

3.1.3 Input Files, Parameter Files, & Data Products

Once `Commander` is installed, the second step in the process is to obtain data files. This can be done using the `Python`-based download utility we have put up. Since we initially stored all source files in `GitLab`, it is located there instead of `GitHub`. The official documentation[14] extensively explains how to obtain and use it. However, in a nutshell, the user should copy the `bp` binary file to a relevant folder[15] and run it via

```
> bp download auxcmd3 l2data
```

which will start downloading both LFI L2Data and `Commander`3 auxiliary files. If, for some reason, the downloading was interrupted by repeating the same command, it will be resumed since the tool checks for file hashes and so “knows” which were downloaded correctly and which were not. Since the data takes more than 1 TB of disk space, the entire process can take a while but is very straightforward in nature.

Once one obtains the data and installs `Commander3` the next step would be to run it. As with any MPI application, this can be done by, e.g.,

```
> mpirun -n {ncore} {path/to/commander3/binary} {param.txt}
```

where `ncore` is the number of cores for the MPI job, and `param.txt` is the human-readable ASCII text file containing configuration for `Commander`3 run (see [54] and documentation for more details). Since the number of accepted parameters is huge, this file tends to become long and hard to read. Therefore, in the latest incarnation of `Commander`3, we introduced support for nested structures that hide unnecessary parameters from the user. The downside of this approach is that (1) an additional layer of abstraction adds up to the complexity of the `Commander`’s steep learning curve, and (2) a couple of new environment variables need to be defined in the user’s shell file. However, once the parameter file is mastered to a reasonable degree, the entire process of running `Commander`3 is fully automated.

As summarised in table 3.2 and described in [54], it takes approximately two hours to produce a single Gibbs sample on a node with 128 cores, which makes the task of full Monte Carlo posterior exploration to be finished within months. The results of the `Commander`3 operation comprise the main data products of BeyondPlanck, which are stored at the Institute of Theoretical Astrophysics’ local cluster and can be downloaded from the official BeyondPlanck website[16].

---

[13] It prefix was pointing to `$HOME/local` then logs of installation will be located inside `$HOME/local/logs`.

[14] See https://cosmoglobe.github.io/Commander/#/01_quick_start/downloader/index

[15] Usually, user executables are located inside `$HOME/.local/bin`. The necessary permission may be required to run the BeyondPlanck download utility. In this case, the user can use, e.g., the `chmod u+x bp` command in the terminal.

[16] https://beyondplanck.science/products/files_v2/

3.2 Simulations and Error Estimation

Once the pipeline is installed, the natural question is how do we test whether it works the way we expect it to? We could have designed the set of tests to check each module/component of `Commander3`, but it may not be enough when we want a full run with complete error propagation. In this case, it is more beneficial to produce synthetic data (via simulating the entire sky and the instrument) and then employ the `Commander3` pipeline to analyze these.

Of course, the *validation* of data processing techniques is not the only thing the simulations are used for. Indeed, end-to-end simulations are employed to ensure future experiments' success (see, e.g., [87]) by forecasting the performance of a particular experimental design. They are also used for error estimation in traditional CMB pipelines. The primary example here is *Planck* LevelS software [130], which was created to produce *Planck* Full Focal Plane (FFP) simulations but was also designed to be a reusable and efficient piece of software used for other experiments as well. The substantial upgrade of the idea was the development of the Time Ordered Astrophysical Scalable Tools (TOAST) specifically created to be run on HPC systems whose FFP simulations were used in Planck 2016 [117] data analysis as the primary mechanism for error propagation in *Planck* 2016 & 2020 releases [114, 121].

In the heart of the modern end-to-end simulations lies the data model defined by eqs.(2.18), (2.19) & (2.20). By defining a set of free parameters, collectively called ω, one evaluates these equations directly with only stochastic terms being correlated and white noise (usually generated via random Gaussian number generators). The evaluation is done under the assumption that ω is perfectly known, which, in reality, is not true. Consequently, the precise specification of ω directly impacts what the simulations inform the scientist about. This brings us to a vital topic of Bayesian and frequentist types of simulations discussed in Paper VI.

Keeping the above in mind, I have effectively converted `Commander3` into a simulator by writing an additional `Fortran` module. I then ran the code with several people on various HPC systems to generate a set of one-year *Planck* LFI data and then analyze those through the `Commander3` pipeline. The idea behind this was twofold. First, it was a good test of reproducibility and, second, the pipeline validation. The only components in the simulation pipeline were CMB sky (drawn from the *Planck* 2018 best-fit ΛCDM model [123] and produced by `synfast` HEALPix [61] utility) and noise term (with noise PSD of eqs.(2.18) & (2.19)). The results of our work are thoroughly explained in Paper VI, and I refer the interested reader to it.

Summary and Outlook

This thesis introduces my work as part of the PhD program at the Institute of Theoretical Astrophysics at the University of Oslo. It revolved around the Cosmic Microwave Background, the light that comes to us from the Big Bang and delivers important information about Early Universe. For instance, it can help answer a fundamental question of topology, i.e., the global shape of the Universe; the topic studied in Papers II and III. Other important conclusions that can be inferred from it are the cosmological parameters that comprise the ΛCDM model of cosmology. This model can be extended by, e.g., the inclusion of relativistic particles, which is done in Paper I where I have generalized the analytical perturbation theory for a flat Universe.

The CMB is the topic of study by a plethora of ground-based, balloon-bourne, and satellite missions. Among these, the one that stands out the most is the *Planck* mission which covered an entire sky in unprecedented resolution. Even though the experiment is finished and the main products are publicly available, it is of interest to use the raw time streams in combination with other datasets to even better constrain the microwave sky. The goal of the BeyondPlanck project was to develop a joint analysis pipeline, called `Commander3`, a single piece of software that starts the analysis from raw time-ordered data and finishes with parameter estimation. It is based on Gibbs sampling and is performed iteratively, meaning the end product is the set of independent samples. The BeyondPlanck project was successfully tested on *Planck* LFI data, and the main results, together with the entire pipeline, are summarized in Paper IV.

My work in BeyondPlanck was mainly connected with reproducibility, Open Science, and pipeline validation. In particular, I have developed the `CMake`-based infrastructure to easily compile `Commander3` with all its dependencies via a set of several different terminal commands. The entire installation process is completely automated, alleviating the pressure from the end user. The details of the `CMake` organization and our other efforts dedicated to reproducibility (i.e., `Commander3` documentation, data download utility, etc.) are summarized in Paper V. The pipeline validation was performed by analyzing a set of 1-year *Planck* LFI simulations created with only two — CMB and noise — components. These simulations were produced on various HPC around the globe within `Commander3` itself with the help of the `Fortran` module I wrote (see Paper VI for details).

Overall, the BeyondPlanck joint analysis framework constitutes a unique way to deal with degeneracies between the instrument and the sky, as well as allows for the global posterior exploration of all model parameters simultaneously. It is done in an iterative fashion with complete uncertainties propagation through the data processing pipeline. The success of BeyondPlanck effectively paved the way for the Cosmoglobe project, whose primary goal is to cover an entire sky in all relevant frequencies within a single analysis framework. Such an approach requires active participation from the community, so we believe our efforts in Open Science and reproducibility will lower the threshold for interested people to enter. Only via the combined effort from CMB scientists will we be able to face a new grand challenge — the detection of large-scale CMB polarization as an imprint of primordial gravitational waves.

Bibliography

[1] Adamek, J. et al. “gevolution: a cosmological N-body code based on General Relativity”. In: *Journal of Cosmology and Astroparticle Physics* vol. 2016, no. 07 (July 2016), p. 053.

[2] Akrami, Y. et al. *The Search for the Topology of the Universe Has Just Begun.* 2022.

[3] Albrecht, A. and Steinhardt, P. J. “Cosmology for Grand Unified Theories with Radiatively Induced Symmetry Breaking”. In: *Phys. Rev. Lett.* vol. 48 (17 Apr. 1982), pp. 1220–1223.

[4] Alpher, R. A., Bethe, H., and Gamow, G. “The Origin of Chemical Elements”. In: *Phys. Rev.* vol. 73 (7 Apr. 1948), pp. 803–804.

[5] Alvey, J. et al. “Cosmic neutrino background detection in large-neutrino-mass cosmologies”. In: *Phys. Rev. D* vol. 105 (6 Mar. 2022), p. 063501.

[6] Andersen et al. “**BeyondPlanck XIII. Intensity foreground sampling, degeneracies and priors**”. In: *A&A, in press* (2023). arXiv: 2201.08188.

[7] Anderson, J. L. “Cosmological Constant and Fundamental Length”. In: *American Journal of Physics* vol. 39, no. 8 (Aug. 1971), pp. 901–904.

[8] Anselmi, S. et al. “What is flat ΛCDM, and may we choose it?” In: *Journal of Cosmology and Astroparticle Physics* vol. 2023, no. 02 (2023), p. 049.

[9] Aslanyan, G. and Manohar, A. V. “Constraints on the global topology and size of the universe from the cosmic microwave background”. In: *Journal of Cosmology and Astroparticle Physics* vol. 2012, no. 06 (2012), pp. 003–003.

[10] Basyrov et al. “**BeyondPlanck X. Planck LFI frequency maps with sample-based error propagation**”. In: *A&A, in press* (2023). arXiv: 2208.14293.

[11] Bennett, C. L. et al. “First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Foreground Emission”. In: *ApJS* vol. 148 (Sept. 2003), pp. 97–117. eprint: astro-ph/0302208.

[12] Bennett, C. L. et al. “Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results”. In: *ApJS* vol. 208, 20 (Oct. 2013), p. 20. arXiv: 1212.5225 `[astro-ph.CO]`.

[13] Bennett, C. L. et al. “The Microwave Anisotropy Probe Mission”. In: *ApJ* vol. 583 (Jan. 2003), pp. 1–23.

[14] BeyondPlanck. “**BeyondPlanck I. Global Bayesian analysis of the *Planck* Low Frequency Instrument data**”. In: *A&A, in press* (2023). arXiv: 2011.05609.

[15] Bielewicz, P. and Banday, A. J. “Constraints on the topology of the Universe derived from the 7-year WMAP CMB data and prospects of constraining the topology using CMB polarization maps”. In: *Proceedings of the XLVIIth Rencontres de Moriond, 2012, eds. E. Auge, J. Dumarchez and J. Tran Thanh Van, published by ARISF* (2012), p. 91. arXiv: 1303.4004 `[astro-ph.CO]`.

[16] Bielewicz, P., Banday, A. J., and Gorski, K. M. “Constraints on the topology of the Universe derived from the 7-yr WMAP data”. In: *MNRAS* vol. 412, no. 3 (Apr. 2011), pp. 2104–2110. arXiv: 1012.3549 `[astro-ph.CO]`.

[17] Borriello, A. and Salucci, P. “The Dark matter distribution in disk galaxies”. In: *MNRAS* vol. 323 (2001), p. 285.

[18] Brilenkov, M., Eingorn, M., and Zhuk, A. “Lattice Universe: examples and problems”. In: *The European Physical Journal C* vol. 75, no. 5 (2015).

[19] Brilenkov et al. “**BeyondPlanck IV. On end-to-end simulations in CMB analysis — Bayesian versus frequentist statistics**”. In: *A&A, submitted* (2023). arXiv: 2209.04437.

[20] Bull, P. et al. “Beyond Λ CDM: Problems, solutions, and the road ahead”. In: *Physics of the Dark Universe* vol. 12 (June 2016), pp. 56–99. arXiv: 1512.05356 `[astro-ph.CO]`.

[21] Burgess, C. P. “The Cosmological Constant Problem: Why it’s hard to get Dark Energy from Micro-physics”. In: *arXiv e-prints*, arXiv:1309.4133 (Sept. 2013), arXiv:1309.4133. arXiv: 1309.4133 `[hep-th]`.

[22] Canay, E. et al. “Scalar and vector perturbations in a universe with nonlinear perfect fluid”. In: *European Physical Journal C* vol. 81, no. 3, 246 (Mar. 2021), p. 246. arXiv: 2011.05914 `[gr-qc]`.

[23] Chu, M. et al. “Cosmological parameter constraints as derived from the Wilkinson Microwave Anisotropy Probe data via Gibbs sampling and the Blackwell-Rao estimator”. In: *Phys. Rev. D* vol. 71, no. 10, 103002 (May 2005), p. 103002. eprint: astro-ph/0411737.

[24] Clifton, T. et al. “Modified gravity and cosmology”. In: *Phys. Rep.* vol. 513, no. 1 (Mar. 2012), pp. 1–189. arXiv: 1106.2476 `[astro-ph.CO]`.

[25] Coc, A. and Vangioni, E. “Primordial nucleosynthesis”. In: *International Journal of Modern Physics E* vol. 26, no. 8, 1741002 (Jan. 2017), p. 1741002. arXiv: 1707.01004 `[astro-ph.CO]`.

[26] Collaboration, K. et al. *Measurements of Degree-Scale B-mode Polarization with the BICEP/Keck Experiments at South Pole*. 2018. arXiv: 1807.02199 `[astro-ph.CO]`.

[27] Collaboration, L. et al. "Probing Cosmic Inflation with the LiteBIRD Cosmic Microwave Background Polarization Survey". In: *Progress of Theoretical and Experimental Physics* (Nov. 2022). ptac150. arXiv: 2202.02773 `[astro-ph.IM]`.

[28] Colombo et al. "**BeyondPlanck XI. Bayesian CMB analysis with sample-based end-to-end error propagation**". In: *A&A, in press* (2023). arXiv: 2208.14276.

[29] Corbelli, E. and Salucci, P. "The Extended Rotation Curve and the Dark Matter Halo of M33". In: *MNRAS* vol. 311 (2000), pp. 441–447.

[30] De Felice, A. and Tsujikawa, S. "f(R) Theories". In: *Living Reviews in Relativity* vol. 13, no. 1, 3 (June 2010), p. 3. arXiv: 1002.4928 `[gr-qc]`.

[31] Delabrouille, J. and Cardoso, J.-F. "Diffuse source separation in CMB observations". In: (2007). arXiv: astro-ph/0702198 `[astro-ph]`.

[32] Delabrouille, J., Cardoso, J.-F., and Patanchon, G. "Multidetector multicomponent spectral matching and applications for cosmic microwave background data analysis". In: *MNRAS* vol. 346 (Dec. 2003), pp. 1089–1102. arXiv: astro-ph/0211504.

[33] Delabrouille, J. et al. "A full sky, low foreground, high resolution CMB map from WMAP". In: *A&A* vol. 493 (Jan. 2009), pp. 835–857. arXiv: 0807.0773.

[34] Deruelle, N. and Uzan, J.-P. *Relativity in Modern Physics.* Oxford University Press, 2018, p. 704.

[35] Dodelson, S. *Modern Cosmology.* 1st edition. Academic Press, Mar. 2003, p. 456.

[36] Draine, B. T. and Hensley, B. S. "The Dielectric Function of "Astrodust" and Predictions for Polarization in the 3.4 and 10 μm Features". In: *ApJ* vol. 909, no. 1, 94 (Mar. 2021), p. 94. arXiv: 2009.11314 `[astro-ph.GA]`.

[37] Dupuy, H. and Bernardeau, F. "Describing massive neutrinos in cosmology as a collection of independent flows". In: *Journal of Cosmology and Astroparticle Physics* vol. 2014, no. 01 (Jan. 2014), p. 030. arXiv: 1311.5487 `[astro-ph.CO]`.

[38] Eingorn, M. "First-order Cosmological Perturbations Engendered by Point-like Masses". In: *ApJ* vol. 825, no. 2, 84 (July 2016), p. 84. arXiv: 1509.03835 `[gr-qc]`.

[39] Eingorn, M. and Brilenkov, R. "Perfect fluids with $\omega = \text{const}$ as sources of scalar cosmological perturbations". In: *Physics of the Dark Universe* vol. 17 (Sept. 2017), pp. 63–67. arXiv: 1509.08181 `[gr-qc]`.

[40] Eingorn, M., Kiefer, C., and Zhuk, A. "Cosmic screening of the gravitational interaction". In: *International Journal of Modern Physics D* vol. 26, no. 12 (2017), p. 1743012.

[41] Eingorn, M., Kiefer, C., and Zhuk, A. “Scalar and vector perturbations in a universe with discrete and continuous matter sources”. In: *Journal of Cosmology and Astroparticle Physics* vol. 2016, no. 09 (2016), pp. 032–032.

[42] Einstein, A. “Zur Elektrodynamik bewegter Körper”. In: *Annalen der Physik* vol. 17 (). [English translation by Jeffery, G. and Perrett, W. (1923). The Principle of relativity. London, Methuen.], pp. 891–921.

[43] Eriksen, H. K. et al. “Cosmic Microwave Background Component Separation by Parameter Estimation”. In: *ApJ* vol. 641 (Apr. 2006), pp. 665–682. arXiv: astro-ph/0508268.

[44] Eriksen, H. K. et al. “Joint Bayesian Component Separation and CMB Power Spectrum Estimation”. In: *ApJ* vol. 676 (Mar. 2008), pp. 10–32. arXiv: 0709.1058.

[45] Fiol, B. and Garriga, J. “Semiclassical unimodular gravity”. In: *J. Cosmology Astropart. Phys.* vol. 2010, no. 8, 015 (Aug. 2010), p. 015. arXiv: 0809.1371 [hep-th].

[46] Fixsen, D. J. “The Temperature of the Cosmic Microwave Background”. In: *ApJ* vol. 707 (Dec. 2009), pp. 916–920. arXiv: 0911.1955 [astro-ph.CO].

[47] Floratos, E. and Leontaris, G. “On topological modifications of Newton's law”. In: *Journal of Cosmology and Astroparticle Physics* vol. 2012, no. 04 (2012), pp. 024–024.

[48] Friedmann, A. “Über die Krümmung des Raumes”. In: *Zeitschrift fur Physik* vol. 10 (Jan. 1922). TL: On the curvature of space, pp. 377–386.

[49] Friedmann, A. “Über die Möglichkeit einer Welt mit konstanter negativer Krümmung des Raumes”. In: *Zeitschrift fur Physik* vol. 21, no. 1 (Dec. 1924), pp. 326–332.

[50] Frieman, J. A., Turner, M. S., and Huterer, D. “Dark energy and the accelerating universe.” In: *ARA&A* vol. 46 (Sept. 2008), pp. 385–432. arXiv: 0803.0982 [astro-ph].

[51] Fromerth, M. et al. “From Quark-Gluon Universe to Neutrino Decoupling: 200<T<2 MeV”. In: *Acta Physica Polonica B.* vol. 43 (2012). arXiv: 1211.4297 [nucl-th].

[52] Fujii, H. and Yoshii, Y. “An improved cosmic crystallography method to detect holonomies in flat spaces”. In: *Astronomy &amp Astrophysics* vol. 529 (2011), A121.

[53] Fujii, H. and Yoshii, Y. “A Search for Nontoroidal Topological Lensing in the Sloan Digital Sky Survey Quasar Catalog”. In: *ApJ* vol. 773, no. 2, 152 (Aug. 2013), p. 152. arXiv: 1306.2737 [astro-ph.CO].

[54] Galloway et al. “**BeyondPlanck III. Computational infrastructure and generalization to other data sets**”. In: *A&A, in press* (2023). arXiv: 2201.03509.

[55] Galloway et al. “**BeyondPlanck VIII. Sidelobe modelling and correction**”. In: *A&A, in press* (2023). arXiv: 2201.03478.

[56] Gannon, D. “Singularities in nonsimply connected space–times”. In: *Journal of Mathematical Physics* vol. 16, no. 12 (1975), pp. 2364–2367.

[57] Gelman, A. et al. *Bayesian Data Analysis*. 3d edition. Chapman and Hall/CRC, Nov. 2013, p. 675.

[58] Gerakakis, E., Brilenkov, M., and others, t. F. In: *The Open Journal of Astrophysics* vol. 6, 1 (Mar. 2023), p. 1. arXiv: 2205.11262 `[astro-ph.CO]`.

[59] Gjerløw, E. et al. “Optimized Large-scale CMB Likelihood and Quadratic Maximum Likelihood Power Spectrum Estimation”. In: *ApJS* vol. 221, no. 1, 5 (Nov. 2015), p. 5. arXiv: 1506.04273 `[astro-ph.IM]`.

[60] Gjerløw et al. “**BeyondPlanck VII. Gain and absolute calibration**”. In: *A&A, in press* (2023). arXiv: 2011.08082.

[61] Górski, K. M. et al. “HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere”. In: *ApJ* vol. 622 (Apr. 2005), pp. 759–771. arXiv: astro-ph/0409513.

[62] Gorski, K. M. “On Determining the Spectrum of Primordial Inhomogeneity from the COBE DMR Sky Maps: Method”. In: *ApJ* vol. 430 (Aug. 1994), p. L85. arXiv: astro-ph/9403066 `[astro-ph]`.

[63] Guth, A. H. “Inflationary universe: A possible solution to the horizon and flatness problems”. In: *Phys. Rev. D* vol. 23 (2 Jan. 1981), pp. 347–356.

[64] Hauser, M. G. and Dwek, E. “The Cosmic Infrared Background: Measurements and Implications”. In: *ARA&A* vol. 39 (Jan. 2001), pp. 249–307. arXiv: astro-ph/0105539 `[astro-ph]`.

[65] Herman et al. “**BeyondPlanck V. Minimal ADC Corrections for *Planck* LFI**”. In: *A&A, in press* (2022). arXiv: 2203.13635.

[66] Herman et al. “**BeyondPlanck XV. Limits on Large-Scale Polarized Anomalous Microwave Emission from Planck LFI and WMAP**”. In: *A&A, in press* (2023). arXiv: 2201.03530.

[67] Hubble, E. “A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae”. In: *Proceedings of the National Academy of Science* vol. 15, no. 3 (Mar. 1929), pp. 168–173.

[68] Hurier, G., Macías-Pérez, J. F., and Hildebrandt, S. “MILCA, a modified internal linear combination algorithm to extract astrophysical emissions from multifrequency sky maps”. In: *A&A* vol. 558, A118 (Oct. 2013), A118. arXiv: 1007.1149 `[astro-ph.IM]`.

[69] Huterer, D. and Turner, M. S. “Prospects for probing the dark energy via supernova distance measurements”. In: *Phys. Rev. D* vol. 60 (8 Aug. 1999), p. 081301.

[70] Ihle et al. “**BeyondPlanck VI. Noise characterization and modelling**”. In: *A&A, in press* (2023). arXiv: 2011.06650.

[71] Joyce, A. et al. “Beyond the cosmological standard model”. In: *Phys. Rep.* vol. 568 (Mar. 2015), pp. 1–98. arXiv: `1407.0059 [astro-ph.CO]`.

[72] Kazanas, D. “Dynamics of the universe and spontaneous symmetry breaking”. In: *ApJ* vol. 241 (Oct. 1980), pp. L59–L63.

[73] Keihänen et al. “**BeyondPlanck II. Optimal mapmaking by Gibbs sampling**”. In: *A&A, in press* (2023). arXiv: 2011.06024.

[74] Khoury, J. *Les Houches Lectures on Physics Beyond the Standard Model of Cosmology*. 2013. arXiv: `1312.2006 [astro-ph.CO]`.

[75] King, S. and Lubin, P. “Circular polarization of the CMB: Foregrounds and detection prospects”. In: *Phys. Rev. D* vol. 94 (2 July 2016), p. 023501. arXiv: `1606.04112 [astro-ph.CO]`.

[76] Lachièze-Rey, M. and Luminet, J.-P. “Cosmic topology”. In: *Physics Reports* vol. 254, no. 3 (1995), pp. 135–214. arXiv: `gr-qc/9605010 [gr-qc]`.

[77] Landau, L. D. and Lifshitz, E. M. *The Classical Theory of Fields*. Course of Theoretical Physics Series, Vol. 2. Oxford Pergamon Press, 2000.

[78] Leach, S. M. et al. “Component separation methods for the PLANCK mission”. In: *A&A* vol. 491 (Nov. 2008), pp. 597–615. arXiv: 0805.0269.

[79] Lehoucq, R., Lachieze-Rey, M., and Luminet, J. P. *Cosmic Crystallography*. 1996. arXiv: `gr-qc/9604050 [gr-qc]`.

[80] Leitch, E. M. et al. “An Anomalous Component of Galactic Emission”. In: *ApJ* vol. 486 (Sept. 1997), pp. L23–L26. eprint: astro-ph/9705241.

[81] Lemaître, G. “The Beginning of the World from the Point of View of Quantum Theory.” In: *Nature* vol. 127, no. 3210 (May 1931), p. 706.

[82] Lesgourgues, J. and Pastor, S. “Massive neutrinos and cosmology”. In: *Physics Reports* vol. 429, no. 6 (2006), pp. 307–379. arXiv: `astro-ph/0603494 [astro-ph]`.

[83] Lewis, A. and Bridle, S. “Cosmological parameters from CMB and other data: A Monte Carlo approach”. In: *Phys. Rev. D* vol. 66, no. 10, 103511 (Nov. 2002), p. 103511. eprint: astro-ph/0205436.

[84] Li, M. et al. “Dark Energy”. In: *Communications in Theoretical Physics* vol. 56, no. 3 (Sept. 2011), pp. 525–604. arXiv: `1103.5870 [astro-ph.CO]`.

[85] Liddle, A. *An Introduction to Modern Cosmology*. 2nd edition. Willey, May 2003, p. 200.

[86] Linde, A. D. “A new inflationary universe scenario: A possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems”. In: *Physics Letters B* vol. 108, no. 6 (Feb. 1982), pp. 389–393.

[87] LiteBIRD Collaboration et al. “Probing Cosmic Inflation with the LiteBIRD Cosmic Microwave Background Polarization Survey”. In: *arXiv e-prints*, arXiv:2202.02773 (Feb. 2022), arXiv:2202.02773. arXiv: `2202.02773 [astro-ph.IM]`.

[88] Longair, M. S. *Galaxy Formation.* 2nd edition. Springer, 2008.

[89] Luminet, J.-P. “Editorial note to: Georges Lemaître, The beginning of the world from the point of view of quantum theory”. In: *General Relativity and Gravitation* vol. 43, no. 10 (Oct. 2011), pp. 2911–2928.

[90] Luminet, J. P. “Past and Future of Cosmic Topology”. In: *Acta Cosmologica* vol. 24 (Jan. 1998), p. 105. arXiv: gr-qc/9804006 `[gr-qc]`.

[91] Luminet, J.-P. *The Shape and Topology of the Universe.* 2008. arXiv: 0802.2236 `[astro-ph]`.

[92] Marsh, D. J. E. et al. “Quintessence in a quandary: Prior dependence in dark energy models”. In: *Phys. Rev. D* vol. 90, no. 10, 105023 (Nov. 2014), p. 105023. arXiv: 1406.2301 `[astro-ph.CO]`.

[93] Martin, J. “Everything you always wanted to know about the cosmological constant problem (but were afraid to ask)”. In: *Comptes Rendus Physique* vol. 13, no. 6-7 (2012), pp. 566–665.

[94] Martínez-González, E. et al. “Cosmic microwave background power spectrum estimation and map reconstruction with the expectation-maximization algorithm”. In: *MNRAS* vol. 345, no. 4 (Nov. 2003), pp. 1101–1109. arXiv: astro-ph/0302094 `[astro-ph]`.

[95] Mastache, J. and de la Macorra, A. “Analytic Fluid Approximation for Warm Dark Matter”. In: *arXiv e-prints*, arXiv:1909.05132 (Sept. 2019), arXiv:1909.05132. arXiv: 1909.05132 `[astro-ph.CO]`.

[96] Mather, J. C. et al. “A Preliminary Measurement of the Cosmic Microwave Background Spectrum by the Cosmic Background Explorer (COBE) Satellite”. In: *ApJ* vol. 354 (May 1990), p. L37.

[97] Mukhanov, V. *Physical Foundations of Cosmology.* Illustrated edition. Cambridge University Press, Dec. 2005, p. 421.

[98] Nemiroff, R. J. and Patla, B. “Adventures in Friedmann cosmology: A detailed expansion of the cosmological Friedmann equations”. In: *American Journal of Physics* vol. 76, no. 3 (Mar. 2008), pp. 265–276. arXiv: astro-ph/0703739 `[astro-ph]`.

[99] Nojiri, S. and Odintsov, S. D. “Unified cosmic history in modified gravity: From F(R) theory to Lorentz non-invariant models”. In: *Phys. Rep.* vol. 505, no. 2 (Aug. 2011), pp. 59–144. arXiv: 1011.0544 `[gr-qc]`.

[100] Olive, K. A., Steigman, G., and Walker, T. P. “Primordial nucleosynthesis: theory and observations”. In: *Physics Reports* vol. 333-334 (2000), pp. 389–407. arXiv: 9905320 `[astro-ph]`.

[101] Padilla, A. and Saltas, I. D. “A note on classical and quantum unimodular gravity”. In: *European Physical Journal C* vol. 75, 561 (Nov. 2015), p. 561. arXiv: 1409.3573 `[gr-qc]`.

[102] Page, L. et al. “Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Polarization Analysis”. In: *ApJS* vol. 170 (June 2007), pp. 335–376. arXiv: astro-ph/0603450.

[103] Paradiso et al. "**BeyondPlanck XII. Cosmological parameter constraints with end-to-end error propagation**". In: *A&A, in press* (2023). arXiv: 2205.10104.

[104] Pasechnik, R. and Sumbera, M. "Phenomenological Review on Quark–Gluon Plasma: Concepts vs. Observations". In: *Universe* vol. 3, no. 1 (Jan. 2017), p. 7. arXiv: 1611.01533 `[hep-ph]`.

[105] Peebles, P. J. E. and Ratra, B. "The cosmological constant and dark energy". In: *Rev. Mod. Phys.* vol. 75 (2 Apr. 2003), pp. 559–606. arXiv: astro-ph/0207347 `[physics]`.

[106] Penzias, A. A. and Wilson, R. W. "A Measurement of Excess Antenna Temperature at 4080 Mc/s." In: *ApJ* vol. 142 (July 1965), pp. 419–421.

[107] Petersen, P. et al. "Cosmic topology. Part I. Limits on orientable Euclidean manifolds from circle searches". In: *Journal of Cosmology and Astroparticle Physics* vol. 2023, no. 01 (2023), p. 030.

[108] Piechocki, W. *Topology of the Universe.* 1999. arXiv: gr-qc/9910055 `[gr-qc]`.

[109] Planck Collaboration I. "*Planck* 2013 results. I. Overview of products and scientific results". In: *A&A* vol. 571 (2014), A1. arXiv: 1303.5062.

[110] Planck Collaboration XII. "*Planck* 2013 results. XII. Diffuse component separation". In: *A&A* vol. 571 (2014), A12. arXiv: 1303.5072.

[111] Planck Collaboration XIV. "*Planck* 2013 results. XIV. Zodiacal emission". In: *A&A* vol. 571 (2014), A14. arXiv: 1303.5074.

[112] Planck Collaboration XXIII. "*Planck* 2013 results. XXIII. Isotropy and statistics of the CMB". In: *A&A* vol. 571 (2014), A23. arXiv: 1303.5083.

[113] Planck Collaboration XXVI. "*Planck* 2013 results. XXVI. Background geometry and topology of the Universe". In: *A&A* vol. 571 (2014), A26. arXiv: 1303.5086.

[114] Planck Collaboration I. "*Planck* 2015 results. I. Overview of products and results". In: *A&A* vol. 594 (2016), A1. arXiv: 1502.01582.

[115] Planck Collaboration X. "*Planck* 2015 results. X. Diffuse component separation: Foreground maps". In: *A&A* vol. 594 (2016), A10. arXiv: 1502.01588.

[116] Planck Collaboration XI. "*Planck* 2015 results. XI. CMB power spectra, likelihoods, and robustness of parameters". In: *A&A* vol. 594 (2016), A11. arXiv: 1507.02704.

[117] Planck Collaboration XII. "*Planck* 2015 results. XII. Full Focal Plane simulations". In: *A&A* vol. 594 (2016), A12. arXiv: 1509.06348.

[118] Planck Collaboration XVI. "*Planck* 2015 results. XVI. Isotropy and statistics of the CMB". In: *A&A* vol. 594 (2016), A16. arXiv: 1506.07135.

[119] Planck Collaboration XVIII. “*Planck* 2015 results. XVIII. Background geometry and topology of the Universe”. In: *A&A* vol. 594 (2016), A18. arXiv: 1502.01593.

[120] Planck Collaboration XXV. “*Planck* 2015 results. XXV. Diffuse, low-frequency Galactic foregrounds”. In: *A&A* vol. 594 (2016), A25. arXiv: 1506.06660.

[121] Planck Collaboration I. “*Planck* 2018 results. I. Overview, and the cosmological legacy of *Planck*”. In: *A&A* vol. 641 (2020), A1. arXiv: 1807.06205.

[122] Planck Collaboration IV. “*Planck* 2018 results. IV. Diffuse component separation”. In: *A&A* vol. 641 (2018), A4. arXiv: 1807.06208.

[123] Planck Collaboration V. “*Planck* 2018 results. V. Power spectra and likelihoods”. In: *A&A* vol. 641 (2020), A5. arXiv: 1907.12875.

[124] Planck Collaboration VI. “*Planck* 2018 results. VI. Cosmological parameters”. In: *A&A* vol. 641 (2020), A6. arXiv: 1807.06209.

[125] Planck Collaboration VII. “*Planck* 2018 results. VII. Isotropy and statistics”. In: *A&A* vol. 641 (2020), A7. arXiv: 1906.02552.

[126] Planck Collaboration XI. “*Planck* 2018 results. XI. Polarized dust foregrounds”. In: *A&A* vol. 641 (2020), A11. arXiv: 1801.04945.

[127] Prézeau, G. and Reinecke, M. “Algorithm for the Evaluation of Reduced Wigner Matrices”. In: *ApJS* vol. 190, no. 2 (Oct. 2010), pp. 267–274. arXiv: 1002.1050 `[astro-ph.IM]`.

[128] Prialnik, D. *An Introduction to the Theory of Stellar Structure and Evolution.* 2nd edition. Cambridge University Press, 2009.

[129] Primack, J. R. “DARK MATTER, GALAXIES AND LARGE SCALE STRUCTURE IN THE UNIVERSE”. In: *Proc. Int. Sch. Phys. Fermi* vol. 92 (1984), p. 140.

[130] Reinecke, M. et al. *Planck Level-S: Planck Simulation Package.* Astrophysics Source Code Library, record ascl:1505.032. May 2015. ascl: 1505.032.

[131] Rogers, K. K. et al. “Spin-SILC: CMB polarization component separation with spin wavelets”. In: *MNRAS* vol. 463, no. 3 (Dec. 2016), pp. 2310–2322. arXiv: 1605.01417 `[astro-ph.CO]`.

[132] Roukema, B. F. “The topology of the universe”. In: *Bulletin of the Astronomical Society of India* vol. 28 (Sept. 2000), pp. 483–497. arXiv: astro-ph/0010185 `[astro-ph]`.

[133] Roukema, B. F. “Topology of the Universe: Background and recent observational approaches”. In: *Pramana* vol. 53, no. 6 (1999), pp. 945–949.

[134] Rubakov, V. and Gorbunov, D. *Introduction To The Theory Of The Early Universe: Hot Big Bang Theory.* 2nd edition. World Scientific Publishing Company, Aug. 2017, p. 596.

[135] Rudjord, Ø. et al. “Cosmic Microwave Background Likelihood Approximation by a Gaussianized Blackwell-Rao Estimator”. In: *ApJ* vol. 692, no. 2 (Feb. 2009), pp. 1669–1677. arXiv: 0809.4624 [astro-ph].

[136] Rybicki, G. B. and Lightman, A. P. *Radiative Processes in Astrophysics.* New York: Wiley, 1985.

[137] Ryden, B. S. “Introduction to Cosmology”. In: Jan. 2006.

[138] Sato, K. “First-order phase transition of a vacuum and the expansion of the Universe”. In: *MNRAS* vol. 195 (May 1981), pp. 467–479.

[139] Schwarzschild, K. “Ueber das zulaessige Kruemmungsmaass des Raumes”. In: *Vierteljahrsschrift der Ast. Ges* vol. 35 (Jan. 1900), p. 337.

[140] Seljebotn, D. S. et al. “Multi-resolution Bayesian CMB component separation through Wiener filtering with a pseudo-inverse preconditioner”. In: *A&A* vol. 627, A98 (July 2019), A98.

[141] Shewchuk, J. R. *An Introduction to the Conjugate Gradient Method Without the Agonizing Pain, Edition* $1\frac{1}{4}$. http://www.cs.cmu.edu/~quake-papers/painless-conjugate-gradient.pdf. USA, 1994.

[142] Slepian, Z. and Portillo, S. K. N. “Too hot to handle? Analytic solutions for massive neutrino or warm dark matter cosmologies”. In: *MNRAS* vol. 478, no. 1 (May 2018), pp. 516–529. arXiv: 1710.01785 [astro-ph.CO].

[143] Sotiriou, T. P. and Faraoni, V. “f(R) theories of gravity”. In: *Reviews of Modern Physics* vol. 82, no. 1 (Jan. 2010), pp. 451–497. arXiv: 0805.1726 [gr-qc].

[144] Starobinsky, A. A. “Dynamics of phase transition in the new inflationary universe scenario and generation of perturbations”. In: *Physics Letters B* vol. 117, no. 3 (1982), pp. 175–178.

[145] Starobinsky, A. A. “Spectrum of relict gravitational radiation and the early state of the universe”. In: *Soviet Journal of Experimental and Theoretical Physics Letters* vol. 30 (1979).

[146] Stewart, J. M., Stewart, M. E., and Schwarzschild, K. “‘On the permissible curvature of space’ by K Schwarzschild”. In: *Classical and Quantum Gravity* vol. 15, no. 9 (Sept. 1998), pp. 2539–2544.

[147] Stompor, R. et al. “Maximum likelihood algorithm for parametric component separation in cosmic microwave background experiments”. In: *MNRAS* vol. 392 (Jan. 2009), pp. 216–232. eprint: arXiv:0804.2645.

[148] Sunyaev, R. A. and Chluba, J. “Signals from the epoch of cosmological recombination”. In: *Astronomical Notes* vol. 330, no. 7 (Aug. 2009), pp. 657–674. arXiv: 0908.0435 [astro-ph.CO].

[149] Sunyaev, R. A and Zeldovich, Y. B. “The Observations of Relic Radiation as a Test of the Nature of X-Ray Radiation from the Clusters of Galaxies”. In: *Comments on Astrophysics and Space Physics* vol. 4 (Nov. 1972), p. 173.

[150] Svalheim et al. “**BeyondPlanck IX. Bandpass and beam mismatch correction**”. In: *A&A, in press* (2023). arXiv: 2201.03417.

[151] Svalheim et al. “**BeyondPlanck XIV. Polarized foreground emission between 30 and 70 GHz**”. In: *A&A, in press* (2023). arXiv: 2011.08503.

[152] Tauber, J. A. et al. “Characterization of the in-flight properties of the *Planck* telescope”. In: *A&A* vol. 622 (2019), A55.

[153] Tegmark, M. “How to Make Maps from Cosmic Microwave Background Data without Losing Information”. In: *ApJ* vol. 480, no. 2 (May 1997), pp. L87–L90. arXiv: astro-ph/9611130 [astro-ph].

[154] Tegmark, M., Taylor, A. N., and Heavens, A. F. “Karhunen-Loève Eigenvalue Problems in Cosmology: How Should We Tackle Large Data Sets?” In: *ApJ* vol. 480, no. 1 (May 1997), pp. 22–35. arXiv: astro-ph/9603021 [astro-ph].

[155] Tsujikawa, S. “Quintessence: a review”. In: *Classical and Quantum Gravity* vol. 30, no. 21, 214003 (Nov. 2013), p. 214003. arXiv: 1304.1961 [gr-qc].

[156] Vega, H. J. de and Sanchez, N. G. “Cosmological evolution of warm dark matter fluctuations. I. Efficient computational framework with Volterra integral equations”. In: *Phys. Rev. D* vol. 85 (4 Feb. 2012), p. 043516. arXiv: 1111.0290 [astro-ph.CO].

[157] Villa, F., Mandolesi, N., and Butler, R. C. “The Planck Project”. In: *Mem. Soc. Astron. Italiana* vol. 74 (2003), p. 223.

[158] Vittorio, N. *Cosmology.* 1st edition. CRC press, Nov. 2017, p. 428.

[159] Wandelt, B. D. and Górski, K. M. “Fast convolution on the sphere”. In: *Phys. Rev. D* vol. 63, no. 12, 123002 (June 2001), p. 123002. arXiv: astro-ph/0008227 [astro-ph].

[160] Watts et al. “From BeyondPlanck to Cosmoglobe: Preliminary WMAP Q-band analysis”. In: *A&A, in press* (2023). arXiv: 2202.11979.

[161] Wright, E. L. “A Cosmology Calculator for the World Wide Web”. In: *PASP* vol. 118, no. 850 (Dec. 2006), pp. 1711–1715. arXiv: astro-ph/0609593 [astro-ph].

[162] Yagi, K., Miake, Y., and Hatsuda, T. “Quark-Gluon Plasma: From Big Bang to Little Bang”. In: Cambridge University Press, 2006.

[163] Zaldarriaga, M. and Seljak, U. “All-sky analysis of polarization in the microwave background”. In: *Phys. Rev. D* vol. 55 (4 Feb. 1997), pp. 1830–1840. arXiv: astro-ph/9609170 [astro-ph].

[164] Zwicky, F. “Die Rotverschiebung von extragalaktischen Nebeln”. In: *Helvetica Physica Acta* vol. 6 (Jan. 1933), pp. 110–127.

Papers

Paper I

Cosmological Perturbations Engendered by Discrete Relativistic Species

M. Brilenkov, E. Canay, and M. Eingorn

Submitted to the *Physics of the Dark Universe*, arXiv: 2206.13495v1

Paper II

Gravitational Interaction in the Chimney Lattice Universe

M. Eingorn, A. McLaughlin II, E. Canay, M. Brilenkov, and A. Zhuk

Published in *MDPI Universe* 2021, 7(4), 101; DOI: 10.3390/universe7040101.

universe

Article

Gravitational Interaction in the Chimney Lattice Universe †

Maxim Eingorn [1,*], Andrew McLaughlin II [1], Ezgi Canay [2], Maksym Brilenkov [3] and Alexander Zhuk [4]

[1] Department of Mathematics and Physics, North Carolina Central University, 1801 Fayetteville St., Durham, NC 27707, USA; amclau12@eagles.nccu.edu
[2] Department of Physics, Istanbul Technical University, Maslak, 34469 Istanbul, Turkey; ezgicanay@itu.edu.tr
[3] Institute of Theoretical Astrophysics, University of Oslo, P.O. Box 1029 Blindern, N-0315 Oslo, Norway; maksym.brilenkov@astro.uio.no
[4] Astronomical Observatory, Odessa I.I. Mechnikov National University, Dvoryanskaya St. 2, 65082 Odessa, Ukraine; ai.zhuk2@gmail.com
* Correspondence: maxim.eingorn@gmail.com or meingorn@nccu.edu

† This paper is an extended version from the proceeding paper: Eingorn, M.; McLaughlin, A., II; Canay, E.; Brilenkov, M.; Zhuk, A. Gravitation in the space with chimney topology. In Proceedings of the 1st Electronic Conference on Universe, Online, 22–28 February 2021.

Citation: Eingorn, M.; McLaughlin, A., II; Canay, E.; Brilenkov M.; Zhuk A. Gravitational Interaction in the Chimney Lattice Universe. *Universe* **2021**, *7*, 101. https://doi.org/10.3390/universe7040101

Academic Editor: James A. Isenberg, Gerald B. Cleaver, Lijing Shao, Gonzalo J. Olmo and Giacomo Tommei

Received: 30 March 2021
Accepted: 12 April 2021
Published: 15 April 2021

Abstract: We investigate the influence of the chimney topology $T \times T \times R$ of the Universe on the gravitational potential and force that are generated by point-like massive bodies. We obtain three distinct expressions for the solutions. One follows from Fourier expansion of delta functions into series using periodicity in two toroidal dimensions. The second one is the summation of solutions of the Helmholtz equation, for a source mass and its infinitely many images, which are in the form of Yukawa potentials. The third alternative solution for the potential is formulated via the Ewald sums method applied to Yukawa-type potentials. We show that, for the present Universe, the formulas involving plain summation of Yukawa potentials are preferable for computational purposes, as they require a smaller number of terms in the series to reach adequate precision.

Keywords: spatial topology; gravitational potential; Yukawa interaction

PACS: 04.25.Nx—Post-Newtonian approximation; perturbation theory; related approximations; 98.80.Jk—Mathematical and relativistic aspects of cosmology

1. Introduction

The shape of the space, whether it is positively curved, negatively curved, or flat, and whether there is a limit to the size of the Universe are all among essential topics of contemporary debate in theoretical physics and cosmology. Spatial topology of the Universe, its function at the very early stages of evolution (in the quantum gravity regime), and in the later process of large scale structure formation are quite interesting questions yet to be answered. General Relativity does not favor any particular topology; hence, on theoretical grounds, the space might be simply connected, in agreement with concordance cosmology, or equally as well, multiply connected. It is worth noting that some "exotic" non-simply connected spacetimes are timelike geodesically incomplete, since they have singularities [1]. Hence, such topologies are not viable.

If the Universe is multiply connected, it may have a finite volume and yet be negatively curved or flat [2]. The current available data cannot reveal the finiteness of its volume if the Universe covers a much wider region than the observable sector. However, a rather smaller volume points at the possibility of finding observational indications of its topological features

[3]. For instance, a photon can travel plenty of times across the volume of multiply connected space and, thus, generate multiple images of the emitting source as a signature [4,5]. Spaces with toroidal topology in one to three dimensions may be presented as common examples of multiply connected spaces. To this class belongs the three-torus $T \times T \times T$, chimney $T \times T \times R$, and slab $T \times R \times R$ topologies.

There are various comprehensive studies in the literature on potential indicators of the shape of the space [6–10], and the majority of research is focused on their relation to the Cosmic Microwave Background (CMB) data. Indeed, there exists a very appealing conjectural relation which suggests that CMB anomalies in large angular scale observations, e.g., the suppression of the quadrupole moment and the quadrupole and octopole alignment, are imprints of spatial topology [11,12]. The weak wide-angle temperature correlations in the CMB can be also explained, e.g., with the help of dodecahedral topology of the Universe [13]. In the present work we study the chimney topology, which admits a single infinite axis subject to interpretations such as the preferred direction of the quadrupole and octopole alignment and the commonly named "axis of evil" [14] (see [15] for additional observable signatures of a preferred axis).

In connection with the investigation of possible topological imprints in CMB observations, Planck 2013 data [2] place the constraint $R_i > 0.71\chi_{\mathrm{rec}}$ on the radius of the largest sphere that may be inscribed in the topological domain for a flat Universe with the equal-sided chimney topology. The parameter χ_{rec} specifies the distance from the recombination surface and it is of the same order with the particle horizon, that is, $\chi_{\mathrm{rec}} \sim 14\,\mathrm{Gpc}$. For the toroidal topologies, the former restrictions on the size of the Universe from the seven- and nine-year WMAP temperature map analyses are presented in [14,16]. The smallest possible size of the fundamental topological domain for flat space, according to the seven-year WMAP results, is $d = 2R_{\mathcal{LSS}}\cos(\alpha_{\min}) \simeq 27.9\,\mathrm{Gpc}$ [12], where $R_{\mathcal{LSS}}$ stands for the distance from the last scattering surface.

In this paper, we study the chimney topology $T \times T \times R$ in terms of the gravitational characteristics of the Universe, manifested in the shape of the gravitational potential and force. In the cosmological setting, the inhomogeneous gravitational field is sourced by fluctuations in the matter density [17] and, as expected, in the Newtonian limit, the potential satisfies the Poisson equation. The form of the gravitational potential in the case of toroidal topologies was previously studied in [18]. Particularly, the authors have shown that there exists no physically justified nontrivial solution of the Poisson equation for the $T \times T \times R$ model. On the other hand, by employing the perturbed Einstein equations from the very beginning, one automatically includes the essential relativistic effects in the formulation and, for the gravitational potential, obtains a Helmholtz-type equation instead of the Poisson one [19–21]. Quite remarkably, as we show in the present work for the chimney topology, it then becomes possible to obtain exact solutions of this equation that are nontrivial and physically meaningful. Herein, we derive distinct expressions for the gravitational potential and force through alternative methods and point out the particular solutions appearing in the form of summed Yukawa potentials as ready-to-use notable sources for numerical computations. It is worth mentioning that, in the above outlined approach, we make no presumptions regarding the spatial distribution of gravitating bodies.

The outline of the paper is as follows. In Section 2, following [22], we introduce the main equations and derive alternative expressions for the gravitational potential using distinct methods, now including the Ewald technique. Subsequently, in Section 3.1, extending the results of [22], we compare these expressions in view of their usefulness for numerical computations. In Section 3.2, we obtain the gravitational force expressions for each form of the potential. We briefly review the results of our work in the concluding Section 4.

2. Methods

The Model and Basic Equations

It is well known that the gravitational potential Φ, created by fluctuations in the matter density, is defined by scalar perturbations of the metric coefficients [23] and that in the framework of General Relativity, it satisfies the linearized Einstein equation (see, e.g., [24,25]). Ignoring peculiar velocities, in the case of the ΛCDM cosmological model, this equation reads [19–21]

$$\Delta\Phi_0 - \frac{3\kappa\bar{\rho}c^2}{2a}\Phi_0 = \frac{\kappa c^2}{2a}(\rho - \bar{\rho})\,, \tag{1}$$

where $\kappa \equiv 8\pi G_N/c^4$ (with G_N and c being the Newtonian gravitational constant and the speed of light, respectively), a is the scale factor, while Δ represents the Laplace operator in comoving coordinates. Here, ρ and $\bar{\rho} = \text{const}$ are the comoving mass density and its averaged value, respectively. As we operate within the ΛCDM model, matter is pressureless, and we consider it in the form of discrete point-like gravitating bodies with masses m_n to represent, e.g., galaxies. Therefore, the comoving mass density

$$\rho = \sum_n m_n \delta(\mathbf{r} - \mathbf{r}_n)\,. \tag{2}$$

The 0 subscript of Φ in Equation (1) refers to the fact that peculiar velocities have been disregarded (see also [26]).

The shifted gravitational potential

$$\widehat{\Phi}_0 \equiv \Phi_0 - \frac{1}{3} \tag{3}$$

fulfils the equation

$$\Delta\widehat{\Phi}_0 - \frac{a^2}{\lambda^2}\widehat{\Phi}_0 = \frac{\kappa c^2}{2a}\rho\,, \tag{4}$$

where the screening length [19]

$$\lambda \equiv \left(\frac{3\kappa\bar{\rho}c^2}{2a^3}\right)^{-1/2}. \tag{5}$$

The presence of the term $\propto \Phi_0$ in Equation (1) (consequently, $\propto \widehat{\Phi}_0$ in Equation (4)) results in the Yukawa-type cutoff of the potential with the characteristic length λ. The term $\propto \Phi$ enters as a summand into energy-momentum fluctuations generating metric perturbations [19].

In what follows, the overhat indicates that the gravitational potential is shifted. A significant bonus of working with the shifted potential is that it is now possible to employ the superposition principle in solving Equation (4): once we find a solution for a single particle that is located at the center of Cartesian coordinates, we may immediately generalize it for a collection of particles at random positions.

We consider the space with chimney topology $T_1 \times T_2 \times R$, where the tori T_1 and T_2 have periods l_1 and l_2 along, e.g., the x- and y-axes, respectively. Hence, each gravitating body has its images positioned away from the original point in multiples of periods l_1 and l_2 along the corresponding axes. Now, let us place a particle with mass m at the center of Cartesian coordinates. For the above indicated topology, the delta functions $\delta(x)$ and $\delta(y)$ may be presented as

$$\delta(x) = \frac{1}{l_1}\sum_{k_1=-\infty}^{+\infty}\cos\left(\frac{2\pi k_1}{l_1}x\right), \quad \delta(y) = \frac{1}{l_2}\sum_{k_2=-\infty}^{+\infty}\cos\left(\frac{2\pi k_2}{l_2}y\right), \tag{6}$$

which implicitly include the contribution from the images of the particle. Consequently, Equation (4) for this particle reads

$$\Delta\widehat{\Phi}_0 - \frac{a^2}{\lambda^2}\widehat{\Phi}_0 = \frac{\kappa c^2}{2a}\frac{m}{l_1 l_2}\sum_{k_1=-\infty}^{+\infty}\sum_{k_2=-\infty}^{+\infty}\cos\left(\frac{2\pi k_1}{l_1}x\right)\cos\left(\frac{2\pi k_2}{l_2}y\right)\delta(z)\,, \tag{7}$$

so, we are motivated to consider the solution

$$\widehat{\Phi}_0 = \sum_{k_1=-\infty}^{+\infty}\sum_{k_2=-\infty}^{+\infty} C_{k_1k_2}(z)\cos\left(\frac{2\pi k_1}{l_1}x\right)\cos\left(\frac{2\pi k_2}{l_2}y\right)\,, \tag{8}$$

where the coefficients $C_{k_1k_2}(z)$ satisfy the equation

$$\begin{aligned}&\sum_{k_1=-\infty}^{+\infty}\sum_{k_2=-\infty}^{+\infty}\left[C''_{k_1k_2}(z) - 4\pi^2\left(\frac{k_1^2}{l_1^2}+\frac{k_2^2}{l_2^2}\right)C_{k_1k_2}(z) - \frac{a^2}{\lambda^2}C_{k_1k_2}(z) - \frac{\kappa c^2}{2a}\frac{m}{l_1 l_2}\delta(z)\right]\\ \times\;&\cos\left(\frac{2\pi k_1}{l_1}x\right)\cos\left(\frac{2\pi k_2}{l_2}y\right) = 0\,.\end{aligned} \tag{9}$$

Using the condition $d^2|z|/dz^2 = 2\delta(z)$, we can easily obtain the explicit expressions for the coefficients $C_{k_1k_2}(z)$, so that the shifted gravitational potential for the selected particle and all its images eventually reads

$$\begin{aligned}\widehat{\Phi}_0 \;=\;& -\frac{\kappa c^2}{4a}\frac{m}{l_1 l_2}\sum_{k_1=-\infty}^{+\infty}\sum_{k_2=-\infty}^{+\infty}\left[4\pi^2\left(\frac{k_1^2}{l_1^2}+\frac{k_2^2}{l_2^2}\right)+\frac{a^2}{\lambda^2}\right]^{-1/2}\\ &\times\exp\left(-\sqrt{4\pi^2\left(\frac{k_1^2}{l_1^2}+\frac{k_2^2}{l_2^2}\right)+\frac{a^2}{\lambda^2}}\,|z|\right)\cos\left(\frac{2\pi k_1}{l_1}x\right)\cos\left(\frac{2\pi k_2}{l_2}y\right).\end{aligned} \tag{10}$$

The above expression has the correct behavior in the Newtonian limit in the neighborhood of the considered particle, where it is no longer possible to distinguish between different (infinite and periodic) axes. For such regions, the summations in (10) may be replaced by the integrals:

$$\begin{aligned}\widehat{\Phi}_0 \;\to\;& -\frac{\kappa c^2 m}{4a}\int_{-\infty}^{+\infty}dk_x\int_{-\infty}^{+\infty}dk_y\left[4\pi^2\left(k_x^2+k_y^2\right)+\frac{a^2}{\lambda^2}\right]^{-1/2}\\ &\times\exp\left(-\sqrt{4\pi^2\left(k_x^2+k_y^2\right)+\frac{a^2}{\lambda^2}}\,|z|\right)\cos\left[2\pi\left(k_x x + k_y y\right)\right]\end{aligned} \tag{11}$$

for $k_x \equiv k_1/l_1$ and $k_y \equiv k_2/l_2$. Introducing the vectors $\mathbf{k} = (k_x, k_y)$, $\mathbf{r} = (x, y)$ with the absolute values $k = \sqrt{k_x^2 + k_y^2}$ and $r = \sqrt{x^2 + y^2}$, and assuming an angle φ between them, we obtain

$$\begin{aligned}\widehat{\Phi}_0 \quad &\to \quad -\frac{\kappa c^2 m}{4a}\int_0^{+\infty} kdk\left[4\pi^2k^2 + \frac{a^2}{\lambda^2}\right]^{-1/2} \exp\left. -\sqrt{4\pi^2k^2 + \frac{a^2}{\lambda^2}}|z|\right) \\ &\qquad \times \int_0^{2\pi} d\varphi \cos(2\pi kr\cos\varphi) \\ &= \quad -\frac{\pi\kappa c^2 m}{2a}\int_0^{+\infty} kdk\left[4\pi^2k^2 + \frac{a^2}{\lambda^2}\right]^{-1/2} \exp\left. -\sqrt{4\pi^2k^2 + \frac{a^2}{\lambda^2}}|z|\right) J_0(2\pi kr) \\ &= \quad -\frac{G_N m}{c^2}\frac{1}{\sqrt{Z^2 + R^2}}\exp\left(-\frac{1}{\lambda}\sqrt{Z^2 + R^2}\right) \to -\frac{G_N m}{c^2}\frac{1}{\sqrt{Z^2 + R^2}}\,, \end{aligned} \tag{12}$$

where $Z = az$ and $R = ar$ represent the physical distances, and the last integration is performed by using the formula 2.12.10(10) of [27].

Evidently, for a system of randomly positioned gravitating bodies, we have

$$\Phi_0 = \frac{1}{3} + \widehat{\Phi}_0\left\{m \to \sum_n m_n;\ x, y, z \to x - x_n, y - y_n, z - z_n\right\}. \tag{13}$$

For linear fluctuations, the averaged value of this expression is equal to zero, as it should be (also see [28]). Indeed,

$$\begin{aligned}\int\limits_0^{l_1} dx \int\limits_0^{l_2} dy \int\limits_{-\infty}^{+\infty} dz \widehat{\Phi}_0 &= -\frac{\kappa c^2}{4a} m \frac{\lambda}{a}\int\limits_{-\infty}^{+\infty} \exp\left(-\frac{a}{\lambda}|z|\right) dz \\ = \frac{\kappa c^2}{2a} m \frac{\lambda^2}{a^2}\exp\left(-\frac{a}{\lambda}z\right)\Big|_0^{+\infty} &= -\frac{\kappa c^2}{2a} m \frac{\lambda^2}{a^2} = -\frac{1}{3}\frac{m}{\overline{\rho}}\,, \end{aligned} \tag{14}$$

and, hence, the spatial average of the total gravitational potential is

$$\overline{\Phi_0} = \frac{1}{3} - \frac{1}{3}\frac{m}{\overline{\rho}} \cdot \frac{N}{l_1 l_2 L_z} = \frac{1}{3} - \frac{1}{3} = 0, \quad \frac{mN}{l_1 l_2 L_z} = \overline{\rho}\,. \tag{15}$$

For the sake of simplicity, here we have considered the particular configuration in which all N bodies in the volume $V = l_1 l_2 L_z$ are assigned identical masses m.

Equation (4) is of Helmholtz type and we can likewise solve it by considering the contribution of periodic images. In this case, the resulting expression consists of summed Yukawa potentials attributed to each one of them:

$$\begin{aligned}\widehat{\Phi}_0 \quad &= \quad -\frac{\kappa c^2 m}{8\pi a}\sum_{k_1=-\infty}^{+\infty}\sum_{k_2=-\infty}^{+\infty} \frac{1}{\sqrt{(x - k_1 l_1)^2 + (y - k_2 l_2)^2 + z^2}} \\ &\quad \times \quad \exp\left. -\frac{a\sqrt{(x - k_1 l_1)^2 + (y - k_2 l_2)^2 + z^2}}{\lambda}\right). \end{aligned} \tag{16}$$

As we have noted previously, the peculiar motion of gravitating bodies is disregarded in Equation (1) and, consequently, in (4). Nevertheless, the significance of such a contribution has recently been pointed out in [29], where the authors have also shown that peculiar velocities may be effectively restored by employing the effective screening length λ_{eff} (given by the

Formula (41) of [29]) instead of the screening length λ in Equations (1) and (4). Specifically, in the matter-dominated epoch, the two quantities λ_{eff} and λ are related to one another as $\lambda_{\mathrm{eff}} = \sqrt{3/5}\lambda$. Returning to our formulation, the effect of peculiar motion is included by replacing λ with λ_{eff} in the Formulas (10) and (16), which yields

$$\begin{aligned}\tilde{\Phi}_{\mathrm{cos}} &\equiv \left(-\frac{\kappa c^2}{8\pi a}\frac{m}{l}\right)^{-1}\hat{\Phi}_{\mathrm{cos}} = \sum_{k_1=-\infty}^{+\infty}\sum_{k_2=-\infty}^{+\infty} k_1^2 + k_2^2 + \frac{1}{4\pi^2\tilde{\lambda}_{\mathrm{eff}}^2}\Bigg)^{-1/2} \\ &\times \ \exp \ -\sqrt{4\pi^2\left(k_1^2+k_2^2\right) + \frac{1}{\tilde{\lambda}_{\mathrm{eff}}^2}}|\tilde{z}|\Bigg)\cos(2\pi k_1\tilde{x})\cos(2\pi k_2\tilde{y}) \end{aligned} \tag{17}$$

and

$$\begin{aligned}\tilde{\Phi}_{\mathrm{exp}} &\equiv \left(-\frac{\kappa c^2}{8\pi a}\frac{m}{l}\right)^{-1}\hat{\Phi}_{\mathrm{exp}} = \sum_{k_1=-\infty}^{+\infty}\sum_{k_2=-\infty}^{+\infty}\frac{1}{\sqrt{(\tilde{x}-k_1)^2+(\tilde{y}-k_2)^2+\tilde{z}^2}} \\ &\times \ \exp\left(-\frac{\sqrt{(\tilde{x}-k_1)^2+(\tilde{y}-k_2)^2+\tilde{z}^2}}{\tilde{\lambda}_{\mathrm{eff}}}\right). \end{aligned} \tag{18}$$

For simpler demonstration, we have assumed $l_1 = l_2 = l$ and introduced the rescaled quantities

$$x = \tilde{x}l, \quad y = \tilde{y}l, \quad z = \tilde{z}l, \quad \lambda_{\mathrm{eff}} = \tilde{\lambda}_{\mathrm{eff}}al\,. \tag{19}$$

Two alternative solutions are labeled with the subscripts "cos" and "exp" in Equations (17) and (18). Now that the peculiar velocities are included in the calculations, the 0 subscript is omitted in the new formulas.

There is also a third way to express the gravitational potential for the given topology. Indeed, Yukawa-type interactions that are subject to periodic boundary conditions can be formulated via Ewald sums, so that the expression for the potential consists of two rapidly converging series, one in each of the real and Fourier spaces. The technique is commonly employed while modeling particle interactions in plasma and colloids, and, in such a context, the corresponding potential for quasi two-dimensional systems, i.e., three-dimensional systems with two-dimensional periodicity, has previously been derived in [30,31]. Being implemented in the cosmological setting considered in our paper, the discussed expression for the "Yukawa–Ewald" potential reads

$$\begin{aligned}\tilde{\Phi}_{\mathrm{mix}} &\equiv \left(-\frac{\kappa c^2}{8\pi a}\frac{m}{l}\right)^{-1}\hat{\Phi}_{\mathrm{mix}} \\ &= \sum_{k_1=-\infty}^{+\infty}\sum_{k_2=-\infty}^{+\infty}\left[\frac{D\left(\sqrt{(\tilde{x}-k_1)^2+(\tilde{y}-k_2)^2+\tilde{z}^2};\alpha;\tilde{\lambda}_{\mathrm{eff}}\right)}{2\sqrt{(\tilde{x}-k_1)^2+(\tilde{y}-k_2)^2+\tilde{z}^2}}\right. \\ &+ \left.\pi\cos[2\pi(k_1\tilde{x}+k_2\tilde{y})]\frac{F\left(\sqrt{4\pi^2(k_1^2+k_2^2)+\tilde{\lambda}_{\mathrm{eff}}^{-2}};\tilde{z};\alpha\right)}{\sqrt{4\pi^2(k_1^2+k_2^2)+\tilde{\lambda}_{\mathrm{eff}}^{-2}}}\right], \end{aligned} \tag{20}$$

where

$$
\begin{aligned}
& D\left(\sqrt{(\tilde{x}-k_1)^2+(\tilde{y}-k_2)^2+\tilde{z}^2};\alpha;\tilde{\lambda}_{\text{eff}}\right) \\
\equiv\ & \exp\left(\frac{\sqrt{(\tilde{x}-k_1)^2+(\tilde{y}-k_2)^2+\tilde{z}^2}}{\tilde{\lambda}_{\text{eff}}}\right) \\
\times\ & \operatorname{erfc}\left(\alpha\sqrt{(\tilde{x}-k_1)^2+(\tilde{y}-k_2)^2+\tilde{z}^2}+\frac{1}{2\alpha\tilde{\lambda}_{\text{eff}}}\right) \\
+\ & \exp\left(-\frac{\sqrt{(\tilde{x}-k_1)^2+(\tilde{y}-k_2)^2+\tilde{z}^2}}{\tilde{\lambda}_{\text{eff}}}\right) \\
\times\ & \operatorname{erfc}\left(\alpha\sqrt{(\tilde{x}-k_1)^2+(\tilde{y}-k_2)^2+\tilde{z}^2}-\frac{1}{2\alpha\tilde{\lambda}_{\text{eff}}}\right) \qquad (21)
\end{aligned}
$$

and

$$
\begin{aligned}
& F\left(\sqrt{4\pi^2(k_1^2+k_2^2)+\tilde{\lambda}_{\text{eff}}^{-2}};\tilde{z};\alpha\right) \\
\equiv\ & \exp\left(\sqrt{4\pi^2(k_1^2+k_2^2)+\tilde{\lambda}_{\text{eff}}^{-2}}\,\tilde{z}\right)\operatorname{erfc}\left(\frac{\sqrt{4\pi^2(k_1^2+k_2^2)+\tilde{\lambda}_{\text{eff}}^{-2}}}{2\alpha}+\alpha\tilde{z}\right) \\
+\ & \exp\left(-\sqrt{4\pi^2(k_1^2+k_2^2)+\tilde{\lambda}_{\text{eff}}^{-2}}\,\tilde{z}\right)\operatorname{erfc}\left(\frac{\sqrt{4\pi^2(k_1^2+k_2^2)+\tilde{\lambda}_{\text{eff}}^{-2}}}{2\alpha}-\alpha\tilde{z}\right). \qquad (22)
\end{aligned}
$$

In these formulas, erfc is the complementary error function and the free parameter α, as indicated in [31], is to be chosen in such a way that a balanced interplay of computational cost and satisfactory precision is achieved. For definiteness, we set α equal to $\tilde{\lambda}_{\text{eff}}$.

In the forthcoming section, we will compare three expressions and present the optimum formula in view of its efficiency in use for numerical analysis.

3. Results

3.1. Gravitational Potentials

Formulas (17), (18) and (20) describe the gravitational potential due to a point-like body, with mass m, placed at $(x, y, z) = (0, 0, 0)$, and by the accompanying images placed at $(x, y, z) = (k_1 l, k_2 l, 0)$, where $k_{1,2} = 0, \pm 1, \pm 2, \ldots$ All three forms of the rescaled potential are composed of infinite series. Hence, for any desired precision, one needs to determine the minimum number of terms that are needed to numerically calculate the potential. The criterion that we use to specify this number n is the following: the ratio $|\text{exact}\ \tilde{\Phi} - \text{approximate}\ \tilde{\Phi}| / |\text{exact}\ \tilde{\Phi}|$ should be less than 0.001. This defines the order of accuracy in our analysis. Evidently, for each form of the potential the number n can be different, so we denote these as n_{exp}, n_{cos} and n_{mix}, correspondingly. The formula requiring the smallest number of terms to define $\tilde{\Phi}$ (i.e., to calculate "approximate $\tilde{\Phi}$") up to the adopted accuracy is clearly the best alternative for numerical computation purposes. In this connection, we are interested in comparing (17), (18) and (20) here. Because the formulas include double series, the accompanying numbers n are to be ascribed the smallest possible number of combinations (k_1, k_2) that can provide the necessary precision. We find these by listing the summands in increasing order for $\sqrt{k_1^2 + k_2^2}$ and assigning to n the total number of terms that are included in the list eventually. This procedure of generating a sequence of combinations (k_1, k_2) and finding n is performed using Mathematica [32].

Tables 1 and 2 show the outputs for eight selected points. As to the adopted accuracy, for all $n \geqslant n_{\text{exp}}$, the approximate value of $\tilde{\Phi}_{\text{exp}}$ (calculated by (18)) differs from the exact value by less than one tenth of a percent. In both tables, the exact value $\tilde{\Phi}$ is calculated from (18) for $n \gg n_{\text{exp}}$. The quantities n_{cos} and n_{mix} indicate the numbers of terms in formulas (17) and (20), respectively, which one needs to keep in order to obtain the same values of the potential at the selected points with the same precision as attained by using (18). In the n_{cos} column, the dash reflects either incorrect outputs that are produced because of complications in the computational process, or the fact that unreasonably large number of summands is necessary. Because our results depend on the ratio of λ_{eff} to the physical size al of the periods of tori, i.e., $\tilde{\lambda}_{\text{eff}} = \lambda_{\text{eff}}/(al)$, we present the results that were obtained for small and large values of $\tilde{\lambda}_{\text{eff}}$ separately in Tables 1 and 2, which include $\tilde{\lambda}_{\text{eff}} = 0.01, 0.1$ and $\tilde{\lambda}_{\text{eff}} = 1, 3$, respectively.

Table 1. Potentials $\tilde{\Phi}$ as well as the numbers $n_{\text{exp}}, n_{\text{cos}}$ and n_{mix} at a selection of points for $\tilde{\lambda}_{\text{eff}} = 0.01$ and $\tilde{\lambda}_{\text{eff}} = 0.1$ in the left and right tables, respectively.

	$\tilde{x}$	$\tilde{y}$	$\tilde{z}$	$\tilde{\Phi}$	n_{exp}	n_{cos}	n_{mix}		$\tilde{x}$	$\tilde{y}$	$\tilde{z}$	$\tilde{\Phi}$	n_{exp}	n_{cos}	n_{mix}
A_1	0.5	0	0.5	5.524×10^{-31}	2	1007	2	A_1	0.5	0	0.5	2.418×10^{-3}	7	40	7
A_2	0.5	0	0.1	2.810×10^{-22}	2	—	2	A_2	0.5	0	0.1	2.398×10^{-2}	6	808	6
A_3	0.5	0	0	7.715×10^{-22}	2	—	2	A_3	0.5	0	0	2.700×10^{-2}	4	—	4
B_1	0.1	0	0.5	1.405×10^{-22}	1	187	1	B_1	0.1	0	0.5	1.203×10^{-2}	4	28	4
B_2	0.1	0	0.1	5.101×10^{-6}	1	2119	1	B_2	0.1	0	0.1	1.719	1	380	1
B_3	0.1	0	0	4.540×10^{-4}	1	—	1	B_3	0.1	0	0	3.679	1	—	1
C_1	0	0	0.5	3.857×10^{-22}	1	236	1	C_1	0	0	0.5	1.353×10^{-2}	4	37	4
C_2	0	0	0.1	4.540×10^{-4}	1	1479	1	C_2	0	0	0.1	3.679	1	490	1

According to these tables, both expressions (18) and (20) seem preferable for numerical calculations in the case $\tilde{\lambda}_{\text{eff}} < 1$ since $n_{\text{exp}}, n_{\text{mix}} \ll n_{\text{cos}}$, although Equation (20) is, of course, much more complicated than Equation (18), and the computation of its every single summand takes longer. However, for $\tilde{\lambda}_{\text{eff}} \gtrsim 1$, the Yukawa–Ewald formula (20) alone becomes superior to the remaining two and the distinction grows as $\tilde{\lambda}_{\text{eff}}$ becomes larger. According to Planck 2013 data [2], the lower limit on the periods of tori (in the case of chimney topology) is ∼20 Gpc. Meanwhile, the current value of the effective cosmological screening length, as indicated in [29], is 2.6 Gpc. Thus, the region that is defined by $\tilde{\lambda}_{\text{eff}} < 1$ depicts the observable Universe and, here, as we have just discussed, Equation (18) is more convenient for numerical analysis.

Table 2. Potentials $\tilde{\Phi}$ as well as the numbers $n_{\text{exp}}, n_{\text{cos}}$ and n_{mix} in a selection of points for $\tilde{\lambda}_{\text{eff}} = 1$ and $\tilde{\lambda}_{\text{eff}} = 3$ in the left and right tables, respectively.

	$\tilde{x}$	$\tilde{y}$	$\tilde{z}$	$\tilde{\Phi}$	n_{exp}	n_{cos}	n_{mix}		$\tilde{x}$	$\tilde{y}$	$\tilde{z}$	$\tilde{\Phi}$	n_{exp}	n_{cos}	n_{mix}
A_1	0.5	0	0.5	3.783	174	9	15	A_1	0.5	0	0.5	15.93	1418	7	6
A_2	0.5	0	0.1	5.067	163	229	15	A_2	0.5	0	0.1	17.60	1379	120	9
A_3	0.5	0	0	5.153	163	—	15	A_3	0.5	0	0	17.71	1377	—	9
B_1	0.1	0	0.5	3.990	171	10	13	B_1	0.1	0	0.5	16.14	1411	8	6
B_2	0.1	0	0.1	9.405	133	164	11	B_2	0.1	0	0.1	22.00	1290	138	9
B_3	0.1	0	0	12.34	123	—	10	B_3	0.1	0	0	24.96	1242	—	9
C_1	0	0	0.5	4.014	170	13	13	C_1	0	0	0.5	16.17	1410	8	7
C_2	0	0	0.1	12.30	123	357	9	C_2	0	0	0.1	24.91	1243	286	9

Concluding this section, we also present Figures 1–4, demonstrating the shape of the rescaled potential $\tilde{\Phi}$ for the same values of $\tilde{\lambda}_{\text{eff}}$ as those picked for Tables 1 and 2. To plot these figures (using Mathematica [32]), we use (18) for $n \gg n_{\text{exp}}$.

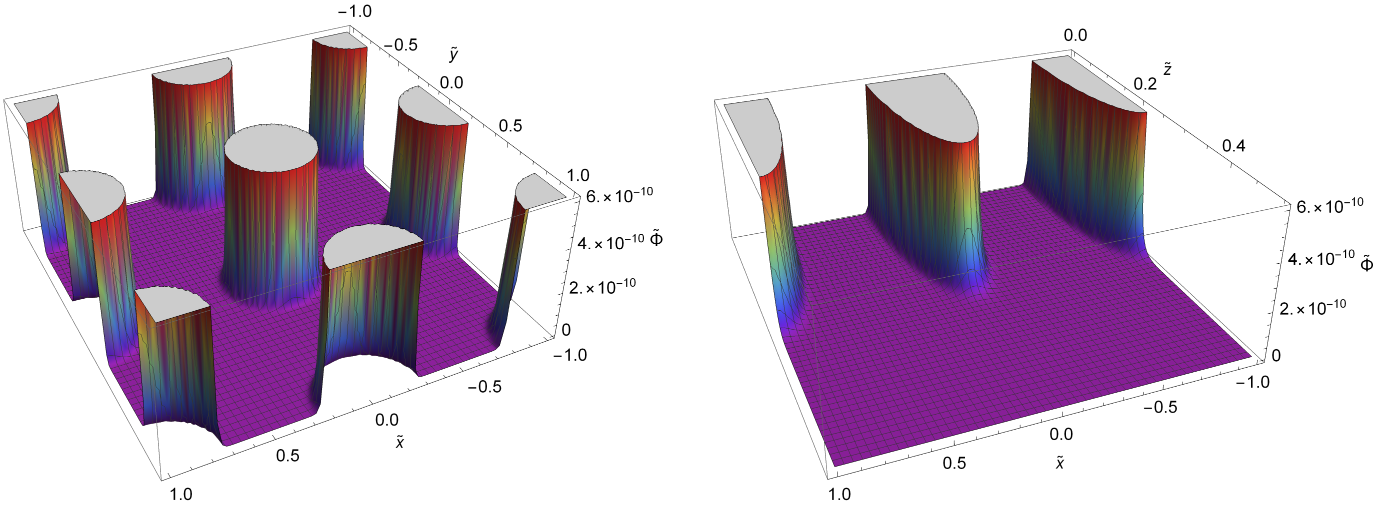

Figure 1. $\tilde{\Phi} = \left[-G_N m/(c^2 al)\right]^{-1}\widehat{\Phi}$ for $\tilde{\lambda}_{\text{eff}} = 0.01$ for the sections $z = 0$ (left panel) and $y = 0$ (right panel).

3.2. Gravitational Forces

It is also interesting to study the forces (per unit mass) associated with the alternative forms of the gravitational potential derived in the previous section. We intend to consider the projections of these forces on the x- and z-axes. Owing to the symmetry of the model, the x and y projections are similar. We calculate the gravitational forces for the same points as for the potentials and, among these, naturally, we only investigate the points at which projections on the axis of interest are nonzero. In this connection, the points A_1, A_2, A_3, C_1, C_2 and the points A_3, B_3 are omitted for the x- and z-components, respectively. The accuracy of force calculations is of the same level as that of the potentials. Here, once again, we compare the number of terms needed to achieve this accuracy, but now for three different forms of the gravitational force presentation.

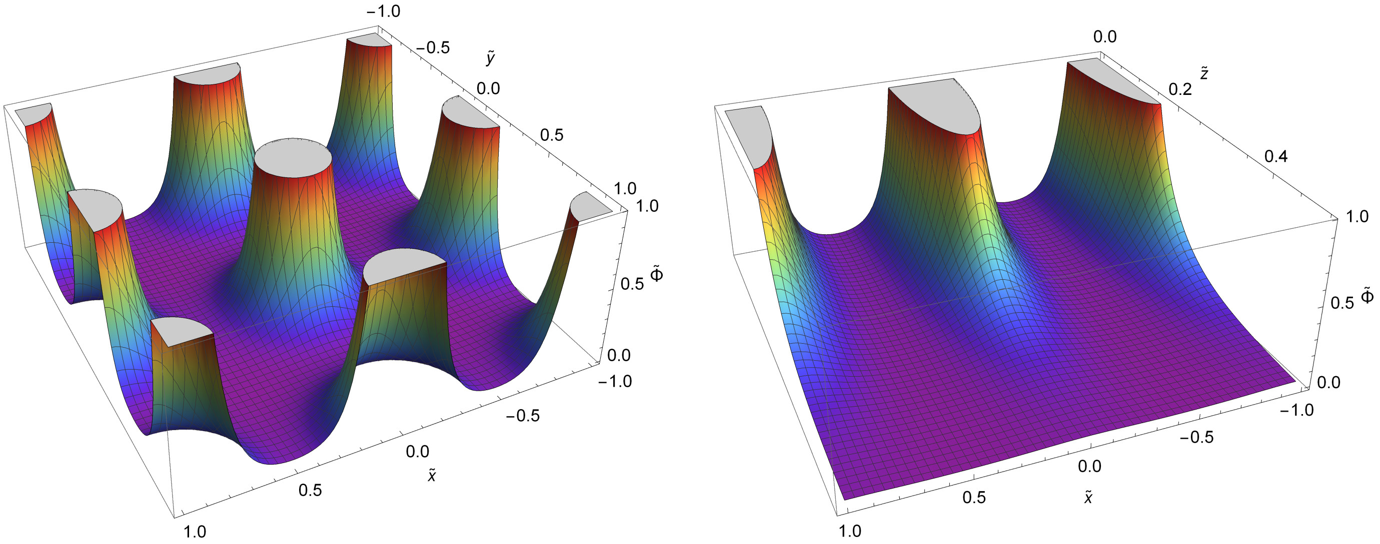

Figure 2. $\tilde{\Phi} = \left[-G_N m/(c^2 a l)\right]^{-1}\widehat{\Phi}$ for $\tilde{\lambda}_{\rm eff} = 0.1$ for the sections $z = 0$ (left panel) and $y = 0$ (right panel).

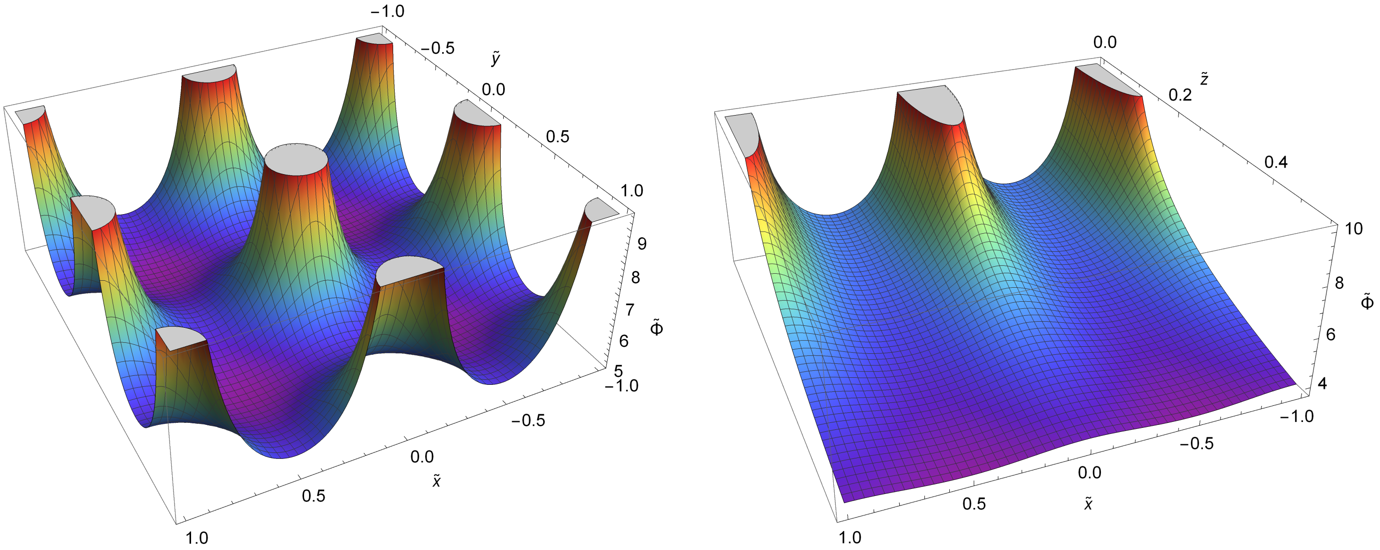

Figure 3. $\tilde{\Phi} = \left[-G_N m/(c^2 a l)\right]^{-1}\widehat{\Phi}$ for $\tilde{\lambda}_{\rm eff} = 1$ for the sections $z = 0$ (left panel) and $y = 0$ (right panel).

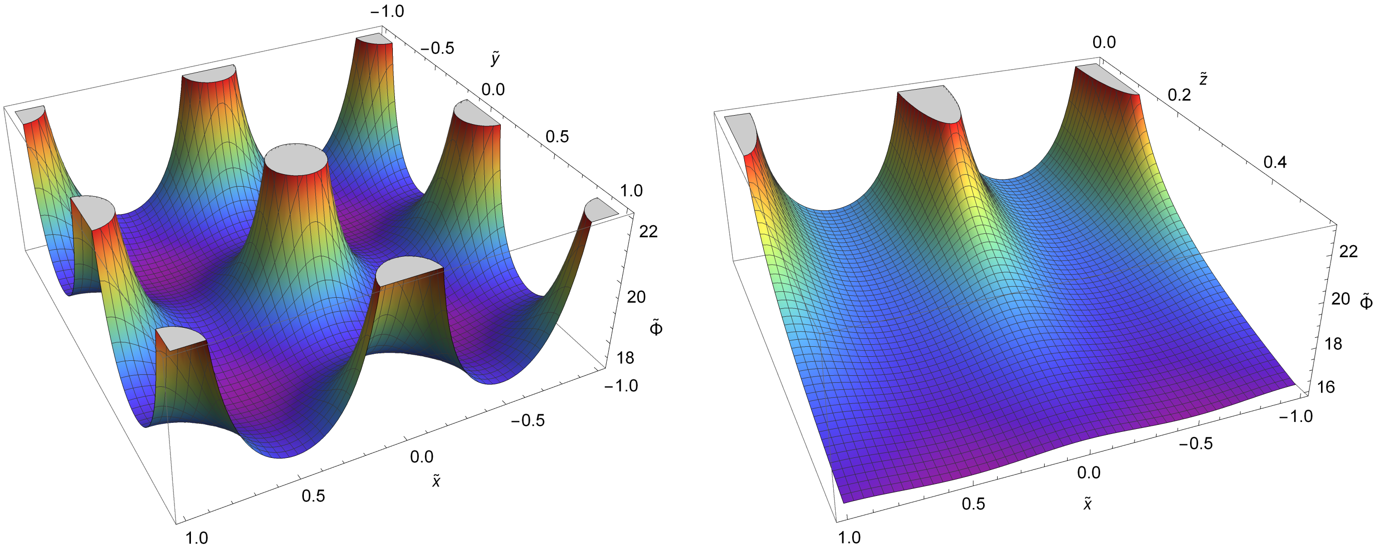

Figure 4. $\tilde{\Phi} = \left[-G_N m/(c^2 a l)\right]^{-1}\widehat{\Phi}$ for $\tilde{\lambda}_{\rm eff} = 3$ for the sections $z = 0$ (left panel) and $y = 0$ (right panel).

3.2.1. x-Component of the Gravitational Force

From (17), (18) and (20), we derive three alternative expressions for the x-component of the rescaled force:

$$\begin{aligned}\frac{\partial}{\partial\tilde{x}}\left(\tilde{\Phi}_{\cos}\right) &= -2\pi\sum_{k_1=-\infty}^{+\infty}\sum_{k_2=-\infty}^{+\infty} k_1^2+k_2^2+\frac{1}{4\pi^2\tilde{\lambda}_{\rm eff}^2}\Bigg)^{-1/2}\\ &\times\ \exp\ -\sqrt{4\pi^2\left(k_1^2+k_2^2\right)+\frac{1}{\tilde{\lambda}_{\rm eff}^2}}|\tilde{z}|\Bigg)\,k_1\sin(2\pi k_1\tilde{x})\cos(2\pi k_2\tilde{y})\,,\end{aligned}\tag{23}$$

$$\begin{aligned}\frac{\partial}{\partial\tilde{x}}\left(\tilde{\Phi}_{\exp}\right) &= -\sum_{k_1=-\infty}^{+\infty}\sum_{k_2=-\infty}^{+\infty}\exp\left(-\frac{\sqrt{(\tilde{x}-k_1)^2+(\tilde{y}-k_2)^2+\tilde{z}^2}}{\tilde{\lambda}_{\rm eff}}\right)\\ &\times\left[\frac{\tilde{x}-k_1}{\left[(\tilde{x}-k_1)^2+(\tilde{y}-k_2)^2+\tilde{z}^2\right]^{3/2}}+\frac{\tilde{x}-k_1}{\tilde{\lambda}_{\rm eff}\left[(\tilde{x}-k_1)^2+(\tilde{y}-k_2)^2+\tilde{z}^2\right]}\right],\end{aligned}\tag{24}$$

$$\begin{aligned}\frac{\partial}{\partial\tilde{x}}\left(\tilde{\Phi}_{\rm mix}\right) &= -\frac{1}{2}\sum_{k_1=-\infty}^{+\infty}\sum_{k_2=-\infty}^{+\infty}\left[\frac{(\tilde{x}-k_1)D\left(\sqrt{(\tilde{x}-k_1)^2+(\tilde{y}-k_2)^2+\tilde{z}^2};\alpha;\tilde{\lambda}_{\rm eff}\right)}{\left[(\tilde{x}-k_1)^2+(\tilde{y}-k_2)^2+\tilde{z}^2\right]^{3/2}}\right.\\ &+\frac{\tilde{x}-k_1}{(\tilde{x}-k_1)^2+(\tilde{y}-k_2)^2+\tilde{z}^2}\cdot C_-\exp\left(-\frac{\sqrt{(\tilde{x}-k_1)^2+(\tilde{y}-k_2)^2+\tilde{z}^2}}{\tilde{\lambda}_{\rm eff}}\right)\\ &+\frac{\tilde{x}-k_1}{(\tilde{x}-k_1)^2+(\tilde{y}-k_2)^2+\tilde{z}^2}\cdot C_+\exp\left(\frac{\sqrt{(\tilde{x}-k_1)^2+(\tilde{y}-k_2)^2+\tilde{z}^2}}{\tilde{\lambda}_{\rm eff}}\right)\\ &\left.+\ 4\pi^2 k_1\sin[2\pi(k_1\tilde{x}+k_2\tilde{y})]\frac{F\left(\sqrt{4\pi^2\left(k_1^2+k_2^2\right)+\tilde{\lambda}_{\rm eff}^{-2}};\tilde{z};\alpha\right)}{\sqrt{4\pi^2\left(k_1^2+k_2^2\right)+\tilde{\lambda}_{\rm eff}^{-2}}}\right],\end{aligned}\tag{25}$$

where

$$\begin{aligned} C_{\mp} &= C_{\mp}\left(\sqrt{(\tilde{x}-k_1)^2+(\tilde{y}-k_2)^2+\tilde{z}^2};\alpha;\tilde{\lambda}_{\rm eff}\right) \\ &\equiv \frac{2\alpha}{\sqrt{\pi}}\exp\left[-\left(\alpha\sqrt{(\tilde{x}-k_1)^2+(\tilde{y}-k_2)^2+\tilde{z}^2}\mp\frac{1}{2\alpha\tilde{\lambda}_{\rm eff}}\right)^2\right] \\ &\pm \frac{1}{\tilde{\lambda}_{\rm eff}}\operatorname{erfc}\left(\alpha\sqrt{(\tilde{x}-k_1)^2+(\tilde{y}-k_2)^2+\tilde{z}^2}\mp\frac{1}{2\alpha\tilde{\lambda}_{\rm eff}}\right). \end{aligned} \tag{26}$$

We present the results of the calculations that were performed in Mathematica [32] in Tables 3 and 4 for $\tilde{\lambda}_{\rm eff} = 0.01, 0.1$ and $\tilde{\lambda}_{\rm eff} = 1, 3$, respectively. As in the case for the gravitational potential, a straightforward analysis shows that Formulas (24) and (25) that are related to the Yukawa and Yukawa–Ewald potentials, respectively, are preferable over (23) for the physically significant case $\tilde{\lambda}_{\rm eff} < 1$ (although the structure of the expression (25) is again much more complicated compared to (24)). Meanwhile, when $\tilde{\lambda}_{\rm eff} \gtrsim 1$, the Yukawa–Ewald force becomes superior. In both tables, Equation (24) was employed for $n \gg n_{\rm exp}$ while computing the values of the rescaled x-component $\tilde{\Phi}_x$.

Table 3. Numerical values of the x-component of the rescaled force $\tilde{\Phi}_x \equiv \partial\tilde{\Phi}/\partial\tilde{x}$ as well as the numbers $n_{\rm exp}, n_{\rm cos}$ and $n_{\rm mix}$ for points B_1, B_2 and B_3 for $\tilde{\lambda}_{\rm eff} = 0.01$ and $\tilde{\lambda}_{\rm eff} = 0.1$ in the left and right tables, respectively.

	$\tilde{x}$	$\tilde{y}$	$\tilde{z}$	$\tilde{\Phi}_x$	$n_{\rm exp}$	$n_{\rm cos}$	$n_{\rm mix}$		$\tilde{x}$	$\tilde{y}$	$\tilde{z}$	$\tilde{\Phi}_x$	$n_{\rm exp}$	$n_{\rm cos}$	$n_{\rm mix}$
B_1	0.1	0	0.5	-2.810×10^{-21}	1	263	1	B_1	0.1	0	0.5	-2.783×10^{-2}	5	54	5
B_2	0.1	0	0.1	-3.862×10^{-4}	1	2448	1	B_2	0.1	0	0.1	-20.75	1	592	1
B_3	0.1	0	0	-4.994×10^{-2}	1	—	1	B_3	0.1	0	0	-73.57	1	—	1

Table 4. Numerical values of the x-component of the rescaled force $\tilde{\Phi}_x \equiv \partial\tilde{\Phi}/\partial\tilde{x}$ as well as the numbers $n_{\rm exp}, n_{\rm cos}$ and $n_{\rm mix}$ for points B_1, B_2 and B_3 for $\tilde{\lambda}_{\rm eff} = 1$ and $\tilde{\lambda}_{\rm eff} = 3$ in the left and right tables, respectively.

	$\tilde{x}$	$\tilde{y}$	$\tilde{z}$	$\tilde{\Phi}_x$	$n_{\rm exp}$	$n_{\rm cos}$	$n_{\rm mix}$		$\tilde{x}$	$\tilde{y}$	$\tilde{z}$	$\tilde{\Phi}_x$	$n_{\rm exp}$	$n_{\rm cos}$	$n_{\rm mix}$
B_1	0.1	0	0.5	-4.730×10^{-1}	130	38	21	B_1	0.1	0	0.5	-4.920×10^{-1}	862	38	21
B_2	0.1	0	0.1	-34.65	20	553	9	B_2	0.1	0	0.1	-34.88	77	552	13
B_3	0.1	0	0	-99.14	19	—	8	B_3	0.1	0	0	-99.49	34	—	9

Additionally, we present Figures 5–8, demonstrating the x-component of the rescaled force $\tilde{\Phi}_x$ for the same values of $\tilde{\lambda}_{\rm eff}$ as those picked for Tables 3 and 4. To plot these figures (using Mathematica [32]), we employ the Formula (24) for $n \gg n_{\rm exp}$.

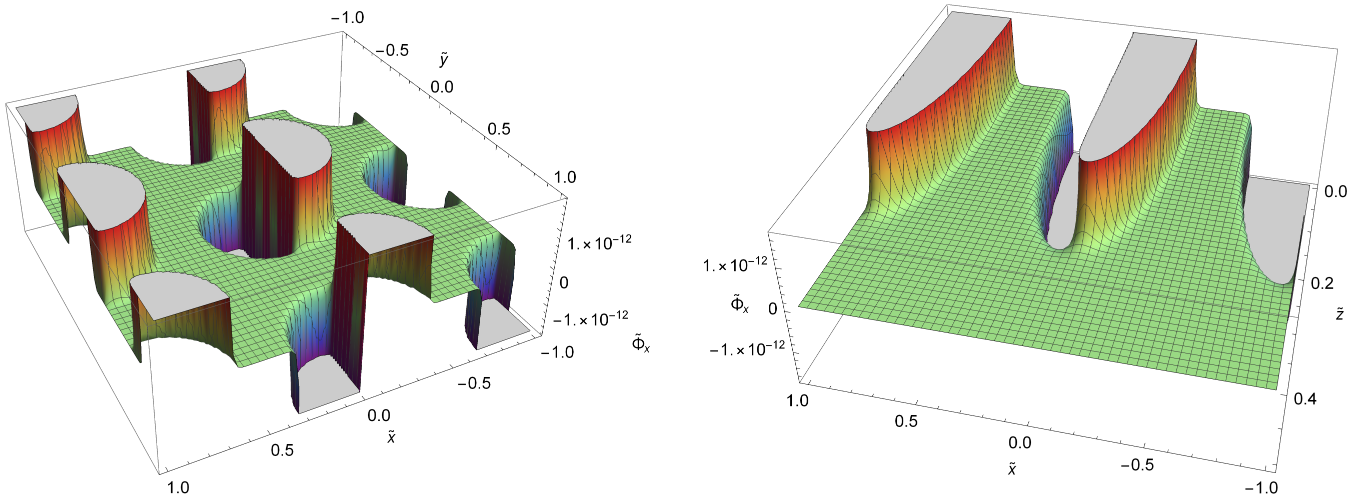

Figure 5. x-component of the rescaled force $\tilde{\Phi}_x$ for $\tilde{\lambda}_{\text{eff}} = 0.01$ for the sections $z = 0$ (left panel) and $y = 0$ (right panel).

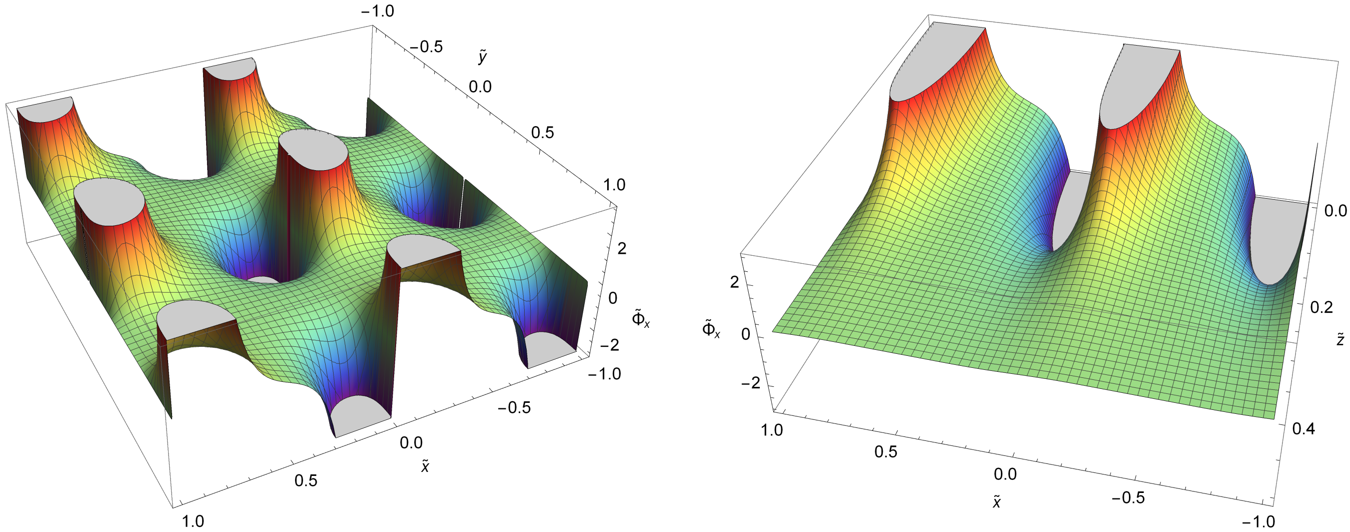

Figure 6. x-component of the rescaled force $\tilde{\Phi}_x$ for $\tilde{\lambda}_{\text{eff}} = 0.1$ for the sections $z = 0$ (left panel) and $y = 0$ (right panel).

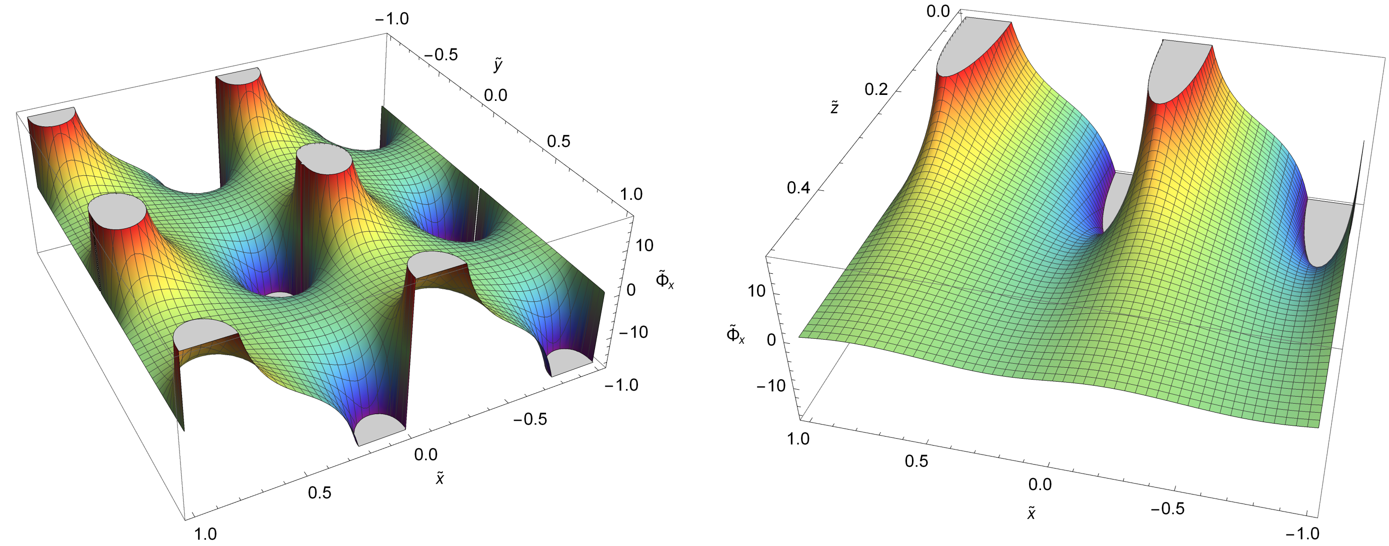

Figure 7. x-component of the rescaled force $\tilde{\Phi}_x$ for $\tilde{\lambda}_{\text{eff}} = 1$ for the sections $z = 0$ (left panel) and $y = 0$ (right panel).

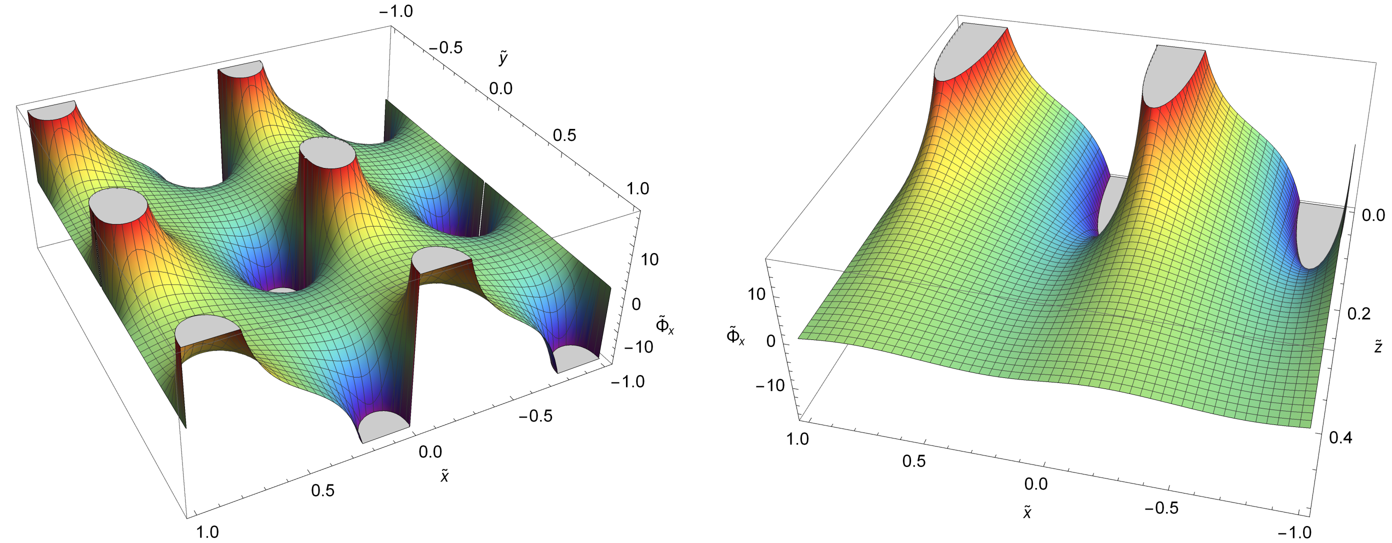

Figure 8. x-component of the rescaled force $\tilde{\Phi}_x$ for $\tilde{\lambda}_{\text{eff}} = 3$ for the sections $z = 0$ (left panel) and $y = 0$ (right panel).

3.2.2. z-Component of the Gravitational Force

For the z-component of the rescaled force, three alternative formulas are:

$$\begin{aligned} \frac{\partial}{\partial \tilde{z}}\left(\tilde{\Phi}_{\cos}\right) &= -2\pi \sum_{k_1=-\infty}^{+\infty} \sum_{k_2=-\infty}^{+\infty} \exp \left. -\sqrt{4\pi^2\left(k_1^2+k_2^2\right)+\frac{1}{\tilde{\lambda}_{\text{eff}}^2}}\,\tilde{z}\right) \\ &\times \cos(2\pi k_1 \tilde{x}) \cos(2\pi k_2 \tilde{y})\,, \end{aligned} \tag{27}$$

where, for simplicity, $\tilde{z} > 0$,

$$\begin{aligned} \frac{\partial}{\partial \tilde{z}}\left(\tilde{\Phi}_{\exp}\right) &= -\sum_{k_1=-\infty}^{+\infty} \sum_{k_2=-\infty}^{+\infty} \exp\left(-\frac{\sqrt{(\tilde{x}-k_1)^2+(\tilde{y}-k_2)^2+\tilde{z}^2}}{\tilde{\lambda}_{\text{eff}}}\right) \\ &\times \left[\frac{\tilde{z}}{\left[(\tilde{x}-k_1)^2+(\tilde{y}-k_2)^2+\tilde{z}^2\right]^{3/2}} + \frac{\tilde{z}}{\tilde{\lambda}_{\text{eff}}\left[(\tilde{x}-k_1)^2+(\tilde{y}-k_2)^2+\tilde{z}^2\right]}\right] \end{aligned} \tag{28}$$

and

$$
\begin{aligned}
&\frac{\partial}{\partial \tilde{z}}\left(\tilde{\Phi}_{\mathrm{mix}}\right) = \\
-\ &\frac{1}{2}\sum_{k_1=-\infty}^{+\infty}\sum_{k_2=-\infty}^{+\infty}\left[\frac{\tilde{z}D\left(\sqrt{(\tilde{x}-k_1)^2+(\tilde{y}-k_2)^2+\tilde{z}^2};\alpha;\tilde{\lambda}_{\mathrm{eff}}\right)}{\left[(\tilde{x}-k_1)^2+(\tilde{y}-k_2)^2+\tilde{z}^2\right]^{3/2}}\right. \\
+\ &\frac{\tilde{z}}{(\tilde{x}-k_1)^2+(\tilde{y}-k_2)^2+\tilde{z}^2}\cdot C_-\exp\left(-\frac{\sqrt{(\tilde{x}-k_1)^2+(\tilde{y}-k_2)^2+\tilde{z}^2}}{\tilde{\lambda}_{\mathrm{eff}}}\right) \\
+\ &\frac{\tilde{z}}{(\tilde{x}-k_1)^2+(\tilde{y}-k_2)^2+\tilde{z}^2}\cdot C_+\exp\left(\frac{\sqrt{(\tilde{x}-k_1)^2+(\tilde{y}-k_2)^2+\tilde{z}^2}}{\tilde{\lambda}_{\mathrm{eff}}}\right) \\
-\ &2\pi\cos[2\pi(k_1\tilde{x}+k_2\tilde{y})]\frac{1}{\sqrt{4\pi^2\left(k_1^2+k_2^2\right)+\tilde{\lambda}_{\mathrm{eff}}^{-2}}} \\
\times\ &\left.\left[F_-\exp\left(-z\sqrt{4\pi^2\left(k_1^2+k_2^2\right)+\frac{1}{\tilde{\lambda}_{\mathrm{eff}}^2}}\right)+F_+\exp\ z\sqrt{4\pi^2\left(k_1^2+k_2^2\right)+\frac{1}{\tilde{\lambda}_{\mathrm{eff}}^2}}\right)\right]\right], \qquad (29)
\end{aligned}
$$

where

$$
\begin{aligned}
F_{\mp} &= F_{\mp}\left(\sqrt{4\pi^2\left(k_1^2+k_2^2\right)+\tilde{\lambda}_{\mathrm{eff}}^{-2}};\tilde{z};\alpha\right)\equiv \\
&\pm\ \frac{2\alpha}{\sqrt{\pi}}\exp\left[-\left(\frac{\sqrt{4\pi^2\left(k_1^2+k_2^2\right)+\tilde{\lambda}_{\mathrm{eff}}^{-2}}}{2\alpha}\mp\alpha\tilde{z}\right)^2\right] \\
&\mp\ \sqrt{4\pi^2\left(k_1^2+k_2^2\right)+\tilde{\lambda}_{\mathrm{eff}}^{-2}}\,\mathrm{erfc}\left(\frac{\sqrt{4\pi^2\left(k_1^2+k_2^2\right)+\tilde{\lambda}_{\mathrm{eff}}^{-2}}}{2\alpha}\mp\alpha\tilde{z}\right), \qquad (30)
\end{aligned}
$$

and $C_{\mp}$ are given by (26).

Now, we employ these formulas to calculate the nonzero z-components of the gravitational force at the previously selected set of points and, again, for the desired precision. The results that were obtained in Mathematica [32] are presented in Tables 5 and 6, which show that while $\tilde{\lambda}_{\mathrm{eff}} < 1$, depicting well the observational restrictions, two Formulas (28) and (29) are favorable (as before, the latter is much more cumbersome). On the other hand, for $\tilde{\lambda}_{\mathrm{eff}} \gtrsim 1$, the Yukawa–Ewald Formula (29) gives the best results. Herein, the quantity $\tilde{\Phi}_z$ stands for the z-component of the rescaled force, calculated from Equation (28) for $n \gg n_{\mathrm{exp}}$. We depict the behavior of this component in Figures 9 and 10 for the section $y = 0$. Obviously, the projection of the gravitational force on the z-axis is absent for the section $z = 0$ due to the symmetry of the model.

Table 5. Numerical values of the z-component of the rescaled force $\tilde{\Phi}_z \equiv \partial\tilde{\Phi}/\partial\tilde{z}$ as well as the numbers $n_{\mathrm{exp}}, n_{\mathrm{cos}}$ and n_{mix} for points A_1, A_2, B_1, B_2, C_1 and C_2 for $\tilde{\lambda}_{\mathrm{eff}} = 0.01$ and $\tilde{\lambda}_{\mathrm{eff}} = 0.1$ in the left and right tables, respectively.

	$\tilde{x}$	$\tilde{y}$	$\tilde{z}$	$\tilde{\Phi}_z$	n_{exp}	n_{cos}	n_{mix}		$\tilde{x}$	$\tilde{y}$	$\tilde{z}$	$\tilde{\Phi}_z$	n_{exp}	n_{cos}	n_{mix}
A_1	0.5	0	0.5	-3.962×10^{-29}	2	1070	2	A_1	0.5	0	0.5	-1.946×10^{-2}	6	47	6
A_2	0.5	0	0.1	-5.620×10^{-21}	2	—	2	A_2	0.5	0	0.1	-5.620×10^{-2}	2	1647	2
B_1	0.1	0	0.5	-1.405×10^{-20}	1	187	1	B_1	0.1	0	0.5	-1.407×10^{-1}	2	33	2
B_2	0.1	0	0.1	-3.862×10^{-4}	1	2228	1	B_2	0.1	0	0.1	-20.75	1	649	1
C_1	0	0	0.5	-3.935×10^{-20}	1	240	1	C_1	0	0	0.5	-1.620×10^{-1}	3	44	3
C_2	0	0	0.1	-4.994×10^{-2}	1	1620	1	C_2	0	0	0.1	-73.58	1	722	1

Table 6. Numerical values of the z-component of the rescaled force $\tilde{\Phi}_z \equiv \partial\tilde{\Phi}/\partial\tilde{z}$ as well as the numbers $n_{\mathrm{exp}}, n_{\mathrm{cos}}$ and n_{mix} for points A_1, A_2, B_1, B_2, C_1 and C_2 for $\tilde{\lambda}_{\mathrm{eff}} = 1$ and $\tilde{\lambda}_{\mathrm{eff}} = 3$ in the left and right tables, respectively.

	$\tilde{x}$	$\tilde{y}$	$\tilde{z}$	$\tilde{\Phi}_z$	n_{exp}	n_{cos}	n_{mix}		$\tilde{x}$	$\tilde{y}$	$\tilde{z}$	$\tilde{\Phi}_z$	n_{exp}	n_{cos}	n_{mix}
A_1	0.5	0	0.5	−3.571	85	21	15	A_1	0.5	0	0.5	−5.072	444	21	11
A_2	0.5	0	0.1	−1.673	64	900	15	A_2	0.5	0	0.1	−2.037	331	863	13
B_1	0.1	0	0.5	−5.045	74	20	13	B_1	0.1	0	0.5	−6.593	397	19	9
B_2	0.1	0	0.1	−35.48	15	444	5	B_2	0.1	0	0.1	−36.04	57	444	8
C_1	0	0	0.5	−5.241	73	26	13	C_1	0	0	0.5	−6.795	392	24	12
C_2	0	0	0.1	−99.97	8	678	4	C_2	0	0	0.1	−100.7	21	677	7

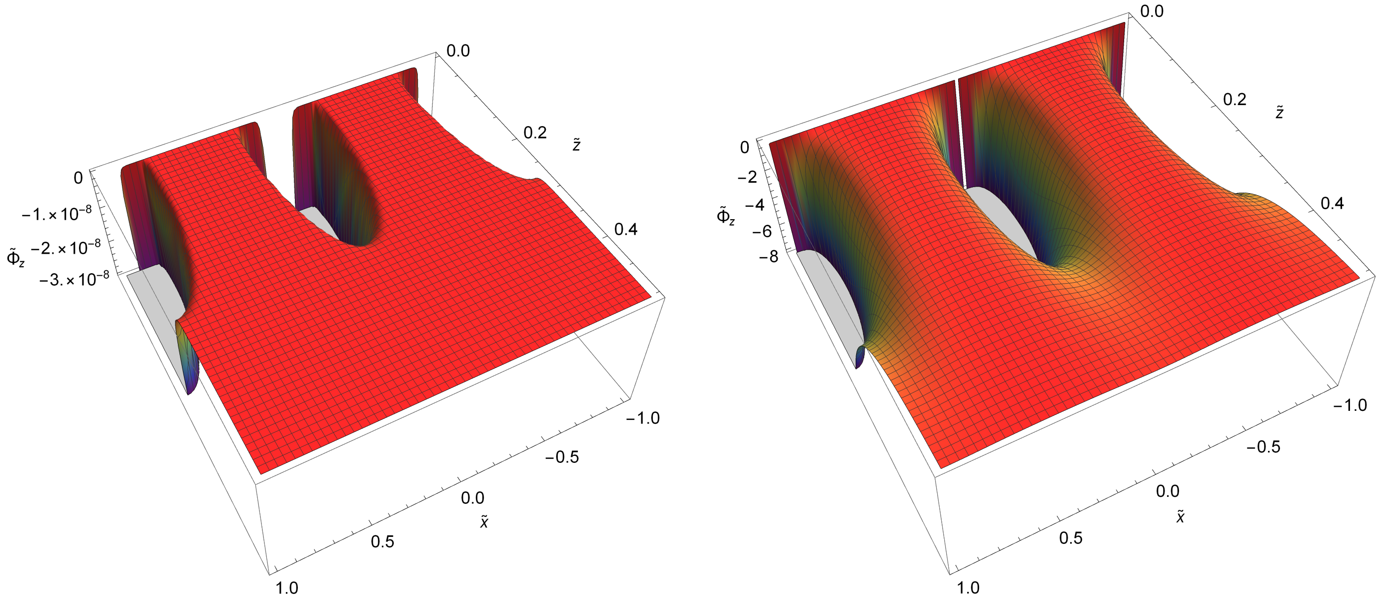

Figure 9. z-component of the rescaled force $\tilde{\Phi}_z$ for $\tilde{\lambda}_{\mathrm{eff}} = 0.01$ and $\tilde{\lambda}_{\mathrm{eff}} = 0.1$ (left and right panels, respectively).

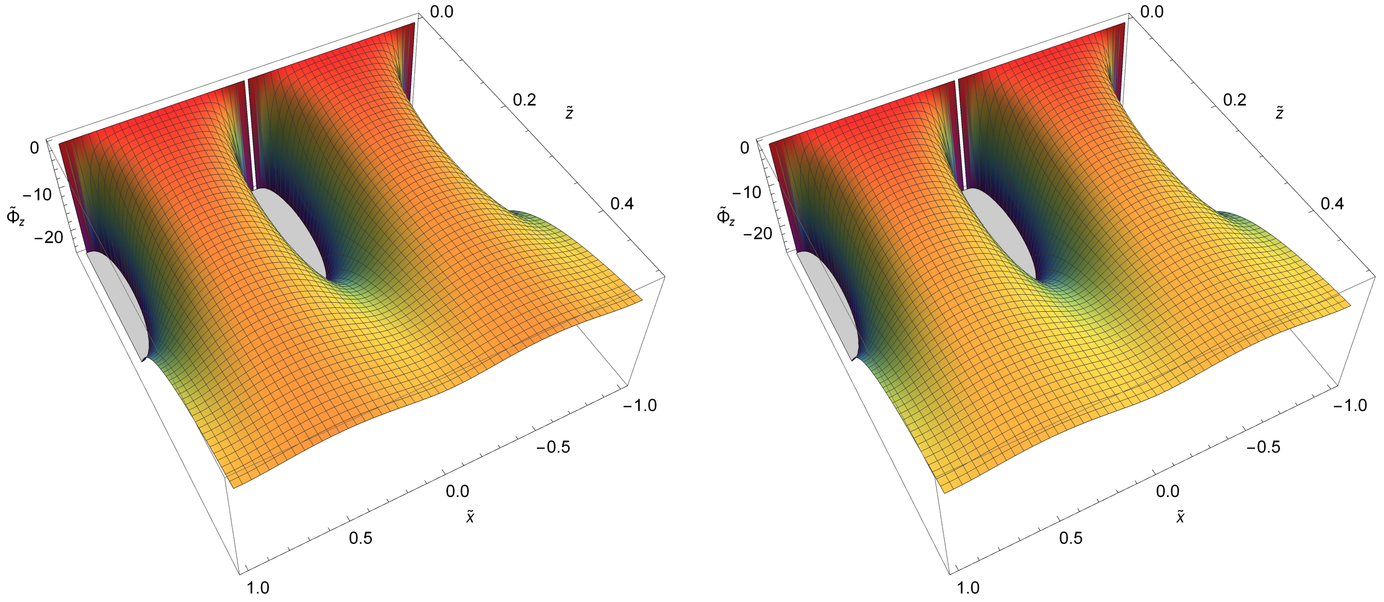

Figure 10. z-component of the rescaled force $\tilde{\Phi}_z$ for $\tilde{\lambda}_{\mathrm{eff}} = 1$ and $\tilde{\lambda}_{\mathrm{eff}} = 3$ (left and right panels, respectively).

4. Conclusions

In this work, we have studied how the chimney topology $T \times T \times R$ of the Universe affects the form of the gravitational potential and, consequently, that of the gravitational force. In this connection, we have proposed three alternative forms for each of the solutions. One of them (see Equation (17)) relies on the Fourier expansion of the delta functions into series while using periodicity in two toroidal dimensions in the model. The second one (see Equation (18)) follows from the summation of solutions of the Helmholtz equation, each in the form of the Yukawa potential, for a source mass and all of its periodic images. Finally, the third form of the potential (see Equation (20)) is formulated via the Ewald sums for Yukawa potentials. Subsequently, we have presented three alternative forms of the gravitational force (see Equations (23)–(25) and (27)–(29) for the x- and z-components, respectively) derived from the potential expressions.

In all three alternative forms, the screening length $\tilde{\lambda}_{\text{eff}}$ serves as a crucial parameter, as it specifies the distance (from the source or the periodic images) where the gravitational potential undergoes exponential cutoff. This fact is most clearly demonstrated in the case where the solution takes on the form of summed Yukawa potentials (18). The observational data show that this parameter should be less than 1 ($\tilde{\lambda}_{\text{eff}} < 1$) in today's Universe.

One of the main goals of this work was to reveal which of the obtained alternative formulas would serve better as a tool to be employed in numerical calculations. Namely, to show which formula would require less terms in the series to reach the desired precision. Our calculations have demonstrated that, for both the gravitational potentials and forces, two expressions involving plain summations of Yukawa potentials are preferable for the physically significant case $\tilde{\lambda}_{\text{eff}} < 1$. However, in the case $\tilde{\lambda}_{\text{eff}} \gtrsim 1$, the Yukawa–Ewald presentation stands out as the best alternative.

Additionally, we have produced Figures 1–10 for $\tilde{\lambda}_{\text{eff}} = 0.01, 0.1, 1, 3$ to provide graphical demonstration of the gravitational potentials and force projections.

Author Contributions: Conceptualization, M.E.; methodology, M.E., E.C. and A.Z.; formal analysis, M.E., A.M.II, E.C., M.B. and A.Z.; investigation, M.E., A.M.II, E.C., M.B. and A.Z.; writing—original draft preparation, A.Z.; writing—review and editing, M.E. and E.C.; visualization, M.E. and A.M.II; supervision, M.E. and A.Z.; project administration, M.E.; funding acquisition, M.E. All authors have read and agreed to the published version of the manuscript.

Funding: The work of M.E. and A.M.II was supported by the National Science Foundation HRD Award number 1954454.

Conflicts of Interest: The authors have no conflicts of interest to declare that are relevant to the content of this article.

References

1. Gannon, D. Singularities in nonsimply connected space-times. *J. Math. Phys.* **1975**, *16*, 2364. [CrossRef]
2. Ade, P.A.R.; et al. [Planck Collaboration]. Planck 2013 results. XXVI. Background geometry and topology of the Universe. *Astron. Astrophys.* **2014**, *571*, A26.
3. Luminet, J.-P. The shape and topology of the Universe. *arXiv* **2008**, arXiv:0802.2236.
4. Roukema, B.F. Topology of the Universe: Background and recent observational approaches. *Pramana* **1999**, *53*, 945–949. [CrossRef]
5. Piechocki, W. Topology of the Universe. *arXiv* **1999**, arXiv:gr-qc/9910055.
6. Bielewicz, P.; Riazuelo, A. The study of topology of the universe using multipole vectors. *Mon. Not. R. Astron. Soc.* **2009**, *396*, 609–623. [CrossRef]
7. Bielewicz, P.; Banday, A.J.; Gorski, K.M. Constraining the topology of the Universe using the polarized CMB maps. *Mon. Not. R. Astron. Soc.* **2012**, *421*, 1064–1072. [CrossRef]
8. Vaudrevange, P.M.; Starkman, G.D.; Cornish, N.J.; Spergel, D.N. Constraints on the topology of the Universe: Extension to general geometries. *Phys. Rev. D* **2012**, *86*, 083526. [CrossRef]
9. Fabre, O.; Prunet, S.; Uzan, J.-P. Topology beyond the horizon: How far can it be probed? *Phys. Rev. D* **2015**, *92*, 043003. [CrossRef]

10. Bucher, M.; et al. [Planck Collaboration]. Planck 2015 results. XVIII. Background geometry and topology. *Astron. Astrophys.* **2016**, *594*, A18.
11. Bielewicz, P.; Banday, A.J. Constraints on the topology of the Universe derived from the 7-year WMAP data. *Mon. Not. R. Astron. Soc.* **2011**, *412*, 2104–2110. [CrossRef]
12. Bielewicz, P.; Banday, A.J.; Gorski, K.M. Constraints on the topology of the Universe derived from the 7-year WMAP CMB data and prospects of constraining the topology using CMB polarization maps. In Proceedings of the XLVIIth Rencontres de Moriond, La Tuile, Italy, 3–10 March 2012; Auge, E., Dumarchez, J., Tran Thanh Van, J., Eds.; ARISF: Paris, France, 2012; p. 91.
13. Luminet, J.P.; Weeks, J.R.; Riazuelo, A.; Lehoucq, R.; Uzan, J.P. Dodecahedral space topology as an explanation for weak wide-angle temperature correlations in the cosmic microwave background. *Nature* **2003**, *425*, 593–595. [CrossRef] [PubMed]
14. Aslanyan, G.; Manohar, A.V. The topology and size of the Universe from the Cosmic Microwave Background. *J. Cosmol. Astropart. Phys.* **2012**, *06*, 003. [CrossRef]
15. Floratos, E.G.; Leontaris, G.K. On topological modifications of Newton's law. *J. Cosmol. Astropart. Phys.* **2012**, *2012*, 24. [CrossRef]
16. Aslanyan, G.; Manohar, A.V.; Yadav, A.P.S. The topology and size of the Universe from CMB temperature and polarization data. *J. Cosmol. Astropart. Phys.* **2013**, *2013*, 9. [CrossRef]
17. Peebles, P.J.E. *The Large-Scale Structure of the Universe*; Princeton University Press: Princeton, NJ, USA, 1980.
18. Brilenkov, M.; Eingorn, M.; Zhuk, A. Lattice Universe: Examples and problems. *Eur. Phys. J. C* **2015**, *75*, 217. [CrossRef] [PubMed]
19. Eingorn, M. First-order cosmological perturbations engendered by point-like masses. *Astrophys. J.* **2016**, *825*, 84. [CrossRef]
20. Eingorn, M.; Kiefer, C.; Zhuk, A. Scalar and vector perturbations in a universe with discrete and continuous matter sources. *J. Cosmol. Astropart. Phys.* **2016**, *2016*, 32. [CrossRef]
21. Eingorn, M.; Kiefer, C.; Zhuk, A. Cosmic screening of the gravitational interaction. *Int. J. Mod. Phys. D* **2017**, *26*, 1743012. [CrossRef]
22. Eingorn, M.; McLaughlin II, A.; Canay, E.; Brilenkov, M.; Zhuk, A. Gravitation in the space with chimney topology. In Proceedings of the 1st Electronic Conference on Universe, online, 22–28 February 2021.
23. Landau, L.D.; Lifshitz, E.M. *The Classical Theory of Fields*; Course of Theoretical Physics Series, V.2; Pergamon Press: Oxford, UK, 2000.
24. Mukhanov, V.F. *Physical Foundations of Cosmology*; Cambridge University Press: Cambridge, UK, 2005.
25. Gorbunov, D.S.; Rubakov, V.A. *Introduction to the Theory of the Early Universe: Cosmological Perturbations and Inflationary Theory*; World Scientific: Singapore, 2011.
26. Eingorn, M. Cosmological law of universal gravitation. *Int. J. Mod. Phys. D* **2017**, *26*, 1750121. [CrossRef]
27. Prudnikov, A.P.; Brychkov, Y.A.; Marichev, O.I. *Integrals and Series, Volume 2. Special Functions*; Gordon and Breach Science Publishers: New York, NY, USA, 1992.
28. Eingorn, M.; Brilenkov, M.; Vlahovic, B. Zero average values of cosmological perturbations as an indispensable condition for the theory and simulations. *Eur. Phys. J. C* **2015**, *75*, 381. [CrossRef] [PubMed]
29. Canay, E.; Eingorn, M. Duel of cosmological screening lengths. *Phys. Dark Universe* **2020**, *29*, 100565. [CrossRef]
30. Mazars, M. Ewald sums for Yukawa potentials in quasi-two-dimensional systems. *J. Chem. Phys.* **2007**, *126*, 056101. [CrossRef] [PubMed]
31. Mazars, M. Yukawa potentials in systems with partial periodic boundary conditions I: Ewald sums for quasi-two dimensional systems. *Mol. Phys.* **2007**, *105*, 1909–1925. [CrossRef]
32. Wolfram Research, Inc. *Mathematica*; Version 11.3; Wolfram Research, Inc.: Champaign, IL, USA, 2018.

Paper III

Gravitational potentials and forces in the Lattice Universe: a slab

M. Eingorn, N. O'Briant, K. Arzu, M. Brilenkov, and A. Zhuk

Published in *The European Physical Journal Plus* 136, 205 (2021); DOI: 10.1140/epjp/s13360-021-01139-y

EPJ manuscript No.
(will be inserted by the editor)

arXiv:2112.14281v1 [gr-qc] 28 Dec 2021

Gravitational potentials and forces in the Lattice Universe: a slab

Maxim Eingorn[1], Niah O'Briant[1], Katie Arzu[1], Maxim Brilenkov[2], and Alexander Zhuk[3]

[1] Department of Mathematics and Physics, North Carolina Central University,
1801 Fayetteville St., Durham, North Carolina 27707, U.S.A. e-mail: `maxim.eingorn@gmail.com`
[2] Institute of Theoretical Astrophysics, University of Oslo, Postboks 1029 Blindern, 0315, Oslo, Norway
[3] Astronomical Observatory, Odessa I.I. Mechnikov National University,
Dvoryanskaya St. 2, Odessa 65082, Ukraine

Abstract. We study the effect of the slab topology $T \times R \times R$ of the Universe on the form of gravitational potentials and forces created by point-like masses. We obtain two alternative forms of solutions: one is based on the Fourier series expansion of the delta function using the periodical property along the toroidal dimension, and another one is derived by direct summation of solutions of the Helmholtz equation for the source particle and all its images. The latter one takes the form of the sum of Yukawa-type potentials. We demonstrate that for the present Universe the latter solution is preferable for numerical calculations since it requires less terms of the series to achieve the necessary precision.

1 Introduction

The question of spatial topology of the Universe belongs to a class of fundamental open questions of cosmology and theoretical physics. What is the shape of the world we live in? Is the Universe finite or infinite? Is its spatial curvature positive, negative or exactly zero? What is the role of topology in the very early epoch (on the quantum gravity arena) as well as for the subsequent large scale structure formation? Since topology is not dictated by general relativity, there is no theoretical hint of whether space is simply connected (as assumed within the concordance cosmological model) or multiply connected. In the latter case the Universe volume can be finite even when the spatial curvature is negative or vanishing [1]. If the Universe volume is much larger than the observable one, the finiteness of the world does not become apparent in the current data. However, if the volume is not too large, it is reasonable to search for observable imprints of its shape [2]. In multiply connected space, a photon emitted by a source can travel many times through the volume resulting in multiple images of the source [3,4]. Typical representatives of multiply connected spaces are spaces with toroidal topology in one or several (maximum three) spatial directions: a slab $T \times R \times R$, an equal-sided chimney $T \times T \times R$ and a three-torus $T \times T \times T$.

The possible imprints (mainly on the CMB data) of the shape of the Universe are carefully studied in literature [5,6,7,8,9]. In particular, it is tempting to interpret the CMB anomalies observed at large angular scales (such as the quadrupole moment suppression as well as the quadrupole and octopole alignment) as topological manifestations [10, 11]. In the present paper, we consider the topology of the Universe in the form of a slab. In this case we have one finite special dimension. It is very interesting whether this direction can be interpreted as a preferred axis of the quadrupole and octopole alignment, or a so-called "axis of evil" [12] (see also [13] for other indications of its existence).

According to Planck 2013 results [1] regarding the search for conjectural topological signatures in the observational data on the CMB radiation, the following restriction is imposed on the radius R_i of the largest sphere, which can be inscribed in the topological domain: in the case of the flat Universe with the slab topology $R_i > 0.50\chi_{rec}$. Planck 2015 results [9] make the above-mentioned restriction tougher: $R_i > 0.56\chi_{rec}$. Here χ_{rec} represents the distance to the recombination surface, which is of the order of the particle horizon, *i.e.* ~ 14 Gpc. Earlier bounds on the Universe size, based on the thorough analysis of 7-year and 9-year WMAP temperature maps, can be found in [12,14]. A lower bound on the size of the fundamental topological domain in the case of the flat Universe, based on the 7-year WMAP data, is $d = 2R_{\mathcal{LSS}} \cos(\alpha_{min}) \simeq 27.9$Gpc [11], where $R_{\mathcal{LSS}}$ is the distance to the last scattering surface (*i.e.* the recombination surface).

In the present paper, we study gravitational properties of the Universe with the slab topology $T \times R \times R$. More precisely, we investigate the effect of such topology on the form of the gravitational potential and force. It is well known that in the Newtonian limit the gravitational potential is defined by the Poisson equation. In the case of cosmology, the matter density fluctuations are the sources of this potential [15]. For toroidal types of topology, this equation was investigated in [16]. It was shown that there is no way to get any physically reasonable and nontrivial solution of this equation in the case of the slab topology. However, if we take into account relativistic effects and derive the equation for the gravitational potential from the perturbed Einstein equations, then we find that the resulting equation has the form of the Helmholtz equation rather than the Poisson one [17,18,19]. As we show in the present paper, this changes the situation drastically and for the considered slab topology we obtain the physically reasonable solutions. Moreover, to achieve it, we do not make any artificial assumptions about the distribution of gravitating masses. We obtain two alternative forms of the gravitational potentials and forces and demonstrate that the solutions in the form of the sum of Yukawa potentials are preferable in the present Universe for numerical calculations.

The paper is structured as follows. In sect. 2, the basic equations are presented and two alternative forms of solutions for the gravitational potential are obtained. They are compared from the point of numerical calculations in sect. 3. The corresponding alternative expressions for the gravitational force are derived in sect. 4. These expressions are also compared from the point of numerical calculations. A brief summary of the main results is given in concluding sect. 5.

2 Basic equations and alternative solutions

If we take into account relativistic effects in the framework of the conventional ΛCDM cosmological model, then the gravitational potential satisfies the following equation [17,18,19]:

$$\Delta\Phi_0 - \frac{3\kappa\overline{\rho}c^2}{2a}\Phi_0 = \frac{\kappa c^2}{2a}\left(\rho - \bar{\rho}\right), \tag{2.1}$$

where $\kappa \equiv 8\pi G_N/c^4$, G_N is the Newtonian gravitational constant, c is the speed of light, a is the scale factor and Δ is the Laplace operator in comoving coordinates. We consider matter in the form of discrete point-like gravitating masses m_n with comoving mass density

$$\rho = \sum_n m_n \delta(\mathbf{r} - \mathbf{r}_n)\,. \tag{2.2}$$

The averaged comoving mass density is constant: $\bar{\rho} = \text{const}$. The subscript 0 for Φ indicates that peculiar velocities were not taken into account in eq. (2.1) (see also [20]).

It can be easily seen that the shifted gravitational potential

$$\widehat{\Phi}_0 \equiv \Phi_0 - \frac{1}{3} \tag{2.3}$$

satisfies the equation

$$\Delta\widehat{\Phi}_0 - \frac{a^2}{\lambda^2}\widehat{\Phi}_0 = \frac{\kappa c^2}{2a}\rho\,, \tag{2.4}$$

where we introduce the screening length [17]

$$\lambda \equiv \left(\frac{3\kappa\overline{\rho}c^2}{2a^3}\right)^{-1/2}. \tag{2.5}$$

Eq. (2.4) now allows us to apply the superposition principle to get its solution. First, we can find a solution for a single particle, let it be a particle m in the center of the Cartesian coordinates, and then we can write a solution for the full system of particles.

In the case of the slab topology $T \times R \times R$, each gravitating mass m_n has its counterparts shifted by a distance multiple of the torus period l (for instance, along the z-axis in the Cartesian coordinates). Thus, for a selected particle m we should take into account all its counterparts. To this end, and also using the topology of the system, we can present the delta function $\delta(z)$ in the form

$$\delta(z) = \frac{1}{l}\sum_{k=-\infty}^{+\infty}\cos\left(\frac{2\pi k}{l}z\right). \tag{2.6}$$

Therefore, for a single point-like mass m in the center of the coordinate system we have

$$\Delta\widehat{\Phi}_0 - \frac{a^2}{\lambda^2}\widehat{\Phi}_0 = \frac{\kappa c^2}{2a}\frac{m}{l}\sum_{k=-\infty}^{+\infty}\cos\left(\frac{2\pi k}{l}z\right)\delta\left(x\right)\delta\left(y\right). \tag{2.7}$$

It is natural to seek for the solution in the following form:

$$\widehat{\Phi}_0 = \sum_{k=-\infty}^{+\infty} C_k(x,y)\cos\left(\frac{2\pi k}{l}z\right) . \tag{2.8}$$

Substitution of this expression into eq. (2.7) gives

$$\sum_{k=-\infty}^{+\infty}\left[\frac{\partial^2}{\partial x^2}C_k(x,y)+\frac{\partial^2}{\partial y^2}C_k(x,y)-\left(\frac{4\pi^2k^2}{l^2}+\frac{a^2}{\lambda^2}\right)C_k(x,y)-\frac{\kappa c^2}{2a}\frac{m}{l}\delta\left(x\right)\delta\left(y\right)\right]\cos\left(\frac{2\pi k}{l}z\right)=0 . \tag{2.9}$$

It is convenient to use the polar coordinates:

$$x=\xi\cos\phi,\quad y=\xi\sin\phi . \tag{2.10}$$

Obviously, due to the symmetry of the model, coefficients C_k depend only on the polar radius ξ, and for $\xi > 0$ satisfy the equation

$$\xi\frac{d^2C_k}{d\xi^2}+\frac{dC_k}{d\xi}-\left(\frac{4\pi^2k^2}{l^2}+\frac{a^2}{\lambda^2}\right)\xi C_k(\xi)=0 . \tag{2.11}$$

The general solution of this equation is superposition of the modified Bessel functions:

$$C_k(\xi)=AI_0\left(\sqrt{b}\xi\right)+BK_0\left(\sqrt{b}\xi\right),\quad b\equiv\frac{4\pi^2k^2}{l^2}+\frac{a^2}{\lambda^2}, \tag{2.12}$$

where A and B are the constants of integration. Hence, omitting the growing mode $I_0\left(\sqrt{b}\xi\right)$, we get

$$C_k(\xi)=BK_0\left(\sqrt{\frac{4\pi^2k^2}{l^2}+\frac{a^2}{\lambda^2}}\xi\right) . \tag{2.13}$$

To define the constant B, we take into account that at small ξ this function should satisfy the two-dimensional Poisson equation with a source proportional to $\delta(x)\delta(y)$. When $\xi\to 0$, we have $C_k(\xi)\to -B\ln\xi$. On the other hand, $\triangle(\ln\xi)=2\pi\delta(x)\delta(y)$. So,

$$-2\pi B=\frac{\kappa c^2}{2a}\frac{m}{l},\quad B=-\frac{\kappa c^2}{4\pi a}\frac{m}{l} . \tag{2.14}$$

Hence,

$$C_k(\xi)=-\frac{\kappa c^2}{4\pi a}\frac{m}{l}K_0\left(\sqrt{\frac{4\pi^2k^2}{l^2}+\frac{a^2}{\lambda^2}}\xi\right) \tag{2.15}$$

and

$$\widehat{\Phi}_0=-\frac{\kappa c^2}{4\pi a}\frac{m}{l}\sum_{k=-\infty}^{+\infty}K_0\left(\sqrt{\frac{4\pi^2k^2}{l^2}+\frac{a^2}{\lambda^2}}\xi\right)\cos\left(\frac{2\pi k}{l}z\right) . \tag{2.16}$$

Obviously, for a system of gravitating masses with arbitrary positions we have

$$\Phi_0=\frac{1}{3}-\frac{\kappa c^2}{4\pi a}\frac{1}{l}\sum_n m_n\left\{\sum_{k=-\infty}^{+\infty}K_0\left(\sqrt{\frac{4\pi^2k^2}{l^2}+\frac{a^2}{\lambda^2}}\left|\boldsymbol{\xi}-\boldsymbol{\xi}_n\right|\right)\cos\left[\frac{2\pi k}{l}\left(z-z_n\right)\right]\right\} . \tag{2.17}$$

It is not difficult to demonstrate that the formula (2.16) has the correct Newtonian limit in the vicinity of the considered gravitating mass. At such small distances all directions should be considered on an equal footing and summation is replaced by integration:

$$\widehat{\Phi}_0\to-\frac{\kappa c^2 m}{4\pi a}\int_{-\infty}^{+\infty}K_0\left(\sqrt{4\pi^2\tilde{k}^2+\frac{a^2}{\lambda^2}}\xi\right)\cos\left(2\pi\tilde{k}z\right)d\tilde{k},\quad \tilde{k}=\frac{k}{l},\ d\tilde{k}=\frac{dk}{l} . \tag{2.18}$$

We proceed with changing again the integration variable: $\tilde{k}=\tilde{k}_\xi/\xi$, $d\tilde{k}=d\tilde{k}_\xi/\xi$. Then, with the help of the formula 2.16.14(1) from [21], we obtain

$$\begin{aligned}\widehat{\Phi}_0 &\to -\frac{\kappa c^2 m}{4\pi a}\frac{1}{\xi}\int\limits_{-\infty}^{+\infty} K_0\left(\sqrt{4\pi^2\tilde{k}_\xi^2+\frac{a^2}{\lambda^2}\xi^2}\right)\cos\left(2\pi\tilde{k}_\xi\frac{z}{\xi}\right)d\tilde{k}_\xi\\ &\to -\frac{\kappa c^2 m}{2\pi a}\frac{1}{\xi}\int\limits_{0}^{+\infty} K_0\left(2\pi\tilde{k}_\xi\right)\cos\left(2\pi\tilde{k}_\xi\frac{z}{\xi}\right)d\tilde{k}_\xi = -\frac{\kappa c^2 m}{2\pi a}\frac{1}{\xi}\frac{1}{4\sqrt{1+z^2/\xi^2}} = -\frac{G_N m}{c^2}\frac{1}{\sqrt{\Xi^2+Z^2}}\,, \qquad (2.19)\end{aligned}$$

where $\Xi=a\xi$ and $Z=az$ are the physical coordinates. This formula is exactly the Newtonian expression.

Now we want to demonstrate another important property of the gravitational potential Φ_0. Since Φ_0 is the linear fluctuation of the metric coefficients, its averaged value should be equal to zero [17] (see also argumentation in [22]). Let us prove it. First, we rewrite Eq. (2.16) as

$$\widehat{\Phi}_0=-\frac{\kappa c^2}{4\pi a}\frac{m}{l}K_0\left(\sqrt{\frac{3\kappa\overline{\rho}c^2}{2a}}\xi\right)-\frac{\kappa c^2}{2\pi a}\frac{m}{l}\sum_{k=1}^{+\infty}K_0\left(\sqrt{\frac{4\pi^2k^2}{l^2}+\frac{a^2}{\lambda^2}}\xi\right)\cos\left(\frac{2\pi k}{l}z\right). \qquad (2.20)$$

Therefore,

$$\begin{aligned}\int\limits_{-\infty}^{+\infty}dx\int\limits_{-\infty}^{+\infty}dy\int\limits_{0}^{l}dz\widehat{\Phi}_0 &= -\frac{\kappa c^2}{4\pi a}m\int\limits_{-\infty}^{+\infty}dx\int\limits_{-\infty}^{+\infty}dyK_0\left(\sqrt{\frac{3\kappa\overline{\rho}c^2}{2a}\left(x^2+y^2\right)}\right)\\ &= -\frac{\kappa c^2}{4\pi a}m\cdot 2\pi\int\limits_{0}^{+\infty}\xi d\xi K_0\left(\sqrt{\frac{3\kappa\overline{\rho}c^2}{2a}}\xi\right) = -\frac{\kappa c^2}{2a}m\left(\frac{3\kappa\overline{\rho}c^2}{2a}\right)^{-1} = -\frac{1}{3}\frac{m}{\overline{\rho}}\,, \qquad (2.21)\end{aligned}$$

where we have used the table integral 2.16.2(2) from [21]. Then, taking into account the relation (2.3), for the averaged value of the total gravitational potential Φ_0 we have

$$\overline{\Phi_0}=\frac{1}{3}-\frac{1}{3}\frac{m}{\overline{\rho}}\cdot\frac{N}{L_xL_yl}=\frac{1}{3}-\frac{1}{3}=0,\quad \frac{mN}{L_xL_yl}=\overline{\rho}\,, \qquad (2.22)$$

where for simplicity we consider the case when all N particles in the volume $V=L_xL_yl$ have the same mass m.

Above we have presented one of the possible ways to solve eq. (2.4). However, since this is the Helmholtz equation, we can also solve it by direct summation over all counterpart contributions which have the form of Yukawa potentials:

$$\widehat{\Phi}_0=-\frac{\kappa c^2}{8\pi a}\frac{m}{l}\sum_{k=-\infty}^{+\infty}\frac{l}{\sqrt{\xi^2+(z-kl)^2}}\exp\left(-\frac{a\sqrt{\xi^2+(z-kl)^2}}{\lambda}\right). \qquad (2.23)$$

As we already mentioned, eq. (2.4) with the corresponding solutions (2.16) and (2.23) does not take into account the peculiar velocities of gravitating masses. However, in [23], the authors argued the importance of such account. It was demonstrated that the peculiar velocities can be included back into consideration effectively in eq. (2.4) by the replacement of the screening length λ with an effective screening length λ_{eff} (defined by the formula (41) in [23]):

$$\triangle\widehat{\Phi}-\frac{a^2}{\lambda_{\mathrm{eff}}^2}\widehat{\Phi}=\frac{\kappa c^2}{2a}\rho\,. \qquad (2.24)$$

In particular, at the matter-dominated stage $\lambda_{\mathrm{eff}}=\sqrt{3/5}\lambda$. Hence, instead of (2.16), (2.17) we have, respectively,

$$\widehat{\Phi}=-\frac{\kappa c^2}{4\pi a}\frac{m}{l}\sum_{k=-\infty}^{+\infty}K_0\left(\sqrt{\frac{4\pi^2k^2}{l^2}+\frac{a^2}{\lambda_{\mathrm{eff}}^2}}\,\xi\right)\cos\left(\frac{2\pi k}{l}z\right), \qquad (2.25)$$

$$\Phi=\frac{1}{3}\left(\frac{\lambda_{\mathrm{eff}}}{\lambda}\right)^2-\frac{\kappa c^2}{4\pi a}\frac{1}{l}\sum_n m_n\left\{\sum_{k=-\infty}^{+\infty}K_0\left(\sqrt{\frac{4\pi^2k^2}{l^2}+\frac{a^2}{\lambda_{\mathrm{eff}}^2}}\,|\boldsymbol{\xi}-\boldsymbol{\xi}_n|\right)\cos\left[\frac{2\pi k}{l}(z-z_n)\right]\right\}. \qquad (2.26)$$

The average value $\overline{\Phi}$ still equals 0.

In the case of the solution in the form of (2.23), we should simply replace λ with λ_{eff}. To distinguish among the derived expressions for the gravitational potential, we introduce the following notation:

$$\tilde{\varPhi}_{\cos} \equiv \left(-\frac{\kappa c^2}{8\pi a}\frac{m}{l}\right)^{-1}\widehat{\varPhi}_{\cos} = 2\sum_{k=-\infty}^{+\infty} K_0\left(\sqrt{4\pi^2k^2+\frac{1}{\tilde{\lambda}_{\text{eff}}^2}}\tilde{\xi}\right)\cos(2\pi k\tilde{z}) \tag{2.27}$$

and

$$\tilde{\varPhi}_{\exp} \equiv \left(-\frac{\kappa c^2}{8\pi a}\frac{m}{l}\right)^{-1}\widehat{\varPhi}_{\exp} = \sum_{k=-\infty}^{+\infty}\frac{1}{\sqrt{\tilde{\xi}^2+(\tilde{z}-k)^2}}\exp\left(-\frac{\sqrt{\tilde{\xi}^2+(\tilde{z}-k)^2}}{\tilde{\lambda}_{\text{eff}}}\right), \tag{2.28}$$

where the rescaled quantities are:

$$z = \tilde{z}l, \quad \xi = \tilde{\xi}l, \quad \lambda_{\text{eff}} = \tilde{\lambda}_{\text{eff}}al\,. \tag{2.29}$$

In what follows, we will explore the benefits of one formula over another from the point of view of numerical analysis.

3 Gravitational potentials

The functions (2.27) and (2.28) represent the contribution to the gravitational potential (scalar perturbation), produced by a point-like mass m located at $z = 0$, $\xi = 0$, and its images located at $z = \pm l, \pm 2l, \ldots$, $\xi = 0$, where l is the (comoving) period of the torus, z is the coordinate along the toroidal dimension, and ξ is the distance from the z-axis.

For numerical calculations, we have to cut off the infinite series (2.27) and (2.28) at some number n. This number of terms depends on the precision with which we want to calculate these expressions. Obviously, the fewer the number of terms needed for this, the better the corresponding formula is suitable for computation. We want to compare the formulas (2.27) and (2.28) from this point of view. The first conclusion, which follows from eq. (2.27), is that this expression for any n does not work for points with $\xi = 0$, since the modified Bessel function K_0 diverges at zero value of the argument. On the other hand, there is no such limitation for eq. (2.28).

Let us introduce the total numbers of terms $n_{\cos}$ and $n_{\exp}$ which we would like to include in the computation, then it follows from (2.27) and (2.28) that

$$\tilde{\varPhi}_{\cos}\Big|_{n_{\cos}} = 2K_0\left(\frac{\tilde{\xi}}{\tilde{\lambda}_{\text{eff}}}\right) + 4\sum_{k=1}^{n_{\cos}-1} K_0\left(\sqrt{4\pi^2k^2+\frac{1}{\tilde{\lambda}_{\text{eff}}^2}}\tilde{\xi}\right)\cos(2\pi k\tilde{z}) \tag{3.1}$$

and

$$\begin{aligned}\tilde{\varPhi}_{\exp}\Big|_{n_{\exp}} &= \frac{1}{\sqrt{\tilde{\xi}^2+\tilde{z}^2}}\exp\left(-\frac{\sqrt{\tilde{\xi}^2+\tilde{z}^2}}{\tilde{\lambda}_{\text{eff}}}\right)\\ &+ \sum_{k=1}^{n_{\exp}-1}\left[\frac{\exp\left(-\sqrt{\tilde{\xi}^2+(\tilde{z}+k)^2}/\tilde{\lambda}_{\text{eff}}\right)}{\sqrt{\tilde{\xi}^2+(\tilde{z}+k)^2}} + \frac{\exp\left(-\sqrt{\tilde{\xi}^2+(\tilde{z}-k)^2}/\tilde{\lambda}_{\text{eff}}\right)}{\sqrt{\tilde{\xi}^2+(\tilde{z}-k)^2}}\right],\end{aligned} \tag{3.2}$$

where we singled out the zero modes. The results of calculations with the help of Mathematica [24] are presented in Tables 1, 2. The values of $n_{\exp}$ in these tables describe the number of terms in (3.2), required to achieve the four-digit accuracy of determining $\tilde{\varPhi}$ (at the point of interest with some coordinates z, ξ). For all $n \geqslant n_{\exp}$ the four-digit value of $\tilde{\varPhi}_{\exp}\Big|_{n}$ does not change. If eq. (3.1) is used instead at the same point, then $n_{\cos}$ defines the number of terms in this formula to get the value of $\tilde{\varPhi}$ with the same accuracy. In the column for $n_{\cos}$, the dash means that the result of calculation is either incorrect (due to the computational difficulties) or indeterminate (because of the divergence of the function K_0). It is clear that the results of calculations depend on the ratio of the effective screening length λ_{eff} and the physical size al of the period of the torus: $\tilde{\lambda}_{\text{eff}} = \lambda_{\text{eff}}/(al)$. Therefore, in Tables 1, 2 we present the numbers obtained for both small and large values of $\tilde{\lambda}_{\text{eff}}$: 0.01, 0.1, 1 and 10, respectively.

These tables demonstrate that the formula (3.2) with Yukawa potentials generally requires much less number of terms ($n_{\exp} \ll n_{\cos}$) when the screening length is less than the period of the torus, *i.e.* $\tilde{\lambda}_{\text{eff}} < 1$. As we mentioned in Introduction, the lower limit on the period of the torus following from the observations is of the order of 16 Gpc [9].

	$\tilde z$	$\tilde\xi$	$\tilde\Phi$	n_{exp}	$n_{\cos}$
A_1	0.5	0.5	5.524×10^{-31}	2	20
A_2	0.5	0.1	2.810×10^{-22}	2	—
A_3	0.5	0	7.715×10^{-22}	2	—
B_1	0.1	0.5	1.405×10^{-22}	1	10
B_2	0.1	0.1	5.101×10^{-6}	1	31
B_3	0.1	0	4.540×10^{-4}	1	—
C_1	0	0.5	3.857×10^{-22}	1	9
C_2	0	0.1	4.540×10^{-4}	1	24

	$\tilde z$	$\tilde\xi$	$\tilde\Phi$	n_{exp}	$n_{\cos}$
A_1	0.5	0.5	2.402×10^{-3}	2	6
A_2	0.5	0.1	2.394×10^{-2}	2	21
A_3	0.5	0	2.695×10^{-2}	2	—
B_1	0.1	0.5	1.201×10^{-2}	2	4
B_2	0.1	0.1	1.719	1	15
B_3	0.1	0	3.679	1	—
C_1	0	0.5	1.350×10^{-2}	2	4
C_2	0	0.1	3.679	1	15

Table 1. Values of the rescaled gravitational potential $\tilde\Phi$ and corresponding numbers n_{exp} and $n_{\cos}$ of terms of series for some selected points in the cases $\tilde\lambda_{\text{eff}} = 0.01$ and $\tilde\lambda_{\text{eff}} = 0.1$ for the left and right tables, respectively.

	$\tilde z$	$\tilde\xi$	$\tilde\Phi$	n_{exp}	$n_{\cos}$
A_1	0.5	0.5	1.740	11	3
A_2	0.5	0.1	2.741	7	14
A_3	0.5	0	2.814	8	—
B_1	0.1	0.5	1.941	7	3
B_2	0.1	0.1	7.068	7	15
B_3	0.1	0	9.986	7	—
C_1	0	0.5	1.965	7	3
C_2	0	0.1	9.958	8	15

	$\tilde z$	$\tilde\xi$	$\tilde\Phi$	n_{exp}	$n_{\cos}$
A_1	0.5	0.5	6.114	64	3
A_2	0.5	0.1	7.297	81	15
A_3	0.5	0	7.378	62	—
B_1	0.1	0.5	6.325	64	3
B_2	0.1	0.1	11.69	49	11
B_3	0.1	0	14.63	46	—
C_1	0	0.5	6.350	61	3
C_2	0	0.1	14.59	40	10

Table 2. Values of the rescaled gravitational potential $\tilde\Phi$ and corresponding numbers n_{exp} and $n_{\cos}$ of terms of series for some selected points in the cases $\tilde\lambda_{\text{eff}} = 1$ and $\tilde\lambda_{\text{eff}} = 10$ for the left and right tables, respectively.

On the other hand, the effective cosmological screening length at the present time is 2.6 Gpc [23]. Therefore, the inequality $\tilde\lambda_{\text{eff}} < 1$ corresponds to the observable Universe, and here eq. (3.2) is preferable for the numerical analysis.

To conclude this section, we present Figs. 1, 2 of the rescaled gravitational potential $\tilde\Phi$ which correspond to four different values of $\tilde\lambda_{\text{eff}}$ selected for Tables 1, 2. To draw these pictures (with the help of Mathematica [24]), we use the formula (3.2) where we choose $n \gg n_{\text{exp}}$.

4 Gravitational forces

Below we present the gravitational forces (per unit mass) corresponding to the potentials considered in the previous section. More precisely, we calculate the projections of these forces on the z-axis and on the polar radius ξ for the selected points. Among these points, it is natural to consider only those ones where the corresponding projections are nonzero. Similarly to the potentials, we calculate these projections up to the fourth digit and find numbers of terms of the series starting from which we achieve this precision. The less the number of terms, the better the formula for the numerical analysis. Thus, we compare two alternative expressions following from (2.27) and (2.28) from this point of view.

4.1 z-component of the gravitational force

First, we consider z-components of the gravitational forces. From the potentials (2.27) and (2.28) we obtain two alternative formulas for the rescaled z-component of the force:

$$\left.\frac{\partial}{\partial\tilde z}\left(\tilde\Phi_{\cos}\right)\right|_{n_{\cos}} = -8\pi\sum_{k=1}^{n_{\cos}} kK_0\left(\sqrt{4\pi^2k^2+\frac{1}{\tilde\lambda^2_{\text{eff}}}}\,\tilde\xi\right)\sin(2\pi k\tilde z) \tag{4.1}$$

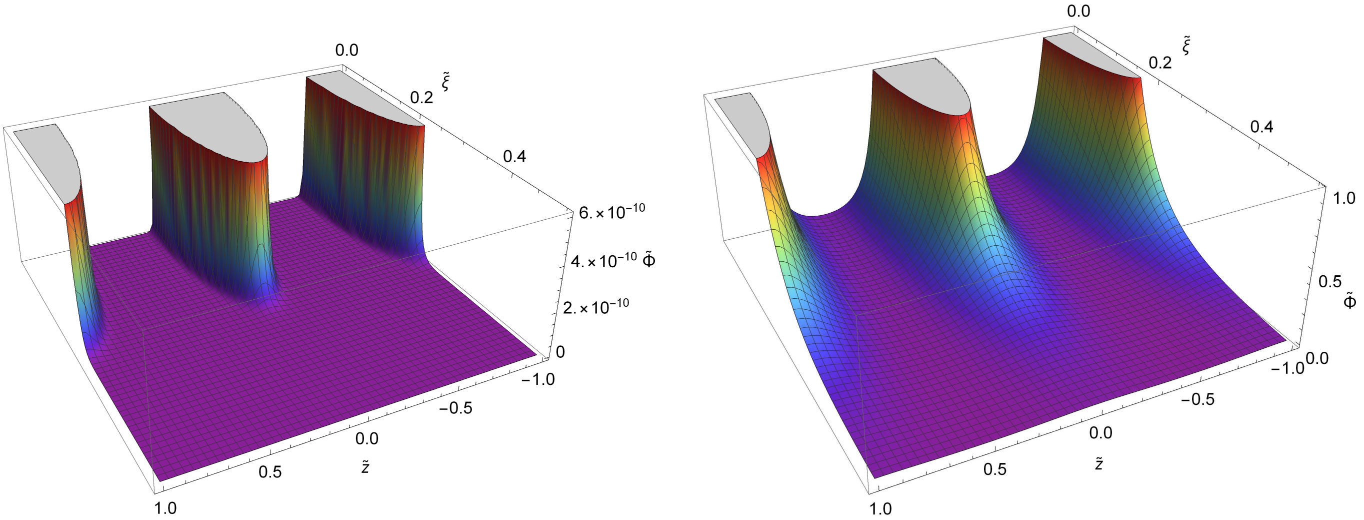

Fig. 1. Rescaled gravitational potential $\tilde\Phi = \left[-G_N m/(c^2 a l)\right]^{-1}\widehat{\Phi}$ in the cases $\tilde\lambda_{\rm eff} = 0.01$ and $\tilde\lambda_{\rm eff} = 0.1$ for the left and right panels, respectively.

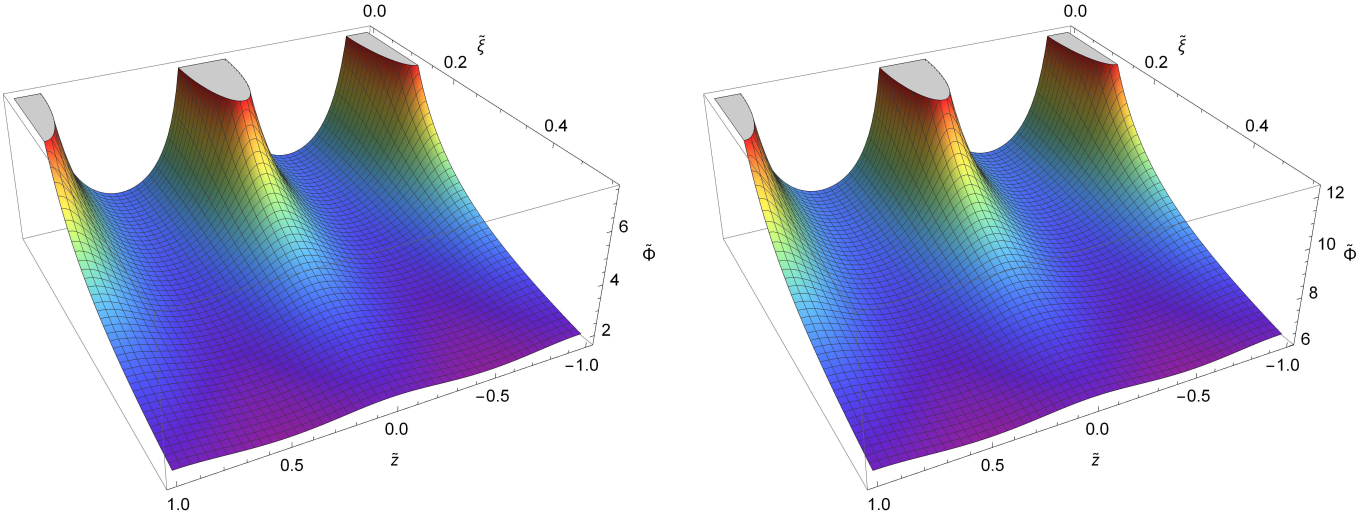

Fig. 2. Rescaled gravitational potential $\tilde\Phi = \left[-G_N m/(c^2 a l)\right]^{-1}\widehat{\Phi}$ in the cases $\tilde\lambda_{\rm eff} = 1$ and $\tilde\lambda_{\rm eff} = 10$ for the left and right panels, respectively.

and

$$\begin{aligned}\frac{\partial}{\partial\tilde z}\left(\tilde\Phi_{\rm exp}\right)\Big|_{n_{\rm exp}} = &-\left(\frac{\tilde z}{\tilde\lambda_{\rm eff}\left(\tilde\xi^2+\tilde z^2\right)} + \frac{\tilde z}{\left(\tilde\xi^2+\tilde z^2\right)^{3/2}}\right)\exp\left(-\frac{\sqrt{\tilde\xi^2+\tilde z^2}}{\lambda_{\rm eff}}\right)\\ &-\sum_{k=1}^{n_{\rm exp}-1}\left[\left(\frac{\tilde z+k}{\tilde\lambda_{\rm eff}\left(\tilde\xi^2+(\tilde z+k)^2\right)} + \frac{\tilde z+k}{\left(\tilde\xi^2+(\tilde z+k)^2\right)^{3/2}}\right)\exp\left(-\frac{\sqrt{\tilde\xi^2+(\tilde z+k)^2}}{\tilde\lambda_{\rm eff}}\right)\right.\\ &\left.+\left(\frac{\tilde z-k}{\tilde\lambda_{\rm eff}\left(\tilde\xi^2+(\tilde z-k)^2\right)} + \frac{\tilde z-k}{\left(\tilde\xi^2+(\tilde z-k)^2\right)^{3/2}}\right)\exp\left(-\frac{\sqrt{\tilde\xi^2+(\tilde z-k)^2}}{\tilde\lambda_{\rm eff}}\right)\right].\end{aligned}\tag{4.2}$$

Obviously, these z-components are equal to zero at the points A_1, A_2, A_3, C_1 and C_2. Therefore, we consider only the points B_1, B_2 and B_3. The results of our computations with the help of Mathematica [24] are presented in Tables 3, 4 for four values of $\tilde\lambda_{\rm eff}$: 0.01, 0.1, 1 and 10, respectively. These results demonstrate that, similarly to the

	$\tilde{z}$	$\tilde{\xi}$	$\tilde{\Phi}_z$	n_{exp}	n_{cos}
B_1	0.1	0.5	-2.810×10^{-21}	1	9
B_2	0.1	0.1	-3.862×10^{-4}	1	32
B_3	0.1	0	-4.994×10^{-2}	1	—

	$\tilde{z}$	$\tilde{\xi}$	$\tilde{\Phi}_z$	n_{exp}	n_{cos}
B_1	0.1	0.5	-2.781×10^{-2}	2	4
B_2	0.1	0.1	-20.75	1	17
B_3	0.1	0	-73.57	2	—

Table 3. Values of the rescaled z-component of the gravitational force $\tilde{\Phi}_z$ and corresponding numbers n_{exp} and n_{cos} of terms of series for points B_1, B_2 and B_3 in the cases $\tilde{\lambda}_{\mathrm{eff}} = 0.01$ and $\tilde{\lambda}_{\mathrm{eff}} = 0.1$ for the left and right tables, respectively.

	$\tilde{z}$	$\tilde{\xi}$	$\tilde{\Phi}_z$	n_{exp}	n_{cos}
B_1	0.1	0.5	-4.618×10^{-1}	8	3
B_2	0.1	0.1	-34.63	4	16
B_3	0.1	0	-99.11	5	—

	$\tilde{z}$	$\tilde{\xi}$	$\tilde{\Phi}_z$	n_{exp}	n_{cos}
B_1	0.1	0.5	-4.819×10^{-1}	39	3
B_2	0.1	0.1	-34.88	6	16
B_3	0.1	0	-99.51	6	—

Table 4. Values of the rescaled z-component of the gravitational force $\tilde{\Phi}_z$ and corresponding numbers n_{exp} and n_{cos} of terms of series for points B_1, B_2 and B_3 in the cases $\tilde{\lambda}_{\mathrm{eff}} = 1$ and $\tilde{\lambda}_{\mathrm{eff}} = 10$ for the left and right tables, respectively.

gravitational potential, the formula (4.2), based on the Yukawa potentials, is preferable for the physically relevant case $\tilde{\lambda}_{\mathrm{eff}} < 1$. In these tables, the values of the rescaled z-component $\tilde{\Phi}_z$ are computed with the help of eq. (4.2) for $n \gg n_{\mathrm{exp}}$.

In addition, we present Figs. 3, 4 of the rescaled z-components of the gravitational force $\tilde{\Phi}_z$ which correspond to four different values of $\tilde{\lambda}_{\mathrm{eff}}$ selected for Tables 3, 4. To draw these pictures (with the help of Mathematica [24]), we use the formula (4.2) where we choose $n \gg n_{\mathrm{exp}}$.

4.2 ξ-component of the gravitational force

Now we turn to the ξ-component of the gravitational force. In this case, two alternative formulas are:

$$\frac{\partial}{\partial\tilde{\xi}}\left(\tilde{\Phi}_{\mathrm{cos}}\right)\Big|_{n_{\mathrm{cos}}} = -\frac{2}{\tilde{\lambda}_{\mathrm{eff}}}K_1\left(\frac{\tilde{\xi}}{\tilde{\lambda}_{\mathrm{eff}}}\right) - 4\sum_{k=1}^{n_{\mathrm{cos}}-1}\sqrt{4\pi^2k^2+\frac{1}{\tilde{\lambda}_{\mathrm{eff}}^2}}\,K_1\left(\sqrt{4\pi^2k^2+\frac{1}{\tilde{\lambda}_{\mathrm{eff}}^2}}\,\tilde{\xi}\right)\cos\left(2\pi k\tilde{z}\right) \tag{4.3}$$

and

$$\begin{aligned}\frac{\partial}{\partial\tilde{\xi}}\left(\tilde{\Phi}_{\mathrm{exp}}\right)\Big|_{n_{\mathrm{exp}}} = &-\left(\frac{\tilde{\xi}}{\tilde{\lambda}_{\mathrm{eff}}\left(\tilde{\xi}^2+\tilde{z}^2\right)}+\frac{\tilde{\xi}}{\left(\tilde{\xi}^2+\tilde{z}^2\right)^{3/2}}\right)\exp\left(-\frac{\sqrt{\tilde{\xi}^2+\tilde{z}^2}}{\tilde{\lambda}_{\mathrm{eff}}}\right)\\ &-\sum_{k=1}^{n_{\mathrm{exp}}-1}\left[\left(\frac{\tilde{\xi}}{\tilde{\lambda}_{\mathrm{eff}}\left(\tilde{\xi}^2+(\tilde{z}+k)^2\right)}+\frac{\tilde{\xi}}{\left(\tilde{\xi}^2+(\tilde{z}+k)^2\right)^{3/2}}\right)\exp\left(-\frac{\sqrt{\tilde{\xi}^2+(\tilde{z}+k)^2}}{\tilde{\lambda}_{\mathrm{eff}}}\right)\right.\\ &\left.+\left(\frac{\tilde{\xi}}{\tilde{\lambda}_{\mathrm{eff}}\left(\tilde{\xi}^2+(\tilde{z}-k)^2\right)}+\frac{\tilde{\xi}}{\left(\tilde{\xi}^2+(\tilde{z}-k)^2\right)^{3/2}}\right)\exp\left(-\frac{\sqrt{\tilde{\xi}^2+(\tilde{z}-k)^2}}{\tilde{\lambda}_{\mathrm{eff}}}\right)\right].\end{aligned} \tag{4.4}$$

Evidently, this component is equal to zero at the points with $\tilde{\xi} = 0$ and nonzero $\tilde{z}$, for instance, at the points A_3 and B_3. Therefore, we consider only the points A_1, A_2, B_1, B_2, C_1 and C_2. The results of numerical calculations with the help of Mathematica [24], based on the formulas (4.3) and (4.4), for four values of $\tilde{\lambda}_{\mathrm{eff}}$ are presented in Tables 5, 6. We calculate the ξ-component up to the fourth digit and find how many terms n_{cos} and n_{exp} of the series in eqs. (4.3) and (4.4) are required for this purpose. $\tilde{\Phi}_\xi$ denotes the values of the rescaled ξ-component computed with the help of eq. (4.4) for $n \gg n_{\mathrm{exp}}$. These tables, similarly to Tables 1-4, demonstrate that in the case $\tilde{\lambda}_{\mathrm{eff}} < 1$, which corresponds to the observational restrictions, the formula (4.4) is preferable.

The behavior of the rescaled ξ-component of the gravitational force $\tilde{\Phi}_\xi$ for four chosen values of $\tilde{\lambda}_{\mathrm{eff}}$ is depicted (with the help of Mathematica [24]) in Figs. 5, 6. To this end, we use the formula (4.4) with $n \gg n_{\mathrm{exp}}$ (where the numbers n_{exp} are given by Tables 5, 6).

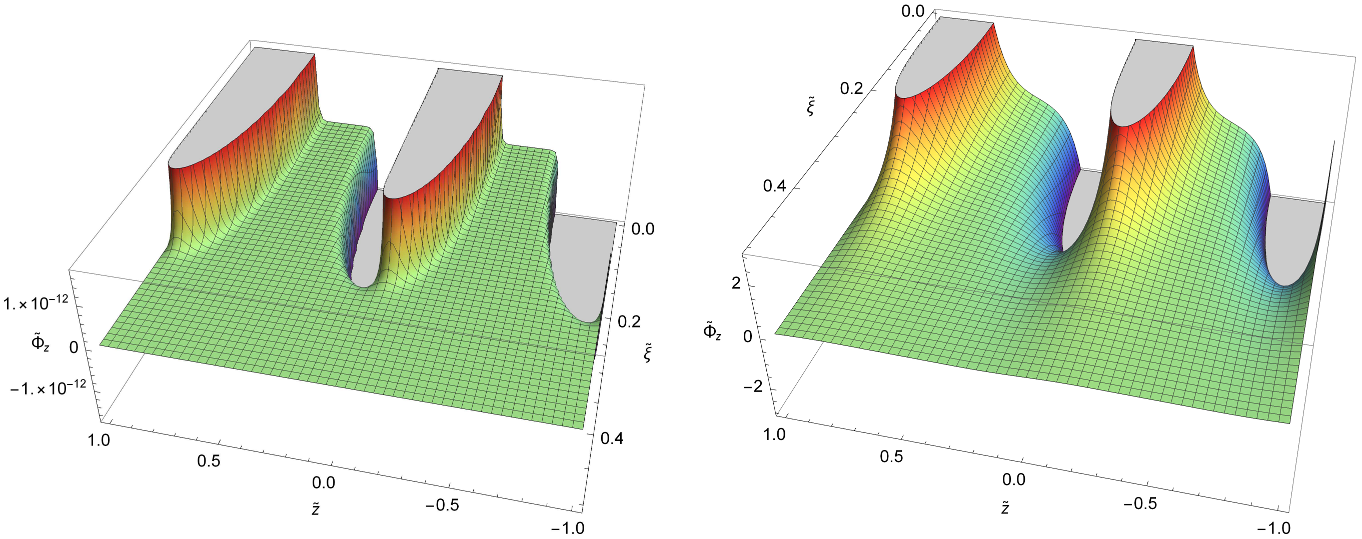

Fig. 3. Rescaled z-component of the gravitational force $\tilde{\Phi}_z \equiv \partial\tilde{\Phi}/\partial z$ in the cases $\tilde{\lambda}_{\rm eff} = 0.01$ and $\tilde{\lambda}_{\rm eff} = 0.1$ (left and right panels, respectively).

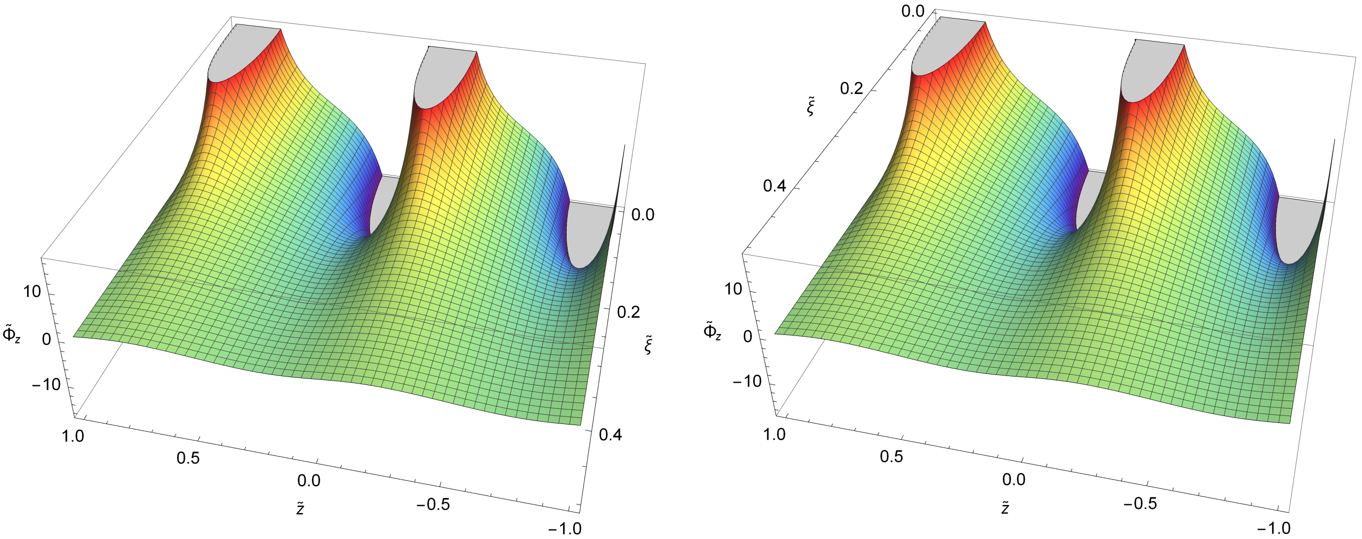

Fig. 4. Rescaled z-component of the gravitational force $\tilde{\Phi}_z \equiv \partial\tilde{\Phi}/\partial z$ in the cases $\tilde{\lambda}_{\rm eff} = 1$ and $\tilde{\lambda}_{\rm eff} = 10$ (left and right panels, respectively).

5 Conclusion

In this paper we have studied the effect of the slab topology $T \times R \times R$ of the Universe on the form of the gravitational potential and force. We have found two alternative forms of the solution: one (see eq. (2.27)) is based on the Fourier series expansion of the delta function using the periodical property along the toroidal dimension, and another one (see eq. (2.28)) is obtained by direct summation of the solutions of the Helmholtz equation for a source particle and all its images. The latter solution takes the form of the sum of Yukawa-type potentials. For both of these alternative presentations, the screening length $\tilde{\lambda}_{\rm eff}$ is an important parameter. The physical meaning of this length can be most clearly seen from the second formula: it defines the distance (from the source or its image) at which the corresponding potential undergoes the exponential cutoff. According to the observations, $\tilde{\lambda}_{\rm eff} < 1$ for the present Universe.

One of the main purposes of the paper was to determine which of the found alternative formulas works better from the point of numerical calculations. "Better" means which of the formulas requires less terms of the series to achieve the necessary precision. Our calculations show that for both gravitational potential and force, the formula with direct summation of Yukawa potentials is preferable in the physically relevant case $\tilde{\lambda}_{\rm eff} < 1$. Additionally, in

	$\tilde{z}$	$\tilde{\xi}$	$\tilde{\Phi}_\xi$	n_{exp}	n_{cos}
A_1	0.5	0.5	-3.962×10^{-29}	2	20
A_2	0.5	0.1	-5.620×10^{-21}	2	—
B_1	0.1	0.5	-1.405×10^{-20}	1	10
B_2	0.1	0.1	-3.862×10^{-4}	1	32
C_1	0	0.5	-3.935×10^{-20}	1	11
C_2	0	0.1	-4.994×10^{-2}	1	25

	$\tilde{z}$	$\tilde{\xi}$	$\tilde{\Phi}_\xi$	n_{exp}	n_{cos}
A_1	0.5	0.5	-1.939×10^{-2}	2	5
A_2	0.5	0.1	-5.615×10^{-2}	2	28
B_1	0.1	0.5	-1.406×10^{-1}	2	4
B_2	0.1	0.1	-20.75	1	17
C_1	0	0.5	-1.618×10^{-1}	2	4
C_2	0	0.1	-73.58	1	21

Table 5. Values of the rescaled ξ-component of the gravitational force $\tilde{\Phi}_\xi$ and corresponding numbers n_{exp} and n_{cos} of terms of series for points A_1, A_2, B_1, B_2, C_1 and C_2 in the cases $\tilde{\lambda}_{\text{eff}} = 0.01$ and $\tilde{\lambda}_{\text{eff}} = 0.1$ for the left and right tables, respectively.

	$\tilde{z}$	$\tilde{\xi}$	$\tilde{\Phi}_\xi$	n_{exp}	n_{cos}
A_1	0.5	0.5	-2.536	6	4
A_2	0.5	0.1	-1.405	4	20
B_1	0.1	0.5	-3.993	5	4
B_2	0.1	0.1	-35.20	3	17
C_1	0	0.5	-4.188	6	4
C_2	0	0.1	-99.69	3	19

	$\tilde{z}$	$\tilde{\xi}$	$\tilde{\Phi}_\xi$	n_{exp}	n_{cos}
A_1	0.5	0.5	-3.177	17	4
A_2	0.5	0.1	-1.587	12	21
B_1	0.1	0.5	-4.686	23	4
B_2	0.1	0.1	-35.60	6	16
C_1	0	0.5	-4.886	16	4
C_2	0	0.1	-100.2	2	14

Table 6. Values of the rescaled ξ-component of the gravitational force $\tilde{\Phi}_\xi$ and corresponding numbers n_{exp} and n_{cos} of terms of series for points A_1, A_2, B_1, B_2, C_1 and C_2 in the cases $\tilde{\lambda}_{\text{eff}} = 1$ and $\tilde{\lambda}_{\text{eff}} = 10$ for the left and right tables, respectively.

Figs. 1-6, we have presented graphically the gravitational potentials and force projections for four screening lengths $\tilde{\lambda}_{\text{eff}} = 0.01, 0.1, 1, 10$, respectively.

Declarations

Funding. The work of M. Eingorn, N. O'Briant and K. Arzu was supported by National Science Foundation (HRD Award #1954454).

Conflicts of interest. The authors have no conflicts of interest to declare that are relevant to the content of this article.

Data availability. All data generated or analyzed during this study are included in this article.

Authors' contributions. Maxim Eingorn: Conceptualization, Methodology, Formal analysis, Investigation, Writing – Review & Editing, Visualization, Supervision, Project administration, Funding acquisition. **Niah O'Briant:** Formal analysis, Investigation, Visualization. **Katie Arzu:** Formal analysis, Investigation, Visualization. **Maxim Brilenkov:** Formal analysis, Investigation. **Alexander Zhuk:** Methodology, Formal analysis, Investigation, Writing – Original Draft, Supervision.

References

1. P.A.R. Ade et al. [Planck Collaboration], *Planck 2013 results. XXVI. Background geometry and topology of the Universe*, A&A **571** (2014) A26; arXiv:1303.5086 [astro-ph.CO].
2. J.-P. Luminet, *The shape and topology of the Universe*; arXiv:0802.2236 [astro-ph].
3. B.F. Roukema, *Topology of the Universe: background and recent observational approaches*, Pramana **53** (1999) 945; arXiv:astro-ph/9904113.
4. W. Piechocki, *Topology of the Universe*; arXiv:gr-qc/9910055.
5. P. Bielewicz and A. Riazuelo, *The study of topology of the universe using multipole vectors*, Mon. Not. R. Astron. Soc. **396** (2009) 609; arXiv:0804.2437 [astro-ph].
6. P. Bielewicz, A.J. Banday and K.M. Gorski, *Constraining the topology of the Universe using the polarized CMB maps*, Mon. Not. R. Astron. Soc. **421** (2012) 1064; arXiv:1111.6046 [astro-ph.CO].

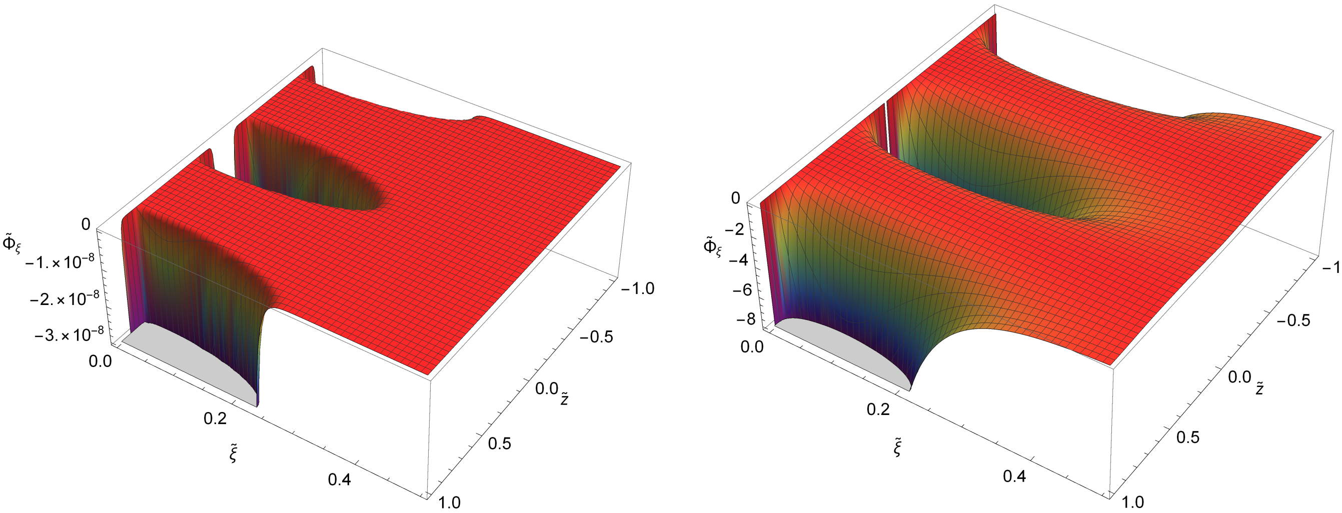

Fig. 5. Rescaled ξ-component of the gravitational force $\tilde{\Phi}_\xi \equiv \partial\tilde{\Phi}/\partial\xi$ in the cases $\tilde{\lambda}_{\rm eff} = 0.01$ and $\tilde{\lambda}_{\rm eff} = 0.1$ (left and right panels, respectively).

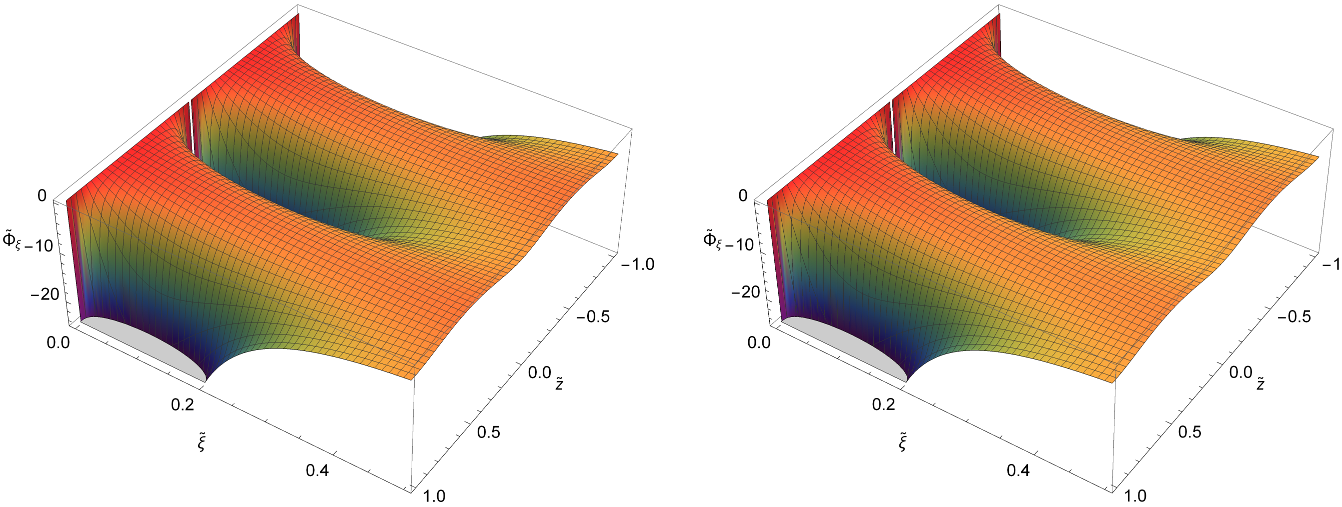

Fig. 6. Rescaled ξ-component of the gravitational force $\tilde{\Phi}_\xi \equiv \partial\tilde{\Phi}/\partial\xi$ in the cases $\tilde{\lambda}_{\rm eff} = 1$ and $\tilde{\lambda}_{\rm eff} = 10$ (left and right panels, respectively).

7. P.M. Vaudrevange, G.D. Starkman, N.J. Cornish and D.N. Spergel, *Constraints on the topology of the Universe: Extension to general geometries*, Phys. Rev. D. **86** (2012) 083526; arXiv:1206.2939 [astro-ph.CO].
8. O. Fabre, S. Prunet and J.-P. Uzan, *Topology beyond the horizon: how far can it be probed?* Phys. Rev. D. **92** (2015) 043003; arXiv:1311.3509 [astro-ph.CO].
9. P.A.R. Ade et al. [Planck Collaboration], *Planck 2015 results. XVIII. Background geometry and topology,* A&A **594** (2016) A18; arXiv:1502.01593 [astro-ph.CO].
10. P. Bielewicz and A.J. Banday, *Constraints on the topology of the Universe derived from the 7-year WMAP data*, Mon. Not. R. Astron. Soc. **412** (2011) 2104; arXiv:1012.3549 [astro-ph.CO].
11. P. Bielewicz, A.J. Banday and K.M. Gorski, *Constraints on the topology of the Universe derived from the 7-year WMAP CMB data and prospects of constraining the topology using CMB polarization maps*, Proceedings of the XLVIIth Rencontres de Moriond, 2012, eds. E. Auge, J. Dumarchez and J. Tran Thanh Van, published by ARISF, p. 91; arXiv:1303.4004 [astro-ph.CO].
12. G. Aslanyan and A.V. Manohar, *The topology and size of the Universe from the Cosmic Microwave Background*, JCAP **06** (2012) 003; arXiv:1104.0015 [astro-ph.CO].
13. E.G. Floratos and G.K. Leontaris, *On topological modifications of Newton's law*, JCAP **04** (2012) 024; arXiv:1202.6067 [astro-ph.CO].
14. G. Aslanyan, A.V. Manohar and A.P.S. Yadav, *The topology and size of the Universe from CMB temperature and polarization data*, JCAP **08** (2013) 009; arXiv:1304.1811 [astro-ph.CO].
15. P.J.E. Peebles, *The large-scale structure of the Universe* (Princeton University Press, Princeton, 1980).

16. M. Brilenkov, M. Eingorn and A. Zhuk, *Lattice Universe: examples and problems*, EPJC **75** (2015) 217; arXiv:1410.3909 [gr-qc].
17. M. Eingorn, *First-order cosmological perturbations engendered by point-like masses*, Astrophys. J. **825** (2016) 84; arXiv:1509.03835 [gr-qc].
18. M. Eingorn, C. Kiefer and A. Zhuk, *Scalar and vector perturbations in a universe with discrete and continuous matter sources*, JCAP **09** (2016) 032; arXiv:1607.03394 [gr-qc].
19. M. Eingorn, C. Kiefer and A. Zhuk, *Cosmic screening of the gravitational interaction*, Int. J. Mod. Phys. D **26** (2017) 1743012; arXiv:1711.01759 [gr-qc].
20. M. Eingorn, *Cosmological law of universal gravitation*, Int. J. Mod. Phys. D **26** (2017) 1750121; arXiv:1709.02264 [gr-qc].
21. A.P. Prudnikov, Yu.A. Brychkov and O.I. Marichev, *Integrals and Series, Volume 2. Special Functions* (Gordon and Breach Science Publishers, New York, 1992).
22. M. Eingorn, M. Brilenkov and B. Vlahovic, *Zero average values of cosmological perturbations as an indispensable condition for the theory and simulations*, EPJC **75** (2015) 381; arXiv:1407.3244 [astro-ph.CO].
23. E. Canay and M. Eingorn, *Duel of cosmological screening lengths*, Phys. Dark Univ. **29** (2020) 100565; arXiv:2002.00437 [gr-qc].
24. Wolfram Research, Inc., Mathematica, Version 11.3, Champaign, IL (2018).

Paper IV

BeyondPlanck I. Global Bayesian analysis of the Planck Low Frequency Instrument data

The BeyondPlanck Collaboration

2023, *A&A*, in press, arXiv: 2011.05609v2

Paper V

From BeyondPlanck to Cosmoglobe: Open Science, Reproducibility, and Data Longevity

S. Gerakakis, M. Brilenkov, M. Ieronymaki, M. San, D. J. Watts, and the BEYONDPLANCK Collaboration.

2023, *The Open Journal of Astrophysics*, Vol. 6, arXiv: 2205.11262v3

Version October 24, 2022
Preprint typeset using LaTeX style openjournal v. 09/06/15

arXiv:2205.11262v2 [astro-ph.IM] 21 Oct 2022

FROM BEYONDPLANCK TO COSMOGLOBE: OPEN SCIENCE, REPRODUCIBILITY, AND DATA LONGEVITY

S. Gerakakis[1], M. Brilenkov[2*], M. Ieronymaki[1], M. San[2], D. J. Watts[2], K. J. Andersen[2], R. Aurlien[2], R. Banerji[2], A. Basyrov[2], M. Bersanelli[3,4,5], S. Bertocco[6], M. Carbone[1], L. P. L. Colombo[3], H. K. Eriksen[2], J. R. Eskilt[2], M. K. Foss[2], C. Franceschet[3,5], U. Fuskeland[2], S. Galeotta[6], M. Galloway[2], E. Gjerløw[2], B. Hensley[7], D. Herman[2], M. Iacobellis[1], H. T. Ihle[2], J. B. Jewell[8], A. Karakci[2], E. Keihänen[9,10], R. Keskitalo[11], J. G. S. Lunde[2], G. Maggio[6], D. Maino[3,4,5], M. Maris[6], S. Paradiso[3,4], M. Reinecke[12], N.-O. Stutzer[2], A.-S. Suur-Uski[9,10], T. L. Svalheim[2], D. Tavagnacco[6,13], H. Thommesen[2], I. K. Wehus[2], A. Zacchei[6]
[1]Planetek Hellas, Leoforos Kifisias 44, Marousi 151 25, Greece
[2]Institute of Theoretical Astrophysics, University of Oslo, Blindern, Oslo, Norway
[3]Dipartimento di Fisica, Università degli Studi di Milano, Via Celoria, 16, Milano, Italy
[4]INAF-IASF Milano, Via E. Bassini 15, Milano, Italy
[5]INFN, Sezione di Milano, Via Celoria 16, Milano, Italy
[6]INAF - Osservatorio Astronomico di Trieste, Via G.B. Tiepolo 11, Trieste, Italy
[7]Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, U.S.A.
[8]Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, California, U.S.A.
[9]Department of Physics, Gustaf Hällströmin katu 2, University of Helsinki, Helsinki, Finland
[10]Helsinki Institute of Physics, Gustaf Hällströmin katu 2, University of Helsinki, Helsinki, Finland
[11]Computational Cosmology Center, Lawrence Berkeley National Laboratory, Berkeley, California, U.S.A.
[12]Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, 85741 Garching, Germany
[13]Dipartimento di Fisica, Università degli Studi di Trieste, via A. Valerio 2, Trieste, Italy

Version October 24, 2022

ABSTRACT

The BeyondPlanck and Cosmoglobe collaborations have implemented the first integrated Bayesian end-to-end analysis pipeline for CMB experiments. The primary long-term motivation for this work is to develop a common analysis platform that supports efficient global joint analysis of complementary radio, microwave, and sub-millimeter experiments. A strict prerequisite for this to succeed is broad participation from the CMB community, and two foundational aspects of the program are therefore reproducibility and Open Science. In this paper, we discuss our efforts toward this aim. We also discuss measures toward facilitating easy code and data distribution, community-based code documentation, user-friendly compilation procedures, etc. This work represents the first publicly released end-to-end CMB analysis pipeline that includes raw data, source code, parameter files, and documentation. We argue that such a complete pipeline release should be a requirement for all major future and publicly-funded CMB experiments, noting that a full public release significantly increases data longevity by ensuring that the data quality can be improved whenever better processing techniques, complementary datasets, or more computing power become available, and thereby also taxpayers' value for money; providing only raw data and final products is not sufficient to guarantee full reproducibility in the future.

1. INTRODUCTION

Reproducibility and replicability are two of the defining features of modern science. Within the field of CMB cosmology, this has most typically been realized in the form of competition between different experiments, each trying to measure the same sky signal but with different instrumentation and analysis techniques.[1] This approach has been tremendously successful and has led to a cosmological concordance ΛCDM model that is able to statistically describe nearly all currently available cosmological observables with only six free parameters (Planck Collaboration VI 2020).

The next major milestone for the CMB field is the potential detection of primordial gravitational waves and large-scale B-mode polarization (e.g., Kamionkowski & Kovetz 2016). If successful, this measurement will have far-reaching implications for our understanding of physics at ultra-high energy scales and the creation of the universe. However, this is also an extremely technologically challenging measurement because of the very faint expected signal amplitude. According to current theories and limits, it is anticipated to account for no more than a few tens of nanokelvin fluctuations on large angular scales, which is to be compared with the amplitude of the CMB solar dipole of 3.4 mK (Fixsen 2009), and with polarized astrophysical foreground contamination of tens of microkelvins (e.g., Planck Collaboration IV 2018). A robust detection will therefore require a relative instrumental calibration better than $O(10^{-5})$ and foreground suppression better than two orders of magnitude (e.g., Gjerløw et al. 2022; Svalheim et al. 2022b).

As discussed by BeyondPlanck (2022), this challenge imposes substantial requirements in terms of analysis and modeling techniques. Most notably, because of the intimate relationship between instrument calibration and astrophysical component separation, it is very likely that the associated parameters must be explored jointly, and it is also quite possible that data from different sources and experiments must be analyzed jointly to break internal degeneracies that exist within each experiment separately. As a concrete example, despite having almost one hundred times as many detectors

---

*maksym.brilenkov@astro.uio.no

[1] See, e.g., `https://en.wikipedia.org/wiki/List_of_cosmic_microwave_background_experiments` for a list of previous, current, and future CMB experiments.

as *Planck* (Planck Collaboration I 2020) and more than one order of magnitude higher map-level large-scale polarization sensitivity, *LiteBIRD*'s (LiteBIRD Collaboration et al. 2022) intensity sensitivity will not match *Planck*'s, and the ultimate *LiteBIRD* data analysis will therefore undoubtedly directly involve *Planck* measurements.

There is every reason to expect this to hold true for virtually all current and planned CMB experiments. The data from these experiments will benefit significantly from, if not depend on, a joint analysis with other datasets within iterative pipelines. Such an approach will maximize the amount of secondary science extracted from the datasets and allow them to achieve their primary science goals. Without exception, every single CMB experiment fielded to date has had parameters to which it was not sufficiently sensitive on its own, whether due to its observation strategy, detector design, or frequency coverage. It has typically required massive efforts to devise algorithmic priors or tricks to self-consistently mitigate these "blind spots" or "poorly measured modes". However, the optimal solution to solving such problems is, of course, by combining datasets with *different* blind spots, such that one experiment can break the degeneracies observed by another. One concrete example of this is the current BeyondPlanck[2] analysis, which re-analyzes the *Planck* LFI observations within a Bayesian end-to-end framework, and uses *WMAP* data to break important degeneracies between large-scale CMB polarization modes and the LFI gain (BeyondPlanck 2022; Gjerløw et al. 2022). Conversely, the ongoing *WMAP* re-analysis by Watts et al. (2022) will hopefully be able to constrain *WMAP*'s transmission imbalance parameters using information from *Planck*, and, if successful, this will improve the data quality of both experiments. Similarly, once *LiteBIRD* data become available, both *WMAP* and *Planck* should be re-analyzed from scratch, exploiting the *LiteBIRD*'s state-of-the-art large-scale polarization information to further improve the gain models of both experiments. In general, we therefore argue that for this type of joint analysis to be possible, it is critically important for all involved experiments to provide both raw data and a fully operational data analysis pipeline that can be re-run by external scientists.

A main goal of Cosmoglobe[3] is to establish a common platform for this type of joint analysis that can process low-level uncalibrated CMB time-ordered data (TOD) from different sources directly into high-level astrophysical component maps and cosmological parameters. The first application of this work is a full re-analysis of the *Planck* LFI observations (BeyondPlanck 2022, and references therein), while extensions to *WMAP* (Watts et al. 2022), *LiteBIRD* (LiteBIRD Collaboration et al. 2022), SPIDER (SPIDER Collaboration et al. 2021), *COBE*-DIRBE (Hauser et al. 1998) and others are on-going. However, for this work to be successful as a community-wide enterprise, it is necessary for all researchers to be able to reproduce the existing work, and integrate their own datasets into the analysis. As such, Open Science and reproducibility plays a critical role in this program.

In this paper, we summarize our efforts on reproducibility within the context of BeyondPlanck and Cosmoglobe. Its main goals are two-fold. First, it outlines the Open Source implementation of these projects and represents a valuable starting point for other experiments aiming to contribute to and build on this framework. Second, we hope that this paper may serve as a reference for any other future astrophysics and cosmology collaboration that wants to perform its work in an Open Source setting; the issues and tasks that need to be addressed in the context of BeyondPlanck and Cosmoglobe are very likely to be similar for any other project of a similar type. As such, a significant fraction of this paper is spent on surveys of tools and topics that were explored during the initial phases of the project but not ultimately chosen simply because this material may be helpful for other collaborations.

The rest of the paper is organized as follows: Starting in Sect. 2, we briefly review the statistical framework used by BeyondPlanck and Cosmoglobe, and we discuss why we believe that these issues will only become increasingly important for all future major cosmology and astrophysical missions. Next, in Sect. 3 we give an overview of various possible productivity tools that might be useful for future experiments, as well as different Open Source licenses. In Sect. 4 we provide an overview of the compilation support facilities implemented for the current software, while in Sect. 5 we summarize our documentation and accessibility efforts. Finally, we conclude in Sect. 6.

2. BeyondPlanck, Cosmoglobe, AND DATA LONGEVITY

2.1. *Breaking degeneracies through joint analysis of complementary datasets*

The fundamental motivation for the BeyondPlanck and Cosmoglobe projects derives directly from the experiences and insights gained within the *Planck* project. Towards the end of that project, it became clear that the main limiting factor with respect to constraining large-scale CMB polarization comes neither from instrumental systematics nor astrophysical foreground modeling as such but rather from the interplay between the two[4]. As formulated in BeyondPlanck (2022),

> *. . . one cannot robustly characterize the astrophysical sky without knowing the properties of the instrument, and one cannot characterize the instrument without knowing the properties of the astrophysical sky.*

One demonstration of this "chicken-and-egg" problem is shown in Fig. 1, reproduced from Gjerløw et al. (2022), where the top panel shows the Stokes Q difference map between *Planck* 2018 30 GHz (Planck Collaboration II 2020) and *WMAP* K-band (Bennett et al. 2013), after scaling the latter by 0.495 to account for the spectral index of synchrotron emission. Here one can see coherent large-scale patterns that massively dominate over the random noise. The origin of these structures is well understood and can, to a considerable extent, be described by the sum of transmission imbalance uncertainties in *WMAP* (middle row; Jarosik et al. 2007) and gain uncertainties in *Planck* (bottom row; Planck Collaboration II 2020). Both the transmission imbalance and the gain estimation rely directly on knowledge about the CMB sky, while estimating the CMB sky relies on knowledge about the transmission imbalance and gain parameters. However, even though their physical origins are well understood, they are still exceedingly difficult to mitigate within each experiment individually, simply because the observation strategy of each experiment leaves them nearly blind to these particular modes. One of the solutions to this problem is to jointly analyze each experiment and use the information in one experiment to break the degeneracies in the other.

---

[2] `https://beyondplanck.science`

[3] `https://cosmoglobe.uio.no`

[4] `https://www.cosmos.esa.int/web/planck/lessons-learned`

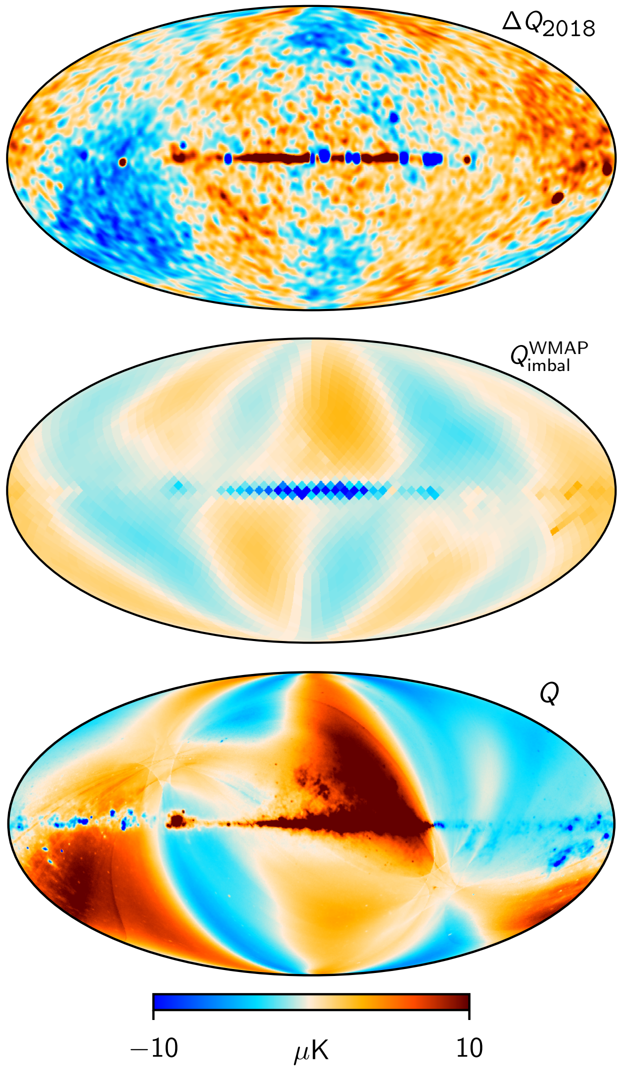

Fig. 1.— (*Top row*:) Stokes Q difference map between the 30 GHz *Planck* 2018 map and the K-band 9-year *WMAP* map, smoothed to a common resolution of 3° FWHM, and the latter has been scaled by a factor of 0.495 to account for different center frequencies; see Gjerløw et al. (2022) for further discussion. (*Middle row*:) *WMAP* transmission imbalance template (Jarosik et al. 2007). (*Bottom row*:) *Planck* 30 GHz gain residual template (Planck Collaboration II 2020).

Such complementary information need not only come from expensive satellite missions but can also come from less expensive ground-based experiments. One example of this is depicted in Fig. 2, reproduced from Ruud et al. (2015), which shows latitude-averaged polarization differences between the 43 GHz QUIET map and the corresponding 9-year *WMAP* (red curve) and *Planck* 2015 (blue curve) maps. In this case, one can see an excess in *Planck* with respect to QUIET outside $|b| < 1^\circ$, while QUIET and *WMAP* agree well. The most likely origin of this excess is bandpass-induced temperature-to-polarization leakage (Ruud et al. 2015; Svalheim et al. 2022a) in the *Planck* map at the level of 0.2 % (dotted line), which is fully consistent with the quoted systematic error of this channel of <1 % (Planck Collaboration III 2016). To reduce this systematic uncertainty below that achievable by *Planck* alone, additional external information is required, and the deep and highly cross-linked Galactic plane measurements made by QUIET are precisely what is needed for this.

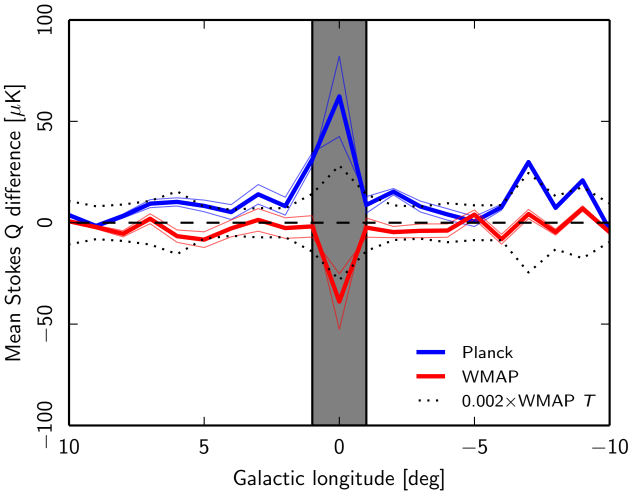

Fig. 2.— Latitude-averaged Stokes Q-band differences between QUIET 43 GHz and *WMAP* Q-band (red) and between QUIET 43 GHz and *Planck* 2015 44 GHz (blue) over the QUIET Galactic center field, evaluated over a latitude band around the Galactic plane of $|b| \leq 1.^\circ5$; reproduced from Ruud et al. (2015). The colored regions indicate the absolute QUIET calibration uncertainty of ±6 %. The dotted lines show the latitude-band-averaged *WMAP* Q-band temperature amplitude scaled by a factor of ±0.002, providing a rough template for a 0.2 % temperature-to-polarization leakage in *Planck*. The gray region marks an area in longitude ±1° around the Galactic center within which all results are dominated by uncertainties in the foreground spectral index.

These are only two relevant examples, and many more could be listed. Nevertheless, such examples motivate our contention that in order to break instrumental and astrophysical degeneracies, all free parameters should be optimized jointly while simultaneously exploiting as many state-of-the-art complementary datasets as possible. As such, the analysis problem should be solved globally, both in a statistical and a research community sense.

2.2. *The* BeyondPlanck *data model and posterior distribution*

As a first proof-of-concept of this global analysis approach, the BeyondPlanck collaboration (BeyondPlanck 2022) was formed with the explicit goal of re-analyzing the *Planck* LFI measurements. This data set represents a significant, but manageable challenge in terms of data volume and systematics complexity. Also, building on the experiences gained through the *Planck* project, we chose to adopt standard Bayesian parameter estimation techniques for our computer codes because of their unique flexibility and fidelity in terms of systematic error propagation. In particular, in the interest of saving costs and development time, we chose the CMB Gibbs sampler called `Commander` (Eriksen et al. 2004, 2008; Seljebotn et al. 2019) as our starting point, which formed a cornerstone in the *Planck* data analysis (Planck Collaboration IV 2018; Planck Collaboration V 2020; Planck Collaboration VI 2020).

The most crucial component in any Bayesian analysis is a parametric model for the data, which may typically take the following symbolic form,

$$\vec{d} = \vec{s}(\omega) + \vec{n}, \tag{1}$$

where $\vec{d}$ denotes a given dataset, $\vec{s}(\omega)$ is a signal model with free parameters ω, and $\vec{n}$ is noise. The posterior distribution, which quantifies the probability distribution of ω as constrained by $\vec{d}$, is given by Bayes' theorem,

$$P(\omega \mid \vec{d}) = \frac{P(\vec{d} \mid \omega)P(\omega)}{P(\vec{d})} \propto \mathcal{L}(\omega)P(\omega), \tag{2}$$

where $\mathcal{L}(\omega) \equiv P(\vec{d} \mid \omega)$ is called the likelihood, and $P(\omega)$ is called the prior; $P(\vec{d})$ is a normalization constant that is irrelevant for our purposes. The likelihood quantifies the constraining power of the actual data, while the prior summarizes our pre-existing knowledge regarding ω before the analysis.

For the *Planck* LFI analysis that is presented in a series of companion papers, we have adopted a parametric model that takes the following form,

$$d_{j,t} = g_{j,t}P_{tp,j}\left[B^{\mathrm{symm}}_{pp',j}\sum_{c} M_{cj}(\beta_{p'}, \Delta^{j}_{bp})a^{c}_{p'} + B^{\mathrm{asymm}}_{pp',j}\left(s^{\mathrm{orb}}_{j,t} + s^{\mathrm{fsl}}_{j,t}\right)\right] + s^{\mathrm{1Hz}}_{j} + n^{\mathrm{corr}}_{j,t} + n^{\mathrm{w}}_{j,t}. \tag{3}$$

For the purposes of the current paper, the specific meaning of each symbol is irrelevant, and we, therefore, refer the interested reader to Sect. 7 in BeyondPlanck (2022) for complete details. Here it is sufficient to note that this expression represents an explicit parametric model of both the instrument, as quantified by $\{g, P, B, \Delta_{bp}, s^{\mathrm{fsl}}, n^{\mathrm{corr}}\}$, and the astrophysical sky as expressed by the sum over components c, which for BeyondPlanck includes CMB, synchrotron, free-free, spinning and thermal dust emission, and compact sources.

It is important to note that this model is far richer than what *Planck* LFI is able to constrain on its own. Simply by counting degrees of freedom alone, we immediately note that the sky model has five free component amplitude parameters per pixel, while the LFI data only provide three independent frequencies. Consequently, the model is massively degenerate, and the LFI data must be augmented with external data. This is done in the current BeyondPlanck analysis by including the *WMAP* 33–61 GHz (Bennett et al. 2013), *Planck* HFI 353 and 857 GHz (Planck Collaboration III 2020), and Haslam 408 MHz (Haslam et al. 1982) measurements in the form of pixelized frequency maps. The advantage of this is that the model is now reasonably well constrained – but a major disadvantage is that these external pixelized sky maps may be associated with their own systematic errors that may compromise the final results. To fully exploit the strengths of each dataset in breaking degeneracies through joint analysis, one should ultimately start from raw TOD for all involved observations, and properly model all potential systematic effects. This is a main goal of the Cosmoglobe effort.

2.3. *Low-level systematics, data longevity, and cost optimization*

Whenever the signal-to-noise ratio of a given dataset increases, new systematic effects become important and must be modeled. This general observation also holds true for the CMB community, which currently targets signals at the nanokelvin level; even minuscule effects need to be accounted for at such low signal levels. This directly increases the importance of external data, as no planned experiment is able to measure all relevant effects internally at the required precision. More typically, each experiment focuses on one piece of the entire puzzle that it does particularly well for technological reasons and relies on other experiments to provide information regarding other free model parameters.

At the most basic level, the reason for this optimization is just a matter of cost and complexity. In particular, modern CMB experiments cost so much that it is unacceptable for funding agencies and taxpayers to repeatedly and needlessly measure the same quantities. For example, the ground-based or sub-orbital experiments typically cost at least about $10 M and involve 20–50 people, while current and next-generation satellite experiments usually cost hundreds of millions of dollars and involve hundreds of people.

To achieve new transformative results in the future within realistic budget limits, these existing million-euro investments must be optimally leveraged and re-used for all future experiments. For this to be possible, however, it is also vital that the systematic error budgets of the old datasets are consistent with the requirement of the new experiments. Unfortunately, this has traditionally been a prohibitive challenge for one simple reason: Until now, most CMB experiments have primarily published frequency maps or angular power spectra—that are static by nature—as their main products. Once the time-ordered data have been co-added into pixelized maps, it is no longer possible to account for a wide range of low-level systematic uncertainties, but only a very limited number of high-level uncertainties, such as white noise, correlated noise on large angular scales (Bennett et al. 2013; Basyrov et al. 2022), a single absolute calibration factor (Planck Collaboration Int. XLVI 2016; Gjerløw et al. 2022), or symmetrized beam uncertainties (Planck Collaboration V 2020). This significantly limits the use of legacy data for future analyses.

There are two noteworthy exceptions to this rule, though, namely *Planck* and *WMAP*. Both published their full uncalibrated time-ordered data as part of their legacy releases. Hence, the corresponding co-added frequency maps may, at least in principle, be continuously improved as new information and complementary datasets become available. However, it is also important to note that neither *Planck* nor *WMAP* released *the data analysis pipelines* that were used to reduce the data.

That is problematic for at least two reasons. First, from a practical point of view, the lack of functional analysis pipelines makes it very difficult and time-consuming for external scientists to repeat and improve the original analyses. Even more problematic, however, is the fact that most modern data reduction pipelines typically employ a significant number of critical ancillary datasets, for instance, ADC correction tables or far-sidelobe models. Since these are only used during low-level processing, few external researchers ask for them. As a result, they may easily be forgotten during the last stages of the main collaboration work and sometimes even lost when the original production computer systems are discontinued.

We argue in this paper that the optimal – if not only – way to ensure full reproducibility and data longevity is to release a complete and functional data processing pipeline together with the raw data, parameter files, high-level products, and documentation. Furthermore, we also argue that such a complete release should be required and supported in terms of dedicated funding for all future CMB experiments by the respective agencies (ESA, NASA, JAXA, NSF, etc.). This is clearly also in the funding agencies' own interests, as it guarantees that their investments may be optimally leveraged in future work.

In addition to sharing data, it is also worth noting that sharing analysis tools may lead to cost optimization of any given new experiment. Indeed, establishing common analysis tools across the field will free up analysis funding that can be better spent on understanding the instrument, exploring ground-breaking theories, or deriving novel secondary science. An important pioneering CMB-related example of this is `HEALPix` (Górski et al. 2005), which both defines a standard pixelization that facilities easy data sharing and comparison across experiments, and provides a wide range of state-of-the-art and user-friendly tools to operate on these data (e.g., Zonca et al. 2019), all published under an Open Source license. In general, common software tools are highly beneficial for the science community, funding agencies, and taxpayers.

2.4. *Open Science: From* BEYONDPLANCK *to* COSMOGLOBE

An important goal for the BeyondPlanck project was to develop and publicly release a complete end-to-end analysis pipeline for one of the essential datasets in contemporary CMB cosmology, namely the *Planck* LFI data (BeyondPlanck 2022). The motivation for this was two-fold. The first aim was to resolve a few notable issues with the LFI data that remained unresolved at the end of the official *Planck* mission, in particular related to the global estimation of the instrumental gain (Planck Collaboration II 2020; Gjerløw et al. 2022). However, this represents only a first step in a much larger process, as embodied within the Cosmoglobe program, whose goal is to develop a general low-level analysis pipeline that would be applicable to a much more comprehensive range of experiments – legacy, current, and future – and at the same time support *joint* analysis of these.

The full scope of this project is massive. For this work to succeed in the long term, it must be firmly based on an Open Science foundation: While a small group of dedicated people may be able to re-analyze one experiment (as BeyondPlanck has done for *Planck* LFI), integrating a wide range of complementary experiments without community contributions is unfeasible for several reasons. First, some important datasets may be proprietary, and the original stakeholders must be leading the analysis for legal reasons alone. Second, the systematic properties of most datasets are quite complicated, and expert knowledge is usually essential to formulate and generalize the data model. Third, the sheer amount of work to be done effectively requires cost-sharing among all interested parties, recognizing the currently constrained funding environment most researchers experience daily.

Fortunately, a solid financial base for the initial phases of this work has been established through three separate but complementary EU-funded projects. First, an ERC Consolidator grant called bits2cosmology (total budget of 2 M€) supports the implementation of time-domain processing in `Commander`, as presented in the current paper suite. bits2cosmology will continue to operate until 2023, and facilitate continued algorithm development activities for *LiteBIRD*, SPIDER, and *WMAP*. The second project is an H2020-funded COMPET-4 project called BeyondPlanck (total budget of 1.5 M€), for which the main goal is to re-process the *Planck* LFI observations in a Bayesian setting. BeyondPlanck ended formally on November 30th, 2020, and its results are currently being reported. The third project is called Cosmoglobe, an ERC Consolidator grant (total budget of 2 M€). It aims to build an Open Science CMB analysis community around these tools and use this to derive a new state-of-the-art model of the radio, microwave, and sub-millimeter sky. Cosmoglobe started in 2019 and will support community-wide activities at least until 2024 based on current funding.

In general, Cosmoglobe will be a hub for which analysis of both time-ordered and map-domain data from different experiments can be integrated into a single framework. A critical goal of Cosmoglobe is to support scientists working on incorporating the data from their own experiments into the larger Cosmoglobe framework and thereby analyzing the data efficiently, robustly, and economically. For the casual user who might be primarily interested in what the microwave sky looks like at some specified frequency, Cosmoglobe will provide a state-of-the-art and user-friendly sky model. By allowing scientists easy access and configuration of their experiment in this framework, Cosmoglobe will become an integral tool for forecasting and planning for experiments.

With a functional codebase in hand, as demonstrated by the current LFI-based data release, we believe that the time is now right for all interested parties to get involved in this work and extend the framework according to their own needs. We anticipate that such contributions will most typically take one of two forms. The first is stand-alone projects, in which the external user simply downloads the software and data and performs some analysis without input from the greater community. In this case, the only formal obligation for the user is to publish all derived codes under an equally permissive software license as the one used for BeyondPlanck (which in practice means a GNU General Public License (GPL)) and to acknowledge previous work through appropriate referencing.

The second mode of operation is active participation in the Cosmoglobe framework. In this case, an external user or project may request direct expert Cosmoglobe support, for instance, in the form of software development and data analysis assistance. That is then likely to increase the chance of success significantly. In exchange for the support, the external user or project must commit to publicly releasing the underlying data after some proprietary period, and all directly contributing Cosmoglobe collaborators must be offered co-authorship, in accordance with standard scientific practices. Required details can be specified in a Memorandum of Understanding (MoU) between the external party and the relevant Cosmoglobe participants before the work commences. Cosmoglobe is intended to be a platform for initiating and supporting mutually beneficial collaborations.

3. REPRODUCIBILITY SURVEY, TOOLS, AND LICENSING

We now turn our attention to the practical aspects of how to build an Open Science-based foundation for this work and examine some of the latest developments on reproducibility in science in general. Next, we identify several tools and services that are available online and aim to provide solutions for reproducible science. Finally, we review the most popular licenses used for Open Source development. These issues cover a variety of topics that constitute the current state of the art in reproducibility and Open Science as of 2018–19.

3.1. *Open Science development tools per 2018*

We collected information regarding available tools that might be useful to strengthen the reproducibility aspects of the project. We found this exercise quite informative and helpful, and we highly recommend future collaborations to conduct similar meta-studies *before* starting the data analysis work, as it is easy to get swamped with scientific problem solving once the main effort begins. We also note that the field moves very quickly, and the state-of-the-art is likely to be quite different only after a few years.

3.1.1. *Workflow definition tools versus integrated software*

We first considered the use of so-called "workflow definition tools" to organize the primary data model and Gibbs sampler discussed in Sect. 2.2. Such workflow managers help scientists define and execute a specific set of tasks, implemented by executing local (or sometimes remote) code, scripts, and other sub-workflows. Each component is only responsible for a small fragment of functionality. Therefore many pieces are working together in a pipeline to achieve the ultimate goal of the workflow, performing a useful task.

There have already been attempts at implementing and utilizing such tools in the CMB community. The most well-known example is the `ProC` workflow manager[5] developed by the Max Planck Institute of Astrophysics for the *Planck* mission. Two popular Open Source and general-purpose tools are `Taverna`[6] and `The Kepler Project`[7].

The main advantage and attraction of such workflow managers is their ability to construct complex, flexible, and reproducible workflows based on well-defined components. At the same time, what makes these managers work is very strict interfaces between the different components. Unfortunately, this strictness adds significant additional burdens on the code developers, both in terms of a steep learning curve to be able to add new features and in terms of restricted flexibility to implement new solutions to unexpected problems; it is difficult to make substantial changes without breaking compatibility with already existing components. A second significant challenge is efficient memory management. Suppose the various pipeline components are written in different programming languages. In this case, one must either resort to data sharing through slow disks or spend great effort on highly non-trivial in-memory communication.

In general, our experience is that general-purpose workflow managers tend to be more practical for well-established and relatively quick routine tasks than for cutting-edge research that relies on high-performance computing. The most critical priority for our Bayesian analysis framework is computational speed, as a factor of six in runtime can make the difference between a two-month runtime (which is painful but doable) and one year (which is prohibitive). Optimal memory and disk management are, therefore, the key. The second most important priority is code flexibility, allowing developers to introduce changes needed to achieve their goals freely.

After careful consideration, we decided to drop the use of workflow managers, as the official *Planck* Data Processing Centers (DPCs) did, to maintain optimal coding agility and flexibility. However, in contrast to the *Planck* DPCs, we instead opted for developing the entire analysis pipeline within one single computer program called `Commander3`, to ensure optimal memory management and computational speed (Galloway et al. 2022a). Two important additional advantages of implementing the entire pipeline within a single code are that the whole collaboration naturally develops a common "language" and knowledge base that are helpful to discuss issues more efficiently, and it also minimizes duplication of effort.

3.1.2. *Online development services*

Another potentially useful class of tools is the so-called "online development services". These offer the possibility to perform all development work to be done online through the use of general-purpose web applications. Some of the major players in this area that we evaluated were `Open Science Framework`[8], `Codeocean`[9], and `Zenodo`[10].

One of the major advantages of these services is that, by definition, all work is performed online. This facilitates very easy dissemination since results may be published in an Open Science manner literally in real-time. However, our evaluation is that they are also associated with three main disadvantages, mirroring the issues discussed in the previous section. First and foremost, online services typically impose a specific and inflexible work style that may not suit everybody in a large collaboration. Second, they have a significant learning curve that may be off-putting to many scientists with busy schedules. Third, depending on the plans offered by each provider, authors can easily run out of hosting space or online computational time, requiring them to update their accounts. While this may make financial sense from the side of the hosting companies, we consider this to be a big disadvantage for authors, just for reproducibility purposes alone. A solution like this might make sense for small projects, but the cost can become prohibitive for larger and heavier collaborations.

At their current stage of development, we, therefore, also decided to avoid the use of integrated online development services and instead leave each collaborator to choose their own development environment individually. We also note that most scientists are, by nature, quite independent-minded and do not necessarily respond well to being imposed on a specific development environment. However, if the available tools offered more obvious advantages, the situation might be different, and we definitely recommend future collaborations to perform a similar survey.

3.1.3. *Software repositories*

One class of software development tools that is critical for a large-scale Open Source effort is efficient *revision control systems* (RCSs). This allows users to collaborate on the same computer program or scientific paper in real-time with a minimum of synchronization problems and is a cornerstone of modern software development. As noted above, 56 % of the user survey respondents already use at least one such system, with `Git` being the most popular.

At the beginning of the program, we quickly settled on `Git` as our main RCS, primarily because it was most widespread in our group, but also because we find that it handles merges and conflicts better than most competitors. The main question was then which (if any) common repository we should use. Three particularly well-known providers are `Bitbucket`[11], `GitHub`[12], and `GitLab`[13].

One advantage of `GitLab` and `Bitbucket` is that they offer free private repositories in addition to public ones. This option might make them a better candidate for users who want to start their project as a private repository but switch to a fully public repository later on, closer to publication time. In addition, all three offer a full suite of online development tools, including bug/issues management, wiki pages, file hosting capabilities, and API access to hosted files.

Initially, we adopted `GitLab` as our main provider, primarily because it allows code to be run remotely on their web

---

[5] `http://planck.mpa-garching.mpg.de/ProC/`
[6] `https://taverna.incubator.apache.org`
[7] `https://kepler-project.org/`
[8] `https://osf.io/`
[9] `https://codeocean.com/`
[10] `https://zenodo.org/`
[11] `https://bitbucket.org`
[12] `https://github.com`
[13] `https://about.gitlab.com/`

hosts. In addition, we considered that it might be helpful for small tasks, such as implementing online tools and calculators or automatically compiling paper drafts after each submission. However, it is important to note that this feature is only free for limited usage. Therefore, we have concluded it was not as useful as initially anticipated. Consequently, halfway through the project, we have switched to `GitHub` for our central software repository, simply because most people in our community already have accounts there and to avoid overhead by maintaining two separate accounts for most users.

3.1.4. *Open Source licenses*

When working in an Open Science setting, it is vital to protect the investments and interests of the various contributors and users. A critical aspect of this is licensing. Today, many Open Source licenses are in active use, and an important task for projects like BeyondPlanck and Cosmoglobe is to choose the appropriate one for the work at hand. In this section, we provide a brief overview of licenses in the most common use today and discuss which one was chosen for our project, given the basic requirements that (1) our software should be Open Source; (2) all derivatives of this work should remain Open Source; and (3) our license should not contradict the licenses of any dependencies (`HEALPix`[14], `FFTW3`[15], etc.).

Although the term *Open Source software* may be intuitively understood to be freely distributable, modifiable, and shareable code written by a single or a group of programmers, it is, in reality, more complex than it may seem at first glance. Generally speaking, when discussing software licenses, one needs to distinguish between several different aspects and concepts. The first aspect concerns basic distribution. On the most restricted side, *proprietary software* is considered private property, and users are not allowed to share, study, change, or reverse engineer the provided software. A variation of this is called *freeware*; in this case, the original software developer retains all rights, and the only difference is that end-users do not need to pay for the basic usage of a given program. *Source available software* is software that allows users to view the source code but does not necessarily give the right to distribute, modify and/or install it on their machines. One example of such a license is the *Commons Clause License*[16], which prohibits users from selling the software. Because of such restrictions, source available licenses are generally not considered to be Open Source. Next, *Public Domain software* or "unlicensed" software is software that waives all the rights of copyright, trademark, or patent. Such software belongs to the "public" that uses it, and it can be freely distributed, modified, and/or sold without attribution to anyone. Examples of such licenses are *Creative Commons* (CC0) and *Unlicense*. During the course of history, experts have been arguing whether this type of license should be considered Open Source or not[17] even though they have many common features. Because of these disagreements, we also do not consider this license to be a proper Open Source license in the present work.

Moving on to what is considered proper Open Source software, *free software* (FS), as defined by the *Free Software Foundation*[18] (FSF), is "software that gives its users the freedom to run, copy, distribute, study, change and improve the software". According to a formal definition, the software is not considered free if it does not respect the following four essential freedoms:[19]

Freedom 0: The freedom to run the program for any purpose.

Freedom 1: The freedom to study how the program works, and change it so it does your computing as you wish. Access to the source code is a precondition for this.

Freedom 2: The freedom to redistribute copies so you can help others.

Freedom 3: The freedom to distribute copies of your modified versions to others. By doing this you can give the whole community a chance to benefit from your changes. Access to the source code is a precondition for this.

Finally, *Open Source licenses* are defined by the *Open Source Initiative*[20] (OSI) and include licenses that "allow the software to be freely used, modified, and shared", and comply with ten distinctive criteria[21] that concern (1) free redistribution; (2) source code; (3) derived works; (4) integrity of the author's source code; (5) no discrimination against persons or groups; (6) no discrimination against fields of endeavor; (7) distribution of license; (8) the license must not be specific to a product; (9) the license must not restrict other software; and (10) the license must be technology-neutral.

Both FS and OSI are considered to be Open Source with only subtle differences.[22] While FS is focused on the user's rights to use, modify and share the program, OSI is focused on the source code being open with unrestricted community driven development.[23] Since the main goal of the Cosmoglobe project is to build a community around a common source code, this strongly suggest that the OSI definition is most suitable for our purposes. However, looking at OSI's list of the most popular and widely used licenses,[24] which includes GNU variants, we are going to use both definitions interchangeably.

In order to select one of these OSI licenses, it is important to recognize that the `Commander` source code does not exist in a vacuum but rather depends directly and indirectly on a variety of different libraries, as visualized in Fig. 3. When one chooses the correct license for the project, the list of dependencies must be considered. Then, typically, the most specific one defines what is allowed for the new software. However, this is not always the case. For example, `cURL`[25] is based on the modified MIT license, but it may or may not be compiled with the support of `MbedTLS`[26] and `LibSSH2`[27] that rely on more complicated licenses. The same applies to,

---

[14] http://healpix.sourceforge.net or https://healpix.sourceforge.io

[15] https://www.fftw.org

[16] https://commonsclause.com/

[17] Lawyer Lorence Rosen has written the essay titled "Why the public domain isn't a license"—faced with strong opposition he has accepted that CC0 can be considered open-source.

[18] https://www.fsf.org/

[19] https://www.gnu.org/philosophy/free-sw.html

[20] https://opensource.org/licenses

[21] https://opensource.org/osd

[22] These are more philosophical in nature. As Richard Stallman, the founder of GNU Project and FSF, stated: *"The term 'open source' software is used by some people to mean more or less the same category as free software... The differences in extension of the category are small: nearly all free software is open source, and nearly all open source software is free."*

[23] https://opensource.org/about

[24] For the full list of open-source licenses, please visit: https://opensource.org/licenses/category

[25] https://github.com/curl/curl/blob/master/COPYING

[26] https://github.com/ARMmbed/mbedtls/blob/development/LICENSE

[27] https://github.com/libssh2/libssh2/blob/master/COPYING

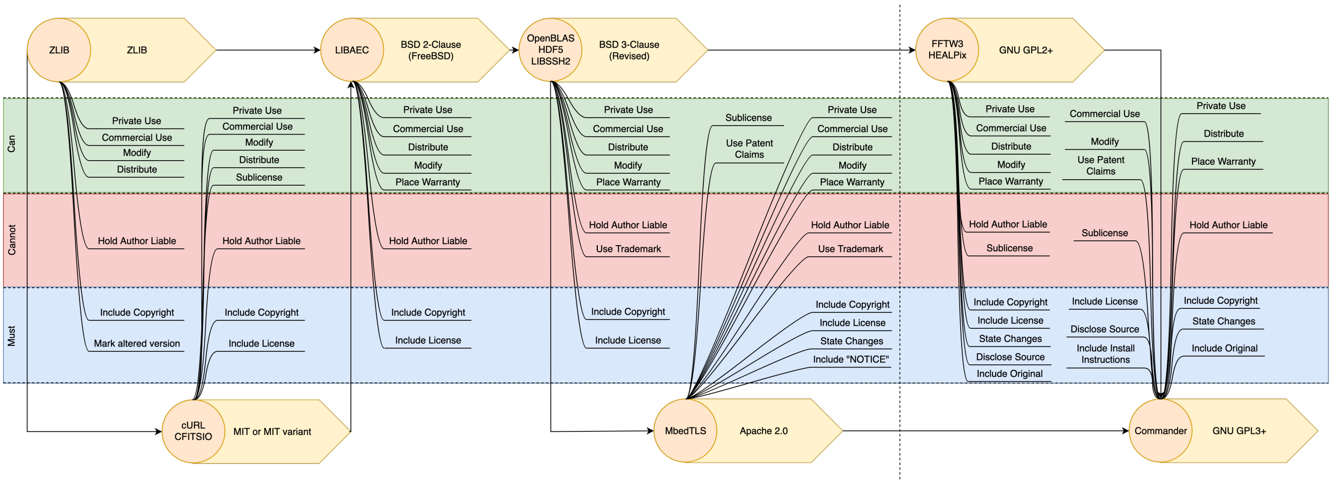

Fig. 3.— Licensing diagram for `Commander` and its dependencies, ordered from left to right according to increasingly specific licenses. Circles indicate libraries and codes, while rectangles show the licenses under which the particular library was issued. The dashed line divides the so-called permissive (i.e., licenses that can be combined and used together with proprietary software) and restrictive licenses (i.e., licenses that enforce the source code to stay open for general public). This diagram was derived from the David A. Wheeler's FLOSS License Slide and the information provided on the `https://tldrlegal.com/`.

e.g., `CFITSIO`[28] which requires `ZLIB`[29] starting from version 4.0.0. A third important example is `Commander` itself, which may be compiled with Intel Parallel Studio and Intel Math Kernel Library (`MKL`), which are issued under one of Intel's proprietary licenses. We do not ship or install this library together with `Commander` in any way, and, thus, such a decision relies solely on the user's judgment.

Our understanding is that as long as one does not explicitly modify the source code of a library issued under a permissive license, the licenses can be used almost interchangeably. Generally, this does not apply to the restrictive ones. However, the GPL licenses are internally compatible as long as it includes a special line[30] that explicitly states that the later version of the license can be used. In practice, this means that if a piece of software is issued under GPLv2 and not GPLv2+, then we cannot use it.

All in all, the license adopted for `Commander` and Cosmoglobe cannot be more restrictive than the most permissive of these, which is set by the GPL2+ license adopted for `HEALPix`. In principle, the same applies to `FFTW3`, but if needed, this library could have been replaced with other FFT implementations. In contrast, `HEALPix` is in practice irreplaceable and therefore determines the license also for the current work. Consequently, our final choice is the GNU General Public License v3, as dictated by the diagram. Of course, for us, this is not only a matter of formality but also of preference; we *want* this software to be and remain Open Source to protect the interests of everybody involved. Therefore, it is crucial for all future participants and developers of the Cosmoglobe framework to familiarize themselves with this license and determine whether this is acceptable for one's work and compliant with potential collaboration policies.

4. COMPILATION SUPPORT

The `Commander` software is primarily intended to be run on `Linux`-based *High-Performance Computing* (HPC) clusters with basic `Fortran` and `MPI` compilers and libraries available and typically tens to thousands of computing cores. Beyond that, it does not impose any specific constraints on the computing platform, neither in the form of processor architecture (AMD, ARM, Cray, Intel, etc.) nor compilers (GNU, Intel, PGI, etc.). At the same time, the code does depend on many external libraries, including `HEALPix`, `FFTW3`, `CAMB`,[31] and `HDF5`;[32] for full specification, see the online `Commander` documentation.[33] The combination of many dependencies and a rich computing platform heterogeneity can, in general, represent a significant challenge and workload in terms of compilation and can be seriously off-putting to many users. Therefore, it is critical to make this process as simple and user-friendly as possible, and automated build systems play a key role in that work.

Build systems come in various flavors and combinations, which makes it non-trivial to choose one among many. As a result, there is no single "best" build system, but each has its advantages and disadvantages. Therefore, when selecting one specific system for `Commander`, we have considered seven different aspects listed in order of importance.

1) Free and Open-source: `Commander` will not be a genuinely Open Source tool if the build system which installs it is itself not free and Open Source.

2) Cross-platform: Although Linux was (and still is) the predominant operating system for modern HPC, there is currently a trend toward a more diverse landscape.[34] In addition, Windows and Macintosh software dominate the PC and laptop market, and these systems are growing rapidly in power and can be used today for productive analysis. Hence, we require that the build system must allow us to compile and run `Commander` on any major operating system, including Linux, Windows, and MacOS.

3) Automation: Many astrophysics departments operate today their own computing cluster. However, experience shows that there is often limited professional support for the installation, tuning, and maintenance of software and libraries.

---

[28] `https://heasarc.gsfc.nasa.gov/docs/software/fitsio/c/f_user/node9.html`

[29] `https://www.zlib.net/zlib_license.html`

[30] *"(...) either version 2 of the License, or (at your option) any later version. (...)"*

[31] `https://camb.info/`

[32] `https://www.hdfgroup.org`

[33] `https://cosmoglobe.github.io/Commander/`

[34] According to top500.org, the share of non-Linux systems running the world's top HPC systems was 2.4 % in November 2010, 3.4 % in November 2014, 7 % in November 2019, and 16 % in November 2021.

This work is often left to the scientists who want to run the code; Ph.D. students, postdocs, and professors. The installation procedure must therefore be both transparent and easy to use. Furthermore, it should be automated, i.e., the build system should be able to check for the existence of specific `Commander` dependency, and, if it doesn't find it on the host system, it should download, verify, compile, and install all missing dependencies, compile `Commander` itself, and link them all together.

4) Long-term support: While Open Source projects have many advantages in current computing, they also tend to have one fundamental flaw: People tend to abandon them in favor of "newer" and "better" codes. This poses an obvious threat to a large and long-term project such as Cosmoglobe, which will require significant community-wide investments over many years. For this reason, the adopted build system should be mature and supported by a dedicated community. Proven stability is more important than cutting-edge.

5) Multi-language support: The bulk of the `Commander` codebase consists of Fortran code, but there are many `Python` scripts and libraries written in the course of the development. In addition, there are also several `C++`-based modules that need to be compiled together with `Commander`, and other languages may become useful in the future. Therefore, the tool should have support for multiple languages used simultaneously in one project.

6) Minimal dependencies: The build system should have as few dependencies as possible, and these should ideally either be already present in the system or easy to install from the source. Preferably, it should be a single binary or a piece of software.

7) IDE Integration: While *Integrated Development Environment* (IDE) support is not a strict requirement, it is certainly a nice bonus feature. In the present day and age, people are using a huge variety of IDEs that can perform syntax highlighting and code checking—having the same features available for the build system gives the programmer an advantage in terms of the code development speed.

With these points in mind, we now provide a survey of possible useful compilation support tools in current use and then describe the implementation chosen for `Commander`.

4.1. *Survey of automatic build systems*

4.1.1. *Low-level build systems –* `Make`

`Make` is a "low-level" build automation tool which uses special instruction files – so-called `Makefiles` – to build and install software from the source code. It has a variety of implementations, with perhaps, the most widespread one being GNU `Make` which is shipped together with most `Linux` and `Unix` distributions. In fact, `Commander1` and `Commander2` were solely built using `Make`. Its advantages are numerous: It is a standard Linux and Unix tool; it is widespread, and the majority of the scientific and open source community knows about it; it will not be deprecated in the foreseeable future since it has a solid community of maintainers; it has a support for a variety of languages such as `C`, `C++`, `Fortran`, `Java`, `Python`, etc.; it allows nested project structures; and, starting from version 3.0, it allows compilation using multiple processes.

However, it also has a few notable disadvantages. First, `Make` has a relatively obscure syntax and associated steep learning curve. Second, for large projects, such as `Commander`, the `Makefiles` tend to become very long and complex and increasingly hard to maintain. Finally, code compilation requires specific instructions for each particular compiler, OS, and architecture, making it a poor solution for cross-platform development. Despite these shortcomings, `Make` remains the most widespread build system in use today, and it has a firmly established user community. Taking into account both its flexibility and maturity, we have chosen `Make` as the primary low-level compilation system for `Commander`.

4.1.2. *High-level build system –* `CMake`

Developed since 1999 under a BSD-3-Clause license, `CMake` is a "meta" or "high-level" build system that is used in conjunction with some other "low-level" build environments, for instance `Make`, `Ninja`, or `Microsoft Visual Studio`. `CMake` may be used to build a software project in a two-step process: First, `CMake` reads in a series of configuration files written in `CMake`'s own scripting language, called `CMakeLists.txt`, and use these to automatically produce a complete low-level build system configuration (e.g., `Makefiles`). Second, these files are used by the low-level native generator (e.g., `Make`) to actually compile and install the project.

`CMake` allows for flexible project structures. For instance, nested directory hierarchies and/or complex library dependencies do not pose a problem, as it can locate a variety of files and executables on the host system. Once such dependencies have been identified, the location data are stored inside a special file called `CMakeCache.txt` that can be manually tuned before the actual build. In addition, it has the functionality to download, verify, unpack and compile archives of missing libraries that utilize non-`CMake` build systems. Furthermore, cross-compilation is straightforward since `CMake` has extensive OS, language, and compiler support.[35] Other important features include, but are not limited to, support for mathematical expressions; string, list, and file manipulation; conditions, loops, functions, and macros; and shell scripting.

Today, `CMake` is the de facto standard tool for `C++` project development,[36] and a variety of Open Source projects use `CMake` as its build system, including several `Commander` dependencies such as `CFITSIO`, `FFTW3`, and `HDF5`. Lastly, `CMake` has good support for IDE integration.

However, despite its many strengths, `CMake` is not perfect. The syntax has a pretty steep learning curve, and the source code may quickly become cumbersome and difficult to read. Also, there does not seem to be a universal approach, or even strict guidelines, for how to structure `CMake` code for large and complicated projects. Finally, the documentation is extensive but non-trivial to navigate and read for newcomers.

Based on its prevalence, mature community, rich feature set, and the fact that many `Commander` dependencies also use `CMake`, we have chosen this as our primary high-level build system, with `Make` as the corresponding low-level system. Specific details regarding the `Commander` `CMake` configuration is described in Sect. 4.2.

4.1.3. *Alternative build systems*

This section provides an overview of other competing systems that were explored during the initial phases of the project

---

[35] `https://cmake.org/documentation/`

[36] According to a 2019 survey by Jet Brains, 42 % of `C++` projects use `CMake`; 33 % use `Makefiles`; 9 % use `Qmake`; 8 % use `Autotools`; and only 1 % use `SCons`. In addition, according to the 2021 Annual C++ Development Survey roughly 80 % of all projects use `CMake` as one of their build systems.

but ultimately not selected. However, several of these may be attractive candidates for future astrophysics and cosmology Open Source projects.

`Ninja`[37] is a low-level build system, similar to `Make`, specifically designed for speed. Like `Make`, it supports a variety of languages, platforms, compilers, and operating systems and, in fact, is meant to eventually replace `Make`. The main reason for not choosing `Ninja` over `Make` is simply the fact that it was designed to be used in combination with high-level build systems and not on its own. In addition, it is still in active development, and this carries a risk of higher – and unnecessary – development overheads for our purposes. This choice is likely to be revisited in the future when `Ninja` has proven itself further in terms of stability and user base.

`QMake` or `makemake` is a build system created by the Qt Company which automates the creation of `Makefiles`, similar to `CMake`. It supports multiple platforms and can produce `Makefiles` tuned for specific operating systems. Although mostly used for `C++` projects, it can incorporate custom compilers (e.g., `gfortran`), and this allows it to work with `Fortran` source files. However, it lacks native support for non-C++ languages, a natural alternative to `Make` build tools, and the ability to incorporate third-party libraries directly into the build.

`XMake` is a lightweight build system that supports multiple languages, tool-chains, and platforms, and it can compile projects both directly and produce configuration files for low-level build systems such as `Make` or `Ninja`. However, native `Fortran` support was added as recently as July 2020 (in version 2.3.6), well after a system had to be chosen for the current `Commander` development. It is fast and has many IDE plug-ins.

GNU `Autotools` is a GNU build system composed of several utility programs that are designed to make source code portable to many `Unix`-like Operating Systems. It is widely used by many free and Open Source projects, including several astrophysical and cosmology ones. In general, `Autotools` generate the distribution archive used to build programs. Once users obtain this package, they need to unpack it and run three simple and well-known commands – `configure`, `make` and `make install` – to compile and install the code using the facilities provided by their host systems. Such an approach, in theory, eliminates the need to install `Autotools` entirely, but, in practice, `Linux` distributions still have it with (sometimes) multiple versions installed, which adds to the confusion. Furthermore, a major drawback of `Autotools` is its complexity; it requires much experience and time to develop robust and user-friendly configuration files. For many future projects, we consider `CMake` to be a more accessible and user-friendly solution.

`Scons` does not implement a new special-purpose and domain-specific language but rather utilizes specific `Python` scripts to build the projects. Thus, the only requirement is to have `Python` installed on the system, and this makes `Scons` cross-platform by default; any system that runs `Python` can install `Scons` using `Python`'s standard installation frameworks, `pip` or `conda`. MIT license, good multi-language support, reliance only on `Python`, and a rich feature set make it a tool worthy of exploring for new projects. A major shortcoming is that there does not appear to be a simple way of integrating other projects or dependencies into the build.

`Waf` is another build system solely based on `Python`. Since `SCons` inspired `Waf`, both tools have many similar features: They rely exclusively on `Python`; are cross-platform; can automatically scan for project dependencies; and have support for multiple languages, including `C`, `C++` and `Fortran`. However, while `SCons` is older and therefore is used in more projects and has better documentation, `Waf` seems to be much faster and provides more user-friendly console output that makes it easier to debug. Additionally, it does not require a separate installation since the tool is designed to be shipped as part of the main project source code. Unfortunately, similar to `SCons`, there is no easy way to incorporate other projects into the build.

`Meson` is the last `Python`-based tool considered here and is the closest `CMake` competitor we have found so far. It is similar to `CMake` in many (if not all) aspects, and both are meta-languages that compile the source code in a two-step process. However, while `CMake` uses `Make` by default, `Meson` uses `Ninja` instead; this makes `Meson` even faster in some cases. It is also an Open Source, cross-platform tool that supports multiple languages, including `Fortran`. In addition, nested hierarchies are not a problem, as well as the incorporation of other projects (both `Autotools`- and `CMake`-based) into the build.

Fortran Package Manager or simply `fpm` is a relatively new[38] initiative by the Fortran-Lang foundation[39] inspired by `Cargo`, the package manager for the `Rust` programming language. It is both a build system and a package manager that can build libraries and applications. In addition, it has native support for unit testing and can include other dependencies (e.g., from `Git`) into the project. While looking very promising, we considered that it was not ready for large-scale production at the time when the current project started. Still, this option should be revisited in the future.

4.2. *`CMake`-based compilation*

As discussed above, we have adopted `CMake` and `Make` as our high- and low-level build systems. In this section, we provide an overview of the `Commander`-specific `CMake` configuration and compilation procedure.

4.2.1. *`Commander`-specific `CMake`-code organization*

Usually, when one is writing `CMake`-code for the project, each subdirectory has its own `CMakeLists.txt` file hence creating a tree-structure with the main `CMakeLists.txt` in the root of the project repository, while the "leafs" are in its respective subdirectories. However, since `Commander` has a substantial codebase already, we have decided to split the `Fortran` and `CMake` codes entirely. This allows for better navigation, makes the directory structure cleaner, and thus easier to maintain and expand. Furthermore, since the number of files involved in compiling `Commander` and its dependencies is significant, we have therefore adopted a so-called "out-of-source" build approach, as recommended by the `CMake` creators. In this organization, the source code and compiled build files are stored in separate locations. In our case, the source code itself is located inside the directory called `commander3`. In contrast, the binary folder is called `build`[40], and it is usually created by the user inside the `root`

---

[37] https://ninja-build.org

[38] Alpha version was released on `GitHub` in November 25, 2020.

[39] https://fortran-lang.github.io/fpm/.

[40] It is worth noting that there is nothing special about this name, and the users can name the directory where the `CMake` build files will be stored in whatever way they prefer. However, the convention is to name it `build`.

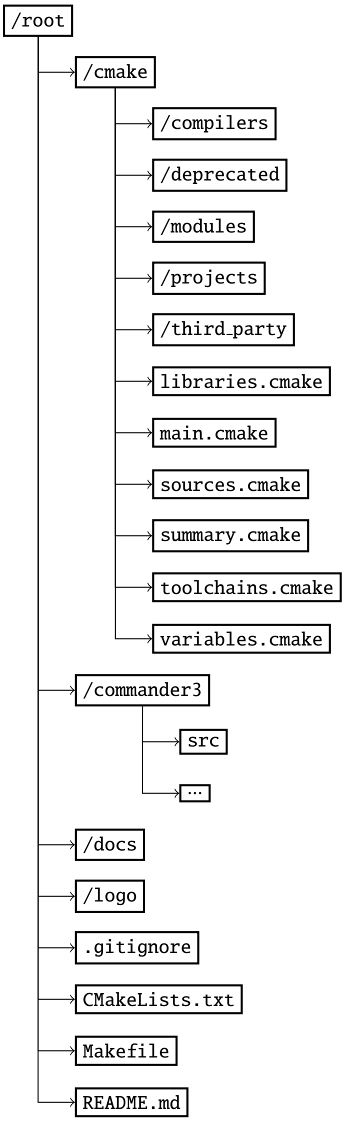

Fig. 4.— Overview of Commander source code directory structure. The CMake-scripts are inside the cmake directory. Here, compilers contains compiler configurations; deprecated contains deprecated code for convenience; modules contains custom written Find<Name>.cmake modules; projects contains configuration for all library subprojects; third_party contains CMake modules taken from other projects; libraries.cmake handles general subprojects' configurations; sources.cmake contains urls and various hashes for the subprojects; summary.cmake module is for debug output of the final CMake configuration on the screen; toolchains.cmake module handles general compiler configuration; variables.cmake has custom-defined variables available for to build Commander as a single project; main.cmake is there to mimic the logic of C++, Fortran, Python etc. codes, and serves as an analogy to the "entry point" of the program.

directory. One advantage of out-of-source compilation is that the users can create whatever amount of build folders they want/require; thus, the same source code tree can be used to produce multiple binaries, corresponding, for instance, to different debug flags or CPU architectures.

All in all, the Commander source code directory structure is visualized in Fig. 4. In this figure, root represents the root folder of the project, created by the original git clone command; cmake contains all CMake related files; commander3 contains all Commander related files; docs contains instructions on how to generate out-of-source documentation; logo contains Commander logos; .gitignore is a git version control file; CMakeLists.txt is the top level CMake configuration file, which serves as the starting point for the compilation process; Makefile is a traditional-style Makefile that may be used to compile Commander3 without CMake; and README.md describes the project on GitHub.

4.2.2. *Commander CMake workflow*

The CMake process works as follows. First, the host system is scanned, and the present/missing libraries are registered inside CMakeCache.txt and other auto-produced files. Then, once the configure step is done and the user has issued the build command, CMake downloads, configures, compiles, and installs the missing dependencies and, together with the ones identified as available, it links all libraries to the compiled Commander.

The CMake module that enables this behavior is called ExternalProject, and its primary purpose is to facilitate downloading and installation of dependencies that are not an internal part of the main project. In this way, the Commander dependencies are treated as entirely independent entities. Such isolation allows the build to be performed in the same way on different platforms, with utterly different build settings (e.g., compiler flags) and/or with a completely different build system (e.g., Autotools). Under the hood, it defines the set of so-called targets, each representing a particular step in the build process of an external project. These steps are then collected under a unified name (in our case, a sub-project name), used later in the code. CMake also remembers information about each performed step, which, if executed successfully, will not be repeated. This allows us to compile all dependencies only once for different Commander build types. An essential feature for debugging this process (if and when something fails) is the CMake logs, which are stored in /usr/local/logs[41].

Command-line arguments determine compiler selection during the scan phase. For instance, -DCMAKE_Fortran_COMPILER=ifort tells CMake to use the Intel ifort compiler. Then, for the most common compilers default optimization flags are defined per (sub-)project in a configuration file called cmake/projects/<project_name>.cmake. When installing this software on a new system with a new, unknown to Commander compiler, these are the configuration files that most likely need to be updated.

Based on the initial system scan and user-specified compilation instructions, CMake proceeds with the following steps for each dependency and for the main Commander source (the two first steps are skipped in the latter case):

Download: The project is downloaded via external links in the form of .zip or .tar.gz archives, or directly from Git repositories. We use MD5 hashes whenever possible to ensure that the correct library versions are downloaded.

Update/Patch: This step applies potential patches to the downloaded archive or, in the case of Git, brings the project up to date. In cases where we download the release versions of the packages, we skip this step.

Configure: This step can use CMake and other build tools alike, depending on the preferences of the authors of the original dependency. In our case, most libraries use Autoconfig

---

[41] The location can be changed by using defined CMAKE_LOG_DIR variable during the configuration process. We refer the interested reader to Commander documentation for further details.

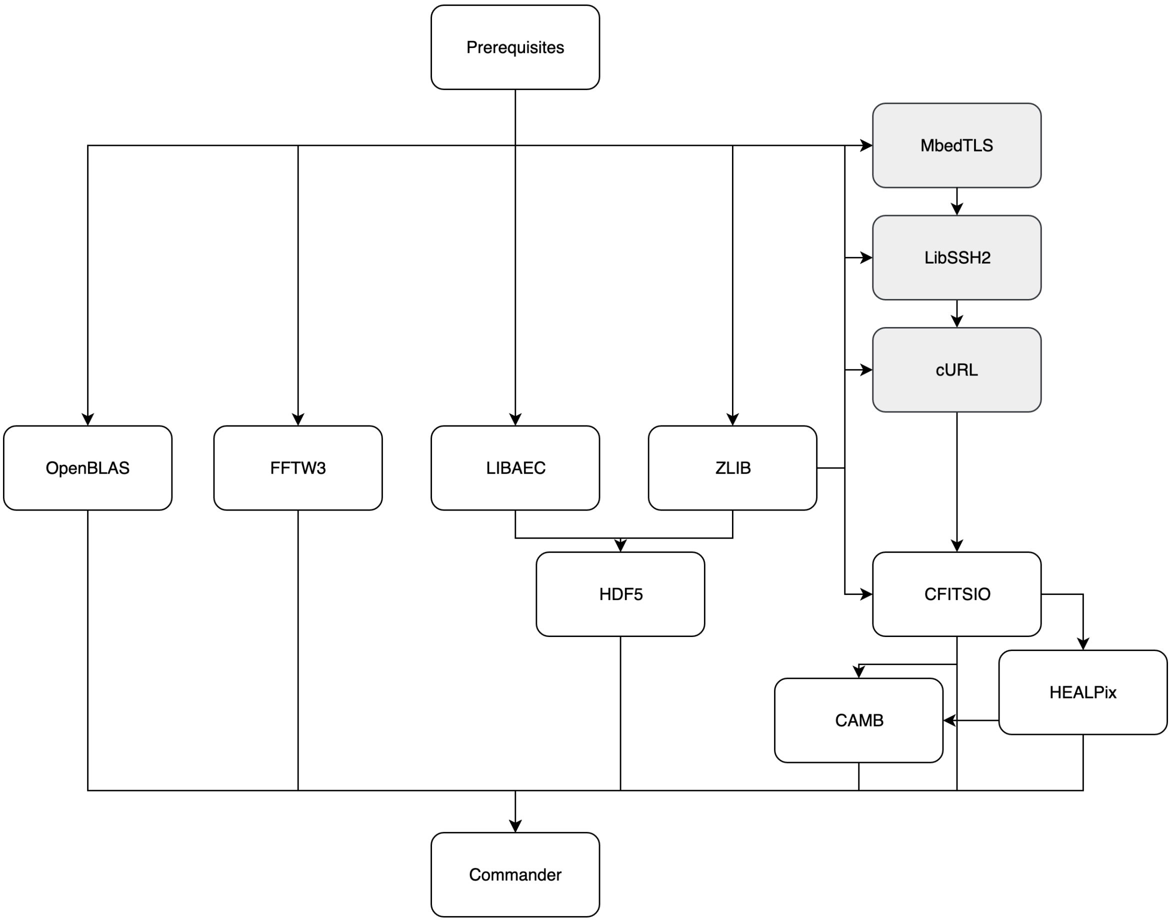

FIG. 5.— An example of the `CMake` `Commander` workflow. The arrows represent the dependence and the linking, while the libraries on the same level will be compiled in parallel. The libraries in grey are optional dependencies. For example, `HDF5` is dependent on both `ZLIB` and `LIBAEC`. The same goes for `CFITSIO` which may (or may not) be compiled with `cURL` support depending on the user's preference. Since `cURL` is dependent on a bunch of other libraries they also need to be incorporated into the build even though `Commander` does not require them explicitly.

scripts and `Makefile` to compile.

Build: In this stage, we use the default build tool as in the rest of the project.

Install: The subproject is installed to a local directory specified by the user during the `CMake` configuration stage. It is worth noting that not all projects (e.g., `HEALPix`) support an explicit install command. We simply copy the compiled binaries and libraries into the specified directory in these cases.

These steps can either be sequential or parallel. In the former case, each library is built sequentially, while the latter allows some libraries to be built in parallel. This idea is illustrated in Fig. 5. It is important to note that this is not the only the way to compile `Commander`since `OpenBLAS` can be substituted for, e.g., Intel `MKL`, which will be detected by `CMake` if present.

4.2.3. *Installation regimes*

`CMake` has various build types defined by default that allows for different optimization categories. We are calling these "installation regimes" with four of these currently supported:

`Release:` Builds `Commander` with the most aggressive optimization flags enabled, tuned for each specific compiler and platform. At the time of writing this paper, only Intel and GNU compilers were supported.

`Debug:` Builds the `Commander` executable without any optimization, but with debug symbols.

`RelWithDebInfo:` ("Release With Debug Information"). A compromise between the two above, building the `Commander` binary with less aggressive optimizations and with debug symbols.

`MinSizeRel:` ("Minimal Size Release"). Builds the `Commander` executable with optimizations that do not increase object code size. However, as the current software targets HPCs with ample disk space, this feature is not used frequently, and it is also accordingly not thoroughly tested.

While we have defined `RelWithDebInfo` to be the default one, the installation regime is determined by the user and his/her needs and can be changed via specifying `CMAKE_BUILD_TYPE` variable in the command line. In general, all external libraries are produced in `Release` format with all optimization flags enabled. This can, however, be changed on a case-by-case basis by editing the `CMake` source files mentioned above.

5. DOCUMENTATION, QUICKSTART GUIDE, AND ACCESSIBILITY TOOLS

For an Open Science project to succeed and continue to grow, making the source code and the data open to the general public is not sufficient. It is also critical that the framework is easy to use and adapt from the user's perspective. Therefore, in this section, we provide a QuickStart guide to the `Commander` documentation and compilation, as well as discuss some valuable tools that make it easy for new users to get "up-to-speed" quickly. We note, however, that this is a continuous work-in-progress; thus, the section is only intended to give a snapshot of the situation at the time of publication.

While the Cosmoglobe framework is designed to simultaneously handle data from different experiments, it is helpful to consider the specific BeyondPlanck pre-processing and analysis pipeline in greater detail, as this will typically serve as a point of reference for most users, whether they want to reproduce the LFI work or generalize the framework to other datasets. In the following, we, therefore, give an overview of the BeyondPlanck installation procedure, but note that most of these steps will be identical for any `Commander`-based analysis.

5.1. *Online documentation*

The documentation for the BeyondPlanck pipeline and `Commander` are currently available on the `Commander` documentation page of the official Cosmoglobe `GitHub` repository[42]. It consists of five main sections, namely (1) Overview (2) QuickStart and installation guide; (3) The `Commander` parameter file; (4) File formats; and (5) Frequently asked questions.

Due to the dynamic nature of the `Commander` and Cosmoglobe projects, this documentation will continually evolve with the addition of new features or experimental datasets. Hence, active participation by the community is essential for its maintenance and expansion. Relatedly, it is also worth noting that because of the high `Commander` development rate, sometimes developers forget to document newly added parameters, and the documentation then becomes outdated. When this happens, the code may request parameters that are not explicitly mentioned anywhere. In these cases, we strongly recommend that the external user notifies the core developers by opening a `GitHub` issue – or, better yet, corrects the documentation and submits the improved version in the form of a `pull request`.

5.2. BeyondPlanck *QuickStart guide*

In the ideal case, installing the BeyondPlanck analysis framework can be done in four simple steps:

```
> git clone https://github.com/Cosmoglobe/Commander.git
> mkdir Commander/build && cd Commander/build
> cmake -DCMAKE_INSTALL_PREFIX=$HOME/local \
        -DCMAKE_C_COMPILER=icc \
        -DCMAKE_CXX_COMPILER=icpc \
        -DCMAKE_Fortran_COMPILER=ifort \
        -DMPI_C_COMPILER=mpiicc \
        -DMPI_CXX_COMPILER=mpiicpc \
        -DMPI_Fortran_COMPILER=mpiifort \
        ..
> cmake --build . --target install -j 8
```

The first line downloads the `Commander` from the official `Git` repository, while the second line creates the aforementioned `build` directory in which compiled binaries and libraries will be stored. The third line collects information regarding the system and auto-generates `Makefiles` inside `build`. Finally, the last line downloads any potentially missing external dependencies and compiles both these and the main source code[43]. The whole process typically takes less than 10 minutes. We have used Intel compilers in this example, but GNU ones have also been successfully tested.

Once `Commander` is installed, the second step is to download the required input data. The number of different files required for a complete BeyondPlanck run can be somewhat intimidating at first sight. To solve the problem, we have implemented a small `Python` utility called `bp` that helps new users to download all required data with a single command:

```
> bp download all
```

This creates a complete directory structure with all required inputs, which amounts to more than 1 TB of data. With a modest download speed of 10 MB/s, this can take some time before completion. The tool also supports downloading individual sub-directories in case the user only requires a subset of the total data.

The third step is to edit the `Commander` parameter file. As described by Galloway et al. (2022a) and in the `Commander` documentation, this is a human-readable ASCII file. It is the step with the steepest learning curve in the process, as the number of `Commander` parameters is quite significant, and a typical parameter file spans several thousands of lines. To address the issue, we have enabled support for the nested include statements allowing for rarely used parameters to be hidden from the user. The downside of this approach is that the special environment variable should be defined in the user's shell[44] for everything to work. Even more so, although very helpful in most cases, such abstraction can still lead to difficult-to-debug errors and become a potential time spender since it requires considerable experience to debug `Commander` parameter files efficiently. A good strategy, in this case, is to start with a well-tested case (such as the final BeyondPlanck parameter file) and only make a few changes between each test run, carefully visually inspecting all outputs at each step while gradually building intuition regarding the code outputs.

The fourth and final step is to actually run the code, which is typically done through an `MPI` runtime environment:

```
> mpirun -n {ncore} path/to/commander param.txt
```

The runtime for a given job varies wildly depending on the parameter file and computing facilities. Still, for the default BeyondPlanck parameter file and a 128-core cluster, it takes about 1 hour and 40 minutes to produce one single sample (Galloway et al. 2022a). For a full Monte Carlo posterior exploration that requires thousands of samples, the end-to-end wall-time is typically on the order of months.

This QuickStart guide represents the ideal case where everything works out of the box. At the time of writing this paper, we estimated that the framework had been successfully installed on at least 20 independent computer systems – and, unfortunately, the default process outlined above worked without modifications in no more than half of these. In the remaining cases, various issues popped up because of compiler idiosyncrasies, missing (or wrong version of) system utilities, insufficient user permissions or disk space, etc. To solve such issues when they arise and improve the current tools, it is vital to have a deeper understanding of all parts of the process, which is the main topic of the rest of the section.

---

[42] https://github.com/Cosmoglobe

[43] After installing `Commander`, the user needs to update the shell to point out the location of the recently installed `HEALPix` library. This is done by `exporting` the `HEALPIX` variable inside `.bashrc` (or similar file).

[44] It is called `COMMANDER_PARAMS_DEFAULT`, and it should point to the `<commander_root>/commander3/parameter_files/defaults` inside `.bashrc` (or other shell files if applicable.)

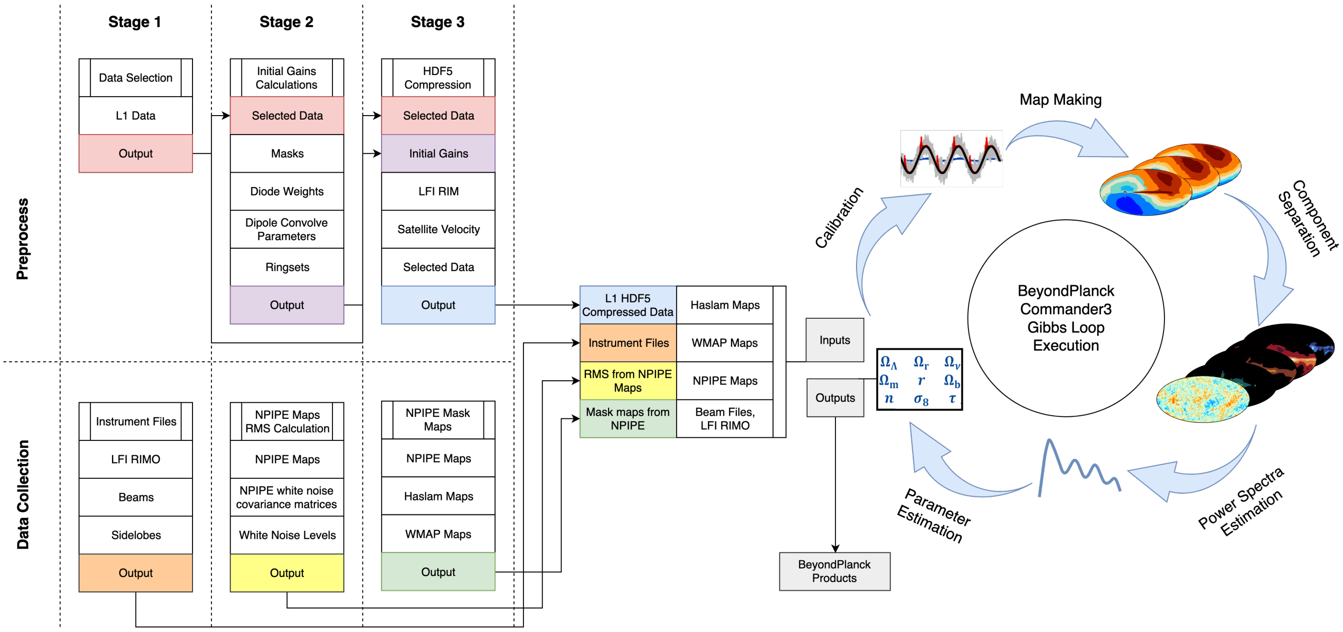

Fig. 6.— Schematic overview of the BeyondPlanck pipeline pre-processing and initialization stages, along with all the input files required for each particular stage.

5.3. BeyondPlanck *pre-processing and initialization*

To understand the whole BeyondPlanck analysis process, it is helpful to take a high-level look at the entire pipeline. This is schematically illustrated in Fig. 6. The heart of this pipeline is the `Commander3` execution, discussed in the previous section and illustrated here by the rightmost analysis loop. This is where the actual posterior sampling takes place (see Sect. 2.2), and it is implemented in terms of a ∼ 60 000 line `Fortran` code, as described by Galloway et al. (2022a)

However, `Commander` requires a significant number of input data objects in order to run, as illustrated by the various small boxes to the left in the figure. These include (1) the raw *Planck* LFI Level-1 data (light blue box; Planck Collaboration II 2020); (2) a so-called "instrument file" (orange box); (3) external and ancillary data that need no pre-processing (white boxes); and (4) external or ancillary data that do need slight preprocessing to conform with `Commander` conventions (colored boxes).

Going through these in order of low to high complexity, the white boxes represent external sky maps and ancillary data that may be used directly in their original forms. These include the frequency maps from *Planck* HFI, *WMAP*, and Haslam 408 MHz, and various instrument characterization such as beam files and the *Planck* LFI Reduced Instrument Model (RIMO). In many cases, these may be simply downloaded directly from external repositories, such as the *Planck* Legacy Archive[45] or LAMBDA[46], and inserted into `Commander` in their original form.

However, some information needs to be slightly preprocessed to match the format expected by the `Commander`. One example is the white noise specification per frequency band (yellow boxes), for which `Commander` expects the user to provide a standard deviation per pixel, whereas the official *Planck* products provide a per-pixel 3 × 3 covariance matrix. As such, the external user needs to reformat the *Planck* format into the `Commander` format (or, better yet, implement support for the native *Planck* format directly into `Commander`, and submit a `Git` pull request).

Another important example is masks (green boxes), which are used at various stages during the `Commander` processing. These may be defined differently whether one is considering correlated noise, gain, bandpass, foreground, or CMB estimation (e.g., Ihle et al. 2022; Gjerløw et al. 2022; Colombo et al. 2022; Paradiso et al. 2022; Andersen et al. 2022; Svalheim et al. 2022b). These masks are typically based both on external sky maps (e.g., *Planck* HFI or *WMAP*) and internal results from a previous `Commander` iteration (e.g., χ^2 maps), and properly optimizing these is an important (and non-trivial) task for any `Commander` user.

The third data collection box represents the so-called `Commander` instrument file. This plays a similar role as the RIMO in *Planck*and contains detailed instrument information for a given frequency channel and detector. This includes objects that are general for all detectors, such as bandpasses and beams (Galloway et al. 2022b; Svalheim et al. 2022a), but also instrument-specific objects such as ADC correction tables (Herman et al. 2022).

However, the most significant and crucial pre-processing step is the preparation of the actual raw time-ordered data, as indicated by the three "Preprocess" stages. These data are stored in compressed `HDF5` files (Galloway et al. 2022a), and include everything from raw detector readouts, pointing, flags, and satellite velocity to initial gain and noise estimates per *Planck* pointing period.

To improve both user-friendliness and reproducibility in each of the above steps, it is useful to define scripts that perform all these tasks for the user. Within the `Commander` repository, we have therefore provided a series of (primarily `Python`) scripts that perform each of these operations, from mask and instrument file generation to full Level-1 data processing. These are intended to serve as useful starting points for users who seek to reproduce the current BeyondPlanck LFI processing and for users who want to analyze a com-

---

[45] https://pla.esac.esa.int/

[46] https://lambda.gsfc.nasa.gov

pletely new dataset with the same framework. If so, it would be greatly appreciated if the new scripts are also committed to the existing repository as part of the community-driven Open Source activities.

Before concluding this section, it is worth noting that `Commander` has, in general, very few means of validating a given input data product. If, say, some instrument specification or compressed `HDF` files are inter-mixed, there is no automatic way for the algorithm to discover this except through visual inspection of the final results and goodness-of-fit statistics. Furthermore, such parameter file errors are likely among the most common and time-consuming errors made when running this code. For most analyses, it is, therefore, useful to start with the set of well-tested parameter and input files provided in the `Commander` repository (Galloway et al. 2022a), which includes individual parameter files for a wide range of common datasets (*Planck* LFI and HFI, *WMAP*, Haslam 408 MHz, etc.) and astrophysical components (CMB, synchrotron, thermal dust emission, etc.). These may be used as "building blocks" when constructing a new analysis configuration.

5.4. *Docker environment for user-friendly data access and code exploration*

We provide a precompiled `Ubuntu`-based `Docker` image for users who are not interested in computationally expensive analyses like BeyondPlanck but simply want to run `Commander` on a small dataset for which computational efficiency is not paramount. This self-contained operating-system-level virtual container can be run on any OS (`Mac`, `Windows`, `Linux`, etc.), and all dependencies are maintained within the `Docker` image itself. Running `Commander` in this mode amounts to one single command line:

```
> docker run -it \
  -v {input_dir}:{input_dir} \
  -v {output_dir}:/output \

  registry.gitlab.com/beyondplanck/r13y-helper/cm3 \
  commander3 {parameter file}
```

where `input_dir` is a directory that contains all required input data, and `output_dir` is an empty directory that will contain the results.

We emphasize, however, that the binary provided in this Docker image is not optimized for any processor type. Therefore, it is computationally less efficient than a natively compiled version. Also, debugging this version is non-trivial since it may be challenging to recompile the binary with different debug flags or source code changes.

6. DISCUSSION

The successes achieved in modern CMB cosmology during the last decades are a solid testament to the ingenuity and dedication of thousands of instrumentalists, observers, data analysts, and theorists. However, these same successes are also a direct product of long-standing and invaluable financial support from ordinary taxpayers. A typical CMB satellite mission costs several hundreds of millions of dollars, euros, or yen. At the same time, ground-based and sub-orbital experiments typically cost from a few to many tens of millions of dollars – and the massive next-generation ground-based CMB-S4 experiment is anticipated to cost $ 600 M.

With steadily rising costs for each generation of experiments, it also becomes even more critical to optimally leverage the investments already made from previous efforts. For instance, it makes very little sense for future experiments to reproduce the temperature sensitivity of *Planck*. Instead, they should aim to provide complementary information that may be combined with the *Planck* measurements, typically in polarization or on small angular scales. Likewise, it makes very little sense for a future satellite mission, such as *LiteBIRD*, to measure small angular scales from space when this can be done much more economically from the ground at a much lower cost, for instance, with CMB-S4.

In this paper, we argue that the most efficient way to move forward as a field is precisely through an integrated joint global analysis of complementary datasets. However, several prerequisites must be in place for this to be possible. First and foremost, researchers in the various teams actually need to have physical access to data from other experiments. Traditionally, this has been achieved through dedicated "Memoranda of Understanding" (MoUs) between pairs of collaborations; the ground-breaking joint analysis of *Planck* and BICEP2 is a well-known example of this (BICEP2/Keck Array and Planck Collaborations 2015). While this works reasonably well for limited two-party cases, we believe that this approach is impractical for future work when many datasets must be involved in the same analysis to obtain optimal results, for instance, when combining proprietary data from *LiteBIRD* (Sugai et al. 2020), CMB-S4 (Abazajian et al. 2019), C-BASS (Jones et al. 2018), QUIJOTE (Génova-Santos et al. 2015), and PASIPHAE (Tassis et al. 2018) with public data from *Planck* and *WMAP*. Rather, we believe the time is overdue to fundamentally change how the CMB field works and move to a fully Open Science mode of operation where both raw data and end-to-end analysis methods are shared between experiments and research groups. This also guarantees that taxpayers and funding agencies get maximum value for money, as it vastly increases the longevity of any given dataset. Hence, the funding agencies should require an Open Source release of raw data, analysis tools, and high-level products for any future experiment.

A second prerequisite is using practical analysis methods to exploit the information stored in these complementary datasets. Establishing one set of such common tools is the primary goal of the BeyondPlanck and Cosmoglobe efforts. Our implementation is based on well-established Bayesian parameter estimation techniques and builds directly on the `Commander` code developed for and used by *Planck*. This software is released under a permissive Open Source GPL license, ensuring that any researchers may use, generalize, and modify the code as they see fit (BeyondPlanck 2022; Galloway et al. 2022a). The only requirement is that these modified versions must also be released under an equally permissive license, ensuring that other scientists may then also benefit from the extended work.

However, it is still not sufficient that the data and analysis codes are publicly available. They must also be *accessible*, and that is the main topic of the current paper: For other scientists to be able to leverage this work in practice, the software must be appropriately documented, and it must be straightforward to install on a range of different computer systems. These practical aspects may seem somewhat mundane compared to the more spectacular topics typically addressed in astrophysics and cosmology papers – but they are no less important in this era of mega-science. These aspects also require significant dedicated resources to be successful, and BeyondPlanck dedicated as much as 20 % of its budget (or 300 k€) to this work. On the other hand, these resources do not scale linearly with the total budget size. However, we still strongly recommend future experiments allocate significant funding for reproducibility and Open Source dissemination in their proposal budgets. We hope that the current paper may serve as a valuable and thought-provoking starting point for future large experiments and collaborations that are likely to face similar issues.

ACKNOWLEDGEMENTS

We thank Prof. Pedro Ferreira and Dr. Charles Lawrence for useful suggestions, comments and discussions. We also thank the entire *Planck* and *WMAP* teams for invaluable support and discussions, and for their dedicated efforts through several decades without which this work would not be possible. The current work has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement numbers 776282 (COMPET-4; BeyondPlanck), 772253 (ERC; bits2cosmology), and 819478 (ERC; Cosmoglobe). In addition, the collaboration acknowledges support from ESA; ASI and INAF (Italy); NASA and DoE (USA); Tekes, Academy of Finland (grant no. 295113), CSC, and Magnus Ehrnrooth foundation (Finland); RCN (Norway; grant nos. 263011, 274990); and PRACE (EU). Some of the results in this paper have been derived using the `healpy` (Zonca et al. 2019) and `HEALPix` (Górski et al. 2005) packages.

REFERENCES

Abazajian, K., Addison, G., Adshead, P., et al. 2019, arXiv e-prints, arXiv:1907.04473
Andersen et al. 2022, A&A, in press [2201.08188]
Basyrov et al. 2022, A&A, submitted [2208.14293]
Bennett, C. L., Larson, D., Weiland, J. L., et al. 2013, ApJS, 208, 20
BeyondPlanck. 2022, A&A, submitted [2011.05609]
BICEP2/Keck Array and Planck Collaborations. 2015, Phys. Rev. Lett., 114, 101301
Colombo et al. 2022, A&A, submitted [2208.14276]
Eriksen, H. K., Jewell, J. B., Dickinson, C., et al. 2008, ApJ, 676, 10
Eriksen, H. K., O'Dwyer, I. J., Jewell, J. B., et al. 2004, ApJS, 155, 227
Fixsen, D. J. 2009, ApJ, 707, 916
Galloway et al. 2022a, A&A, in press [2201.03509]
Galloway et al. 2022b, A&A, in press [2201.03478]
Génova-Santos, R., Rubiño-Martín, J. A., Rebolo, R., et al. 2015, MNRAS, 452, 4169
Gjerløw et al. 2022, A&A, submitted [2011.08082]
Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759
Haslam, C. G. T., Salter, C. J., Stoffel, H., & Wilson, W. E. 1982, A&AS, 47, 1
Hauser, M. G., Arendt, R. G., Kelsall, T., et al. 1998, ApJ, 508, 25
Herman et al. 2022, A&A, in press [2203.13635]
Ihle et al. 2022, A&A, in press [2011.06650]
Jarosik, N., Barnes, C., Greason, M. R., et al. 2007, ApJS, 170, 263
Jones, M. E., Taylor, A. C., Aich, M., et al. 2018, MNRAS, 480, 3224
Kamionkowski, M. & Kovetz, E. D. 2016, ARA&A, 54, 227
LiteBIRD Collaboration, Allys, E., Arnold, K., et al. 2022, arXiv e-prints, arXiv:2202.02773
Paradiso et al. 2022, A&A, submitted [2205.10104]
Planck Collaboration III. 2016, A&A, 594, A3
Planck Collaboration I. 2020, A&A, 641, A1
Planck Collaboration II. 2020, A&A, 641, A2
Planck Collaboration III. 2020, A&A, 641, A3
Planck Collaboration IV. 2018, A&A, 641, A4
Planck Collaboration V. 2020, A&A, 641, A5
Planck Collaboration VI. 2020, A&A, 641, A6
Planck Collaboration Int. XLVI. 2016, A&A, 596, A107
Ruud, T. M., Fuskeland, U., Wehus, I. K., et al. 2015, ApJ, 811, 89
Seljebotn, D. S., Bærland, T., Eriksen, H. K., Mardal, K. A., & Wehus, I. K. 2019, A&A, 627, A98
SPIDER Collaboration, Ade, P. A. R., Amiri, M., et al. 2021, arXiv e-prints, arXiv:2103.13334
Sugai, H., Ade, P. A. R., Akiba, Y., et al. 2020, Journal of Low Temperature Physics, 199, 1107
Svalheim et al. 2022a, A&A, in press [2201.03417]
Svalheim et al. 2022b, A&A, in press [2011.08503]
Tassis, K., Ramaprakash, A. N., Readhead, A. C. S., et al. 2018, arXiv e-prints, arXiv:1810.05652
Watts et al. 2022, A&A, in press [2202.11979]
Zonca, A., Singer, L., Lenz, D., et al. 2019, Journal of Open Source Software, 4, 1298

Paper VI

BeyondPlanck IV. On end-to-end simulations in CMB analysis — Bayesian versus frequentist statistics

M. Brilenkov, K. S. F. Fornazier, L. T. Hergt, G. A. Hoerning, A. Marins, T. Murokoshi, F. Rahman, N.-O. Stutzer, Y. Zhou, F. .B. Abdalla, and the BEYONDPLANCK Collaboration.

Submitted to *A&A*, arXiv: 2209.04437v1.