Maximum a-posteriori Lensing Reconstruction for CMB Science

Belkner, Sebastian

How to cite

BELKNER, Sebastian. Maximum a-posteriori Lensing Reconstruction for CMB Science. Doctoral Thesis, 2024. doi: 10.13097/archive-ouverte/unige:181650

This publication URL: https://archive-ouverte.unige.ch/unige:181650
Publication DOI: 10.13097/archive-ouverte/unige:181650

UNIVERSITÉ DE GENÈVE
Section de Physique
Department de Physique Théorique

FACULTÉ DES SCIENCES
Professeur Julien Carron

Maximum a-posteriori Lensing Reconstruction for CMB Science

THÈSE

présentée à la Faculté des sciences de l'Université de Genève en vue de l'obtention du grade de Docteur ès sciences en physique

par
Sebastian Belkner
de
Genève (Switzerland)

Thèse N°5856

GENÈVE
Atelier de reproduction de la Section de Physique
2024

UNIVERSITÉ
DE GENÈVE
FACULTÉ DES SCIENCES

DOCTORAT ÈS SCIENCES, MENTION PHYSIQUE

Thèse de Monsieur Sebastian BELKNER

intitulée :

«Maximum a-posteriori Lensing Reconstruction for CMB Science»

La Faculté des sciences, sur le préavis de

Monsieur J. CARRON, professeur assistant et directeur de thèse
Département de physique théorique

Monsieur A. W. RIOTTO, professeur ordinaire
Département de physique théorique

Monsieur J. ERRARD, docteur
Astroparticule et cosmologie, Université Paris Cité, Paris, France

Monsieur A. LEWIS, professeur
Department of Physics & Astronomy, University of Sussex, Brighton, United Kingdom

autorise l'impression de la présente thèse, sans exprimer d'opinion sur les propositions qui y sont énoncées.

Genève, le 6 novembre 2024

Thèse - 5856 -

La Doyenne

N.B. - La thèse doit porter la déclaration précédente et remplir les conditions énumérées dans les "Informations relatives aux thèses de doctorat à l'Université de Genève".

Abstract

This thesis is dedicated to advancing our understanding of the origin and evolution of the Universe by studying the cosmic microwave background (CMB), a light travelling through the cosmos since the dawn of time. The CMB contains rich information about the earliest time of the Universe, and informs us about the Universe as a whole due to it containing an imprint of the mass distribution of the Universe. We study this imprint through the development and application of a novel curved sky Bayesian analysis method. This comprehensive approach, inclusive of carefully treating biases, is designed to effectively handle real observational data and is aimed at tackling cosmological phenomena captured through the CMB.

Central to this thesis is the improvement of a Bayesian method, dubbed the Maximum a-posterior (MAP) lensing reconstruction. The MAP method optimally recovers the lensing information and is therefore an improvement over standard quadratic estimators (QE). Our Bayesian approach iteratively finds the MAP point of the deflection field, which makes this approach computationally more expensive than the QE. This is in particular true due to the full sky treatment for which spherical harmonics are the basis functions; their numerical calculation is more expensive than their flat sky counterpart, the Fourier basis.

An important element of our approach is the lensing operator, whose optimization is vital for the efficiency and accuracy and which is not present in QE method. The lensing operator describes the mapping between the unlensed and the lensed CMB field. In this context, we discuss the notion of non-uniform Spherical Harmonic Transforms (SHT), and non-uniform Fast Fourier Transforms (nuFFT), and show how to approximate the former at machine precision accuracy. We introduce two novel algorithms that significantly expedite this calculation: one optimized for CPU and the other for GPU architectures. Both implementations achieve machine precision accuracy and fast computations, marking a substantial improvement over existing methods.

We demonstrate our approach in the context of the analysis of early-time physics. Our methodology plays a pivotal role in probing inflationary theories by mitigating the contamination known as B-lensing. With many current and upcoming experiments probing scales of the CMB polarization that are obscured by B-lensing, an efficient and optimal mitigation, a so called delensing, is desired. To reach the accuracy needed to rule out current viable inflationary models, and to fully utilize the information accessible by current experiments sensitivity, Bayesian methods are preferred over QE. Delensing is crucial for refining measurements of the tensor-to-scalar ratio, a fundamental parameter in cosmology. Our analyses demonstrate the application onto realistic simulations for upcoming experiments and its integration into state-of-the-art CMB analysis pipelines: we present the evaluation on different sky models using various foregrounds, inclusive of inhomogeneous noise, and a masked sky. We provide means to mitigate mask effects such as the mean-field.

We also demonstrate our approach in the context of late-time physics that covers non-linear structure growth, and post-Born effects and analyse their impact on the Bayesian reconstruction. By reconstructing a non-Gaussian lensing power spectrum, we assess its induced biases, providing new insights into the reconstruction capabilities of our Bayesian approach.

Furthermore, we present `Delensalot`, a software further developed during this thesis, crafted to enable the broader cosmology community to use optimal lensing reconstruction. `Delensalot` is designed with the intent to lower the barrier to entry for researchers not specializing in lensing or unfamiliar with Bayesian techniques. While further enhancements are possible, `Delensalot` stands as a robust, user-friendly tool that is ready for practical application in real-world scenarios.

The integration of Bayesian methods into the CMB data analysis pipelines of major collaborations like CMB-S4 and PICO underscores the timeliness and relevance of this work. Our contributions not only push forward the frontier of cosmological research but also equip the scientific community with powerful, accessible tools to explore the cosmos.

Acknowledgements

I would like to express my deepest gratitude to my advisor, Professor Julien Carron, for his invaluable support, patience, and guidance throughout my research. His insight and encouragement have been instrumental in the completion of this work. Juliens' profound knowledge and expertise in cosmology have greatly shaped my understanding of the field. His unwavering commitment to excellence has inspired me to push the boundaries of my own work.

I am particularly thankful for the countless hours he spent discussing ideas, reviewing drafts, and providing thoughtful feedback. His ability to challenge me while offering support and encouragement helped me grow both as a researcher and as an individual. His mentorship extended far beyond the technical aspects of my research; he has been a constant source of motivation and has played a crucial role in my professional development.

Juliens's optimism and belief in my abilities never wavered. I am truly fortunate to have had the opportunity to work under his supervision, and I will carry the lessons learned from him throughout my career. For all these reasons and more, I am deeply indebted to him and express my sincerest thanks.

I would also like to thank the postdoctoral researchers in our group, Louis Legrand, and Omar Darwish, for their continuous support, stimulating discussions, and assistance in various aspects of this research. Their expertise and willingness to help have greatly contributed to the progress of this work.

Additionally, I extend my appreciation to the members of the CMB-S4 collaborations. Their collaborative efforts, discussions, and shared knowledge have provided a strong foundation and enriched my understanding of the field. Special thanks to Clement Pryke, Jacques Delabrouille, Colin Bischoff, Reijo Keskitalo, Kimmy Wu, Dominic Beck, Federico Bianchini for their contributions and feedback.

Furthermore, I want to express my appreciation to the members of the PICO collaboration. In particular, Shaul Hanany, Shamik Ghosh, and Jacques Delabrouille.

My heartfelt thanks go to my collaborators Martin Reinecke, Nathanael Schaeffer, and Adriaan Duivenvoorden, for their invaluable contributions, insightful discussions, and dedicated efforts. Working with you has been a rewarding experience.

I extend my appreciation to my colleagues in Geneva and colleagues from the Simons Foundation and their working staff for their support and collaborative spirit. Your contributions and shared knowledge have been greatly beneficial.

I would also like to acknowledge the people I met at various conferences for their inspiring conversations and the new perspectives they brought to my research.

A special mention goes to my colleagues in Geneva, Goran Yelcic, Davide Piras, Niccolò Muttoni, Francesco Sorrenti, Ahmadreza Nourizonoz, Martin Pijenburg, and Francesco Iacovelli. Our volleyball games, bouldering sessions, barbecues, and various social gatherings have provided me with a much-needed balance between work and relaxation. Your friendship and camaraderie have been a source of joy and motivation.

I cannot thank enough the other pre-docs of the pre-doctoral program of the Simons Foundation for their friendship during our time in New York. Harrison Souchereau, Arianna Arna, Andrew Powers, Richard Brooks, Gabe Schumm, Hang Le, Leah Zuckermann, Marc Ritter, Rachel Zang, Caprice Phillips, I could not think of a better group to share this experience together, as we so quickly built a strong friendship. I will never forget our road trip to the solar eclipse, the visits to the Broadway shows, and our breakfasts at work.

A special acknowledgment goes to my best friend, Dirk Schoene, and Torsten Schoene. Their support and encouragement were invaluable throughout this journey. Dirk's friendship and Torsten's advice have been a constant source of strength and motivation, for which I am deeply grateful.

Lastly, I wish to extend my deepest gratitude to a very special person, Zhenming Peng, for his unwavering support and encouragement throughout this thesis journey. Your presence has been a constant source of comfort and strength, providing me with the resilience needed to persevere through challenging times. Your support, patience, and understanding played an essential role in helping me stay focused and balanced during this demanding process. For that, I am deeply thankful.

This work would not have been possible without the collective efforts and support of everyone involved. I am truly grateful for the opportunity to work with such a dedicated and talented group of individuals, and to have such great mentors, colleagues and friends.

Jury Members

- **Prof. Julien Carron**

  Département de Physique Théorique & Center for Astroparticle Physics, Université de Genève, 24 Quai E. Ansermet, CH-1211 Genève 4, Switzerland

- **Prof. Antonio Riotto**

  Département de Physique Théorique & Center for Astroparticle Physics, Université de Genève, 24 Quai E. Ansermet, CH-1211 Genève 4, Switzerland

- **Josquin Errard, PhD**

  Université de Paris, CNRS, Astroparticule et Cosmologie, F-75013 Paris, France

- **Prof. Antony Lewis**

  Department of Physics & Astronomy, University of Sussex, Brighton BN1 9QH, UK

Thank you all for having accepted to be part of the jury for my thesis defense.

List of Publications

The following works have been considered as part of this thesis:

[1] S. Belkner, A. J. Duivenvoorden, J. Carron, N. Schaeffer, and M. Reinecke. "cunuSHT: GPU Accelerated Spherical Harmonic Transforms on Arbitrary Pixelizations". In: (2024). arXiv: `2406.14542 [astro-ph.IM]`.

[2] S. Belkner et al. "CMB-S4: Iterative Internal Delensing and r Constraints". In: *Astrophys. J.* 964 (2024), p. 148. DOI: `10.3847/1538-4357/ad2351`. arXiv: `2310.06729 [astro-ph.CO]`.

[3] O. Darwish, S. Belkner, L. Legrand, J. Carron, and G. Fabbian. "Non-Gaussian deflections in iterative optimal CMB lensing reconstruction". In: (2024). arXiv: `2407.00228 [astro-ph.CO]`.

[4] M. Reinecke, S. Belkner, and J. Carron. "Improved cosmic microwave background (de-)lensing using general spherical harmonic transforms". In: *Astron. Astrophys.* 678 (2023), A165. DOI: `10.1051/0004-6361/202346717`. arXiv: `2304.10431 [astro-ph.CO]`.

[5] R. Aurlien, M. Remazeilles, S. Belkner, J. Carron, J. Delabrouille, H. K. Eriksen, R. Flauger, U. Fuskeland, M. Galloway, K. M. Gorski, et al. "Foreground separation and constraints on primordial gravitational waves with the PICO space mission". In: *JCAP* 06 (2023), p. 034. DOI: `10.1088/1475-7516/2023/06/034`. arXiv: `2211.14342 [astro-ph.CO]`.

Résumé

Cette thèse est dédiée à l'avancement de notre compréhension de l'origine et de l'évolution de l'Univers en étudiant le fond diffus cosmique (CMB), une lumière qui voyage à travers le cosmos depuis la nuit des temps. Le CMB contient de riches informations sur les premiers instants de l'Univers et nous renseigne sur l'Univers dans son ensemble car il porte une empreinte de la distribution de masse de l'Univers. Nous étudions cette empreinte à travers le développement et l'application d'une nouvelle méthode d'analyse bayésienne pour le ciel courbé. Cette approche globale, incluant un traitement soigné des biais, est conçue pour gérer efficacement les données d'observation réelles et vise à aborder les phénomènes cosmologiques capturés à travers le CMB.

L'amélioration d'une méthode bayésienne, appelée reconstruction de lentille par Maximum A-Posteriori (MAP), constitue le cœur de cette thèse. La méthode MAP récupère de manière optimale l'information sur les lentilles, constituant ainsi une amélioration par rapport aux estimateurs quadratiques standards (QE). Notre approche bayésienne trouve itérativement le point MAP du champ de déviation, ce qui rend cette méthode plus coûteuse en termes de calcul par rapport à l'estimateur quadratique, surtout pour le traitement du ciel complet où les harmoniques sphériques servent de fonctions de base.

Un élément crucial de notre approche est l'opérateur de lentille, dont l'optimisation est cruciale pour l'efficacité et la précision, et qui n'est pas présent dans la méthode QE. L'opérateur de lentille décrit la correspondance entre le champ du CMB non lentillé et le champ lentillé. Dans ce contexte, nous discutons des notions de transformée harmonique sphérique non uniforme (SHT) et de transformée de Fourier rapide non uniforme (nuFFT), et nous montrons comment approximer la première avec une précision de machine. Nous introduisons deux nouveaux algorithmes qui accélèrent considérablement ce calcul : l'un optimisé pour les architectures CPU et l'autre pour les architectures GPU. Les deux implémentations atteignent une précision de machine et des calculs rapides, ce qui représente une amélioration substantielle par rapport aux méthodes existantes.

Nous démontrons notre approche dans le contexte de l'analyse de la physique des premiers instants de l'Univers. Notre méthodologie joue un rôle essentiel dans l'étude des théories inflationnistes en atténuant la contamination connue sous le nom de *B-lensing.* Avec de nombreuses expériences actuelles et à venir sondant les échelles de

polarisation du CMB qui sont obscurcies par le *B-lensing*, une atténuation efficace et optimale, appelée delensing, est souhaitée. Pour atteindre la précision nécessaire à l'élimination des modèles inflationnistes viables actuels, et pour utiliser pleinement l'information accessible par la sensibilité des expériences actuelles, les méthodes bayésiennes sont préférées au QE. Le delensing est crucial pour affiner les mesures du rapport tenseur-scalaire, un paramètre fondamental en cosmologie. Nos analyses démontrent l'application sur des simulations réalistes pour les expériences à venir et son intégration dans les pipelines d'analyse CMB de pointe : nous présentons l'évaluation sur différents modèles de ciel utilisant divers avant-plans, y compris le bruit inhomogène, et un ciel masqué. Nous proposons des moyens d'atténuer les effets de masque tels que le champ moyen.

Nous démontrons également notre approche dans le contexte de la physique tardive qui couvre la croissance non linéaire de la structure et les effets post-Born et analysons leur impact sur la reconstruction bayésienne. En reconstruisant un spectre de puissance de lentille non gaussien, nous évaluons les biais induits, ce qui donne un nouvel aperçu des capacités de reconstruction de notre approche bayésienne.

En outre, nous présentons *Delensalot*, un logiciel développé au cours de cette thèse, conçu pour permettre à tout le monde d'utiliser la reconstruction optimale de l'effet de lentille. *Delensalot* est conçu dans l'intention de réduire toutes les difficultés que pourraient rencontrer les chercheurs qui ne sont pas spécialisés dans les lentilles ou qui ne sont pas familiers avec les techniques bayésiennes. Bien que d'autres améliorations soient possibles, *Delensalot* est un outil robuste et agréable qui est prêt pour une application pratique dans des scénarios du monde réel.

L'intégration des méthodes bayésiennes dans les pipelines d'analyse des données CMB de collaborations majeures telles que CMB-S4 et PICO souligne l'actualité et la pertinence de ce travail. Nos contributions ne font pas que repousser les frontières de la recherche cosmologique, elles équipent également la communauté scientifique d'outils puissants et accessibles pour explorer le cosmos.

Contents

0

Cosmic Microwave Background lensing Analysis

This chapter introduces the Cosmic Microwave Background (CMB), CMB lensing, lensing reconstruction, and the data analysis pipeline for the search of the tensor-to-scalar ratio r.

0.1 Cosmic Microwave Background

The CMB is the oldest light in the Universe, first theorized by [6] and discovered later by [7]. Originating from the hot, dense plasma of the early Universe, the CMB now fills the cosmos, observed from all directions at a mere few Kelvin. Under the hot Big Bang model, proposed by [8], the Universe began in a hot and dense state and has since expanded to its current size. Initially, the Universe's high temperatures prevented the formation of atoms, leading to a state where electrons and protons existed as a plasma. As the Universe cooled, these particles combined to form neutral hydrogen—and during Big Bang Nucleosynthesis, trace amounts of helium and lithium isotopes—allowing CMB photons to decouple and travel largely undisturbed. This process, known as recombination, occurred approximately 380,000 years after the Big Bang, marking the photons' decoupling from the plasma. This event also defines the last scattering surface, from which these photons now reach us as a sphere with a radius of about 46 billion light years.[1] Additionally, due to Compton scattering and the presence of a quadrupole moment during decoupling, the CMB is slightly polarized, providing a wealth of information. Serving as the earliest electromagnetic snapshot of the Universe, the CMB carries invaluable data about the Universe's infancy, including insights into the inflationary epoch proposed by [9] in 1980.

Inflation describes a commonly acknowledged paradigm proposing that the very early Universe, when it was younger than about 10^{-32} years old, underwent exponential expansion. This explains the observed isotropy of the Universe as detected by various cosmological probes. During this period, primordial microscopic fluctua-

---

[1] While today, the CMB has travelled about 13.4 billion years to reach Earth, cosmic expansion increased its size.

tions were present and expanded to macroscopic scales, contributing to the first order anisotropies seen in the CMB. These fluctuations also serve as the seeds for growth of cosmic structures. Two types of perturbations resulting from the primordial fluctuations - scalar and tensor - are (potentially) observable today. The microscopic scalar fluctuations that were expanded to macroscopic scales initiated the formation of large-scale structures, and leave an anisotropic imprint in the CMB's intensity at the level of $\mathcal{O}(10^{-5})$ and due to the presence of a quadrupole also in polarization. This quadrupole arises due to the scalar perturbations in combination with the expansion of the Universe.

Tensor perturbations that are directly related to primordial gravitational waves [10, 11] are proportional to the energy scale of inflation, a critical parameter for the characterisation of inflationary models. They also leave an imprint in the CMB; due to their nature as spin-2 fields, they also generate a quadrupole in the density distribution of the primordial plasma. This quadrupole is different than the one described above. Thus, primordial gravitational waves generate a very distinct polarization pattern during recombination, different from that generated by scalar perturbations.

Primordial anisotropies in the CMB are analyzed using the CMB power spectrum, a critical statistical tool that quantifies the variance of temperature fluctuations at various angular scales. The power spectrum is computed from the distribution of CMB photons over the celestial 2-sphere. This computation involves spherical harmonic transforms (SHT). In this context, ℓ describes the angular momentum number, representing different angular scales, and m, ranging from $-\ell$ to ℓ specifies each mode's orientation. By averaging the squared amplitudes over all m for each ℓ, the power spectrum is obtained, and illustrates the scale dependence of the variance in the CMB.

The CMB power spectra are fundamental observables providing insights into the condition of the early Universe. Their forms are directly linked to the initial conditions and are shaped by the primordial perturbations. Assuming the Lambda Cold Dark Matter (ΛCDM) model [12], these fluctuations evolved into the observed large-scale structures, and also describes how matter and energy content influence its evolution. It describes the Universe's total energy content as a mix of ordinary matter and radiation, dark matter, and dark energy. This composition affects the expansion and matter waves that propagated through the primordial plasma which thus leaves an imprint in the CMB as temperature and polarization anisotropies.

The connection between the primordial fluctuations and the CMB power spectra is through the Boltzmann equations, which describes the evolution under the influence of gravity and pressure. The equations track how perturbations in density, velocity, and potential evolve over time. By using the parameters from the ΛCDM model to solve the Boltzmann equations, one can simulate the early Universe and obtain the shapes of the power spectra of the CMB, and thus compare this with the observational data.

Through time, the Universe formed rich and complex structures due to the macroscopic perturbations and emerging gravitational pull, rendering the Universe anisotropic and creating mass concentrations that formed into stars, galaxies, galaxy

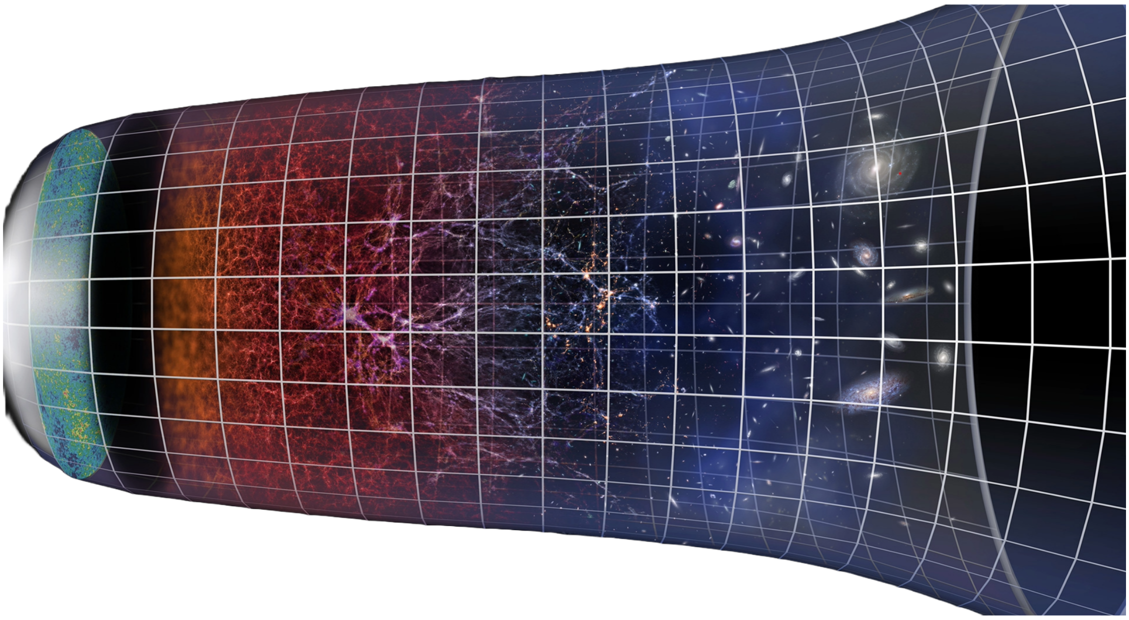

Figure 1: The Cosmic Microwave Background and its journey through the anisotropic cosmos. Copyright: ESO/M. Kornmesser

clusters, and Large Scale Structure filaments. For better or worse, this gravitational pull also affects photons. Due to the interaction via gravitational pull of the intermediate mass distribution, it thus imprints a second order anisotropy in the CMB in the distribution of the CMB photons. This is implicitly shown in Fig. (1), where I show a cutout through the Universe over time. On the far left, we see the primordial CMB. The evolving matter of the Universe becomes more and more anisotropic over time, giving rise to the deflection of the CMB.

CMB lensing is an important tool for cosmology for several reasons listed below and motivates why CMB lensing is appealing for today's research.

Primordial tensor perturbations that give rise to primordial B-modes are one major science goal of many upcoming experiments. CMB lensing supports this by setting tighter constraints on their existence, that can in return also inform about the validity of inflationary models. Our findings are presented for the CMB-S4 and PICO collaboration in chapter 3.

The mapping of the mass distribution is particularly interesting due to CMB lensing being sensitive to dark matter, an elusive content of the Universe that does otherwise not interact lightly. CMB lensing provides a way of mapping the distribution of matter content in the Universe. We discuss this in the context of the CMB-S4 and PICO collaboration and on the search of non-Gaussian biases in chapter 3 and 4.

Neutrino mass impacts the smallest scales of the power spectra of the lensing potential due to their damping effect. Lighter neutrinos, being more relativistic, slow the growth of structures as they escape the gravitational potential well, and thereby smooth out small-scale features in the matter power spectrum and lensing potential. These effects can be quantified with CMB lensing and used to inform us about the neutrino mass, among others. This is discussed in the context on the search of non-Gaussian biases in chapter 4.

Cluster lensing offers a direct method for mapping the distribution of dark matter within clusters. As it is sensitive to the total matter content, it thus also offers means to study the mass profile of galaxy clusters, constraint neutrino mass and the matter density, and informs on feedback processes and cluster dynamics.

Cross-correlation to large-scale-structure and galaxy surveys enhances our understanding of the Universe. This synergy improves measurement accuracies and helps resolve inherent observational degeneracies and pin down systematics. As both are sensitive to similar phenomena, combining these two sources of information can improve uncertainties in the measurements, and break degeneracies due to their intrinsic different nature.

CMB lensing slightly changes the CMB statistics, and therefore obscures a clear picture of the primordial Universe. Accumulating the combined matter along the line-of-sight gives the deflection field $\boldsymbol{\alpha}$ that is sourced by any potential that might be present. We decompose the potentials into the gradient and curl part, $\boldsymbol{\alpha} = \nabla\phi + \nabla\times\Omega$, where ϕ is the lensing potential, and Ω is the curl potential.

Introducing the gradient of the gravitational potential, $\nabla\Psi$, the total deflection is a weighted projection along the line of sight $\hat{\boldsymbol{n}}$ [13],

$$\boldsymbol{\alpha}(\hat{\boldsymbol{n}}) = -2\int_0^{\chi_*} \mathrm{d}\chi W(\chi,\chi_*)\nabla\Psi(\hat{\boldsymbol{n}},\chi)) = \nabla\underbrace{\left(-2\int_0^{\chi_*} \mathrm{d}\chi W(\chi,\chi^*)\Psi(\hat{\boldsymbol{n}},\chi))\right)}_{:=\phi,\ \text{Born approx.}}, \quad (1)$$

with χ the comoving distance and χ_* the source distance. The weighting comes from the lensing efficiency W, which in a flat Universe, is given by,

$$W = \frac{\chi_* - \chi}{\chi_*\chi}\,. \quad (2)$$

To lowest order, and good accuracy, lensing can be described within the Born approximation, which assumes that the integration in Eq. (1) can be performed along the unperturbed photon path. In this approximation, the deflection field is necessarily curl-free, $\nabla\times\boldsymbol{\alpha} = 0$. Post-Born corrections can be included by perturbatively treating the effect onto the background geodesic, which to second order is,

$$\Psi(\hat{\boldsymbol{n}}) = \Psi(\hat{\boldsymbol{n}}_0 + \boldsymbol{\alpha}(\hat{\boldsymbol{n}}_0)) = \Psi(\hat{\boldsymbol{n}}_0) + (\alpha^a(\hat{\boldsymbol{n}}_0)\nabla_a\Psi)(\hat{\boldsymbol{n}}_0)\,, \quad (3)$$

and the index a denotes the coordinate - either the Cartesian $a \in (x,y)$ on the flat sky, or angular positions $a \in (\theta,\phi)$ on the curved sky. Plugging this into Eq. (1) gives an equation that can be solved iteratively for the deflection field. With the zeroth order being the Born approximation, each subsequent iteration gives higher order post-Born corrections [14]. Concretely to second order,

$$\begin{aligned}\alpha^a(\hat{\boldsymbol{n}}_0) =& -2\int_0^{\chi_*} d\chi W(\chi,\chi_*)\left[\nabla^a\Psi(\chi) - 2\int_0^{\chi} d\chi' W(\chi',\chi)\nabla^a\nabla_b\Psi(\chi)\nabla^b\Psi(\chi')\right](\hat{\boldsymbol{n}}_0)\\ &+\mathcal{O}(\Psi^3).\end{aligned} \quad (4)$$

The first term, again, is the Born approximation. The second term is the lens-lens coupling that is the primary source of lensing rotation, and also gives a correction to the convergence.

Post-Born corrections have an almost negligible impact on the estimate of the deflection field. And while lensing rotation may be detectable by upcoming CMB-S4 experiments, I ignore the subdominant lensing curl in the following. We return to these effects later again, when we study the lensing curl in chapter 3 (where we also discuss its negligible contribution to B-lensing at current and near-future experiment sensitivities, an effect that I describe further below) and for convergence in chapter 4.

The general impact of CMB lensing on a CMB field is that the lensed field observed at location $\hat{\boldsymbol{n}}$ is the un-deflected field at another location $\hat{\boldsymbol{n}}' = \hat{\boldsymbol{n}} + \nabla\phi$;[2] the majority of the effect can be described by a remapping of the CMB,

$$\begin{aligned} \tilde{T}(\hat{\boldsymbol{n}}) &= T(\hat{\boldsymbol{n}} + \nabla\phi)\,, \\ \left(\tilde{Q} \pm \mathrm{i}\tilde{U}\right)(\hat{\boldsymbol{n}}) &= (Q \pm \mathrm{i}U)\,(\hat{\boldsymbol{n}} + \nabla\phi)\,. \end{aligned} \tag{5}$$

$\tilde{Q} \pm \mathrm{i}\tilde{U}$ denotes the two lensed polarization maps as measured by experiments in terms of the coordinate-dependent Stokes parameters and are spin-2 fields. We may decompose them into the spin-0 coordinate-independent gradient (E) and curl (B) part, as fundamental physics can be well described and distinguished by this composition, here on flat sky,

$$\begin{pmatrix} E_\ell \\ B_\ell \end{pmatrix} = \begin{pmatrix} \cos(2\beta_\ell) & \sin(2\beta_\ell) \\ -\sin(2\beta_\ell) & \cos(2\beta_\ell) \end{pmatrix} \begin{pmatrix} Q_\ell \\ U_\ell \end{pmatrix}\,. \tag{6}$$

$\beta_\ell = \arctan\left({}^{\ell_y}/_{\ell_x}\right)$ is the angle between two modes.

It is often helpful to define the magnification matrix $\boldsymbol{A}$, that is valid in the lensing regime for which it is everywhere invertible, and its determinant close to unity,

$$\boldsymbol{A} = \frac{\partial \boldsymbol{\alpha}}{\partial \hat{\boldsymbol{n}}} = \begin{pmatrix} 1 + \kappa + \gamma_Q & \gamma_U + \omega \\ \gamma_U - \omega & 1 + \kappa - \gamma_Q \end{pmatrix}\,, \tag{7}$$

and is parameterised by the convergence κ, the two shear components that stretch and compress the image along x and y and vice versa, γ_Q and γ_U, and the field rotation ω. The convergence and field rotation can directly be derived from the deflection field,

$$\kappa = -\frac{1}{2}\nabla\boldsymbol{\alpha}\,, \qquad\qquad \omega = -\frac{1}{2}\nabla \times \boldsymbol{\alpha}\,. \tag{8}$$

Explicitly, and going to Fourier space, (remember that curl is a second order effect in the expansion of the gravitational potential) [14],

---

[2] There are subtle effects other than remapping that we can safely ignore here: time delay, and the emission angle at the last scattering surface.

$$\kappa(\boldsymbol{L}) = L^2 \int_0^{\chi_*} d\chi W(\chi, \chi_*) \Psi(\boldsymbol{L}, \chi) \,, \tag{9}$$

$$\omega(\mathbf{L}) = -2 \int_0^{\chi_*} d\chi W(\chi, \chi_*) \int_0^{\chi} d\chi' W(\chi', \chi) \times \int \frac{d^2 l}{(2\pi)^2} (\boldsymbol{\ell} \cdot \boldsymbol{\ell}')[\boldsymbol{\ell} \times \mathbf{L}] \Psi(\boldsymbol{\ell}, \chi) \Psi(\boldsymbol{\ell}', \chi') \,. \tag{10}$$

Returning to Eq. (5): in the flat-sky approximation, ignoring the lensing curl, and assuming that deflection is a perturbation to the unlensed field, a Taylor series expansion for the temperature case gives up to second order,

$$\tilde{T}(\hat{\boldsymbol{n}}) = T(\hat{\boldsymbol{n}}) + \nabla T(\hat{\boldsymbol{n}}) \cdot \nabla \phi(\hat{\boldsymbol{n}}) + \frac{1}{2} \nabla^a \nabla^b T(\hat{\boldsymbol{n}}) \nabla_a \phi(\hat{\boldsymbol{n}}) \nabla_b \phi(\hat{\boldsymbol{n}}) + \mathcal{O}(\phi^3) \,. \tag{11}$$

I note here that this perturbative approach is only approximately correct,[3] but comes in very handy for a general understanding of the effect. One way to investigate the effect of CMB lensing is by analysing the summary statistics such as the power spectrum. We calculate the auto power spectrum of the temperature map as the 2-point correlator of the field in harmonic space, that are obtained either via a SHT on the curved sky, or Fourier transform (FT) on the flat sky. The type-2 transforms are given by,

$$\mathcal{F}[\tilde{X}(\hat{\boldsymbol{n}})] = \tilde{X}(\boldsymbol{\ell}) = \int \mathrm{d}\hat{\mathbf{n}} \, \tilde{X}(\hat{\boldsymbol{n}}) \mathrm{e}^{-i\boldsymbol{\ell} \cdot \hat{\mathbf{n}}} \,, \qquad \text{(FT)} \tag{12}$$

$$\mathcal{S}[\tilde{X}(\hat{\boldsymbol{n}})] = \tilde{X}(\boldsymbol{\ell}) = \int \mathrm{d}\hat{\mathbf{n}} \, \tilde{X}(\hat{\boldsymbol{n}}) Y_\ell^{m,\dagger}(\hat{\boldsymbol{n}}) \,. \qquad \text{(SHT)} \tag{13}$$

For brevity, in Fourier space $\boldsymbol{\ell}$ are the two Fourier modes in x and y-direction, respectively; in spherical harmonic space, $\boldsymbol{\ell}$ are the two multipole numbers ℓ and m. The temperature power spectrum is thus,

$$\langle \tilde{T}(\boldsymbol{\ell}_1) \tilde{T}^*(\boldsymbol{\ell}_2) \rangle \propto \delta(\boldsymbol{\ell}_1 - \boldsymbol{\ell}_2) \tilde{C}_\ell^{TT} \,, \tag{14}$$

that, after plugging Eq. (11) into this equation gives that it is proportional to a term with suppression factor $\gamma = 1 - \ell^2 R^\phi$, with R^ϕ defined as half the total deflection power and a term that causes smoothing of the peaks [13]. Overall, CMB lensing smooths the oscillations in the power spectrum of the temperature and polarization, and shifts the power onto smaller scales.

In polarization, CMB lensing has another very important effect: B-lensing. B-lensing describes the conversion from E-modes into B-modes, and can directly be seen by repeating Eq. (14) for polarization: assuming that there is no primordial B, we find for the lensed B-power spectrum on the flat sky [13],

$$\tilde{C}_\ell^{BB} = \int \frac{\mathrm{d}^2 \boldsymbol{\ell}'}{(2\pi)^2} \left[\boldsymbol{\ell}' \cdot (\boldsymbol{\ell} - \boldsymbol{\ell}')\right]^2 C_{|\boldsymbol{\ell}'|}^{EE} C_{|\boldsymbol{\ell} - \boldsymbol{\ell}'|}^{\phi\phi} \sin^2(2(\beta_{\boldsymbol{\ell}'} - \beta_{\boldsymbol{\ell}})) \,. \tag{15}$$

---

[3] The perturbative approach breaks down when reconstructing the actual B-modes, as B-lensing is the primary signal in the observed maps (ignoring foregrounds and noise). However, power spectrum estimates still work well [15].

Hence, even in the absence of primordial B-modes there will be lensed B-power, sourced by the deflection field and primordial E-modes.

It is exactly this signal that is the main driver of this thesis: our hunt to measure the faint primordial B-modes is hindered by the much larger B-lensing signal, a contamination of the size of about 5μK-arcmin. The primordial B-mode signal is currently confirmed to be at least roughly 30 times smaller (at the relevant scales) than B-lensing itself. By removing the variance of the B-lensing signal, we are able to measure and constraint the tensor-to-scalar ratio r, which measures the size of the primordial tensor perturbations (that directly relate to the B-modes) relative to the scalar perturbations. Sophisticated tools such as Bayesian methods and presented in this thesis are needed to reach this level of cleaning. The currently best upper bound on r comes from a combination of BICEP/Keck, baryonic acoustic oscillation measurements, and Planck: $r < 0.032$ at 95% confidence [16]. The variance of r, $\sigma^2(r)$, can be calculated via the inverse of the information matrix I,

$$\frac{1}{\sigma^2(r)} = I = \sum_{\ell=\ell_{\min}}^{\ell_{\max}} \frac{\partial C_\ell^{\text{tot}}}{\partial r} \text{Cov}_\ell^{-1} \frac{\partial C_\ell^{\text{tot}}}{\partial r} \,, \tag{16}$$

and I use only the diagonal of the covariance and the sum goes over the multipole range of interest. Assuming that the total power spectrum C_ℓ^{tot} is given by the sum of the foreground, noise, B-lensing (and parameterized by A_{lens}), and the primordial (tensor) signal itself, $C_\ell^{\text{tot}} = C_\ell^{BB,\text{fg}} + C_\ell^{BB,\text{noise}} + A_{\text{lens}} C_\ell^{BB,\text{lens}} + r C_\ell^{BB,\text{tensor}}$, and using $\text{Cov}_\ell = \Delta \text{C}_\ell^{\text{tot}} = \frac{2(\text{C}_\ell^{\text{tot}})^2}{(2\ell+1)\text{f}_{\text{sky}}}$, with f_{sky} the observed sky fraction, I find,

$$\frac{1}{\sigma^2(r)} = \sum_{\ell=\ell_{\min}}^{\ell_{\max}} \frac{(2\ell+1) f_{\text{sky}}}{2} \left(\frac{C_\ell^{BB,\text{tensor}}}{C_\ell^{BB,\text{fg}} + C_\ell^{BB,\text{noise}} + A_{\text{lens}} C_\ell^{BB,\text{lens}} + r C_\ell^{BB,\text{tensor}}} \right)^2 . \tag{17}$$

We return to lensing in more detail in chapter 2 where I discuss highly accurate lensing operators, and when I present the embedding of lensing into our software `Delensalot` in chapter 1.1.4.

0.2 Quadratic Estimator

CMB lensing is a second order anisotropy in the CMB. The standard way of extracting the effect of CMB lensing is by quadratic estimators (QE). A QE can in principle be built for any source that causes (anisotropic) correlations. If the source is the lensing potential, we make use of the induced correlation between different scales $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$, hereby allowing to distinguish between the primordial anisotropies that are assumed uncorrelated between different scales, and the effect of the lensing potential. Correlating squared combinations of the temperature and polarization maps, first proposed by [17], allows to uncover the imprint of this second order anisotropy by measuring the correlations. The key is to determine the response $f^{XY}(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2)$ of

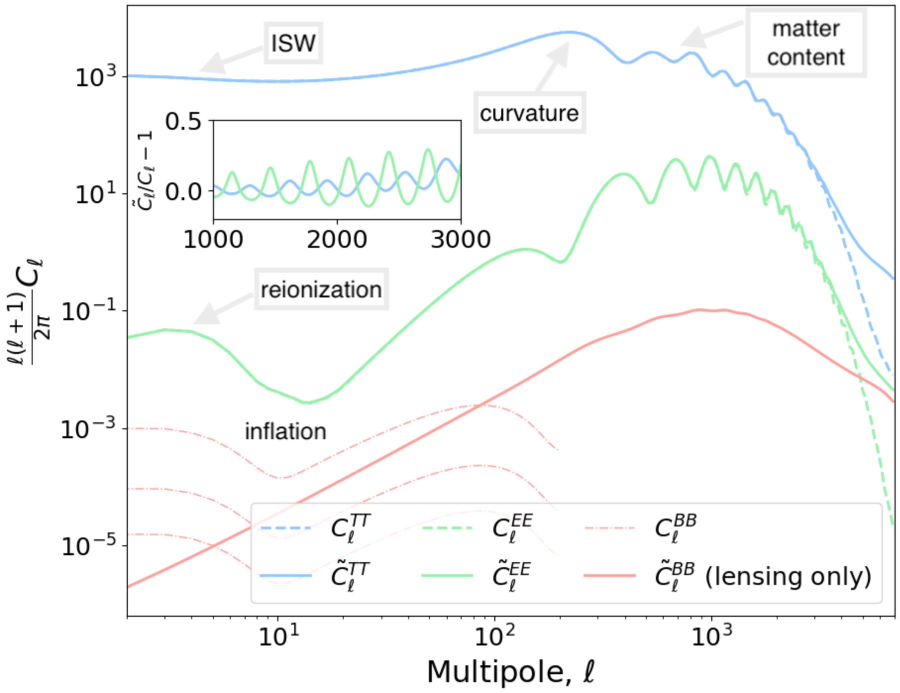

Figure 2: The CMB temperature and polarization power spectra and how key features relate to the underlying primordial perturbations and the ΛCDM model.

the quadratic combination of the fields X and Y to the source of the anisotropy.[4]

While this thesis focuses on the full sky, the flat sky equivalents are often easier to follow and lighter in notation which is why flat sky equations are mainly discussed in the following. Occasionally, I provide full sky equations. Lensing reconstruction is done on observed fields that contains the map-level experiment noise n and transfer function $\mathcal{B}$, $\tilde{X}^{\mathrm{obs}} = \mathcal{B}\tilde{X} + n$ and we assume the noise to be uncorrelated with the fields.

The perturbative, curved sky response function is given by [18],

$$\left\langle \tilde{X}^{\mathrm{obs}}(\boldsymbol{\ell}_1)\tilde{Y}^{\mathrm{obs}}(\boldsymbol{\ell}_2)\right\rangle_{\mathrm{CMB}} = \\ C^{XY}_{\boldsymbol{\ell}_1}\delta_{\ell_1\ell_2}\delta_{m_1-m_2}(-1)^{m_1} + \sum_{LM}(-1)^M \begin{pmatrix} \ell_1 & \ell_2 & L \\ m_1 & m_2 & -M \end{pmatrix} f^{XY}_{\ell_1 L \ell_2}\phi_{LM}\,, \tag{18}$$

and I have introduced the average over a fixed deflection field that only averages over the CMB, $\langle\rangle_{\mathrm{CMB}}$.[5] The 2×3 structure in the brackets is the Wigner-3j symbol. On the flat sky this becomes much easier to track,

$$\langle \tilde{X}^{\mathrm{obs}}(\boldsymbol{\ell}_1)\tilde{Y}^{\mathrm{obs}}(\boldsymbol{\ell}_2)\rangle_{\mathrm{CMB}} = \delta(\boldsymbol{L}-\boldsymbol{\ell}_1-\boldsymbol{\ell}_2)f^{XY}(\boldsymbol{\ell}_1,\boldsymbol{\ell}_2)\phi(\boldsymbol{L})\,, \qquad \boldsymbol{L} = \boldsymbol{\ell}_1+\boldsymbol{\ell}_2\,. \tag{19}$$

---

[4] To follow convention of the literature, functions with fields as superscripts are without diacritical mark: however, this does not mean that they are a function of the unlensed field

[5] An ensemble average, that is the average over both the CMB and deflection field would give zero, as the deflection field is statistically zero.

In the following, I drop the superscript obs for brevity. Plugging the equations Eq. (5) into Eq. (19) directly gives the desired response function, and are direct functions of the power spectra of the CMB fields. The more accurate, non-perturbative response function can be obtained by tracking the impact of ϕ onto the lensed fields non-perturbatively,

$$\left\langle \frac{\delta}{\delta\phi(\boldsymbol{L})}\tilde{X}(\boldsymbol{\ell}_1)\tilde{Y}(\boldsymbol{\ell}_2)\right\rangle = \delta(\boldsymbol{L}-\boldsymbol{\ell}_1-\boldsymbol{\ell}_2)f^{XY,\mathrm{np}}(\boldsymbol{\ell}_1,\boldsymbol{\ell}_2)\,, \tag{20}$$

and gives a solution to further reduce the variance of the estimator: $f^{XY,\mathrm{np}}$, for most field combinations, is a function of the power spectrum of the gradient of the field, $C_\ell^{X\nabla Y}$ [19].

A quadratic estimate of the lensing potential on the full sky is [18],

$$\hat{\phi}^{XY}_{LM} = \sum_{\ell_1 m_1 \ell_2 m_2}\left[\frac{A^{XY}_L}{L(L+1)}(-1)^M\begin{pmatrix}\ell_1 & \ell_2 & L\\ m_1 & m_2 & -M\end{pmatrix}F^{XY}_{\ell_1 L \ell_2}\right]\tilde{X}_{\ell_1 m_1}\tilde{Y}_{\ell_2 m_2}\,. \tag{21}$$

A^{XY}_L is the normalization, and $F^{XY}_{\ell_1 L \ell_2}$ are the appropriate quadratic weights. More intuitively on the flat sky this is,

$$\hat{\phi}^{XY}(\boldsymbol{L}) = \frac{A^{XY}_L}{L^2}\int_{\boldsymbol{\ell}_1+\boldsymbol{\ell}_2=\boldsymbol{L}}\tilde{X}(\boldsymbol{\ell}_1)\tilde{Y}(\boldsymbol{\ell}_2)F^{XY}(\boldsymbol{\ell}_1,\boldsymbol{\ell}_2)\,, \qquad \boldsymbol{L}=\boldsymbol{\ell}_1+\boldsymbol{\ell}_2\,, \tag{22}$$

and I have introduced the notation $\int_{\boldsymbol{\ell}_1+\boldsymbol{\ell}_2=\boldsymbol{L}} = \int \frac{d\boldsymbol{\ell}_1}{(2\pi)^2}\frac{d\boldsymbol{\ell}_2}{(2\pi)^2}(2\pi)\delta(\boldsymbol{L}-\boldsymbol{\ell}_1-\boldsymbol{\ell}_2)$ for brevity. A^{XY}_L is the corresponding flat-sky normalization but using the same symbol for brevity. We find the normalization by requiring this estimator to be unbiased,

$$\langle\hat{\phi}(\boldsymbol{L})\rangle_{\mathrm{CMB}} = \phi(\boldsymbol{L})\,, \tag{23}$$

thus,

$$A^{XY}_L = L^2\left(\int_{\boldsymbol{\ell}_1+\boldsymbol{\ell}_2=\boldsymbol{L}} f^{XY}(\boldsymbol{\ell}_1,\boldsymbol{\ell}_2)F^{XY}(\boldsymbol{\ell}_1,\boldsymbol{\ell}_2)\right)^{-1}\,. \tag{24}$$

Minimizing the variance of the power spectrum of this estimator, $\langle\hat{\phi}\hat{\phi}\rangle - \langle\hat{\phi}\rangle\langle\hat{\phi}\rangle$, and rearranging gives the handy interpretation of using inverse variance filtered maps (denoted by the bar), and at the same time simplifies as $F=f$,

$$\hat{\phi}^{XY}(\boldsymbol{L}) = \frac{A^{XY}_L}{L^2}\int_{\boldsymbol{\ell}_1}\bar{\tilde{X}}(\boldsymbol{\ell}_1)\bar{\tilde{Y}}(\boldsymbol{\ell}_2)f^{XY}(\boldsymbol{\ell}_1,\boldsymbol{\ell}_2)\,. \tag{25}$$

The inverse variance filtering is,

$$\bar{\tilde{X}} = \mathcal{B}\mathrm{Cov}_\alpha^{-1}\tilde{X}\,. \tag{26}$$

In the case of isotropic sky, the inverse variance filtering is achieved by simply dividing the maps by the corresponding power spectra [18].

The response functions are functions of the power spectra of the CMB fields, and Eq. (25) can also be reformulated to give the interpretation of, instead, inverse variance filtering one of the maps, and Wiener-filtering the other. A Wiener-filtered (WF) map $X^{\mathrm{WF}} = \tilde{C}_\ell^{XX}\left(\tilde{C}_\ell^{XX} + C_\ell^{XX,N}\right)^{-1}\tilde{X}$, with $C_\ell^{XX,N}$ the noise power spectrum, is our best estimate of the primordial, unlensed field.[6] With this at hand, we can build a B-lensing template, that is our best estimate of the B-lensing map. This can be achieved by calculating the B-lensing map with the Wiener-filtered E-field and the estimate on the deflection field, in principle following Eq. (11) but for polarization. Delensing then requires the subtraction of this template from the actual B map, which, when sufficiently correlated with the underlying lensing signal, will reduce the variance, and thus undo the B-lensing. It is worth noting that the QE delensing performs better by using a perturbative version of the B-lensing template [15].

It is instructive to discuss the variance of the deflection field estimator as we can derive important properties that we review in the following. The isotropic noise of the estimator is,

$$\langle(\hat{\phi}-\phi)(\boldsymbol{L}_1)(\hat{\phi}-\phi)^*(\boldsymbol{L}_2)\rangle \propto N(\boldsymbol{L}_1)\delta(\boldsymbol{L}_1-\boldsymbol{L}_2)\,, \tag{27}$$

and the average goes over both the CMB and deflection fields. The complete variance is the ensemble average,

$$\langle\langle\hat{\phi}\hat{\phi}\rangle_{\mathrm{CMB}} - \langle\hat{\phi}\rangle_{\mathrm{CMB}}\langle\hat{\phi}\rangle_{\mathrm{CMB}}\rangle_{\mathrm{LSS}}\,. \tag{28}$$

In the following, we assume a Gaussian lensing potential, unlensed fields, and detector noise. Later in this section, we discuss a bias that arises from the non-Gaussianity of the lensing potential itself.

The variance on $\hat{\phi}$ is important whenever we study the power spectrum. The power spectrum $C_L^{\hat{\phi}\hat{\phi}}$ is a 4-point function in the map and picks up contributions beyond the signal because it is itself a measure of the variance. We need to distinguish between the actual signal, and other contributions, in the following referred to as reconstruction bias.

For starter I am only interested in the leading order contribution to the reconstruction bias, but will discuss higher order contributions shortly after.

To calculate the variance, we can decompose it into all possible contractions. Assuming Gaussianity, contractions of an odd number of fields are trivially zero. The remaining contractions can be split into disconnected correlators that are 2-pairings of contractions, and the fully connected, 4-pairing contraction,

$$\langle\tilde{X}\tilde{X}\tilde{X}\tilde{X}\rangle = \underbrace{\langle\tilde{X}\tilde{X}\tilde{X}\tilde{X}\rangle_{\mathrm{C}}}_{\text{signal+bias}} + \underbrace{\langle\tilde{X}\tilde{X}\tilde{X}\tilde{X}\rangle_{\mathrm{DC}}}_{\text{bias}}\,. \tag{29}$$

For brevity I have omitted the dependence of the field to the multipole $\boldsymbol{\ell}_i$ and used only one character for the 4 different fields. From context I assume that it is clear to the reader that $\langle\tilde{X}\tilde{X}\tilde{X}\tilde{X}\rangle \equiv \langle\tilde{X}(\boldsymbol{\ell}_1)\tilde{X}'(\boldsymbol{\ell}_2)\tilde{X}''(\boldsymbol{\ell}_3)\tilde{X}'''(\boldsymbol{\ell}_4)\rangle$.

---

[6] For anisotropic sky or noise, the implicit change to Fourier space, as was done above, is not as trivial, resulting in the Wiener-filtering to become more involved than just dividing by the power spectra, but can be solved with iterative and approximate methods like conjugate gradient (CG) [20].

The disconnected contractions make up the primary bias, denoted $N_L^{(0)}$, and relates to the 2-pairings through Wick's theorem,[7]

$$\langle \tilde{X}\tilde{X}\tilde{X}\tilde{X} \rangle_{\mathrm{DC}} = \sum_P \langle \tilde{X}\tilde{X} \rangle \langle \tilde{X}\tilde{X} \rangle \,, \tag{30}$$

where P refers to all possible pairings (here, there are 3).

The signal of interest is in the connected correlator that gives a trispectrum, $C_L^{\hat{\phi}^{\mathrm{QE}}\hat{\phi}^{\mathrm{QE}}} \propto \langle \tilde{X}\tilde{X}\tilde{X}\tilde{X} \rangle_{\mathrm{C}} = \mathcal{T}(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \boldsymbol{\ell}_3, \boldsymbol{\ell}_4)$.

The complete correlator is proportional to the signal and all biases,

$$\langle \hat{\phi}\hat{\phi} \rangle \propto C_L^{\phi\phi} + N_L^{(0)} + N_L^{(1)} + N_L^{(2)} + \dots \,. \tag{31}$$

and I have grouped the noise contributions in terms of their dependence on $C_L^{\phi\phi}$: $N_L^{(1)}$ has a linear dependence on it. The ellipsis denotes higher order variance contributions. All contributions higher than zeroth order arise from higher-order contractions inside the connected graph.

It is an important result that $N_L^{(0)}$ is present even in the absence of lensing. Concretely, and making use of Wick's theorem, I find,

$$\begin{aligned} &N^{(0),XY}(L) \\ &= \frac{(A_L^{\mathrm{XY}})^2}{L^2} \int_{\boldsymbol{\ell}_1+\boldsymbol{\ell}_2=\boldsymbol{L}} F^{XY}(\boldsymbol{\ell}_1,\boldsymbol{\ell}_2) \left(F^{XY}(\boldsymbol{\ell}_1,\boldsymbol{\ell}_2) C^{XX}_{\boldsymbol{\ell}_1} C^{YY}_{\boldsymbol{\ell}_2} + F^{XY}(\boldsymbol{\ell}_2,\boldsymbol{\ell}_1) C^{XY}_{\boldsymbol{\ell}_1} C^{XY}_{\boldsymbol{\ell}_2} \right) \,, \end{aligned} \tag{32}$$

which directly falls out of the calculation for the disconnected graph when we calculate the variance of the estimator: there are 3 possible contractions, with one of them cancelling with the second term of the equation of the variance. The superscript XY denotes any quadratic combination of the two fields $(XY) \in \{T, E, B\}^2$. It is independent of the lensing potential and only depends implicitly on it due to the total power spectrum.

We now review the origin of higher order contractions in the trispectrum, with a rigorous derivation of the individual biases provided in [21]. The quantum field theory analog to higher-order corrections can also be developed in terms of Feynman diagrams [22].

Starting with the variance, only averaging over CMB realizations, we find (assuming a Gaussian deflection field thus, $\boldsymbol{L} = \boldsymbol{L}'$),

$$\begin{aligned} &\langle \hat{\phi}^{*XY}(\boldsymbol{L}) \hat{\phi}^{X'Y'}(\boldsymbol{L}') \rangle_{\mathrm{CMB}} - \langle \hat{\phi}^{*XY}(\boldsymbol{L}) \rangle_{\mathrm{CMB}} \langle \hat{\phi}^{X'Y'}(\boldsymbol{L}') \rangle_{\mathrm{CMB}} \\ &\propto \delta(\boldsymbol{L}-\boldsymbol{L}') \boldsymbol{L}\cdot\boldsymbol{L}' \frac{A_L^{XY} A_{L'}^{X'Y'}}{L^2 L'^2} \int_{\boldsymbol{\ell}_1+\boldsymbol{\ell}_2=\boldsymbol{\ell}'_1+\boldsymbol{\ell}'_2} \langle \tilde{X}\tilde{Y}\tilde{X}'\tilde{Y}' \rangle_{\mathrm{CMB}} F^{XY}(\boldsymbol{\ell}_1,\boldsymbol{\ell}_2) F^{X'Y'}(\boldsymbol{\ell}'_1,\boldsymbol{\ell}'_2) \\ &- \phi^*(\boldsymbol{L})\phi(\boldsymbol{L}') \,, \end{aligned} \tag{33}$$

---

[7] We will see shortly that the superscript (0) refers to the dependence of powers in $C_L^{\phi\phi}$.

with $\boldsymbol{L} = \boldsymbol{\ell}_1 + \boldsymbol{\ell}_2$, $\boldsymbol{L}' = \boldsymbol{\ell}_1' + \boldsymbol{\ell}_2'$, and the integral goes over all modes with the constraint $\boldsymbol{L} = \boldsymbol{L}'$. The 4-point correlator inside the integral is quite involved, see Eq. (18) of [21]. To obtain the terms that contribute to $N^{(1)}$, the 4-point function can be evaluated, similar to calculating the response function, using Eq. (5) up to second order.

Averaging Eq. (33) over realizations of the deflection field,

$$\langle\langle\hat{\phi}^*\hat{\phi}\rangle_{\mathrm{CMB}} - \langle\hat{\phi}^*\rangle_{\mathrm{CMB}}\langle\hat{\phi}\rangle_{\mathrm{CMB}}\rangle_{\mathrm{LSS}}\,, \tag{34}$$

gives the final result. For temperature only, the 4-point correlator inside the integral (again, assuming Gaussianity of the deflection field, thus $\boldsymbol{\ell}_1 + \boldsymbol{\ell}_2 = \boldsymbol{\ell}_1' + \boldsymbol{\ell}_2'$, and that $\boldsymbol{L} = 0$ is unobservable) is defined as,

$$\begin{split}&\langle\langle\tilde{T}\tilde{T}\tilde{T}\tilde{T}\rangle_{\mathrm{CMB}}\rangle_{\mathrm{LSS}} \propto \\ &\quad C^{TT}_{\ell_1}C^{TT}_{\ell_2}\left(\delta(\boldsymbol{\ell}_1' - \boldsymbol{\ell}_1)\delta(\boldsymbol{\ell}_2' - \boldsymbol{\ell}_2) + \delta(\boldsymbol{\ell}_2' - \boldsymbol{\ell}_1)\delta(\boldsymbol{\ell}_1' - \boldsymbol{\ell}_2)\right) + \mathcal{T}(-\boldsymbol{\ell}_1, -\boldsymbol{\ell}_2, \boldsymbol{\ell}_1', \boldsymbol{\ell}_2')\,,\end{split} \tag{35}$$

Plugging in the term with the δ-function into Eq. (34), again, gives the leading contribution of the reconstruction bias, $N^{(0)}$. These δ-functions, in essence act as a splitting of the 4-point function into 2-point functions, and are thus equivalent of Wick's theorem.

The trispectrum can be written as [21],

$$\begin{split}\mathcal{T}(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \boldsymbol{\ell}_3, \boldsymbol{\ell}_4) =&C^{\phi\phi}_{|\boldsymbol{\ell}_1+\boldsymbol{\ell}_2|} f^{TT}(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2) f^{TT}(\boldsymbol{\ell}_3, \boldsymbol{\ell}_4) + \\ &C^{\phi\phi}_{|\boldsymbol{\ell}_1+\boldsymbol{\ell}_3|} f^{TT}(\boldsymbol{\ell}_1, \boldsymbol{\ell}_3) f^{TT}(\boldsymbol{\ell}_2, \boldsymbol{\ell}_4) + \\ &C^{\phi\phi}_{|\boldsymbol{\ell}_1+\boldsymbol{\ell}_4|} f^{TT}(\boldsymbol{\ell}_1, \boldsymbol{\ell}_4) f^{TT}(\boldsymbol{\ell}_2, \boldsymbol{\ell}_3)\,,\end{split} \tag{36}$$

which helps in identifying the individual contributions, and from which we can directly see that it is a function of the lensing potential power spectrum (which should somehow be obvious due to it being a trispectrum). N-th order reconstruction biases come from N-th order in $C^{\phi\phi}_{\ell}$ term of the trispectrum after subtracting the actual signal $\langle\langle\hat{\phi}^*(\boldsymbol{L})\rangle_{\mathrm{CMB}}\langle\hat{\phi}(\boldsymbol{L}')\rangle_{\mathrm{CMB}}\rangle_{\mathrm{LSS}}$ (the second term in Eq. (34)).[8] Explicitly for temperature [21],

$$\begin{split}&N^{(1),TTTT}(L) = \\ &\delta(\boldsymbol{L} - \boldsymbol{L}')\boldsymbol{L}\cdot\boldsymbol{L}'\frac{A^{TT}_{L}A^{TT}_{L'}}{L^2L'^2}\int_{\boldsymbol{\ell}_1+\boldsymbol{\ell}_2=\boldsymbol{\ell}_1'+\boldsymbol{\ell}_2'} F^{TT}(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2)F^{TT}(\boldsymbol{\ell}_1', \boldsymbol{\ell}_2') \\ &\times\left\{C^{\phi\phi}_{|\boldsymbol{\ell}_1-\boldsymbol{\ell}_1'|} f^{TT}(-\boldsymbol{\ell}_1, \boldsymbol{\ell}_1') f^{TT}(-\boldsymbol{\ell}_2, \boldsymbol{\ell}_2') + C^{\phi\phi}_{|\boldsymbol{\ell}_1-\boldsymbol{\ell}_2'|} f^{TT}(-\boldsymbol{\ell}_1, \boldsymbol{\ell}_2') f^{TT}(-\boldsymbol{\ell}_2, \boldsymbol{\ell}_1')\right\}.\end{split} \tag{37}$$

In some sense, $N^{(1)}$ can be considered a signal due to its direct proportion to $C^{\phi\phi}_{\ell}$. But it is important to note that $N^{(1)}$ arises from an integration across different combination of scales than $C^{\phi\phi}_{L}$ itself, and may rather be understood as an interference.

---

[8] Note that there are no higher order than $N^{(1)}$ in this trispectrum shown here due to our lensing perturbation that we terminated after the second order in ϕ.

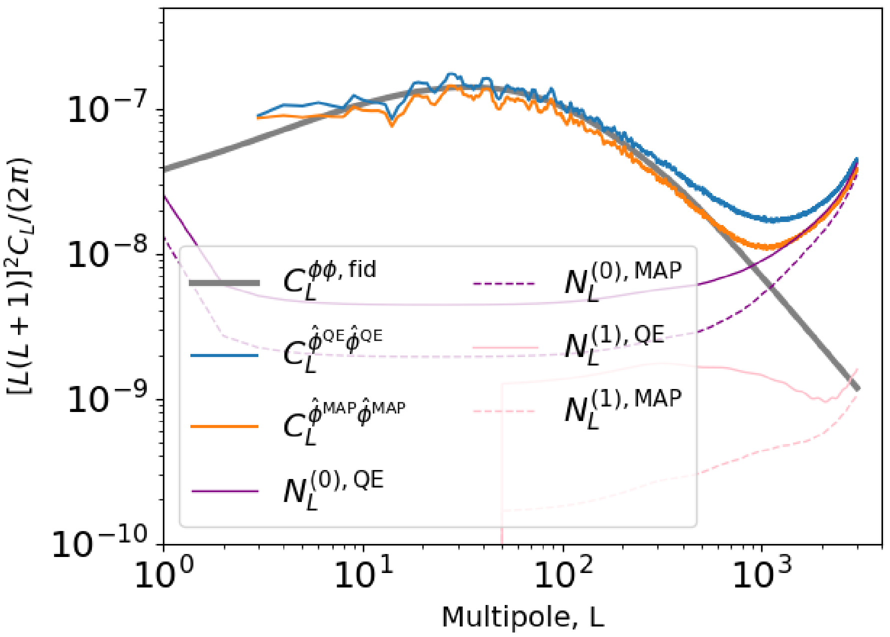

Figure 3: CMB lensing reconstruction and associated reconstruction biases $N^{(0)}$ and $N^{(1)}$ for a CMB-S4 like configuration. See text for discussion.

As a side note, the same scheme can be used to obtain the $N^{(2)}$ biases, for which consequently Eq. (5) would have to be taken up to fourth order, which generates the needed higher order terms in the 4-point correlator Eq. (35).

Typical reconstruction biases are shown in Fig. (3) in purple ($N^{(0)}$) and pink ($N^{(1)}$) for a CMB-S4 like configuration. QE is shown as solid lines. The figure also shows results on the optimal approach (dashed lines), and reconstruction that I discuss in the next section. These lines come from the forecasting method developed in [23]. Example reconstruction estimates are shown in blue and orange, binned over a few multipoles.

In the case of a non-Gaussian lensing potential, additional biases arise. The leading contribution is dubbed $N_L^{(3/2)}$. This is because the leading contribution comes from the 3-point function of the lensing field that is now non-zero (as opposed to the Gaussian case, in which correlators across odd number of fields become zero). A thorough derivation is given in [24]. Using Edgeworth series, we find that the primary contribution comes from the 3-point correlator and can be interpreted as the response of the variance to the 3-point correlator [25]. The first term, the Gaussian part, can be calculated in the usual manner and gives the result discussed above. This three-point correlator gives rise to a bispectrum and is non-zero because the following now non-vanishing correlators need to be considered for the calculation of the variance,

$$N_L^{(3/2)} \propto \langle \delta\tilde{X}\delta\tilde{X}\delta\tilde{X}\tilde{X}\rangle + \langle \delta^2\tilde{X}\delta\tilde{X}\tilde{X}\tilde{X}\rangle + \langle \delta^3\tilde{X}\tilde{X}\tilde{X}\tilde{X}\rangle\,, \tag{38}$$

and I have introduced the n-th order perturbation on the field $\tilde{X}$ through lensing as $\tilde{X} = X + \sum_n \delta^n X$. Analytic results are presented in [26].

Another effect that adds a bispectrum is the post-Born evaluation itself, Eq. (4), as $\boldsymbol{\alpha}$ is not Gaussian. The resulting bias has similar shape and signal to the one discussed above, but quantitatively opposite sign [25]. Both effects therefore somewhat cancel each other, which makes this an important source of $N_L^{(3/2)}$, and hence needs to be accounted for. We discuss both in the context of iterative lensing reconstruction in chapter 4 and show that they can be treated in similar ways.

QEs have a long and successful history and are a crucial ingredient of today's research. A rigorous prescription on the flat sky was presented in [27], and later extended to the full sky [18]. [28] showed that using lensed spectra in the filter effectively leads to cancellation between the $N^{(1)}$ and $N^{(2)}$ bias on large scales, and [19] derived the non-perturbative response function, which is more correct due to the non-perturbative nature of lensing, further reducing the variance.[9]

With real data come real foregrounds. In this context, impacts from second order anisotropies onto the reconstructions and possible methods for mitigation have been studied [29, 30], and bias-hardened approaches [31] to reduce the impact of foregrounds that correlate with the lensing field.

QE methods, and CMB lensing cross galaxy surveys have successfully been applied to data by many collaborations and experiments [32, 33, 34, 35, 36, 37] where they contribute significantly to constraining cosmological parameters and help to break degeneracies [38], and are also a key ingredient in measuring the sum of the neutrino mass [39], and constraint dark energy [40].

Recently, an improved, *generalized* gMV has been proposed [41] that further reduces the variance of the combined, minimum variance estimator.

Due to their comparably straightforward implementation, and analytic tractability for many problems, QEs will likely remain an important tool for decades to come. QE is the perfectly suited entry level for studying CMB lensing effects and cover much of the needed properties for today's research.

Despite its practicality, the QE does have limitations. Due to the squared combination of the maps, a QE probes correlations that are at most two-point in the map. Lensing, however, also induces higher order correlations such as the bispectrum, trispectrum, and higher orders. Probing them is crucial for a complete disentanglement of the anisotropies.

QE analysis is therefore not optimal. The next section discusses an optimal estimator via a Bayesian approach.

0.3 Maximum a posteriori Estimator

The first optimal estimator for extracting the deflection field was proposed in the pioneering paper [42], and is based on a maximum likelihood estimator (MLE). Our approach [43] follows their idea and iteratively reconstructs the deflection field by calculating the gradient of the likelihood. This section aims to introduce the MLE

---

[9] non-perturbative response functions are particular helpful in temperature and for low noise experiments.

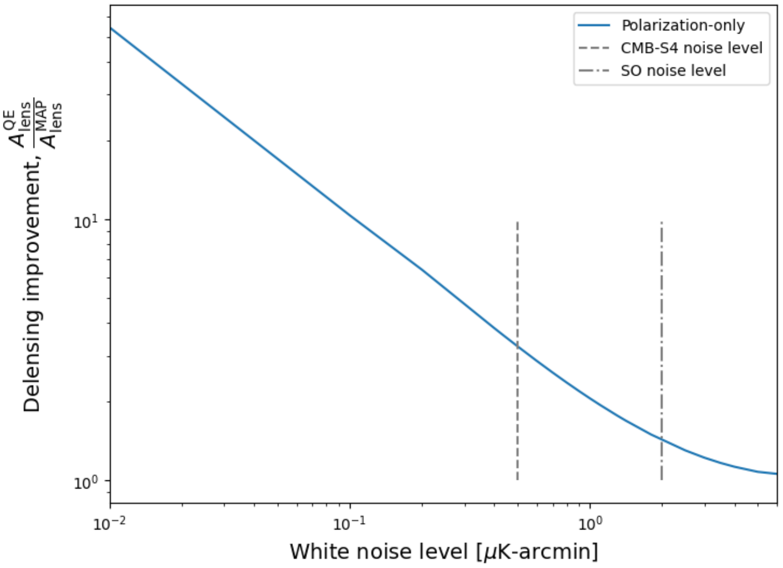

Figure 4: Improvement forecast of MAP compared to QE for the residual lensing amplitude A_{lens} in the perturbative, large-scale limit of the B-modes [13]. This forecast assumes full sky reconstruction and an intermediate beam-width for all noise levels. The sensitivity, the noise level shown in the x-axis shows a highly optimistic range; for CMB-S4 an about 4-times improvement compared to QE will likely be achieved. This forecast assumes that we have foregrounds under control on all scales.

approach and how it is applied in the context of CMB lensing reconstruction, and lays out arguments why it works and why it is better than QE.

The quadratic estimator of the previous section would be (close to) optimal if lensing was a perturbative effect. However, CMB lensing does induce higher-order correlations, therefore only tracing the 2-point function does not recover all the information. In the following, I switch to the estimation of α rather than ϕ which has no real impact other than the notation. An optimal estimator that takes into account higher-order correlations such as $\langle\tilde{X}\tilde{X}\tilde{X}\rangle$, $\langle\tilde{X}\tilde{X}\tilde{X}\tilde{X}\rangle$, up to N fields, in principle would be,

$$\hat{\alpha}(\boldsymbol{L}) = \sum_{n}^{N} \frac{A_L^{(n)}}{L^n} \int_{\partial\boldsymbol{\ell}^n} \left(\prod_{i}^{n} \tilde{X}^i(\boldsymbol{\ell}_i)\right) F^{(n)}(\{\boldsymbol{\ell}\}_n)\,, \tag{39}$$

with $F(\{\boldsymbol{\ell}\}_n)$ being the weighting function with the set over all n modes, $A_L^{(n)}$ the normalization for the n-th correlator, and $\tilde{X}^i$ the i-th field, and the integral goes over the respective surface constraint by the conditions $\partial\boldsymbol{\ell}^n : \boldsymbol{L} = \sum_n \boldsymbol{\ell}_n$

A MLE finds the maximum value of the likelihood. This point is optimal due to the likelihood principle, stating that all information is contained in the likelihood function, and is optimal as the MLE achieves the lowest possible variance.[10] In fact,

---

[10] There are various measures that could be considered optimal, examples being the mean of the a posteriori or the maximum of it. This thesis solely uses the maximum.

what we eventually use is the maximum a posteriori (MAP) estimator. The main difference is that MAP estimators also provide means to include a prior. I start by defining the negative log likelihood using Gaussian fields and Gaussian experiment noise,

$$-2\ln(\mathcal{L}(\tilde{X}|\boldsymbol{\alpha})) = \ln(\det(\mathrm{cov}_{\boldsymbol{\alpha}})) + \tilde{X}^{\dagger}\mathrm{cov}_{\boldsymbol{\alpha}}^{-1}\tilde{X}\,. \tag{40}$$

$\ln(\mathcal{L}(\cdot))$ is Gaussian in $\tilde{X}$, and $\mathrm{cov}_{\boldsymbol{\alpha}}$ is the Gaussian covariance, and $\boldsymbol{\alpha}$ is the deflection field with respect to both Cartesian directions.

$$\mathrm{cov}_{\boldsymbol{\alpha}} = \langle \tilde{X}\tilde{X}^{\dagger}\rangle = \mathcal{B}\mathcal{D}_{\boldsymbol{\alpha}}\mathcal{Y}C^{XX}\mathcal{Y}^{\dagger}\mathcal{D}_{\boldsymbol{\alpha}}^{\dagger}\mathcal{B}^{\dagger} + N\,, \tag{41}$$

$\mathcal{Y}$ is a short-hand matrix notation for a SHT, and I have introduced the data model $\tilde{X}(\boldsymbol{n}) = \mathcal{B}\mathcal{D}_{\boldsymbol{\alpha}}X + n$, with $\mathcal{B}$ again the transfer function, and $\mathcal{D}_{\boldsymbol{\alpha}}$ the deflection operator. With this, the likelihood becomes directly dependent on $\boldsymbol{\alpha}$. Applying Bayes' theorem, we obtain the log posterior up to an irrelevant constant,

$$\mathcal{P}(\boldsymbol{\alpha}|\tilde{X}) \equiv \ln\left(p(\boldsymbol{\alpha}|\tilde{X})\right) = \ln\left(\mathcal{L}(\tilde{X}|\boldsymbol{\alpha})P(\boldsymbol{\alpha})\right)\,, \tag{42}$$

with P the prior on $\boldsymbol{\alpha}$. The posterior is a scalar that depends on each and every mode that is reconstructed and is thus a high-dimensional problem. We find the MAP point by calculating the gradient $g = \frac{\partial L}{\partial \boldsymbol{\alpha}}$, and setting the condition $g = 0$. This gradient has the dimension of the number of modes that are reconstructed.

If the posterior $\mathcal{P}(\boldsymbol{\alpha}|\tilde{X})$ were Gaussian in $\boldsymbol{\alpha}$,

$$\mathcal{P}(\boldsymbol{\alpha}|\tilde{X}) = \frac{1}{2}(\boldsymbol{\alpha} - \boldsymbol{\alpha}^{\mu})H^{-1}(\boldsymbol{\alpha} - \boldsymbol{\alpha}^{\mu})\,, \tag{43}$$

with $\boldsymbol{\alpha}^{\mu}$ the mean, and H being the curvature matrix of the log likelihood, $[H^{-1}]^{ab}_{\boldsymbol{L}\boldsymbol{L}'} = \frac{\partial^2\ln(\mathcal{L}(\tilde{X}|\alpha))}{\partial\alpha^a(\boldsymbol{L})\partial\alpha^{*b}(\boldsymbol{L}')}$, and ab the indices of the two coordinates, one step would be enough to instantly reach the MAP point,

$$\hat{\boldsymbol{\alpha}}^{\mathrm{MAP}} = H^{-1}g\,. \tag{44}$$

However, the posterior is not Gaussian in $\boldsymbol{\alpha}$: while not directly visible, Eq. (40) depends on $\boldsymbol{\alpha}$ in a non-trivial way through the covariance that may or may not give a quadratic dependence on it. Therefore, an iterative scheme must be applied. The first to propose this were [42]. [20] present a flat sky implementation using the above likelihood, and is also the starting point for the curved sky implementation presented in chapter 3.

The iterative scheme works as follows. Calculating the gradient of the likelihood with an initial guess of $\boldsymbol{\alpha}$, we calculate $\mathcal{D}_{\boldsymbol{\alpha}}X$.[11] The gradient of the likelihood contains a Wiener-filtering that is a function of $\mathcal{D}_{\boldsymbol{\alpha}}$ and returns an estimate of the unlensed fields. With the gradient obtained, we get a new $\boldsymbol{\alpha}$ using Eq. (44). This new $\boldsymbol{\alpha}$ can be used to update the likelihood model with a new $\mathcal{D}_{\boldsymbol{\alpha}}X$, that in return gives an even better estimate on the unlensed fields through Wiener-filtering.

---

[11] The calculation of the lensed field merely adds computational complexity, as this can be solved at machine precision accuracy and can therefore be considered exact.

Repeating this until convergence gives the desired MAP point. The reason that every iteration becomes better is that we obtain better and better unlensed fields. The MAP points' convergence is only limited by the instrumental noise for polarization estimators, due to the data model depending on the unlensed spectra: there is no primordial EB correlation due to parity, and primordial B is confirmed to be very small (perhaps even zero).[12] By iteratively following the gradient, we can therefore follow the non-Gaussian features of the likelihood (such as skewness and kurtosis) to eventually reach the MAP point.

The gradient can naturally be split into three pieces: the quadratic term, the prior term, and the "mean-field" that comes from the determinant. Only looking at the "quadratic" part of the gradient, we see a clear similarity between the QE and iterative estimate in position space. Here, using spin-weight notation (see [18, 44] for QE),

$$\hat{\boldsymbol{\alpha}}^{\mathrm{QE}} = \sum_{s=0,\pm 2} \left[{}_s\bar{X}^{\mathrm{res,mod}} \partial \, {}_sY^{\mathrm{WF}} \right] , \tag{45}$$

$$\hat{\boldsymbol{\alpha}}^{\mathrm{MAP},i+1} = \hat{\boldsymbol{\alpha}}^{\mathrm{MAP},i} + H^{-1} \sum_{s=0,\pm 2} \left[{}_s\bar{X}^{\mathrm{res}} \mathcal{D}_{\boldsymbol{\alpha}} \partial \, {}_sY^{\mathrm{WF}} \right] , \tag{46}$$

with $\bar{X}^{\mathrm{res,mod}}$ and $\bar{X}^{\mathrm{res}}$ being the inverse noise weighted "residual" for the QE and MAP data model, respectively [20],

$$\bar{X}^{\mathrm{res,mod}} = \mathcal{B}^{\dagger} N^{-1} (X - \mathcal{B} X^{\mathrm{WF}}) , \tag{47}$$

$$\bar{X}^{\mathrm{res}} = \mathcal{B}^{\dagger} N^{-1} (X - \mathcal{B} \mathcal{D}_{\boldsymbol{\alpha}} X^{\mathrm{WF}}) , \tag{48}$$

and are a direct result of applying some matrix operations onto Eq. (26). As the residual approaches zero, $\hat{\alpha}^{\mathrm{MAP},i+1}$ will more and more closely match $\hat{\alpha}^{\mathrm{MAP},i}$ as the Wiener-filtering does not improve further. If we were to only use the quadratic part (and assuming that the mean-field term can be ignored), this would result in an illicit behaviour: the likelihood of the noisy modes, the smallest scales of the deflection field is not very constraining (has a large width), and motivates introducing the prior. This is one of the reasons why we use a MAP estimator as opposed to an MLE. Modes that are too noisy would otherwise drive the reconstruction into a regime that renders the result (possibly) unusable.

With this description, we are ready to understand that the Bayesian approach can truly trace the higher order correlations. We will see shortly, that the higher order correlations originate from the deflection operator and Wiener-filtering. To see this, consider first the QE. For brevity, I use $\alpha \equiv \boldsymbol{\alpha}$, and assume that the estimator can be written as a correlator to show its dependence on the fields: $\hat{\alpha}^{\mathrm{QE}} = \langle \tilde{X}^{\mathrm{mod}} \tilde{X}^{\mathrm{mod}} \rangle = \hat{\alpha}^{(0)}$, but should really be understood as a short-hand notation for

---

[12] In the case of jointly reconstructing the lensing potential and curl potential, one might wonder if both fields can be reconstructed simultaneously as the number of unknowns essentially doubles in this case. We show in section 3.2, that, at least for the noise levels and resolutions considered there, we can indeed jointly reconstruct both; to leading order, these fields do not share the same parameters for the reconstruction as their $N^{(0)}$ reconstruction noise do not correlate.

Eq. (22) (and of course rewriting to account for α). The second equality holds as the MAP approach uses the QE as starting point. Here $\tilde{X}^{\rm mod} = \mathcal{B}\tilde{X} + n$ is the QE model. Then, for the first iteration, we use the QE estimate to derive the deflection operator $\mathcal{D}_{\hat{\alpha}^{\rm QE}}$ and estimate a new lensing field, $\hat{\alpha}^{(1)}$:

$$\begin{aligned} \hat{\alpha}^{\rm QE} &= \langle \tilde{X}^{\rm mod} \tilde{X}^{\rm mod} \rangle \,, \\ \hat{\alpha}^{(1)} &= \langle \tilde{X}^{\hat{\alpha}^{\rm QE}} \tilde{X}^{\hat{\alpha}^{\rm QE}} \rangle \,, \\ &\vdots \end{aligned} \tag{49}$$

and the ellipsis denotes further iterations. $\tilde{X}^{\hat{\alpha}^{\rm QE}}$ uses the optimal data model with $D_{\hat{\alpha}^{\rm QE}}$. The deflection operator, to first order, is linear in the lensing potential and acts along the gradient of the Wiener-filtered field,

$$\mathcal{D}_{\hat{\alpha}^{\rm QE}} X^{\rm WF} \propto X^{\rm WF} + \hat{\alpha} \nabla X^{\rm WF} \,, \tag{50}$$

I find,

$$\begin{aligned} \langle \tilde{X}^{\alpha^{\rm QE}} \tilde{X}^{\alpha^{\rm QE}} \rangle &\propto \big\langle (X^{\rm WF} + \hat{\alpha}^{\rm QE} \nabla X^{\rm WF})(X^{\rm WF} + \hat{\alpha}^{\rm QE} \nabla X^{\rm WF}) \big\rangle \\ &\propto \langle \tilde{X}^{\rm mod} \tilde{X}^{\rm mod} \nabla X^{\rm WF} \tilde{X}^{\rm mod} \tilde{X}^{\rm mod} \nabla X^{\rm WF} \rangle \,, \end{aligned} \tag{51}$$

and the second line holds by only collecting all terms that are explicitly second order in $\hat{\alpha}$ (hereby assuming $X^{\rm WF}$ is close to the primordial signal). Applying Wick's theorem, I already find one type of pairing that traces the trispectrum,

$$\begin{aligned} &\langle \tilde{X}^{\rm mod} \tilde{X}^{\rm mod} \nabla X^{\rm WF} \tilde{X}^{\rm mod} \tilde{X}^{\rm mod} \nabla X^{\rm WF} \rangle \\ &\qquad \ni \underbrace{\langle \tilde{X}^{\rm mod} \tilde{X}^{\rm mod} \tilde{X}^{\rm mod} \tilde{X}^{\rm mod} \rangle}_{\text{trispectrum}} \langle \nabla X^{\rm WF} \nabla X^{\rm WF} \rangle \,. \end{aligned} \tag{52}$$

This shows in simple ways that an iterative approach captures higher-order correlations. Needless to say, this schematic derivation is too simplified to make any quantitative statements. Further, the result is after one iteration, for a perturbative deflection operator, and I have omitted many more contractions. A rigorous study would become quickly much more involved and include non-perturbative deflection, and the non-linear Wiener-filtering in α that further induces higher order correlations.

It is for this reason that analytic descriptions of iterative lensing reconstructions are currently somewhat ‘underdeveloped’. It would certainly be interesting to describe the reconstruction biases analytically by calculating the variance of the estimator as in the previous section, and understand - at least qualitatively - the individual higher-order terms that enter. However, the iterative approach makes it difficult to trace exactly the reconstruction biases, and the intermediate, iterative steps depends on the path that is chosen towards the MAP point, that is the concrete Wiener-filter implementation and gradient update. Luckily, at the MAP point, things become clearer again and the reconstruction agrees well with a forecast method developed in [45]. They show in essence that the above iterative scheme boils down to using partially delensed spectra for calculating the reconstruction biases.

Iterative lensing reconstruction comes at a cost. It is computationally much more demanding, partly as it also requires the reconstruction of both the primordial fields and the lensing potential, and calculation of the deflected maps, which is not the case for QE.

Nonetheless, optimal methods improve the lensing reconstruction compared to QE, which is an interesting, and sometimes much needed aspect for many science cases as discussed in the following.

0.3.1 A New Light on Science Cases

The MAP approach can be seen as an improvement to QE for two reasons. First, it generates data products that have a higher cross correlation to the true input (that is, to the underlying deflection field, and primordial CMB fields). Second, the data products are built using an estimator model that traces higher order correlation than two point and thus contain entirely new information that QE does not provide. Another important aspect is that the reconstruction error are in principle unbound and only depend on the experiment noise [42]; a very different behaviour to QE which precision is additionally bound by the lensed spectra.

This has various implications on the final data products and on the CMB science cases:

First and foremost are the improved constraints on the tensor-to-scalar ratio r that we obtain with MAP and for the polarization fields due to the reconstruction error being unbound. A forecast on this is shown in Fig. (4), where I show the improvement on the residual lensing amplitude A_{lens} as a function of the noise level of the experiment.[13]

Second, in the case of the T and E maps the reconstructed fields resemble a better estimate of the earliest times of the Universe than for QE. This can be beneficial; if the estimate better resembles the early Universe, they will correlate less with late time physics, perhaps resulting in lower biases, among other.

Third, as the MAP estimated deflection field traces higher order correlations of the underlying CMB fields, any external tracer that also contains higher order correlations may potentially show different behaviour than QE and inform us in completely new ways. An example of this may be studies on extra galactic foregrounds (or external tracers coming from large-scale-structure probes and galaxy cluster surveys), that have non-Gaussian distributions, and are also correlated with the deflection field. As an example, we study the impact of higher order correlations of a non-Gaussian deflection field onto our MAP estimator in chapter 4 and find that it is less biased than QE. In some sense, optimal methods such as the one presented in this thesis can provide means to study higher order correlations like never before, albeit in a brute-force manner.

---

[13] Measuring neutrino properties can also benefit from the the reconstruction error being unbound as their properties leave a direct imprint in CMB lensing. However, the effect of the neutrino mass on the deflection field reconstruction is degenerate with other parameters such as the depth of reionization τ, the amount of baryonic matter, and the scalar amplitude A_s. The CMB-S4 science goals on the sum of the neutrino mass can actually be reached without MAP [45, 46].

0.3.2 Some Unexplored Charters

One avenue is to optimally and simultaneously reconstruct the deflection field, and foregrounds that correlate with the CMB, or other fields such as the cosmic infrared background. It is left to explore how to accurately incorporate non-Gaussian fields into the likelihood, however. Related to this, there has been no study so far on how to mitigate biases in the MAP approach that arise from extra galactic foregrounds that are correlated with the underlying deflection field. It would be interesting to perform lensing reconstruction jointly on the individual frequency maps rather than the component separated maps. Also, further studies on real data would be highly beneficial for the further development of `Delensalot`.

Another path would be to include non-Gaussianity by starting with non-Gaussian probability functions to build the likelihood model. However, the primordial CMB appears to be Gaussian, and the amount of primordial non-Gaussianity, parameterized by f_{NL}, is constrained to be rather small [47], making this effort somewhat difficult to motivate. And Gaussian probability distribution in the likelihood, Gaussian covariance, and using a Gaussian prior for the deflection field does not prevent us from reconstructing non-Gaussian features, as we show in chapter 4. Furthermore, tailored approaches to probe non-Gaussianity might be better suited by directly building a bispectrum estimator [48, 49]. On the bright side, delensing can still help reduce lensing-induced biases on the estimation of f_{NL}.

Some technical topics that come to mind: it would be interesting to explore other starting points than QE. As the posterior clearly is non-Gaussian, a QE starting point, implicitly assuming Gaussianity of the posterior, is not optimal. A machine learning approach could potentially help here, and could serve as a provider of starting points from which the exact algorithm then converges onto the MAP point.

The MAP approach raises the question if there are "interference-terms" between the different iterations that result in sub-optimal directions for the descent along the lensing posterior. A thorough study on the implementation of the Wiener-filter, or in the update on the MAP point in the gradient step Eq. (44) could explore this. This also leads to the question about knowing whether or not we have actually reached the MAP point.

A final source of complication is that the estimator is numerically not exact due to the Wiener-filtering being applied. We use a conjugate gradient (CG) solver. Depending on the accuracy, numerical inaccuracies lead to inaccuracies on the filtered modes, which is why the accuracy parameters of the CG step is often a much varied parameter in the reconstruction. It would be interesting to study the impact of the CG accuracy onto the estimates of the iterative estimator more thoroughly in the sense of,

$$\hat{\phi}^{\mathrm{QE,CG}} \approx \hat{\phi}^{\mathrm{QE,exact}} - \delta\hat{\phi}^{\mathrm{QE}} \,. \tag{53}$$

$\delta\hat{\phi}^{\mathrm{QE}}$ is a function of the multipole and generally depends on the accuracy of the method and the method itself. This allows to study terms of the form $\langle \delta\hat{\phi}^{\mathrm{QE}}\delta\hat{\phi}^{\mathrm{QE}}\rangle$: if some of the scales of this correlator were insensitive to the actual reconstruction, they could be traced and forced to converge faster. Additionally, with the conjugate gradient being an iterative approach, it might be worth exploring if the initial iterations can be replaced by machine learning models. This is in particular true for the pre-conditioner, a technique that helps with the convergence by binning the problem across scales first, before solving the full problem, if applied.

We study the iterative lensing reconstruction in more detail in chapter 3, where we discuss the full Bayesian approach, and how to achieve this on the full sky inclusive of foreground, masking, and inhomogeneous noise in the context of tensor-to-scalar ratio searches. A thorough discussion on the development of `Delensalot` that was done after the implementation of the curved sky code is given in chapter 1. We discuss another application that studies non-Gaussian deflection fields in chapter 4.

0.4 CMB Data Analysis Pipeline

I discuss the pipeline needed for testing and validation of the analysis of the CMB science goals, and how lensing reconstruction is integrated.

Real data is challenging. As the true input is unknown, we need to make sure to clearly understand the error budget and systematics, and each and every step of the pipeline so that results become interpretable. This is generally achieved by running the analysis pipeline on simulations for which the true input is known. In the following, I describe a general CMB lensing data analysis pipeline on simulations.

Simulation of the primordial fields: these are the CMB fields after the evolution of the Boltzmann equation, and may also include the effect of (patchy) reionization. To generate these simulations, parameters of the inflationary phase are chosen, and the Universe evolves using a Boltzmann solver [50, 51] to eventually obtain the auto and cross power spectra of the CMB fields. Additional parameters have an influence such as the optical depth τ of reionization. Sampling realizations of the final spectra then gives the primordial CMB fields.

Simulation of the sky: during their journey across the Universe, the CMB photons are slightly deflected which results in a change of the distribution of the anisotropies as discussed in the previous sections. A realization of the mass distribution is required which can be simulated either by highly realistic many-body-simulations and ray tracing [52] (which has the advantage of including post-Born effects), or by sampling of the expected power spectrum of an assumed lensing potential itself. This deflection field can then be used to lens the primordial CMB. Most notably, this can be done by [53, 54, 1, 55]. Secondary anisotropies such as the integrated Sachs Wolfe (ISW) effect [56], the Rees-Sciama effect [57], the thermal and kinematic Sunyaev-Zel'dovich effect [58], cluster lensing, and reionization may be of interest and can additionally be simulated and included, as well as galactic foregrounds that do not correlate with the deflection field, but add to the overall noise budget. Mainly, there exist dust, and synchrotron foregrounds emitted from our own galaxy and simulation suites such as `PySM` [59] can be used to include them. Ultimately, different galactic foreground models are analysed simultaneously to quantify impact onto the analysis.

Simulation of the atmosphere and human activity: For ground-based telescopes, additional complications such as atmospheric noise or satellites passing by the field of view potentially have to be taken into account (in the future).

Simulation of the experiment: this includes modelling of the transfer function that describes the actual readout of the measured sky and includes the beam, and the scan strategy of the experiment, as well as the simulation of the noise. The experiment may consist of multiple detectors at different frequencies, for which the simulations need to be generated accordingly.

Component Separation: CMB lensing reconstruction, in its current form, requires some kind of pre-cleaning of the data. Generally, this is done by what is called component separation. Through component separation, the different maps of the experiment at different frequencies are combined in a way to reduce variance, and to filter out components other than the CMB. This works in particular well for components that are frequency dependent, which makes them detectable and distinguishable from the CMB, as the experiments are calibrated in a way that the CMB has the same power across all frequencies of an CMB observation [60, 61, 62, 63]. Component separation can be validated using likelihood searches to recover the input foreground parameters, which helps understanding biases that enter through component separation.

Reconstruction of the lensing field, primordial CMB, and B-lensing templates QE and iterative lensing reconstruction are performed on the component separated maps to extract the effect of CMB lensing. This is the main topic of this thesis, discussed above, and in more detail in the subsequent chapters.

Estimation of cosmological parameter: the effect of B-lensing on the observed B map can be removed with the B-lensing template, which reduces the overall variance. As the power spectra are sensitive to the cosmology, they are fed into Markov Chain Monte Carlo (MCMC) solvers to determine the parameters that best fit the data [64, 65]. They are arguably quite sensitive to the likelihood at choice, so caution is advised. Eventually, this then gives results on r, $\sigma(r)$, the neutrino mass, and all initial parameters that we started with.

1

Developing the Software: `Delensalot`

Figure 1.1 shows the `Delensalot` logo, a byproduct of the initial work of this thesis: the duck is a reference to the underlying SHT solver `DUCC`[1], the scarf emphasises the integration of `DUCC` into our code; this integration was initially done by our Bachelor student Samuel Šimko in the scope of his seminar project. The magnifier refers to the task of lensing, and the sword is a reference to the software name `Delensalot` that is phonetic similar to "Sir Lancelot", one of the knights of the round table.

Figure 1.1: The `Delensalot` logo.

1.1 Motivation

With the advent of CMB experiments reaching sensitivity surpassing 5μK-arcmin, QE have become sub-optimal, necessitating more advanced methods for lensing reconstruction. One such method is developed and implemented through `Delensalot`, the first publicly available software using a Bayesian approach for iterative lensing reconstruction at the map level that performs the task on the full-sky in reasonable time. This novel approach is essential for optimally reconstructing the lensing field in current and upcoming CMB experiments, particularly those sensitive enough to resolve B-lensing.

---

[1] `https://pypi.org/project/ducc0/`

The development of `Delensalot` began prior to this thesis, initially demonstrated with a flat-sky implementation that also included masking [20]. It was subsequently upgraded to curved-sky applications, and its effectiveness on realistic sky models with inhomogeneous noise was demonstrated in a CMB-S4 collaboration paper in the scope of this thesis (see chapter 3.3.2.1).

`Delensalot` is not yet at a software maturity level (SML) that would make it broadly accessible to a wide range of users with low risk (SML-5). Currently, it is mainly used by experts due to the complexity of the task it addresses and the specialized nature of its application; most research relies on QE, and only a mere few CMB observations that are currently running (such as the South Pole Telescope (SPT)) provide data sensitive enough to fully benefit from the MAP reconstruction and thus from `Delensalot`. A reasonable reconstruction also requires the right choice of many parameters, including those related to the data, the experiment, and the reconstruction itself. With around 100 parameters involved, the process is far from automatic.

As with every software, an execution calls underlying functions in the right sequence, in the following referred to as the pipeline. Versions of `Delensalot` prior to this thesis had one access point; the usage of what I call a *parameter file*. In essence, a parameter file is the whole pipeline: raw function calls, intervened with the necessary parameters, exceeding hundreds of lines of code. The logical connection of a pipeline is merged with the parameters themselves, and using a parameter file requires the knowledge of the underlying functions and logical connection. While suitable for experts and the development of different reconstruction pipelines, this entry point presents a high barrier for broader use, in particular if we want to provide `Delensalot` to the greater cosmologists community. Thus, `Delensalot` was rather positioned around SML-2: expert knowledge needed and software poses somewhat high risk.

`Delensalot` is proven to be feasible [2, 3, 5, 23, 43, 45]; its results are robust and promising and it was ready to be elevated onto the next level on the SML scale and to make it more accessible to everyone. This means lowering the threshold, removing the need of expert knowledge to produce sensible results, and reducing overall risk of using the software by introducing validation and testing. This chapter describes the efforts that aim to reach this.

Central to this chapter is the implementation of the `Delensalot` language model (`DLM`) and the "job-stream". The `DLM` simplifies the complex task of some of the pipelines that may include iterative lensing reconstruction. In particular, it provides almost one-click solutions for standard applications like reconstructing the lensing field or building the B-lensing template for measuring the tensor-to-scalar ratio r, as required by many collaborations. The `DLM` organizes parameters in a structured, human-readable format and connects them to the core functions of `Delensalot` via model transformers, significantly reducing the complexity of running these pipelines. `Delensalot` provides well optimized pipelines that can automatically be run by the user and be controlled by a set of parameters. At the same time, `Delensalot` is developed with backward-compatibility for parameter files, so that existing workflows that depend on them are not affected. The job-stream combines minimal worker units, dubbed "jobs", that are executed sequentially.

The shortest possible code for a complete `Delensalot` analysis can be as short as three lines of code! This is a severe reduction of complexity for the usage of `Delensalot` that previously required about 250 lines of code to run. As a concrete example, the sequence of instructions to run a full analysis is shown in the following snippet:

```
from .dlensalot_mm import DLENSALOT_Model as DM
from delensalot.run import run
dmodel = DM(defaults_to='CMBS4_fullsky_polarization')
drunner = run(config=dmodel)
drunner.run()
ana = drunner.collect_model()
```

The first line (after the imports) creates a `DLM`: for each parameter that is not explicitly chosen by the user (in this case here, all), `Delensalot` automatically defaults to the values defined for a general full sky polarization reconstruction for CMB-S4-like experiments, `defaults_to='CMBS4_fullsky_polarization'`. The second line validates the `DLM` and initializes the analysis. The third line executes the analysis.

The additional fourth line returns the full `Delensalot` analysis inclusive of convenience functions to access the results. Of course, a concrete analysis requires a correct configuration of the `dmodel`, and details will be discussed in the following sections.

This new architecture also allows for the development of robust convenience functions. In particular for the idealized of a full sky analysis, `Delensalot` now also provides functions to calculate map-delensed fields, lensing templates, delensed B-modes, and residual B-mode power spectra, with one line of code, as shown in the following snippet:

```
import delensalot
blt = delensalot.map2tempblm(obs_map)
delblm = delensalot.map2delblm(obs_map)
Cldel = delensalot.anafast(obs_map)
```

Most complication is thus hidden and greatly simplified. While not necessarily providing optimal reconstruction, these functions allow for very simple and fast access to results that were otherwise difficult to reach.

This chapter gives a high level introduction into the `Delensalot` architecture. Furthermore, I discuss the following contributions:

- a user-friendly simulation library and handler for generating and managing CMB realizations, and handling external input, see Sec. 1.4.
- implementation of a new interface using the `DLM`, see Sec. 1.3.
- introduction of model-transformers to map between `DLM` and other models, see Sec. 1.5.

- a structured approach to managing default settings.
- a single entry point for executing `Delensalot` and analyzing results, see Sec. 1.3.
- introduction of a job-stream, see Sec. 1.5.
- introduction of a collect-run structure for handling message passing interface (MPI) calls. See Sec. 1.5.
- enhanced tracking and validation of analyses ensuring robust and safe resumption, see Sec. 1.3.
- continuous integration (CI) for testing and validation, see Sec. 1.7.
- comprehensive documentation and tutorials, see Sec. 1.6.

1.2 Major Milestones

Throughout the last years, `Delensalot` has seen many changes to the code base. Fig. 1.2 puts the most important achievements chronologically into context. The first "flat sky" implementation was demonstrated in 2017. At the start of this PhD, "itercurv", the first curved sky pipeline, was developed. In summer 2021, our Bachelor student Samuel Simko developed and integrated "`lenscarf`", allowing for SHTs on custom geometries. The "DLM" was introduced at the start of 2022, and marks the birth of `Delensalot`. Shortly after, the simulation library "simlib" was built to simplify simulation management. "`Lenspyx2.0`" was integrated in 2023, significantly speeding up the deflection operation and increasing accuracy. As an outlook, I would like to mention the integration of a GPU implementation of the deflection operator, "cunuSHT".

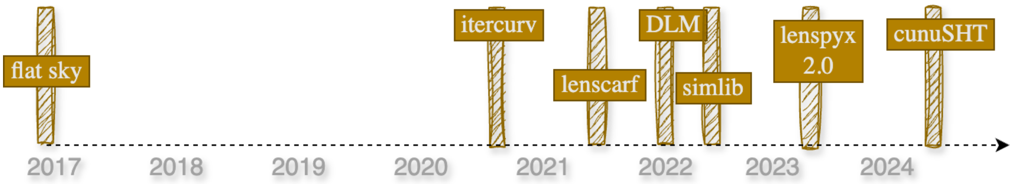

Figure 1.2: Milestones during the development of `Delensalot`.

1.3 The `Delensalot` Language Model

A `Delensalot` language model (DLM) consists of 14 data structures, each containing various parameters, and in total defines a complete `Delensalot` analysis. Its capabilities reach beyond standard iterative lensing reconstruction: it automates the generation of simulations, QE analysis, delensing and ϕ power spectrum analysis, and Wiener-filter maps, as well as mean-fields. The data structures are shown in Fig. 1.3, and the full model inclusive of all parameters is shown in Appendix A.

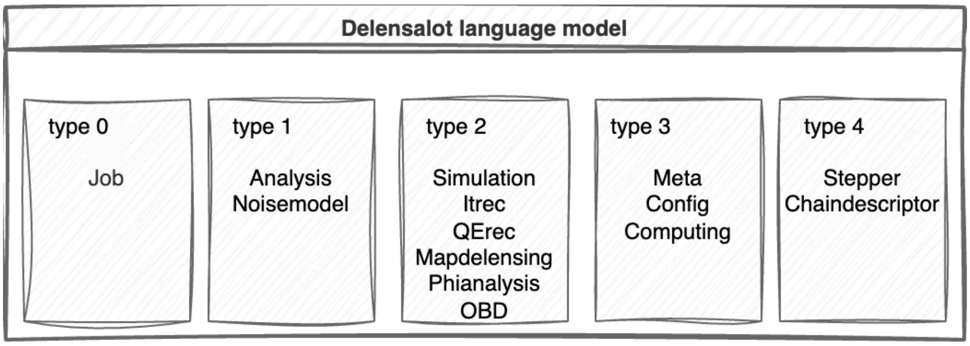

Figure 1.3: The `Delensalot` language model and its data structures. Each structure contains a set of parameters. A complete configuration consists of all data structures and their parameters.

For the sake of this chapter, each structure is grouped into types.

There is only one **type 0** structure, "Job". It controls which job to execute in the analysis: Building the overlapping B-mode deprojection matrix, generating Simulations, QE reconstruction, MAP reconstruction, Map-level delensing, and the analysis of the power spectrum of ϕ.

In addition, there are four types of data structures:

- **type 1** controls modelling aspects: Analysis, Noisemodel.
- **type 2** configures the individual jobs: Simulation, OBD, QErec, Itrec, Mapdelensing, Phianalysis.
- **type 3** controls behaviour and does not affect the analysis: Meta, Config, Computing.
- **type 4** controls very specific, underlying behaviour of the gradient update, and conjugate gradient method: Stepper, Chaindescriptor.

The third and fourth type does not generally need to be changed by the user and can be left unconfigured; `Delensalot` takes care of it anyway.

The first and second type of data structures are most important for everyday usage.

The data structure "Analysis" contains the configuration of the overall run such as the type of quadratic combination of fields (e.g. TT, EB, etc.), filter, and detector settings.

The data structure "Noisemodel" collects parameters for the modelling of the noise: how to build the noise variance map, and how to deal with masking, if applicable.

The type 2 data structures configure the individual jobs. In the case of the QE and iterative reconstruction, jobs are divided in tasks and controlled by the "tasks" parameter. A task is the calculation of the deflection field, the B-lensing template, and the meanfield. Further, there are parameters that define the behaviour of the reconstruction and configure the Wiener-filtering, among others.

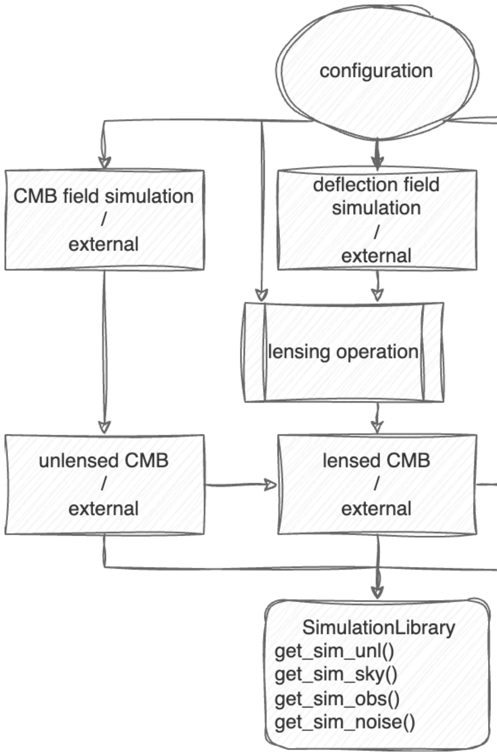

Figure 1.4: The simulation library. "external" refers to the possibility of it being replaceable by external data

For each analysis that is executed, `Delensalot` reads the `DLM` and validates the parameters. If a provided parameter does, for example, not match the expected data type or range, validation will not succeed and provide feedback to the user. If a parameters, however, is not set by the user, `Delensalot` falls back to a set of predefined default values, thereby handily simplifying the amount of parameters that need to be set. The user can choose which of the predefined default values `Delensalot` should fall back to, with many for typical configuration scenarios to choose from (such as analysing data on the full sky, masked sky, for CMB-S4 like experiments, only using polarization estimator, among other).

Configurations are also tested against existing analyses, guaranteeing that reconstruction results are not mixed up. For each analysis, a directory is created which contains a copy of the `DLM`, intermediate results, and the analysis results. If analyses folders match, the already existing and currently executed `DLM`'s will be compared against each other. Mismatching parameters result in a failed validation and termination of the execution for a subset of parameters.

I refer to Appendix A for a complete list of parameters, and to the next sections for examples of how to use it. I also provide an online documentation, and tutorials, as discussed in Sec. 1.6.

1.4 The Simulation Library

This section describes the implementation of the type 2 data structure "Simulation" and how to use it. One of the possible jobs of `Delensalot` is the generation and handling of simulations. While a simulation library already existed since the development of `plancklens`[2], building a new simulation library was motivated to seamlessly integrate it with the DLM.

`Delensalot` takes a single CMB field (or a combination of two different fields) and performs iterative lensing reconstruction to optimally reconstruct the underlying deflection field. This is either done on real data, or simulations. In many cases, `Delensalot` needs to generate its own simulations.

There is some amount of freedom how data or simulations are usually provided, and a performant library must be able to safely handle all of them:

- The CMB field can either be polarization, or temperature. In the case of polarization, fields may come as spin-2 (Q,U), or spin-0 (E,B) fields.
- The simulations may come as either harmonic coefficients or as real-space maps. It is also possible to start with mere power spectra.
- CMB fields can either be provided as unlensed, lensed, or observed maps. For unlensed maps, this means that the simulation library must be able to perform lensing on the data and apply the transfer function to generate sky, and observed maps, respectively. It must be capable of generating and applying noise.

Each of these come with constraints and instructions, that are partly shared and dependent on each other. It thus offers means to automate most of it. The motivation behind a new simulation library is to handle all the above in a consistent and user-friendly manner, hiding most of the detailed operations.

The simulation library is implemented as follows, and shown in Fig. 1.4. I implement 3 classes, each handles a different stage of the simulation: unlensed, lensed, and observed sky. They are interconnected to share common parameters. Data provisioning uses simple data structures that define the properties (configuration) for the individual generation of realizations (CMB field simulation, deflection field simulation, noise simulation). Configuration further contains instructions for the lensing operation and experiment configuration. An outward facing class (SimulationLibrary) provides high-level functions for accessing simulations. For the generation and access of the simulations, each high-level function comes with the same set of parameters:

`simidx` is an identifier of a particular realization.
`field` refers to either temperature or polarization fields.
`space` refers to either power spectra, harmonic coefficients, or real space maps.
`flavour` refers to the fields either being unlensed, lensed, or observed maps.

---

[2] `plancklens` provides routines for QE analyses and is integrated into `Delensalot`: `https://github.com/carronj/plancklens`.

`spin` refers to the polarization maps either being Q/U (spin-2) or E/B (spin-0 maps). In case of temperature, spin is always assumed zero.

Five outward facing functions facilitate easy access to different simulation products and stages:

```
get_sim_unl(**kwargs)
get_sim_sky(**kwargs)
get_sim_obs(**kwargs)
get_sim_noise(**kwargs)
get_sim_phi(**kwargs)
```

with `**kwargs` being the same set of parameters as described above.

Default configurations hide underlying parameters that may be irrelevant to the user. The simulation library is built to always generate observed CMB fields, independent of the simulation stage that the configuration provides. If this is not possible, the configuration is not valid and the library returns feedback to the user. The library, of course, cannot go the other way round: if only observed maps are provided by the configuration, it is still a valid configuration. The library simply cannot provide lensed, or unlensed maps from it.

A typical configuration of the simulation data that starts with unlensed CMB maps is shown in the following code snippet. Note, that this data structure is also part of the `DLM`.

```
simulationconfig = DLENSALOT_Simulation(
    space = 'map',
    flavour = 'unl',
    lmax = 4096,
    phi_lmax = 5120,
    spin = 2,
    libdir = data_dir,
    libdir_noise = data_dir_noise,
    fnsnoise = {
        'E':'E_mc{{:04d}}.fits',
        'B':'B_mc{{:04d}}.fits'
    },
    fns = {
        'Q':'Q_{{:04d}}.fits',
        'U':'U_{{:04d}}.fits'
    },
    CMB_modifier = func,
    transfunction = gauss_beam(2.3/180/60 * np.pi, lmax=4096),
    geominfo = ('healpix', {'nside': 2048}),
)
```

The next snippet initializes the simulation library with the above configuration using the outward-facing `Simhandler`.

```
from delensalot.sims.sims_lib import Simhandler
s = Simhandler(**simulationconfig.__dict__)
```

The following snippet shows example calls of how to access the simulations from the simulation library above, for a few combinations of parameters:

```
Tu  = s.get_sim_unl(0, spin=0, space='alm', field='temperature')
EBu = s.get_sim_unl(0, spin=0, space='alm', field='polarization')

phi = s.get_sim_phi(0, space='alm')

Ts  = s.get_sim_sky(0, spin=0, space='alm', field='temperature')
To  = s.get_sim_obs(0, spin=0, space='alm', field='temperature')
QUo = s.get_sim_obs(0, spin=2, space='alm', field='polarization')
Tn  = s.get_sim_noise(0, spin=0, space='alm',field='temperature')
```

This library does not automatically store simulations on disk. Instead, this can be achieved by running a `Delensalot`-job called `generate_sim`,

```
from .dlensalot_mm import DLENSALOT_Model as DM
from delensalot.run import run
dmodel = DM(simulationdata = simulationconfig)
drunner = run(config=dmodel, job='generate_sim')
drunner.run()
ana = drunner.collect_model()
```

Note that this is a completely valid `DLM`. It is the exact same entry point as one would use for running iterative lensing reconstruction. In the example above, it is merely used for generating the simulations. Hence, all irrelevant parameters do not have to be provided.

1.5 Implementation

The configuration file (that contains an instance of a `DLM`) contains requested jobs, for example "simulation generation", "MAP reconstruction", "delensing". Executing `Delensalot` with a configuration file executes the jobs as a "job-stream"; as jobs have an implicit hierarchy, they are automatically executed sequentially in the right order. While the jobs are executed sequentially, each job is parallelized using the message passing interface (MPI). The parallelization is generally done across simulations. This is shown in Fig. 1.5: pointing solid arrows show the dependency of one job to another. It is possible that the configuration only requests the highest hierarchy, in which case `Delensalot` internally builds the job-stream by adding all required lower-hierarchy jobs. If the user requests MAP reconstruction, `Delensalot` first builds the B-mode deprojection (OBD) matrix and generates the simulations, and then executes QE reconstruction.

`Delensalot`, upon execution, validates the configuration, sets default values for unconfigured parameters, and builds the job-stream using the configuration as provided by the `DLM`. The job-stream and `DLM` data structures correspondence is shown in Fig. 1.5. Some data structures affect the whole job-stream (analysis, meta, config, computing), while others configure a particular job (type 2 data structures).

Some jobs contain tasks, which can be thought of as sub-jobs. For each task in

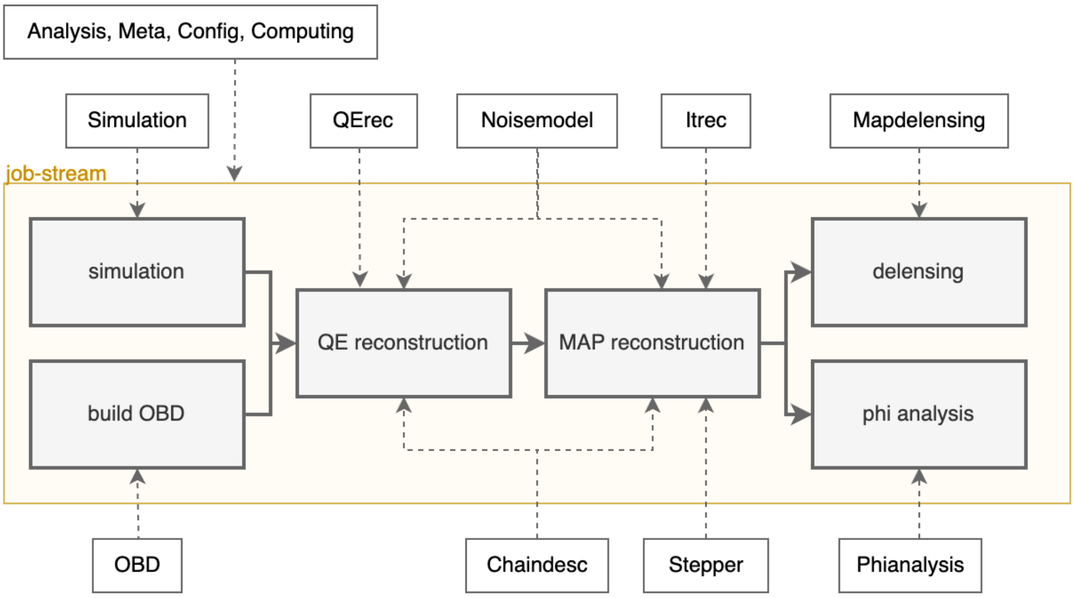

Figure 1.5: The `DLM` data structures and connection to the job-stream.

a job, all simulations are collected that need execution, which is decided based on the configuration and existing files in the analysis folder. Then, with the collected execution list, these are executed in parallel using the message passing interface (MPI).

1.5.1 Transformer

The configuration of an analysis in `Delensalot` can be represented in three distinct models, as shown in Fig. 1.6. These models are used to manage and transform data in stages, making the system flexible and efficient. Let's break down each model and its role:

1. Outward-facing `DLM`: This is the most basic model, containing simple data structures. It provides an initial high-level view of the configuration and is easy to manipulate by users.
2. Core Model: This is an internal, validated representation of the `DLM`. It holds all the necessary parameters and lower-level instructions that the core functions of `Delensalot` can process.
3. Job Model: The final representation is the job model, which is derived from the core model. It contains all the detailed instructions and parameters needed for executing specific jobs within `Delensalot`.

The system works by mapping the parameters from the `DLM` to the core model and then transforming them into the job model, ensuring that all instructions are correctly passed along at each stage.

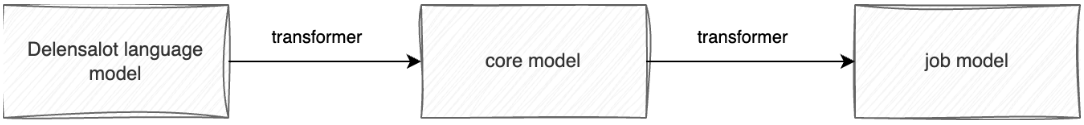

Figure 1.6: The different model representations and their correspondence.

To transition between these different models, `Delensalot` uses a model transformer. A transformer, denoted as T, converts one model M into another model M' based on specific instructions I.

$$M' = T(M, I)\,. \tag{1.1}$$

In some cases, these instructions might simply be another model or an identifier. This transformation process allows the system to convert high-level user input into the appropriate format for execution.

This approach is implemented in `Delensalot` to handle transformations between the `DLM`, core model, and job model. Below is an example of how to transform a core model into a job model using the function `l2d_T()` and an identifier, `Job_id`, which is part of a set of defined jobs:

```
job_model = transform3d(configuration_model, job_id, l2d_T())
```

Here, the job model is created as a fully defined model, containing all the necessary instructions for execution. This model can then be passed directly to the job class for running the task.

Working with transformers offers several advantages,

1. Code Simplification: Wherever a model is needed, it can be generated on the fly using a transformer and a base model. For example, generating a job model is straightforward when starting with the core model.

2. Hidden Parameters: Parameters for different models are encapsulated within the transformation instructions, ensuring proper initialization. This also reduces the clutter in the code by moving parameter handling outside of the main structure.

Beyond handling the core, `DLM`, and job models, transformers in `Delensalot` serve various purposes. They can be used to generate models for filters: you can transform a job model into specific MAP filters by using particular instructions. An example for creating MAP filters is shown in the following snippet.

```
MAP_filters = transform(job_model, MAP_transformer())
MAP_filter = transform(job_model, MAP_filters())
```

Transformers are also used to generate iterators: different iterators can be generated for the MAP reconstruction. For instance, you can create an iterator model from a job model, where the instructions depend on the `DLM` and the simulation index, as shown in the following snippet.

```
iterator_model = transform(job_model, it_T(job_model, 0, Dlm))
```

One final use-case worth mentioning is the ease of validation. Since the instructions used in transformations generally remain unchanged, the output model (M') can be easily validated against expectations. If M' does not match the expected output, then the input model M is likely faulty, allowing for efficient debugging.

1.6 Usage, Documentation and Tutorials

Using this new interface, the workflow goes as follows:

1. Define a `DLM` (`Delensalot` can receive this as a standalone configuration file). Choose the jobs you would like to execute.
2. Call the `Delensalot` runner with the configuration file. The configuration file will be validated, and the job-stream will be set up.
3. Run the analysis. The job-stream gets executed.
4. Upon finishing the job-stream, `Delensalot` terminates.
5. The `Delensalot` analysis-collector can now be called to return the analysis results.

A code snippet of this workflow is shown in Sec. 1.1. Concrete examples can be found on our GitHub page[3]. We provide various tutorials for,

- running quadratic estimator and iterative lensing reconstruction on the full and masked sky.
- delensing B-modes.
- mitigating the impact of the meanfield and internal lensing bias.
- forecasting residual lensing amplitudes.
- . . . and many more!

Additionally, a documentation can be found at `https://delensalot.readthedocs.io/en/latest/`.

1.7 Testing and Validation

Within frameworks like `Delensalot`, robust testing and validation mechanisms are important to ensure the reliability and accuracy of an ever-changing and evolving software. As the code base exceeds thousand of lines of code, a structured approach for tracking code changes and providing means to validate changes are crucial.

`Delensalot` employs continuous integration (CI) to automate the testing and validation process, streamlining development and ensuring that the integration of

---

[3] `https://github.com/NextGenCMB/delensalot/tree/main/first_steps`

new features meets quality standards and does not degrade existing features. We utilize GitHub's CI feature to provide rapid feedback to defects introduced in the code base.

We use CI to validate various components of `Delensalot`. This includes the validation of the job models, and verifies that the correct computational tasks are performed for the specified configuration files. This step verifies that all components of the job model are correctly instantiated. Among other, it validates that the filter models within the reconstruction jobs.

We additionally use CI to run a full job-stream, and for various estimators. This goes beyond testing individual components and verifies that the job-stream operates as expected. This full job-stream starts with the generation of simulation data, and eventually performs delensing, hereby not only testing the result against the expected outcome, but also provides means to check its performance.

`Delensalot` is shipped with a suite of pre-configured tests that cover a wide range of scenarios and edge cases. With these tests, users can check the performance individually, if desired. This is in particular helpful for users and developers who are modifying the code.

The complete list of tests can be found on the GitHub[4] page.

1.8 Outlook

Building the `DLM`, introducing the minimal worker units "jobs" of `Delensalot`, and structuring the execution of `Delensalot` via a job-stream introduced new means for testing and validation, and simplified the every day's task of generating B-lensing templates for CMB-S4 and PICO's analysis pipeline on simulations. It even provides automated ways for delensing, analysing lensing potential power spectra, and accessing the results.

The `DLM` is far from optimal and has many sharp edges that need honing. Also, the workflows are not as intuitive as one might wish for, thus still requiring an expert providing necessary information for support.

Another limitation is that I do not currently provide means to build a custom job-stream. To achieve this, in the end, one currently has to return to parameter files.[5] Tailored job-streams for the study of, for example, the mean-field, the likelihood itself, or the correlation to external tracers would likely be interesting to have. However, the `DLM` is developed with the simplification of supporting r-analysis, putting this effort somewhat out of scope. Nonetheless, a future development of `Delensalot` can include this, if needed.

The question arises what are the most sensitive minimal worker units that provide enough flexibility. Instead of "MAP reconstruction", it might be better to have the underlying functions such as calculating the Wiener-filter and gradient of the

---

[4] `https://github.com/NextGenCMB/delensalot/tree/main/tests`

[5] In fact, a change in the pipeline can be achieved by building a new job and adding it to the job-stream in the corresponding place, but requires knowledge of the underlying architecture of `Delensalot`. This could, for example, be an extra-galactic foreground cleaning step if this is needed.

likelihood to provide better flexibility. A job stream built with more fine-grained jobs like this may be more advantageous and flexible.

This makes me wonder if iterative lensing reconstruction software, once more mainstream, will be built around a single tool like `Delensalot`. More likely, the individual steps such as calculating the Wiener-filter, the deflection operation, the calculation of the gradient, and the gradient stepper will be standalone tools. The final application of the actual iterative lensing reconstruction might as well just be a simple likelihood maximiser with the above mentioned building blocks maintained and optimized as standalone packages, similar to modern software development where each and every part is a standalone service. This will highly increase robustness, maintainability, and flexibility.

2

Accelerating Spherical Harmonic Transforms

2.1 Motivation

One aspect of the Bayesian method for lensing reconstruction is the introduction of the deflection operator, $\mathcal{D}_{\boldsymbol{\alpha}}$. The deflection operator maps the unlensed field X to the lensed one, $\tilde{X} = \mathcal{D}_{\boldsymbol{\alpha}} X$. On the curved sky, this operation includes a non-uniform SHT. Numerical evaluation is necessarily performed on a grid, and there is some freedom in how to choose the grid for SHT, with some reducing computational complexity of the problem. But, either the undeflected, or deflected field will have to be evaluated at positions that are off-grid, and of course, non-uniform. This means, that the deflection operators task actually involves the more complicated job of calculating a SHT on a non-uniform grid, and no fast algorithm exists for this.[1] Furthermore, our Bayesian approach requires many such transforms: $\mathcal{O}(10^2)$ simulations for sensible error estimation, $\mathcal{O}(10^1)$ MAP iterations per simulation for finding the MAP point, and $\mathcal{O}(10^2)$ conjugate-gradient solver iterations per MAP iteration to find the Wiener-filtered map that matches the accuracy threshold. The Wiener-filtered field is, schematically $X^{\mathrm{WF}} = \frac{C}{C^{-1}+N_{\alpha}^{-1}} X$, where C is the power spectrum "signal", and N_{α}^{-1} is the partially delensed noise covariance matrix. Concretely for the E-mode,

$$N_{\alpha}^{-1} = {}_2\mathcal{Y}^{\dagger}\mathcal{D}_{\boldsymbol{\alpha}}^{\dagger}\mathcal{B}^{\dagger}N^{-1}\mathcal{B}\mathcal{D}_{\boldsymbol{\alpha}}\,{}_2\mathcal{Y}\,, \tag{2.1}$$

where ${}_2\mathcal{Y}$ is a short-hand notation for a spin-2 spherical harmonic transform, $\mathcal{B}$ is the transfer function, and N^{-1} is the noise covariance matrix. Thus, $N_{\boldsymbol{\alpha}}^{-1}$ contains a forward and backward non-uniform SHT. This is repeated for each quadratic combination of the CMB fields, each change in the mask, or the noise model, to name a few. All combined, this often accumulates to about $\mathcal{O}(10^6)$ non-uniform SHT. With SHTs often being the bottleneck of the computation, it is thus quite natural to search for efficient ways of calculating the deflection operator. In addition, a higher accuracy than what we could provide with the previous approach is also desired in some cases.

---

[1] In fact, some faster algorithms are current active research, but none have proven feasible so far due to their parameter limitations or memory usage, among others.

This chapter discusses two paper that aim to efficiently and accurately solve the deflection operator using a much better suited interpolation scheme between the ring-ordered grids: non-uniform fast Fourier transforms.

The first paper here presents `lenspyx 2.0`, a paper in collaboration with Martin Reinecke and Julien Carron, that tremendously improves the previous CPU algorithm of `Delensalot`. While Julien and Martin implemented the CPU routines in their respective packages `DUCC` and `lenspyx 2.0`, I tested the implementations, and ran benchmarks. We decided to compare the new approach not only to our previous solution, but also to many other existing, yet sometimes outdated implementations of the deflection operation. These implementations sometimes had outdated package dependencies and compiler requirements, compliacting installation. I contributed to the paper by writing the "benchmark" section, and integrating `lenspyx 2.0` into `Delensalot`, where it is now extensively used.

The second paper presents `cunuSHT`, a paper in collaboration with Nathanael Schaeffer, Martin Reinecke, Julien Carron, and Adriaan Duivenvoorden, that studies the feasibility of the deflection operator on GPUs.

The opportunity to work abroad during my PhD led to Julien proposing the idea of porting our `lenspyx 2.0` code to GPU, supported by a pre-doctoral program at the Simons Foundation. This initiative allowed me to relocate to New York for five months to lead the project, despite my initial lack of GPU programming experience.

My role involved reimplementing `lenspyx 2.0` on the GPU, a task demanding optimal programming and minimizing data transfers between host and device to compete with established CPU algorithms. The GPU implementation was developed using the structural framework from Martin's and Julien's CPU work, along with Nathanael's SHT implementation from `SHTns`[2], to ensure high performance. My experience with `lenspyx 2.0` provided a strong foundation, enabling a fast development and benchmarking of the new solution.

Motivated by the encouraging results, I aimed to include this work in my thesis and publish it as a standalone paper.

2.2 `lenspyx 2.0`

This section contains the `lenspyx` paper as published on *Astronomy and Astrophysics*.

---

[2] https://nschaeff.bitbucket.io/shtns/

Astronomy & Astrophysics manuscript no. 46717
October 26, 2024

Improved cosmic microwave background (de-)lensing using general spherical harmonic transforms

Martin Reinecke[1], Sebastian Belkner[2], and Julien Carron[2]

[1] Max-Planck Institut für Astrophysik, Karl-Schwarzschild-Str. 1, 85748 Garching, Germany
[2] Université de Genève, Département de Physique Théorique et CAP, 24 Quai Ansermet, CH-1211 Genève 4, Switzerland

October 26, 2024

ABSTRACT

Deep cosmic microwave background polarization experiments allow a very precise internal reconstruction of the gravitational lensing signal in pinciple. For this aim, likelihood-based or Bayesian methods are typically necessary, where very large numbers of lensing and delensing remappings on the sphere are sometimes required before satisfactory convergence. We discuss here an optimized piece of numerical code in some detail that is able to efficiently perform both the lensing operation and its adjoint (closely related to delensing) to arbitrary accuracy, using nonuniform fast Fourier transform technology. Where applicable, we find that the code outperforms current widespread software by a very wide margin. It is able to produce high-resolution maps that are accurate enough for next-generation cosmic microwave background experiments on the timescale of seconds on a modern laptop. The adjoint operation performs similarly well and removes the need for the computation of inverse deflection fields. This publicly available code enables de facto efficient spherical harmonic transforms on completely arbitrary grids, and it might be applied in other areas as well.

Key words. Cosmology – Cosmic Microwave Background – Gravitational lensing

1. Introduction

Weak gravitational lensing by large-scale structure affects radiation from the cosmic microwave background (CMB) in subtle but important ways (Lewis & Challinor 2006) by distorting and smoothing the primordial near isotropic two-point statistics and introducing a large trispectrum that can now easily be detected with very high significance (Aghanim et al. 2020; Carron et al. 2022; Qu et al. 2023). Extraction of the signal is now an important piece of the science case of many CMB experiments (Abazajian et al. 2016; Challinor et al. 2018; Ade et al. 2019) because the lensing potential power spectrum, which probes the formation of structure at high redshift, is thought to be a particularly clean probe of the neutrino mass scale (Lesgourgues & Pastor 2006; Hall & Challinor 2012; Allison et al. 2015). Another reason is that removal of the lensing signal (delensing) from the CMB polarization B-mode (Zaldarriaga & Seljak 1998; Knox & Song 2002; Kesden et al. 2002) is now compulsory in order to place the best constraints on a primordial background of gravitational waves from inflation (Ade et al. 2021; Tristram et al. 2022).

It has long been known that deep high-resolution observations of the CMB polarization in principle allow an extremely good internal reconstruction of the lensing signal (Hirata & Seljak 2003b). Recent years have seen works trying to achieve this goal of capturing a signal-to-noise ratio that is as high as possible in realistic experimental configurations (Carron & Lewis 2017; Millea et al. 2019, 2021; Millea & Seljak 2022; Legrand & Carron 2022, 2023; Aurlien et al. 2022), and some attempts were made on data as well (Adachi et al. 2020; Millea et al. 2021). These methods have in common that they use a likelihood model of the lensing signal, which allowed them to outperform the now standard quadratic estimators (Hu & Okamoto 2002; Okamoto & Hu 2003; Maniyar et al. 2021), which are limited by the amount of lensing B-mode power in the data.

These likelihood-based methods form the main motivation for this work. They are typically much more expensive than current quadratic estimator analysis: Properly modeling the subtle, ~ 2 arcmin remapping effect of gravitational lensing or delensing on CMB maps requires working at high resolution, and these operations must be performed many times before convergence. Numerically speaking, harmonic transforms of data distributed on the sphere are significantly costlier than in flat space. Only a handful of works were so far able to run such optimized reconstructions on large sky fractions and take the sky curvature into account (Aurlien et al. 2022; Legrand & Carron 2022, 2023).

In order to be slightly more concrete, we consider for example the problem of recovering the unlensed CMB from data and an estimate of the lensing deflection. The unlensed CMB likelihood conditioned on the lensing map is Gaussian, but with a covariance that is anisotropic owing to the deflection. If no further approximation is made, recovery of the delensed CMB will involve the inverse covariance (Wiener-filtering), which today can only be performed with iterative methods such as a conjugate gradient, and needs approximately ten iterations to converge. Each iteration requires two remapping operations (one forward operation, and the second operation is the adjoint operation). The construction of optimal lensing mass maps internally from the CMB proceeds by iteratively applying these delensing steps and measuring residual lensing (Hirata & Seljak 2003a,b), and at best, it requires about approximately ten delensing iterations. Hence (again, in the absence of approximations), it seems difficult to reconstruct a single best-lensing mass map with fewer than approximately 200 remapping operations. Sampling methods (Millea et al. 2019, 2021) are in principle orders of magnitude costlier still.

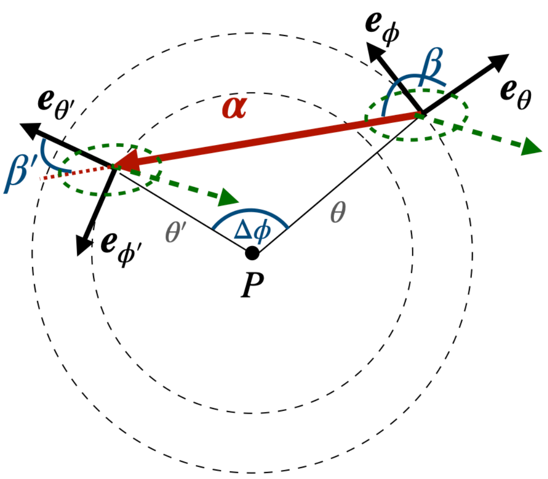

Fig. 1. Lensing geometry and notation near the north pole. The sky curvature is suppressed for clarity. The deflection vector $\boldsymbol{\alpha}(\hat{n})$ lies in the plane tangent to the observed coordinate $\hat{n}$ at latitude θ and longitude φ, and points toward the unlensed coordinate $\hat{n}'$, lying a distance $\alpha = |\boldsymbol{\alpha}(\hat{n})|$ away along the great circle generated by $\boldsymbol{\alpha}$. The lensing remapping for parallel-transported spin-weighted fields like the dashed green vector or ellipse receives a phase correction $e^{is(\beta-\beta')}$ from the rotation of the local θ and φ basis axes.

For these reasons, we have developed an optimized piece of code that is able to perform the deflection operation and its adjoint (the operations are properly defined below) efficiently. While several pieces of software are publicly available to perform the forward-deflection operation, we find that our implementation outperforms them by far, and the adjoint operation (equally as important for likelihood-based methods) appears to be new.

2. Lensing and delensing the cosmic microwave background

To a very good approximation that is valid for next-generation CMB experiments such as CMB-S4[1] (Abazajian et al. 2016), the effect of gravitational lensing is that of a remapping of points on the sphere. The observed CMB intensity signal at position $\hat{n}'$ is related to that of an unlensed position by the relation

$$T^{\rm len}(\hat{n}) = T^{\rm unl}(\hat{n}'), \tag{1}$$

where $\hat{n}'$ is located at a distance α from $\hat{n}$ along the great circle generated by the deflection field $\boldsymbol{\alpha}$. Fig. 1 shows the geometry and our notation. In polarization, and more generally, for any spin-weighted field ${}_s\mathcal{T}$, there is an additional phase factor that is sourced by the change in the local basis between the deflected and undeflected position (hence mostly relevant only near the poles) (Challinor & Chon 2002). This may be written

$${}_s\mathcal{T}^{\rm len}(\hat{n}) = e^{is\chi(\hat{n})}\, {}_s\mathcal{T}^{\rm unl}(\hat{n}')_, \tag{2}$$

where $\chi = \beta - \beta'$ on Fig. 1. We are primarily interested in efficient implementations of both the deflection operator, $\mathcal{D}_{\boldsymbol{\alpha}}$, which from a band-limited set of harmonic modes results in the lensed map on some arbitrary locations, or pixels, $\hat{n}_i$, and of its adjoint $\mathcal{D}^\dagger_{\boldsymbol{\alpha}}$. The forward operation can be written explicitly (for spin-0 fields) as

$$\left[\mathcal{D}_{\boldsymbol{\alpha}} T^{\rm unl}\right]_i \equiv \sum_{\ell=0}^{\ell_{\rm max}} \sum_{m=-\ell}^{\ell} T^{\rm unl}_{\ell m}\, Y_{\ell m}(\hat{n}'_i) \quad \text{(spin-0)} \tag{3}$$

and is thus closely related to the problem of finding an efficient forward spherical harmonic transform to an irregular grid. The adjoint takes a map as input, together with a set of deflected coordinates, to produce harmonic coefficients as follows (again, here for a spin-0 field):

$$\left[\mathcal{D}^\dagger_{\boldsymbol{\alpha}} T^{\rm len}\right]_{\ell m} \equiv \sum_i T^{\rm len}(\hat{n}_i)\, Y^*_{\ell m}(\hat{n}'_i) \quad \text{(spin-0)} \tag{4}$$

for $|m| \leq \ell \leq \ell_{\rm max}$. In the most general situation, the points $\hat{n}_i$ and $\hat{n}'_i$ are completely arbitrary, such that the code presented here forms in fact a spherical harmonic transform (SHT) pair that works on any pixelization of the sphere.

In situations like those encountered in CMB lensing, the points $\hat{n}_i$ cover the sphere according to a reasonable sampling scheme (e.g., a Healpix (Górski et al. 2005) or Gauss-Legendre grid), and $\hat{n}'_i$ are the deflected coordinates given by Fig. 1. When quadrature weights are added to Eq. (4), the sum becomes an approximation to an integral over the observed coordinate,

$$\left[\mathcal{D}^\dagger_{\boldsymbol{\alpha}} T^{\rm len}\right]_{\ell m} \sim \int d^2n\, T^{\rm len}(\hat{n})\, Y^*_{\ell m}(\hat{n}') \quad \text{(spin-0).} \tag{5}$$

This is different to the operation inverse to Eq. (3) (delensing), as we discuss now.

If the remapping of the sphere is invertible (which is always the case in the weak-lensing regime), we can perform a variable transformation to the unlensed coordinate $\hat{n}'$ and obtain

$$\left[\mathcal{D}^\dagger_{\boldsymbol{\alpha}} T^{\rm len}\right]_{\ell m} \sim \int d^2n' \left(\frac{T^{\rm len}(\hat{n})}{|A(\hat{n})|}\right) Y^*_{\ell m}(\hat{n}') \quad \text{(spin-0),} \tag{6}$$

where $|A(\hat{n})| = |d^2n'/d^2n|$ is the Jacobian (magnification) determinant of the lensing remapping.

Eq. (6) now has the form of a standard SHT of $(T/|A|)(\hat{n})$, where $\hat{n}$ is matched to $\hat{n}'$ as in Fig. 1 (hence, $T^{\rm len}/A$ is first delensed, and then mapped back to harmonic space). The choice of an isolatitude grid for $\hat{n}'$ provides one way to calculate this integral quickly with a standard backward SHT (Aurlien et al. (2022); Legrand & Carron (2022, 2023)) implemented $\mathcal{D}^\dagger$ in this way. However, significant overhead can remain with this method because it requires calculating the inverse deflection angles $\hat{n}(\hat{n}')$ on this grid. In a standard situation, the angles $\hat{n}'(\hat{n})$ are easily obtained from a standard SHT of the deflection field on an isolatitude grid sampling the observed coordinate $\hat{n}$, which does not provide $\hat{n}(\hat{n}')$ when the unlensed coordinate $\hat{n}'$ itself is sampled on such a grid. Moreover, usage of Eq. (6) requires the additional calculation of the magnification determinant, which has the cost of several forward SHTs (see Appendix B). The algorithm presented here bypasses this additional work and drops the requirement of an invertible deflection field.

For spin-weighted fields, the situation is almost identical. The harmonic space coefficients are split into a gradient ($G_{\ell m}$) and a curl term ($C_{\ell m}$), and the deflection operation is defined through

$${}_s\left[\mathcal{D}_{\boldsymbol{\alpha}}\, \mathcal{T}\right]_i = -e^{is\chi(\hat{n}_i)} \sum_{\ell=0}^{\ell_{\rm max}} \sum_{m=-\ell}^{\ell} \left(G_{\ell m} + iC_{\ell m}\right)\, {}_sY_{\ell m}(\hat{n}'_i) \tag{7}$$

---

[1] `www.cmb-s4.org`

for s > 0. The sign is not consistent with the spin-0 case to accommodate for the most prevalent conventions in the community. This creates a complex map of spin weight s, whose complex conjugate can be referred to with the subscript $-s$.

The adjoint takes this complex map as input and calculates the two sets of coefficients

$$ {}_{\pm s}\left[\mathcal{D}^\dagger_\alpha \mathcal{T}\right]_{\ell m} = \sum_i e^{\mp i s \chi(\hat{n}_i)}\ {}_{\pm s}\mathcal{T}(\hat{n}_i)\ {}_{\pm s}Y^*_{\ell m}(\hat{n}'_i), \tag{8} $$

which are decomposed as usual into gradient and curl modes,

$$ \begin{aligned} &-\frac{1}{2}\left({}_{s}\left[\mathcal{D}^\dagger_\alpha \mathcal{T}\right]_{\ell m} + (-1)^s\ {}_{-s}\left[\mathcal{D}^\dagger_\alpha \mathcal{T}\right]_{\ell m}\right) \quad (G_{\ell m}) \\ &-\frac{1}{2i}\left({}_{s}\left[\mathcal{D}^\dagger_\alpha \mathcal{T}\right]_{\ell m} - (-1)^s\ {}_{-s}\left[\mathcal{D}^\dagger_\alpha \mathcal{T}\right]_{\ell m}\right) \quad (C_{\ell m}). \end{aligned} \tag{9} $$

The phase χ is quite specific to CMB lensing applications and is absent in the general-purpose interpolation routines. The relation between the adjoint and inverse is unchanged from the case of spin-0 fields.

3. Implementation

Of the many implementations of the forward operation that were tested over the years, our approach is closest to that of Basak et al. (2008). The fundamentals are quite straightforward: The key point is that on the sphere parameterized by co-latitude θ and longitude φ, a band-limited function can be exactly written as a two-dimensional discrete Fourier series in θ and φ. While any function on the sphere is naturally periodic in the longitude coordinate, we must artificially extend the θ range to $[0, 2\pi)$ to obtain a doubly periodic function. We can then apply nonuniform fast Fourier transform (NUFFT) techniques (Barnett et al. 2019) to this function to perform the desired interpolation.

The main steps of the forward operations can be summarized as follows:

synthesis Synthesis of the map from $T^{\text{unl}}_{\ell m}$ on a rectangular equidistant grid with $n_\theta \geq \ell_{\max} + 2$ and $n_\varphi \geq 2\ell_{\max} + 2$ points with a standard SHT, with one isolatitude ring on each pole. Both dimensions are chosen in a way to make FFTs of lengths n_φ and $2n_\theta - 2$ efficient; in addition, n_φ must be an even number to enable the subsequent doubling step. For a standard CMB lensing application requiring $\ell_{\max} \sim 5000$, this corresponds to a sampling resolution close to 2 arcmin. The asymptotic complexity is $O(\ell^3_{\max})$.

doubling Doubling the θ-range from $[0, \pi]$ to $[0, 2\pi)$, by creating a $(2n_\theta - 2, n_\varphi)$ map, with the original map in the first half, and mirrored image $\theta \rightarrow 2\pi - \theta$ and $\varphi \rightarrow \varphi + \pi$ in the second half. In the case of odd spin values, the mirrored image also takes a minus sign.

FFT Going to Fourier space with the help of a standard forward 2D FFT. These coefficients by construction contain all information necessary to evaluate the map perfectly at any location. The asymptotic complexity is $O(\ell^2_{\max} \log \ell_{\max})$.

NUFFT Finally, from these Fourier coefficients and the coordinates $\hat{n}'$, interpolation proper, performed with a uniform-to-nonuniform (or type 2) NUFFT. The asymptotic complexity is $O(\ell^2_{\max} \log \ell_{\max}) + O(n_{\text{points}})$, where n_{points} is the number of points on the irregular grid.
This is the only step that incurs inaccuracies beyond those introduced by the finite precision of floating-point arithmetic. These inaccuracies are controlled by a user-specified parameter ϵ that was described in some detail in Arras et al. (2021).

The first three steps given here construct the 2D Fourier coefficients of the doubled-sphere representation of the same map from the spherical harmonic coefficients. Our approach to performing this (*synthesis-doubling-FFT*) starts with an SHT (*synthesis*). Consistent with the given asymptotic complexities, we find that this typically dominates the execution time overall. There are alternative approaches, however: By manipulating Wigner small-d matrices, we can build a more explicit representation of the relation between spherical and 2D Fourier harmonics that can be implemented via a well-behaved three-term recurrence formula without any Fourier transforms and with a similar $O(\ell^3_{\max})$ theoretical complexity. This is how Huffenberger & Wandelt (2010) implemented their SHTs, for example, and how Basak et al. (2008) implemented their CMB remapping. In place of a Legendre transform for each θ coordinate, a Legendre transform is required only at the equator, but one spin-weighted transform is required for each spin between 0 and $\ell_{\max}$. In our case, previous measurements (see, e.g., Section 2 of Galloway et al. 2022) showed that it was difficult to bring this recursion to speeds on CPUs that were comparable to the highly optimized standard Legendre transforms derived from the `libsharp` library (Reinecke & Seljebotn 2013) we are using. While we cannot exclude that there is some room for improvements provided the recursion can be optimized in a similar fashion (or, possibly, on GPUs), these are likely to be minor.

For the adjoint operation, the steps naturally go backwards (the individual complexities stay unchanged):

NUFFT From the input map and deflected coordinates, we perform a nonuniform-to-uniform (or type 1) NUFFT, resulting in the 2D FFT Fourier coefficients.

FFT We remap them to position space using a standard backward FFT on the same doubled Fourier sphere of shape $(2n_\theta - 2, n_\varphi)$ as for the forward operation.

undoubling The doubling of the Fourier sphere is undone by adding its mirror image to (or, for odd spins, subtracting from) the part $[0, \pi]$.

adjoint synthesis We perform a standard adjoint SHT on this new map of shape (n_θ, n_φ). This gives us the desired spherical harmonic coefficients.

Most of these steps are well established[2] in the astrophysical community or can be understood intuitively, with the possible exception of the NUFFT, whose purpose and structure we therefore outline (for more theoretical and technical details, see Potts et al. (2001); Greengard & Lee (2004); Barnett et al. (2019)). When given the Fourier coefficients of an n-dimensional function, it is trivial to obtain the function values on a regular grid in real space, in (almost) linear time. This is achieved using the fast Fourier transform, potentially after zero-padding the Fourier coefficients to increase the grid resolution. If the function values are required at irregularly spaced positions, however, this is not possible, and naive calculation of the Fourier series at each point is typically prohibitively slow in practice. One mathematically correct approach is to perform an FFT to a regular grid and then convolve these points with a sinc kernel centered on the desired

---

[2] Fast Fourier transforms are handled with code derived from the `pocketFFT` library (`https://gitlab.mpcdf.mpg.de/mtr/pocketfft`), which in turn is a descendant of FFTPACK (`https://netlib.org/fftpack/`), which was heavily modified for an improved accuracy and performance.
The standard ring-based spherical harmonic transforms are derived from the `libsharp` (Reinecke & Seljebotn 2013) library (`https://gitlab.mpcdf.mpg.de/mtr/libsharp`).

locations (sinc-interpolation). This is just as prohibitive, but approximate solutions can be found by choosing an alternative and more suitable convolution kernel. In particular, a kernel with compact support will be chosen (in our case, it takes the form of eq. (29) of Arras et al. 2021), and we divide the Fourier coefficients of the function by the Fourier coefficients of the kernel. This is the deconvolution step, and is followed by zero-padding the Fourier coefficients which increases the size of each dimension by a factor of roughly 1.2 to 2, and perform an FFT of the resulting array. At last comes the convolution step: for each irregularly spaced point, we perform a sum over its neighborhood and weight it by the kernel function.

Depending on the chosen kernel shape, zero-padding factor, and kernel support, it is possible to achieve accuracies close to machine precision; tuning the algorithm for higher tolerances is also possible and will improve the run time considerably. Good kernels compromise between the conflicting properties of being fast to evaluate and having small support and a quickly decaying Fourier transform. We used the variations of the kernel proposed in Barnett et al. (2019), as discussed in Arras et al. (2021).

Because all steps mentioned above are linear operations, the adjoint of the described NUFFT is obtained by executing the adjoint steps in reverse order, which is the nonuniform-to-uniform, or type 1, NUFFT.

Finally, for CMB-lensing applications, the unlensed angles $\hat{n}'$ can easily be gained from the representation

$$\hat{n}' = \cos\alpha\,\hat{n} + \frac{\sin\alpha}{\alpha}\left(\alpha_\theta \boldsymbol{e}_\theta + \alpha_\varphi \boldsymbol{e}_\varphi\right), \tag{10}$$

where at each point, the deflection vector is obtained from the spin-1 transform of the gradient (ϕ) and curl (Ω, if present) lensing potentials,

$$\begin{aligned}&\alpha_\theta(\hat{n}) + i\alpha_\varphi(\hat{n})\\ &= -\sum_{LM} \sqrt{L(L+1)}\,(\phi_{LM} + i\Omega_{LM})\ {}_1Y_{LM}(\hat{n}).\end{aligned} \tag{11}$$

Here, we follow the convention of using capital letters LM to refer to the spherical harmonic coefficients of the lensing potential.

4. Benchmark

This section discusses the computational cost, the scaling with threading, and memory usage of the forward and adjoint operation, and it compares our work to a few publicly available implementations of the forward operation, such as `LensPix`[3] (Lewis 2005), `lenspyx`[4] (Aghanim et al. 2020), `pixell`[5] (Naess et al. 2021), and `Flints`[6] (Lavaux & Wandelt 2010).

A typical task in the CMB lensing context is to compute lensed CMB spherical harmonic coefficients ($a^{\rm len}_{\ell m}$) starting from unlensed ones ($a^{\rm unl}_{\ell m}$). This can be achieved by first performing a forward remapping onto a suitable pixelization of the sphere, and computing the spherical harmonics coefficients by a standard backward SHT. The adjoint of this entire operation is first built out of a forward SHT, followed by the adjoint remapping. Key parameters impacting the execution time are the band limit $\ell^{\rm unl}_{\rm max}$ of $a^{\rm unl}_{\ell m}$ and the requested maximum multipole $\ell^{\rm len}_{\rm max}$ of the lensed CMB. In the applications that motivated this work, $\ell^{\rm len}_{\rm max}$ is at most the multipole above which the information on the lensing signal becomes negligible. For example, *Planck* reconstructions (Aghanim et al. 2020; Carron et al. 2022) used $\ell^{\rm len}_{\rm max} = 2048$, and the recent ACT results (Madhavacheril et al. 2023) $\ell^{\rm len}_{\rm max} = 3000$. For a very deep future polarization experiment such as the CMB-S4 deep survey (Abazajian et al. 2016), this is closer to $\ell^{\rm len}_{\rm max} = 4096$, a number we use often as reference in this section. Then $\ell^{\rm unl}_{\rm max}$ must typically be taken slightly higher than $\ell^{\rm len}_{\rm max}$ in order to account for the mode mixing by lensing.

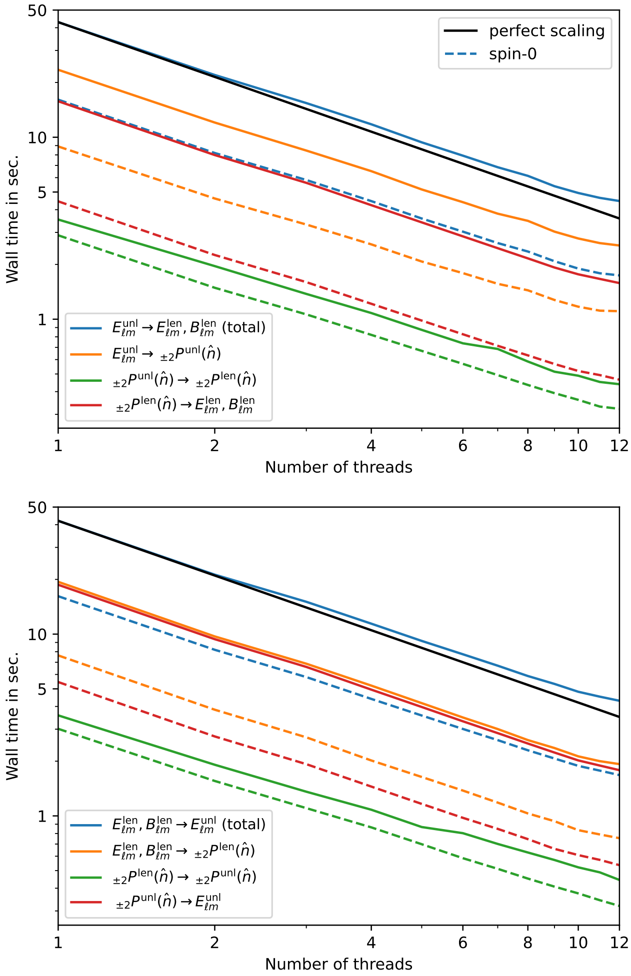

Fig. 2. Execution time of different transforms as a function of number of threads used for the calculation. *Upper panel:* Scaling of a single forward-lensing execution time, producing lensed spherical harmonic coefficients from their unlensed counterparts. The solid blue line shows the total time of producing $a^{\rm len}_{\ell m}$ from $a^{\rm unl}_{\ell m}$ in the polarized spin-2 case, and the solid orange, green, and red curves show the synthesis, interpolation proper and final backward SHT contributions. The black line shows a perfect scaling with the inverse thread number for comparison and the scaling of each and every operation is almost perfect. Dashed lines show the spin-0 results, for which harmonic transforms are substantially faster. The forward-lensing operation is built out of remapping on a pixelized sphere (orange and green, the latter being the remapping step, properly speaking), and sending it back to harmonic space with a standard spin-2 spherical harmonic transform (red). See the text for the precise specifications. These curves do not include the cost of calculating the deflected angles from the deflection field spherical harmonic coefficients, which is comparable to that of a spin-1 forward spherical harmonic transform. *Lower panel:* Corresponding results for the adjoint operation (note that while the adjoint operation is closely related to delensing, this is not the operation inverse to forward lensing; see text). The dashed lines show the corresponding results for the spin-0 case.

---

[3] `https://cosmologist.info/lenspix`
[4] `https://github.com/carronj/lenspyx`
[5] `https://pixell.readthedocs.io`
[6] `https://bitbucket.org/glavaux/flints`

When lensed spherical harmonic coefficients are built in this way, another important parameter is the intermediate grid pixelization. Because the lensed CMB is not band limited, there is no exact quadrature rule, and this choice can strongly impact the accuracy of the recovered $a^{\mathrm{len}}_{\ell m}$ at high multipoles. We show in Fig. 2 the execution time of these tasks for the forward (upper panel) and adjoint (lower panel) cases as a function of the number of threads (strong scaling). We picked $\ell^{\mathrm{unl}}_{\max} = 5120$, $\ell^{\mathrm{len}}_{\max} = 4096$, and a longitude-thinned Gauss-Legendre grid with 6144 rings as intermediate grid ($\sim 5 \cdot 10^7$ pixels). This is a conservative configuration that in our experience allows a robust and accurate lensing reconstruction for very deep stage IV CMB observations. This was performed on a CSCS Piz Daint XC50 compute node[7], with 12 physical cores. Results on more recent processors can sometimes be up to twice as fast in our experience. The top panel shows the result for the synthesis operation, while the lower panels shows the adjoint of it, and we see that each operation scales almost perfectly. We have excluded the cost of building the undeflected angles from the timing, which we briefly discuss below, as is suitable in the context of optimal lensing reconstruction, where many maps are deflected with the same set of angles. The accuracy of the interpolation was set to 10^{-5}, which value is good enough for most purposes, as also discussed in more detail below. The interpolation is very efficient. Only a minor part of the total time is dominated by the pair of SHTs that is involved.

We now discuss the accuracy and memory usage of our implementation. Our code uses simple heuristics to allow the user to choose a target-relative accuracy, $\epsilon_{\mathrm{target}}$, of at least 10^{-13}. For the forward operation, schematically $a^{\mathrm{unl}}_{lm} \rightarrow \tilde{m}$, we calculated a map-level effective relative accuracy, ϵ_{eff}, as follows: We took the difference between the true and estimated lensed maps, $\tilde{m}^{\mathrm{true}}$, $\tilde{m}^{\mathrm{est}}$, respectively, and normalized by the total power of the true lensed maps,

$$\epsilon_{\mathrm{eff}} = \frac{1}{P^{\mathrm{true}}} \sqrt{\sum_i \left(\tilde{m}^{\mathrm{true}}_i - \tilde{m}^{\mathrm{est}}_i\right)^2}. \tag{12}$$

Here, $P^{\mathrm{true}} = \sqrt{\sum_i (m^{\mathrm{true}}_i)^2}$, and the sum index i run over pixels as well as the map components for nonzero spin (Q and U in polarization). The maps $\tilde{m}^{\mathrm{true}}$ were determined by a brute-force approach: We remapped the unlensed map to the desired grid by calculating eq. (3) explicitly. As this can be fairly expensive, we calculated 10^5 exactly lensed pixels at most for a few numbers of isolatitude rings close to the equator.

The right panel in Figure 3 shows the computational cost of the lensing routine for spin-2 fields at $\ell^{\mathrm{unl}}_{\max} = 4096$, using four CPUs, mapping onto a Healpix pixelization with $N_{\mathrm{side}} = 2048$, as a function of ϵ_{eff} (bottom axis) and $\epsilon_{\mathrm{target}}$ (top axis). The effective accuracy is typically slightly higher than requested, except in the vicinity of 10^{-13}. This is not a problem of the NUFFT interpolation, however, but rather of the true SHT calculations themselves, which lose several digits in accuracy at this band limit. The diagram shows the split of the total computational cost into its most relevant steps described earlier. The SHT steps for the angle and doubled Fourier sphere always dominate the cost in this configuration, and the choice of accuracy has a fairly minor impact. If relevant, the SHTs can be performed with a gradient-only setting for nonzero spin fields (an implementation specific to the case of a vanishing curl component), further reducing the computational cost of these tasks by about 25%. The scale of the peak memory usage is indicated on the right y-axis of the right panel, and is shown as connected green square data points. The timings, memory, and accuracy calculations were made for a set of five analyses, and their standard deviation is negligible.

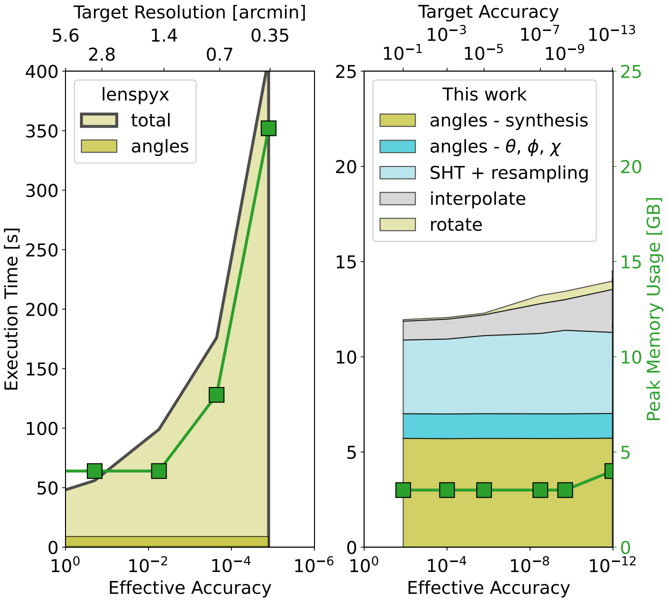

Fig. 3. Full-sky forward operation execution times for spin-2 fields of our implementation compared with `lenspyx`, which uses a popular bicubic spline interpolation technique. The top panels show `lenspyx` (left) and the new implementation (right) as a function of the effective relative accuracy for $\ell_{\max} = 4096$, producing lensed CMB maps in a Healpix geometry with $N_{\mathrm{side}} = 2048$. The effective accuracy is calculated for a few isolatitude rings close to the equator. All calculations were made with four CPUs. We indicate the execution time (shaded areas, left y-axes), peak memory usage (green data points, right y-axis), and the accuracy parameters (target resolution, target accuracy) at the top x-axes. For `lenspyx`, we divided the full sky map into two bands to reduce the memory consumption. The differently colored areas show the main steps of the operation. The execution time and memory consumption both grows rapidly for `lenspyx` because it requires high-resolution grids for an accurate interpolation.

The left panel shows the corresponding results for a code using a popular interpolation scheme, `lenspyx`. This code (just as `LensPix` and `pixell`) uses bicubic spline interpolation on a intermediate grid obtained by cylindrical projection. The resolution of this grid then essentially sets the accuracy of the result and memory usage. While these interpolation schemes are perfectly fine for the purpose of producing a set of lensed CMB maps, they have strong limitations for more intensive tasks, or when higher accuracies are imperative. For low target resolutions of about 1.4 arcmin, the new implementation speeds up the execution time approximately 7 times for a similarly effective accuracy of about $\epsilon_{\mathrm{eff}} \approx 10^{-2}$. This increases to a speed-up of 30 times for an effective accuracy of 10^{-5}. Higher accuracies are almost not manageable for `lenspyx` in both time and memory, whereas the new implementation can easily reach accuracies as low as 10^{-12}.

Figure 4 shows the execution time and peak memory usage of the new implementation for various N_{side} and $\ell^{\mathrm{unl}}_{\max}$ configurations and a target accuracy of $\epsilon_{\mathrm{target}} = 10^{-13}$. The memory usage for the most challenging benchmarks is still below 64 GB, and we would like to note that the memory consumption is only slightly larger than the memory needed for the three maps (the deflection angles) for the respective N_{side}.

---

[7] https://www.cscs.ch/computers/piz-daint/

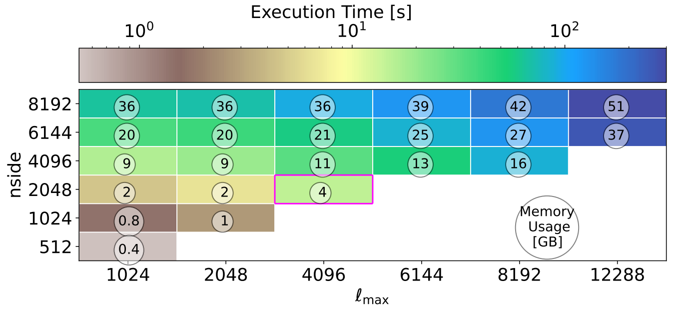

Fig. 4. Benchmark of the execution time and peak memory consumption for a wide range of Healpix resolution $N_{\rm side}$ and $\ell_{\rm max}$, for a target accuracy of 10^{-13}, and for our new implementation. The cell highlighted in magenta is the configuration we used for the accuracy benchmark in Fig. 3. The execution time ranges from 0.55 seconds ($N_{\rm side} = 512, \ell_{\rm max} = 1024$) to 336 seconds ($N_{\rm side} = 8192, \ell_{\rm max} = 12288$), while the memory consumption increases from 0.4 GB to 51 GB. The memory consumption closely follows the memory needed to store the Healpix maps of that particular size.

We now compare our results to other implementations that benchmark the spin-0 forward operation. The effective accuracy was calculated as in the above case, replacing the data by the spin-0 true and estimated temperature maps. For `LensPix` and `Flints`, we estimated the costs by using the natively C-implemented lensing routine, and with `OpenMP` support and four CPUs. We measured the tasks directly related to lensing, such as the calculation of deflection angles, the deflection, the interpolation, and rotation. To farily compare all implementations, we did not cache any calculations in `Flints`, even though they are available. For `pixell`, we used the full-sky lensing routine implemented in Python, and measured the costs within Python.

Almost all implementations provide an accuracy parameter in form of the intermediate grid pixel resolution. We chose them to provide approximately comparable results.

For `Flints`, `LensPix`, and `lenspyx`, we generated input CMB and lensing potential realizations and the true and estimated lensed maps on a Healpix grid. For `pixell`, the full-sky lensing routine was evaluated onto a CAR geometry, and we calculated the effective accuracy using our own CAR geometry to calculate the true lensed map.

We used $\ell_{\rm max}^{\rm unl} = 4096$ and an HPC machine with four CPUs and 60 GB of memory. To reduce memory usage, the data for `LensPix` were split into sets and were calculated individually.

Table 1 summarizes the computational costs and memory usage for the pixel resolution and effective accuracy. `LensPix` and `Flints` are well-established algorithms and perform the task in reasonable time and for very reasonable memory usage. `LensPix` allows us to control the intermediate grid pixel resolution, whereas with `Flints`, this can be indirectly controlled by the choice of $N_{\rm side}$. If needed, `Flints` execution time could be reduced by caching some of the calculations, making them available for repeated runs. `pixell` supports splitting of the full sky data into bands, which can be used to reduce the memory usage, if needed. The algorithm in this paper, shown in the bottom row, is vastly more efficient both in execution time and memory consumption, even more so in the high-accuracy regime.

Finally, we comment more specifically on the relevance of this code for a maximum a posteriori lensing reconstruction (Hirata & Seljak 2003b; Carron & Lewis 2017). In these methods, the reconstruction proceeds by finding an approximate solution to the Wiener-filtered delensed CMB using a conjugate gradient solver at each iteration that involves applying the forward and adjoint operation. As mentioned earlier, this typically involves several hundred such operations. The benefits for a reconstruction strategy like this are twofold. Faster operations directly speed up the lensing reconstruction, and accurate operations can prevent the conjugate gradient solver from building up errors or showing instability.

We saw this explicitly by testing the full-sky lensing reconstruction with CMB-S4-like settings ($(\ell_{\rm max}^{\rm unl}, \ell_{\rm max}^{\rm len}) = (4500, 4000)$, a polarization noise level of $\sqrt{2}$ arcmin, and 4 CPUs), reconstructing the lensing potential using polarization-only maps, using both `lenspyx`, and the new implementation for the forward and adjoint operations. With `lenspyx`, we observed that a target resolution of about 1.0 arcmin is beneficial for a stable reconstruction of the very largest modes of the lensing potential in experimental configurations like these, resulting in execution times for a single adjoint operation of about 3 minutes. This is significantly longer than for the new implementation, for which we find execution times of only 8 seconds for better accuracy. Analogously, we find for the execution time of a typical full-iteration step of a lensing reconstruction $\sim$ 25 minutes, or $\sim$ 4 minutes, respectively, and the speed-up is even larger for lower $\ell_{\rm max}$ analyses ($\sim$ 24 minutes vs. $\sim$ 2 for $(\ell_{\rm max}^{\rm unl}, \ell_{\rm max}^{\rm len}) = (3200, 3000)$). It is worth mentioning that with `lenspyx` at this resolution, outliers can occasionally occur. They can take more than one hour or do not converge at all. This is directly due to the lower accuracy, in particular very close to the poles, where the bicubic spline method is less accurate. This has not been observed with the new implementation, whose error map is uniform across the entire sphere. We observed something similar for temperature-only reconstruction, where an effective relative accuracy of 10^{-7} generally appears to be required for a successful convergence of the lensing-map search on the very largest scales. This accuracy is accessible only with the implementation discussed here.

5. Conclusion

We have described an optimized implementation of the spherical transform pair of an arbitrary spin-weight that can be used on any pixelization of the sphere, such as regular grids distorted by CMB lensing. A C++ implementation and comprehensive Python front-end is available, together with the low-level algorithms (FFT, NUFFT, and SHT), under the terms of the GNU General Public License and named `DUCC`[8]. The code is written with the goal of portability and does not depend on external libraries. It supports multithreading via the C++ standard threading functionality and will make use of CPU vector instructions on `x86` and `ARM` hardware, if the compiler supports the respective language extensions. The Python interface is kept deliberately general and flexible to allow use in the widest possible range of scientific applications. As a consequence, parts of the interface are somewhat complex and are perhaps best used by higher-level more application-specific packages to hide unnecessary details from the end user.

For users interested in applications specific to CMB lensing, the `lenspyx`[9] Python package has been updated in this spirit to include these developments and provide additional wrappers to these routines. This results in an improvent of some orders of magnitude in execution time and accuracy over currently publicly available tools.

---

[8] https://gitlab.mpcdf.mpg.de/mtr/ducc

[9] https://github.com/carronj/lenspyx

Table 1. spin-0, $\ell_{\max} = 4096$, and $N_{\text{side}} = 2048$ (if applicable) forward-operation execution time and memory consumption for different implementations on the full sky, using 4 CPUs. Additionally, we split the full sky map for `pixell` and `lenspyx` into two bands to reduce memory consumption.

implementation	grid resolution	effective accuracy	computation time	memory peak usage
`pixell`	1.0 arcmin	$6 \cdot 10^{-3}$	3 min 30 sec	20 GB
`LensPix`	0.85 arcmin	$3 \cdot 10^{-4}$	9 min 30 sec	3.4 GB
`Flints`	N.A.	$3 \cdot 10^{-3}$	4 min	6.5 GB
`lenspyx`	0.70 arcmin	$1 \cdot 10^{-4}$	2 min	8 GB
	0.35 arcmin	$1 \cdot 10^{-5}$	5 min	22 GB
This work	N.A.	$2 \cdot 10^{-6}$	9 sec	2.8 GB
	N.A.	$4 \cdot 10^{-12}$	10 sec	2.8 GB

Acknowledgements. We thank Guilhem Lavaux, Antony Lewis, and Mathew Madhavacheril for discussions on `Flints`, `LensPix` and `pixell` respectively. JC and SB acknowledge support from a SNSF Eccellenza Professorial Fellowship (No. 186879). This work was supported by a grant from the Swiss National Supercomputing Centre (CSCS) under project ID s1203.

References

Abazajian, K. N. et al. 2016 [`arXiv:1610.02743`]
Adachi, S. et al. 2020, Phys. Rev. Lett., 124, 131301
Ade, P. et al. 2019, JCAP, 02, 056
Ade, P. A. R. et al. 2021, Phys. Rev. Lett., 127, 151301
Aghanim, N. et al. 2020, Astron. Astrophys., 641, A8
Allison, R., Caucal, P., Calabrese, E., Dunkley, J., & Louis, T. 2015, Phys. Rev. D, 92, 123535
Arras, P., Reinecke, M., Westermann, R., & Enßlin, T. A. 2021, Astron. Astrophys., 646, A58
Aurlien, R. et al. 2022 [`arXiv:2211.14342`]
Barnett, A. H., Magland, J., & af Klinteberg, L. 2019, SIAM Journal on Scientific Computing, 41, C479
Basak, S., Prunet, S., & Benabed, K. 2008, in 12th Marcel Grossmann Meeting on General Relativity, 2213–2215
Carron, J. & Lewis, A. 2017, Phys. Rev. D, 96, 063510
Carron, J., Mirmelstein, M., & Lewis, A. 2022, JCAP, 09, 039
Challinor, A. & Chon, G. 2002, Phys. Rev. D, 66, 127301
Challinor, A. et al. 2018, JCAP, 04, 018
Galloway, M., Reinecke, M., Andersen, K. J., et al. 2022, arXiv e-prints, arXiv:2201.03478
Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759
Greengard, L. & Lee, J.-Y. 2004, SIAM Review, 46, 443
Hall, A. C. & Challinor, A. 2012, Mon. Not. Roy. Astron. Soc., 425, 1170
Hirata, C. M. & Seljak, U. 2003a, Phys. Rev., D67, 043001
Hirata, C. M. & Seljak, U. 2003b, Phys. Rev., D68, 083002
Hu, W. & Okamoto, T. 2002, Astrophys. J., 574, 566
Huffenberger, K. M. & Wandelt, B. D. 2010, Astrophys. J. Suppl., 189, 255
Kesden, M., Cooray, A., & Kamionkowski, M. 2002, Phys. Rev. Lett., 89, 011304
Knox, L. & Song, Y.-S. 2002, Phys. Rev. Lett., 89, 011303
Lavaux, G. & Wandelt, B. D. 2010, Astrophys. J. Suppl., 191, 32
Legrand, L. & Carron, J. 2022, Phys. Rev. D, 105, 123519
Legrand, L. & Carron, J. 2023 [`arXiv:2304.02584`]
Lesgourgues, J. & Pastor, S. 2006, Phys. Rept., 429, 307
Lewis, A. 2005, Phys. Rev., D71, 083008
Lewis, A. & Challinor, A. 2006, Phys. Rept., 429, 1
Lewis, A., Challinor, A., & Turok, N. 2002, Phys. Rev. D, 65, 023505
Madhavacheril, M. S. et al. 2023 [`arXiv:2304.05203`]
Maniyar, A. S., Ali-Haïmoud, Y., Carron, J., Lewis, A., & Madhavacheril, M. S. 2021, Phys. Rev. D, 103, 083524
Millea, M., Anderes, E., & Wandelt, B. D. 2019, Phys. Rev. D, 100, 023509
Millea, M. & Seljak, U. 2022, Phys. Rev. D, 105, 103531
Millea, M. et al. 2021, Astrophys. J., 922, 259
Naess, S., Madhavacheril, M., & Hasselfield, M. 2021, Pixell: Rectangular pixel map manipulation and harmonic analysis library, Astrophysics Source Code Library, record ascl:2102.003
Okamoto, T. & Hu, W. 2003, Phys. Rev., D67, 083002
Potts, D., Steidl, G., & Tasche, M. 2001, Fast Fourier Transforms for Nonequispaced Data: A Tutorial, ed. J. J. Benedetto & P. J. S. G. Ferreira (Boston, MA: Birkhäuser Boston), 247–270
Qu, F. J. et al. 2023 [`arXiv:2304.05202`]
Reinecke, M. & Seljebotn, D. S. 2013, Astron. Astrophys., 554, A112
Tristram, M. et al. 2022, Phys. Rev. D, 105, 083524
Zaldarriaga, M. & Seljak, U. 1998, Phys. Rev. D, 58, 023003

Appendix A: Map analysis with a minimal number of rings

In the text above, we regularly use equiangular spherical grids with $\ell_{\max} + 2$ rings (first and last ring located at the poles) to represent functions with a band limit of $\ell_{\max}$ (inclusive). This number of rings may sound insufficient to fully represent a function with the given band limit because the minimum number of rings required for an accurate map analysis via a quadrature rule on this layout is $2\ell_{\max} + 2$ (Clenshaw-Curtis quadrature).

However, we can again make use of the double Fourier sphere technique that was introduced in Sect. 3, that is, we can follow a meridian from the north pole to the south pole, and then back again on the opposite side. A full meridian like this has $2\ell_{\max} + 2$ points in total. Because we assumed the function on the sphere to be band limited, the θ-dependent function along each of these full meridians is the sum of associated Legendre polynomials of degrees up to $\ell_{\max}$ in $\cos(\theta)$. In other words, it can be expressed as a Fourier series $\sum_{k=0}^{\ell_{\max}} A_k \cos(k\theta)$. This function in turn is completely determined by $2\ell_{\max} + 1$ equidistant samples in the range $\theta = [0; 2\pi)$, that is, one less than we actually have. The same is true in azimuthal direction, where we also have at least $2\ell_{\max} + 2$ pixels on each ring.

As a consequence, the actual function value at any θ and φ can be obtained using a combination of fast Fourier transforms and phase-shifting factors, or (in an approximate fashion) via NUFFT. One way of extracting the spherical harmonic coefficients from this map would therefore be by first computing the function values at a shifted set of isolatitude rings located exactly between the existing ones, increasing the number of rings to $2\ell_{\max} + 3$. This is then followed by applying the appropriate Clenshaw-Curtis quadrature weights to the full set of rings. Finally, running an adjoint spherical harmonic synthesis operation on the full set of weighted rings gives the desired result.

It is even possible to shift the newly generated rings back to the original positions after weighting, which again reduces the number of rings in the adjoint synthesis operation to $\ell_{\max} + 2$. This speeds up the SHT considerably.

Appendix B: Adjoint and inverse lensing

In the most general situation, invertibility of the deflection field is not necessarily always achieved at all points in a likelihood search under nonideal conditions, where lensing estimators can react strongly to signatures of anisotropies unrelated to lensing. When the lensing deflection is invertible, the inverse lensing operation can still be useful. To this end, we can use the same adjoint operation $\mathcal{D}^\dagger$, but with input $T^{\rm len}(\hat{n}) \cdot |A|(\hat{n})$ instead of $T^{\rm len}(\hat{n})$. Eq. (6) in the main text shows that the result then is $T^{\rm unl}_{\ell m}$.

The magnification determinant may be obtained as follows: With ${}_1\alpha = \alpha_\theta + i\alpha_\varphi$ as in Eq. (11), let the convergence (κ), field rotation (ω) and shears (γ) be

$$
\begin{aligned}
\kappa + i\omega &= \frac{1}{2}\tilde{\eth}\,{}_1\alpha \\
\gamma_1 + i\gamma_2 &= \frac{1}{2}\eth\,{}_1\alpha,
\end{aligned}
\qquad \text{(B.1)}
$$

where $\eth$ and $\tilde{\eth}$ are the spin raising and lowering operators (see the first appendix of Lewis et al. (2002) for a discussion in the context of the CMB). It holds

$$
\begin{aligned}
|A|(\hat{n}) = {} & \frac{\sin\alpha}{\alpha}\left((1-\kappa)^2 + \omega^2 - \gamma^2\right) \\
& + \left(\cos\alpha - \frac{\sin\alpha}{\alpha}\right)(1 - \kappa - \cos(2\beta)\gamma_1 - \sin(2\beta)\gamma_2)\,.
\end{aligned}
\qquad \text{(B.2)}
$$

All of these quantities can be computed from the harmonic coefficients of the deflection field with the help of spin-weighted spherical harmonic transforms.

Informally, ignoring technical issues of band limits, quadrature weights, and so on, we may write

$$
\left[\mathcal{D}_{\boldsymbol{\alpha}}\mathcal{D}^\dagger_{\boldsymbol{\alpha}}\right](\hat{n}_i, \hat{n}_j) \sim \frac{\delta^D(\hat{n}_i - \hat{n}_j)}{|A(\hat{n}_i)|}, \quad (\boldsymbol{\alpha} \text{ invertible}), \qquad \text{(B.3)}
$$

where δ^D is the Dirac delta. Similarly, the operator $\mathcal{D}^\dagger\mathcal{D}$ produces the spherical harmonic coefficients of $\left(T^{\rm unl}/|A|\right)(\hat{n})$ from those of $T^{\rm unl}$.

2.3 `cunuSHT`

This section contains the `lenspyx` paper as approved on *RASTI*.

`cunuSHT`: GPU Accelerated Spherical Harmonic Transforms on Arbitrary Pixelizations

Sebastian Belkner,[1,2] Adriaan J. Duivenvoorden,[2] Julien Carron,[1] Nathanael Schaeffer,[3] Martin Reinecke[4]
[1] *Université de Genève, Département de Physique Théorique et CAP, 24 Quai Ansermet, CH-1211 Genève 4, Switzerland*
[2] *Center for Computational Astrophysics, Flatiron Institute, 162 5th Avenue, New York, NY 10010, USA*
[3] *ISTerre, Université de Grenoble 1, CNRS, F-38041 Grenoble, France*
[4] *Max-Planck Institut für Astrophysik, Karl-Schwarzschild-Str. 1, 85748 Garching, Germany*

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

We present `cunuSHT`, a general-purpose Python package that wraps a highly efficient CUDA implementation of the nonuniform spin-0 spherical harmonic transform. The method is applicable to arbitrary pixelization schemes, including schemes constructed from equally-spaced iso-latitude rings as well as completely nonuniform ones. The algorithm has an asymptotic scaling of $O(\ell_{\max}^3)$ for maximum multipole $\ell_{\max}$ and can be made to achieve machine precision accuracy, considering band-limited transforms for which $N \approx \ell_{\max}^2$ (where N is the number of pixels in the map). While `cunuSHT` is developed for applications in cosmology in mind, it is applicable to various other interpolation problems on the sphere. We outperform the fastest available CPU algorithm at problem sizes $\ell_{\max} \sim 4 \cdot 10^2$ and larger. The speed-up increases with the problem size and reaches a factor of up to 5 for problems with a nonuniform pixelization and $\ell_{\max} > 4 \cdot 10^3$ when comparing a single modern GPU to a modern 32-core CPU. This performance is achieved by utilizing the double Fourier sphere method in combination with the nonuniform fast Fourier transform and by avoiding transfers between the host and device. For scenarios without GPU availability, `cunuSHT` wraps existing CPU libraries. `cunuSHT` is publicly available and includes tests, documentation, and demonstrations.

keywords: Nonuniform Spherical Harmonic Transform, Nonuniform Fast Fourier transform, Cosmic Microwave Background Weak Lensing, CUDA

1 INTRODUCTION

Spherical harmonic transforms (SHTs) are a key ingredient in signal processing for data sets on the 2-sphere. They are extensively used in active research fields such as cosmology (both in studying the cosmic microwave background (CMB) (Hivon et al. 2002; Aghanim et al. 2020; Ade et al. 2021, 2018; Qu et al. 2024; Hanany et al. 2019) and large-scale structure of the Universe Scharf & Lahav (1993); Hikage et al. (2011)), gravitational waves Deppe et al. (2024), meteorology Wedi et al. (2013), solar physics Brun & Rempel (2009), or solving partial differential equations on the sphere Browning et al. (1989).

Modern applications of harmonic analysis on the sphere routinely require the evaluation of the SHT up to to a maximum multipole $\ell_{\max} = O(10^4)$ onto the $N \sim \ell_{\max}^2$ points that are required to represent a signal of harmonic band-limit $\ell_{\max}$ on the sphere. A direct evaluation of the SHT scales as $O(\ell_{\max}^4)$, which is intractable for large problems. Reducing the computational complexity is thus crucial. One standard optimization is achieved by pixelizing the sphere into rings of constant latitude with equi-angular spaced samples, reducing the problem to $O(\ell_{\max}^3)$. We will refer to this setup as the ring spherical harmonic transform (rSHT).[1] See Reinecke & Seljebotn (2013); Schaeffer (2012) for modern implementations of the rSHT. For cases where the transform has to be evaluated on irregularly sampled pixels, ring sampling is not possible. We will refer to this more general setup as the nonuniform spherical harmonic transform (nuSHT). Notable applications of the nuSHT are CMB weak lensing Lewis & Challinor (2006), ray tracing Fabbian et al. (2018); Ferlito et al. (2024), or fields where the pixelization changes over time. It should be noted that due to their computational complexity, the rSHT

---

[1] There exist algorithms for the rSHT that asymptotically scale as $O(\ell_{\max}^2 \log^2(\ell_{\max}))$ Driscoll & Healy (1994); Potts et al. (1998a,b); Keiner et al. (2009); Seljebotn (2012) or $O(\ell_{\max}^2 \log^2(\ell_{\max})/\log\log(\ell_{\max}))$ Hale & Townsend (2015). While there have been improvements over the last few years Slevinsky (2019); Yin et al. (2019) such methods require high memory usage and significant pre-computations. A general purpose implementation that is competitive with the $O(\ell_{\max}^3)$ counterpart has yet to be published.

and nuSHT often become the bottleneck in iterative algorithms that repeatedly apply the transforms Wandelt et al. (2004); Carron & Lewis (2017).

In the field of cosmology, nuSHTs are routinely solved via an rSHT and subsequent interpolation to the nonuniform points using bicubic splines Lewis (2005) or a Taylor series expansion Næss & Louis (2013). These methods scale as $O(\ell_{\max}^3)$ but only reach relatively low accuracy. A different $O(\ell_{\max}^3)$ nuSHT method, proposed by Basak et al. (2008), makes use of the double Fourier sphere (DFS) method (see Boer & Steinberg (1975); Merilees (1973) for the backward or type 2 transform, McEwen (2011) for the forward or type 1, and Huffenberger & Wandelt (2010a); McEwen & Wiaux (2011) for both) in combination with the nonuniform fast Fourier transform (nuFFT) to achieve an accurate nuSHT algorithm. Several other implementations of this setup exist Townsend et al. (2016); Keiner et al. (2009); Potts & Van Buggenhout (2017). Recently, the `lenspyx`[2] and `DUCC`[3] libraries Reinecke & Seljebotn (2013); Reinecke et al. (2023) implemented a highly efficient and machine precision accurate implementation of the nuSHT based on the DFS method.

In recent years, there has been a dramatic increase in the availability and usage of graphics processing units (GPUs) dedicated to scientific computing. This development drives the need for GPU based codes and provides an opportunity for increased performance of existing methods. GPUs are optimized for single-instruction multiple-data (SIMD) applications and provide multi-threading well beyond what is achievable with CPUs. Highly parallelizable algorithms can thus greatly benefit from the GPU architecture. In principle, the evaluation of the SHT allows for a large amount of parallel computation, making it a natural target for a GPU implementation. The work presented in Hupca et al. (2012); Szydlarski et al. (2013) was one of the first that explored the use of GPUs for SHTs in the context of cosmology.

Robust and efficient rSHT GPU implementations have been developed in recent years. Notable examples are `SHTns` Schaeffer (2012), supporting the spin-0 and 1 rSHTs, and `S2HAT`[4] Szydlarski et al. (2013), which supports spin-n transforms. `S2FFT` Price & McEwen (2024) is a recent `JAX` implementation of the spin-n rSHT that provides differentiable transforms. A first implementation of the nuSHT in a cosmological context on GPUs was presented in Baleato Lizancos & White (2024).

We present `cunuSHT`, a CUDA accelerated nuSHT algorithm on the GPU. This is, to our knowledge, the first publicly available nuSHT GPU algorithm that can reach machine precision accuracy and achieves significant speed-up compared to the fastest CPU algorithms. We achieve this by carefully combining existing robust and efficient GPU implementations of the rSHT and nuFFT algorithm. `cunuSHT` does not require memory allocation or calculation on the host, which allows it to be incorporated in GPU-based software.[5]

The remainder of the paper is organized as follows. In Section 2 we introduce notation and definitions, and present in qualitative terms our implementation. Section 3 discusses the implementation on the GPU. Section 4 shows benchmarks and results. We conclude in Section 5. A series of appendices collect further details.

2 NONUNIFORM SPHERICAL HARMONIC TRANSFORM

We introduce our notation and conventions in 2.1, and define the spherical harmonic transform operations that we implement in this paper. We describe the double Fourier sphere method in 2.2.

2.1 Definition and properties

The (spin-0) spherical harmonic functions $Y_\ell^m(\theta,\phi)$, with quantum numbers ℓ and m, with $-\ell \leq m \leq \ell$, are given by,

$$Y_\ell^m(\theta,\phi) = P_\ell^m(\theta) e^{im\phi}, \tag{1}$$

where $P_\ell^m(\theta)$ are the associated Legendre polynomials. A "general" (or "nonuniform") spherical harmonic transform (nuSHT) is a linear transformation between a set of spherical harmonic coefficients and field values defined at arbitrary locations on the sphere. We distinguish two types of transforms, with nomenclature inspired by nonuniform Fourier transform literature[6] Barnett et al. (2019):

- *Type 1* (also "adjoint nuSHT", the adjoint operation to type 2 below): given as input a set of N values f_i, and N locations (θ_i,ϕ_i), desired are the coefficients $c_{\ell m}$ defined by,

$$c_{\ell m} = \sum_{i=1}^{N} f_i \, Y_\ell^{\dagger m}(\theta_i,\phi_i)\,, \tag{2}$$

for ℓ up to some band-limit $\ell_{\max}$. We want the result to match a target accuracy ϵ requested by the user.

- *Type 2*: given as input a set of harmonic coefficients $c_{\ell m}$ up to some band-limit $\ell_{\max}$, and a set of N locations (θ_i,ϕ_i), desired are the field values,

$$f_i = \sum_{\ell=0}^{\ell_{\max}} \sum_{m=-\ell}^{\ell} c_{\ell m} Y_\ell^m(\theta_i,\phi_i)\,, \tag{3}$$

again respecting a target accuracy ϵ as requested by the user.

In matrix notation, type 2 may be written as,

$$\mathbf{f} = \mathbf{Y}\mathbf{c}, \tag{4}$$

where the vector $\mathbf{c}$ collects the harmonic coefficients, the vector $\mathbf{f}$ the

---

[2] https://github.com/carronj/lenspyx

[3] https://pypi.org/project/ducc0/

[4] https://apc.u-paris.fr/APC_CS/Recherche/Adamis/MIDAS09/software/s2hat/s2hat.html

[5] The current implementation does not support automatic differentiation.

[6] The `DUCC` package uses the names `adjoint_synthesis_general` and `synthesis_general` for type 1 and type 2.

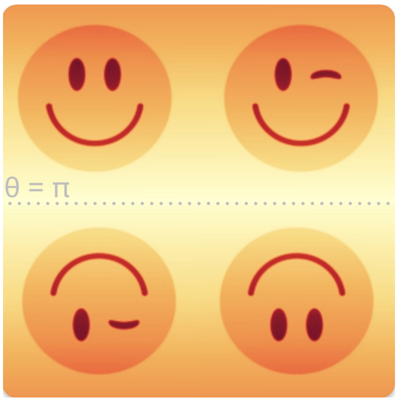

Figure 1. Illustration of the doubling step in the DFS method. The upper half shows a spherical map calculated on a rectangular grid and is mirrored along the $\theta = \pi$ axis. The mirrored image is split in half across the ϕ-direction in the center, and swapped. The result is a function on the torus, with a 2D ordinary Fourier series having the exact same band-limit as the spherical harmonic series of the original map defined on the sphere. This allows the use of efficient nonuniform Fast Fourier Transform for accurate interpolation.

output field values, and the entries of the matrix $\mathbf{Y}$ are the spherical harmonics. Type 1 is represented by the adjoint matrix $\mathbf{Y}^\dagger = \left[\mathbf{Y}^t\right]^*$. t denotes transposition and $*$ complex conjugation of the matrix Y. We also use the qualifiers type 1 and type 2 for the analogous nonuniform (or uniform) Fourier transforms, where the spherical harmonics and coefficients are replaced by their plane wave counterparts.

In typical applications, the total number of points N is comparable to the squared band-limit, $\ell_{\rm max}^2$. In this case the naive computational complexity of these operations is $O(\ell_{\rm max}^4)$.

2.2 Double Fourier sphere method

We adopt the approach proposed in Reinecke et al. (2023) and implement type 1 and type 2 transforms using the DFS method. In this approach, the matrix $\mathbf{Y}$ of the type 2 nuSHT is decomposed into 4 matrices,

$$\mathbf{Y} = \mathbf{NFDS}\,. \tag{5}$$

The final operation $\mathbf{N}$, which represents the main novel aspect to our approach, is a nonuniform Fourier transform of type 2 to the given locations. The role of **FDS** is to produce the needed Fourier coefficients.

The matrix $\mathbf{S}$ is an iso-latitude rSHT, that transforms the input harmonic coefficients onto an equi-spaced grid in both ϕ and θ, covering the entire sphere. $\mathbf{D}$ is a “doubling” operation, that extends the range of θ from $[0, \pi]$ to $[0, 2\pi)$, see Fig. 1. The doubling is performed by extending the meridians across the south pole back up to the north pole. The essential point is that the resulting map, seen as a map on the doubly-periodic torus, has a standard Fourier series with exactly the same Fourier band-limit[7] as the spherical harmonic band-limit of the input array $c_{\ell m}$. Finally, $\mathbf{F}$ is simply the standard 2D Fourier transform that produces the Fourier coefficients input to $\mathbf{N}$ from the doubled map.

The adjoint operator $\mathbf{Y}^\dagger$, or the type 1 nuSHT is, by definition,

$$\mathbf{Y}^\dagger = \mathbf{S}^\dagger\mathbf{D}^\dagger\mathbf{F}^\dagger\mathbf{N}^\dagger\,. \tag{6}$$

$\mathbf{N}^\dagger$ is a nuFFT of type 1 that produces Fourier frequencies from the input locations and field values. $\mathbf{F}^\dagger$ produces from these frequencies a 2D map on the torus. $\mathbf{D}^\dagger$ (the adjoint doubling matrix) effectively “folds” this doubled map. The resulting map is 2π-periodic in the ϕ-direction, and the θ-direction goes again from 0 to π. The map is then transformed to harmonic coefficients with $\mathbf{S}^\dagger$, an iso-latitude type 1 rSHT.

3 IMPLEMENTATION

We discuss the concrete GPU implementation. Readers interested in the CPU equivalent may consult Reinecke et al. (2023).

A GPU is designed to efficiently apply a single instruction on multiple data (SIMD). On the hardware side, it achieves this with Streaming Multiprocessors (SMs) (at the order of 100), that contain a number of simple processors for arithmetic operations (at the order of 100) that execute “warps” of 32 threads in parallel.

On the software side, a GPU accelerated program is executed via a number of threads that are arranged in thread blocks. The GPU is responsible for distributing the thread blocks across the SMs. High throughput is achieved by overloading SMs with many threads as to hide data latency and by ensuring that memory is accessed in multiples of the warp size.

We differentiate between the GPU memory that is “close” to the processor units and can be accessed fast by the device, and host memory, that is managed by the host system of the GPU, and which is generally slow to access by the GPU. We show data transfer benchmarks in Appendix A. Our implementation avoids data transfer and usage of host memory altogether; intermediate results are kept in GPU memory. This is realized by `CuPy`-arrays in combination with a `C++`-binding `nanobind`, Jakob (2022), handily providing a `nanobind-CuPy` interface.

Our implementation of the individual operators $\mathbf{N}$, $\mathbf{D}$, $\mathbf{F}$, and $\mathbf{S}$ and their adjoints are realised as follows.

For the (adjoint) synthesis ($\mathbf{S}^\dagger$) $\mathbf{S}$, we use the highly efficient software package `SHTns` Schaeffer (2012), and calculate the SHTs onto a Clenshaw-Curtis (CC) grid. The GPU implementation uses `vkFFT` Tolmachev (2023) for the Fourier transforms, and requires the sample size to be divisible by 4, see Appendix B for details. For the (adjoint) synthesis ($\mathbf{S}^\dagger$) $\mathbf{S}$, we use the highly efficient software

---

[7] This may be seen for example from the well-known Fourier representation of the Wigner d-matrices Risbo (1996); Huffenberger & Wandelt (2010b), and using the relation $P_\ell^m(\theta) = d_{m0}^\ell(\theta)\sqrt{(2\ell+1)/4\pi}$.

package `SHTns` Schaeffer (2012), and calculate the SHTs onto a Clenshaw-Curtis (CC) grid.

For iso-latitude rings, in order to achieve best efficiency for the Legendre transform part (the θ part of the transform, which is the critical part), modern top-performing CPU and GPU codes like `SHTns` use on-the-fly calculation of $P_\ell^m(\theta)$ using efficient recurrence formulas put forward recently Ishioka (2018). This allows to keep memory usage low: indeed, only the recurrence coefficients that are independent of θ need to be stored, which requires only $O(\ell_{\max}^2)$ memory, the same order as the data. It leaves two dimensions along which to parallelize: θ and m, and requires a sequential loop over ℓ to compute the $P_\ell^m(\theta)$ recursively. When $\ell_{\max}$ is larger than 500 to 1000, this leaves enough parallelization opportunities to efficiently use all of the GPU compute units. The computational complexity stays $O(\ell_{\max}^3)$.

The (adjoint) doubling ($\mathbf{D}^\dagger$) $\mathbf{D}$ is implemented via CUDA, and we write the arrays in a θ-contiguous memory layout, as required by `SHTns` to keep high efficiency for the Legendre transform. The computational complexity is $O(\ell_{\max}^2)$.

For the type-1 and type-2 nuFFT in 2-dimensions ($\mathbf{N}^\dagger$, $\mathbf{N}$), we use `cufinuFFT` Shih et al. (2021) in double precision, even for single-precision accuracy final results. This choice is done to robustly and consistently reach the requested accuracies for single precision, which would otherwise not always be the case using single-precision `cufinuFFT`. The nuFFT method works by utilizing the Fourier transform convolution theorem, and interpolation or convolution onto a slightly finer, up-sampled grid. Highly accurate versions use kernels whose error ϵ decreases exponentially as a function of the up-sampling factor. The computational complexity (without planning phase) is $O(\ell_{\max}^2 \log(\ell_{\max}) + \ell_{\max}^2 |\log^2(\epsilon)|)$ in 2 dimensions. It is worth mentioning that we use the guru[8] interface to `cufinuFFT` to initialize the nuFFT plans. The plans allow for repeated and fast transforms, without re-initialization. However, this planning step is the most time consuming operation and is therefore done before calling the functions.

For the FFT operations $\mathbf{F}$ (type 2) and $\mathbf{F}^\dagger$ (type 1), we use the package `CuPyx` Okuta et al. (2017) and its `cuFFT`[9] integration therein. For the type 1 Fourier synthesis, we use double-precision accuracy. For the type 2 Fourier synthesis, we use single precision accuracy for $\epsilon \leq 10^{-6}$, which increases speed at the cost of a negligible decrease in effective accuracy on the final result. For $\epsilon < 10^{-6}$, we again use double-precision accuracy. The computational complexity is $O(\ell_{\max}^2 \log(\ell_{\max}))$.

FFTs become particularly fast if the prime factorization for the sample size gives many small prime numbers, in the following referred to as a *good number*. If additional constraints are put on the sample size, good numbers may be more difficult to find, see Appendix B for a discussion and concrete definition.

Ignoring the scaling with accuracy, the overall asymptotic computational complexity (for both type 1 and 2) is,

$$O(\ell_{\max}^3) + O(\ell_{\max}^2 \log(\ell_{\max})) + O(\ell_{\max}^2)\,. \tag{7}$$

We assume that the number of uniform, or equi-spaced grid in (θ, ϕ), and the number of nonuniform points are about the same. In the case of CMB weak lensing, the number of equi-spaced grid points is set by the band limit of the unlensed map, and the number of non-uniform grid points can be assumed close to this number for robust reconstructions. We provide a pointing routine for calculating the deflected locations, and implemented via `CUDA`, see Appendix C. The pointing routine provides the deflected positions (θ', ϕ') from the initial undeflected positions and the deflection field of the CMB. For use cases that are not CMB weak lensing, `cunuSHT` supports the use of custom user-provided non-uniform locations as the nuSHT routines merely require a set of new deflected locations.

$\mathbf{Y}$ (and $\mathbf{Y}^\dagger$) could be further optimized: $\mathbf{S}$ and $\mathbf{F}$ both contain a Fourier transform in ϕ-direction, effectively cancelling out each other. Avoiding this reduces $\mathbf{F}$ to a 1-dimensional Fourier transform and $\mathbf{S}$ to a Legendre transformation. This optimization is implemented in the CPU implementation in `DUCC`. We leave this optimization to a future study for the GPU implementation. We only expect a large speed-up for the $\mathbf{F}$ operator in `cunuSHT`, which takes about 10 to 20% of the total runtime.

In Appendix D, we show example code snippets that demonstrate the usage of `cunuSHT`.

4 BENCHMARK

We present the scaling and execution time of our implementation. We take as a use case an application in CMB weak lensing Lewis & Challinor (2006). CMB weak lensing describes the deflection of primordial CMB photons by mass fluctuations along the line of sight as they travel through the Universe. These small deflections (their root-mean-square is of the order of a few arcminutes) are large enough to be detectable in CMB sky maps Aghanim et al. (2020); Qu et al. (2024); Pan et al. (2023). For some applications, it is necessary to simulate these deflections accurately and efficiently Hirata & Seljak (2003); Carron & Lewis (2017); Belkner et al. (2024).

Owing to the deflections, the CMB intensity field $\tilde{f}(\theta, \phi)$ observed at location (θ, ϕ) is the un-deflected field f at another location,

$$\tilde{f}(\theta, \phi) = f(\theta', \phi')\,, \tag{8}$$

where (θ', ϕ') depends on (θ, ϕ) in a smooth way. Eq. (8) describes the relation between the spherical harmonic transforms and the remapping operation: a remapping is a SHT (Eq. (2) and Eq. (3)) with the locations (θ, ϕ) replaced by the new, deflected locations

---

[8] This is an interface that allows for more fine grained configuration of the nuFFT pipeline.

[9] `https://developer.nvidia.com/cufft`

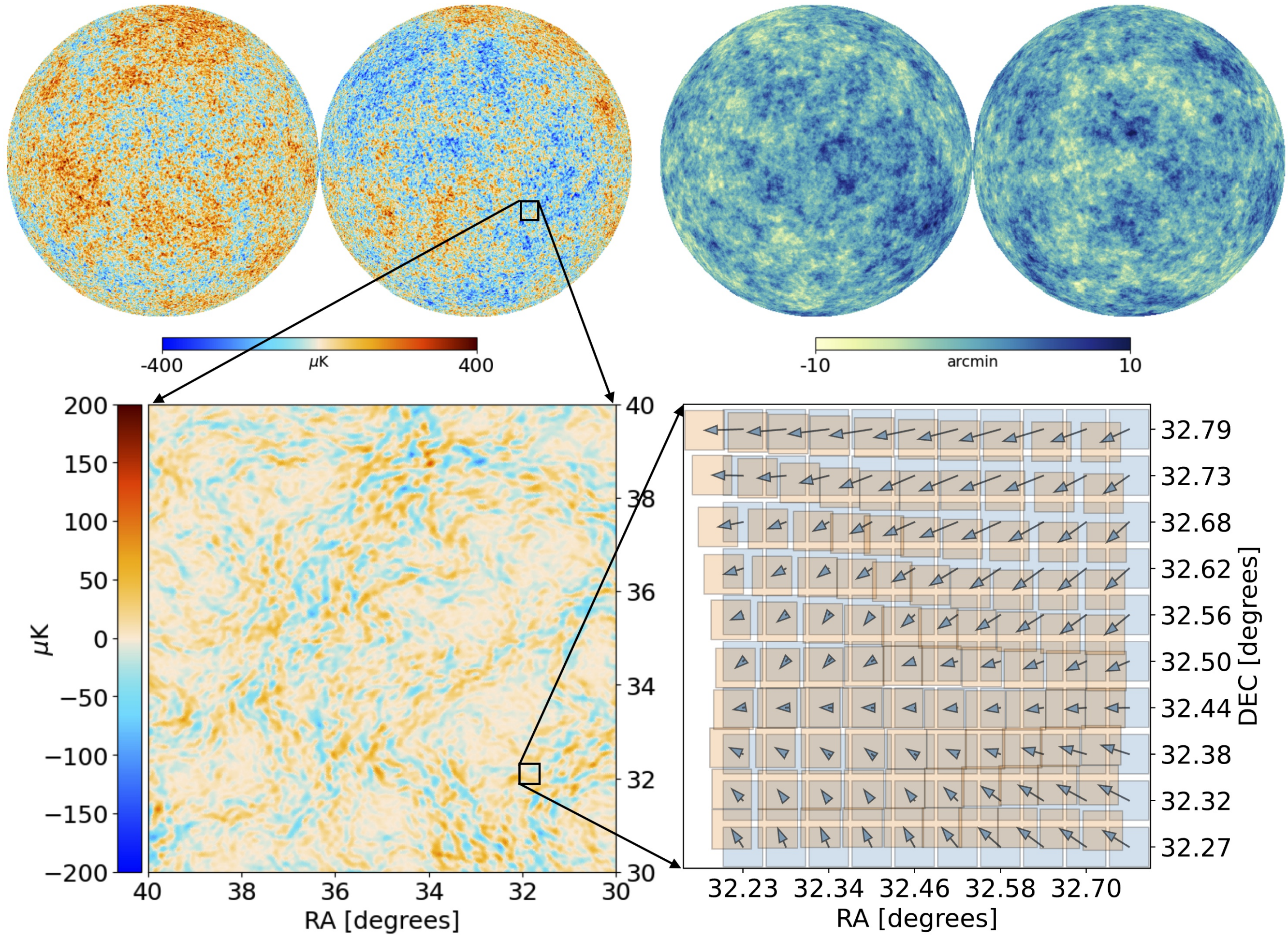

Figure 2. Setup and simulated data, here for a problem size of $\ell_{\max} = 3095$, $N \approx 2 \cdot 10^7$. The top left (right) panel shows a typical CMB temperature map (a typical deflection field) in orthographic projection. The bottom left plot shows a 10×10 degree detail view of the difference between the undeflected and deflected CMB in Cartesian projection. On the bottom right, we show a 0.5×0.5 degree detail view comparison between the uniform grid (blue squares), nonuniform grid (orange squares), and the relation between them as indicated by the black arrows.

(θ', ϕ'). Concretely for type 1,

$$c_{\ell m} = \sum_{i=1}^{N} \tilde{f}_i \, Y_\ell^{\dagger m}(\theta'_i, \phi'_i) \,, \tag{9}$$

where $\tilde{f}$ is the set of values of the lensed field. For the type 2 nuSHT, the inputs are these same set of angles and the spherical harmonic coefficients of the un-deflected field f. Details on how these angles relate to each other are given in Appendix C. The set up of our use case is shown in Fig. 2. The top left panel shows the input CMB map, the deflection field is shown on the top right. Both are shown in orthographic projection. The bottom maps show a detail view of 10×10 degree of the difference between the input and deflected map (left panel) and a detail view of approximately 0.5×0.5 degree (right panel) of the nonuniform point (orange squares) relative to the uniform point locations (blue squares). We use the same setup for the adjoint operation, in which case the deflected map becomes the input, and the result becomes the adjoint SHT coefficients.

Benchmarks are run on an NVIDIA A-100 GPU with 80 GB of memory and an Intel Xeon Gold 8358 processor with 32 cores. We set the number of threads to 32 for the CPU benchmarks. If not stated otherwise, we choose the following parameters (that mostly affect the nuFFT): an up-sampling factor of 1.25 to reduce the memory usage at a small price of increased computation time[10], a `gpu_method` utilizing the hybrid scheme, called *shared memory*, and the default kernel evaluation method. The resulting map (or input map in the case of the adjoint) is calculated on a Gauss-Legendre grid.

With this implementation, we can solve problem sizes of up to $\ell_{\max} \sim 9000$ on an A-100 with 80 GB, by using the pre-computed nuFFT plans and keeping all necessary intermediate results in memory.

Fig. 3 shows the speed-up of the GPU algorithm compared to the CPU as a function of $\ell_{\max}$ for different accuracies. The left panel

---

[10] This somewhat low up-sampling factor reduces the maximal possible accuracy, in double precision, to 10^{-10} due to its dependence on the size of the up-sampled grid, but can easily be changed by the user, if needed.

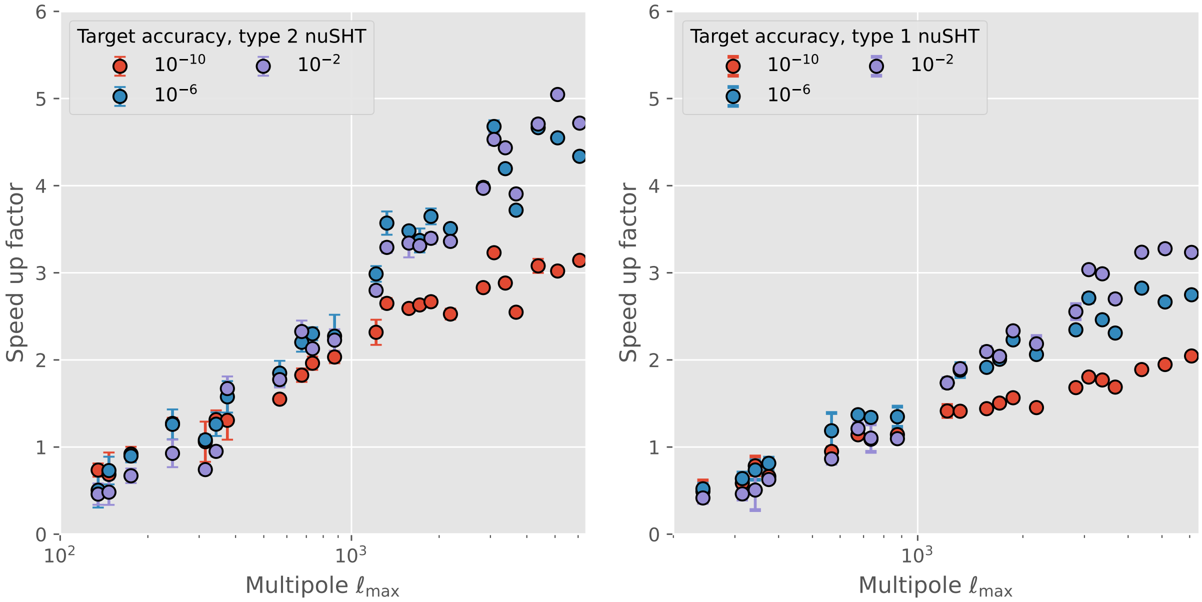

Figure 3. `cunuSHT` execution time comparison against `DUCC` for type 2 (type 1) nuSHT shown in the left (right) panel, comparing one A-100 against one Intel Xeon Gold 8358 processor with 32 cores. A speed-up factor > 1 means that the GPU is faster, and we show the result for different target accuracies. The error bars show the $\pm 1\sigma$ variance calculated from 5 runs. The GPU takes over at $\ell_{\rm max} \sim 3 \cdot 10^2$ for type 2, and at $\ell_{\rm max} \sim 4 \cdot 10^2$ for type 1. The speed-up increases up to 5 (3) for large $\ell_{\rm max}$ left (right) panel.

shows the evaluation of Eq. (5), the right panel shows Eq. (6). The $\pm 1\sigma$ variance from 5 runs is indicated by error bars.

For type 2 nuSHT, we reach a speed-up between 1 and 5 times for single precision and 1 to 3 times for double precision, with the speed-up increasing with increasing $\ell_{\rm max}$. This increase is expected due to better parallelizability for GPUs for higher $\ell_{\rm max}$. The speed-up is mostly independent of the target accuracy, with the smallest accuracies tending to perform better. For type 1 nuSHT, we find lower speed-up factors, which may indicate further improvements. However, type 1 directions are generally expected to perform worse on the GPU due to the sheer amount of threads that have to be written concurrently to the same memory location. Nevertheless, the GPU code either performs almost as good as the CPU code (small $\ell_{\rm max}$), or better by a factor of up to 3 for large problem sizes. For high $\ell_{\rm max}$, the speed-up depends on the accuracy, with lower accuracies performing better. For double precision ($\epsilon = 10^{-10}$), the speed-up is diminished due to the double-precision penalty that we pay in our implementation.

We can get a better understanding of the resulting speed-up factors by looking at the time spent on each of the operators on the GPU. This breakdown is shown in the top panels of Fig. 4 for the type 2 (left panel) and type 1 (right panel) nuSHT, as a function of $\ell_{\rm max}$ for different accuracies. The respective total execution times are shown in the bottom panels together with the empirically fitted computational complexity model, Eq. (E1). At the top panels, for each problem size $\ell_{\rm max}$, each bar represents a benchmark with an accuracy of (from left to right) 10^{-10}, 10^{-6}, 10^{-2}.[11] Each bar represent the mean over 5 runs. For small problem sizes, doubling dominates and becomes almost negligible for large $\ell_{\rm max}$. **S** only takes about 20% of the execution time for large $\ell_{\rm max}$, even though it has the worst asymptotic computational complexity. This highlights the quality of the rSHT implementation by `SHTns`.

The choice of accuracy has an impact on the total execution time, as seen in the bottom panels of Fig. 4. The highest accuracy ($\epsilon = 10^{-10}$, red line) takes at most twice as long compared to the low accuracy ($\epsilon = 10^{-2}$, purple line). Our results suggests that improvements in **F** might be possible: due to FFT being in principle a bandwidth limited routine, we would expect the execution time of **F** to be comparable to that of **D**, and only a fraction of that of **S**. For type 1 (right column), the total execution time overall takes longer. The breakdown shows that for the low and intermediate accuracy cases, less time is spent in the $\mathbf{F}^\dagger$ call. This is a consequence of our use of single precision arithmetic FFTs for the $\epsilon \leq 10^{-6}$ case, which, as mentioned before, is possible for the type 1 nuSHT. The lower panels show an empirical fit of the data to the computational complexity model, see Appendix E.

It is interesting to compare this breakdown to the CPU implementation of `DUCC`. While we cannot expect both breakdowns to be exactly the same due to the different nature of the hardware, large differences may imply possible improvements. Fig. 5 shows the

---

[11] The highest possible accuracy for double precision is 10^{-14} and tests show that this can well be reached with our code. We have, however not included them in our benchmarks.

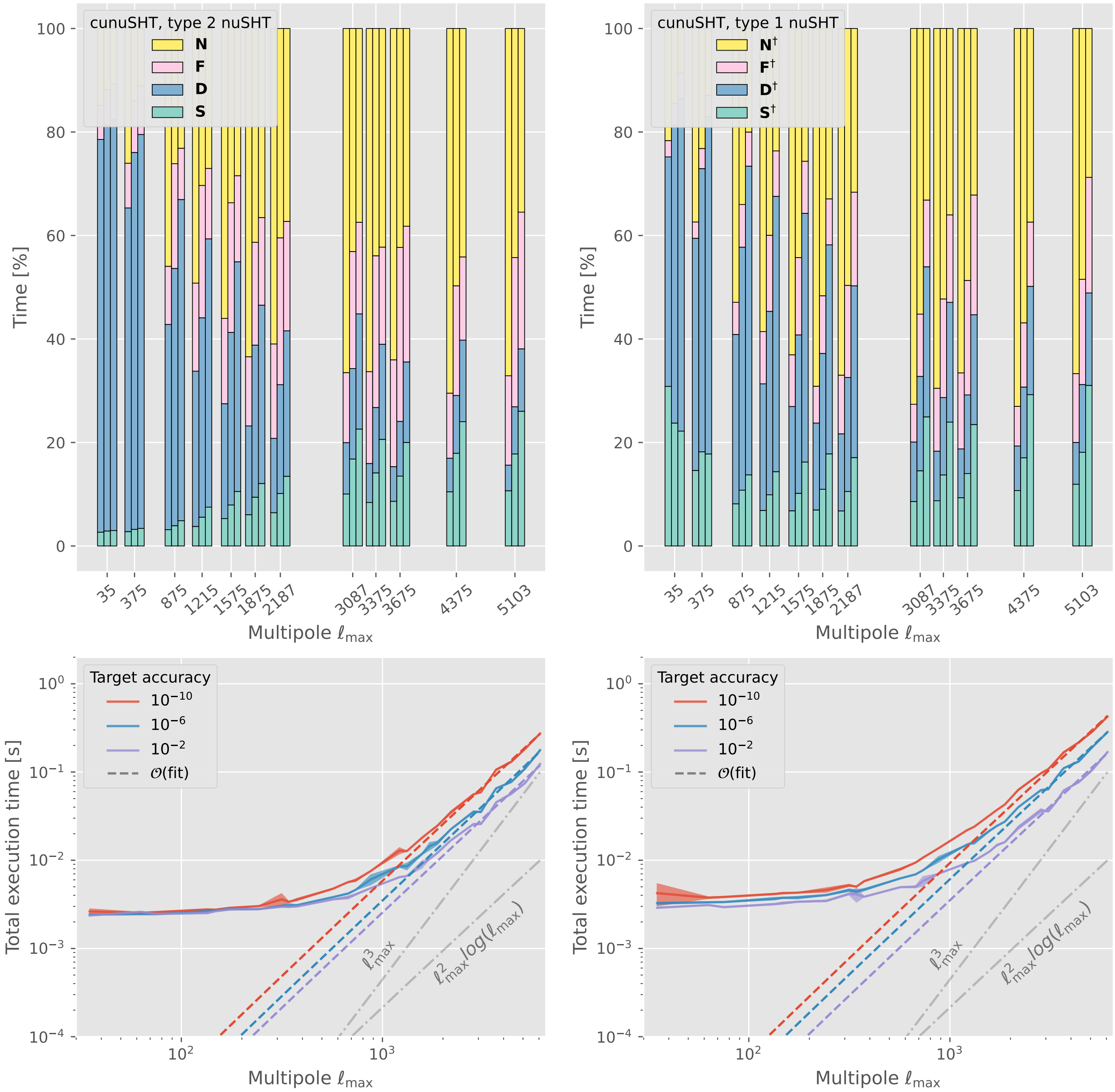

Figure 4. Breakdown and total execution time of the GPU implementation `cunuSHT`. The top panels show the percentage of time spent with the individual operators for type 2 (left column) and type 1 (right column) nuSHT. For each of the problem sizes we show this for different accuracies (10^{-10} on the left, 10^{-6} in the center, 10^{-2} on the right). The bottom panels show the total execution time, with $\pm 1\sigma$ variance as shaded area, for different accuracies and the empirically fitted computational complexity model, Eq. (E1). The grey dashed lines show the theoretical scaling for different computational complexities.

computation time of the CPU implementation of `DUCC` for the type 2 nuSHT (left column) and type 1 nuSHT (right column) as a function of $\ell_{\max}$ for different accuracies. Looking at the top left panel, we see that **S** increases with increasing $\ell_{\max}$, as expected from its $O(\ell^3_{\max})$ scaling. The bottom panel of the left figure shows that the total execution time of the high-accuracy run is at most twice as long as the low accuracy one.

Many optimizations have gone into `DUCC`'s and `SHTns`'s SHT routines (**S**, $\mathbf{S}^\dagger$) over the years and it is safe to assume that they are close to optimal. Comparing the breakdown of the CPU and GPU implementation for **S** (green bars) shows that `DUCC` spends more than twice as long with this operator. This hints to potential sub-optimalities in the implementations of the GPU operators. This also becomes apparent when we look at the operator **N** (and $\mathbf{N}^\dagger$). The scaling with the problem size is much more pronounced for the GPU code.

Fig. 6 shows the effective accuracy as a function of target accuracy for both CPU and GPU. The effective accuracy ϵ_{eff} is calculated by

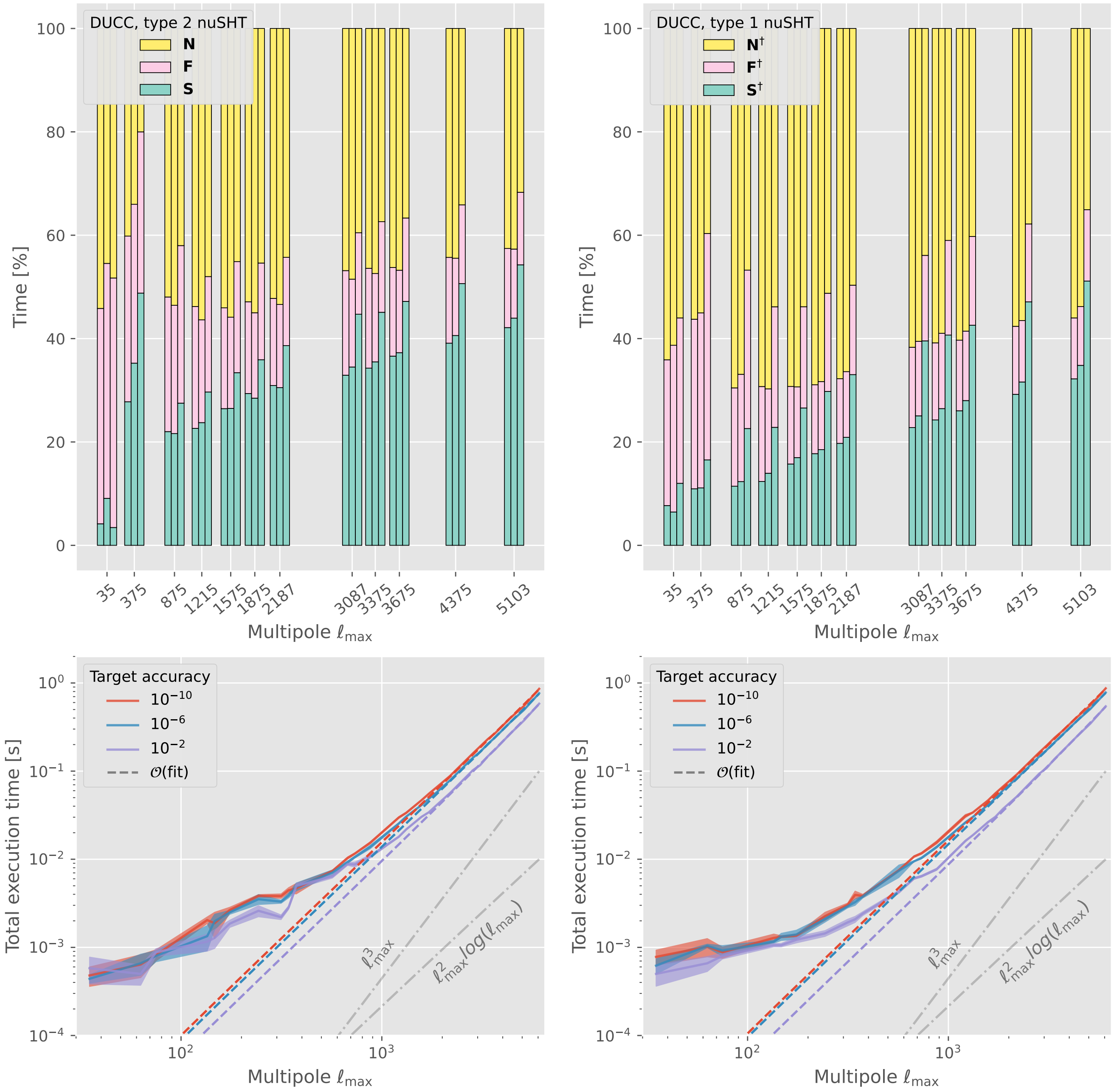

Figure 5. Same as Fig. 4, but for the CPU implementation `DUCC`. It is important to note that `DUCC` implements the individual operators more efficiently by avoiding redundant Fourier transforms and by effectively combining the doubling and Fourier transform operations into one. To reflect this, we have grouped the **D** and **F** contributions in the top panel.

solving Eq. (3) in a brute-force manner (giving us the values f^{true}) and comparing it against Eq. (5) (giving us the values $\hat{f}^{\mathrm{est}}$),

$$\epsilon_{\mathrm{eff}} = \frac{1}{\bar{f}^{\mathrm{true}}}\sqrt{\sum_{i=1}^{N_t}\left(f_i^{\mathrm{true}} - \hat{f}_i^{\mathrm{est}}\right)^2}, \tag{10}$$

with

$$\bar{f}^{\mathrm{true}} = \sqrt{\sum_{i=1}^{N_t}\left(f_i^{\mathrm{true}}\right)^2}. \tag{11}$$

We choose $N_t = 10^6$ random points on the sphere, giving us good sampling across the full sphere, and a sufficiently low variance on ϵ_{eff}. Both algorithms achieve good effective accuracies, with the CPU code being more conservative. For the GPU implementation, we note that only low effective accuracies, $\epsilon_{\mathrm{eff}} > 10^{-4}$, can be achieved when nuFFT is executed in single precision. When an accuracy of $\epsilon = 10^{-6}$ is desired, a single precision nuFFT evaluation is thus insufficient. We find that executing **N** and $\mathbf{N}^\dagger$ in double precision solves this. There is an associated penalty in efficiency due to the

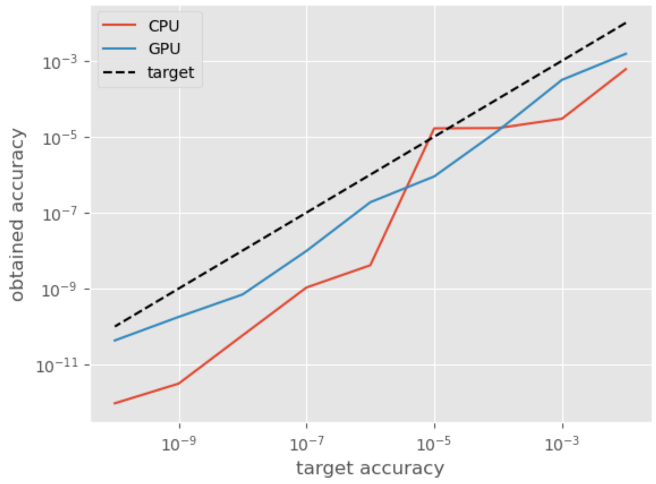

Figure 6. Effective accuracy as a function of target accuracy for the CPU (blue), and GPU (orange) implementation.

double-precision calculations being approximately twice as slow. The nuFFT implementation in `DUCC` circumvents this by allowing for double-precision accuracy on the pointing and intermediate results while using single precision accuracy on the Fourier coefficients, hereby reducing the execution time for single precision accuracies.

5 CONCLUSION

We presented `cunuSHT` Belkner (2024)[12], a GPU accelerated implementation of the spherical harmonic transform on arbitrary pixelization that is, to our knowledge, the first of its kind to achieve faster execution when compared against CPU based algorithms. `cunuSHT` can be made to achieve machine precision accuracy by transforming the problem of interpolating on the sphere into a problem of computing a nonuniform fast Fourier transform on the torus. Comparing our implementation executed on an A-100 to the fastest available CPU implementations to date running on a single Intel Xeon Gold 8358 processor with 32 cores, we find that our implementation is up to 5 times faster. We used highly efficient, publicly available packages that are well tested and robust: `SHTns` for rSHTs, `cufinuFFT` for nuFFTs. We found that, although it has the highest asymptotic complexity, the high-quality rSHT implemented in SHTns is not the bottleneck. Our code does not require intermediate transfers between host and device, allowing it to be incorporated within larger GPU-based algorithms. Many applications in cosmology that we have in mind typically require spin-1 to 3 transforms. We thus plan to extend this package to spin-n transforms in the near future. There are, in principle, no obstacles to the generalization to spin-n by implementing the corresponding Wigner-d transform.

`cunuSHT` is a general purpose package distributed via pypi, and also works on standard pixelization schemes such as `HEALPix`, and can also perform rSHTs. Demonstrations of our package on GitHub present how to integrate it into exisiting pipelines.

ACKNOWLEDGMENT

The authors thank Alex Barnett for helpful discussions about nuFFT, Lehman Garrison for help on the `CuPy-nanobind` implementation, and Libin Lu for general discussion and upgrades to `cufinuFFT`. We also thank the anonymous reviewers for helpful comments and suggestions to improve this paper. This work is supported by the Research Analyst grant from the Simons Foundation and the computing resources of the Flatiron Institute. The Flatiron Institute is supported by the Simons Foundation. SB and JC acknowledges support from a SNSF Eccellenza Professorial Fellowship (No. 186879).

DATA AVAILABILITY

No new data were generated or analysed in support of this research.

References

Ade, P. A. R. et al. 2018, Phys. Rev. Lett., 121, 221301
Ade, P. A. R. et al. 2021, Phys. Rev. D, 103, 022004
Aghanim, N. et al. 2020, Astron. Astrophys., 641, A8
Baleato Lizancos, A. & White, M. 2024, JCAP, 05, 010
Barnett, A. H., Magland, J., & af Klinteberg, L. 2019, SIAM Journal on Scientific Computing, 41, C479
Basak, S., Prunet, S., & Benabed, K. 2008 [[:arXiv]0811.1677]
Belkner, S. 2024, cunuSHT
Belkner, S., Carron, J., Legrand, L., et al. 2024, Astrophys. J., 964, 148
Boer, G. J. & Steinberg, L. 1975, Atmosphere, 13, 180
Browning, G. L., Hack, J. J., & Swarztrauber, P. N. 1989, Monthly Weather Review, 117, 1058
Brun, A. S. & Rempel, M. 2009, "Space Science Reviews", 144, 151
Carron, J. & Lewis, A. 2017, Phys. Rev. D, 96, 063510
Deppe, N., Throwe, W., Kidder, L. E., et al. 2024, `SpECTRE v2024.06.18`, 10.5281/zenodo.12098412
Driscoll, J. & Healy, D. 1994, Advances in Applied Mathematics, 15, 202
Fabbian, G., Calabrese, M., & Carbone, C. 2018, Journal of Cosmology and Astroparticle Physics, 2018, 050–050
Ferlito, F., Davies, C. T., Springel, V., et al. 2024, Ray-tracing vs. Born approximation in full-sky weak lensing simulations of the MillenniumTNG project
Hale, N. & Townsend, A. 2015, IMA Journal of Numerical Analysis, 36, 1670
Hanany, S. et al. 2019 [[:arXiv]1902.10541]
Hikage, C., Takada, M., Hamana, T., & Spergel, D. 2011, MNRAS, 412, 65
Hirata, C. M. & Seljak, U. 2003, Phys. Rev., D68, 083002
Hivon, E., Górski, K. M., Netterfield, C. B., et al. 2002, ApJ, 567, 2
Huffenberger, K. M. & Wandelt, B. D. 2010a, The Astrophysical Journal Supplement Series, 189, 255
Huffenberger, K. M. & Wandelt, B. D. 2010b, Astrophys. J. Suppl., 189, 255

---

[12] `https://github.com/Sebastian-Belkner/cunuSHT`

Hupca, I. O., Falcou, J., Grigori, L., & Stompor, R. 2012, Spherical Harmonic Transform with GPUs (Springer Berlin Heidelberg), 355–366

Ishioka, K. 2018, Journal of the Meteorological Society of Japan, 96, 241

Jakob, W. 2022, nanobind: tiny and efficient C++/Python bindings, https://github.com/wjakob/nanobind

Keiner, J., Kunis, S., & Potts, D. 2009, ACM Trans. Math. Software, 36, Article 19, 1

Lewis, A. 2005, Phys. Rev., D71, 083008

Lewis, A. & Challinor, A. 2006, Phys. Rept., 429, 1

McEwen, J. D. 2011, All Res. J. Phys, 1, 4

McEwen, J. D. & Wiaux, Y. 2011, IEEE Transactions on Signal Processing, 59, 5876–5887

Merilees, P. E. 1973, Atmosphere, 11, 13

Næss, S. K. & Louis, T. 2013, Journal of Cosmology and Astroparticle Physics, 2013, 001–001

Okuta, R., Unno, Y., Nishino, D., Hido, S., & Loomis, C. 2017, in Proceedings of Workshop on Machine Learning Systems (LearningSys) in The Thirty-first Annual Conference on Neural Information Processing Systems (NIPS)

Pan, Z. et al. 2023, Phys. Rev. D, 108, 122005

Potts, D., Steidl, G., & Tasche, M. 1998a, Linear Algebra and its Applications, 275-276, 433, proceedings of the Sixth Conference of the International Linear Algebra Society

Potts, D., Steifdl, G., & Tasche, M. 1998b, Mathematics of Computation, 67

Potts, D. & Van Buggenhout, N. 2017, in 2017 International Conference on Sampling Theory and Applications (SampTA), 82–86

Price, M. A. & McEwen, J. D. 2024, Journal of Computational Physics, 510, 113109

Qu, F. J. et al. 2024, Astrophys. J., 962, 112

Reinecke, M., Belkner, S., & Carron, J. 2023, Astron. Astrophys., 678, A165

Reinecke, M. & Seljebotn, D. S. 2013, Astronomy & Astrophysics, 554, A112

Risbo, T. 1996, Journal of Geodesy, 70, 383

Schaeffer, N. 2012, Geochemistry, 14, 751

Scharf, C. A. & Lahav, O. 1993, MNRAS, 264, 439

Seljebotn, D. S. 2012, The Astrophysical Journal Supplement Series, 199, 5

Shih, Y.-h., Wright, G., Andén, J., Blaschke, J., & H. Barnett, A. 2021, 2021 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW), 688

Slevinsky, R. M. 2019, Applied and Computational Harmonic Analysis, 47, 585

Szydlarski, M., Esterie, P., Falcou, J., Grigori, L., & Stompor, R. 2013 [[:arXiv]1106.0159]

Tolmachev, D. 2023, IEEE Access, PP, 1

Townsend, A., Wilber, H., & Wright, G. B. 2016, SIAM Journal on Scientific Computing, 38, C403

Wandelt, B. D., Larson, D. L., & Lakshminarayanan, A. 2004, Phys. Rev. D, 70, 083511

Wedi, N. P., Hamrud, M., & Mozdzynski, G. 2013, Monthly Weather Review, 141, 3450

Yin, F., Wu, J., Song, J., & Yang, J. 2019, Mathematics, 7

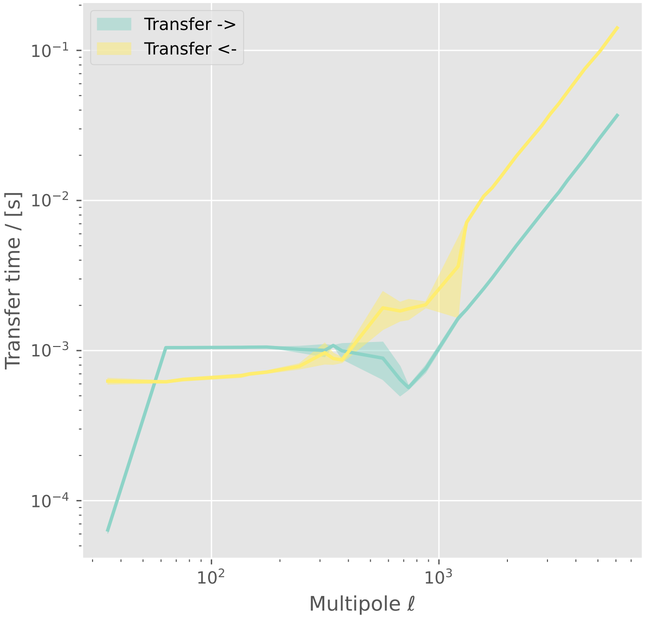

Figure A1. type 2 nuSHT transfer times “Host-to-Device” (H2D) (green) and “Device-to-Host” (D2H) (yellow) as a function of $\ell_{\max}$. The shaded areas show the $\pm 1\sigma$ variance in transfer time, calculated from ten runs. Note that these transfer times are not generally part of the routine and can completely be avoided. H2D transfer contains the SHT coefficients, while D2H contains the much larger map. For type 1 nuSHT, the transfer times are equally long, but the H2D and D2H are swapped due to the input and output of the function.

APPENDIX A: DATA TRANSFER

All benchmarks are performed without accounting for the time required to transfer data to and from the device (the GPU). This can be a large part of the overall computation and should be avoided, as is shown in Fig. A1 for different problem sizes. We note here that `cunuSHT` provides means to keep everything on the GPU without having to transfer the data. Thus, these transfer times can in principle be avoided.

APPENDIX B: GOOD NUMBERS

For fast Fourier transforms, execution time depends strongly on the largest prime factor of the transform length; the smaller it is, the better. Luckily enough, many good numbers for CPU algorithms exist. Assuming Clenshaw-Curtis quadrature, the constraint on the number of rings N_r to sufficiently sample a map with band limit $\ell_{\max}$ is $N_r \geq \ell_{\max} + 1$. Therefore, after applying the double Fourier sphere method, we must have at least $2N_r + 2$ samples, and we are free to choose an arbitrarily larger number which happens to be a good FFT size. For the GPU algorithm, however, there is a caveat: due to the additional constraint on N_r to be divisible by 4 for the rSHT operation, the selection of good numbers is reduced. While there are still good numbers for FFT satisfying both constraints, they are sparser. The full list of N_r that are multiples of 4 up to $N_r \leq 10000$

for which $\ell_{\max} = N_r - 1$ can be factored into primes up to 11 is 4, 8, 12, 16, 28, 36, 56, 64, 76, 100, 136, 148, 176, 232, 244, 276, 316, 344, 364, 376, 496, 540, 568, 676, 736, 848, 876, 892, 1156, 1216, 1324, 1332, 1376, 1576, 1716, 1816, 1876, 2080, 2188, 2476, 2696, 2836, 3088, 3268, 3376, 3676, 4236, 4376, 4456, 4852, 5104, 5776, 6076, 6616, 6656, 6876, 7204, 7624, 7876, 8020, 8576, 9076, 9376.

The numbers with the highest prime factor being 11 are underlined; they tend to lead to slightly less fast computation times in the FFT parts, and have not been used in this study. They could nonetheless be used, as well as numbers factored by higher primes.

It is important to note that during the review process of this paper, `SHTns`' was upgraded. Its next release will support arbitrary N_r, removing the constraint on N_r to be divisible by 4. This increases the list of available good numbers. Hence, restrictions on the number of good numbers are expected to go away with the next release of `SHTns`.

APPENDIX C: CMB WEAK LENSING POINTING

Let $\hat{e}_\theta, \hat{e}_\phi$ and $\hat{n}$ form the right-handed unit basis vectors at the point on the sphere parametrized by θ, ϕ (so, the components of $\hat{n}$ are $\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta$). In our benchmark CMB lensing application, the deflected positions $\hat{n}'$ that define the angles θ' and ϕ' at which the CMB field must be evaluated are given by,

$$\hat{n}' = \cos(\alpha)\hat{n} + \frac{\sin(\alpha)}{\alpha}\left(\alpha_\theta \hat{e}_\theta + \alpha_\phi \hat{e}_\phi\right) \,, \tag{C1}$$

where $\alpha = \sqrt{\alpha_\theta^2 + \alpha_\phi^2}$, and α_θ and α_ϕ form the gradient of the lensing potential Φ,

$$\alpha_\theta(\hat{n}) + \mathrm{i}\alpha_\phi(\hat{n}) = \left(\frac{\partial}{\partial\theta} + \frac{\mathrm{i}}{\sin\theta}\frac{\partial}{\partial\phi}\right)\Phi(\hat{n}). \tag{C2}$$

`cunuSHT` first obtains α_θ and α_ϕ from this equation using `SHTns`. Then, threading across each ring, we solve for θ', ϕ' in Eq. (C1) on the fly. A benchmark as a function of $\ell_{\max}$ is shown in Fig. C1.

APPENDIX D: CODE EXAMPLES

The following code block shows a minimum working example for calculating a type 2 and type 1 nuSHT on the GPU for SHT coefficients `alm` that are deflected by a deflection field `dlm_scaled`, for an accuracy of `epsilon` and band limit `lmax`. Both routines will set up the plans, so that the actual `nusht2dX()` call may be called repeatedly. The `lenjob_geominfo` structure identifies the geometry that is used by `cunuSHT`, here `'gl'` stands for "Gauss-Legendre".

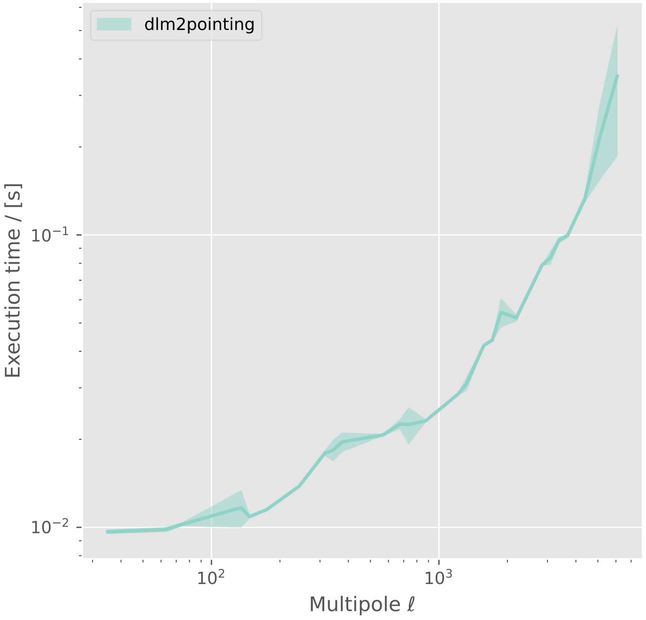

Figure C1. Execution time of the pointing routine on the GPU as a function of the problem size. The shaded area shows the ±1σ uncertainty calculated from 5 runs.

```python
import cunusht as cu

lenjob_geominfo = ('gl',{'lmax': lmax})
kwargs = {
    'geominfo_deflection': lenjob_geominfo,
    'epsilon': epsilon,
    'nuFFTtype': 2,
}
t = cu.get_transformer(backend='GPU')(**kwargs)
ptg = t.dlm2pointing(dlm_scaled)
lenmap = t.nusht2d2(alm, ptg, lmax, lenmap)
```

```python
import cunusht as cu

lenjob_geominfo = ('gl',{'lmax': lmax})
kwargs = {
    'geominfo_deflection': lenjob_geominfo,
    'epsilon': epsilon,
    'nuFFTtype': 1,
}
t = cu.get_transformer(backend='GPU')(**kwargs)
ptg = t.dlm2pointing(dlm_scaled)
alm = t.nusht2d1(alm, ptg, lmax, lenmap)
```

[]

Table E1. Empirical fits of Eq. (7) for type 2 and type 1 nuSHT, for different accuracies for both CPU and GPU.

type 2	Target acc.	α	β
CPU	10^{-10}	$1.52 \cdot 10^{-1}$	$6.96 \cdot 10^{-1}$
	10^{-6}	$1.41 \cdot 10^{-1}$	$6.16 \cdot 10^{-1}$
	10^{-2}	$1.67 \cdot 10^{-1}$	$4.11 \cdot 10^{-1}$
GPU	10^{-10}	$1.04 \cdot 10^{-6}$	$2.72 \cdot 10^{-1}$
	10^{-6}	$1.06 \cdot 10^{-2}$	$1.60 \cdot 10^{-1}$
	10^{-2}	$1.39 \cdot 10^{-4}$	$1.19 \cdot 10^{-1}$
type 1	Target acc.	α	β
CPU	10^{-10}	$1.24 \cdot 10^{-1}$	$7.34 \cdot 10^{-1}$
	10^{-6}	$1.06 \cdot 10^{-1}$	$6.67 \cdot 10^{-1}$
	10^{-2}	$1.68 \cdot 10^{-1}$	$3.72 \cdot 10^{-1}$
GPU	10^{-10}	$1.31 \cdot 10^{-8}$	$4.33 \cdot 10^{-1}$
	10^{-6}	$1.07 \cdot 10^{-6}$	$2.84 \cdot 10^{-1}$
	10^{-2}	$2.16 \cdot 10^{-5}$	$1.64 \cdot 10^{-1}$

APPENDIX E: COMPUTATIONAL COMPLEXITY

Table E1 shows the results for the empirical fit of the type 2 and type 1 nuSHT computational complexity models,

$$C(\ell) = \alpha \frac{\ell^3}{\ell_{\mathrm{norm}}^3} + \beta \frac{\ell^2 \log(\ell)}{\ell_{\mathrm{norm}}^2 \log(\ell_{\mathrm{norm}})} \,, \tag{E1}$$

with $\ell_{\mathrm{norm}} = 6067$ the normalization for the unknown prefactors. Compared to Eq. (7) we only account for the terms with the largest complexities. The fits are shown in Fig. 3, Fig. 4, and Fig. 5.

3

Searching for the Tensor-to-Scalar Ratio

3.1 Motivation

One of the primary CMB Science goals of current and future CMB experiments is the detection of the tensor-to-scalar ratio r, or to give tighter upper bounds. Two experiments that are currently designed and studied for this purpose are the CMB stage-4 (CMB-S4) and probe of inflation and cosmic origin (PICO).

Both depend on the delensing of the B-maps, a task that reduces the variance of the observed B-modes by removing the contamination coming from B-lensing. For this purpose, we have integrated `Delensalot` into the analysis pipeline of these two experiments. This integration had also been the driver for developing the `DLM` as discussed in the previous section.

This chapter discusses our contributions to the CMB lensing analysis for these two collaborations.

3.2 CMB-S4

CMB-S4 is a future CMB observation, currently in design phase and supported by the Department of Energy (DOE) Office of Science and the National Science Foundation (NSF) of the United States. It is the next generation ground-based CMB experiment, building and operating highly sensitive telescopes both at the South Pole, and the Chilean Atacama desert, and aims to cross critical thresholds in our understanding of the origin and evolution of the Universe. Recommended by the 2014 Particle Physics Project Prioritization Panel (P5), and the by the 2015 National Academies report, CMB-S4 is strongly supported and aiming for success. The CMB-S4 collaboration was established in 2018, and finally approved by the DOE critical decision (CD-0) in July 2019: today, the collaboration has 100s of members, from hundreds of different institutions, coming from almost 20 countries on 6 continents.

CMB-S4 four science cases are spectacular: first, the search for primordial gravitational waves and other evidence of inflation. Its sensitivity will probe physics at the highest energy scales and cross important thresholds for constraining inflation. Second, to study the dark Universe and search for imprints of relic particles including neutrinos. CMB-S4 will shed light on the earliest times, reaching 10 000 times further back than current experiments. Third, the mapping of the matter in the cosmos to help constraint dark energy, the sum of neutrino masses and test general relativity, and fourth, explore the time-variable millimeter-wave sky, and thus provide a powerful complement to other surveys for the task of studying the transient Universe.

CMB-S4 will provide an unprecedented dataset, setting a new baseline for testing models: we have only just begun to access the information that is available in CMB polarization, lensing, and secondaries, all of them providing a large discovery space, with CMB-S4 going to lead with a data set to discover and test new physics.

To achieve its science goals, CMB-S4 requires an enormous increase in sensitivity compared to previous experiments, totaling half a million superconducting detectors that will measure the sky at different sky fractions: the "deep" field, covering about 5% of the sky aims to reach a noise level as low as approximately 0.5μK-arcmin, ultimately giving the best data for allowing to remove about 95% of the B-lensing contaminant. The "wide" field will cover about 60% of the sky and help constraint neutrino parameters, will map the matter of the Universe, detect secondaries, and enable time-variable millimeter-wave science.

I joined CMB-S4 in 2021, and became a full member in 2023. With my contributions mostly centered around the low-ell-BB group that focuses on the r-estimate analysis (the first science case in the list above), we were able to support CMB-S4 for various analyses and contributed to many data challenges as well as to different studies: the analysis of alternatives, the search for optimal component separation, and the validation of the general analysis pipeline for the search of r that resulted in a collaboration paper.

3.2.1 Collaboration paper

This section reviews the collaboration paper that describes the functionality of `Delensalot` on a curved sky, and showcases its capabilities by reaching the CMB-S4 science goals using realistic simulation data.

This analysis was performed in the context of one of the data challenges of CMB-S4, a systematic approach to test the current r-estimate analysis pipeline. This comprehensive process includes the generation of simulations inclusive of various foregrounds and inhomogeneous noise, performing component separation, reconstructing lensing, and estimating cosmological parameters.

The successful completion of the pipeline was the result of a collaborative effort involving numerous co-authors and additional support from both the low-ell-BB group and the broader collaboration.

My contribution to this project is the execution of `Delensalot` using the component separated maps that were provided by the group and the analysis of the data products such as the B-lensing template. I also investigated the reconstruction and its dependence on the mean-field, and developed a simple forecasting tool to validate our findings on the residual lensing amplitude as a function of residual noise and foreground, and sky area. I performed comparative reconstructions for different field surveys (form the South Pole and Chile) to evaluate performance variations in the lensing reconstruction. My contribution to the paper is the writing of the result section.

The analysis began with 100 simulations for each of three sky models, eventually expanding to 500. Managing an increasing number of simulations introduced complexities, and manual adjustments of reconstruction parameters were sometimes necessary. Initially, analyses were run using parameter files, which made modifying reconstruction parameters challenging and prone to errors.

The duration and complexity of the project were considerable, largely due to the extensive data and various tests applied, requiring significant computational resources from NERSC in Berkeley. During the project, NERSC transitioned from the Cori system to Perlmutter, a newer high-performance computing system. This change led to frequent down-times and required all computations to be migrated to the new system, occasionally resulting in the repetition of some analyses.

This project was a key motivator for developing the `DLM` to ensure better control over parameters, analysis status, and automated result inspection tools.

The project ended successfully with the convergence on final results, the introduction of a new interface for `Delensalot`, and the acceptance of the paper for publication. The following contains the paper as published in The Astrophysical Journal (Astrophys. J).

Draft version October 27, 2024
Typeset using LaTeX **twocolumn** style in AASTeX631

CMB-S4: Iterative internal delensing and r constraints

Sebastian Belkner,[1] Julien Carron,[1] Louis Legrand,[1] Caterina Umiltà,[2] Clem Pryke,[3] Colin Bischoff,[4] and
CMB-S4 Collaboration

[1] *Université de Genève, Département de Physique Théorique et CAP, 24 Quai Ansermet, CH-1211 Genève 4, Switzerland*
[2] *Department of Physics, University of Illinois Urbana-Champaign, 1110 West Green Street, Urbana, IL 61801, USA*
[3] *Department of Physics, University of Minnesota, Minneapolis, MN 55455, USA*
[4] *Department of Physics, University of Cincinnati, Cincinnati, OH 45221, USA*

ABSTRACT

The tightest constraints on the tensor-to-scalar ratio r can only be obtained after removing a substantial fraction of the lensing B-mode sample variance. The planned CMB-S4 experiment[a)] will remove the lensing B-mode signal internally by reconstructing the gravitational lenses from high-resolution observations. We document here a first lensing reconstruction pipeline able to achieve this optimally for arbitrary sky coverage. We make it part of a map-based framework to test CMB-S4 delensing performance and its constraining power on r, including inhomogeneous noise and two non-Gaussian Galactic polarized foreground models. The framework performs component-separation of the high-resolution maps, followed by the construction of lensing B-mode templates, which are then included in a parametric small-aperture maps cross-spectra-based likelihood for r. We find that the lensing reconstruction and framework achieve the expected performance, compatible with the target $\sigma(r) \simeq 5 \cdot 10^{-4}$ in the absence of a tensor signal, after an effective removal of 92% to 93% of the lensing B-mode variance, depending on the simulation set. The code for the lensing reconstruction can also be used for cross-correlation studies with large-scale structures, lensing spectrum reconstruction, cluster lensing, or other CMB lensing-related purposes. As part of our tests we also demonstrate joint optimal reconstruction of the lensing potential with the lensing curl potential mode, second-order in the density fluctuations.

1. INTRODUCTION

According to the inflationary paradigm, the same mechanism responsible for the seeds of structure in the Universe also may produce a sizeable background of primordial gravitational waves (PGWs). These waves, or tensor perturbations, leave a signature on the polarization of the Cosmic Microwave Background (CMB), which is potentially observable with current technology (Seljak & Zaldarriaga 1997; Seljak 1997; Kamionkowski et al. 1997). So far, several fundamental predictions of the simplest inflationary models have been confirmed by observations, including a nearly but not exactly scale invariant spectrum of almost completely Gaussian and adiabatic scalar perturbations (Akrami et al. 2020). A detection of the tensor signal would provide us a closer view than ever of the earliest moments of our Universe, and would give strong support for large-field inflation, with important implications for our understanding of the inflationary phase and physics at very high energies (see (Kamionkowski & Kovetz 2016; Achúcarro et al. 2022) or (Lyth & Riotto 1999; Baumann & McAllister 2015) for reviews).

The tensor signal is most distinctive on the degree scales of the CMB polarization B-mode, owing to the absence of first-order scalar contributions and their associated cosmic variance (Seljak & Zaldarriaga 1997; Kamionkowski et al. 1997; Hu & White 1997). Several experiments have been conducted or are currently undertaking measurements of the B-mode power. The primordial component is usually parameterized by the tensor-to-scalar ratio r. The tightest constraint to date[1] is $r_{0.05} < 0.032$ at 95% CL., and comes from a combination of *Planck* and BICEP/Keck measurements (Ade et al. 2021; Tristram et al. 2022).

The limiting factor in this constraint is now the leading non-linear effect in the CMB: gravitational lensing. During their journey across the Universe, the CMB photons are deflected by the pull of large-scale structures. The light is deflected by a few arc minutes at all frequencies by large-scale lenses distributed across a wide range of redshifts (Lewis & Challinor 2006). In particular, this distorts the primordial polarization pattern of the CMB slightly: a pure E polarization pattern will produce a sizeable amount of B-modes af-

sebastian.belkner@unige.ch

julien.carron@unige.ch

---

a) cmb-s4.org

[1] The pivot scale in this constraint is $k_\star = 0.05\mathrm{Mpc}^{-1}$ and tensor spectral index $n_t = 0$.

ter lensing (Zaldarriaga & Seljak 1998; Knox & Song 2002; Kesden et al. 2002). On the scales relevant for r measurements, this lensing effect appears as a white-noise power with amplitude close to 5 μK-arcmin, which is well above the current sensitivity of dedicated experiments. As a result, the lensing sample variance significantly contributes to the error of the measurement of r (Ade et al. 2021).

The Cosmic Microwave Background Stage 4 experiment, or CMB-S4, is a next-generation planned ground-based experiment. Among its major science goals (Abazajian et al. 2016) is detecting PGWs at greater than 5σ, provided r is larger than 0.003, or putting an upper limit $r < 0.001$ at 95% C.L. in the absence of a signal (Abazajian et al. 2022), which would either confirm or reject some of the most popular inflationary models. This can be achieved through deep observations of the recombination peak of the B-mode power, in combination with high-resolution observations dedicated to the reconstruction of the lensing signal. These tight constraints on r from degree scales can only be achieved if the lensing sample variance is removed to some fraction of its initial value (Abazajian et al. 2022). By providing a high-quality tracer of the lenses on a broad range of scales centered on lensing multipole $L \sim 500$, one can combine this with knowledge of the E-mode on similar scales to produce a template of the lensing B-mode. This is illustrated in Fig. 1, where we have used the well-known perturbative, large-scale limit for the white lensing B-mode power (Lewis & Challinor 2006; Challinor et al. 2018),

$$C_\ell^{BB,\mathrm{len}} \simeq 2\sum_L \frac{2L+1}{4\pi} C_L^{\kappa\kappa} C_L^{EE}\,. \tag{1}$$

Here, $C_L^{\kappa\kappa}$ is the lensing convergence power spectrum and C_L^{EE} is the CMB E-mode power. The template-delensing approach is essentially optimal ((Baleato Lizancos et al. 2021a)) when compared to complex and costly Bayesian techniques that extract all of the information (Carron 2019). As can be seen in the Figure, the deep CMB-S4-like configuration allows for about 95% delensing, leaving a delensed kernel centered on $L \sim 1000$ (shown as the dashed line, again with an arbitrary normalization). External lensing tracers including ambitious future galaxy surveys are not expected to exceed a cross-correlation of ~ 0.7 (Manzotti 2018) on the relevant scales, leaving little room for further improvements to internal delensing. This is in contrast to shallower experiments such as Simons Observatory (Ade et al. 2019; Namikawa et al. 2022) which can benefit from external delensing. In a practical situation, it is often the largest lenses that are hardest to reconstruct, but as can be seen, they contribute little to the B-mode power.

The standard tools to extract the lensing deflections from the CMB are the quadratic estimators originally designed by (Hu & Okamoto 2002; Okamoto & Hu 2003) (see also the more recent (Maniyar et al. 2021)). By construction, the lensing reconstruction noise of a quadratic estimator is determined by the observed CMB total power spectra, inclusive of the lensing component and instrumental noise. While instrumental noise can be lowered with the help of more sensitive observations, the lensing component sets a fundamental limit to the precision of the estimator. However, with deep observations of the sub-degree scale CMB polarization, where the primordial B-mode is negligible, a simple field-counting argument[2] suggests that the lensing deflection field can be reconstructed internally from the CMB with high S/N. Methods to overcome this limit in idealized experimental settings have long been known (Hirata & Seljak 2003b,a), where usage of the CMB likelihood function is the key to improve the lensing measurement. The output is the Maximum A Posteriori (MAP) lensing map, conditioned on the likelihood model. More recently, this method has been revisited, cleaned of approximations and extended to more realistic conditions (Carron & Lewis 2017), and was tested on deep POLARBEAR data (Adachi et al. 2020). Its applicability remained, how-

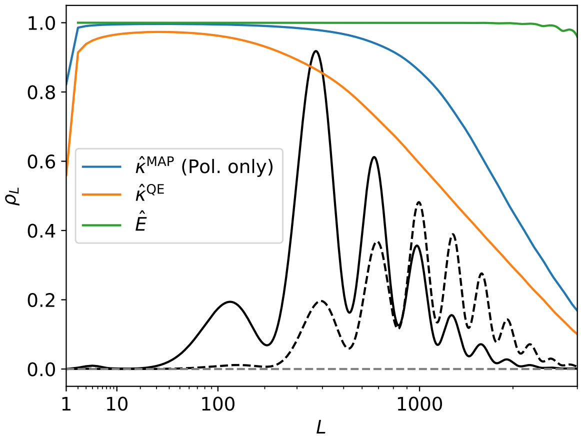

Figure 1. Expected fidelity of an optimal internal lensing reconstruction from idealized deep CMB-S4-like polarization data with noise level 0.5 μK-arcmin (blue). Plotted is the expected cross-correlation coefficient, ρ_L, of the reconstruction to the true lensing. Quadratic estimator reconstruction is also shown for comparison (orange). The solid black line shows the contribution of lensing multipoles to the large-scale lensing B-mode power, and is derived from the lensing-perturbative, large-scale limit formula, Eq. (1), with an arbitrary normalization. The shape of the black line shows that good lens reconstruction is relevant over the approximate range $30 \lesssim L \lesssim 2000$. In this approximation, the contribution of the E-mode power to the large-scale B-mode power is given by the exact same curve, and the E-mode on the relevant scales for lensing reconstruction purposes is essentially noise-free. The expected cross-correlation coefficient to the true E-mode is shown in green, and ignores possible complications on large scales owing to the intricacies of ground-based observations.

---

[2] In the absence of significant tensor modes, and since the deflection is almost a pure gradient, there are two unknown fields; the unlensed E-modes and the lensing potential, which in the low-noise limit can be reconstructed from the same number of observed fields, the lensed E and B (Hirata & Seljak 2003a), at least in principle.

ever, limited to very small fractions of the sky, since it used the flat sky approximation.

As part of the CMB-S4 collaboration planning effort (Abazajian et al. 2022), we extended the MAP lensing reconstruction method presented in (Carron & Lewis 2017) to arbitrary large sky fractions and to curved sky geometry, and used it to test the performance of a map-based internal delensing forecast. The forecast starts from component-separation of simulated CMB maps all the way to the inference of r. The CMB-S4 simulations include inhomogeneous noise that our internal delensing algorithm takes into account accordingly. Preliminary versions of these efforts towards optimal iterative lensing reconstruction on component-separated data were recently demonstrated by (Legrand & Carron 2022, 2023), focussing on the lensing auto-spectrum reconstruction. Further, this algorithm was applied on the simulated full sky polarization maps of the Probe of Inflation and Cosmic Origins (PICO) (Aurlien et al. 2022). PICO is a probe-scale space mission concept (Hanany et al. 2019) targeting $\sigma(r) \simeq 1 \times 10^{-4}$. A successful implementation of map-based internal delensing after component-separation has been demonstrated, reaching the required levels of $\sigma(r)$ for most foreground models with homogeneous noise.

In this paper, we use polarization maps only. The motivation for this is simple. Although there is some additional independent information on the lensing B-mode that can be gleaned by including the temperature maps, when the polarization noise is low, the expected gain is small.[3] Including temperature would also substantially complicate the analysis, if only due to the expected small-scale extragalactic foregrounds (such as the Cosmic Infrared Background (CIB), the Sunyaev–Zeldovich (SZ) effect and point-source signals), which are correlated with the lensing signal and to which standard quadratic estimators react strongly (Osborne et al. 2014; van Engelen et al. 2014; Ferraro & Hill 2018; Darwish et al. 2023; Baleato Lizancos & Ferraro 2022; Omori 2022; MacCrann et al. 2023). Nevertheless, the lensing reconstruction software, `delensalot` (Belkner & Carron 2023),[4] which is now publicly available, can also analyze temperature data alone or in combination with polarization data.

This paper is structured as follows. After a general discussion dedicated to lensing reconstruction beyond the quadratic estimator, Section 2 presents the lensing reconstruction scheme. This is the most technical section of the paper, which can be omitted for readers interested in results obtained on CMB-S4 simulations. In its most general implementation, the reconstruction is optimal. We close that section with a few tests assuming full-sky coverage. Among those tests, we demonstrate in Section 2.9 a successful joint-reconstruction of the lensing gradient and post-Born lensing curl deflection modes that is detectable with CMB-S4. In Sec. 3 we discuss the generation of the CMB-S4 simulations. Besides the lensed CMB signal, the simulations contain inhomogeneous noise. In addition, we also test different foreground models that contain non-Gaussianity signatures extending to high multipoles, both aspects potentially relevant for the lensing reconstruction. Results are presented in Sec. 4. We first discuss the component-separation of the simulations generated in Sec. 3, we then present the delensing performance seen on these maps, and compare to predictions, and finally, we feed the lensing templates to a parametric r-inference pipeline. We conclude in Sec. 5. The appendix collects curved-sky geometry calculational details relating to Sec. 2.

2. ITERATIVE CMB LENSING RECONSTRUCTION

Intuitively, optimal lensing reconstruction works iteratively until the two following steps have converged i) estimation of the lensing signal with a quadratic estimator, and ii) construction of a version of the CMB data that is delensed using that estimated lensing signal. This is what the pioneering papers on likelihood-based reconstructions effectively do (Hirata & Seljak 2003b,a), albeit not very transparently.

We use a straightforward and economical approach, where gradients of a CMB lensing likelihood and prior are calculated and used to progress towards the most probable lensing map, using a variant of Newton's optimization method. However, this is certainly not the only possible way to eventually achieve efficient delensing. Several variants of 'Bayesian-lensing' are being developed, that can potentially also be very powerful and most useful by the time CMB-S4 starts collecting data. For example, instead of finding the most probable lensing map, the CMB lensing posterior can also be probed through sampling (Anderes et al. 2015; Millea et al. 2020), as has also been demonstrated recently on data (Millea et al. 2021). In this case, keeping the computational cost under control is clearly a challenge at the moment. Reconstructing the joint most probable CMB fields and lensing map (Millea et al. 2019), instead of just the lensing map, is another possibility, which should have in principle a numerical complexity comparable to the one of this work. The lensing map obtained from the joint posterior has some unphysical features on large scales (Millea et al. 2019), but this is of moderate importance for delensing, and could potentially also be cured. Finally, the MUSE method (Millea & Seljak 2022; Bianchini & Millea 2023) is another iterative scheme that addresses the CMB lensing reconstruction problem which could also be efficient, but was not tested here.

Before we turn to details, let us discuss, first in general terms, the analogies and differences of our method to quadratic estimation, and the physical interpretation of the CMB lensing likelihood gradients.

Crucially, for a fixed set of lenses, a deflected Gaussian field remains Gaussian, and only displays an anisotropic power spectrum, or, equivalently, an anisotropic covariance. The log-likelihood function of the lenses, $\ln p(X^{\rm dat}|\boldsymbol{\alpha})$

---

[3] On the South Pole deep-patch configuration, the main focus of this paper, we have estimated that including temperature data could reduce the residual lensing B-mode power by only 0.2 percentage, provided the effects of atmospheric noise and foregrounds are perfectly under control.

[4] https://github.com/NextGenCMB/delensalot

(where $\boldsymbol{\alpha}$ is the deflection vector field), when seen as a function of $X^{\rm dat}$, has a Gaussian form, and therefore always contains a term quadratic in the CMB fields. The variation with respect to $\boldsymbol{\alpha}$ of this term has a form that is close to a (unnormalized) standard quadratic estimator acting on maps delensed by $\boldsymbol{\alpha}$ (Hanson et al. 2010; Carron & Lewis 2017, see also later on in this paper, where we derive all the relevant terms in this discussion), and is sensitive to the residual lensing signal beyond $\boldsymbol{\alpha}$. Further, the anisotropic covariance explicitly depends on $\boldsymbol{\alpha}$. Hence, the log-likelihood gradient also contains a term independent of the data (the variation of the log-determinant of the covariance function). This is the direct analog to the 'mean-field' of quadratic estimation that removes signatures of anisotropy, unrelated to lensing, from the quadratic piece. Let us call these two pieces $g^{\rm QD}_{\boldsymbol{\alpha}}$ and $g^{\rm MF}_{\boldsymbol{\alpha}}$, with the total log-likelihood gradient being $g^{\rm QD}_{\boldsymbol{\alpha}} - g^{\rm MF}_{\boldsymbol{\alpha}}$, as introduced in (Carron & Lewis 2017). For $\boldsymbol{\alpha} = 0$ this gradient is precisely[5] the quadratic estimator. For non-zero $\boldsymbol{\alpha}$, this is a 'quadratic' estimator probing residual lensing beyond $\boldsymbol{\alpha}$ from $\boldsymbol{\alpha}$-delensed maps. We note that the usage of 'quadratic' here is an abuse of language, since $\boldsymbol{\alpha}$ itself depends on the data in a quite convoluted manner.

In a practical situation, the quadratic estimator mean-field typically strongly dominates at low lensing multipoles, but has a very red spectral profile and therefore becomes irrelevant for smaller scale lenses (see (Hanson et al. 2009; Benoit-Levy et al. 2013), or(Aghanim et al. 2020a, figure B.1) for a detailed characterization of the components of the *Planck* reconstruction mean-field). For perfectly idealized conditions, it is zero (because with $\boldsymbol{\alpha} = 0$ and an isotropic covariance there is no special vector-field to point to). At non-zero $\boldsymbol{\alpha}$, $g^{\rm MF}_{\boldsymbol{\alpha}}$ differs in two ways from the quadratic estimator mean-field. On the one hand, the CMB maps are partially delensed by $\boldsymbol{\alpha}$. This, however, does not essentially change anisotropies coming from masking or other anisotropies, that leave a similar signature, but would have a different amplitude according to the changes in the data spectra. On the other hand, when producing the delensed maps, the instrumental noise is unavoidably also 'delensed' by $\boldsymbol{\alpha}$. This creates a new source of anisotropy in the delensed data, which was not present before delensing, and therefore induces a mean-field even for perfectly idealized conditions. To the best of our knowledge, this specific characterization of the mean-field component is new to this paper and is therefore discussed in detail in Sec. 2.5.

If $p(X^{\rm dat}|\boldsymbol{\alpha})$ were also Gaussian when seen as a function of $\boldsymbol{\alpha}$, then from the gradient calculated at $\boldsymbol{\alpha} = 0$, a single step of Newton's method would bring us directly to the maximum; in other words, the quadratic estimator would be perfect. This is clearly not the case when the lensing effect on the total map spectra is strong, which then requires an iterative method to find the maximum point. The main technical difficulties one encounters in doing so can be split as follows:

i) to get $g^{\rm QD}_{\boldsymbol{\alpha}}$ we must, using the data and $\boldsymbol{\alpha}$, estimate the partially delensed CMB in the presence of masking and (possibly wildly) inhomogeneous noise etc.;

ii) a naive estimate of $g^{\rm MF}_{\boldsymbol{\alpha}}$ in quadratic estimation must go through simulations, but this is too costly to perform at each and every iteration; and

iii) all this is costly enough numerically speaking that we must devise a scheme with fast convergence to the maximal likelihood point.

We found that converting the flat-sky tools of (Carron & Lewis 2017) to the curved sky without essential modifications does deal with point i) and iii) to a satisfactory level: lensing reconstruction is to a large extent a localized operation in position space, and flat-sky implementations can inform in a mostly reliable manner the performance of their curved-sky versions[6]. The iterative scheme we use to progress towards the maximum point is the very same L-BFGS scheme (Nocedal 1980) as described in (Carron & Lewis 2017). Point ii) has the potential of being more than a technical annoyance: even if a maximally-efficient method to calculate an accurate mean-field with just one simulation existed (which is in fact the case under some conditions, as we test later on), this would still approximately double the numerical cost. The good news is that in this paper, we account for the mean-field $g^{\rm MF}_{\boldsymbol{\alpha}}$ at no additional cost compared to a standard quadratic estimator (QE) analysis, by using the QE mean-field. The mean-field dependence on $\boldsymbol{\alpha}$ is weak, because it is almost entirely sourced by the EE part of the quadratic piece. Typical mean-field sources look like lensing convergence rather than shear, but large-scale convergence modes do not produce EB-type signatures from a pure-E-polarization pattern. Since the relative importance of the lensing contribution to the EE CMB spectrum is much weaker than for BB, the change in the mean-field with iterations is relatively weak. Additionally, low lensing multipoles where the mean-field is largest contribute little to degree-scale lensing B-modes (see Fig. 1), so inaccuracies have minor impacts.

The rest of this section is dedicated to a detailed and self-contained description of our lensing reconstruction method in curved-sky geometry. Sec. 2.1 reviews the basis of lensing remapping on the sphere, which allow us to introduce some notation as well. Sec. 2.2 defines more precisely our lensing likelihood model and priors. Sec. 2.3 obtains the lensing magnification matrix induced by the deflection on the curved sky, which our implementation requires. We then discuss $g^{\rm QD}_{\boldsymbol{\alpha}}$ in more detail in Sec. 2.4, and $g^{\rm MF}_{\boldsymbol{\alpha}}$ in Sec. 2.5. Finally, the construction of the lensing-induced CMB polarization templates from the lensing reconstruction outputs are discussed in Sec. 2.6. We make heavy use throughout of spin-weighted spherical harmonics transforms, and of the spin-

---

[5] Up to optimization of the quadratic estimator weights to include non-perturbative effects (Hanson et al. 2011; Fabbian et al. 2019), which are irrelevant for the discussion, and up to the normalization of the estimate.

[6] After accounting of course for the different scaling of spherical harmonic transforms compared to Fourier transforms etc.

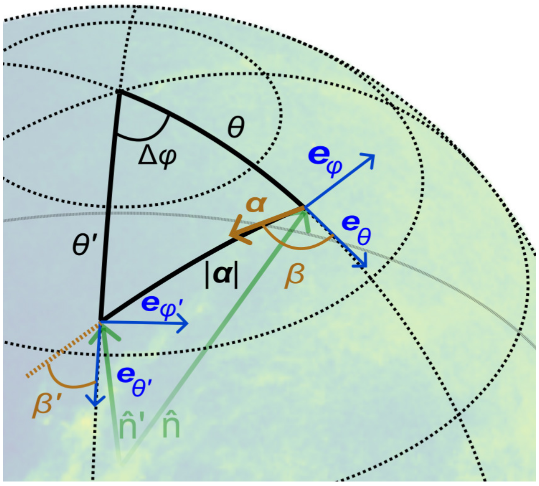

Figure 2. Lensing geometry on the curved sky and our notation, here for a greatly exaggerated deflection. The deflection vector $\boldsymbol{\alpha}$ lies in the plane tangent to the unit sphere at observed position $\hat{n}$, pointing towards the undeflected position $\hat{n}'$, lying an angle $\alpha = |\boldsymbol{\alpha}|$ away on the great circle generated by $\boldsymbol{\alpha}$. The change of the local reference axes from $\hat{n}$ to $\hat{n}'$ owing to the sky curvature is characterized by the difference $\beta - \beta'$, the change in the angle between the great circle and the local basis vector $\boldsymbol{e}_{\theta}$ at $\hat{n}$ and $\hat{n}'$ respectively (see Eq. (2)).

raising and spin-lowering operators on the sphere, with the necessary mathematics reviewed in appendix A.

2.1. *Lensing remapping*

To a good approximation,[7] the only effect of gravitational lensing on the Stokes parameters of the CMB is a remapping of points on the sphere (Lewis et al. 2017; Lewis & Challinor 2006). Within this approximation, the lensed temperature field at position $\hat{n}$ matches that of the unlensed one at $\hat{n}'$, where $\hat{n}'$ is defined lying at distance α from $\hat{n}$ along the geodesic starting there in the direction $\boldsymbol{\alpha}$. Here $\boldsymbol{\alpha}(\hat{n})$ is the deflection vector field, living in the space tangent to the unit sphere at $\hat{n}$. The geometry and angles are shown on Fig. 2. For polarization or other fields with non-zero spin, the curvature of the sphere sources a typically small additional phase shift, caused by the change in the reference axes at $\hat{n}$ and $\hat{n}'$. For a spin-weighted CMB-field (${}_s\mathcal{T}$), we can generally write the deflection operation ($\mathcal{D}_{\boldsymbol{\alpha}}$) compactly as

$$\left[\mathcal{D}_{\boldsymbol{\alpha}}\ {}_s\mathcal{T}\right](\hat{n}) \equiv e^{is(\beta-\beta')}\ {}_s\mathcal{T}(\hat{n}'). \tag{2}$$

Actual implementation of the remapping on the sphere requires explicit formulae relating $\hat{n}'$, β' to $\hat{n}$, β. These are easily gained from spherical trigonometry (Lewis 2005; Lavaux & Wandelt 2010, for example).

Let $\eth$ (or ∂^+) and $\bar{\eth}$ (or ∂^-) be the spin-raising and spin-lowering operators (reviewed very briefly in appendix A). We can use the gradient and curl scalar potentials, ϕ and Ω, to describe the two degrees of freedom of $\boldsymbol{\alpha}$, with spin-weight components,

$${}_1\alpha(\hat{n}) = -\partial^+\phi(\hat{n}) - i\partial^+\Omega(\hat{n})\,. \tag{3}$$

To first order and for a flat Universe,

$$\phi(\hat{n}) = -2\int_0^{\chi_*} d\chi \left(\frac{\chi_* - \chi}{\chi\chi_*}\right)\Psi(\hat{n},\chi)\,, \tag{4}$$

where Ψ is the Weyl potential and χ_* the comoving radial distance to the last scattering surface. There is no first-order contribution of Ψ to Ω, because to linear order in the scalar perturbations, the lensing curl potential Ω vanishes. We do not consider non-scalar sources here, since they are expected to be small. At second order, the scalar curl has, however, some relevance at CMB-S4 depth and is discussed further in Sec. 2.9. Locally, the distortions induced by $\boldsymbol{\alpha}$ can also be characterized through the (scalar) convergence κ, the (scalar) field rotation ω, and the spin-2 shear ${}_2\gamma = \gamma_1 + i\gamma_2$:

$$\begin{aligned}\kappa(\hat{n}) + i\omega(\hat{n}) &= \frac{1}{2}\partial_1^-\alpha(\hat{n})\,;\\ \gamma_1(\hat{n}) + i\gamma_2(\hat{n}) &= \frac{1}{2}\partial_1^+\alpha(\hat{n})\,.\end{aligned} \tag{5}$$

Plugging (3) into (5), we see that the gradient and curl modes of the shear are uniquely set by the gradient and curl potential, respectively. It also holds that $\kappa = -\frac{1}{2}\Delta\phi$ and $\omega = -\frac{1}{2}\Delta\Omega$, where Δ is the spherical Laplacian.

Finally, for future reference, we provide the perturbative version of the lensing remapping (2), which is valid for tiny deflections (Challinor & Chon 2002):

$$\left[\mathcal{D}_{\boldsymbol{\alpha}}\ {}_s\mathcal{T}\right](\hat{n}) \sim -\frac{1}{2}\left[{}_{-1}\alpha(\hat{n})\partial_s^+\mathcal{T}(\hat{n}) + {}_1\alpha(\hat{n})\partial_s^-\mathcal{T}(\hat{n})\right]. \tag{6}$$

In ΛCDM cosmologies, however, this perturbative description is not sufficient, and (2) must be applied.

2.2. *Lensing likelihood model and priors*

A concrete calculation of the CMB likelihood must utilize several fiducial ingredients, which we now discuss.

We model the likelihood according to the remapping approximation, Eq. (2), as follows: let ${}_2P^{\rm unl}$ be the unlensed polarization maps. The observed Stokes Q and U polarization maps are then

$$X^{\rm dat} \equiv \begin{pmatrix} Q^{\rm dat} \\ U^{\rm dat}\end{pmatrix} = \mathcal{B}\mathcal{D}_{\boldsymbol{\alpha}}\ {}_2P^{\rm unl} + \text{noise}\,, \tag{7}$$

where $\mathcal{B}$ describes the instrument beam and transfer functions, along with the projection of the lensed polarization onto its Q and U components. In all realistic cosmological

---

[7] For degree scale B-modes, leading deviations from the remapping approximation are expected to come from the combination of emission-angle and time-delay effects, relevant only for r well below 10^{-5} (Lewis et al. 2017).

models, the unlensed polarization is massively dominated by the E-mode signal. The degree scales at which the B-mode signal, sourced from gravitational waves, could potentially be observed, are typically excluded from the lensing reconstruction analysis (see Sec. 2.7). This motivates the use of a pure E-mode unlensed polarization as the baseline, in which case

$$ {}_2P^{\rm unl}(\hat{n}) = -\sum_{\ell m} E^{\rm unl}_{\ell m}\, {}_2Y_{\ell m}(\hat{n})\,. \tag{8}$$

The power spectrum of this Gaussian unlensed E-mode, the $C_\ell^{EE,\rm unl}$ spectrum, is another fiducial ingredient of the analysis. In principle, one could attempt to jointly reconstruct this spectrum along with $\boldsymbol{\alpha}$ and $E^{\rm unl}$. However, within the ΛCDM model, this spectrum is already tightly constrained (and will be even more so by the time of CMB-S4), and introducing this additional flexibility is unlikely to improve the delensing while complicating the analysis by a fair amount. Therefore, we fix it to our fiducial cosmology.

The covariance of the data conditioned on $\boldsymbol{\alpha}$, ${\rm Cov}_{\boldsymbol{\alpha}}$, is obtained by squaring (7) and averaging over the unlensed CMB, while keeping the lensing deflection fixed:

$$\begin{aligned}{\rm Cov}_{\boldsymbol{\alpha}} &\equiv \left\langle X^{\rm dat} X^{\rm dat,\dagger}\right\rangle_{\boldsymbol{\alpha}} \\ &= \mathcal{B}\mathcal{D}_{\boldsymbol{\alpha}}\, {}_2\mathcal{Y}\, C^{EE,\rm unl}\, {}_2\mathcal{Y}^\dagger \mathcal{D}^\dagger_{\boldsymbol{\alpha}} \mathcal{B}^\dagger + N\,.\end{aligned} \tag{9}$$

Here, we have used the compact matrix times vector notation ${}_2P^{\rm unl} = {}_2\mathcal{Y}E^{\rm unl}$ to represent Eq. (8), and assumed that the noise term is statistically independent of the signal, with noise covariance matrix N. The lensing likelihood then takes the form

$$ -2\ln p(X^{\rm dat}|\boldsymbol{\alpha}) = X^{\rm dat,\dagger}\, {\rm Cov}^{-1}_{\boldsymbol{\alpha}} X^{\rm dat} + \ln\det {\rm Cov}_{\boldsymbol{\alpha}}. \tag{10}$$

Of course, a likelihood that relies on several fiducial ingredients will only be approximate. For example, the noise matrix is, at least in part, only crudely known in practice. Unmodeled anisotropies in the beams or induced by the scanning strategy can also potentially introduce uncertainties in the lensing estimation. These challenges do not generally prevent the reconstruction of the lensing map, but can certainly result in a suboptimal performance of the resulting tracer. These complications are beyond the scope of this paper and left for future work.

The lensing likelihood does not carry enough information to constrain the small-scale modes of the lensing deflection field. Supplementing the likelihood with prior information is necessary for an iterative search to converge. Including a prior is in fact not only necessary, but also desirable, since the prior down-weights noisy lensing modes, thereby preventing lensing reconstruction noise to add too much B power to the lensing B-mode template. Non-linear effects affecting the lensing map are fairly weak on most scales relevant to delensing, making the simplest choice of a Gaussian prior on each mode both natural and close to optimal.[8] This requires introducing another fiducial ingredient, the spectrum of the lensing map. Using κ and ω to parametrize the field, the posterior probability density function to be maximized can then be written

$$\begin{aligned}-2\ln p(\boldsymbol{\alpha}|X^{\rm dat}) =& -2\ln p(X^{\rm dat}|\boldsymbol{\alpha}) \\ &+ \sum_{LM} \frac{|\kappa_{LM}|^2}{C_L^{\kappa\kappa,\rm fid}} + \sum_{LM}\frac{|\omega_{LM}|^2}{C_L^{\omega\omega,\rm fid}},\end{aligned} \tag{11}$$

where the likelihood on the right-hand side is given by Eq. (10), and $C_L^{\kappa\kappa,\rm fid}$, $C_L^{\omega\omega,\rm fid}$ are the fiducial spectra. In our baseline analyses the curl is assumed to vanish, in which case only the convergence term appears in this equation.

2.3. *Lensing magnification for macroscopic deflections*

Gravitational lensing changes the solid angle under which a CMB patch is observed. In Fig. 2, a lensed region of solid angle $d\Omega$ observed at $\hat{n}$ occupies an unlensed area $d\Omega|A(\hat{n})|$ at $\hat{n}'$, where

$$ |A(\hat{n})| \equiv \left|\frac{d^2\hat{n}'(\hat{n},\boldsymbol{\alpha}(\hat{n}))}{d^2\hat{n}}\right|\,, \tag{12}$$

is the determinant of the magnification matrix (the Jacobian of the sphere remapping induced by $\boldsymbol{\alpha}$). The well-known main effect is given by the convergence, as $|A| \sim 1-2\kappa$ to first order (Lewis & Challinor 2006, e.g.). For this reason, delensing the CMB will slightly demagnify (for $\kappa > 0$) or magnify (for $\kappa < 0$) regions of the data, inevitably, for the purpose of reducing statistical anisotropies in the CMB signal. However, this (de)magnification induces new anisotropies in the noise map. Based on this, and for deflections larger than the coherence length of the noise, we expect the noise covariances to crudely follow the magnification,

$$ N_{\text{[delensed map]}} \sim (1-2\kappa(\hat{n}))\, N_{\text{[raw map]}}\,. \tag{13}$$

As a result, the noise covariance becomes inhomogeneous after delensing, even if it were perfectly uniform in the data. The noise covariance decreases(increases) slightly in regions of positive(negative) convergence. This introduces a novel source of anisotropy in the iterative search for the optimal lensing map, referred to as the 'delensed-noise mean-field'. Quadratic estimates aimed at capturing residual lensing must, in principle, subtract this 'delensed-noise mean-field'. Further details on this effect are discussed in Sec. 2.5.

In our reconstruction algorithm, which is based on the likelihood function (see Eq. (9) and (10)), the operator $\mathcal{D}^\dagger_{\boldsymbol{\alpha}}$ contains the delensing operations, and its explicit form is as follows: it can be observed (using for example an explicit matrix

---

[8] Usage of a prior ignoring the non-Gaussianity of the deflection field does not prevent reconstruction of non-Gaussian features, as long as the likelihood is constraining enough. Reconstructions on simulations with realistic, non-Gaussian input lensing maps converge just as well. A detailed study will be presented in a different paper. Using a Gaussian prior, results are insensitive to the prior spectrum shape, since this is already tightly constrained from observations (Aghanim et al. 2020b; Carron et al. 2022; Qu et al. 2023).

representation for $\mathcal{D}_{\boldsymbol{\alpha}}$ defined in 2), that for deflection fields that are weak enough to be invertible, the following relation holds:

$$\left[\mathcal{D}^{\dagger}_{\boldsymbol{\alpha}}\ {}_s\mathcal{T}\right](\hat{n}') = |A_{\boldsymbol{\alpha}^{-1}}(\hat{n}')|\left[\mathcal{D}_{\boldsymbol{\alpha}^{-1}}\ {}_s\mathcal{T}\right](\hat{n}'). \tag{14}$$

In this equation, $\boldsymbol{\alpha}^{-1}$ is the deflection field inverse[9] to $\boldsymbol{\alpha}$, and $|A_{\boldsymbol{\alpha}^{-1}}|$ is the determinant of its magnification matrix.

In practice, delensing is always applied through $\mathcal{D}^{\dagger}_{\boldsymbol{\alpha}}$, on maps that are first inverse-noise-variance weighted. At the time this analysis started, the implementation of $\mathcal{D}^{\dagger}_{\boldsymbol{\alpha}}$ used relation (14), and therefore required explicit calculation of these determinants, which we now discuss.[10]

The exact form of the Jacobian determinant, and allowing for macroscopic deflections, is derived in Appendix C. The result is

$$\begin{aligned} A = &\frac{\sin\alpha}{\alpha}\left[(1-\kappa)^2 + \omega^2 - \gamma_1^2 - \gamma_2^2\right] \\ &+ \left(\cos\alpha - \frac{\sin\alpha}{\alpha}\right)\left(1 - \kappa - \cos(2\beta)\gamma_1 - \sin(2\beta)\gamma_2\right), \end{aligned} \tag{15}$$

at each point on the sphere. The entire second line is a correction that is at least quadratic in the deflection angle α. It tends to reduce the Jacobian for large deflections, since neighboring geodesics get closer together on the sphere. Sanity checks of this formula are provided in Appendix C. Generally though, the difference of $|A|$ to the leading order result $1 - 2\kappa$ is of course small for reasonably sized deflections. This linearization is accurate over the full sky to 0.3% at worst for a ΛCDM deflection field with $L_{\max} = 4000$. For the same configuration, neglecting the sky curvature (i.e. only considering the square bracket in Eq. (15)), introduces a negligible error of $\mathcal{O}(10^{-7})$ in the determinant. This is what we used in practice.

2.4. *Likelihood gradients, quadratic part*

The gradients with respect to the lensing map of the lensing likelihood are a necessary calculation to search for the maximum a-posteriori deflection field. We find position-space convenient for analytical work: lensing is a local effect on the CMB and this naturally leads to compact expressions. We proceed following Appendix A of (Carron & Lewis 2017), which we complete with the exact calculation of the gradients. We consider the complex variation,

$$2\frac{\delta}{\delta_{\pm 1}\alpha(\hat{n})} \equiv \frac{\delta}{\delta\alpha_\theta(\hat{n})} \mp i\frac{\delta}{\delta\alpha_\varphi(\hat{n})}. \tag{16}$$

These two gradients will lower (for $\delta_1\alpha(\hat{n})$), or raise (for $\delta_{-1}\alpha(\hat{n})$) the spin of the field acted upon by one unit, respectively. We can then define the spin-1 likelihood gradient as:

$$2\frac{\delta \ln p(X^{\rm dat}|\boldsymbol{\alpha})}{\delta_{-1}\alpha(\hat{n})} \equiv {}_1g^{\rm QD}_{\boldsymbol{\alpha}}(\hat{n}) - {}_1g^{\rm MF}_{\boldsymbol{\alpha}}(\hat{n}), \tag{17}$$

where we have split the total likelihood gradient into its contribution from the quadratic part ($g^{\rm QD}_{\boldsymbol{\alpha}}$), and from the log-determinant ($-g^{\rm MF}_{\boldsymbol{\alpha}}$). The decomposition of ${}_1g_{\boldsymbol{\alpha}}(\hat{n})$ into gradient and curl components gives the lensing potential and curl potential gradients.

It is apparent that the relevant variations entering Eq. (17) are those of deflected spin-weighted fields $\mathcal{D}_{\boldsymbol{\alpha}}\ {}_s\mathcal{T}$ under a change in the deflection vector. On the flat sky, where the remapping equation at position $\hat{\boldsymbol{n}}$ is merely $\left[\mathcal{D}_{\boldsymbol{\alpha}}\ {}_s\mathcal{T}\right](\hat{\boldsymbol{n}}) = {}_s\mathcal{T}(\hat{\boldsymbol{n}} + \boldsymbol{\alpha})$, the gradient with respect to $\boldsymbol{\alpha}$ is simply the gradient of the field evaluated at the deflected position. On the curved sky, we expect a similar-looking result, with (when working with the spin-weighted components) the standard gradient replaced by minus the spin-raising operator. If the deflection is large enough that the sky curvature is relevant between $\hat{n}$ and $\hat{n}'$, there will be possible corrections. As we now discuss, it turns out that they can be calculated exactly, but are also safely negligible.

The concrete calculation is deferred to appendix D. The result is

$$\begin{aligned} 2\frac{\delta\left[\mathcal{D}_{\boldsymbol{\alpha}}\ {}_s\mathcal{T}\right](\hat{n})}{\delta_{\pm 1}\alpha(\hat{n})} = &- \left[\mathcal{D}_{\boldsymbol{\alpha}}\,\eth^{\mp}_s\mathcal{T}\right](\hat{n}) \\ &+ f(\alpha(\hat{n}))\left(\left[\mathcal{D}_{\boldsymbol{\alpha}}\,\eth^{\mp}_s\mathcal{T}\right](\hat{n}) - e^{\mp 2i\beta}\left[\mathcal{D}_{\boldsymbol{\alpha}}\,\eth^{\pm}_s\mathcal{T}\right](\hat{n})\right). \end{aligned} \tag{18}$$

Here, ${}_sT$ is the corresponding spin-s field, and we have omitted its spin subindex on the right hand side of the equation. The first line is the direct analog of the flat-sky result, and the second line is the correction term. The prefactor $f(\alpha)$ is $\frac{1}{2}\left(1 - \frac{\sin\alpha}{\alpha}\right) \sim \alpha^2/12$ for small α. The second line accounts for the focusing effect of the spherical geometry on nearby geodesics. For example, all geodesics of length π starting from a point meet again at the antipodal point. This provides a useful sanity check of this formula, performed in the appendix, which is that any variation of $\boldsymbol{\alpha}$ at $\hat{n}$ that leaves the deflection angle $\alpha(\hat{n})$ invariant and equal to π must result in a vanishing gradient (18). For any realistically-sized ΛCDM CMB weak lensing deflection field of a couple of arcmin, $f(\alpha) \simeq 5 \times 10^{-8}$, hence is a negligible correction. There are no practical difficulties including this correction term and we tested its impact for both temperature and polarization estimators and on the full sky and found that it has negligible impact: the relative improvement on the cross correlation coefficient improves only marginally.

Safely neglecting this correction, the precise expression for the likelihood gradient takes a form perfectly analogous to the case of the quadratic estimator. As is well known (Okamoto & Hu 2003; Lewis & Challinor 2006; Maniyar et al. 2021), the quadratic estimator constructs an estimate of the lensed CMB from the data, and weights its gradient in real-space by inverse-variance-weighted residuals. In our context, the only difference is that the lensing

---

[9] The deflection field $\boldsymbol{\alpha}^{-1}$ inverse to $\boldsymbol{\alpha}$ is defined as the one undoing the remapping induced by $\boldsymbol{\alpha}$, when this is possible. In the notation of Eq. (2), $\mathcal{D}_{\boldsymbol{\alpha}^{-1}} = [\mathcal{D}_{\boldsymbol{\alpha}}]^{-1}$.

[10] At the time of publishing this paper, the delensing operation has been improved by (Reinecke et al. 2023), and explicit calculation of the determinant is not a requirement anymore.

likelihood uses the current estimate of lensing map $\boldsymbol{\alpha}$ as an additional piece of knowledge. The calculation of the gradient can be described as follows:

In a first step, we estimate $\boldsymbol{\alpha}$ using a standard QE, and construct an estimate of the unlensed CMB, referred to as $E^{\rm WF}_{\boldsymbol{\alpha}}$. This is in contrast to the QE for which the lensed CMB is used. We call this CMB reconstruction step 'Wiener-filtering', since it has the precise meaning of reconstructing the Maximum a Posterior point of the assumed Gaussian $E^{\rm unl}$ signal, conditioned on $\boldsymbol{\alpha}$ and the other likelihood ingredients[11] being the truth. Its explicit form is

$$E^{\rm WF}_{\boldsymbol{\alpha}} = \left[C^{EE,{\rm unl},-1} + N^{-1}_{\boldsymbol{\alpha}}\right]^{-1} {}_2\mathcal{Y}^{\dagger}\mathcal{D}^{\dagger}_{\boldsymbol{\alpha}}\mathcal{B}^{\dagger}N^{-1}X^{\rm dat}\,, \tag{19}$$

where $N_{\boldsymbol{\alpha}}$ is the delensed E-noise covariance matrix,

$$N^{-1}_{\boldsymbol{\alpha}} \equiv {}_2\mathcal{Y}^{\dagger}\mathcal{D}^{\dagger}_{\boldsymbol{\alpha}}\mathcal{B}^{\dagger}N^{-1}\mathcal{B}\mathcal{D}_{\boldsymbol{\alpha}}\,{}_2\mathcal{Y}. \tag{20}$$

In practice, the large bracketed matrix inverse is performed with a conjugate-gradient solver. In the second step, we construct the inverse-variance weighted polarization residuals, which we write ${}_2\bar{P}_{\boldsymbol{\alpha}}(\hat{n})$. The residuals are the difference between the data maps ($X^{\rm dat}$) and the prediction of the likelihood model assuming that the deflection and Wiener-filtered CMB are the truth. This conditioned prediction is the beamed and deflected polarization $\mathcal{B}\mathcal{D}_{\boldsymbol{\alpha}}P^{\rm WF}_{\boldsymbol{\alpha}}$, where the Wiener-filtered Stokes polarization $P^{\rm WF}_{\boldsymbol{\alpha}}$ is obtained from its E-mode $E^{\rm WF}_{\boldsymbol{\alpha}}$ in the usual way, Eq. (8). Inverse-variance weighting the residuals gives

$${}_2\bar{P}_{\boldsymbol{\alpha}}(\hat{n}) \equiv \left[\mathcal{B}^{\dagger}N^{-1}\left(X^{\rm dat} - \mathcal{B}\mathcal{D}_{\boldsymbol{\alpha}}P^{\rm WF}_{\boldsymbol{\alpha}}\right)\right](\hat{n})\,. \tag{21}$$

With this in hand, the quadratic piece of the likelihood gradient is

$${}_1g^{\rm QD}_{\boldsymbol{\alpha}}(\hat{n}) = -\sum_{s=\pm 2} {}_{-s}\bar{P}_{\boldsymbol{\alpha}}(\hat{n})\left[\mathcal{D}_{\boldsymbol{\alpha}}\,\partial^{+}_{s}P^{\rm WF}_{\boldsymbol{\alpha}}\right](\hat{n})\,, \tag{22}$$

where, for the reasons discussed above, we have only included the first line of Eq. (18) in the brackets. This equation is the exact analog of the standard unnormalized, quadratic estimators built from polarization (Hu & Okamoto 2002), when written in the spin-weight formalism (Aghanim et al. 2020b). Since $\bar{P}_{\boldsymbol{\alpha}}$ contains both E and B components, this gradient is the combination of an EE and an EB quadratic piece, with the EB piece containing most of the signal. There is no BB piece since our fiducial model assumes vanishing primordial B-modes.

2.5. *Likelihood gradient, 'mean-field'*

Our goal in this section is to quantify the importance of the 'delensed-noise' mean-field which we introduced briefly in Sec. 2.3. We will demonstrate that it plays a very minor role in the configurations of interest in this paper.

The mean-field is the part of the likelihood gradient that originates from the determinant term in Eq. (10). As is the case with standard quadratic estimator analyses, its role is to remove signatures of anisotropies picked up by the quadratic piece that are unrelated to the lensing signal of interest: this can be seen from the identity

$${}_1g^{\rm MF}_{\boldsymbol{\alpha}}(\hat{n}) = \left\langle {}_1g^{\rm QD}_{\boldsymbol{\alpha}}(\hat{n})\right\rangle\,, \tag{23}$$

where the average is over realizations of the data according to the likelihood model $p(X^{\rm dat}|\boldsymbol{\alpha})$, with the deflection field set to $\boldsymbol{\alpha}$ in each of these realizations.[12] In the case of the quadratic estimator (for which $\boldsymbol{\alpha} = 0$ in Eq. (23)), and in the absence of non-idealities like masking or noise inhomogeneities, the mean-field vanishes. For the iterative process, this is not true anymore, because $\boldsymbol{\alpha}$, the current best estimate of the lensing field, sources itself a mean-field: as discussed in Sec. 2.3, by moving the data slightly around, it effectively magnifies or demagnifies certain areas, and the delensed noise on the E-mode becomes inhomogeneous even if it was homogeneous before.

We calculate and discuss this contribution in the absence of other sources of anisotropies in Appendix B. For high-resolution experiments, with polarization noise spectrum N_ℓ, there is the simple result

$$g^{{\rm MF},\kappa}_{LM} \simeq -2\kappa_{LM}\sum_{\ell}\frac{2\ell+1}{4\pi}\left(\frac{C^{EE,{\rm unl}}_{\ell}}{C^{EE,{\rm unl}}_{\ell}+N_\ell}\right)\,. \tag{24}$$

This is valid for low lensing Ls, but does not fail too badly at high Ls either. The sign is negative, because a magnified ($\kappa > 0$) CMB patch will have its noise map appear demagnified after delensing. The curl component $g^{{\rm MF},\omega}$ is much smaller, since the lensing curl potential affects the magnification to second order.

Equation (24) comes from the EE-only part of the iterative quadratic estimator. Although the EB estimator captures much more signal, its contribution to the convergence mean-field is small. This can be understood as follows: a large converging lens does not locally change the pure E property of the unlensed E-modes. Hence, the EB part does not generally respond to large-scale convergence-like anisotropies.

To assess the importance of this mean-field component, it is instructive to compare its magnitude to that of the other gradients. Doing so allows us now to give an analytical argument that this mean-field component is small enough that it can only play a role when the MAP iterative search has already converged to very good accuracy, but is negligible on the onset of the iterations.

For deep experiments like CMB-S4, the number of noise-free E-modes is high. For a beam with full width at half

---

[11] Explicitly, these ingredients are the inverse noise matrix, the transfer function, and the unlensed E-mode power spectrum.

[12] The above equation follows from the fact that under any parameter variation δ we must have $0 = \delta\left\langle 1\right\rangle = \left\langle \delta \ln p(X^{\rm dat}|\boldsymbol{\alpha})\right\rangle = \left\langle g^{\rm QD}\right\rangle - \left\langle g^{\rm MF}\right\rangle = \left\langle g^{\rm QD}\right\rangle - g^{\rm MF}$. The last equality holds since the gradient of the log-determinant is data-independent.

maximum (FWHM) of 1 arcmin and polarization noise of 0.5 μK-arcmin, we find from Eq. (24),

$$g^{\rm MF,\kappa}_{LM} \simeq (-3\cdot 10^6)\kappa_{LM}\,. \tag{25}$$

If the mean-field component is neglected, then a fully converged iterative solution must have the prior and quadratic pieces in equilibrium, that is,

$$g^{\rm PR,\kappa}_{LM} = -\frac{\kappa_{LM}}{C_L^{\kappa\kappa,\rm fid}} = -g^{\rm QD,\kappa}_{LM}. \tag{26}$$

At the peak of the lensing power of $L\sim 30$, we find a similar value:

$$-\frac{\kappa_{LM}}{C_{L=30}^{\kappa\kappa,\rm fid}} \simeq (-4.5\cdot 10^6)\kappa_{LM}. \tag{27}$$

On the other hand, in the first few iterations, the quadratic piece still completely dominates. The first iteration gradient is the (unnormalized) standard quadratic estimator, $g^{QE,\kappa}_{LM}$. The standard EB quadratic estimator response in the large lens limit is

$$g^{QE,\kappa}_{LM} \simeq \\ 2\kappa_{LM}\sum_\ell \frac{2\ell+1}{4\pi}\left(\frac{C_\ell^{EE,\rm len}}{C_\ell^{EE,\rm len}+N_\ell}\right)\left(\frac{C_\ell^{EE,\rm len}}{C_\ell^{BB,\rm len}+N_\ell}\right). \tag{28}$$

This equation holds for $L\geq 2$.[13] Since the quadratic estimator already resolves the spectrum at low L with good signal to noise, we can directly compare the result to Eq. (25), finding

$$g^{QE,\kappa}_{LM} \simeq (1.7\cdot 10^8)\kappa_{LM}\,, \tag{29}$$

for the same, deep configuration. This suggests, that at low L, the delensed noise mean-field piece is small. Only if the MAP method reduces the difference between the true and estimated convergence by a factor of about 70 compared to the QE estimate, does this mean-field become a bias on the reconstruction which may be taken into account. At higher L, it is irrelevant in any case. In practice, the mean-field is heavily dominated by masking, as well as other instrumental non-idealities that are either poorly known or more difficult to include in the likelihood model for practical reasons. Most relevant contributions of the mean-field peak at low L, but they contribute little to B-mode delensing. This motivates a simple baseline solution as follows (and comes at zero additional cost): we perform the iterations using a constant mean-field, calculated for example from quadratic estimators on a set of simulation with accurate noise maps and varying lensing fields. If necessary, after the iterative scheme has converged to an approximate MAP solution, $\hat{\boldsymbol{\alpha}}^{\rm MAP}$, we can correct for residual mean-field contamination using

$$\hat{\boldsymbol{\alpha}}^{\rm MAP} - \left\langle\hat{\boldsymbol{\alpha}}^{\rm MAP}\right\rangle_{\rm MC}\,, \tag{30}$$

instead of $\hat{\boldsymbol{\alpha}}^{\rm MAP}$ for the delensing step, where the average is built from the set of simulations for which the iterative solutions are constructed. Since these simulations have varying input lensing, Eq. (30) does not contain the $\boldsymbol{\alpha}$-induced part, which is neglected altogether in this approach. Secs. 2.8 and 4 contain explicit tests of this and of other options, confirming that the mean-field is not a major issue for the purpose of B-mode delensing.

2.6. *Construction of the B-polarization lensing templates*

We now discuss our construction of the lensing B-mode template from the reconstructed lensing and E-mode tracers.

At the end of the iterative process, we can calculate the prediction of the lensing-induced CMB polarization as follows. The Wiener-filtered E-mode from Eq. (19) (with, in that equation, $\boldsymbol{\alpha}=\boldsymbol{\alpha}^{\rm MAP}$) is our best estimate of the unlensed E-mode, which we write $\hat{E}^{\rm unl}$. We first build its Stokes polarization maps, ${}_2\hat{P}^{\rm unl}$, and then remap it,

$${}_2\hat{P}^{\rm len}(\hat{n}) = \left[\mathcal{D}_{\hat{\boldsymbol{\alpha}}^{\rm MAP}}\,{}_2\hat{P}^{\rm unl}\right](\hat{n})\,, \tag{31}$$

where

$${}_2\hat{P}^{\rm unl}(\hat{n}) = -\sum_{\ell m}\hat{E}^{\rm unl}_{\ell m}\,{}_2Y_{\ell m}(\hat{n}). \tag{32}$$

The B-mode of the deflected polarization ${}_2\hat{P}^{\rm len}$ of Eq. (31) is the B-mode lensing template, $\hat{B}^{\rm LT,MAP}$. In several parts of this work, we will use the following shorter notation for these operations:

$$\hat{B}^{\rm LT,MAP} \equiv \hat{\boldsymbol{\alpha}}^{\rm MAP}\circ\hat{E}^{\rm unl}\,. \tag{33}$$

We will often compare the performance of iterative delensing to that of quadratic estimator delensing. One possibility is to proceed in the same manner,

$$\hat{B}^{\rm LT,QE} \equiv \hat{\boldsymbol{\alpha}}^{\rm QE}\circ\hat{E}^{\rm len}\,, \tag{34}$$

where the E-mode map $\hat{E}^{\rm len}$ is the Wiener-filtered lensed CMB. There is a known subtelty, though, that when combining the *lensed* E-mode with the quadratic estimator, it is more powerful to apply the remapping to leading order in the deflection, rather than exactly (Baleato Lizancos et al. 2021a). This is due to a beneficial cancellation of non-perturbative effects, which does not occur anymore when the exact remapping Eq. (31) is used on the lensed E-mode. Hence, for quadratic estimator delensing, we also consider the perturbative action of lensing (given by Eq. (6)) on the E-mode.

There is less flexibility in the choice of the E-template in the case of the iterative lensing solution: by design of the lensing reconstruction process, the reconstruction noise in $\hat{\boldsymbol{\alpha}}^{\rm MAP}$ and $\hat{E}^{\rm unl}$ have a strong statistical dependence: when the CMB likelihood model is inaccurate, both $\hat{E}^{\rm unl}$ and $\hat{\boldsymbol{\alpha}}^{\rm MAP}$ can be somewhat wrong, but still manage to provide a reliable B-mode template. In this case, independent adjustments to $\hat{E}^{\rm unl}$ or $\hat{\boldsymbol{\alpha}}^{\rm MAP}$ can break this dependence and degrade the delensing performance. The construction described by Eq. (31) consistently gives the best results.

---

[13] The polarization estimator dipole response is much weaker, since a lensing dipole does not produce any shear.

2.7. *Low-ℓ B-mode deprojection*

The search for primordial B-mode power focuses on large, degree scales, where B-lensing is weaker and the primordial signal is expected to be strongest. To remove B-lensing, we construct a B-lensing template from high-resolution data, which provides the small-scale information needed for reconstructing the deflection field and primordial E modes. This template is then applied to data measuring the largest CMB scales, effectively reducing variance from non-primordial signals. Although large-scale information from high-resolution data could technically improve the lensing template, this approach generally introduces a statistical dependence between the deflection field's reconstruction noise and the delensed map (Teng et al. 2011; Carron et al. 2017; Namikawa 2017; Baleato Lizancos et al. 2021b). At the QE level, this appears notably as a very strong disconnected 4-point function, given by the Gaussian pairings of

$$\hat{C}_\ell^{B^{\rm dat} B^{\rm LT,QE}} \ni B^{\rm dat} \cdot \hat{\boldsymbol{\alpha}}^{\rm QE}(E^{\rm dat}, B^{\rm dat}) \circ \hat{E}^{\rm len} \,, \qquad (35)$$

with a tendency to decrease both the B-mode power and its variance, however, it is unrelated to true delensing. At low noise levels, a 6-point function also becomes relevant, as studied in some detail in (Namikawa 2017; Baleato Lizancos et al. 2021b), and predictions of the spectra are rendered much more complicated, particularly in the presence of tensor modes for which $B^{\rm unl} \neq 0$.

Less is known in the case of the iterative estimator, but this bias is even larger there (as we show later in Sec. 2.8). The lensing reconstruction gives little weight to these large scale B-modes. Hence, a simple solution which we adopt here is to exclude these overlapping modes from the beginning. We proceed here in a quite careful manner with the curved-sky version of the 'overlapping B-mode deprojection' (OBD) technique introduced in (Adachi et al. 2020). We exclude modes as follows: let $\mathcal{P}_{\ell m}(\hat{n}_i)$ be the Stokes polarization pattern on pixel $\hat{n}_i$ produced by a CMB-sky $B_{\ell m}$-mode. We assign a high noise level to a set of these patterns by augmenting the noise matrix in the following way, $N \to N + \mathcal{P}^\dagger \sigma^2 \mathcal{P}$, where σ^2 is a diagonal matrix of dimension the number of patterns considered, with values on the diagonal the assigned noise variances of the patterns. According to the Woodbury matrix identity (Hager 1989), this has the effect of replacing the inverse noise matrix model N^{-1} by

$$N^{-1} - N^{-1}\mathcal{P}^\dagger \left[\frac{1}{\sigma^2} + \mathcal{P}N^{-1}\mathcal{P}^\dagger\right]^{-1} \mathcal{P}N^{-1} \,. \qquad (36)$$

In this way, all patterns included in $\mathcal{P}$ are perfectly masked from the analysis in the limit of infinite σ^2. A high but finite value of σ^2 is often required for practical reasons but works equally well. We deproject all modes $B_{\ell m}$ with $2 \leq \ell \leq 200$, for a tiny loss of signal to noise[14], using $\sigma^2 = \left(10^3 \mu{\rm K}\right)^2$. The matrix inside the square brackets in this expression is therefore a dense matrix with about ℓ_B^4 elements. The matrix is precomputed and stored in memory during the reconstruction. An alternative to this computationally rather expensive calculation of the OBD-matrix would be to filter the data maps by setting the transfer function for these modes to zero. On the full sky with homogeneous noise, the two procedures are exactly equivalent. On the masked sky, it is less straightforward to cleanly exclude this set of sky modes, potentially sourcing residual internal delensing biases. We have not pursued the alternative approach in this work and leave it to a future study.

2.8. *Idealized full sky delensing*

As a first step, we demonstrate the effectiveness of our lensing reconstruction method and of the B-mode delensing procedure on the full-sky. We generate full-sky curved-sky lensed CMB data with no primordial B-modes, and use different ϕ and E-modes for each realization, synthesized from a standard ΛCDM cosmology. The simulations use an idealized, low-noise, high-resolution configuration: we add an isotropic Gaussian noise of $\Delta_{\rm P} = 0.5\ \mu$K arcmin and convolve the maps with a beam FWHM of 1 arcmin. We reconstruct the lensing potential from the E and B maps using both QE and MAP methods. We restrict to the multipole range $\ell \in [2, 3000]$ for the E-modes and $\ell \in [200, 3000]$ for the B-modes to avoid internal delensing bias, as discussed in Sec. 2.7.

For the QE B-mode template estimation, we compare two approaches, as described in Sec. 2.6. The first approach is a perturbative remapping of the Wiener-filtered and lensed E-mode at first order, following Eq. (6). The deflection field is taken as the Wiener-filtered QE estimate,

$$\kappa_L^{\rm WF} = \hat{\kappa}_L^{QE} \, \frac{C_L^{\kappa\kappa}}{C_L^{\kappa\kappa} + N_L^{(0),\kappa{\rm QE}}} \,, \qquad (37)$$

with $N_L^{(0),\kappa{\rm QE}}$ the dominant term in the variance of the reconstructed QE power spectrum, arising from the Gaussian contractions of the four-point function of the lensed CMB fields.[15] The second B-mode template is estimated with the non-perturbative lensing remapping of Eq. (31), where the unlensed polarization field is given by Eq. (19) and the deflection field is the QE field. In practice, this corresponds to the first iteration of the MAP algorithm. We found that both approaches give a similar residual B-mode lensing power. As shown in (Baleato Lizancos et al. 2021a), the perturbative remapping of the observed (Wiener-filtered) polarization field has a similar residual B-mode lensing power as the non-perturbative remapping of the unlensed polarization.

For the MAP approach, the B-mode template is derived using Eq. (31), hence, by remapping the unlensed polariza-

---

[14] In the CMB-S4 South Pole deep configuration of this paper, deprojection of these modes causes the predicted residual B power amplitude after delensing to rise formally to 0.053 from 0.052

[15] We find that including $N_L^{(1),\kappa{\rm QE}}$ to the Wiener filter reduces the residual lensing amplitude by about 2 percent.

tion field by the MAP lensing estimate.[16] As discussed in Sec. 2.5, the QE mean-field is zero, but the MAP mean-field is not. We estimate the mean-field at first order in the lensing field estimate, as described in the appendix B of (Carron & Lewis 2017). Figure 3 shows the power spectrum of the lensing B-mode (orange), and residual lensing B-modes after the subtraction of the QE (dark blue), and MAP (light blue) B-mode templates. The QE method removes significant power of the lensed B-modes, thereby decreasing the mean C_ℓ^{BB} amplitude in the range $\ell \in [2, 200]$ from $2.1 \times 10^{-6} \mu\mathrm{K}^2$ to $4.8 \times 10^{-7} \mu\mathrm{K}^2$. The MAP method reduces it further, leaving about $1.3 \times 10^{-7} \mu\mathrm{K}^2$ of the B power. For illustration purposes, we also estimate the MAP B-mode template for which the B-modes $\ell \leq 200$ are included in the lensing reconstruction (solid gray), and additionally show the power spectrum of the simulation's white noise (dashed gray). Including these modes results in an internal delensing bias (see Eq. (35)), as can clearly be seen: the residual B-mode power is below the noise level of the polarization maps. In the standard analysis, where we reconstruct the lensing potential from the B-modes with $\ell \in [200, 3000]$, this internal delensing bias appears at scales $\ell > 200$ for both the QE and the MAP methods, albeit it is much smaller in the former case, as we can see in Fig. 3.

We can compare our results with the prediction of the residual lensing B-mode power. To build predictions we generally follow the procedure detailed in (Legrand & Carron 2022), a variant on the original procedure for the EB-estimator put forward by (Smith et al. 2007). The MAP solution is analytically intractable even under idealized conditions, and this simplified procedure is not free from some ambiguities, as we now discuss.

Predictions are estimated by computing the delensed power spectrum $C_\ell^{BB,\mathrm{delens}}$ from the unlensed E spectrum $C_\ell^{EE,\mathrm{unl}}$ and a lensing power spectrum $C_L^{\kappa\kappa,\mathrm{delens}}$ reduced by some fraction:

$$C_L^{\kappa\kappa,\mathrm{delens}} = (1 - \epsilon_L)\, C_L^{\kappa\kappa} \ , \tag{38}$$

where the delensing efficiency ϵ_L is

$$\epsilon_L = \frac{C_L^{\kappa\kappa}}{C_L^{\kappa\kappa} + N_L^{\kappa\kappa}} \ . \tag{39}$$

This variance of the reconstructed lensing potential $N_L^{\kappa\kappa}$ contains the Gaussian contribution $N_L^{(0)}$, as well as terms higher-order in $C_L^{\kappa\kappa}$. For the QE, we found that the $N_L^{(1)}$ term (the secondary connected contractions at first order in $C_L^{\kappa\kappa}$) has an impact of roughly 20% on the amplitude of the predicted residual B-mode, from $3.9 \times 10^{-7} \mu\mathrm{K}^2$ when $N_L^{\kappa\kappa} = N_L^{(0)}$ to $4.8 \times 10^{-7} \mu\mathrm{K}^2$ when $N_L^{\kappa\kappa} = N_L^{(0)} + N_L^{(1)}$. The case including $N_L^{(1)}$ provides a better fit to the observed delensed power. For the MAP prediction, we compute theses lensing spectrum biases iteratively following (Legrand & Carron 2022): starting with the noise of the QE reconstruction, $N_L^{\kappa\kappa}$, we compute the delensed spectrum of Eq. 38, and the corresponding partially lensed E power spectrum. These spectra are then inserted as the weights and lensing response in the analytical expression of $N_L^{\kappa\kappa}$. Finally, we obtain a new delensing efficiency ϵ_L and repeat the calculation with an iterative procedure until convergence. Contrary to (Smith et al. 2012), we do not include the imperfect knowledge of the E-mode power in the iteration procedure. We found that it does not impact the predictions, due to the low noise level in polarization considered here. It appears that the prediction of the MAP delensing residual provides a better fit when we only include the $N_L^{(0)}$ bias in the iterations. By including the $N_L^{(1)}$ bias in all iteration steps we obtain a mean B-mode amplitude of $1.6\times10^{-7}\ \mu\mathrm{K}^2$ across the relevant scales. However, not including it gives a mean B-mode amplitude of $1.3\times10^{-7}\ \mu\mathrm{K}^2$, which matches the observed delensing residual from our simulation. Nevertheless, we do not have an analytical argument for whether or not the $N_L^{(1)}$ bias should be taken into account for the lensing power spectrum noise of the MAP algorithm. We take this as a measure of our uncertainty in the prediction, and show the predicted delensing amplitudes as shaded areas in Fig. 3. The lower limit corresponds to $N_L^{\kappa\kappa} = N_L^{(0)}$, while the upper limit corresponds to $N_L^{\kappa\kappa} = N_L^{(0)} + N_L^{(1)}$.

The reconstruction described above used a perturbative approximation to the delensed noise mean-field, which is the only source of mean-field in this idealized reconstruction. We now test the importance of this term explicitly. Figure 4 shows the spectrum of the quadratic piece $g^{\mathrm{QD}}_{\boldsymbol{\alpha}}$ of the gradients at the starting point (the quadratic estimator, in blue) and at iteration 15 in orange, where we often stop the reconstruction for CMB-S4-like configurations. We show results for $L \geq 2$ because the lensing dipole behaves differently.[17] The displayed spectra are the gradients normalized by our prediction of the iterative estimator response, $\mathcal{R}_L^{\mathrm{MAP}}$,

$$\frac{g^{\mathrm{QD}}_{\boldsymbol{\alpha},LM}}{\mathcal{R}_L^{\mathrm{MAP}}} \ , \tag{40}$$

such that, very crudely, the normalized map corresponds to the expected Newton increment to the lensing map whenever the prior is irrelevant. The prior term is shown in purple. The two gradients are mostly in equilibrium at $L \gtrsim 2000$.[18] As can be seen in the Figure, for $L \lesssim 1000$, the prior is

---

[16] We find that using the perturbative remapping, together with the lensed E-map, results in an almost identical MAP residual lensing B-mode spectrum, in agreement to the findings discussed in (Baleato Lizancos et al. 2021a)

[17] On the one hand, the lensing dipole has much higher noise than for $L \geq 2$, because it is insensitive to the shear part of the signal. On the other hand, the dipole receives a strong contribution from the aberration of the CMB caused by our motion relative to the CMB frame (Aghanim et al. 2014).

[18] In this reconstruction, the size of the Newton step in the iterative procedure was suppressed for $L \geq 1000$ compared to smaller L's, in order to avoid too large steps. This causes the feature seen at $L \sim 1000$ in the orange line.

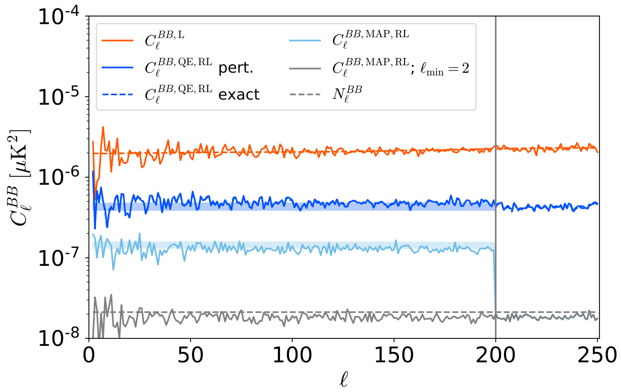

Figure 3. Residual B-mode power after subtracting the B-template estimated with the QE (blue line) or with the MAP (light blue line), in an idealized full-sky reconstruction, on the scales relevant for inference on r. The dashed blue line overlaps with the solid blue and is therefore invisible; it corresponds to using the lensing remapping of the unlensed E obtained at the first iteration of the MAP and it matches almost exactly the perturbative remapping of the lensed E-modes. Colored bands show the corresponding fiducial delensed CMB spectra (see discussion in the text). The orange line is the B power of the map before template subtraction (including noise) and the dashed orange line is the fiducial lensed B power including noise. The gray dashed line shows the instrumental polarization noise power. To calculate the light blue line, the B-modes below $\ell = 200$ are not used to estimate the B-templates. This avoids internal delensing bias; the result of not treating internal delensing bias is illustrated by the gray line, in which case the B-template is estimated using all modes from $\ell = 2$. Interpreted naively, the line shows unrealistically efficient delensing with delensed B-mode power slightly below the instrumental noise power of the map. The light blue line follows the gray line for $\ell \geq 200$.

substantially smaller, which suggests that there is still some amount of information in the likelihood that can be collected, although it is small.

The estimated mean-field for iteration step 15 is shown in green. To compute it, we used a trick presented in the appendix 2 of (Carron & Lewis 2017): instead of $\langle g^{\rm QD}_{\boldsymbol\alpha}\rangle$ we calculate the average of a similar map,

$$ {}_1\hat g^{\rm MF}_{\boldsymbol\alpha}(\hat n) = -\sum_{s=\pm2}\left\langle {}_{-s}\bar p_{\boldsymbol\alpha}(\hat n)\left[\mathcal{D}_{\boldsymbol\alpha}\partial^+_s p^{\rm WF}_{\boldsymbol\alpha}\right](\hat n)\right\rangle , \qquad (41)$$

where $\bar p_{\boldsymbol\alpha} = \mathcal{B}^\dagger x$, $p^{\rm WF}_{\boldsymbol\alpha} = C^{EE,\rm unl}\mathcal{D}^\dagger_{\boldsymbol\alpha}\mathcal{B}^\dagger{\rm Cov}^{-1}_{\boldsymbol\alpha}x$, x is a map of unit variance, independent Gaussian variables, and s is the spin of the field. Assuming the CMB likelihood model correctly describes the data, the average, Eq. (41), is the same as $\langle g^{\rm QD}_{\boldsymbol\alpha}\rangle$, but requires orders of magnitude less Monte-Carlo simulations for the same precision of the result. This is because the $N^{(0)}$ noise of this quadratic estimator is much lower than that of the one with the correct weights: see appendix 2 of (Carron & Lewis 2017) for more details. As derived in Sec. 2.5, the mean-field shape follows that of the convergence (shown in black). At the lowest multipoles, it appears negative, but this may be attributed to residual Monte-Carlo noise in that range. The mean-field remains much smaller than the quadratic gradient piece. Only at iteration 35 (red) does the latter reach an amplitude comparable to the former. Between iterations 15 and 35, the delensing efficiency has hardly grown, from 94.5% to 95.0%.

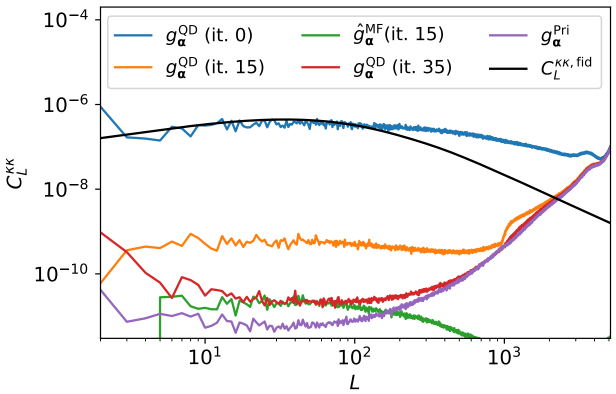

Figure 4. Comparison of normalized CMB likelihood gradients for a full-sky idealized reconstruction. The blue and orange curves show the quadratic piece spectrum at iteration 0 (the quadratic estimator) and 15, where we often stop the iterations. The prior term at the same iteration is shown in purple. The delensed noise mean-field term, shown in green and obtained with the steps given in the text, is very small and has negligible impact on the quality of the reconstruction. We also show the quadratic piece after 35 iterations in red. Only then does it become comparable to the mean-field on large-scales. As discussed in the text, the improvement achieved in going from 15 to 35 iterations is tiny, however.

Finally, using Eq.(41), it is possible to run the entire full-sky reconstruction process using a single Monte-Carlo simulation to obtain the mean-field. As expected from Fig. 4, we find no significant difference. All of this demonstrates that in general the delensed-noise part of the mean-field can be safely neglected in the reconstruction from polarization data.

2.9. *Idealized joint ϕ-Ω lensing reconstruction*

Even though the lensing curl vanishes to first order for density perturbations,[19] it is well known that the combined effect of two shearing lenses (with misaligned shears) does induce field rotation. The leading contribution to quadratic order can be written as,

$$\omega(\hat n) = -4\int_0^{\chi_*} d\chi\left(\frac{\chi_*-\chi}{\chi\chi_*}\right)\int_0^{\chi} d\chi'\left(\frac{\chi-\chi'}{\chi'\chi}\right) \cdot\left[\gamma_1(\hat n,\chi)\gamma_2(\hat n,\chi')-\gamma_2(\hat n,\chi)\gamma_1(\hat n,\chi')\right] , \qquad (42)$$

---

[19] We ignore here non-scalar contributions, such as lensing by gravitational waves, which are expected to result in negligible signals (Cooray et al. 2005).

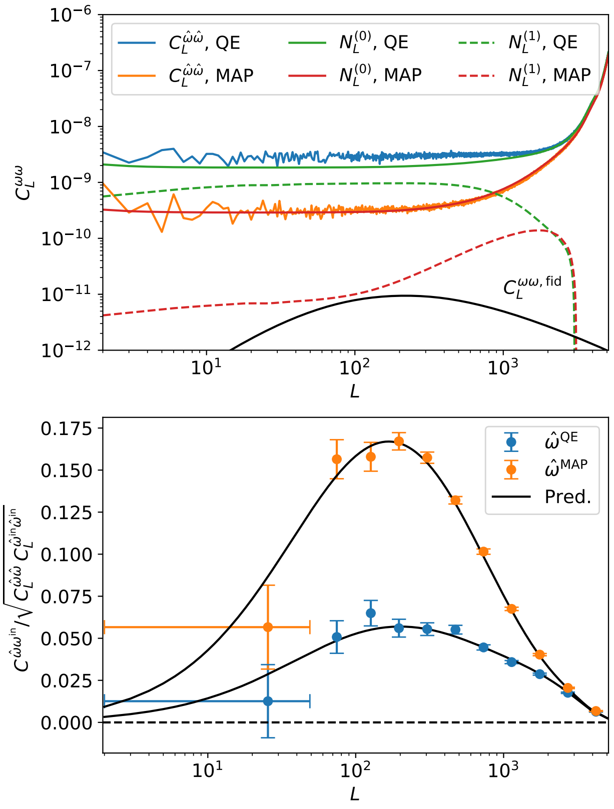

Figure 5. Lensing field rotation reconstruction results from our full-sky joint lensing gradient and curl reconstruction. *Upper panel:* Spectra of the curl quadratic estimator and iterative reconstruction are shown in blue and orange, after rescaling by the predicted Wiener-filter amplitude, as described in the text. The predicted $N^{(0)}$ biases are shown in green and red. The black line shows the field rotation power spectrum in our fiducial cosmology. The dashed lines show the corresponding κ-induced $N^{(1)}$ lensing biases (the ω-induced $N^{(1)}$ are much smaller and negligible), which is mostly relevant in the QE case. *Lower panel:* Cross-correlation coefficient of this reconstruction to the input curl modes, together with the predictions built from the $N^{(0)}$ and $N^{(1)}$ biases. Error bars are obtained from the scatter within bins assuming independent estimates at each and every L. The auto-spectrum detection S/N jumps by a factor of 5 with the iterative solution.

and is referred to as 'lens-lens coupling', see (Hirata & Seljak 2003a) or (Pratten & Lewis 2016) for recent discussions. Here, ${}_2\gamma(\hat n,\chi)$ is the spin-2 shear as defined in the usual way (see Eq. (5)), but acting on the Weyl potential along the ray. At the noise levels of CMB-S4, this lensing curl mode can be detected in cross-correlation to external data (Robertson & Lewis 2023). In this section, we test the impact of the expected curl signal by performing a joint reconstruction of the lensing gradient and curl modes. This will allow us to confirm that the lensing curl is not a source of worry for CMB-S4 delensing. Although this joint analysis doubles the number of modes to be reconstructed simultaneously, we note that this is not expected to significantly degrade the quality of the gradient reconstruction: in the absence of a parity-violating C_ℓ^{EB} or C_ℓ^{TB} signal, the lensing gradient and curl CMB quadratic reconstruction share no leading Gaussian covariance[20] $N^{(0)}$ (Namikawa et al. 2012). It is true though that each component (here convergence and curl) induces a bias of $N^{(1)}$-type (Kesden et al. 2002) on the lensing spectrum of the other component,[21] but this effect is smaller than their individual $N^{(0)}$ noise variance levels.

For the test conducted in the following, we add a Gaussian realization of the post-Born lensing curl component to a simulation. This Gaussian realization comes from the Eq. (42)-induced power spectrum calculated by CAMB,[22] and is shown as the black line in Fig. 5. The lensing curl is second order in the scalar perturbations. A consistent framework would require the inclusion of the corresponding post-Born corrections to the lensing gradient potential (this is, however, smaller than 0.2% and we can therefore neglect it here), as well as the usage of a non-Gaussian simulated lensing potentials. However, the latter goes beyond the scope of this paper and we stick to Gaussian realizations, which still provides an important test of the joint reconstruction.

We jointly reconstruct the convergence κ and field rotation ω on the full sky by starting from their Wiener-filtered quadratic estimate.[23] The reconstruction is done for $1 \leq L \leq 5120$ for κ and $2 \leq L \leq 5120$ for ω. The same lensing curl spectrum used for the simulation generation is used as the prior on the curl modes for the reconstruction in Eq. (11). We pick a homogeneous polarization noise level of 0.5μK-arcmin, use modes of the data up to $\ell_{\rm max} = 4096$, and reconstruct the unlensed CMB E-mode up to $\ell_{\rm max,unl} = 5120$.

Figure 5 shows in the upper panel the spectrum of the reconstructed curl in orange, with Blue being QE for comparison. We have rescaled the MAP spectrum by the isotropic normalization, $(\mathcal{W}_L^\omega)^2$, with

$$\mathcal{W}_L^\omega = \frac{C_L^{\omega\omega,\rm fid}}{C_L^{\omega\omega,\rm fid} + N_L^{(0),\omega\rm MAP}}\,. \tag{43}$$

Regarding the lensing gradient, (Legrand & Carron 2022) demonstrates that the approximation of this normalization, and applied to the MAP solution, is accurate within a few percent. This appears to also be a good fit for the curl mode normalization, and the resulting spectrum can nicely be described by the expected reconstruction noise $N_L^{(0),\omega\rm MAP}$, as shown in red. In this joint reconstruction, there are two types of $N^{(1)}$ bias: the standard one induced by κ, and the one induced by the non-zero ω. The latter, however, is much

---

[20] Contrary claims are given, for example in (Cooray et al. 2005; Pratten & Lewis 2016). However, this is likely due to small numerical inaccuracies in (Cooray et al. 2005).

[21] See for example appendix A of (Aghanim et al. 2020a) for the general form of $N^{(1)}$-type contributions.

[22] https://camb.readthedocs.io/en/latest/postborn.html

[23] We use the same `Plancklens` package to build the curl-mode quadratic estimators

smaller than the κ-induced bias and is negligible. For the QE, the $N^{(1)}$ bias (dashed green) plays a relevant role in the auto-spectrum, but is strongly suppressed for the MAP solution (dashed red), since both the lensing B power and lensing power are reduced by large factors. In the lower panel of Fig. 5, the points show binned estimates of the empirical cross-correlation coefficient to the input lensing curl modes,

$$\rho_L^{\omega} \equiv \frac{C_L^{\hat{\omega}\omega^{\rm in}}}{\sqrt{C_L^{\hat{\omega}\hat{\omega}} C_L^{\omega^{\rm in}\omega^{\rm in}}}} \, . \tag{44}$$

The error bars were calculated from the empirical scatter within each bin, assuming statistically independent fluctuations across multipoles. The black lines show the predictions

$$\rho_L^{\omega,\rm pred} = \left(\frac{C_L^{\omega\omega,\rm fid}}{C_L^{\omega\omega,\rm fid} + N_L^{(0),\omega} + N_L^{(1),\omega}} \right)^{1/2} , \tag{45}$$

for the QE and MAP. The improvement brought by the iterative method is clearly substantial: the full-sky QE signal to noise,

$$(S/N)^2 \equiv \frac{1}{2} \sum_L (2L+1) \left(\rho_L^{\omega,\rm pred} \right)^2 , \tag{46}$$

on the auto-spectrum detection is $S/N \sim 2.1$ for this configuration, but reaches $S/N \sim 10.4$ for the MAP method. Of course, at the polarization noise level of 0.5μK-arcmin, CMB-S4 would probably observe only a small fraction of the sky, so that a purely internal detection in this configuration still remains challenging.

Finally, the impact on the residual lensing B-mode amplitude is small. Comparing delensing efficiencies obtained on this same simulation with and without curl modes, we find that the amplitude is reduced by 0.4% compared to the former case. This allows us to conclude that our reconstruction tools perform as expected, even in the presence of the lensing curl. For simplicity, the curl mode is set to zero for the rest of the paper.

2.10. *A faster scheme under idealized conditions*

The construction of the Wiener-filtered delensed E-mode map, as given by Eq. (19), is generically the most expensive step of the entire lensing reconstruction procedure. This equation may be rewritten

$$E^{\rm WF}_{\boldsymbol{\alpha}} = C^{EE,\rm unl} \left[C^{EE,\rm unl} + N_{\boldsymbol{\alpha}} \right]^{-1} E^{\rm del}_{\boldsymbol{\alpha}} \, , \tag{47}$$

where $E^{\rm del}_{\boldsymbol{\alpha}}$ is the delensed CMB,

$$E^{\rm del}_{\boldsymbol{\alpha}} \equiv N_{\boldsymbol{\alpha}} \, {}_2\mathcal{Y}^\dagger \mathcal{D}^\dagger_{\boldsymbol{\alpha}} \mathcal{B}^\dagger N^{-1} X^{\rm dat} , \tag{48}$$

and the prefactor to it is the Wiener-filtering matrix. The interpretation of $E^{\rm del}_{\boldsymbol{\alpha}}$ as the delensed E-mode can be justified by noting that the signal part of $E^{\rm del}_{\boldsymbol{\alpha}}$ is the unlensed CMB, provided that the true transfer function matches that of the likelihood model and that $\boldsymbol{\alpha}$ is the true lensing. While still tractable, these optimized weights and filters are not trivial to apply, even under idealized conditions, because $\boldsymbol{\alpha}$ breaks the isotropy of the delensed noise, making $N_{\boldsymbol{\alpha}}$ a dense matrix. However, this is a small effect, and, under idealized conditions, one can still expect to recover the bulk of the signal with suitable approximations using much faster filtering (as suggested by the pioneering papers (Hirata & Seljak 2003b,a)). At each iteration, it is quite natural to replace Eq. (47) with

$$\left[E^{\rm WF}_{\boldsymbol{\alpha}} \right]_{\ell m} \to \frac{C_\ell^{EE,\rm unl}}{C_\ell^{EE,\rm unl} + N_\ell^{EE}} \left[{}_2\mathcal{Y}^{-1} \mathcal{D}^{-1}_{\boldsymbol{\alpha}} \mathcal{B}^{-1} X^{\rm dat} \right]_{\ell m} , \tag{49}$$

where the inverse beam and transfer operation, $\mathcal{B}^{-1}$, is a simple rescaling by the isotropic beam, $1/b_\ell$, in harmonic space. N_ℓ^{EE} is the isotropic instrument noise, and ${}_2\mathcal{Y}^{-1}$ is the projection onto the E-mode. There is no conjugate gradient inversion to perform here, which makes this version significantly faster, numerically.

We tested this approximation in the same configuration as in the previous section. At each step, the quadratic gradient calculation takes the slightly suboptimal Wiener-filtered E-mode as input, but is otherwise unchanged. We found that the first iteration worsened the solution significantly below $L = 30$, which we solved simply by fixing these large lenses to the QE solution. On the other hand, smaller scales appear to be recovered almost equally well. After 15 iterations, we found that the cross-correlation coefficient is within 1% of the solution without these approximations for $100 \leq L \leq 5120$. While we do not pursue this further here, it clearly means that the simplified gradient calculation can be valuable for theoretical investigations using idealized conditions. Work is ongoing to understand to what extent faster yet efficient methods can also be devised in realistic configurations.

3. SIMULATIONS

This section discusses the individual components of the simulated maps and how they are generated and combined.

3.1. *CMB signal simulations*

The simulations used in this work are generated using the lensed CMB maps from the *Planck* FFP10 simulation suite as input. These CMB maps, in harmonic space, were built by one of the corresponding authors in preparation for the 2018 *Planck* release, and were later made public on NERSC and used by the CMB-S4 collaboration. The maps contain a series of small lensing-related defects (and another one was also introduced at the very beginning of this work), which are completely innocuous at most noise levels and on large scales relevant for r-inference. However, in precise tests performed at low polarization noise levels, these defects occasionally caused troubles. In the following, we provide a record of these issues along with the relevant lensing properties.

- For the simulations, the rotation of the polarization that is required to match the change in reference axes caused by the sky curvature (the phase in Eq.(2)) was not performed. As a consequence, the polarization right at the poles is just wrong, since there this phase is not small. This is something to which our reconstruction pipeline reacts (for noise below the μK-arcmin level) by assigning absurd values to the lensing potentials at the poles, or failing to converge altogether. However, only a few arcmin away from the poles, we find that this is not an issue any longer.
- Due to a bug in the interpolating code that was used at the time, a set of about ~ 20 pixels at $N_{\rm side} = 2048$ that are located in the southern hemisphere, well separated from one another, and close to the meridian defined by $\varphi = 0°$, were just wrong.
- In addition to the massive CMB dipole, our velocity relative to the CMB frame also causes more subtle modulation and aberration of the anisotropies along the velocity direction (Aghanim et al. 2014). This aberration results in an almost perfect dipole lensing signal, approximately 5 times larger than the expected ΛCDM lensing potential dipole, and is easily detected with reasonably wide sky coverage (Aghanim et al. 2014). Utilized for this were CMB frequency maps that contain the frequency-independent part of the modulation plus the aberration, for all *Planck* channels.
- After synthesizing the FFP10 lensed $E_{\ell m}$ and $B_{\ell m}$ into real-space maps, the CMB-S4 team adds foreground maps defined in equatorial coordinates to these maps. However, the *Planck* team worked in Galactic coordinates, and this leads to the aberration signal mentioned just above having a inconsistent direction to the pair of statistically anisotropic galactic foreground models that we will be using.

3.2. *Instrument noise simulations*

Noise is generated in the same manner as was done in our previous CMB-S4 r forecast paper (Abazajian et al. 2022). Uniform full sky realizations of noise are generated and then divided by the “Pole Deep” and “Chile Full” hits patterns, which are shown in Figure 7 of that paper. The normalization of the noise is adjusted such that when re-analyzing these realizations with inverse-variance weighting one obtains the power spectrum noise levels given in Table 1. These hits pattern come from full-blown time-ordered-data level scanning simulations for the CMB-S4 small aperture telescopes (SATs). The noise levels in Table 1 come from a preliminary set of CMB-S4 measurement requirements. For the present we assume that the large aperture telescopes (LATs) have the same hits pattern—as could be (approximately) arranged in practice. Excess low frequency $1/f$ noise is included with the knee and slope parameters given in the table. These parameters are derived from BICEP/Keck and SPT data (Abazajian et al. 2022). Map noise contours are shown in Figs. 9 and 18. Experiments at the South-Pole, or more generally ground-based experiments, can suffer from large-scale noise that is highly anisotropic in harmonic space, sourced by atmospheric noise and the scanning strategy. This type of anisotropy is not included in this series of simulations and the study of its impact on the iterative lensing map is left for future work.

3.3. *Foreground models*

We run the curved-sky lensing reconstruction algorithm on three different simulation sets with varying degrees of foreground complexity. We refer to the three sets containing foreground and CMB only as *Sky Model 00, 07, 09*, and as the *Simulation Set* ($\rm M_{00}$, $\rm M_{07}$, $\rm M_{09}$), if it includes all components (foregrounds, CMB, and noise). In our previous paper (Abazajian et al. 2022), more foreground models were studied. However, the other models did not extend to the small scales relevant to the lensing reconstruction.

$\rm M_{00}$ includes the simplest possible (and completely unrealistic) foreground model; Gaussian realizations of synchrotron and dust with uniform amplitude across the full sky. For the synchrotron A_s, and dust amplitude A_d, we use $A_s = 3.8\ \mu{\rm K}^2$ and $A_d = 4.25\ \mu{\rm K}^2$, for the corresponding α parameters, we use $\alpha_s = -0.6$ and $\alpha_d = -0.4$, and for the β parameters $\beta_s = -3.1$ and $\beta_d = 1.6$. For the temperature of the dust T_d we choose $T_d = 19.6$ K (see (Ade et al. 2018) for a description of the foreground emission laws). There are good reasons to include such a model in this study; the mean-field from this simulation set will not pick up foreground residuals (other than foreground inhomogeneities arising from the magnifying and demagnifying of the delensing procedure, analogous to the induced noise inhomogeneity, as discussed in Sec. 2.3), which will be useful for assessing the impact of the mean-field on our results.

$\rm M_{07}$ includes an amplitude-modulated foreground model; we take the Gaussian fields from $\rm M_{00}$ and modulate their amplitude across the sky according to a map of degree-scale B-mode power measured from small patches of *Planck*'s 353 GHz map in a similar manner to Fig. 8 of (Adam et al. 2016).

$\rm M_{09}$ includes a realization of the multi-layer Vansyngel model (Vansyngel et al. 2017). Each layer has the same intensity (constrained by the Planck intensity map), but different magnetic field realizations. It produces Q and U maps by integrating along the line of sight over these multiple layers of magnetic fields.

The complete simulation sets are generated as follows. For each, we sum the noise, lensed CMB, and the respective foreground for each simulated frequency band. All components are generated in equatorial coordinates, with the exception of the FFP10 lensed CMB. This is not expected to pose an issue for lensing reconstruction, as discussed in the previous section. The component $\rm fg_{09}$ is rotated to equatorial coordinates before adding it. Each realization contains a different realization of the lensed CMB and noise. For $\rm M_{00}$ and $\rm M_{07}$, each realization contains a different foreground realization, with the amplitude modulation of $\rm fg_{07}$ being the

Table 1. Simulated SAT and LAT map noise levels.

	Band	Beam	EE			BB		
			white	ℓ-knee	slope	white	ℓ-knee	slope
	[GHz]	[arcmin]	[μK-arcmin]			[μK-arcmin]		
SAT	30	72.8	3.74	60	-2.2	3.53	60	-1.7
	40	72.8	4.73	60	-2.2	4.46	60	-1.7
	85	25.5	0.93	60	-2.2	0.88	60	-1.7
	95	22.7	0.82	60	-2.2	0.78	60	-1.7
	145	25.5	1.25	65	-3.1	1.23	60	-3.0
	155	22.7	1.34	65	-3.1	1.34	60	-3.0
	220	13.0	3.48	65	-3.1	3.48	60	-3.0
	270	13.0	8.08	65	-3.1	5.97	60	-3.0
LAT	20	11.0	13.16	150	-2.7	13.16	150	-2.7
	30	7.3	6.50	150	-2.7	6.50	150	-2.7
	40	5.5	4.15	150	-2.7	4.15	150	-2.7
	95	2.3	0.63	150	-2.6	0.63	150	-2.6
	145	1.5	0.59	200	-2.2	0.59	200	-2.2
	220	1.0	1.83	200	-2.2	1.83	200	-2.2
	270	0.8	4.34	200	-2.2	4.34	200	-2.2

same in each realization. For M_{09} there is only one realization; each and every map therefore contains the exact same foreground component. The foreground models do not contain extragalactic foregrounds, and our analysis focuses on polarization data only. Extragalactic polarized foreground contamination is generally expected to be weak and not to affect the lensing reconstruction significantly, dominated by the EB signal. Source masking can potentially bias the lensing reconstruction if the mask cross-correlates substantially to the lensing signal. However, this has also been shown to be small (Lembo et al. 2022) in our configuration. Therefore, we do not include point source masks in our analysis.

4. RESULTS

The discussion on the full pipeline of the r-analysis and the analysis of the robustness of our iterative curved-sky lensing reconstruction is split as follows: in Sec. 4.1, we discuss the component-separation technique used to produce a clean CMB map from the LAT simulated maps. In Sec. 4.2 and 4.3, we discuss the performance of the lensing reconstruction and the generation of the B-lensing template to obtain the best estimate of the B-lensing signal. In Sec. 4.4, we discuss the construction of the likelihood for r and resulting constraints.

4.1. *component-separation*

We apply a harmonic internal linear combination (ILC) on the LAT frequency maps (Tegmark et al. 2003). We use for this all frequencies with the exception of the 20GHz channel, which is placed on the LAT but has the primary function of aiding in the constraint of the synchrotron component in the likelihood for r; the beam of the 20GHz channel is too large to make any meaningful contribution for the purpose of lensing reconstruction and is therefore not taken into account for the LAT component separation. We obtain the weights ω_ℓ analytically by creating a covariance matrix R_ℓ and applying the formula.

$$\omega_\ell = \frac{a^T R_\ell^{-1}}{a^T R_\ell^{-1} a}, \tag{50}$$

where a is the emission law of the CMB in the chosen units. The covariance matrix R_ℓ is composed of input theoretical spectra and cross-spectra of simulation-set M_{00} maps. We do this for the E-mode and B-mode independently, but apply a single set of weights to all simulation sets. These weights are derived analytically from the input power spectra of the CMB + noise + fg_{00}. We compare the weights obtained from the input spectra, and from the spectra obtained from the realizations and find that for each simulation set and on the scales relevant to the lensing reconstruction they differ by a margin we consider acceptable. At most, the weights between the input and empiricial spectra varied 10 percent. The above approach is, by no means, an optimized approach and more sophisticated component-separation techniques are currently studied. The weights are applied as follows. We first convert the Stokes parameters Q and U to E and B with a standard spin-2 transform using the `Healpix` package (Górski et al. 2005), and re-beam the maps to a common resolution of 2.3 arcmin FWHM. In order to avoid too large E-to-B leakage, we first multiply the maps with an apodization mask. Here we simply use the hits map of the scanning strategy, smoothed by a 0.5 degree Gaussian beam. We then isotropically weight the harmonic coefficients and re-project onto Q

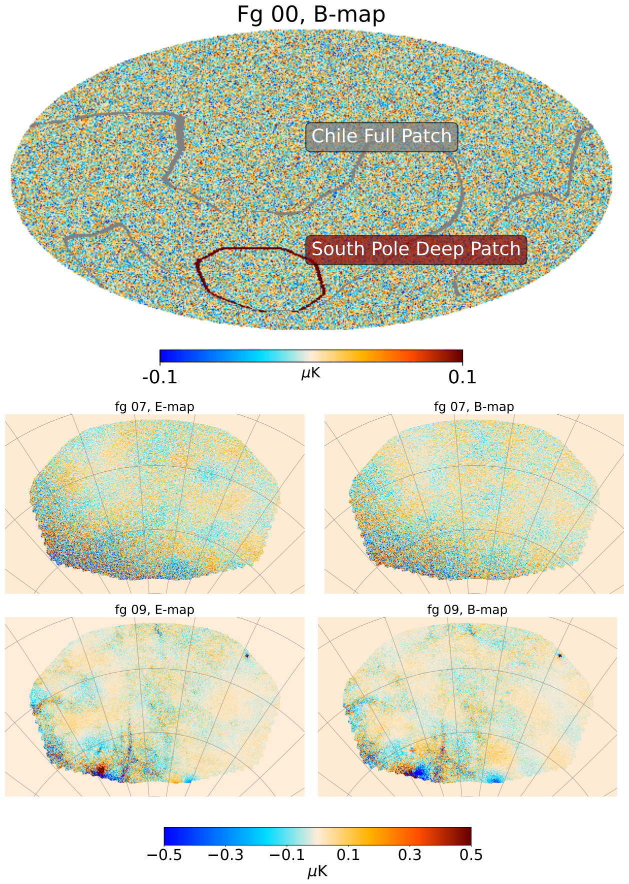

Figure 6. Foreground residual maps for fg_{00} (top), fg_{07} (center), and fg_{09} (bottom), band-passed with $10 < \ell < 3000$, in equatorial coordinates obtained via combining the individual foreground-only frequency maps with the ILC weights. For fg_{00}, we only show the B-map in Mollweide projection, whereas for fg_{07} and fg_{09} we show the E-map (left column), and the B-map (right column) in Cartesian projection and about the South Pole Deep Patch. The graticules in the Cartesian projection plots have a spacing of 15 degrees in both RA and DEC direction. The edges of the South Pole Deep Patch are indicated by the red honeycomb shape in the top panel, and the gray wiggly line indicates the Chile Full Patch.

and U. Owing to masking, the resulting maps are suppressed close to the patch edges. The lensing estimation pipeline always assumes that the local response of the data map to the CMB is isotropic. We approximately account for this by simply rescaling the ILC maps by the inverse of the mask.

An estimate of each of the foreground model residuals is shown in Fig. 6, and is calculated by combining all frequency maps with the ILC weights from the component-separation step, and for the foreground-only maps. We also indicate the boundaries of the South Pole Deep and Chile Full hits pattern indicated by the dark red and gray contours.

Figure 7 shows the EE (BB) noise, CMB, and foreground power spectra estimates on the South Pole Deep Patch in the top (bottom) panel and for a single realization. The fiducial CMB signal is shown in red. The noise is shown in blue, and becomes the dominant component at $L \gtrsim 4000$ in E, and at $L \gtrsim 2800$ in B. We also show a theoretical purely white-noise power spectrum of $0.45\ \mu$K-arcmin in thin gray. The different foregrounds fg_{00} (dash dotted), fg_{07} (solid), and fg_{09} (dashed) are shown in green; The component fg_{07} has the largest amplitude in both E and B and on most scales. The black data points show the binned power spectra of the component-separated maps, and the error bars indicate the uncertainties on the bins.

4.2. *South Pole Deep Patch Map-based Delensing*

We now discuss the QE and iterative lensing reconstruction, which we use to delens the simulations on the South Pole Deep Patch (SPDP). We use 500 ILC component-separated maps of simulation sets M_{00}, M_{07}, and M_{09} on the SPDP, which covers about five percent of the sky. Our final products are QE and MAP method Wiener-filtered E-maps, QE and MAP method lensing potential maps, and QE and MAP B-lensing templates. We quantify the performance and accuracy of our B-lensing templates by calculating the residual lensing amplitude A_{lens} and split the SPDP into various sky patches of different noise levels. We analyse A_{lens} as a function of these sky patches that we consider, and analyze the robustness of the lensing reconstruction by changing parts of the pipeline and dependencies. If not stated otherwise, calculation of a power spectrum is always done with our own implementation of the `PolSpice` (Chon et al. 2004) algorithm.

4.2.1. *B-lensing templates*

As discussed in detail in Sec. 2.2, the reconstructions (both QE and iterative solution) take a number of fiducial ingredients as input, notably a suitable noise-covariance matrix model. We do this by calculating a central noise level $\hat{n}_{\mathrm{lev}}$ and scale it with the hits-count map as follows. To obtain $\hat{n}_{\mathrm{lev}}$, we calculate the empirical power spectrum of the component-separated noise maps for the central area of the patch, $\hat{C}_\ell^{\mathrm{noise}}$, and then convert that result to a white noise level in μK-arcmin, using

$$\hat{n}_{\mathrm{lev}} = \sqrt{\langle \hat{C}_\ell^{\mathrm{noise}} \rangle_\ell \frac{60 \cdot 180}{\pi}}\,, \tag{51}$$

where $\langle\rangle_\ell$ denotes the average over the multipole range $1800 < \ell < 2000$. We choose this range (which is not particularly critical) because lensing reconstruction utilizes information of the CMB at the smallest scales possible, which are, at the same time, not too noisy. While the E-mode is always close to being effectively noise-free, this range is roughly at the scales at which, for the first iterations in the B-mode map, noise becomes the dominant component (see Fig. 7). With this, we find $\hat{n}_{\mathrm{lev}} = 0.42\ \mu$K-arcmin. We scale this number with the inverse root of the hits-count map to obtain our fiducial inhomogeneous noise map across the full

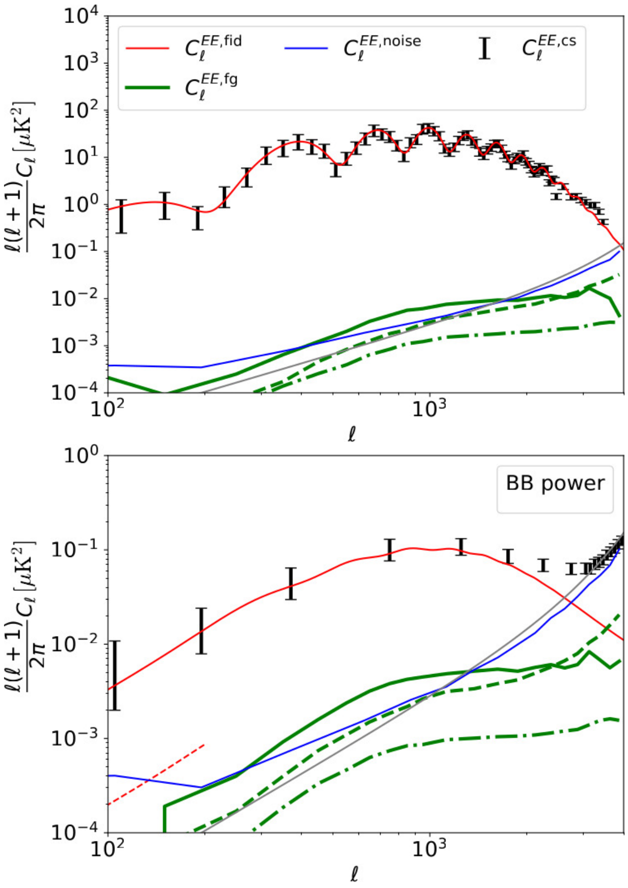

Figure 7. Binned LAT residual power spectra on the South Pole Deep Patch for a sky fraction of about two percent about the central area. The component-separated M_{07} CMB (black) with a simple error estimation shown as error bars, foreground residuals (green) for foreground model fg_{00} (dash-dotted), fg_{07} (solid), and fg_{09} (dashed), and noise (blue) are shown in both panels. The component fg_{07} has the largest power compared to all foreground models. The gray line shows the theoretical white-noise level power spectrum of $0.45\ \mu$K-arcmin and is the estimated average of that patch, (see Fig. 9). *Top panel:* E-mode power spectra. The red line is the fiducial lensed E-mode power spectrum. Noise becomes the dominant component at about $\ell = 4000$. *Bottom panel:* B-mode power spectra. The red line is the fiducial B-lensing power spectrum, the dashed red line shows the theoretical B power spectrum with a residual lensing amplitude of 6 percent (we only show multipoles $\ell < 200$, as these are the ones that our procedure can produce). Noise becomes the dominant component at about $\ell = 2800$.

patch.[24] The resulting SPDP noise variance map is shown in Fig. 9, which ultimately is used for both reconstruction methods. The procedure described above does not account for the residual foregrounds in the maps, and therefore results in slightly suboptimal reconstructions. However, the analytical forecast tool described earlier predicts this to be a small effect[25] and we neglect it in this paper. We also neglect any effective noise correlation introduced by the component-separation procedure, foregrounds, or $1/f$-noise.

Our fiducial model transfer function is the same isotropic Gaussian beam with a FWHM of 2.3 arcmin used for the component-separation, along with the appropriate pixel window function, and the fiducial spectra for $C_L^{\kappa\kappa,\mathrm{fid}}$ and $C_\ell^{EE,\mathrm{unl}}$ are the input FFP10 cosmology spectra.

The recovery of the unlensed E-modes with $\ell_E < 10$ is not attempted in the baseline reconstructions. This is taken care of by setting the fiducial transfer function in the likelihood for these multipoles to zero. We apply the overlapping B-mode deprojection on the noise model as discussed in Sec. 2.7 for scales $\ell_B \leq 200$, suppressing all these B-modes by a factor of 10^4. Further, a prior much smaller than $C_L^{\kappa\kappa,\mathrm{fid}}$ is set on the lenses with $L < 4$. This is done to reduce the size of the Newton steps for these scales and for all iterations, so that the algorithm becomes numerically more stable on these scales.

The quadratic estimator is built in the same way as for the latest *Planck* analyses (Aghanim et al. 2020b; Carron et al. 2022), using `Plancklens`. The QE mean-field $\hat{\boldsymbol{\alpha}}^{\mathrm{QE,MF}}$ is calculated by averaging over 200 reconstructions of $\hat{\boldsymbol{\alpha}}^{\mathrm{QE}}$, less than the available 500 simulations for historical reasons.[26] The QE lensing potential is mean-field subtracted, and then isotropically[27] weighted by

$$\mathcal{W}_L^{\kappa} = \frac{C_L^{\kappa\kappa,\mathrm{fid}}}{C_L^{\kappa\kappa,\mathrm{fid}} + N_L^{(0),\kappa}}\ , \tag{52}$$

and this serves as the starting point for the iterative method. The QE uses CMB multipoles up to 3000, but the MAP iterations use modes up to 4000.[28] We assume that the unlensed E-mode is band-limited to 4000, and neglect the transfer of power caused by lensing beyond this scale.

The mean-field part of the gradient, $g_{\boldsymbol{\alpha}}^{\mathrm{MF}}$ is set to $\hat{\boldsymbol{\alpha}}^{\mathrm{QE,MF}}$ at each iteration and therefore kept invariant across iterations. By doing this, we neglect the delensed noise mean-field as discussed earlier, as well as the changes in the mask and noise mean-fields induced by the changes in power of the maps

---

[24] For a real data analysis, the same simulation-based approach might not necessarily be as accurate, but the precise value is not critical and other ways are possible, for example using splits of the data.

[25] In a comparable experimental configuration we found a predicted relative increase in residual lensing B-mode power of 3% to 12% depending on the foreground model.

[26] This is not critical; results using 100 simulations only for the QE mean-field are extremely close to our baseline 200.

[27] We do not include the QE a-posteriori inhomogeneous weighting called 'κ-filtering'(Mirmelstein et al. 2019), which could in principle slightly improve the reconstruction and QE lensing templates close the mask edges.

[28] The values for QE and MAP are different only for historical reasons. The main purpose of the QE estimate is to have a starting point for the MAP estimator; using more multipoles to build the QE keeps the converged MAP solution unchanged. In the baseline configuration discussed in this paper, the predicted residual lensing amplitude for QE-delensing reduces from 23.2 to 20.6 if the multipole range is increased to 4000, which could have been done without any caveats.

through the iterations. Later in Sec. 4.2.4, we show that this is an adequate approximation.

During each iteration, we must perform a number of deflection operations $\mathcal{D}_{\alpha}$. This remapping, in principle, is done on the full sky. However, the reconstructed lensing map is very weak outside of the SPDP. It is therefore a valid simplification to neglect the deflection field sufficiently far away from the SPDP. We define this region to be larger by 5 degrees than the SPDP itself. Outside that region, we effectively assume that the Wiener-filtered, unlensed E-mode is zero. We perform the lensing remapping on a reduced Gauss-Legendre geometry (Reinecke & Seljebotn 2013) using a bicubic spline interpolation on an equidistant, cylindrical grid of resolution 1.7 arcmin. This resolution is not always sufficient for precise recovery of the largest lenses, but is good enough for our purposes. A higher resolution increases the robustness of the reconstruction at these low multipoles at the expense of an increased computational cost.[29]

We calculate the B-lensing templates following Eq. (33), and for the quadratic estimator with the perturbative version of Eq. (34), given by Eq. (6). Figure 8 shows an example B-lensing map before (top panel), and the residual B-lensing maps after QE (middle panel), and MAP delensing (bottom panel) on the patch area. As expected, most power is removed at the center of the patch where the data noise is lowest and therefore lensing reconstruction performs best.

4.2.2. *Delensing Performance*

We now characterize the fidelity of the obtained B-lensing templates and quantify the reduction of the B-lensing signal after internal delensing in terms of the residual lensing amplitude A_{lens}. The QE and MAP residual B-lensing maps are calculated by subtracting the respective B-lensing templates from the B-lensing maps by using Eq. (33). We obtain A_{lens} by calculating the ratio between the B-lensing power spectrum and the residual B-lensing power spectrum, and compare our results to the analytical predictions obtained by our tools described earlier in Sec. 2.8.

Both the noise and foreground residuals are inhomogeneous across the patch. For this reason, the quality of the lensing reconstruction varies across the observed area. To quantify the residual lensing amplitudes as a function of position, we define sky patches inside the SPDP following its hits-count map: patch p_t with hits-count ratio threshold t includes the observed pixel i with the hits-count h_i if,

$$h_i < \frac{1}{t}\,. \tag{53}$$

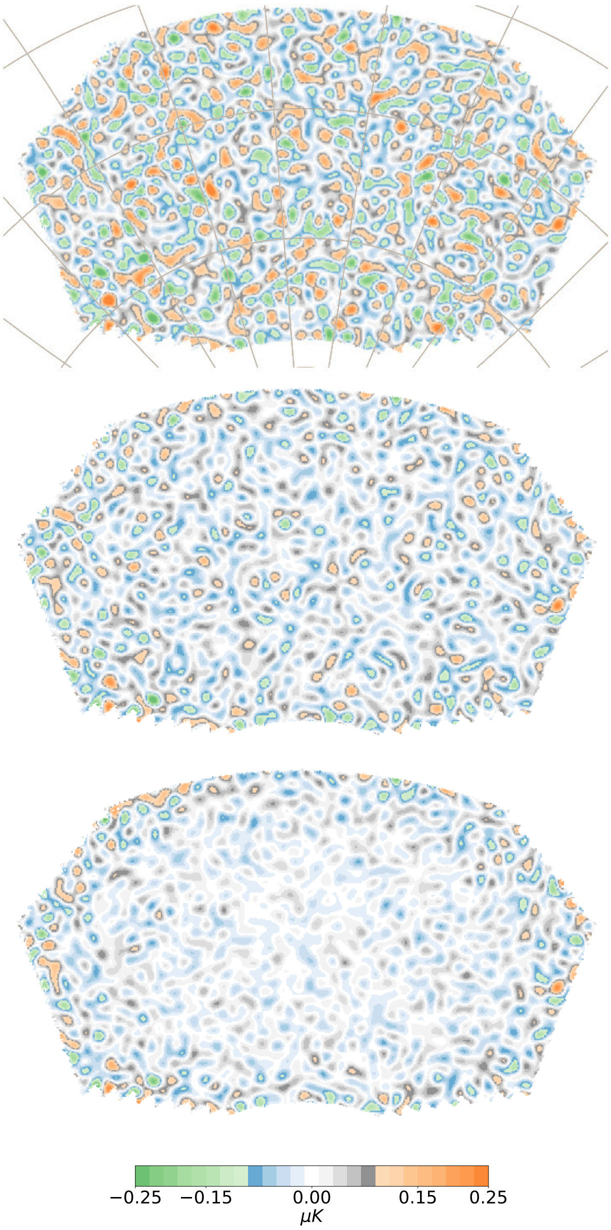

Figure 8. CMB B-lensing maps of the South Pole Deep Patch, before (top panel) and after quadratic estimator (middle panel) and iterative delensing (bottom panel), using the B-lensing templates from simulation set M_{00}. The maps are band-passed to only show $30 < \ell < 200$. Delensing works best in the center of the patch, where the improved performance of the iterative solution is clearly visible. Close to the mask edges the instrumental noise increases, reducing the quality of the lensing reconstruction and of the delensing. The graticules have a spacing of 15 degrees in both RA and DEC directions.

Here, h_i is normalized by the maximum number of hits. The conversion between h_i and its noise level $n_{\mathrm{lev},i}$ is straightfor-

---

[29] As we neared completion of this paper, our lensing remapping procedure has seen major improvements (Reinecke et al. 2023) in accuracy and execution time, allowing better recovery of the large-scale lenses and removing the need for the simplifications just discussed. While the entire set of maps of this analysis has not been reprocessed, dedicated tests on M_{00} are indicating that our latest and most ambitious version of the code can lead to visible improvements, reaching a residual lensing amplitude of 0.059, compared to 0.069 obtained in this paper in Sec. 4.2.2.

ward, if the central noise level $\hat{n}_{\mathrm{lev}}$ is known:

$$n_{\mathrm{lev},i} = \frac{\hat{n}_{\mathrm{lev}}}{\sqrt{h_i}} \, . \tag{54}$$

Thus, an area of the sky observed 10 times less frequently is about 3.3 times noisier. We choose $t = [1.2, 2, 10, 50]$ and obtain patches with sky fractions between 0.07% and 4.4% (see Fig. 9). The blue patch is the essentially homogeneous central region, with an average noise level of $0.44\ \mu$K-arcmin and its edge pixels noise level is about $0.46\ \mu$K-arcmin. This is indicated as a blue line inside the color bar at the bottom. A summary of the average and maximum noise levels together with its sky fraction and the hits-count ratio thresholds is shown in Table 2. The B-lensing signal is about $5\ \mu$K-arcmin. With a central noise level of $\hat{n}_{\mathrm{lev}} = 0.42\ \mu$K-arcmin, we therefore expect the iterative lensing analysis to become ineffective at ~ 10 times the central noise value, as B-lensing is an $\sim 5\ \mu$K-arcmin contamination that cannot be resolved anymore at these levels. This boundary is close to the green contour that can be seen in Fig. 8, where delensing is much weaker and the iterative method brings little improvement.

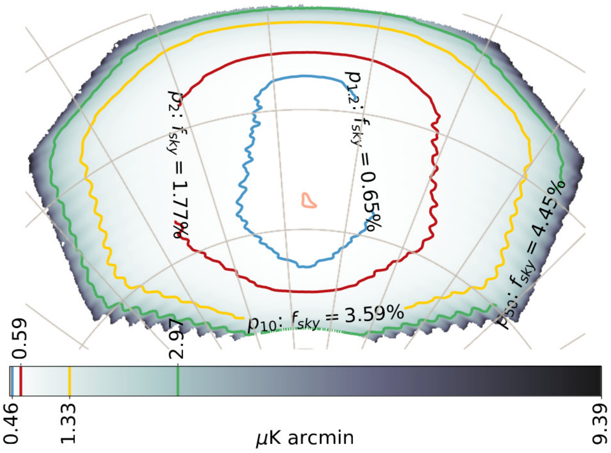

Figure 9. South Pole Deep Patch noise-level map used in the lensing likelihood model, derived from calculating a central noise level and scaling it according to the hits-count map. The contours show edges which contain the masks $p_{1.2}$ (blue), p_2 (red), p_{10} (green), and p_{50} (yellow). Each mask covers the full area inside of it, for example, $p_{1.2}$ is contained in p_2. The lowest (highest) noise of about $0.42(9.39)\ \mu$K-arcmin is indicated by the pink (dark gray) contour. The edge pixel noise levels of the contours are indicated by lines in the color bar, the pink contour is not shown. The graticules have a spacing of 15 degrees in both RA and DEC directions.

The power spectra of the B-lensing and residual B-lensing maps of M_{07} for each of the sky patches are shown in Fig. 10, using a binning window of $\Delta\ell = 30$ for the calculation. Each panel shows the ensemble average and standard deviation of the B-lensing (orange), and QE (dark blue), and MAP (light blue) delensed power spectra for the patch indicated by the legend referring to Fig. 9. The orange line shows the input B-lensing signal, and the blue lines show the analytic predictions of the delensed power spectra. Later, we give more details about the calculation of the predictions. As can be seen, the delensed power spectra are largely scale-independent and increase with increasing sky fraction due to the increase in noise. At the same time, the standard deviation decreases due to the larger sky fraction reducing the sample variance. We do not show the first bin of the delensed power spectra, since the simulations do not contain simulated signals on the reionization peak.[30]

Table 2. Average and maximum noise levels for the patches chosen in this analysis. f_{sky} is the ratio of the number of considered pixels to the total in patch. See text for discussion.

Identifier	**Contour color**	f_{sky} **[%]**	**Average noise level [μK arcmin]**	**Maximum noise level [μK arcmin]**
p_1	pink		0.42	0.43
$p_{1.2}$	blue	0.65	0.44	0.46
$p_{2.0}$	red	1.77	0.48	0.59
p_{10}	yellow	3.59	0.55	1.33
p_{50}	green	4.45	0.64	2.97
p_{∞}		4.97	0.67	9.39

Across all simulation sets, M_{07} gives the highest ensemble variance, and is consistent with M_{07} having the largest foreground residuals, see Fig 7.

We now calculate A_{lens} as the ratio of the ensemble mean of the power spectra using the relation between the power spectrum and its Gaussian variance. This approximates the decrease in the B power fluctuations, since for Gaussian fields, the spectrum variance $\left(\Delta C_\ell^{BB}\right)^2$ is quadratic in the power itself,

$$(\Delta C_\ell^{BB})^2 \approx \frac{2(C_\ell^{BB})^2}{f_{\mathrm{sky}}(2\ell+1)} \, , \tag{55}$$

where f_{sky} is the sky fraction from which the power spectrum is computed. We quote the QE and MAP delensed A_{lens} using

$$A_{\mathrm{lens}}^{\mathrm{QE/MAP}} = \frac{\left\langle C_\ell^{BB,\mathrm{QE/MAP,RL}} \right\rangle_{30\leq\ell\leq200}}{\left\langle C_\ell^{BB,\mathrm{L}} \right\rangle_{30\leq\ell\leq200}} \, , \tag{56}$$

where $C_\ell^{BB,\mathrm{QE/MAP,RL}}$ ($C_\ell^{BB,\mathrm{L}}$) is the (residual) B-lensing power spectrum. The results are shown in Fig. 11. The left (right) panel shows the empirical MAP (QE) result and all

---

[30] By construction, the CMB-S4 simulation suite used in this work did not contain signals below $\ell = 30$, in order to force conservative forecasts based on the recombination peak signal only.

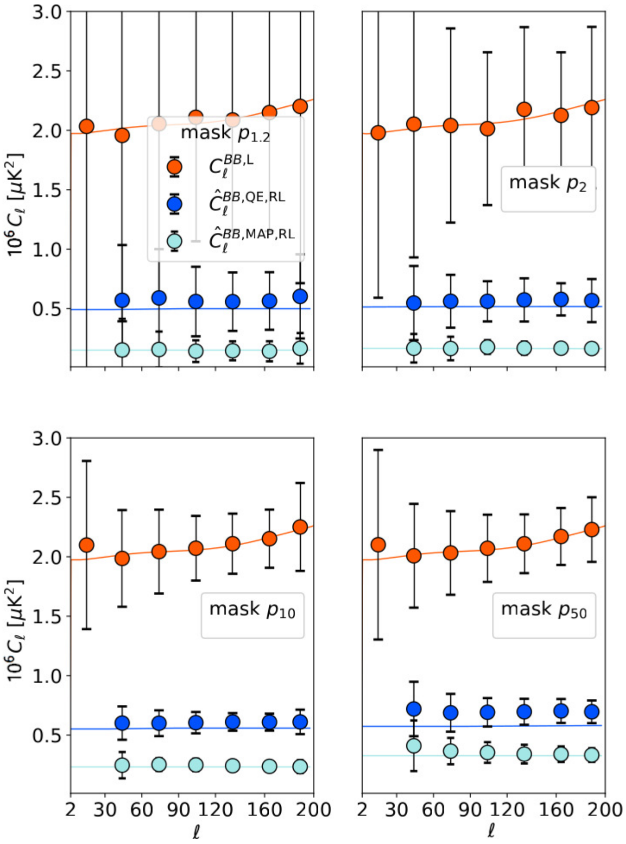

Figure 10. Ensemble average and standard deviation of the B-lensing (orange), QE delensed (dark blue), and MAP delensed (light blue) B-lensing power spectrum, binned with $\Delta\ell = 30$, for simulation set M_{07}. The power spectra are calculated on four different sky patches (see Fig. 9), one per panel. The solid lines show the fiducial B-lensing signal (orange), and analytical prediction for the QE (dark blue) and MAP residual power spectrum. Delensing performs almost equally well across all scales but shows a dependence on the sky patch owing to inhomogeneous noise and foreground residuals. Results on the other foreground models are qualitatively similar. See Fig. 11 for more detailed comparisons.

three simulation sets in light blue (dark blue), where M_{00} is shown in the top two columns, M_{07} in the middle two columns, and M_{09} in the bottom two columns. The error bars display the standard deviation of this ratio across the simulation set.

Each panel shows the empirical results on $\hat{A}_{\mathrm{lens}}$ for each sky patch, and includes semi-analytical predictions (orange and pink), which we discuss in the next section. The QE results show the values obtained via perturbative remapping of the E-mode, Eq. (6). We note that perturbative lensing remapping results in lower residual lensing amplitudes by about 4 percent compared to non-perturbative results, for all sky patches and simulation sets, in agreement with (Baleato Lizancos et al. 2021a). Table 3 reproduces the numbers for the central, most relevant region.

Table 3. Residual lensing amplitudes for $30 \leq \ell \leq 200$ found in this study on the region p_2 of the South Pole Deep patch, see Fig. 9.

	M_{00}	M_{07}	M_{09}
$A_{\mathrm{lens}}^{\mathrm{QE}}$	0.232 ± 0.015	0.238 ± 0.015	0.235 ± 0.015
$A_{\mathrm{lens}}^{\mathrm{MAP}}$	0.069 ± 0.005	0.079 ± 0.005	0.073 ± 0.005

Unsurprisingly, M_{00} gives the lowest values for both QE and MAP, since it has the lowest foreground contamination, and, in addition, contains homogeneous Gaussian noise across the SPDP. M_{07} contains the highest foreground contamination, therefore resulting in the highest A_{lens}. M_{09} is the most complex of all three simulation sets, but has slightly less foreground residual power compared to M_{07}, leading to a slightly lower residual lensing amplitude.

4.2.3. *Comparison to predictions*

We now discuss the calculation of our predictions, starting with simple assumptions, and gradually adding components to increase realism.

In the most optimistic case, we can calculate predictions based solely on the central white noise level of 0.42 μK-arcmin, and ignore foregrounds, sky cuts, and the scan strategy of the experiment. By using the same cuts as in the reconstruction, ($200 < \ell_B < 4000$, $10 \leq \ell_E \leq 4000$) for MAP, and ($200 < \ell_B < 3000$, $10 \leq \ell_E \leq 3000$) for QE, we predict $A_{\mathrm{lens}}^{\mathrm{MAP}} = 0.053$ for the polarization-only iterative delensing estimator and $A_{\mathrm{lens}}^{\mathrm{QE}} = 0.23$ for the QE case. For sky-model M_{00} on the sky patch p_2 we find 0.069 and 0.23, see Table 3. This indicates that in contrast to the QE case, foreground residual power plays a non-negligible role overall in the iterative reconstruction. This is of course expected to some extent, since the size of the foreground residuals relative to the B-lensing power is much greater for the MAP reconstruction.

We can attempt to take foreground residuals into account by assuming they effectively act as additional noise. Foreground models fg_{07} and fg_{09} contain some level of non-Gaussianity, which can in principle affect the reconstruction in less trivial ways. In the baseline reconstructions, this non-Gaussianity is removed by subtracting the mean-field from the lensing-potential estimates (recall that for these foreground models the same non-Gaussian realization enters all simulations). However, the next section shows that iterative reconstructions, for which the mean-field does not match the input foreground model, gives the almost same results.

We construct residual foreground maps by combining the single frequency simulation-input foreground maps with the weights derived from the component-separation step in the previous section, and with the appropriate beams of each detector. By repeating the exercise of calculating the power spectrum for each patch on the sky for an appropriate binning, we obtain the corresponding residual foreground power spectra.

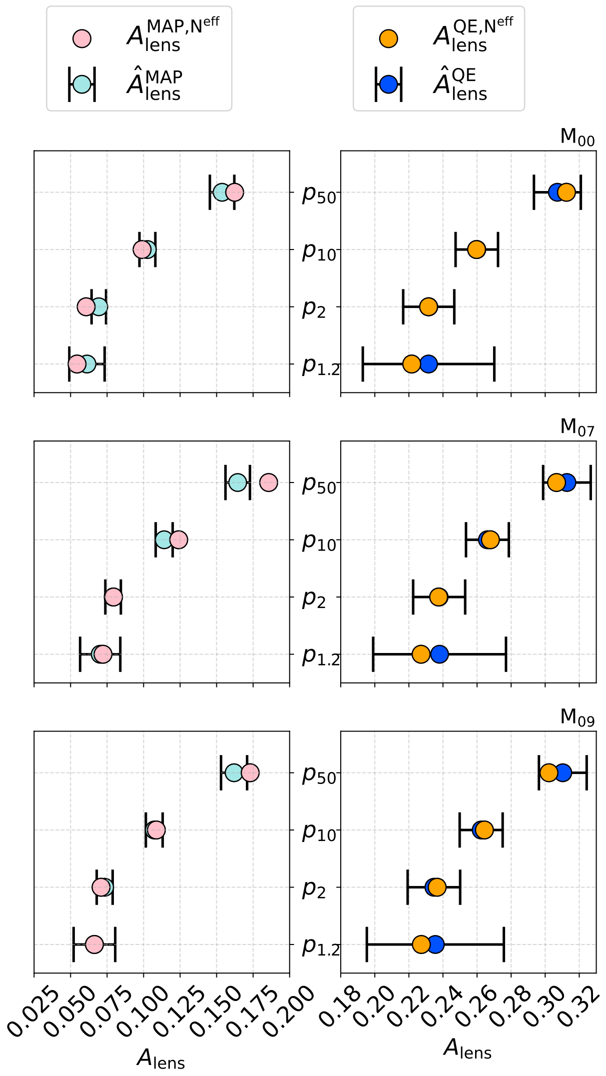

Figure 11. Ensemble average and standard deviation of the QE (right panels, dark blue), and MAP (left panels, light blue) residual lensing amplitude $A_{\rm lens}$ binned across the recombination peak, for $\rm M_{00}$ (top two panels), $\rm M_{07}$ (middle two panel), and $\rm M_{09}$ (bottom two panel), and predicted residual lensing amplitudes shown as pink (MAP) and orange (QE) data points. The dark blue data points are the empirical results using perturbative lensing remapping. In each panel, data point across the y-axis shows the result for the different sky regions, as indicated by the labels on the y-axis (see Fig. 9). $\rm M_{07}$ is always slightly larger than $\rm M_{09}$ and we find the lowest ensemble mean $A_{\rm lens}$ for MAP delensing in patch $p_{1.2}$. The lowest ensemble standard deviation, however, comes from patch p_2. All predictions consider foreground power in the noise estimate.

We refer to the input-noise in our prediction as *noise-only* when we calculate the noise power spectrum from the empirical noise maps, and as *effective noise* if we add the foreground residuals power spectrum on top of it. Calculating the power spectrum across a sky patch effectively gives the average noise power spectrum. This may be valid for the central area, for which the hits-count map is close to being homogeneous, but it is a poor approximation for the areas outside the central part of the SPDP. It is therefore advisable to include the noise inhomogeneities for the calculation of the residual lensing amplitude. We do this as follows. Let $f_{\rm sky}(\hat{n}_{\rm lev})$ be the observed sky fraction with local noise level $\hat{n}_{\rm lev}$, and $A_{\rm lens}(\hat{n}_{\rm lev})$ the predicted residual local lensing amplitude. A weighted average across the SPDP is

$$\frac{1}{f_{\rm sky,tot}}\int d\hat{n}_{\rm lev} f_{\rm sky}(\hat{n}_{\rm lev}) A_{\rm lens}(\hat{n}_{\rm lev})\,. \qquad (57)$$

Here, $f_{\rm sky,tot}$ is the total observed sky fraction. $f_{\rm sky}(\hat{n}_{\rm lev})$ can simply be obtained from the hits-count map, and $A_{\rm lens}(\hat{n}_{\rm lev})$ is approximated as a polynomial function and calculated from various white-noise and noise-only level predictions. With this, we find that including inhomogeneous noise, $A_{\rm lens}$ increases by about 0.01 at p_{10}, and by about 0.03 at p_{50}.

Since $A_{\rm lens}$ is close to linearly dependent on the noise for small $A_{\rm lens}$, we include this correction by simply adding it on top of our predictions for the *noise-only* and *effective noise* case. Figure 12 shows the predictions for $\rm M_{09}$ in purple (noise-only) and pink (effective noise) in the left panel for the MAP solution, and similarly in yellow and orange in the right panel for the QE. The effective noise predictions all seem consistent with our empirical findings, and we see a similar behavior for sky models $\rm M_{00}$ and $\rm M_{09}$ as shown in Fig. 11.

Our findings show that for the relevant sky areas, inclusive of the area in which the noise is not too inhomogeneous any more, the residual foregrounds effectively act as a source of noise in the lensing reconstruction, and a simple modelling can accurately predict the reduction in B-mode power.

4.2.4. *Mean-field tests*

The lensing reconstruction likelihood model contains several ingredients, and their accuracy affects the quality of the B-lensing template. One factor that relies heavily on good simulations of the data is the mean-field, which reflects our (mis-)understanding of induced anisotropies from, among other factors, noise, foreground residuals, and masking. The better we can simulate our noise and foregrounds, the more we can remove their contribution from our estimate. At the same time, it is interesting to understand the impact of having simulations with only limited accuracy. In this section, we discuss the robustness of our lensing reconstruction by altering the mean-field, and obtain lensing potential estimates and residual lensing amplitudes that are modified compared to the 'baseline' results, which were discussed in the previous section.

Similar to the quadratic estimator mean-field, we define the mean-field of the iterative solution using an average over a set of converged iterative reconstructions,

$$\hat{\boldsymbol{\alpha}}^{\rm MF} \equiv \left\langle \hat{\boldsymbol{\alpha}}^{\rm MAP} \right\rangle_{\rm MC}\,. \qquad (58)$$

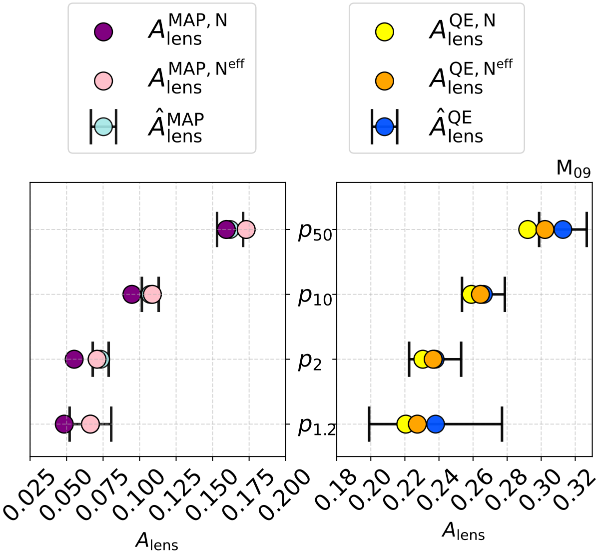

Figure 12. Residual lensing amplitude for simulation set $\rm M_{09}$ at the recombination peak, as calculated for different regions inside the observed patch (see Fig. 9) from their empirical instrumental noise and foreground residuals. If we include the residual foregrounds as additional noise in our prediction, the iterative prediction increases by a few percent (pink on the left panel), and agrees well with our empirical findings, with the exception of the largest sky region where the noise is highest and least homogeneous, and our prediction scheme is less effective. The impact of the foregrounds on the QE results (right panel) is milder owing to the increased importance of the lensing B power relative to the foreground residuals.

We remind the reader, that all lensing estimates $\hat{\boldsymbol{\alpha}}^{\rm MAP}$are built by subtracting a constant QE mean-field at each iteration. Hence, $\hat{\boldsymbol{\alpha}}^{\rm MF}$ is a residual mean-field that has not been accounted for by this term: if we could incorporate the mean-field term $g^{\rm MF}_{\boldsymbol{\alpha}}$ from Eq. (23) perfectly and at all iterations, then $\hat{\boldsymbol{\alpha}}^{\rm MF}$ defined by Eq. (58) would vanish.

Figure 13 shows the lensing mean-field deflection field,

$$\hat{\alpha}^{\rm MF}_{LM} \equiv \sqrt{L(L+1)}\hat{\phi}^{\rm MF}_{LM}\,, \tag{59}$$

for QE (MAP) in the right (left) column, and for $\rm M_{00}$ (top panels), $\rm M_{07}$ (middle panels), and $\rm M_{09}$ (bottom panels). Fig. 14 shows their power spectra in green. In Fig. 13, in the QE case, only the largest modes are visible, and are sourced by the mask and noise anisotropies. Their amplitudes are much larger than the expected true lensing signal, and are not visible in the MAP mean-field (seen as the wildly different scale), showing that the usage of the approximate QE mean-field in the iterations is successful at removing these large signatures. This might be surprising at first, since the delensed B-mode power is wildly different to the lensed one, which could in principle create large changes in the mean-field. However, the EE spectrum is much less affected by delensing, and the mean-field is predominantly sourced by the EE quadratic piece. The reason for this is that all mean-field sources in these simulations appear as a convergence-like lensing signal rather than shear-like. Small-scale signatures in these maps have large Monte-Carlo noise originating from the finite number of simulations (only 100 simulations were used for the tests in this section) used in Eq.(58). The $\rm M_{09}$ QE mean-field (bottom left panel) has a feature in the lower left area of the patch. A similar structure and a couple of bright spots can also be seen for the MAP case.

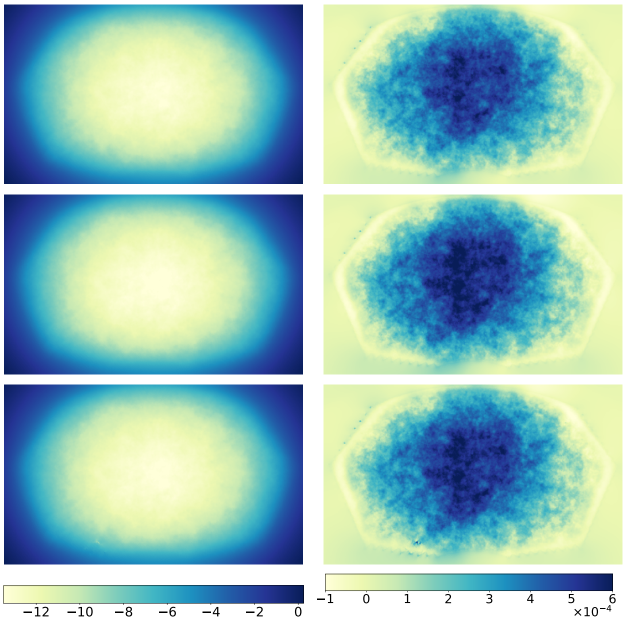

Figure 13. QE mean-field (left column) and MAP residual mean-field (right column) for simulation sets $\rm M_{00}$ (top panels), $\rm M_{07}$ (middle panels), and $\rm M_{09}$ (bottom panels). The MAP mean-field is about three orders of magnitude smaller than the QE mean-field. In fact, most of the small-scale fluctuations are merely noise, but for $\rm M_{09}$ some foreground-induced artefacts are visible in the lower left part of the patch.

The mean-field difference maps, between simulation sets (obtained by differencing maps obtained from Eq.(59)) are shown in Fig. 15. The top two panels show the deflection-mean-fields of $\hat{\boldsymbol{\alpha}}^{\rm MF}(\rm M_{07}) - \hat{\boldsymbol{\alpha}}^{\rm MF}(\rm M_{00})$, while the bottom two panels show $\hat{\boldsymbol{\alpha}}^{\rm MF}(\rm M_{09}) - \hat{\boldsymbol{\alpha}}^{\rm MF}(\rm M_{00})$. QE is shown in the left panels, MAP on the right. These differences are $\sim 10\%$ of the total mean-field. The top panels shows signatures of the foreground amplitude modulation of $\rm fg_{07}$ in the bottom left part of the patch, and in the bottom panels, signatures from foreground filaments of $\rm fg_{09}$ can be seen in the southern part.

In Fig. 14, the pink lines show the spectrum of the lensing potential estimates for $\rm M_{00}$ (dot-dashed), $\rm M_{07}$ (solid), and $\rm M_{09}$ (dashed), $C_L^{\tilde{\phi}\tilde{\phi}}$, before subtracting their respective mean-field. The foregrounds only become relevant in the MAP case: without delensing, the lensing B power dominates the QE reconstruction noise over the foregrounds and all curves are basically indistinguishable.

The calculation of the mean-field from an average of simulations like that given in Eq.(58) inevitably leads to some level of residual Monte-Carlo noise in the estimate due to the finite number of simulations (with expected spectrum $\sim N^{(0)}/N_{\rm MC}$, where $N_{\rm MC}$ is the number of simulations). To avoid this type of noise in the mean-field spectra (indicated by the green lines), we have plotted the cross-spectrum

$$C_L^{\rm MF_x} \equiv \frac{1}{f_{\rm sky}(2L+1)} \sum_M \hat{\phi}_{LM}^{\rm MF_1} (\hat{\phi}_{LM}^{\rm MF_2})^* \,, \qquad (60)$$

where the two estimates come from two independent sets of simulations, and we show them in green. Here, $f_{\rm sky}$ is calculated from the ratio of number of pixels between the masked and unmasked sky. The Monte-Carlo noise enters the iterative reconstruction via the subtracted mean-field of the QE, the starting point of the iterations. We show the auto-spectrum of the mean-field estimate in gray, and the spectrum of this noise is clearly seen, as the curve departs from the green at higher multipoles. Of course, the mean-field template can be set to zero when the noise dominates, and accelerating methods as discussed on the full sky in Sec. 2.8 can also be used in this setting (though the accuracy of the mean-field recovered with these methods relies on that of the likelihood model).

We now turn to three quantitative tests on the impact of the mean-field on the delensing efficiency. The changes applied are the following.

(I) A posteriori mean-field subtraction: we obtain the average of the MAP reconstructions obtained from Eq.(58) and subtract it from each lensing potential estimate (a procedure similar to the QE mean-field subtraction).

(II) Swapped mean-field: we use the simplest simulation set available (the Gaussian foreground model $\rm M_{00}$) to build the mean-field template for the reconstructions on all simulation sets. Hence, the mean-field subtraction will not remove the non-Gaussianity of the foregrounds (and also be wrong in a broader sense, since the different models have different power).

(III) Mean-B-lensing template subtraction (similar to point (I), but at the level of the B-lensing templates): we subtract from each B-template the average of templates obtained on a set of reconstructions.

In all of these tests, we calculate the residual lensing amplitude on the large, green contoured sky patch p_{50} (see Fig. 9). This includes some mean-field signals in the outskirt of the SPDP, yielding conservative results due to the small weight of the SPDP edge for the r-inference. For convenience, we only use 100 simulations, with almost exactly same configuration as described in Sec. 4.2 (here, we boosted the numerical accuracy parameters to obtain more precise large scale lenses). We obtain $A_{\rm lens}$ from the ratio of power, following Eq. (56), along with the ensemble standard deviation. We summarize our findings in terms of a single value $\frac{\Delta A_{\rm lens}}{A_{\rm lens}}$ in

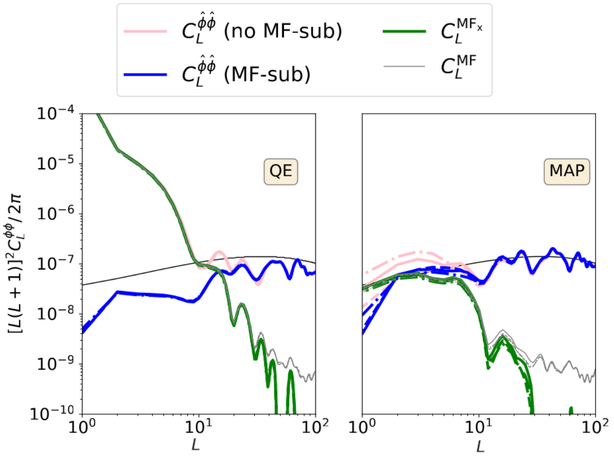

Figure 14. Lensing potential estimate pseudo-power spectra for QE (left panel) and MAP (right panel) as obtained from the CMB-S4 LAT simulations. The solid dark line is the fiducial lensing potential. We show the result for all three simulation sets: $\rm M_{00}$ in dot dashed, $\rm M_{07}$ solid, and $\rm M_{09}$ dashed. However, their spectra often overlap and are barely distinguishable, particularly so in the QE case. Pink (blue) shows the QE or MAP lensing potential before (after) mean-field subtraction, gray (green) shows the QE mean-field or MAP residual mean-field with (without) Monte-Carlo noise. The mean-field drops off quickly for multipoles larger than $L \simeq 10$. The QE mean-field (green and gray) contains a large signal at large angular scales, and is the lensing potential's (pink) main component on that scale. On large scales, the residual MAP mean-field signal is of the order of the lensing potential, many orders of magnitude smaller than its QE counter part (green, left panel). This means that at each iteration, the QE mean-field is a good proxy for the true mean-field term.

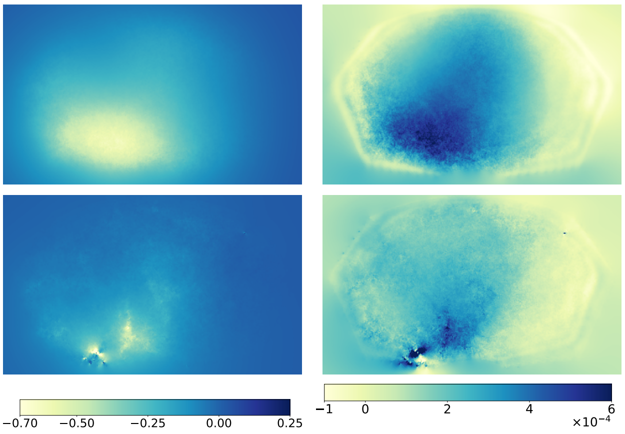

Figure 15. Absolute difference maps for the QE (left column) and MAP (right column) deflection-mean-fields, $\rm MF_{07} - MF_{00}$ (top panels), and $\rm MF_{09} - MF_{00}$ (bottom panels). There are clear features seen in every panel, which originate at least in part from the foreground residuals.

Table 4 that compares the tests to the baseline configuration. Here, $\Delta A_{\rm lens}$ is the difference in residual lensing amplitude between the baseline and modified analyses.

(I): Mean-field subtraction:

For test (I), we subtract the MAP mean-field from the lensing-potential estimates prior to the lensing operation, and calculate the B-lensing templates, denoted $B^{\rm LT}_{\rm MF-}$. Explicitly,

$$B^{\rm LT,MAP}_{\rm MF-} = \left(\hat{\boldsymbol{\alpha}}^{\rm MAP} - \hat{\boldsymbol{\alpha}}^{\rm MF}\right) \circ \hat{E}^{\rm unl}. \tag{61}$$

The result is shown in the first row of Table 4. The impact is always small, and mostly consistent with zero (at 2σ for simulation set $\rm M_{00}$). The given error bars are the one-simulation-equivalents, and therefore, averaging across all simulations, we would detect a tiny improvement in all cases. Nevertheless, this confirms that simply subtracting the QE mean-field as calculated for the starting point is sufficient.

(II): Swapped mean-field:

We now perform the lensing reconstructions with a mean-field different to the baseline. For simulation sets $\rm M_{07}$ and $\rm M_{09}$, we perform reconstructions where the mean-field gradient term is set to that of model $\rm M_{00}$ (doing so also changes the starting point for the quadratic estimator but this alone does not affect the converged solution).

Thus, the lensing reconstruction is performed on an inhomogeneous and non-Gaussian foreground simulation set, but its mean-field is calculated from a simpler simulation set (with differences shown in Fig. 15), with homogeneous and Gaussian foregrounds. The result is shown in Table 4 in the second row, and is consistent with zero. This demonstrates that the choice of mean-field in the lensing reconstruction does only have a small impact. The mean-field from $\rm M_{00}$ only traces the patch mask and noise variance map, yet it nevertheless yields similar results for all simulation sets.

(III): Mean-B-lensing Template Subtraction:

Figure 16 shows an example B-lensing template for $\rm M_{07}$ and $\rm M_{09}$. The MAP mean-B-lensing templates for $\rm M_{07}$ and $\rm M_{09}$, together with the MAP difference maps, $\rm M_{07} - M_{00}$, and $\rm M_{09} - M_{00}$ are shown in Fig. 17. For $\rm M_{09}$ (bottom left panel) we see in Fig. 17 clear features along the edge of the patch, hinting at contamination from foreground residuals. $\rm M_{07}$ does not show any dominant features. The difference plots (right panels) show that for $\rm M_{07}$ minus $\rm M_{00}$, the mean-B-lensing template follows the foreground amplitude modulation of $\rm M_{07}$. For $\rm M_{09}$ minus $\rm M_{00}$, we see bright spots at about the same areas where foreground contamination appears in the mean-field of $\rm M_{09}$, also clearly visible in the mean-B-lensing template in Fig. 16. We also see a small amplitude ringing effect along the latitude of the SPDP, which we believe is an artefact of the map-building procedure, introduced when modes below 30 were excised from the maps.

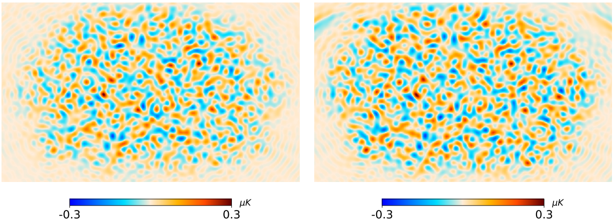

Figure 16. B-lensing templates for $\rm M_{07}$ (left panel) and $\rm M_{09}$ (right panel). All plots are band-passed to $30 < \ell < 200$.

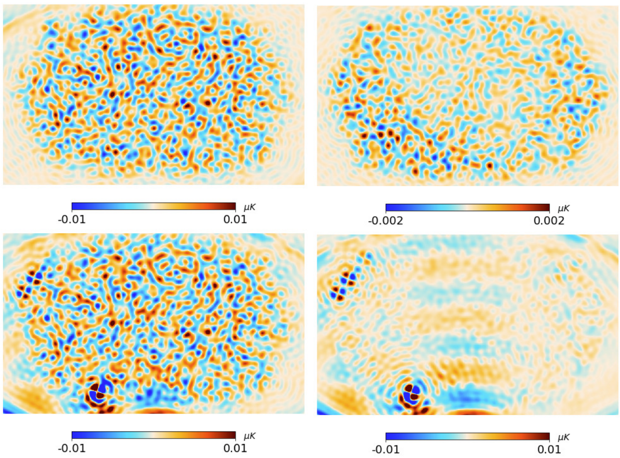

Figure 17. *Left column*: MAP mean B-lensing templates for $\rm M_{07}$ (top) and $\rm M_{09}$ (bottom). *Right column*: Difference maps between the mean B-lensing templates $\rm M_{07}$ and $\rm M_{09}$ with $\rm M_{00}$. All plots are band-passed to $30 < \ell < 200$.

Table 4. Mean-field test results. We give the relative change in terms of residual lensing amplitude, compared to the 'baseline' configuration, $\frac{\Delta A_{\rm lens}}{A_{\rm lens}}$, a negative number indicating an improvement. See text for the definition of the individual tests.

Test	$\rm M_{00}$	$\rm M_{07}$	$\rm M_{09}$
I:	-0.0028 ± 0.0014	-0.0005 ± 0.0014	-0.0009 ± 0.0014
II:		0.0001 ± 0.004	0.0039 ± 0.01
III:	-0.0005 ± 0.02	-0.0002 ± 0.015	-0.0028 ± 0.03

We test for the impact of these features by subtracting the mean-B-lensing template from each B-lensing template,

$$B^{\rm LT,MAP}_{\ell m} \rightarrow B^{\rm LT,MAP}_{\ell m} - \left\langle B^{\rm LT,MAP}_{\ell m}\right\rangle_{\rm MC}, \tag{62}$$

where in the average on the right-hand side we always exclude the simulation index for which $B^{\rm LT,MAP}$ is calculated.

However, similar to the QE and MAP mean-field, subtracting does not only remove contaminants, but also adds some level of Monte-Carlo noise. For this reason, we calculate power spectra by dividing the simulation sets into two splits and taking cross-spectra. The result is shown in the third column of Table 4. We do not find any improvement, and the error bars are rather large.

Summarizing these tests, they confirm that our results are robust against changes to the mean-field treatment, and more generally that the impact of the mean-field is small.

4.3. *Chile Deep Patch Delensing*

As a last test, we repeat the calculation of the B-lensing template for a different CMB-S4 configuration. Specifically, we assume that the Deep Field would be observed from Chile, referring to this as the Chile Deep Patch (CDP). Observing the sky from Chile would, among other things, change the scan strategy of the telescopes, therefore impacting the noise levels and observed sky fractions. Figure 18 shows the derived noise-level map after component-separation, in equatorial coordinates. The central noise level is $\hat{n}_{\rm lev} = 0.6$ μK-arcmin and therefore about 50% higher than for SPDP, and the scan area is much broader.

We generate two simulation sets, that is, the foreground models $\rm M_{00}$ and $\rm M_{07}$, and repeat lensing reconstruction for 100 realizations of the CDP component-separated maps ($\rm M_{09}$ was not available over the larger area at this time owing to practical difficulties). We update our noise model and overlapping B-mode deprojection matrix to this case, and leave other ingredients of our algorithm essentially the same.

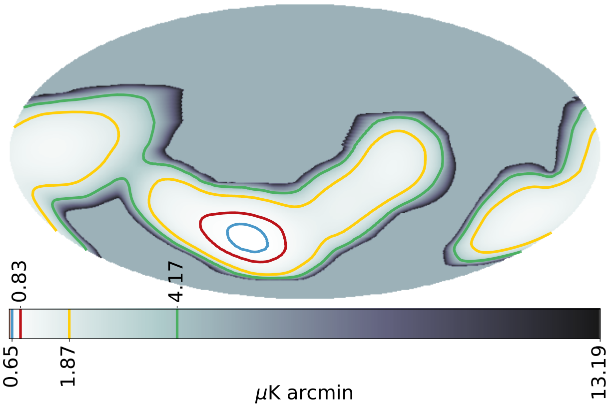

Figure 18. This is a full sky Noise-level map for the Chile Configuration and in equatorial coordinates. The contours show edges that contain the masks A (blue), B (red), C (green), and D (yellow), and the numbers in the color bar refer to the contour noise level. The lowest (highest) noise of about 0.65 (13.19) μK-arcmin is indicated by the blue (dark gray) contour and all are derived from the Chile hits-count map.

To improve numerical stability for the iterative lensing reconstruction step, we have increased the mask by excluding areas for which the noise is more than 100 times larger than the central noise value. These are mostly pixels near the galactic plane with high foreground residuals that can safely be discarded. Iterative lensing reconstruction succeeds on the entire resulting area.

The bulk of the constraining power comes from the red region in Fig. 18. The larger sky patches indicated by the yellow and green contours partly cover the Galactic plane. Although lensing reconstruction does converge there as well, the foreground residuals are too large in these areas, leading to uninteresting delensing results.

In the red region and for $\rm M_{07}$, we find[31] $A_{\rm lens}^{\rm QE} = 0.30 \pm 0.02$, $A_{\rm lens}^{\rm MAP} = 0.121 \pm 0.007$. For $\rm M_{00}$, we find $A_{\rm lens}^{\rm QE} = 0.29 \pm 0.02$, $A_{\rm lens}^{\rm MAP} = 0.115 \pm 0.006$.

These empirical results for iterative delensing are in reasonable agreement with our theoretical predictions, $A_{\rm lens}^{\rm MAP} = 0.11$ for $\rm M_{07}$, and $A_{\rm lens}^{\rm MAP} = 0.10$ for $\rm M_{00}$, matching as well those expected from the CMB-S4 r-forecast paper (Abazajian et al. 2022).

4.4. *Constraints on r*

In our previous CMB-S4 r constraint paper (Abazajian et al. 2022) one of the two re-analysis methods used was the parametric multi-component cross-spectral likelihood that was first introduced in (BICEP2/Keck Collaboration et al. 2015), and which has been used in all BICEP/Keck analyses up to the most recent (Ade et al. 2021). Here we take this likelihood and add the lensing templates derived in the previous section as an additional pseudo-frequency band, in a similar manner to the joint analysis of BICEP/Keck and SPT data presented in (BICEP/Keck Collaboration et al. 2021).

To obtain sufficient E/B separation purity we calculate the auto- and cross-spectra between the maps using the `S²HAT` library (Grain et al. 2009). The resulting bandpowers are then compared against a likelihood model built using the Hamimeche-Lewis (Hamimeche & Lewis 2008) approximation with the following parameters: the tensor-to-scalar ratio r; the amplitudes of the dust and synchrotron B-mode power spectra, their frequency spectral indices, and their spatial power law indices; the frequency-independent spatial correlation between dust and synchrotron; and the dust correlation between 217 and 353 GHz. See appendix G of (Ade et al. 2018) for a full description of the parameterization.

The likelihood requires statistical characterization of the signal and noise parts of the lensing template auto-spectrum and of their cross-spectra to the SAT frequency maps. In this first analysis, we do this in the following crude (but effective) manner. Making the assumption that the template may be written as a filtered version of the true CMB lensing B-mode plus an additive noise term, we compute an effective, isotropic filter function from 100 cross-spectra to their known simulation CMB inputs (neglecting the precise position dependence of the filter). This allows us to build predictions for the cross-spectra to the SAT maps. Further, we obtain the lensing template auto-spectrum noise power by subtracting from the templates spectra the signal term as predicted from the isotropic filter model.

For each of the 500 simulation realizations, and each of the three foreground models, we perform a maximum-likelihood (ML) search, both with and without the lensing template. We use the mean and the standard deviation of a given parameter value to assess the bias and experimental uncertainty of that parameter—where the parameter r is of course by far of the greatest interest. The SPDP results are shown in Fig. 19

---

[31] The results for QE are calculated using non-perturbative remapping of the E-mode, improving performance slightly, as before.

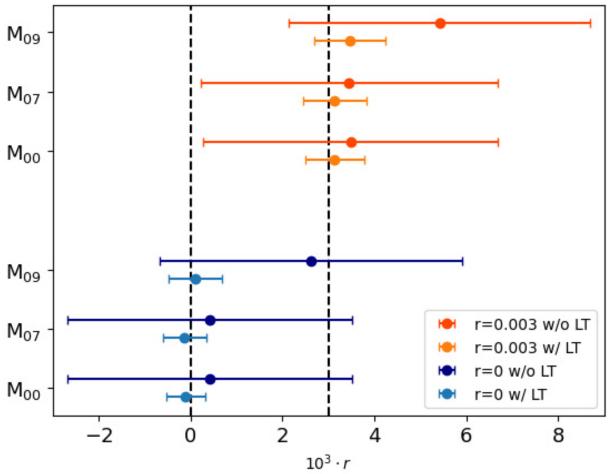

Figure 19. Results from reanalysis of map-based simulations of CMB-S4 observing the South Pole Deep Patch. The mean and standard deviation of the maximum likelihood r values evaluated over ensembles of simulation realizations are shown as the central point and horizontal error bar, respectively. Blue and orange points show the results for $r = 0$ and $r = 0.003$ realizations, with the vertical lines indicating these values. The darker points are results without the lensing template, while the lighter points show the substantial benefit of including the iteratively-reconstructed templates described in this paper.

Table 5. Results from reanalysis of map-based simulations of CMB-S4 observing the South Pole Deep Patch, including the reconstructed lensing templates described in this paper.

r value	Sky model	$\sigma(r) \times 10^4$	r bias $\times 10^4$
0	0	4.3	-1.0
	7	4.7	-1.3
	9	5.8	1.1
0.003	0	6.4	1.3
	7	6.9	1.4
	9	7.8	4.7

for the realizations with $r = 0$ and those with $r = 0.003$. As expected, the addition of the lensing template (effective delensing) dramatically reduces the uncertainty σ on r, here denoted $\sigma(r)$. In the case of the non-Gaussian model M_{09}, a significant bias is also removed. Table 5 gives the with-template results in numerical form. These can be compared to the results in the rightmost column of Table 2 of our previous paper (Abazajian et al. 2022). The results for M_{00} are most directly comparable. Previously for $r = 0$ we had $\sigma(r) = 5.7 \times 10^{-4}$, to be compared to the present 4.3×10^{-4}, the latter being below the $\sigma(r) = 5 \times 10^{-4}$ science goal of CMB-S4 (Abazajian et al. 2019).

We have preliminary ML search runs for the CDP simulations as well. The uniform amplitude M_{00} model is a gross oversimplification for this larger sky patch. Nevertheless, for input $r = 0$ this model gives $\sigma(r) = 5.2 \times 10^{-4}$, which is consistent with the expectation that the best sensitivity will be obtained with the most concentrated sky coverage. For M_{07} the effective foreground amplitude averaged over the patch is extremely high and the likelihoods result is not robust. We find that $\sigma(r)$ is only increased to 7.0×10^{-4}, but there is significant bias—further investigations are ongoing. No results are available for M_{09} at the time of writing.

5. CONCLUSION

In this paper, we have presented an optimal lensing reconstruction software on the curved sky, usable under realistic conditions, inclusive of inhomogeneous noise, and on sky patches of arbitrary size. With it, we achieve the CMB-S4 science goal on r by removing 92% to 93% of the B-lensing power, depending on the simulation set.

The method mostly generalizes that of (Carron & Lewis 2017) to curved-sky geometry. We discussed carefully all ingredients, and showed that the few approximations that are made are perfectly under control given the moderate size of ΛCDM deflection angles (as is often the case, the sky curvature can be neglected on the scale of the deflection), and that the reconstruction should be optimal provided that the CMB likelihood model is a fair representation of the data.

We found most difficulties of practical nature in the reconstruction process at low lensing multipoles. There, the mean-field is prominent, and more generally, the behavior of the reconstruction of the largest scale lenses is less stable, and also often requires higher accuracy in the most important steps (lensing remappings, recovery of the delensed E-mode) than for smaller-scale lenses. We have used here as the baseline a compromise, where the mean-field and largest lenses are not precisely reconstructed, which for most applications is perfectly satisfactory: these large lenses contribute little to B-mode delensing on degree scales, and, for the purpose of lensing-spectrum reconstruction, cosmic variance dominates there anyway, so that the quadratic estimator solution can be satisfactory. Recent work (Reinecke et al. 2023) allowed us to solve most of these issues, making upcoming analyses easier as well as slightly improving the lensing reconstruction further.

New to this paper is the characterization of the noise anisotropies induced by delensing as the only source of the mean-field, which has no direct analog in quadratic-estimator theory. We showed that it is a negligible source of complication, in particular for situations where uncertainties in the statistical properties of the data overshadow its contribution. One is naturally lead to drastic simplifications in the treatment of the mean-field, which becomes as simple as for the standard quadratic estimator.

The iterative reconstruction usually converges in a reasonable number of steps (rarely did we use more than 15 iterations), but is by no means a smooth process. In fact, the non-Gaussian nature of the likelihood is such that for low noise levels full Newton steps are much too strong, resulting generically in unrealistic, non-invertible deflection-field estimates. We found that reducing the step in a scale-dependent manner can correct that problem, but this requires undesirable ad-hoc tuning, with little improvements each time, except for the first few iterations; calculation of the gradients is costly and generally improvement of the convergence after a few steps would be highly desirable and is left for future work. For the purpose of delensing, it helps that recovery of the lensing B-mode appears easier than that of the lensing deflection field and E-mode map individually: errors in both tracers display statistical dependence in order to still produce the right level of B-modes seen in the data.

Also new to this paper is the joint reconstruction of the lensing post-Born curl together with the standard lensing-gradient mode. The joint reconstruction brought no complications with it, and the output is correctly described by our predictions of the iterative $N^{(0)}$ and $N^{(1)}$ biases. At the current time the lensing curl is really only useful as a sanity check of the lensing reconstruction, but at the CMB-S4 depth probed here, it will be also be a signal, whose detectability is boosted by large factors with our optimal method.

We focused mostly on simulations of a very deep patch, using semi-realistic inhomogeneous CMB-S4 noise maps, together with two non-Gaussian foreground models. Our baseline reconstructions on ILC-cleaned maps are not perfect near the edges and on large scales. Nevertheless this has little impact overall, and the tracer is able to remove close to 93% of the lensing over the relevant area, as expected. Of course, available Galactic foreground models are still crude on small scales, but we find encouragingly that the main effect of the presence of the foregrounds on the lensing templates appears to be the increased effective noise of the maps. This extends cautious optimism of (Beck et al. 2020) to the case of iterative delensing. For the most complex model M_{09} a non-Gaussian feature is present close to the edges, which is visible when averaging the templates.

We built predictions for the delensing capabilities and lensing bias calculations, using a slightly modified version of the well-known algorithm introduced in (Smith et al. 2012). It is quite noteworthy how well these predictions (which do not exactly follow from first principles), seem to work even under non-idealized conditions, and it is reassuring to see that forecasts with these tools can probably be trusted.

To summarize, (Kamionkowski & Kovetz 2016) writes that B-mode delensing is an ‘ambitious, sophisticated, and challenging endeavor’: we believe that this paper makes the prospect of an eventually successful delensing much more realistic. Of course, much work remains to be done. Our noise simulations were inhomogeneous but otherwise did not have the full complexity expected from a ground-based experiment like CMB-S4. While at least some of such non-idealities are expected to impact predominantly scales larger than those most relevant for lensing reconstruction, it is difficult to assess the importance of these effects for delensing without more realistic simulations. This is left for future work. Other areas that need development include the study of systematic effects, better foreground models and cleaning methods, non-linear large-scale structure simulation inputs, and further code improvements.

The authors thank Victor Buza, Benjamin Racine for previous work relevant to this paper. CMB-S4 is supported by the Director, Office of Science, Office of High Energy Physics of the U.S. Department of Energy under Contract No.DE-AC02-05CH11231; by the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility under the same contract; and by the Divisions of Physics and Astronomical Sciences and the Office of Polar Programs of the U.S. National Science Foundation under Mid-Scale Research Infrastructure award OPP-1935892. Considerable additional support is provided by the many CMB-S4 team members and their institutions. This research used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. JC, SB, and LL acknowledge support from a SNSF Eccellenza Professorial Fellowship (No. 186879). CB acknowledges support from the U.S. Department of Energy Office of Science (DE-SC0011784). This work was supported by a grant from the Swiss National Supercomputing Centre (CSCS) under project ID s1203 and sm80. Some computations in this paper were run on the Odyssey/Cannon cluster supported by the FAS Science Division Research Computing Group at Harvard University.

REFERENCES

Abazajian, K., et al. 2019. https://arxiv.org/abs/1907.04473

—. 2022, Astrophys. J., 926, 54, doi: 10.3847/1538-4357/ac1596

Abazajian, K. N., et al. 2016. https://arxiv.org/abs/1610.02743

Achúcarro, A., et al. 2022. https://arxiv.org/abs/2203.08128

Adachi, S., et al. 2020, Phys. Rev. Lett., 124, 131301, doi: 10.1103/PhysRevLett.124.131301

Adam, R., et al. 2016, Astron. Astrophys., 586, A133, doi: 10.1051/0004-6361/201425034

Ade, P., et al. 2019, JCAP, 02, 056, doi: 10.1088/1475-7516/2019/02/056

Ade, P. A. R., et al. 2018, Phys. Rev. Lett., 121, 221301, doi: 10.1103/PhysRevLett.121.221301

—. 2021, Phys. Rev. Lett., 127, 151301, doi: 10.1103/PhysRevLett.127.151301

Aghanim, N., et al. 2014, Astron. Astrophys., 571, A27, doi: 10.1051/0004-6361/201321556

—. 2020a, Astron. Astrophys., 641, A8, doi: 10.1051/0004-6361/201833886

—. 2020b, Astron. Astrophys., 641, A8, doi: 10.1051/0004-6361/201833886

Akrami, Y., et al. 2020, Astron. Astrophys., 641, A10, doi: 10.1051/0004-6361/201833887

Anderes, E., Wandelt, B., & Lavaux, G. 2015, Astrophys. J., 808, 152, doi: 10.1088/0004-637X/808/2/152

Aurlien, R., et al. 2022. https://arxiv.org/abs/2211.14342

Baleato Lizancos, A., Challinor, A., & Carron, J. 2021a, Phys. Rev. D, 103, 023518, doi: 10.1103/PhysRevD.103.023518

—. 2021b, JCAP, 03, 016, doi: 10.1088/1475-7516/2021/03/016

Baleato Lizancos, A., & Ferraro, S. 2022, Phys. Rev. D, 106, 063534, doi: 10.1103/PhysRevD.106.063534

Baumann, D., & McAllister, L. 2015, Inflation and String Theory, Cambridge Monographs on Mathematical Physics (Cambridge University Press), doi: 10.1017/CBO9781316105733

Beck, D., Errard, J., & Stompor, R. 2020, JCAP, 06, 030, doi: 10.1088/1475-7516/2020/06/030

Belkner, S., & Carron, J. 2023, delensalot, Zenodo, doi: 10.5281/zenodo.10352863. https://doi.org/10.5281/zenodo.10352863

Benoit-Levy, A., Dechelette, T., Benabed, K., et al. 2013, Astron. Astrophys., 555, A37, doi: 10.1051/0004-6361/201321048

Bianchini, F., & Millea, M. 2023, Phys. Rev. D, 107, 043521, doi: 10.1103/PhysRevD.107.043521

BICEP2/Keck Collaboration, Planck Collaboration, et al. 2015, PhRvL, 114, 101301, doi: 10.1103/PhysRevLett.114.101301

BICEP/Keck Collaboration, SPTpol Collaboration, et al. 2021, PhRvD, 103, 022004, doi: 10.1103/PhysRevD.103.022004

Carron, J. 2019, Phys. Rev. D, 99, 043518, doi: 10.1103/PhysRevD.99.043518

Carron, J., & Lewis, A. 2017, Phys. Rev. D, 96, 063510, doi: 10.1103/PhysRevD.96.063510

Carron, J., Lewis, A., & Challinor, A. 2017, JCAP, 1705, 035, doi: 10.1088/1475-7516/2017/05/035

Carron, J., Mirmelstein, M., & Lewis, A. 2022, JCAP, 09, 039, doi: 10.1088/1475-7516/2022/09/039

Challinor, A., & Chon, G. 2002, Phys. Rev. D, 66, 127301, doi: 10.1103/PhysRevD.66.127301

Challinor, A., et al. 2018, JCAP, 04, 018, doi: 10.1088/1475-7516/2018/04/018

Chon, G., Challinor, A., Prunet, S., Hivon, E., & Szapudi, I. 2004, Mon. Not. Roy. Astron. Soc., 350, 914, doi: 10.1111/j.1365-2966.2004.07737.x

Cooray, A., Kamionkowski, M., & Caldwell, R. R. 2005, Phys. Rev. D, 71, 123527, doi: 10.1103/PhysRevD.71.123527

Darwish, O., Sherwin, B. D., Sailer, N., Schaan, E., & Ferraro, S. 2023, Phys. Rev. D, 107, 043519, doi: 10.1103/PhysRevD.107.043519

Fabbian, G., Lewis, A., & Beck, D. 2019, JCAP, 10, 057, doi: 10.1088/1475-7516/2019/10/057

Ferraro, S., & Hill, J. C. 2018, Phys. Rev., D97, 023512, doi: 10.1103/PhysRevD.97.023512

Goldberg, J. N., MacFarlane, A. J., Newman, E. T., Rohrlich, F., & Sudarshan, E. C. G. 1967, J. Math. Phys., 8, 2155, doi: 10.1063/1.1705135

Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759, doi: 10.1086/427976

Grain, J., Tristram, M., & Stompor, R. 2009, PhRvD, 79, 123515, doi: 10.1103/PhysRevD.79.123515

Hager, W. W. 1989, SIAM Review, 31, 221, doi: 10.1137/1031049

Hamimeche, S., & Lewis, A. 2008, Phys. Rev. D, 77, 103013, doi: 10.1103/PhysRevD.77.103013

Hanany, S., et al. 2019. https://arxiv.org/abs/1902.10541

Hanson, D., Challinor, A., Efstathiou, G., & Bielewicz, P. 2011, Phys. Rev. D, 83, 043005, doi: 10.1103/PhysRevD.83.043005

Hanson, D., Challinor, A., & Lewis, A. 2010, Gen. Rel. Grav., 42, 2197, doi: 10.1007/s10714-010-1036-y

Hanson, D., Rocha, G., & Gorski, K. 2009, Mon. Not. Roy. Astron. Soc., 400, 2169, doi: 10.1111/j.1365-2966.2009.15614.x

Hirata, C. M., & Seljak, U. 2003a, Phys. Rev., D68, 083002, doi: 10.1103/PhysRevD.68.083002

—. 2003b, Phys. Rev., D67, 043001, doi: 10.1103/PhysRevD.67.043001

Hu, W., & Okamoto, T. 2002, Astrophys. J., 574, 566, doi: 10.1086/341110

Hu, W., & White, M. J. 1997, New Astron., 2, 323, doi: 10.1016/S1384-1076(97)00022-5

Kamionkowski, M., Kosowsky, A., & Stebbins, A. 1997, Phys. Rev. D, 55, 7368, doi: 10.1103/PhysRevD.55.7368

Kamionkowski, M., & Kovetz, E. D. 2016, Ann. Rev. Astron. Astrophys., 54, 227, doi: 10.1146/annurev-astro-081915-023433

Kesden, M., Cooray, A., & Kamionkowski, M. 2002, Phys. Rev. Lett., 89, 011304, doi: 10.1103/PhysRevLett.89.011304

Knox, L., & Song, Y.-S. 2002, Phys. Rev. Lett., 89, 011303, doi: 10.1103/PhysRevLett.89.011303

Lavaux, G., & Wandelt, B. D. 2010, Astrophys. J. Suppl., 191, 32, doi: 10.1088/0067-0049/191/1/32

Legrand, L., & Carron, J. 2022, Phys. Rev. D, 105, 123519, doi: 10.1103/PhysRevD.105.123519

—. 2023. https://arxiv.org/abs/2304.02584

Lembo, M., Fabbian, G., Carron, J., & Lewis, A. 2022, Phys. Rev. D, 106, 023525, doi: 10.1103/PhysRevD.106.023525

Lewis, A. 2005, Phys. Rev., D71, 083008, doi: 10.1103/PhysRevD.71.083008

Lewis, A., & Challinor, A. 2006, Phys. Rept., 429, 1, doi: 10.1016/j.physrep.2006.03.002

Lewis, A., Challinor, A., & Turok, N. 2002, Phys. Rev. D, 65, 023505, doi: 10.1103/PhysRevD.65.023505

Lewis, A., Hall, A., & Challinor, A. 2017, JCAP, 08, 023, doi: 10.1088/1475-7516/2017/08/023

Lyth, D. H., & Riotto, A. 1999, Phys. Rept., 314, 1, doi: 10.1016/S0370-1573(98)00128-8

MacCrann, N., et al. 2023. https://arxiv.org/abs/2304.05196

Maniyar, A. S., Ali-Haïmoud, Y., Carron, J., Lewis, A., & Madhavacheril, M. S. 2021, Phys. Rev. D, 103, 083524, doi: 10.1103/PhysRevD.103.083524

Manzotti, A. 2018, Phys. Rev. D, 97, 043527, doi: 10.1103/PhysRevD.97.043527

Millea, M., Anderes, E., & Wandelt, B. D. 2019, Phys. Rev. D, 100, 023509, doi: 10.1103/PhysRevD.100.023509

—. 2020, Phys. Rev. D, 102, 123542, doi: 10.1103/PhysRevD.102.123542

Millea, M., & Seljak, U. 2022, Phys. Rev. D, 105, 103531, doi: 10.1103/PhysRevD.105.103531

Millea, M., et al. 2021, Astrophys. J., 922, 259, doi: 10.3847/1538-4357/ac02bb

Mirmelstein, M., Carron, J., & Lewis, A. 2019, Phys. Rev. D, 100, 123509, doi: 10.1103/PhysRevD.100.123509

Namikawa, T. 2017, Phys. Rev. D, 95, 103514, doi: 10.1103/PhysRevD.95.103514

Namikawa, T., Yamauchi, D., & Taruya, A. 2012, JCAP, 01, 007, doi: 10.1088/1475-7516/2012/01/007

Namikawa, T., et al. 2022, Phys. Rev. D, 105, 023511, doi: 10.1103/PhysRevD.105.023511

Newman, E. T., & Penrose, R. 1966, J. Math. Phys., 7, 863, doi: 10.1063/1.1931221

Nocedal, J. 1980, Mathematics of Computation, 35, 773

Okamoto, T., & Hu, W. 2003, Phys. Rev., D67, 083002, doi: 10.1103/PhysRevD.67.083002

Omori, Y. 2022. https://arxiv.org/abs/2212.07420

Osborne, S. J., Hanson, D., & Doré, O. 2014, JCAP, 03, 024, doi: 10.1088/1475-7516/2014/03/024

Pratten, G., & Lewis, A. 2016, JCAP, 1608, 047, doi: 10.1088/1475-7516/2016/08/047

Qu, F. J., et al. 2023. https://arxiv.org/abs/2304.05202

Reinecke, M., Belkner, S., & Carron, J. 2023. https://arxiv.org/abs/2304.10431

Reinecke, M., & Seljebotn, D. S. 2013, Astronomy & Astrophysics, 554, A112, doi: 10.1051/0004-6361/201321494

Robertson, M., & Lewis, A. 2023, JCAP, 08, 048, doi: 10.1088/1475-7516/2023/08/048

Seljak, U. 1997, Astrophys. J., 482, 6, doi: 10.1086/304123

Seljak, U., & Zaldarriaga, M. 1997, Phys. Rev. Lett., 78, 2054, doi: 10.1103/PhysRevLett.78.2054

Smith, K. M., Hanson, D., LoVerde, M., Hirata, C. M., & Zahn, O. 2012, JCAP, 06, 014, doi: 10.1088/1475-7516/2012/06/014

Smith, K. M., Zahn, O., & Dore, O. 2007, Phys. Rev., D76, 043510, doi: 10.1103/PhysRevD.76.043510

Tegmark, M., de Oliveira-Costa, A., & Hamilton, A. 2003, Phys. Rev. D, 68, 123523, doi: 10.1103/PhysRevD.68.123523

Teng, W.-H., Kuo, C.-L., & Wu, J.-H. P. 2011. https://arxiv.org/abs/1102.5729

Tristram, M., et al. 2022, Phys. Rev. D, 105, 083524, doi: 10.1103/PhysRevD.105.083524

van Engelen, A., Bhattacharya, S., Sehgal, N., et al. 2014, Astrophys. J., 786, 13, doi: 10.1088/0004-637X/786/1/13

Vansyngel, F., Boulanger, F., Ghosh, T., et al. 2017, Astron. Astrophys., 603, A62, doi: 10.1051/0004-6361/201629992,10.1051/0004-6361/201629992e

Zaldarriaga, M., & Seljak, U. 1998, Phys. Rev. D, 58, 023003, doi: 10.1103/PhysRevD.58.023003

APPENDIX

A. SPIN-WEIGHTED FIELDS AND HARMONICS

Throughout this paper, we make heavy use of spin-weighted spherical harmonic (Newman & Penrose 1966; Goldberg et al. 1967; Lewis et al. 2002) transforms with spins between 0 and 3. A spin-weighted field ${}_sf(\hat{n})$ ($s \geq 0$) on the sphere and parametrized with colatitude and longitude (θ, φ) is defined with reference to the local axes, $\boldsymbol{e}_\theta$ and $\boldsymbol{e}_\varphi$. Our conventions are that $\boldsymbol{e}_\theta$ and $\boldsymbol{e}_\varphi$ point southwards and eastwards, respectively.

Under a clockwise rotation at $\hat{n}$ by some angle ψ, the field transforms, by definition, as

$$ {}_sf(\hat{n}) \rightarrow e^{is\psi(\hat{n})} \; {}_sf(\hat{n}). \tag{A1}$$

A spin-s field can conveniently be decomposed into its gradient (G) and curl (C) harmonics by using the spin-weighted spherical harmonics, ${}_sY_{\ell m}(\hat{n})$. The expansion has the form

$$ {}_{\pm s}f(\hat{n}) = -(\pm 1)^s \sum_{\ell m} (G_{\ell m} \pm iC_{\ell m}) \; {}_{\pm s}Y_{\ell m}(\hat{n}). \tag{A2}$$

Further, let $\eth^+$ and $\eth^-$ be the spin-raising and spin-lowering operators (Newman & Penrose 1966; Goldberg et al. 1967),

$$ \eth^\pm_s f = -(\sin\theta)^{\pm s} \left(\partial_\theta \pm \frac{i}{\sin\theta}\partial_\varphi \right) (\sin\theta)^{\mp s} \; {}_sf\,. \tag{A3}$$

We follow generally the conventions of (Lewis et al. 2002), and refer to the appendices of that paper for a discussion in the context of the CMB. With the help of these operators, one can introduce orthonormal basis functions (the spin-weighted spherical harmonics ${}_sY_{\ell m}$) for spin-weighted fields. Explicitly,

$$ \eth^\pm_s Y_{\ell m} = \pm\sqrt{(\ell \mp s)(\ell \pm s + 1)} \; {}_{s\pm 1}Y_{\ell m}. \tag{A4}$$

A lensing remapping on the sphere can be described by a vector field (concretely, a field of spin-1), and this vector field can, in turn, be described by two scalar (spin-0) lensing potentials, $\phi(\hat{n})$ and $\Omega(\hat{n})$, with the spin-1 deflection vector defined as follows:

$$ \begin{aligned} {}_1\alpha(\hat{n}) &= -\eth^+ \left(\phi(\hat{n}) + i\,\Omega(\hat{n})\right) \\ &= -\sum_{\ell m} \sqrt{\ell(\ell+1)}\,(\phi_{\ell m} + i\Omega_{\ell m}) \; {}_1Y_{\ell m}(\hat{n})\,. \end{aligned} \tag{A5}$$

The second line follows from Eq. (A4). As can be seen by comparing this to Eq. (A2), ϕ defines the gradient and Ω the curl component of the deflection vector. For a pure gradient deflection field, its real and imaginary components at $\hat{n}$ are simply $\partial_\theta\phi(\hat{n})$ and $\csc(\theta)\partial_\varphi\phi(\hat{n})$, and are generally denoted α_θ and α_φ. If convenient, we also make use of the polar representation of the vector, with α the magnitude of the field, and β the angle to the θ coordinate direction, such that, at each point on the sphere, the following relation holds:

$$ {}_{\pm 1}\alpha(\hat{n}) \equiv \alpha_\theta(\hat{n}) \pm i\alpha_\varphi(\hat{n}) \equiv \alpha(\hat{n})e^{\pm i\beta(\hat{n})}. \tag{A6}$$

B. DELENSED-NOISE MEAN-FIELD

This section gives details on the delensed-noise mean-field (the variation of the covariance matrix log-determinant with $\boldsymbol{\alpha}$) in the case of an otherwise isotropic configuration. Using the Woodbury determinant relation, the relevant log-determinant is

$$ \frac{1}{2}\ln\det\left[C^{EE,\mathrm{unl},-1} + N_{\boldsymbol{\alpha}}^{-1}\right] \equiv \frac{1}{2}\ln D\,, \tag{B7}$$

with the inverse delensed E-mode noise matrix

$$ N_{\boldsymbol{\alpha}}^{-1} = {}_2\mathcal{Y}^\dagger \mathcal{D}_{\boldsymbol{\alpha}}^\dagger \mathcal{B}^\dagger N^{-1} \mathcal{B}\mathcal{D}_{\boldsymbol{\alpha}}\,{}_2\mathcal{Y}. \tag{B8}$$

There are two variations to be considered. The first variation of the log-determinant can be written in the following way, using simple matrix algebra,

$$ \begin{aligned} \delta\ln D &= \mathrm{Tr}\, C^{EE,\mathrm{unl}}(C^{EE,\mathrm{unl}} + N_{\boldsymbol{\alpha}})^{-1}\delta\ln N_{\boldsymbol{\alpha}}^{-1} \\ &\equiv \mathrm{Tr}\,\mathcal{W}_{\boldsymbol{\alpha}}\delta\ln N_{\boldsymbol{\alpha}}^{-1}\,. \end{aligned} \tag{B9}$$

For convenience, we have used the notation $\delta \ln N_{\boldsymbol{\alpha}}^{-1}$ for the matrix $N_{\boldsymbol{\alpha}} \delta N_{\boldsymbol{\alpha}}^{-1}$. This matrix has roughly constant contributions on all scales. It is contracted with the Wiener-filtering matrix (defined as $\mathcal{W}_{\boldsymbol{\alpha}}$ in the second line), thereby suppressing contributions from noisy E-modes. The impact of $\boldsymbol{\alpha}$ in the Wiener-filter matrix is, however, only significant when the noise is comparable to $C^{EE,\mathrm{unl}}$. Hence, for low noise experiments, the isotropic Wiener filter $\mathcal{W}_\ell = C_\ell/(C_\ell + N_\ell)$, which is the leading-order contribution, already gives an almost exact result. We now discuss $\delta \ln N_{\boldsymbol{\alpha}}^{-1}$ in more detail, and then look at the second variation as well.

The action of lenses much larger than the beam commutes with the beam to good accuracy. It is useful to consider this approximation first, which results in the correct low-L behavior, and which also renders the results more transparent. This assumption also corresponds to the exact result in the case of zero beamwidth (but arbitrary large white noise level, so that this still corresponds to a somewhat physically relevant situation). Under this assumption, we can swap the order of appearance of the beam and deflection operations in Eq. (B8). Assuming homogeneous noise with full sky coverage, and using the exact relation $[\mathcal{D}_{\boldsymbol{\alpha}}^\dagger \mathcal{D}_{\boldsymbol{\alpha}}][\hat{n}, \hat{n}'] = |A_{\boldsymbol{\alpha}^{-1}}|(\hat{n}) \delta^D(\hat{n}, \hat{n}')$, the noise matrix greatly simplifies.

For convenience, we now adopt notation convenient for intensity measurements by using spin-0 spherical harmonics. However, the simulations discussed in this paper include polarized data and have no primordial B-modes, and we generally must use spin-2 harmonics to correctly describe it. These two differences lead to slight complications in the argument below. They are, nevertheless, only formal, and lead to identical leading-order and similar second-order results, and will be discussed at the end.

The inverse noise matrix, in the approximation discussed above, becomes

$$\left[N_{\boldsymbol{\alpha}}^{-1}\right]_{\ell m, \ell' m'} = \int d\hat{n} \frac{|A_{\boldsymbol{\alpha}^{-1}}|(\hat{n})}{N_\ell^{1/2} N_{\ell'}^{1/2}} Y_{\ell m}^\dagger(\hat{n}) Y_{\ell' m'}(\hat{n}) \,. \tag{B10}$$

Assuming that the bandlimits give a sufficiently large range, the spherical harmonics project onto a complete set, and the matrix has the explicit inverse

$$\left[N_{\boldsymbol{\alpha}}\right]_{\ell m, \ell' m'} = \int d^2\hat{n} \frac{N_\ell^{1/2} N_{\ell'}^{1/2}}{|A_{\boldsymbol{\alpha}^{-1}}|(\hat{n})} Y_{\ell m}^\dagger(\hat{n}) Y_{\ell' m'}(\hat{n}) \,. \tag{B11}$$

With this, we find that the variation $N_{\boldsymbol{\alpha}} \delta N_{\boldsymbol{\alpha}}^{-1}$ in Eq. (B9) is

$$\begin{aligned} &\left[\delta \ln N_{\boldsymbol{\alpha}}^{-1}\right]_{\ell m, \ell' m'} \\ &= \frac{N_\ell^{1/2}}{N_{\ell'}^{1/2}} \int d^2\hat{n}\, \delta \ln |A_{\boldsymbol{\alpha}^{-1}}|(\hat{n}) Y_{\ell m}^\dagger(\hat{n}) Y_{\ell' m'}(\hat{n}). \end{aligned} \tag{B12}$$

Further, we can then write the leading term of Eq. (B9) as

$$\delta \ln D = \left(\delta \int \frac{d^2\hat{n}'}{4\pi} \ln A_{\boldsymbol{\alpha}^{-1}}(\hat{n}')\right) \left(\sum_\ell (2\ell+1) \mathcal{W}_\ell\right) \,. \tag{B13}$$

By using $\ln A_{\boldsymbol{\alpha}^{-1}}(\hat{n}') = -\ln A_{\boldsymbol{\alpha}}(\hat{n})$, where $\hat{n}'$ and $\hat{n}$ are the deflected and undeflected positions, respectively, and performing the change of parameterization in the integral, the prefactor is

$$-\delta \int \frac{d^2\hat{n}}{4\pi} |A_{\boldsymbol{\alpha}}|(\hat{n}) \ln |A_{\boldsymbol{\alpha}}|(\hat{n}) \sim -\delta \int \frac{d^2\hat{n}}{4\pi} 2\kappa^2(\hat{n}) \,. \tag{B14}$$

By definition, $g_{LM}^{\mathrm{MF},\kappa} = \frac{1}{2} \delta \ln D / \delta \kappa_{LM}$, and we recover

$$g_{LM}^{\mathrm{MF},\kappa} = -2\kappa_{LM} \left(\sum_\ell \frac{2\ell+1}{4\pi} \mathcal{W}_\ell\right) \,, \tag{B15}$$

as given in the text in Eq. (24).

Let us now discuss the second variation. It is natural to continue working with $\ln A_{\boldsymbol{\alpha}^{-1}}$, since the second variation $\delta^2 \ln N_{\boldsymbol{\alpha}}^{-1}$ vanishes (see Eq. B12). For this reason, the only contribution is from the change of the Wiener filter. This may be written as

$$\delta^2 \ln D = \left[\mathrm{Tr}\, \mathcal{W}_{\boldsymbol{\alpha}} \delta \ln N_{\boldsymbol{\alpha}}^{-1} (1 - \mathcal{W}_{\boldsymbol{\alpha}}) \delta \ln N_{\boldsymbol{\alpha}}^{-1}\right] \,. \tag{B16}$$

The full result following from this is given in Eq. (B18). The squeezed limit, however, (low L, high ℓ, and $\ell \sim \ell'$) is easily extracted:

$$\left(\int \frac{d^2\hat{n}}{4\pi} |A_{\boldsymbol{\alpha}}| (\ln |A_{\boldsymbol{\alpha}}|)^2\right) \left(\sum_\ell (2\ell+1) \mathcal{W}_\ell (1 - \mathcal{W}_\ell)\right) \,. \tag{B17}$$

The prefactor is $\int (2\kappa)^2 +$ higher orders. This is of the same order in κ as the first variation, but with much lower relevance due to $\mathcal{W}_\ell(1-\mathcal{W}_\ell)$ and taking only contributions when $C_\ell \sim N_\ell$. The exact second variation result for spin-0 intensity is

$$\begin{aligned} \frac{1}{2}\ln D \ni & \frac{1}{2}\sum_{LM}|2\kappa_{LM}|^2 \\ & \cdot 2\pi \int_{-1}^{1} d\mu \, d^L_{00}(\mu)\xi^{\mathcal{W}}_{00}(\mu)\xi^{(1-\mathcal{W})}_{00}(\mu)\,, \end{aligned} \tag{B18}$$

with, for any spins a, b, and filter F,

$$\xi^F_{ab}(\mu) = \sum_\ell \frac{2\ell+1}{4\pi} F_\ell \, d^\ell_{ab}(\mu). \tag{B19}$$

For polarization, and in the limit of vanishing $C_\ell^{BB,\mathrm{unl}}$, the differences are the following. First, the Wiener-filter $\mathcal{W}_{\boldsymbol{\alpha}}$ becomes pure EE to leading order, but $(1-\mathcal{W}_{\boldsymbol{\alpha}})$ has both an EE and a BB component. This last term originates from the contribution of the EB part of the quadratic estimator. Second, we must use the spin-2 harmonics. The first variation is unchanged, but the second becomes

$$\begin{aligned} &\frac{1}{2}\ln D \ni \frac{1}{2}\sum_{LM}|2\kappa_{LM}|^2 \times \\ &\Bigg\{ 2\pi \int_{-1}^{1} d^L_{00}(\mu)\frac{1}{2}\left[\xi^{\mathcal{W}}_{2,2}(\mu)\xi^{1-\mathcal{W}}_{-2,-2}(\mu) + \xi^{\mathcal{W}}_{-2,2}(\mu)\xi^{1-\mathcal{W}}_{2,-2}(\mu)\right] \\ &+2\pi \int_{-1}^{1} d^L_{00}(\mu)\frac{1}{2}\left[\xi^{\mathcal{W}}_{2,2}(\mu)\xi^{1}_{-2,-2}(\mu) - \xi^{\mathcal{W}}_{-2,2}(\mu)\xi^{1}_{2,-2}(\mu)\right] \Bigg\}\,, \end{aligned}$$

where the second line is the EB contribution. Combining this with Eq. (B15) gives the final result for the mean-field for tiny beams,

$$\begin{aligned} g^{\mathrm{MF},\kappa}_{LM} &= -2\kappa_{LM}\Bigg\{\left[\sum_\ell \frac{2\ell+1}{4\pi}\mathcal{W}_\ell\right] \\ &- 2\pi \int_{-1}^{1} d^L_{00}(\mu)\frac{1}{2}\left[\xi^{\mathcal{W}}_{2,2}(\mu)\xi^{1-\mathcal{W}}_{-2,-2}(\mu) + \xi^{\mathcal{W}}_{-2,2}(\mu)\xi^{1-\mathcal{W}}_{2,-2}(\mu)\right] \\ &-2\pi \int_{-1}^{1} d^L_{00}(\mu)\frac{1}{2}\left[\xi^{\mathcal{W}}_{2,2}(\mu)\xi^{1}_{-2,-2}(\mu) - \xi^{\mathcal{W}}_{-2,2}(\mu)\xi^{1}_{2,-2}(\mu)\right] \Bigg\}\,. \end{aligned} \tag{B20}$$

The first two lines are the leading and sub-leading EE contributions, while the third line is the one from the EB-part of the quadratic estimator, which is basically zero at low L, as expected.

C. MAGNIFICATION DETERMINANT

In this section, we describe a series of steps to derive the exact form of the remapping Jacobian determinant in Eq. (15).

In standard polar coordinates, we use the local orthonormal basis at $\hat{n}$, yielding $\boldsymbol{e}_r = \hat{n}$, $\boldsymbol{e}_\theta$, and $\boldsymbol{e}_\varphi$. Variations along $\boldsymbol{e}_\theta$ and $\boldsymbol{e}_\varphi$ induce at the deflected position $\hat{n}'$ shifts $\partial_\theta \hat{n}'$ and $\frac{1}{\sin\theta}\partial_\varphi \hat{n}'$ in the plane tangent to the sphere at $\hat{n}'$. We can write these shifts using $\boldsymbol{e}_r$, $\boldsymbol{e}_\theta$, and $\boldsymbol{e}_\varphi$.

With this, the Jacobian determinant can be written as

$$|A| = |\partial_\theta \, \hat{n}' \times \partial_{\tilde{\varphi}} \, \hat{n}'|\,, \tag{C21}$$

where we have used the notation $\partial_{\tilde{\varphi}}$ for $\frac{1}{\sin\theta}\partial_\varphi$. The third deflection vector $\boldsymbol{d} = \hat{n}' - \hat{n}$ can be written in this basis as

$$\hat{n}' = \hat{n} + \boldsymbol{d} \equiv (1+\tilde{\alpha}_r)\boldsymbol{e}_r + \tilde{\alpha}_\theta \boldsymbol{e}_\theta + \tilde{\alpha}_\varphi \boldsymbol{e}_\varphi\,, \tag{C22}$$

with

$$1+\tilde{\alpha}_r = \cos\alpha, \quad \tilde{\alpha}_\theta = \frac{\sin\alpha}{\alpha}\alpha_\theta, \quad \tilde{\alpha}_\varphi = \frac{\sin\alpha}{\alpha}\alpha_\varphi. \tag{C23}$$

The variation of $\hat{n}'$ can be calculated from the variations of ${}_1\tilde{\alpha}$, and those of $\boldsymbol{e}_r$, $\boldsymbol{e}_\theta$, and $\boldsymbol{e}_\varphi$. A straightforward, albeit somewhat lengthy calculation, gives

$$\begin{aligned} &\partial_\theta \hat{n}' \times \partial_{\tilde{\varphi}} \hat{n}' \\ &= \begin{pmatrix} \partial_\theta \tilde{\alpha}_r - \tilde{\alpha}_\theta \\ 1 + \tilde{\alpha}_r + \partial_\theta \tilde{\alpha}_\theta \\ \partial_\theta \tilde{\alpha}_\varphi \end{pmatrix} \times \begin{pmatrix} \partial_{\tilde{\varphi}} \tilde{\alpha}_r - \tilde{\alpha}_\varphi \\ \partial_{\tilde{\varphi}} \tilde{\alpha}_\theta - \tilde{\alpha}_\varphi \cot\theta \\ 1 + \tilde{\alpha}_r + \tilde{\alpha}_\theta \cot\theta + \partial_{\tilde{\varphi}} \tilde{\alpha}_\varphi \end{pmatrix} . \end{aligned} \tag{C24}$$

With the definitions of the convergence and field rotation, $\tilde{\kappa} + i\tilde{\omega} = \frac{1}{2}\tilde{\eth}\, {}_1\tilde{\alpha}$, shear $\tilde{\gamma} = \frac{1}{2}\eth\, {}_1\tilde{\alpha}$, as well as a new spin-1 field ${}_1f = f_1 + if_2 \equiv {}_1\tilde{\alpha} + \eth\tilde{\alpha}_r$, we may write

$$\partial_\theta \hat{n}' \times \partial_{\tilde{\varphi}} \hat{n}' = \begin{pmatrix} (1 + \tilde{\alpha}_r - \tilde{\kappa})^2 + \tilde{\omega}^2 - \tilde{\gamma}^2 \\ -f_1(1 + \tilde{\alpha}_r - \tilde{\kappa} + \tilde{\gamma}_1) + f_2(-\tilde{\gamma}_2 - \tilde{\omega}) \\ -f_1(-\tilde{\gamma}_2 + \tilde{\omega}) + f_2(1 + \tilde{\alpha}_r - \tilde{\kappa} - \tilde{\gamma}_1) \end{pmatrix} . \tag{C25}$$

The first ($\boldsymbol{e}_r$) component reduces to the usual magnification matrix determinant $1 - 2\kappa$ for infinitesimal deflections. The $\boldsymbol{e}_\theta$ and $\boldsymbol{e}_\varphi$ components form the real and imaginary parts of a spin-1 field, giving a correction quadratic in the deflection angle.
To simplify all the calculations, we introduce and calculate with spin-0 quantities $\tilde{\eta}, \eta$ and $\tilde{\xi}, \xi$, built from the 3-dimensional and tangential deflection , respectively,

$$\begin{aligned} &\tilde{\eta} \equiv 1 + \tilde{\alpha}_r - \tilde{\kappa} + i\tilde{\omega}, \quad \eta \equiv 1 - \kappa + i\omega\,, \\ &\tilde{\xi} \equiv e^{-2i\beta}\tilde{\gamma}, \quad \xi \equiv e^{-2i\beta}\gamma\,. \end{aligned} \tag{C26}$$

A short calculation, applying $\eth$ on $\tilde{\eta}$ and $\tilde{\xi}$ shows that they are quite simply related to their tangential counterparts:

$$\begin{aligned} \begin{pmatrix} \tilde{\eta} \\ \tilde{\xi} \end{pmatrix} &= \begin{pmatrix} j_0(\alpha) - \frac{\alpha}{2} j_1(\alpha) & \frac{\alpha}{2} j_1(\alpha) \\ \frac{\alpha}{2} j_1(\alpha) & j_0 - \frac{\alpha}{2} j_1(\alpha) \end{pmatrix} \begin{pmatrix} \eta \\ \xi \end{pmatrix} \\ &- \frac{\alpha}{2} j_1(\alpha) \begin{pmatrix} 1 \\ 1 \end{pmatrix} , \end{aligned} \tag{C27}$$

with $j_0(\alpha) = \sin(\alpha)/\alpha$ and $\alpha j_1(\alpha) = -\cos(\alpha) + \sin(\alpha)/\alpha$ the first two spherical Bessel functions. With this notation in place, the $\boldsymbol{e}_r$ component of the vector product is

$$\begin{aligned} &(1 + \tilde{\alpha}_r - \tilde{\kappa})^2 + \tilde{\omega}^2 - \tilde{\gamma}^2 = |\tilde{\eta}|^2 - |\tilde{\xi}|^2 \\ &= \cos\alpha \left(j_0(\alpha) \left(|\eta|^2 - |\xi|^2 \right) - \frac{\alpha}{2} j_1(\alpha)(\eta + \eta^* - \xi - \xi^*) \right) . \end{aligned} \tag{C28}$$

The contributions of the $\boldsymbol{e}_\theta$ and $\boldsymbol{e}_\varphi$ can also be expressed easily in terms of η and ξ, with the exact same result as Eq. (C28), but with a $\sin\alpha$ prefactor replaced by $\cos\alpha$. The full Jacobian determinant then collects a factor $\cos^2\alpha$ and $\sin^2\alpha$ from the component parallel and perpendicular to $\hat{n}$, respectively, with the result

$$|A| = \left| j_0(\alpha) \left(|\eta|^2 - |\xi|^2 \right) - \frac{\alpha}{2} j_1(\alpha)(\eta + \eta^* - \xi - \xi^*) \right| , \tag{C29}$$

equivalent to the expression given in Eq. (15).

The first line of Eq. (15) looks familiar in the context of lensing, arguably less so the second. We can illustrate its role with the case of a locally constant deflection field, for which $\kappa = \omega = \gamma = 0$ locally, with the result $A(\hat{n}) = \cos\alpha$. Consider a tiny disk centred at the North Pole, with a locally constant deflection field there. If we increase the magnitude of the deflection field there gradually from zero, the circle starts to move towards the equator. The diameter of the disk parallel to the deflection vector stays unchanged, since the points are lying on the same great circle. The diameter transverse to the deflection gets smaller owing to focussing. Upon reaching the equator ($\alpha = \pi/2$), the disk has become a line, with the component parallel to the equator being squeezed to zero, consistent with $A = 0$. Increasing α further, at the South Pole ($\alpha = \pi$) the line has becomes a disk again, but with $A = -1$, since the transverse geodesics did cross each other on the equator.

D. CURVED-SKY LIKELIHOOD GRADIENTS

In this section, we give details on the derivation of the curved-sky likelihood gradients, Eq. (18). For this, we consider a small variation $\boldsymbol{\epsilon}$ to the deflection vector $\boldsymbol{\alpha}(\hat{n})$ in the plane tangent to $\hat{n}$ in Fig. 2. The deflection vector $\boldsymbol{\alpha}$ gives the coordinates of the deflected position $\hat{n}'$. If we were to slightly change the deflection vector, for example to $\boldsymbol{\alpha} + \boldsymbol{\epsilon}$, this would lead us to a position $\hat{n}'_\epsilon$. We can picture this as a squeezed spherical triangle, with two large sides, one of length $|\boldsymbol{\alpha}|$ joining $\hat{n}$ to $\hat{n}'$, the other of length $|\boldsymbol{\alpha}(\hat{n}) + \boldsymbol{\epsilon}(\hat{n})|$ joining $\hat{n}$ to $\hat{n}'_\epsilon$. The remaining, squeezed side of the triangle joining $\hat{n}'$ to $\hat{n}'_\epsilon$ lies, for practical purposes, in the plane tangent to $\hat{n}'$ and can be treated to first order in $\boldsymbol{\epsilon}$.

Our strategy is then as follows. Working in spin-weight components, we can associate to the change in deflection ${}_1\epsilon(\hat{n})$ at $\hat{n}$ a deflection field ${}_1\epsilon(\hat{n}')$ at $\hat{n}'$, and the remapping from $\hat{n}'$ to $\hat{n}'_\epsilon$ that it induces can be treated perturbatively. Doing so, and using Eq. (2) for the 'large' deflection and Eq. (6) for the small deflection, we can write

$$\begin{aligned} &2\frac{\delta}{\delta_{\pm1}\alpha(\hat{n})}\left[\mathcal{D}_{\boldsymbol{\alpha}}\ {}_s\mathcal{T}\right](\hat{n}) \\ &= -e^{is(\beta-\beta')}\left[\frac{\partial\ {}_{-1}\epsilon(\hat{n}')}{\partial\ {}_{\pm1}\epsilon(\hat{n})}\eth^+_s\mathcal{T}(\hat{n}') + \frac{\partial\ {}_1\epsilon(\hat{n}')}{\partial\ {}_{\pm1}\epsilon(\hat{n})}\eth^-_s\mathcal{T}(\hat{n}')\right]. \end{aligned} \tag{D30}$$

In this way, we only need to obtain the linear, explicit dependency of the components ${}_1\epsilon(\hat{n}')$ to ${}_1\epsilon(\hat{n})$.

It is convenient to split $\boldsymbol{\epsilon}(\hat{n}) = \boldsymbol{\epsilon}_\parallel(\hat{n}) + \boldsymbol{\epsilon}_\perp(\hat{n})$ in the components parallel and perpendicular to $\boldsymbol{\alpha}(\hat{n})$. Using, for example, simple spherical trigonometry, it is not too difficult to see that the parallel component gives a contribution $\boldsymbol{\epsilon}_\parallel(\hat{n})$ to the squeezed side of the spherical triangle, but the perpendicular contribution is reduced by the sky curvature to $\boldsymbol{\epsilon}_\perp(\hat{n})\,(\sin\alpha/\alpha)$.

The spin-weight components of $\boldsymbol{\epsilon}(\hat{n})$ at $\hat{n}$ are given by

$${}_1\epsilon(\hat{n}) = e^{i\beta}(\epsilon_\parallel(\hat{n}) + i\epsilon_\perp(\hat{n}))\,. \tag{D31}$$

At $\hat{n}'$, for the reason just discussed above, the components of the deflection vector pointing to $\hat{n}'_\epsilon$ become instead

$${}_1\epsilon(\hat{n}') = e^{i\beta'}\left(\epsilon_\parallel(\hat{n}) + i\left(\frac{\sin\alpha}{\alpha}\right)\epsilon_\perp(\hat{n})\right)\,. \tag{D32}$$

We can rearrange these two relations, and find one piece that is linear in the deflection angle α, and a second piece that is quadratic:

$$\begin{aligned} {}_1\epsilon(\hat{n}') = &\ e^{-i(\beta-\beta')}\ {}_1\epsilon(\hat{n}) \\ &- e^{-i(\beta-\beta')}\frac{1}{2}\left(1 - \frac{\sin\alpha}{\alpha}\right)\left[\ {}_1\epsilon(\hat{n}) - e^{2i\beta}\ {}_{-1}\epsilon(\hat{n})\right]. \end{aligned} \tag{D33}$$

Plugging this into Eq. (D30), and using the definition of the remapping by $\boldsymbol{\alpha}$, Eq. (2), recovers Eq. (18), as given in the text, after identification of $e^{\pm2i\beta}$ to $({}_{\pm1}\alpha(\hat{n})/\alpha(\hat{n}))^2$.

For $\alpha = \pi$, the perpendicular component is reduced to zero, since all great circles starting at $\hat{n}$ meet at the antipodal point. This provides a simple sanity check of these expressions: consider varying the deflection phase β, then the gradient of the deflected fields must exactly vanish. Since, by definition, ${}_{\pm1}\alpha(\hat{n}) = \alpha(\hat{n})e^{\pm i\beta(\hat{n})}$, we may write in general

$$\begin{aligned} &\frac{1}{i}\frac{\delta\left[\mathcal{D}_{\boldsymbol{\alpha}}\ {}_s\mathcal{T}\right](\hat{n})}{\delta\beta(\hat{n})} = \\ &{}_1\alpha(\hat{n})\frac{\delta\left[\mathcal{D}_{\boldsymbol{\alpha}}\ {}_s\mathcal{T}\right](\hat{n})}{\delta_1\alpha(\hat{n})} - {}_{-1}\alpha(\hat{n})\frac{\delta\left[\mathcal{D}_{\boldsymbol{\alpha}}\ {}_s\mathcal{T}\right](\hat{n})}{\delta_{-1}\alpha(\hat{n})}\,. \end{aligned} \tag{D34}$$

The spin-weighted gradients (Eq. 18) themselves become

$$\begin{aligned} &\frac{2\delta}{\delta_{\pm1}\alpha(\hat{n})}\left[\mathcal{D}_{\boldsymbol{\alpha}}\ {}_s\mathcal{T}\right](\hat{n}) \\ &\overbrace{=}^{\alpha(\hat{n})=\pi} -\frac{1}{2}[\mathcal{D}_{\boldsymbol{\alpha}}\eth^\mp_s\mathcal{T}](\hat{n}) - \frac{1}{2}\left(\frac{{}_{\mp1}\alpha(\hat{n})}{\alpha(\hat{n})}\right)^2[\mathcal{D}_{\boldsymbol{\alpha}}\eth^\pm_s\mathcal{T}](\hat{n})\,. \end{aligned} \tag{D35}$$

It is not difficult to see that the β gradient Eq. (D34) always vanishes, as it should.

E. OTHER MISCELLANEOUS FORMULAE

Generally, the undeflected angles can easily be obtained from the relation

$$\hat{n}' = \cos\alpha\,\hat{n} + \frac{\sin\alpha}{\alpha}\left(\alpha_\theta\boldsymbol{e}_\theta + \alpha_\varphi\boldsymbol{e}_\varphi\right)\,. \tag{E36}$$

The explicit relations are (Lewis 2005, for example)

$$\begin{aligned}\cos\theta' &= -\alpha_\theta \frac{\sin\alpha}{\alpha}\sin\theta + \cos\alpha\cos\theta\,,\\ \sin\theta'\sin(\varphi'-\varphi) &= \alpha_\varphi\frac{\sin\alpha}{\alpha}\,,\\ \sin\theta'\cos(\varphi'-\varphi) &= \alpha_\theta\frac{\sin\alpha}{\alpha}\cos\theta + \cos\alpha\sin\theta\,.\end{aligned} \tag{E37}$$

The explicit dependence on the deflection amplitude is always second order and weak, since $\alpha \simeq 2$ arcmin $= 5.8\times10^{-4}$. Conversely, the components of a deflection field can be obtained from the angles using

$$\begin{aligned}\alpha_\varphi\left(\frac{\sin\alpha}{\alpha}\right) &= \sin(\varphi'-\varphi)\sin\theta'\,,\\ \alpha_\theta\left(\frac{\sin\alpha}{\alpha}\right) &= \sin(\theta'-\theta) - 2\sin^2\left(\frac{1}{2}(\varphi'-\varphi)\right)\cos\theta\sin\theta'\,.\end{aligned} \tag{E38}$$

To leading order the phase of the lensing remapping is given by[32]

$$\beta-\beta' = \alpha_\varphi\left[\cot\theta - \frac{1}{2}\mathrm{csc}^2\theta(1+\cos^2\theta)\alpha_\theta + \cdots\right]\,, \tag{E39}$$

and can also easily be calculated in general from

$$\beta' = \mathrm{arctan}_2(\alpha_\varphi, \alpha\sin\alpha\cot\theta + \alpha_\theta\cos\alpha)\,. \tag{E40}$$

---

[32] (Challinor & Chon 2002) lacks a factor $\mathrm{csc}(\theta)$ in their Eq. 8.

3.2.2 Analysis of Alternatives

With the NSF requesting to study various experiment configuration of CMB-S4 to analyse feasibility of reaching the CMB-S4 science goals, we continuously work on optimizing the experiment capabilities.

As delensing is a critical part of the pipeline, we want to confirm that changes in the experiment configuration support the CMB-S4 science goals for map based delensing.

In this context, we explored the distribution of the number of telescopes across South Pole and Chile. For each of the final three configurations, we simulated noise maps and a sky scanning strategy, and applied different foreground models. We then generated component separated maps for each of them. I performed iterative lensing reconstruction with `Delensalot` and calculated the B-lensing templates for 100 simulations for each of the sky models and configurations.

I tested the results by calculating the residual lensing amplitudes that we would obtain with these sky models and configurations, and found that they were not in alignment with the expected delensing efficiencies, in particular for two out of three of the configurations.

Near the end of performing lensing reconstruction, it became clear that the noise simulations were erroneous. This error had not directly shown up in the estimation of the noise model due to a combination of bugs somewhat cancelling each other.

My work on this is summarized on the CMB-S4 Confluence website. They can be found with the following links (membership required):

1. https://cmb-s4.atlassian.net/wiki/spaces/XC/pages/1452343297/AoA+-+delensing+all+foregrounds+alternative+1
2. https://cmb-s4.atlassian.net/wiki/spaces/XC/pages/1453260808/AoA+-+delensing+all+foregrounds+alternative+2
3. https://cmb-s4.atlassian.net/wiki/spaces/XC/pages/1453817858/AoA+-+delensing+all+foregrounds+alternative+3

Ultimately, we decided to not move forward with any of the analysis results and instead redo the generation of the maps. On the bright side, it was good to test `Delensalot` for this task, hereby proving that the `DLM` indeed simplifies this task, allowing to easily handle the 9 different analyses, for which each configuration requires a different noise model, and overlapping B-mode deprojection matrix. Compared to the time of working on the collaboration paper, `Delensalot` is now much easier to control.

3.2.3 Optimal Component Separation

The goal of this project is to compare different pipelines for estimating r using map-based simulations, and we use the bias and sensitivity on r as a measure. While the CMB-S4 configuration and noise map is the same as in the collaboration paper of the previous section, we changed the foregrounds to various PySM models.

Foreground cleaning methods consist of a parametric maximum likelihood method using auto- and cross-power spectra between frequencies, a parametric maximum likelihood method based on frequency maps, and a harmonic-space internal linear combination based on frequency maps. This project uses `Delensalot` to generate lensing B-mode templates from LAT frequency maps.

We initially started with two simulations each. A quick sanity check uncovered a much better cross correlation coefficient to the input lensing potential. The reason for it was quickly identified as a better reconstruction from `Delensalot`: `Lenspyx 2.0` allows for a much more accurate deflection operation and subsequently also increases accuracy at the Wiener-filtering step.

However, two simulations were not enough for interpretable results which is why we ended up reconstructing 100 simulations per sky model.

In addition, to asses the bias from higher order correlations from these PySM foreground models on the lensing template, we additionally generated three sky models containing specific foreground parameterisation; they are constructed assuming that all information can be obtained from the two-point spectrum measurements. Compared to previous Gaussian foreground models, these are constructed, among other, using isotropic decorrelation and isotropic spectral energy distribution (SED) parameters.

My contribution to this project is the generation of the B-lensing templates for all sky models.

At the time of writing the thesis, this project is still ongoing. Results may be presented in a standalone paper in the near future.

3.3 PICO

PICO will scan the sky for 5 years in 21 frequency bands spread between 21 and 799 GHz from space; It will produce full-sky surveys of temperature and polarization with a resulting noise that would perform as 6400 Planck missions. With these capabilities, PICO will address seven science objectives. First, PICO will determine the energy scale of inflation and give a first, direct probe of quantum gravity. The mission will attempt to detect the tensor-to-scalar ratio r signal at a level of $r = 5 \cdot 10^{-4}$ at 5σ. This level is 100 times lower than current upper limits. PICO will constrain a wide range of classes of inflationary models, exclude models at 5σ for which the characteristic scale in the potential is the Planck scale and will be able to distinguish between reheating scenarios at 3σ. Combining PICO data with large-scale-structure (LSS) surveys (such as LSST) could rule out all models of slow-roll single-field inflation, marking a significant contribution to science. Second, the mission will have a strong impact on particle physics by measuring the minimum sum of neutrino masses at 4σ, a precision that can only be met with an instrument like PICO measuring the polarization of the CMB on the largest angular scales, while at the same time being able to remove large amounts of foreground contaminants by deploying a wide range of frequency detectors. An independent and equally competitive measure of the sum of neutrino masses can be achieved by using the cluster counts provided by PICO in combination with LSS data and in combination of the reconstructed lensing potential. Third, the PICO will either detect or strongly constrain deviations from the standard model of particle physics by counting the number of light particle species N_{eff} in the early Universe. The constraint of $\Delta N_{\mathrm{eff}} < 0.06(2\sigma)$ will move the allowed decoupling temperature of a hypothetical new vector particle to temperatures that are 400 times higher than the Planck constraints. Fourth, PICO will measure the optical depth to reionization τ with an error $\sigma(\tau) = 0.002$, a measurement that will be limited by cosmic variance from the largest polarization scales. Fifth, PICOs galactic dust polarization maps will make maps of the galactic magnetic field. Seventh, PICO will enable a search for primordial magnetic fields with good sensitivity: PICO will rule them out as the sole source for the largest observed galactic magnetic fields. PICO will improve constraint on polarization rotation from early Universe magnetic fields causing cosmic birefringence by a factor of 300, and thus constrain string theory-motivated axions.

I joined PICO in 2022, and supported the collaboration with the iterative lensing reconstruction and calculation of the residual lensing amplitudes for the simulations they provided. The PICO maps are clearly different compared to Planck or other CMB experiments: quite astonishing is the feature-less galactic plane, a feature only reached by the shear amount of different frequency channels that PICO can provide for the component separation.

For the paper that is presented in the next section, we studied various component separation and sky models and assessed PICOs capabilities in constraining r for two different configurations: the baseline configuration, and the best estimate configuration. Further, we divided the sky into sky patches of different residual foreground levels, to analyse the improvements on r with this approach.

With all the above being based on forecast approaches, we decided to validate our findings using map-based delensing. For one of the sky models, we generated 10 high-resolution simulations to perform map-based delensing.

With this exceptionally effective component separation from PICO, we decided to work on full-sky maps without a galactic mask, which turned out to give very good results, well in agreement with forecasts.

My contributions to this paper is the QE and iterative reconstruction, and the analysis of the map-level delensing. I supported the paper by contributing to writing the chapter about map-level delensing.

The following contains the publication as appeared in the Journal of Cosmology and Astroparticle Physics (JCAP).

Draft version October 26, 2024
Typeset using LaTeX default style in AASTeX631

Foreground Separation and Constraints on Primordial Gravitational Waves with the PICO Space Mission

Ragnhild Aurlien[1], Mathieu Remazeilles[2], Sebastian Belkner[3], Julien Carron[3], Jacques Delabrouille[4], Hans Kristian Eriksen[1], Raphael Flauger[5], Unni Fuskeland[1], Mathew Galloway[1], Krzysztof M. Górski[6,7], Shaul Hanany[8], Brandon S. Hensley[9], J. Colin Hill[10,11], Charles R. Lawrence[12], Clem Pryke[8], Alexander van Engelen[13] and Ingunn Kathrine Wehus[1]

[1] *Institute of Theoretical Astrophysics, University of Oslo, Blindern, Oslo, Norway*
[2] *Instituto de Fisica de Cantabria (CSIC-UC), Avda. los Castros s/n, 39005 Santander, Spain*
[3] *Université de Genève, Département de Physique Théorique et CAP, Genève 4, Switzerland*
[4] *Centre Pierre Binétruy International Research Laboratory, CNRS, UC Berkeley and LBNL, Berkeley, CA 94720, USA*
[5] *UC San Diego, La Jolla, CA, 92093, USA*
[6] *Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA*
[7] *Warsaw University Observatory, Aleje Ujazdowskie 4, 00-478 Warszawa,Poland*
[8] *University of Minnesota - Twin Cities, 115 Union St. SE, Minneapolis, MN, 55455, USA*
[9] *Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA*
[10] *Department of Physics, Columbia University, New York, NY 10027, USA*
[11] *Center for Computational Astrophysics, Flatiron Institute, New York, NY 10010, USA*
[12] *Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109*
[13] *School of Earth and Space Exploration, Arizona State University, Tempe, AZ 85287*

ABSTRACT

PICO is a concept for a NASA probe-scale mission aiming to detect or constrain the tensor to scalar ratio r, a parameter that quantifies the amplitude of inflationary gravity waves.

We carry out map-based component separation on simulations with five foreground models and input r values $r_{\rm in} = 0$ and $r_{\rm in} = 0.003$. We forecast r determinations using a Gaussian likelihood assuming either no delensing or a residual lensing factor $A_{\rm lens}$= 27%. By implementing the first full-sky, post component-separation, map-domain delensing, we show that PICO should be able to achieve $A_{\rm lens}$= 22% - 24%. For four of the five foreground models we find that PICO would be able to set the constraints $r < 1.3 \times 10^{-4}$ to $r < 2.7 \times 10^{-4}$ (95%) if $r_{\rm in} = 0$, the strongest constraints of any foreseeable instrument. For these models, $r = 0.003$ is recovered with confidence levels between 18σ and 27σ. We find weaker and, in some cases, significantly biased upper limits when removing a few low or high-frequency bands. The fifth model gives a 3σ detection when $r_{\rm in} = 0$ and a 3σ bias with $r_{\rm in} = 0.003$. However, by correlating r determinations from many small 2.5% sky areas with the mission's 555 GHz data we identify and mitigate the bias. This analysis underscores the importance of large sky coverage. We show that when only low multipoles $\ell \leq 12$ are used, the non-Gaussian shape of the true likelihood gives uncertainties that are on average 30% larger than a Gaussian approximation.

Keywords: (cosmology:) cosmic microwave background — (cosmology:) early universe — (cosmology:) diffuse radiation — (cosmology:) observations — methods: data analysis

1. INTRODUCTION

Observations of the cosmic microwave background (CMB) from three space missions, COBE-DMR, WMAP, and Planck have established ΛCDM as the widely accepted standard model of cosmology, have give a wealth of information about the evolution of structures in the Universe, and have been used to support numerous astrophysical investigations. Several missions have been proposed to take the next leap in understanding the evolution of the Universe, including LiteBIRD (LiteBIRD Collaboration et al. 2022), which has been approved by the Japanese space agency JAXA, CORE (Delabrouille et al. 2018), CMB-Bhārat (Adak et al. 2021), and PICO (Hanany et al. 2019). PICO, The Probe of Inflation and Cosmic Origins, a concept for a NASA-led probe-scale space mission (Hanany et al. 2019), is the most sensitive of the proposed next generation space missions. For example, relative to LiteBIRD it would give maps that are deeper by at least a factor of 2.5 with nearly four times higher angular resolution in overlapping bands. PICO is expected to give the tightest constraints on cosmological parameters compared to all forthcoming or planned CMB instruments.

PICO has seven science objectives, five of which are to be extracted from the highest ever signal-to-noise ratio (S/N) full-sky maps of the CMB. After PICO's prime mission of 5 years, the combined map-noise level over the entire sky would be 0.6 μK arcmin. Since the mission has no liquid cryogens, longer lifetime and lower noise are likely. Of the numerous probe-scale mission concepts submitted to the Astro2020 panel, a next-generation CMB mission was one of three recommended for further development this decade for a possible flight in the 2030s (National Academies of Sciences, Engineering, and Medicine 2021).

PICO's most demanding requirement is the level of constraint on the tensor to scalar ratio r, which quantifies the amplitude of gravitational waves produced during the epoch of inflation shortly after the big bang. For the case of a null detection, the requirement is to achieve $r \leq 0.0002\,(95\%)$. If $r \neq 0$ the requirement is to achieve 5σ detection of $r = 0.0005$. A priori, in the context of inflation, the expected value of r can range over many orders of magnitude, including values far too small to be detected. However, the simplest models of inflation that have a single inflaton, and that naturally explain the observed value of the scalar spectral index, only have a single free parameter, the distance in field space over which the potential varies appreciably. In several well-motivated scenarios, this scale shares a common origin with the scale of gravity, typically referred to as the Planck scale. Such models predict values $r \gtrsim 0.001$ (Abazajian et al. 2016; Linde 2017). The absence of detection at this level would definitively rule out this particularly well-motivated class of inflationary models (Shandera et al. 2019). Conversely, an unambiguous detection would definitively establish inflation as the source of primordial perturbations, determine the energy scale at which inflation took place, and would give a first direct probe of quantum gravity. The current upper limit is $r < 0.032\,(95\%)$ (Ade et al. 2021; Tristram et al. 2021).

CMB determinations of r rely on measuring the polarization Q and U Stokes parameters, converting them to E- and B-modes, and forming the corresponding EE and BB angular power spectra. The level of r is linearly proportional to the BB power spectrum, $r = 0.15 \cdot \mathcal{D}^{\rm BB}_{\ell=90}/(0.1\,\mu{\rm K})^2$, where $\mathcal{D}^{\rm BB}_{\ell=90} = \ell(\ell+1)C^{\rm BB}_{\ell}/2\pi$ is the level at $\ell = 90$. Reaching PICO's constraint on inflationary gravitational waves is demanding – even with its high S/N maps – because the B-mode foreground emission within the Milky Way is known to have amplitudes much larger than the levels of r targeted by PICO (Planck Collaboration et al. 2020a; Ade et al. 2021). Separating the cosmological signal from the foregrounds, a process commonly called "component separation," requires low-noise multi-frequency observations, or accurate prior knowledge, that are unlikely to be available before the mission flies, because it would take a PICO-type mission to map the sky with the requisite S/N.

Gravitational lensing of CMB photons by large scale structures between the surface of last scattering and our telescopes presents another challenge (Lewis & Challinor 2006). Lensing scatters photons off their original paths, distorting slightly the original pattern of the CMB anisotropy. In polarization, the primary effect is the conversion of high ℓ E-mode power to lower ℓ B-mode power. The sample variance of this lensing-induced B-mode acts as a source of approximately white noise. However, with sufficiently high S/N polarimetric measurements extending to $\ell \gtrsim 1000$, CMB maps can be delensed (Zaldarriaga & Seljak 1998; Hu 2002; Seljak & Hirata 2004) and delensing improves constraints on cosmological parameters. PICO's resolution of few arcmin at the main CMB bands near 200 GHz makes internal delensing – the use of its own data to delens – possible (Hu & Okamoto 2002; Okamoto & Hu 2003), which is a key element in achieving the stringent science goal for r.

In this paper we use simulations to study whether data from PICO's 21 frequency bands (Section 2) would enable foreground cleaning such that the required level of constraint on r can be achieved. We construct sky maps matching five possible foreground models, all of them broadly consistent with current Planck data (Section 3). They span a broad range of complexity and input assumptions. The sky maps, which include CMB signals and noise (Section 4), approximately represent maps that would be obtained with PICO. The maps are approximate because we assume a spatial noise distribution that is pixel-independent and scale-free with ℓ, an appropriate approximation at this stage of mission development. We apply both parametric and blind component separation methods and estimate r. Details of the methods and results from each are given in Sections 5 and 6. Delensing of the sky maps is handled in two ways, an analytic, power spectrum domain approach, and using an iterative, map domain delensing algorithm (Section 7). We discuss and summarize in Sections 8 and 9.

Testing the component separation capability of various instruments has been carried out in the past. The recent most relevant reports were those provided in the context of CORE (Remazeilles et al. 2018), LiteBIRD and PICO (LiteBIRD Collaboration et al. 2022; Remazeilles et al. 2021), CMB-Bharat (Adak et al. 2021), Simons Observatory (Ade et al. 2019), and CMB-S4 (Abazajian et al. 2022). This paper is unique because with the PICO data we strive to set unprecedented constraints on cosmological parameters, because we analyze a diverse set of five sky foreground models spanning a realistic range of possibilities, and because this is the first paper to report on map-level iterative delensing as an integral part of the analysis pipeline.

2. PICO INSTRUMENT PARAMETERS

The PICO mission concept has a single instrument, an imaging polarimeter, operating in 21 frequency bands between 21 and 799 GHz. The instrument consists of a two mirror telescope with a 1.4 m diameter entrance aperture, which is based on an 'open-Dragone' design (Young et al. 2018). The telescope feeds a focal plane populated with 13,000 bolometers and gives a resolution between 1.1 arcmin at the highest frequency and 38.4 arcmin at the lowest. The bolometers are operated from a bath temperature of 0.1 K. PICO will conduct observations from the L2 Lagrange point with a scan pattern that covers the full sky within 6 months. The prime mission duration is 5 years giving 10 redundant full sky surveys, but because there are no consumables, mission lifetime could extend significantly longer. The mission concept has required and estimated noise levels of 0.87 and 0.61 μK arcmin, respectively, accounting only for the 5 yr prime mission. In this paper we use the estimated map noise levels, as given in Table 1. The simulations that are used in this work began before the concept report was finalized and due to late iterations on the design of the focal plane some of the values in the table differ slightly from the values quoted in the more definitive final report (Hanany et al. 2019). The differences are minor and do not change the combined map noise level. The frequency bands and their noise levels as given in Table 1 were not optimized for component separation. Such optimization is left for future work; see Section 8.

Table 1. Parameters of the PICO mission. The map noise is assumed white and the numbers given are for the polarized maps (Q/U or E/B).

Frequency (GHz)	21	25	30	36	43	52	62	75	90	108	129	155	186	223	268	321	385	462	555	666	799
FWHM (arcmin)	38.4	32.0	28.3	23.6	22.2	18.4	12.8	10.7	9.5	7.9	7.4	6.2	4.3	3.6	3.2	2.6	2.5	2.1	1.5	1.3	1.1
Noise† (μK-arcmin)	16.9	11.8	8.1	5.7	5.8	4.1	3.8	2.9	2.0	1.6	1.6	1.3	2.6	3.0	2.1	2.9	3.5	7.4	34.6	144	896

Combined Map Noise 0.61 (μK-arcmin)

† Polarization noise in CMB thermodynamic units

3. FOREGROUND MODELS

We consider five models for Galactic foreground emission, as described in the following sections and summarized in Table 2. For ease of reference, the models are given the names and short names as shown in the first column. For levels of r relevant for upcoming CMB instruments including PICO, $0.5 \times 10^{-3} \leq r \leq 5 \times 10^{-3}$, the B-mode signal is much smaller than Galactic foregrounds, see Figure 1. The figure shows the level of foregrounds expected at three frequency bands with the five models and the level of CMB B-modes as a function of ℓ. Models 2MBB, PhysDust, and MHD, are identical to those used in Abazajian et al. (2022).

3.1. *Model Planck Baseline*

The 'Baseline' model is based on polarized Galactic emission provided by the Python Sky Model (PySM2; Thorne et al. 2017), specifically the "d1" model of polarized dust emission and the "s1" model of polarized synchrotron emission. In each sky pixel the dust emission is characterized by Planck's 353 GHz polarized intensity in Q and U (Planck Collaboration X 2016) and by an opacity law index β_d and dust temperature T_d to describe the frequency dependence. The Q and U 353 GHz amplitude maps, as well as the T_d and β_d maps, are based on the Planck Collaboration X (2016) component separation analysis with the Commander framework. Gaussian fluctuations are added to low-pass filtered amplitude maps at angular scales $\ell > 69$, where the data are not constraining, to produce final maps with power at scales up to $\ell \sim 1500$ (Thorne et al. 2017).

Table 2. Summary of Polarized Foreground Models

Model Name (Short Name)	Dust Model	Synchrotron Model	Other Components
Planck Baseline (Baseline)	PySM d1: modified blackbody with spatially varying T_d and β_d	PySM s1: power law spectrum with spatially varying β_s	None
Dust: Two Modified Black Bodies (2MBB)	PySM d4: two component dust model of Meisner & Finkbeiner (2015)	PySM s3: power law spectrum with spatially varying β_s and sky-constant curvature	PySM a2 AME model: Spatially varying spectrum with fixed 2% polarization fraction
Physical Dust (PhysDust)	PySM d7: physical dust model of Hensley (2015) including magnetic dipole emission	PySM s3	PySM a2 AME model
MHD (MHD)	Modified blackbody dust emission in each cell of a TIGRESS MHD simulation (Kim & Ostriker 2017; Kim et al. 2019), integrated along the line of sight	Power law synchrotron spectrum with amplitude coupled to B-fields in a TIGRESS MHD simulation (Kim & Ostriker 2017; Kim et al. 2019)	None
Multi-Layer Dust (Multi-Layer)	"MKD" dust model (Martínez-Solaeche et al. 2018) based on multiple modified blackbody emission laws in each pixel	PySM s3	PySM a2 AME model

The synchrotron emission in each pixel is likewise described by an amplitude in each of Q and U and by a power law index β_s to describe the frequency dependence. The Q and U amplitudes are based on the 9-yr WMAP 23 GHz Q and U maps smoothed to a resolution of 3°. The β_s map is taken from "Model 4" of Miville-Deschênes et al. (2008), who used the Haslam 408 MHz survey data (Haslam et al. 1982) and 3-yr WMAP 23 GHz data (Hinshaw et al. 2007) to derive synchrotron spectral indices.

3.2. *Model Dust: Two Modified Black Bodies*

Model 2MBB is also based on polarized Galactic emission models from PySM, but differs from the Baseline in the frequency scalings of both dust and synchrotron emission as well as inclusion of a polarized anomalous microwave emission (AME) component. Specifically, model 2MBB employs the "d4" dust model, "s3" synchrotron model, and "a2" AME model.

The dust emission in model d4 is based on the parametric fits of Meisner & Finkbeiner (2015), who employed a two component dust model to describe a combination of Planck and DIRBE/IRAS data. The emission in each pixel is specified by an amplitude in each of Q and U and by a dust temperature for each of the two components to describe the frequency dependence. The temperature maps are taken directly from the fits of Meisner & Finkbeiner (2015). The Q and U amplitude maps are generated by scaling the model to 353 GHz in total intensity, then applying the same polarization angle and polarization fraction as used in the d1 model. Gaussian small scale fluctuations are added as in the d1 model. Thus, the 353 GHz Q and U amplitude maps of the d4 model differ in detail from those of the d1 model.

The s3 synchrotron model is in all respects identical to the s1 model described in the previous section with the addition of a curvature term to the frequency scaling. The synchrotron curvature parameter is taken to be constant over the full sky.

The AME emission in the a2 model is based on the parametric fits of Planck Collaboration X (2016), who describe the AME spectrum as the sum of two spinning dust components. The a2 model in total intensity is described by three parameter maps specifying the amplitudes of each of these components and the peak frequency of one. The

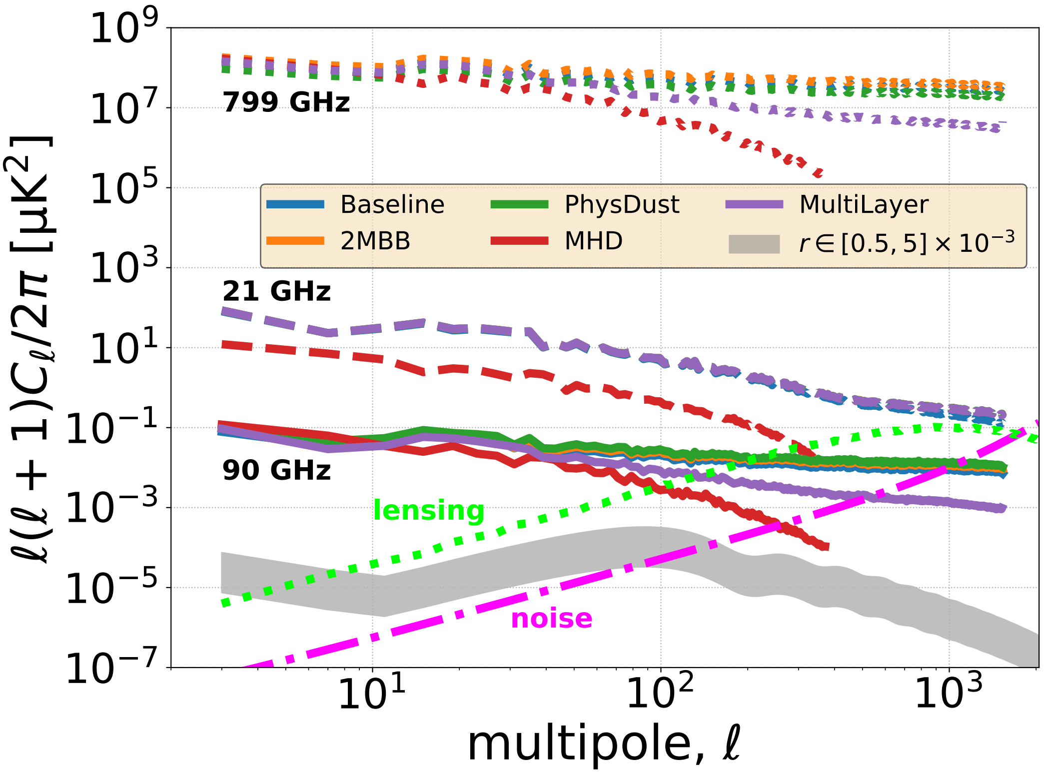

Figure 1. B-mode power spectra from dust and synchrotron at 21 GHz (dash), 90 GHz (solid), and 799 GHz (dots) for different foreground models on 46% of the sky compared to the primordial CMB spectrum (gray) with tensor-to-scalar ratio values $0.5 \times 10^{-3} \leq r \leq 5 \times 10^{-3}$, lensing (light green dots), and to the PICO instrument noise level (magenta dash dot). All models except MHD have information up to $\ell \simeq 1500$. Model MHD, based on a 3D MHD simulation with finite resolution, has information only up to $\ell \simeq 380$. Synchrotron emission dominates at 21 GHz and the two PySM synchrotron models used for models Baseline, 2MBB, PhysDust, and MultiLayer overlap (purple dash). Dust emission dominates at 799 GHz and the different models are constrained to match available data at low ℓ.

peak frequency of the other is held fixed across the sky. The parameter maps are taken directly from the Planck Collaboration X (2016) fits. To construct Q and U maps, it is assumed that the AME has the same polarization angle as the Planck Collaboration X (2016) maps of 353 GHz polarized dust emission and a sky-constant polarization fraction of 2%.

3.3. *Model Physical Dust*

Model PhysDust is identical in all respects to 2MBB with the exception of the model for polarized dust emission, where the "d7" model is used instead of d4. Polarized dust emission in the d7 model is based on a physical model of interstellar grains with a distribution of sizes and temperatures (Hensley 2015). This model notably includes grains with magnetic iron inclusions, which emit significant magnetic dipole radiation at frequencies $\lesssim 100$ GHz (Draine & Hensley 2013).

The Q and U dust amplitude maps of the d7 model are identical to those in the d1 model (employed in our model Baseline). The frequency scaling is based on a single parameter $\mathcal{U}$ quantifying the intensity of radiation heating the grains. A map of $\mathcal{U}$ is constructed from the Planck Collaboration X (2016) dust temperature maps assuming $\mathcal{U} \propto T_d^{4+\beta_d}$ and that $T_d = 20$ K corresponds to $\mathcal{U} \simeq 1$. The frequency scaling is based on tabulated model calculations and is not simply expressed with a parametric formula.

3.4. *Model MHD*

Unlike the Baseline, 2MBB, and PhysDust models that are taken from PySM, model MHD is based on the output of the large-scale TIGRESS magnetohydrodynamic (MHD) simulations (Kim & Ostriker 2017). Kim et al. (2019)

constructed full sky maps of Galactic emission from these simulations, and we use maps generated in a similar way here. Model MHD includes only polarized dust and synchrotron components.

The polarized dust emission in model MHD is based on the gas density and magnetic field orientation in each grid cell of the MHD simulation. The frequency scaling is taken to be a modified blackbody in each cell, with the dust temperature set by the local radiation field and a β_d value having a dependence on the local gas density. The 3D simulation cube is then integrated along the line of sight to produce full sky maps following Kim et al. (2019). The fact that the dust frequency spectrum can vary along the line of sight introduces line of sight frequency decorrelation (Tassis & Pavlidou 2015; Martínez-Solaeche et al. 2018; Pelgrims et al. 2021) in addition to decorrelation induced by spectral parameters that vary across the sky, which is present in the Baseline, 2MBB, and PhysDust models.

The polarized synchrotron emission in model MHD is based on the 3D magnetic field geometry of the simulation coupled with a simple parametric model of the cosmic ray electron number density with Galactic scale height (see Choi & Page 2015). The synchrotron frequency spectrum is taken to be a power law with constant spectral index across the sky.

3.5. *Model Multi-Layer Dust*

Model MultiLayer is a realization of thermal dust emission with the multi-layer dust model of Martínez-Solaeche et al. (2018), updated to be included in the Planck Sky Model (version 2.2.3, Delabrouille et al. 2013). The key idea of the model is that if parameters describing the frequency scaling of dust emission vary across the sky, they must also vary along the line of sight. Similar to model MHD, this introduces line of sight frequency decorrelation. The total emission at 353 GHz is modelled as the sum of emissions from six dust template maps (loosely associated with six layers of distance from the observer), the sum of which is constrained to the total dust emission at this frequency in the Planck intensity and polarization maps (Planck Collaboration et al. 2016, 2020a) as obtained with the GNILC component separation method (Remazeilles et al. 2011a).

Explicitly, in each pixel

$$Q_\nu^{\rm MKD} = \sum_k A_{d,k}^Q \left(\frac{\nu}{\nu_0}\right)^{\beta_{d,k}} B_\nu\left(T_{d,k}\right) \tag{1}$$

$$U_\nu^{\rm MKD} = \sum_k A_{d,k}^U \left(\frac{\nu}{\nu_0}\right)^{\beta_{d,k}} B_\nu\left(T_{d,k}\right) \quad , \tag{2}$$

where index k, running from 1 to 6, identifies a layer of emission. In practice, at high Galactic latitude only the first three layers contribute to the total emission. Closer to the Galactic plane, all six layers have non-vanishing contributions because even the more distant layer contain a significant amount of dust. These dust maps are added to PySM maps of other Galactic components to produce a full emission model enumerated dms3a2f1.

4. COMPONENT SEPARATION APPROACHES AND SKY MAPS

There are two broad classes of component separation approaches, parametric and blind (e.g., Leach et al. 2008; Delabrouille & Cardoso 2009). With the parametric approach one parametrizes the spectral dependence of the emission law for a given foreground emission component, estimates the free parameters given the data, and produces a map of the cosmological signal by marginalizing over the foreground parameters. The premise with the blind approach is not to assume any specific model for the foreground emission. Rather, one exploits the statistical independence of emission from different physical origins to separate them using observations at different frequencies. In particular, assuming that the frequency scaling of only the CMB is known, one can form linear combinations of the various observations that minimize residual foreground and noise.

We use the blind approach with all five sky models. The specific implementation is NILC (Delabrouille et al. 2009). We use the parametric approach only with Baseline model, and the specific implementation is Commander1 (Eriksen et al. 2004, 2008). In Commander1 component separation is carried out on each map pixel. While the algorithm is computationally fast, it requires uniform angular resolution for all frequency maps. More advanced versions of Commander (Seljebotn et al. 2017; Galloway et al. 2022) can account for beam size differences, but require significantly more compute time. For estimating r, most of the relevant information comes from $\ell < 150$, and for the very low values of r that PICO targets and at which the lensing signal dominates, the reionization peak at $\ell \lesssim 12$ is particularly important. In this paper, the Commander analysis is restricted to $\ell \leq 12$.

We make simulated PICO-observed Q, and U sky maps for NILC by adding the Galactic foreground models described above to realizations of CMB and experimental noise, using the same methods as used by Abazajian et al. (2022). Realizations of lensed-ΛCDM are borrowed from the Planck FFP10 simulations (Planck Collaboration et al. 2020b), as are realizations of tensor modes with $r = 0$ and $r = 0.003$. The foreground and CMB maps are smoothed assuming Gaussian beam shapes with the FWHM values given in Table 1. Gaussian noise is generated in harmonic space with the levels given in Table 1. The noise is assumed to have a flat spectrum in C_ℓ. The foregrounds, CMB, and noise are added together to produce simulated observed sky maps. While there is only one realization for each of the foreground models, we add these to multiple realizations of CMB and noise. The simulations are done at a HEALPix (Górski et al. 2005) resolution $N_{\rm side} = 512$, except when we are doing lensing reconstruction and map-based delensing, for which additional maps are rendered at $N_{\rm side} = 2048$; see Section 7. The foreground component uses the same ℓ space information at $N_{\rm side} = 2048$ as at $N_{\rm side} = 512$.

Maps for Commander are constructed largely the same way as they are for NILC except for the following differences: (1) isotropic, homogeneous Gaussian noise is generated in pixel space, not harmonic space; (2) because Commander1 requires a common beam, the CMB and foreground signals have been smoothed with a Gaussian beam of 40 arcmin FWHM for all frequency bands. The noise is not smoothed; (3) as stated earlier, with Commander we only use model Baseline.

5. COMPONENT SEPARATION - NILC

5.1. *Methodology*

We compute the spherical harmonic transform of the PICO-observed Q_ν and U_ν full-sky maps at frequency ν to obtain the harmonic coefficients $a^E_{\ell m,\nu}$ and $a^B_{\ell m,\nu}$. For the analysis and results presented in this section we maintain only the B-mode coefficients $a^B_{\ell m,\nu}$, calculate their inverse spherical harmonic transform, and obtain full-sky B-mode maps B_ν at each frequency. The subsequent component separation process only requires the B_ν maps and its goal is to produce a best estimate of the underlying CMB B-mode map $\hat{B}^{\rm NILC}$. However, the map-domain delensing analysis presented in Section 7 also requires a best estimate component-separated CMB E-mode map $\hat{E}^{\rm NILC}$, and to produce this map we use the $a^E_{\ell m,\nu}$ coefficients to obtain E-mode maps E_ν. Although subsequent paragraphs in this section refer to component separation using B_ν, the process is identical for E_ν. Using full-sky maps prevents E-to-B leakage.

With NILC, the data d_ν at each frequency ν and sky pixel p are assumed to be the sum of the CMB component, whose **SED!** (**SED!**) a_ν is known but fluctuation amplitude s is not, and a global nuisance term n_ν which includes all other foreground emission components and instrument noise

$$d_\nu \;=\; a_\nu \, s \,+\, n_\nu \,. \tag{3}$$

In this equation and most subsequent equations the dependence on sky pixel p has been suppressed to simplify notation. The data d_ν represent the observed B-mode signal at frequency ν, that is $d_\nu \equiv B_\nu$. Operating on B-mode maps instead of using the Stokes Q_ν and U_ν maps optimizes foreground cleaning by having NILC minimize the variance of the foreground B-mode signal directly. The variance of the Q and U maps is dominated by E-modes.

The first step in the NILC algorithm is to perform a wavelet decomposition of the B-mode data on a needlet frame (Narcowich et al. 2006) as follows. The full-sky B-mode maps d_ν are bandpass-filtered in harmonic space using the seven window functions $(j) = (1), ..., (7)$ shown in Figure 2. This provides seven maps $d^{(j)}_\nu$, each exhibiting fluctuations of a specific range of angular scales (j), for each frequency band. For each range of scales (j) we compute an estimate $\hat{s}^{\,(j)}$ of the CMB B-mode anisotropy s in each pixel by forming a weighted linear combination of the frequency band data $d^{(j)}_\nu$

$$\hat{s}^{\,(j)} = \sum_\nu w^{(j)}_\nu \, d^{(j)}_\nu \,, \tag{4}$$

using the specific weights

$$w^{(j)}_\nu = \frac{\sum_{\nu'} \, {\rm C^{-1}}^{(j)}_{\nu\nu'} \, a_{\nu'}}{\sum_{\nu'} \sum_{\nu''} \, a_{\nu'} \, {\rm C^{-1}}^{(j)}_{\nu'\nu''} \, a_{\nu''}} \,. \tag{5}$$

${\rm C^{-1}}^{(j)}_{\nu\nu'}$ are the elements of matrix ${\rm C}^{-1\,(j)}$ which is the inverse of the empirical covariance matrix of the data ${\rm C}^{(j)}$ in each pixel p and needlet scale (j), whose elements are computed as

$${\rm C}^{(j)}_{\nu\nu',p} = \sum_{p'} W^{(j)}_{p,p'} \, d^{(j)}_{\nu,p'} \, d^{(j)}_{\nu',p'} \,. \tag{6}$$

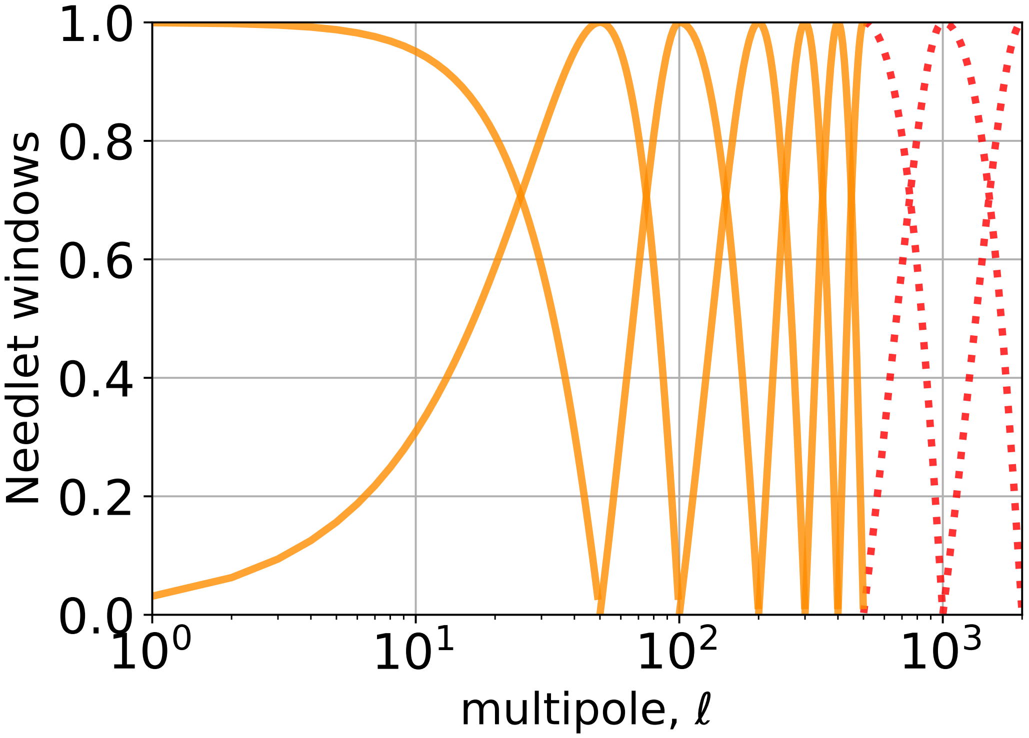

Figure 2. Cosine-shaped needlet bandpasses in harmonic space. Most of the analysis uses the seven bandpasses that peak at $\ell = 0$, 50, 100, 200, 300, 400, and 500 (solid orange lines). Two additional bandpasses (red dots), one peaking at ℓ =1000, and another peaking at $\ell = 500$ and $\ell = 2000$ are only used for the high-resolution maps of Section. 7.

The matrix coefficient $C^{(j)}_{\nu\nu',p}$ is the result of a convolution in pixel space and for this reason we show the explicit dependence on p. The Gaussian convolution kernel $W^{(j)}$ in Equation (6) defines the spatial domain over which the average of the product of maps $d^{(j)}_{\nu} d^{(j)}_{\nu'}$ for a given pair of frequencies is computed. The size of the convolution kernel is set by the number of independent modes within each spatial domain, which requires a FWHM of the Gaussian function $W^{(j)}$ of 4.0, 1.6, 0.85, 0.56, 0.46, 0.40, and 0.51 radians for needlet scale $(j) = (1), \dots, (7)$, respectively. These values guarantee sufficient spatial localization of the convolution kernel in pixel space while keeping the number of pixels large enough to ensure that the so-called 'ILC bias' (Delabrouille et al. 2009) is lower than a percent. We then combine the seven estimated CMB B-mode maps $\hat{s}^{(j)}$ from each needlet scale to form the final CMB B-mode full sky map $\hat{B}^{\rm NILC}$. The best estimate CMB E-mode map $\hat{E}^{\rm NILC}$ is produced the same way using E_ν as input. No attempt was made to optimize the needlet window functions as a function of the effectiveness of component separation; see Section 8.

For each foreground model and r value, we take ten realizations of CMB and noise. For each realization, r value, and foreground model, we construct the map $B^{\rm NILC}$, apply the same Galactic mask corresponding to a sky fraction $f_{\rm sky} = 0.46$, and calculate the B-mode angular power spectrum $\widehat{C}^{BB}_{\ell}$ using MASTER (Hivon et al. 2002). The mask was built from sky model Baseline by nulling the pixels in which the variance of the 40′-smoothed observed B-mode map at 555 GHz is the largest, until 50% of the sky is masked out, see Figure 3. Apodization of the mask border leaves an effective sky fraction of 46%. This mask choice has not been optimized, although we will see later that it appears to be conservative for foreground models Baseline-MHD. Power spectra error bars are derived analytically using Equation (A7). Angular power spectra of residual foreground and noise are obtained by applying the NILC weights Equation (5) to foreground-only and noise-only maps, then following the same process that generates $\hat{B}^{\rm NILC}$.

Using the B-mode power spectrum of $\hat{B}^{\rm NILC}$ and the corresponding noise power spectrum, we calculate a likelihood for r including information from multipoles $2 \leq \ell \leq 300$. We assume the Gaussian likelihood given in Equation (A6), which is a common approximation for the exact likelihood. We discuss the approximation in Appendix A. To account for PICO's capability to improve r constraints through delensing, we subtract 73% of the cosmic variance of the lensing

B-mode signal from the covariance matrix of the likelihood, that is, in Equations (A6) and (A7) we replace $\widehat{C}_\ell^{BB,\mathrm{NILC}}$ with $\widehat{C}_\ell^{BB,\mathrm{NILC}} - 0.73\, C_\ell^{\mathrm{lens}}$. We justify the level of 73% in Section 7.1. When confidence intervals are quoted, they are calculated by integrating the likelihood to encompass 68% when $r \neq 0$, or to 95% when $r = 0$.

5.2. *NILC results*

Figure 3 gives two B-mode map-domain examples for the component separation results. The left column is for the Baseline model, which is also prototypical of the results with 2MBB, PhysDust, and MHD, and the right column is for the MultiLayer model. At 90 GHz, and at all other PICO frequency bands, the input sky map (first row) is entirely dominated by foregrounds. The input CMB map, containing no inflationary signal and only lensing induced B-modes (second row), has a scale ten times lower than the full emission map. Nevertheless, the component separation produces an output CMB map (third row) that is visually nearly indistinguishable from the input, including along the Galactic plane. With model MultiLayer, the Galactic plane is clearly visible and there is a larger variance across sky. The residual maps showing the difference of output and input (fourth row) more clearly reveal the same conclusions.

Figure 4 gives the NILC component separation results. The top row shows results for each of the ten realizations with the Baseline model with the two r_{in} values. The recovered CMB signal matches the input, which is dominated by lensing for $\ell \gtrsim 10$; the noise post-component separation is well below the signal; and the residual foregrounds are well below the noise. The middle panels show the average power spectra over the ten realizations and demonstrate that these conclusions hold for all models except MultiLayer. The inflection point in the residual noise spectra at $\ell \sim 800$, most prominently visible for the MultiLayer model (purple triangles), is a typical consequence of NILC weighting (Planck Collaboration IV 2020). The outcome of the algorithm is to minimize the foreground variance at low ℓ at the expense of higher noise. At high ℓ there is minimization of noise variance. For all models except MultiLayer the residual foreground spectra are equivalent to levels below $r = 5 \times 10^{-4}$ for $2 \leq \ell \leq 150$. The lower row gives likelihoods for r calculated using the average spectra , and the posteriors are rescaled to make it easier to compare them with one another. With the exception of the MultiLayer model, both $r = 0.003$ and an upper limit on $r = 0$ are obtained without bias. With MultiLayer the residual foreground is larger than with other sky models and it reaches a level of $r = 0.005$ near $\ell = 10$. The large residual leads to biased posterior distributions for both $r = 0$ and $r = 0.003$. We discuss the MultiLayer model in more detail in Section 5.2.1.

Table 3 summarizes the inferred values of r for input values $r_{\mathrm{in}} = 0.003$ and $r_{\mathrm{in}} = 0$. For all sky models except MultiLayer, $r = 0.003$ is detected with insignificant bias and more than 5σ significance without accounting for delensing, and with 18σ significance or more with 73% delensing. Similarly, with $r_{\mathrm{in}} = 0$ the inferred values are consistent with $r = 0$, and after delensing the 95% confidence upper bounds are between 1.3 and 2.7×10^{-4}. With model MultiLayer there is a 3σ bias with both r values after 73% delensing. Residual foregrounds after component separation lead to biased estimates.

Table 3. r forecasts using blind component separation with $r_{\mathrm{in}} = 0.003$ and $r_{\mathrm{in}} = 0$ and using 21 frequency bands. $r_{95\%}$ denotes the 95% upper limit, while $r/\sigma(r)$ gives the value of r at the peak of the likelihood relative to the 1σ interval, calculated by integrating the likelihood from its peak to 68% probability.

	$r_{\mathrm{in}} = 0.003$		$r_{\mathrm{in}} = 0$			
	$[r \pm \sigma(r)]\,/10^{-3}$		$r_{95\%}/10^{-4}$	$r/\sigma(r)$	$r_{95\%}/10^{-4}$	$r/\sigma(r)$
Model	No delensing	73% delensing	No delensing		73% delensing	
Baseline	3.12 ± 0.54	3.15 ± 0.16	3.7	< 1	2.6	1.2
2MBB	3.07 ± 0.50	3.09 ± 0.13	2.5	< 1	1.5	< 1
PhysDust	3.07 ± 0.50	3.09 ± 0.11	2.3	< 1	1.3	1.2
MHD	3.09 ± 0.53	3.09 ± 0.17	4.1	< 1	2.7	1.2
MultiLayer	3.90 ± 0.79	3.93 ± 0.32	14.8	1.6	13.2	2.8

The results presented in Table 3 and in Figure 4 use all the frequency bands for the component separation. We have analyzed the need for the wide frequency range by carrying out the component separation and removing some edge bands. A PICO-HF configuration assumes bands only between 43 and 799 GHz. In this configuration the four lowest PICO bands have been removed. A PICO-LF configuration assumes bands only between 21 and 462 GHz. In this

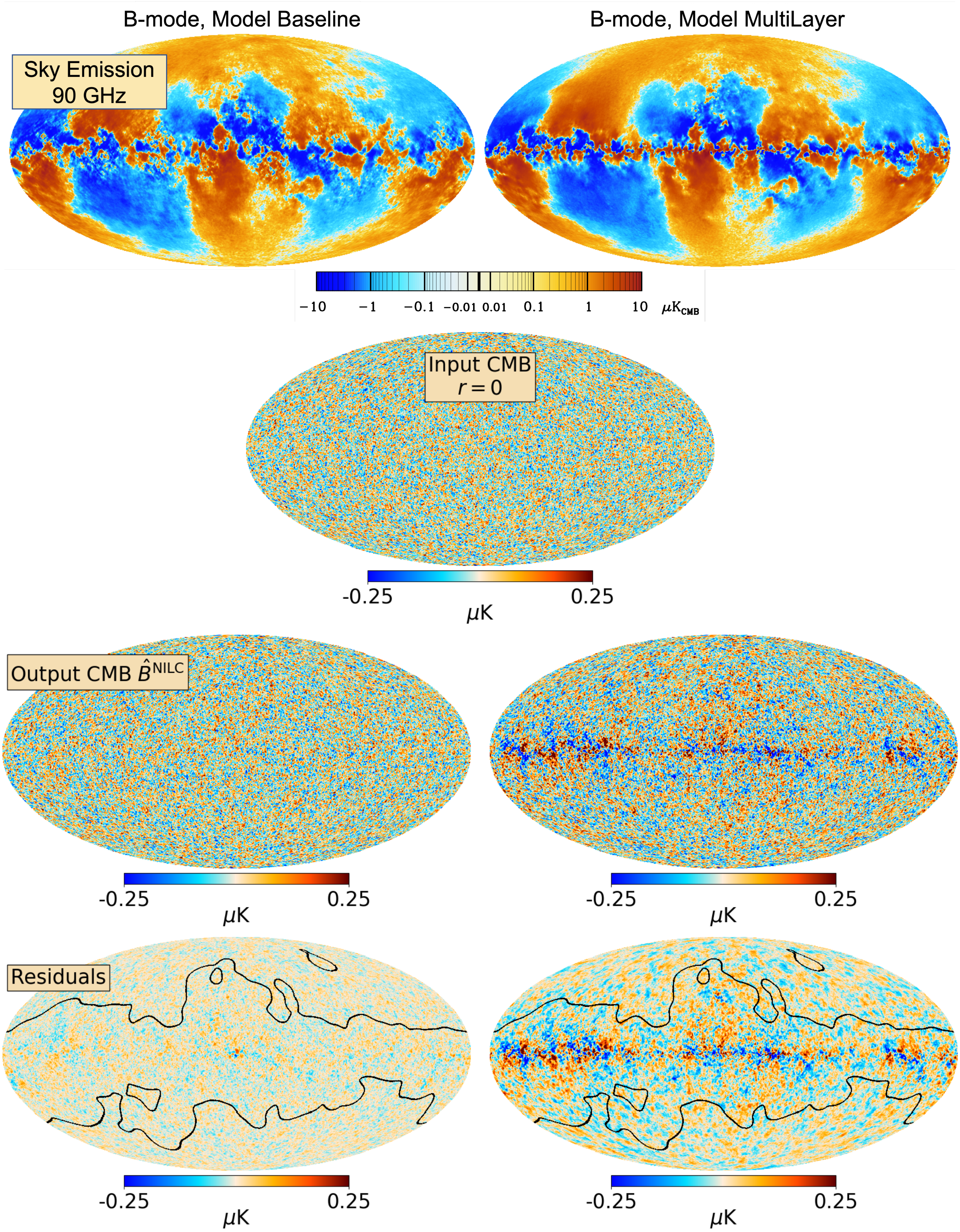

Figure 3. B-mode maps smoothed to 40' fwhm for models Baseline (left) and MultiLayer (right) before and after component separation with NILC. From top to bottom: sky map at 90 GHz; the input CMB with $r = 0$; the output CMB after component separation; and the CMB residual map = output CMB - input CMB, and outline of the 50% mask that determines the portion the sky for power spectra calculations.

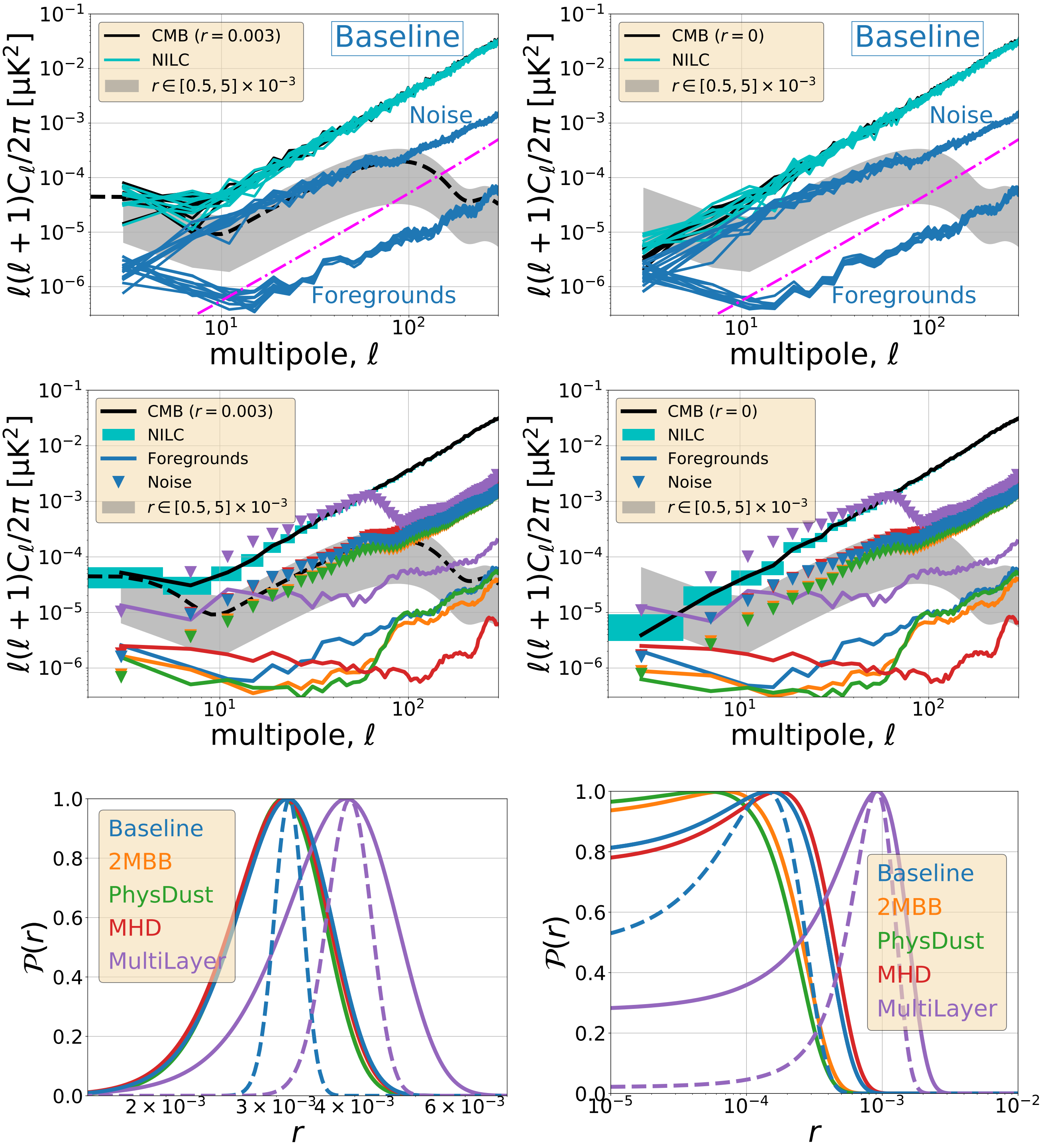

Figure 4. Power spectra and r likelihood after NILC component separation with $r_{\rm in} = 0.003$ (left) and $r_{\rm in} = 0$ (right). Top: Recovered B-mode power spectra of CMB (cyan), residual foregrounds (blue) and noise (blue) for sky model Baseline and ten realizations of CMB and noise. For reference, the black solid lines show the power spectra of the input CMB realisations while the gray shaded area shows the theoretical power spectrum of the primordial CMB B-mode for r ranging from 5×10^{-4} to 5×10^{-3}. Middle: Power spectra averaged over all 10 realisations of residual foregrounds (solid coloured lines) and noise (coloured triangles) for all sky models: Baseline (blue), 2MBB (orange), PhysDust (green), MHD (red) and MultiLayer (purple). The cyan boxes show the average recovered CMB B-mode power spectrum and standard deviations for model Baseline. Bottom: Recovered posterior distributions of r without delensing (solid) and with $A_{\rm lens}$= 27% for models Baseline and MultiLayer (dash). To improve readability, the posterior distributions are rescaled by their maximum value to have a maximum value of 1.

configuration the three highest PICO bands have been removed; see Table 1. The PICO-HF and -LF configurations are intended to give a coarse assessment about the relative importance of accounting for synchrotron vs. Galactic dust and to inform the level of constraints attainable with a narrower frequency range. A frequency range narrower than PICO's baseline are planned for next generation ground-based instruments and have been proposed for a future space mission (Abazajian et al. 2016; Delabrouille et al. 2018; Hazumi et al. 2020). Results from this analysis, which was only carried out for $r_{\rm in} = 0$, are given in Figure 5 and Table 4.

For either band-restricted configuration and for all sky models the level of residual foregrounds increases (Figure 5 left column), degrading constraints on r and in some cases producing statistically significant biases. For example, with models Baseline and MHD, the value of r at the peak of the likelihood is approximately 3σ away from $r = 0$. For sky models 2MBB and PhysDust, which include synchrotron curvature, the constraints on r degrade approximately equally with PICO-HF and -LF indicating an approximate equal role for separating Galactic dust and synchrotron. In contrast, with model Baseline, which has no synchrotron curvature, or with models MHD and MultiLayer, in which Galactic dust has more complex emission properties, PICO-HF gives slightly less biased r constraints compared to PICO-LF, indicating that separating Galactic dust is more important than separating the relatively simpler synchrotron emission model. Overall, a 21-band PICO instrument would reduce the 95% upper limit on $r = 0$ by 20-50% relative to PICO-HF and by 40-54% relative to PICO-LF, see Table 4.

Table 4. NILC r forecasts with $r_{\rm in} = 0$ and without either the low frequencies (LF) or the high frequencies (HF), and assuming 73% delensing.

$r_{\rm in} = 0$, 73% delensing				
	PICO-HF (43-799 GHz)		PICO-LF (21-462 GHz)	
Model	$r_{95\%}/10^{-4}$	$r/\sigma(r)$	$r_{95\%}/10^{-4}$	$r/\sigma(r)$
Baseline	4.3	1.5	5.6	1.6
2MBB	2.8	1.4	2.5	1.3
PhysDust	2.6	1.7	2.7	1.3
MHD	3.8	1.3	4.8	1.6
MultiLayer	16.7	3.4	22.8	4.4

5.2.1. *Model MultiLayer and Local Estimates of r*

With model MultiLayer, the NILC component separation process gives biased estimates of r. In Section 8 we discuss potential improvements for the component separation process that might mitigate the biases.

One way to identify the existence of biased estimates during the analysis is to compare independent constraints on r from independent sections of the sky. To carry out this analysis, we partition the sky to 26 sections of equal area outside the Galactic plane, each with a 2.5% area of the sky, see Figure 6. Masks for individual regions are directly applied to the $\hat{B}^{\rm NILC}$ map, with no apodization, to compute local B-mode spectra and infer local constraints on r. No apodization is necessary because there is no E-to-B leakage and no significant contamination of small scale power from large angular scale modes. With models Baseline and MultiLayer we plot the constraints as a function of the rms dust polarization intensity $\sigma_{P_{555}}$ as determined by the 555 GHz data in each section prior to component separation

$$\sigma_{P_{555}} = \sigma\left(P_{555} = \sqrt{Q_{555}^2 + U_{555}^2}\right) , \tag{7}$$

where Q_{555} and U_{555} are the pixel values in each section. There is no difference in the component separation process itself; it is conducted on the full sky. The results are shown in Figure 6 for $r_{\rm in} = 0$ (left panel) and $r_{\rm in} = 0.003$ (right panel). In both cases we show r estimates after 73% delensing. For both input r values we identify a common trend that is best observed by following $\sigma_{P_{555}}$ from higher to lower values. As $\sigma_{P_{555}}$ decreases the values of the estimated r values on the individual sky sections are decreasing. For the lowest $\sigma_{P_{555}}$ values, those corresponding to the cleanest sections of the sky, r converges to a stable value; that is, the same value or upper limit is obtained within statistical errors. The trend indicates that, as expected, sky areas with high polarized intensity from dust correlate with biased r values. As the dust polarization intensity decreases, the bias decreases and at a certain level becomes insignificant.

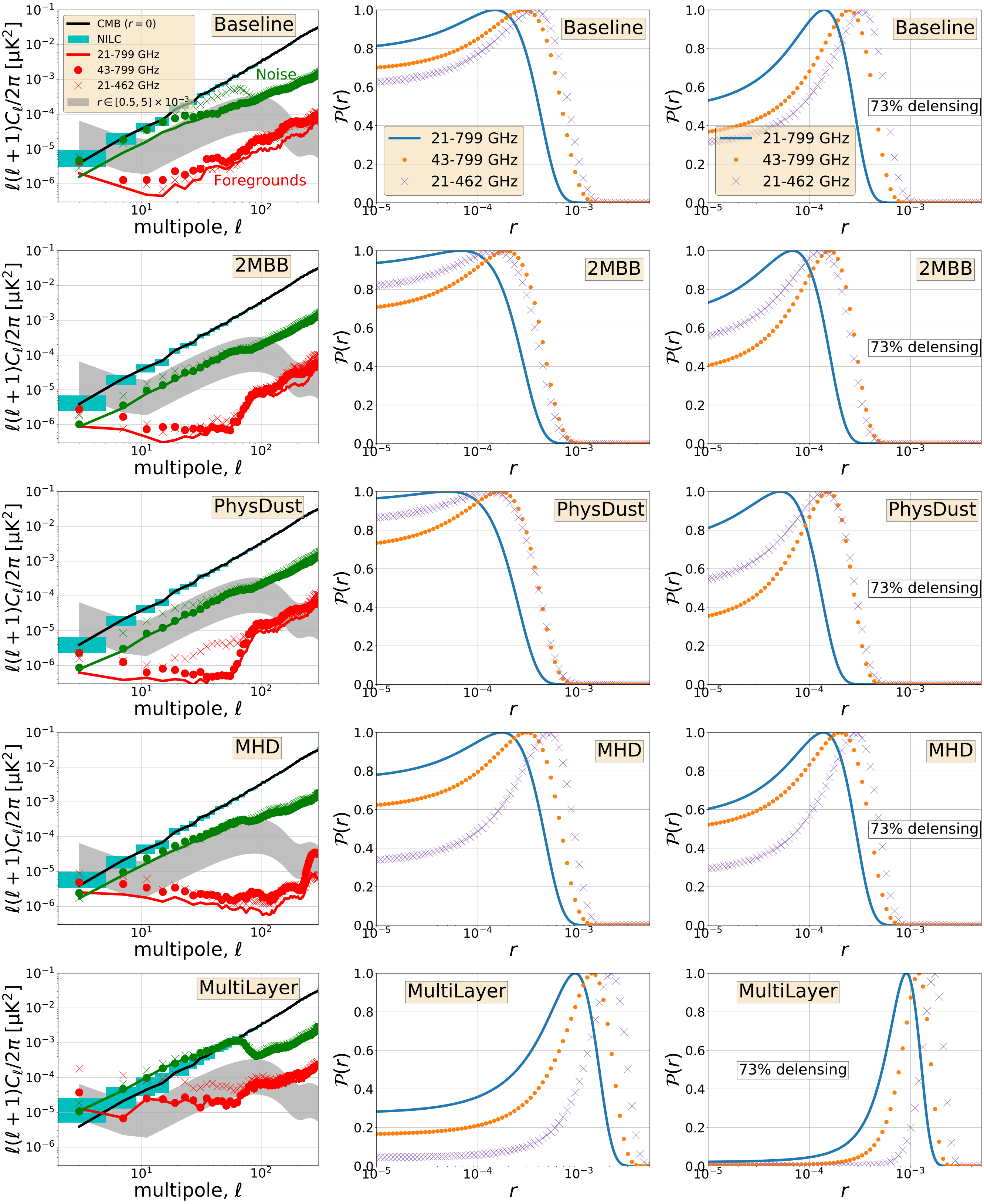

Figure 5. Power spectra (left column) and r likelihoods without and with delensing (middle and right columns) when $r_{\mathrm{input}} = 0$ comparing three configurations: all frequency bands (21 GHz - 799 GHz, solid); without the low frequency bands (43 GHz - 799 GHz, dots); and without the high frequency bands (21 GHz - 462 GHz, crosses). Legends for the panels are given in the first row. To improve readability, the posterior distributions are rescaled by their maximum value to have a maximum value of 1.

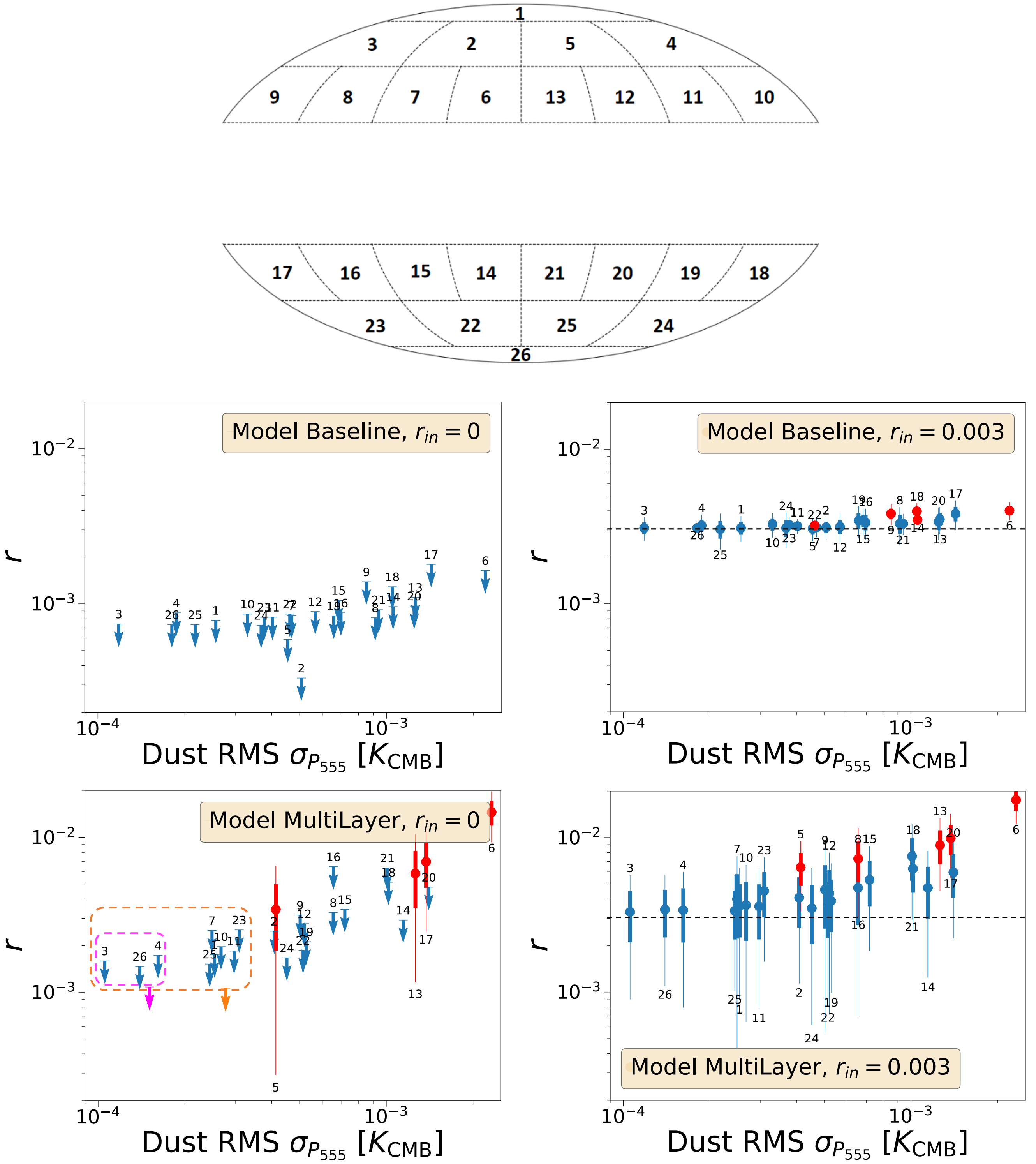

Figure 6. The top panel shows the 26 equal area sky sections with $f_{\rm sky}$= 2.5% each. The middle and bottom rows show the section-by-section results from NILC for models Baseline and MultiLayer, respectively. The left and right panels give 95% confidence limits (blue bars with arrows) with $r_{\rm in} = 0$, and r likelihood confidence intervals (thick blue is 68%, thin blue is 95%) with $r_{\rm in} = 0.003$ (black dashed line), respectively, both as a function of P_{555}. Some patches produce biased detections at a confidence level larger than 95% (red points), both when $r = 0$ (lower left, for Multilayer) and when $r = 0.003$ (right column, Baseline and Multilayer). For biased detections we give the 68% (thick red) and 95% (thin red) confidence intervals. With model MultiLayer and $r_{\rm in} = 0$ we also give 95% upper limits after aggregating the sections with the three lowest values (lower left, magenta arrow) and with the nine lowest values (lower left, orange arrow) of dust polarization. The vertical scale in all panels is identical to give a visual comparison of the results.

When $r_{\text{in}} = 0$, most patches give upper limits at 95% confidence (bars with arrows) even when dust contamination is more pronounced. However, with model MultiLayer, which has higher dust residual after component separation than Baseline, four patches give biased $r \neq 0$ detections at levels exceeding 2σ (red thin error bars). When $r_{\text{in}} = 0.003$ the higher dust residual of model MultiLayer gives higher r levels at high $\sigma_{P_{555}}$ and larger error bars compared to Baseline. The final result is that in both models there are five patches that give biased detections of $r > 0.003$ by more than 2σ significance, and in both models many lower dust level patches are consistent with $r_{\text{in}} = 0.003$. Most of the patches that give erroneous detections are near the Galactic plane, although not all, and all are at levels of 4 times the lowest $\sigma_{P_{555}}$ or above.

Aggregating several areas of the sky that have low dust polarized intensity, and for which r upper limits have converged, decreases sample variance and decreases the upper limit. By aggregating the three (magenta) and nine lowest P_{555} (orange) sky areas we find with model MultiLayer 95% upper limits of 1.9×10^{-3} and 1.6×10^{-3}, respectively.

6. COMPONENT SEPARATION - COMMANDER

6.1. *Methodology*

The Commander data model d_ν at each frequency ν and sky pixel is given by the sum of the signal components and a noise term n_ν,

$$\begin{aligned} d_\nu &= a_{\text{CMB}} \\ &+ a_{\text{s}}\,\gamma(\nu)\left(\frac{\nu}{\nu_{0,\text{s}}}\right)^{\beta_s} \\ &+ a_{\text{d}}\,\gamma(\nu)\frac{e^{\frac{h\nu_{0,d}}{kT_{\text{d}}}}-1}{e^{\frac{h\nu}{kT_{\text{d}}}}-1}\left(\frac{\nu}{\nu_{0,d}}\right)^{\beta_{\text{d}}+1} \\ &+ n_\nu . \end{aligned} \tag{8}$$

At each pixel the free parameters are the CMB, synchrotron, and thermal dust amplitudes, $\{a_{\text{CMB}}, a_{\text{s}},\ a_{\text{d}}\}$; the synchrotron and thermal dust spectral indices, β_s and β_{d}; and the dust temperature, T_{d}. In addition, k and h are Boltzmann's and Planck's constants, respectively, $\gamma(\nu) = (e^x-1)^2/(x^2e^x)$ with $x = h\nu/kT_{\text{CMB}}$, is the unit conversion factor between Rayleigh-Jeans brightness temperature and the CMB thermodynamic temperature, and $\nu_{0,\text{s}}$ and $\nu_{0,\text{d}}$ are reference frequencies for the synchrotron and thermal dust components.

The aim is to fit all the modelled parameters $\omega = \{a_i, \beta_i, T_d\}$ to the input data, which in a Bayesian framework means computing the posterior distribution $P(\omega|d)$, where $d = \{d_\nu\}$. With Bayes' rule this is given as the product of the likelihood $P(d|\omega)$ and a set of priors $P(\omega)$. The amplitudes $a = \{a_{\text{CMB}}, a_{\text{d}}, a_{\text{s}}\}$ and spectral indices $\theta = \{\beta_{\text{s}}, \beta_{\text{d}}, T_{\text{d}}\}$ are fitted with Gibbs sampling iteratively in steps described by

$$\begin{aligned} a^{i+1} &\leftarrow P(a|\theta^i, d) \\ \theta^{i+1} &\leftarrow P(\theta|a^{i+1}, d) \end{aligned} \tag{9}$$

where the parameters to the left of $\leftarrow$ are sampled from the distribution to the right, keeping the spectral indices fixed while sampling the amplitudes and vice versa. As discussed by Eriksen et al. (2008), the amplitude distribution $P(a|C_l^i, \theta^i, d)$ is a simple multi-variate Gaussian in the various component amplitudes, while the spectral index distribution $P(\theta|C_l^i, a^{i+1}, d)$ can be sampled efficiently by mapping out the corresponding one-dimensional distribution by brute force per pixel and parameter. We adopt a standard Blackwell-Rao estimator for likelihood and cosmological parameter estimation (Chu et al. 2005).

Commander is first run for the full sky, fitting component amplitudes and spectral parameters at each pixel. We define a per-pixel χ^2 statistic

$$\chi^2 = \sum_\nu \left(\frac{d_\nu - s_\nu}{\sigma}\right)^2, \tag{10}$$

where s_ν is the total signal model and σ are the instrumental noise rms per pixel. This χ^2 map is smoothed to 8° FWHM and conservatively thresholded at $4\,\sigma$ to remove pixels with a significant model errors. The resulting confidence mask, outside of which the CMB is presumed reconstructed at higher confidence, removes 25 % of the low-latitude and high-foreground sky; it is shown in gray in Figure 7.

We use a Blackwell-Rao estimator (Chu et al. 2005) to estimate r. The estimator requires a series of full-sky CMB map samples, rather than the partial sky samples as are available at this stage. We therefore run another Gibbs chain for just the CMB map of the form,

$$\begin{aligned} a_{\mathrm{CMB}}^{i+1} &\leftarrow P(a_{\mathrm{CMB}}|d, a_{\mathrm{s}}, a_{\mathrm{d}}, \theta, C_l^i) \\ C_l^{i+1} &\leftarrow P(C_l^i|a_{\mathrm{CMB}}^{i+1}) \end{aligned} \tag{11}$$

in which the presence of the power spectrum C_l in the first distribution ensures that the masked pixels in the CMB sky map sample are effectively "in-painted" with a constrained Gaussian realization (Eriksen et al. 2004). The second distribution $P(C_l|a_{\mathrm{CMB}}^{i+1})$ is an inverse Wishart distribution, which has a known, simple sampling algorithm (Larson et al. 2007). The explicit expression for the Blackwell-Rao estimator given these full-sky CMB samples is given in Appendix A. No correction for delensing is done with the Commander analysis, and to the extent that comparison is done with NILC, the comparison should be with the no-delensing results.

In principle, the two Gibbs chains described by Equations (9) and (11) can be merged into one chain with three steps. However, splitting it into two independent steps has several advantages (Colombo et al. 2022; Paradiso et al. 2022). The most important concerns the amplitude sampling step in Equation (9). Solving for both foreground amplitudes and a power spectrum-constrained CMB component leads to a computationally costly non-local linear system with a high condition number, and increases the total runtime of the full algorithm by orders of magnitude (Seljebotn et al. 2017). A second advantage is that the χ^2-based CMB confidence mask can be defined after the component separation step, but before estimating the CMB power spectrum. The main disadvantage of a split chain is a slightly higher white noise level, as the foreground amplitudes are allowed to explore a larger posterior volume in the first step, when the CMB component is unconstrained by the power spectrum (Colombo et al. 2022). A total of ten simulations with independent CMB and noise realizations are processed with this algorithm for both $r = 0$ and $r = 0.003$.

6.2. *Commander Results*

The upper panels in Figure 7 show the Stokes Q and U for one single reconstructed CMB realization with model Baseline, smoothed to an effective resolution of 5° FWHM. Commander was only used with model Baseline. The gray region is the confidence mask. The bottom panel, which shows the difference between the reconstructed and the true input CMB map with a color scale that is hundred times smaller than the signal, demonstrates the high signal-to-noise ratio maps that PICO will generate. At a visual level the residual map appears statistically consistent with white noise, and there are no discernible foreground residuals around the mask edge.

The likelihood results from Commander are summarized in Figure 8. The left and right columns shows r constraints for simulations with input values of $r = 0$ and $r = 0.003$, respectively. The top row gives individual marginal posterior distributions for ten realizations of each of the r inputs. In both cases, the recovered distributions are consistent with the true inputs; for $r = 0$, three out of ten realizations peak at $r = 0$, while the remaining seven realizations peak at a slightly positive value, but none has a likelihood ratio greater than $\mathcal{L}(r_{\mathrm{max}})/\mathcal{L}(r = 0) \approx 3$, as expected for random variations. The probability of observing seven realizations with a positive value by random chance is 17 % from pure binomial statistics. Similar conclusions are obtained for the $r = 0.003$ case, with four (six) realizations having a lower (higher) peak value than the true input.

The bottom left panel of Figure 8 shows the cumulative distribution functions for each of the ten realizations. The black vertical dashed line shows the median upper limit of $r < 8.9 \times 10^{-4}$ (95 % confidence limit). This limit accounts for the non-Gaussian shape of the posterior distributions, and it is a factor of 2.4 higher than the corresponding NILC limit, which uses a Gaussian likelihood. The NILC estimate, however, includes a factor of ~ 20 broader multipole range, $\ell < 200$, and is therefore expected to give tighter constraints. The bottom right panel shows the posterior standard deviation σ_r when $r_{\mathrm{in}} = 0.003$ for each of the ten realizations, and the dashed black line shows their mean of $\sigma_r = 1.1 \times 10^{-3}$. For comparison, the dashed red line shows the corresponding NILC estimate for model Baseline without delensing $\sigma = 0.54 \times 10^{-3}$.

7. LENSING AND DELENSING

PICO constraints on r depend on the effectiveness of delensing. In Section 5 we assume A_{lens}= 0.27, amounting to a delensing factor of 73%. This estimate is based on an angular power spectrum domain calculation rather than on an actual map domain delensing exercise. In Section 7.1 we explain how A_{lens} was estimated. For one foreground model we have also conducted map domain delensing. We describe it in Section 7.2 and show that the map domain delensing level is in close agreement with the power spectrum-based estimate.

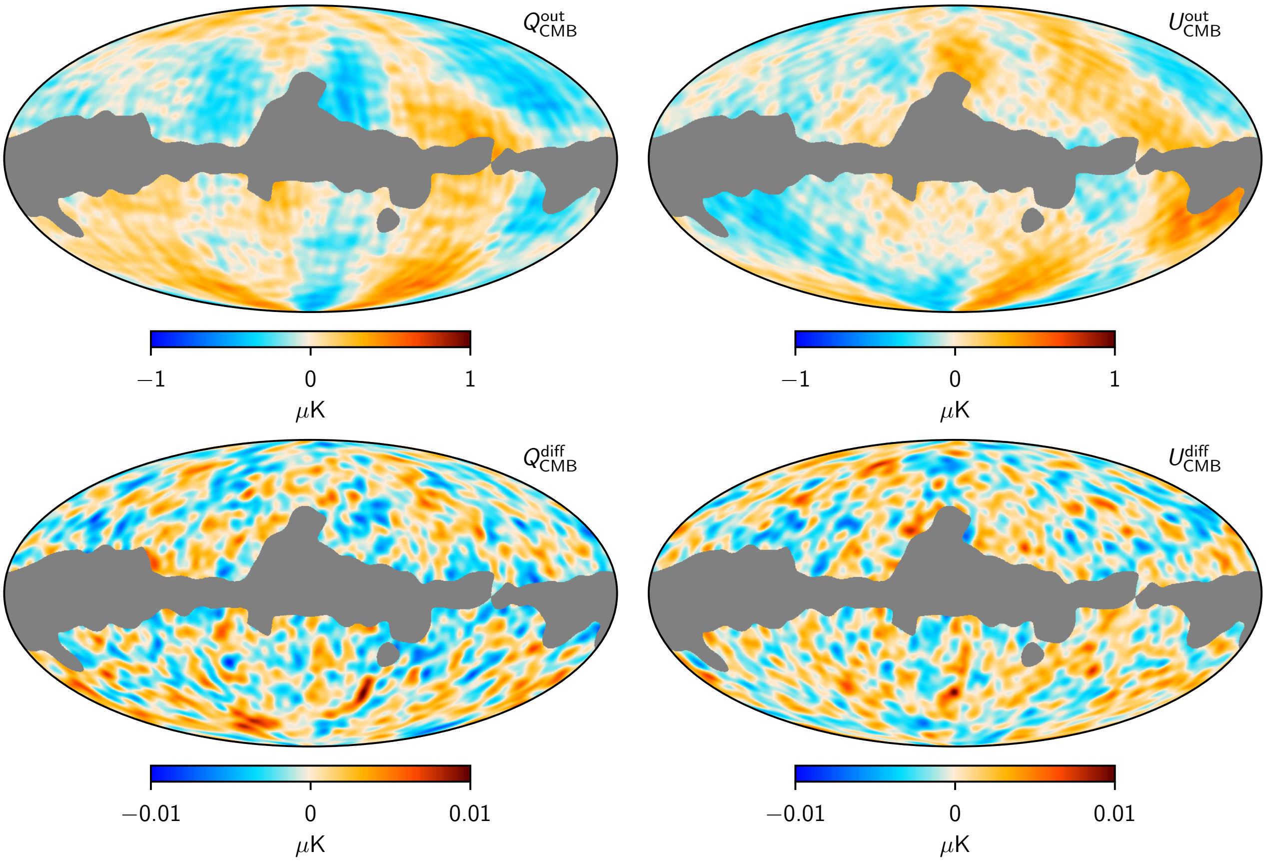

Figure 7. Top row: Stokes Q (left) and U (right) maps reconstructed with Commander for one realization with model Baseline. Bottom row: Difference between the reconstructed map and the input CMB map. The gray region is the confidence mask derived from the χ^2 cut. All maps are smoothed to a common angular resolution of 5° FWHM. Note that the color scale in the bottom row is 100 times smaller than the top row.

7.1. *Power Spectrum Domain Estimate of Delensing*

Given PICO's noise level and beam size we forecast the beam-deconvolved noise power spectrum of the CMB polarization after component separation. For this, we use a harmonic-domain internal linear combination (ILC) approach identical to the one described by Ade et al. (2019), apart from minor changes described in the following. The sky model includes analytic models for the polarized power spectra of dust, synchrotron, the dust-synchrotron cross-correlation, radio sources, and CMB (Ade et al. 2019). The amplitudes of the dust and synchrotron power spectra at $\ell = 80$ are chosen to approximately match those measured on large sky fractions by Planck (Planck Collaboration XI 2020) at 353 and 30 GHz, respectively. The assumed correlation coefficient of the dust and synchrotron fields is 40%. The computed E- and B-mode power spectra at all PICO frequencies are then passed through the harmonic ILC code to obtain the post-component-separation noise on the CMB component. On the largest angular scales $\ell < 30$, we use inverse-covariance-weighted noise curves rather than the post-ILC noise, due to potential ILC biases at these low multipoles (Delabrouille et al. 2009); this has negligible impact on the derived lensing noise curves, and thus on the delensing performance. We implement standard minimum-variance ILC as well as constrained ILC methods (Remazeilles et al. 2011b) that deproject components with a fiducial polarized dust SED, assumed to be a modified blackbody with $\beta_{\mathrm{d}} = 1.59$, a fiducial synchrotron SED that is a power-law in antenna temperature with $\beta_{\mathrm{s}} = -3.1$, or both. This deprojection leads to larger post-ILC noise, but would significantly reduce biases due to foregrounds, and thus conservatively brackets the expected post-ILC sensitivity.

Given the post-ILC polarization noise power spectra, we use the forecasting approach of Smith et al. (2012) to obtain the expected delensed level of the lensing B-mode power spectrum. This method involves iteratively computing the

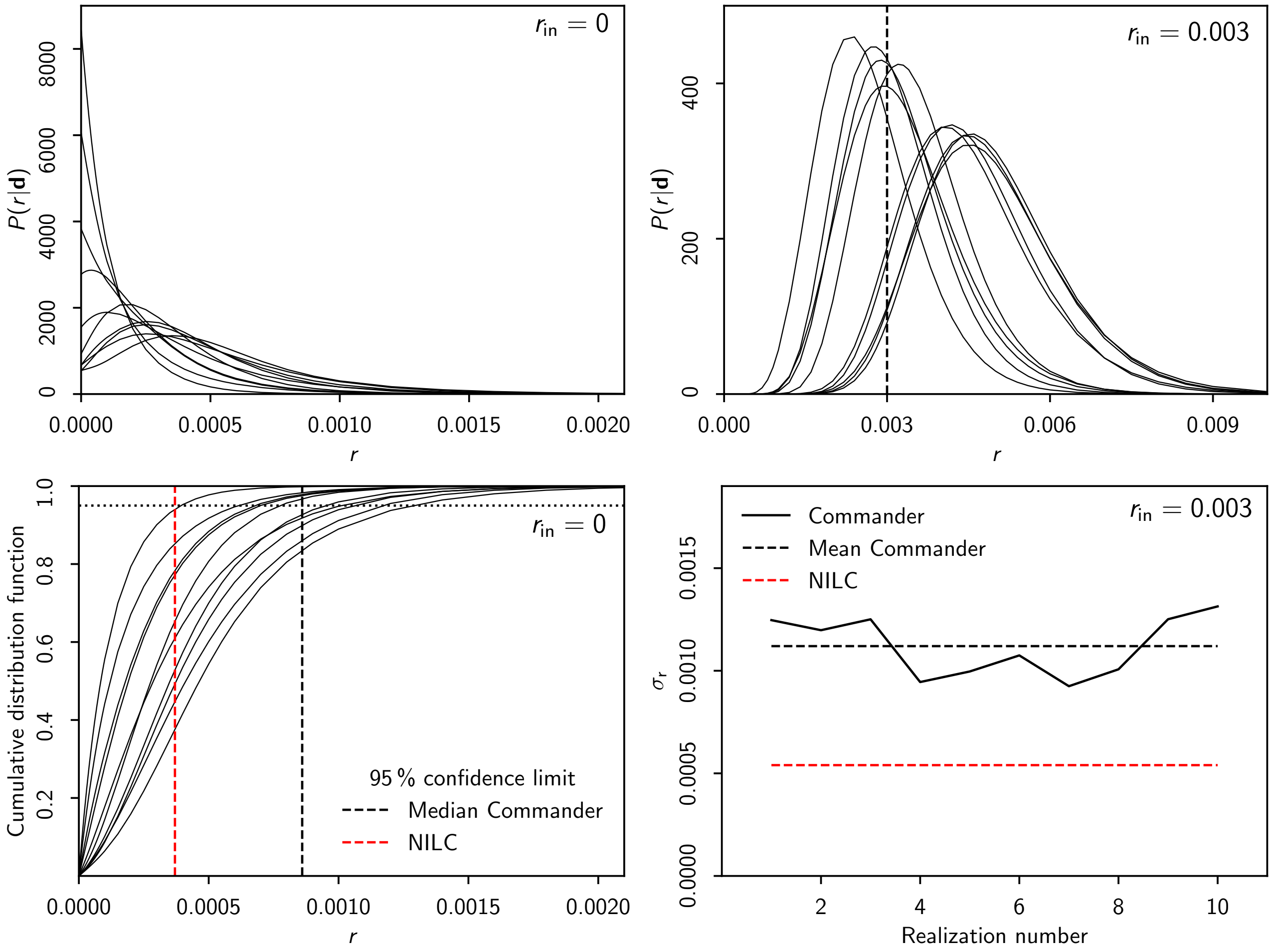

Figure 8. Top row: Commander posterior distributions for $r = 0$ (left) and $r = 0.003$ (right) with ten realizations. Bottom left: Cumulative probability distributions for ten realizations with $r = 0$. The vertical dashed black line shows the median of the 95 % confidence limits derived with Commander, while the vertical red line shows the corresponding 95 % confidence limit computed with NILC and a Gaussian likelihood. Bottom right: Posterior standard deviation σ_r for each of the ten Commander realizations with $r = 0.003$ (solid black). For comparison, the dashed black line shows the corresponding mean, while the dashed red line shows the corresponding estimate derived with NILC coupled to a Gaussian likelihood (red dash). Note that NILC results use all simulations to generate one combined estimate while the Commander results are based on an independent estimate per realization.

quadratic estimator-based lensing noise and then the B-mode power after delensing with the associated lensing map. With this entirely angular power spectrum-based forecasting approach, we necessarily neglect the spatial variation of the foregrounds, both for the initial harmonic-ILC step and for the subsequent iterated steps of forecasted lensing reconstruction and delensing.

Table 5 shows the delensing factor $1-A_{\rm lens}$, defined in Equation (A5), as a function of deprojection choices for two values of map noise. The lower noise is the value we use in this paper; see Table 1. The higher noise is the level PICO is required to achieve (Hanany et al. 2019). We emphasize that the harmonic ILC procedure uses each band's noise level and beam size, and that the combined map noise levels are only a short-hand to indicate either configuration. The NILC component separation of Section 5 conservatively uses $1-A_{\rm lens}= 0.73$, obtained when deprojecting both synchrotron and dust components with a noise level that is $\sqrt{2}$ higher than assumed in this paper. With the level of noise we use here we expect $1-A_{\rm lens}= 0.78$. In the next section, we describe a map-based delensing exercise that give $A_{\rm lens}$ values closely reproducing the results of the power spectrum-based forecasting approach.

Table 5. Forecast delensing factor $1-A_{\rm lens}$ with two map noise levels for different ILC analysis assumptions (Ade et al. 2019; Remazeilles et al. 2011b). The delensing factor is defined in Equation (A5).

ILC Assumptions	Map Noise Level (μK arcmin) 0.87	0.61
No foregrounds	0.80	0.85
No deprojection; standard ILC	0.80	0.84
Polarized dust deprojected	0.80	0.84
Polarized synchrotron deprojected	0.73	0.78
Polarized dust & synchrotron deprojected	0.73	0.78

7.2. *Map Domain Delensing*

To perform map-domain delensing we made maps of model Baseline with $r = 0$ at $N_{\rm side}$= 2048 (see Section 4). We conducted NILC component separation as described in Section 5, this time with an output resolution of $8'$, and the high-resolution output E- and B-mode CMB maps $\hat{E}^{\rm NILC}$ and $\hat{B}^{\rm NILC}$ were used to reconstruct the lensing field $\hat{\phi}$ and to produce a best estimate delensed E-mode map. This map and the field $\hat{\phi}$ are then used to estimate two additional maps, the delensed, primordial CMB B-mode map and a B-mode lensing map. Ten simulations have been conducted beginning with generation of $\hat{E}^{\rm NILC}$ and $\hat{B}^{\rm NILC}$ high resolution maps and culminating with these four products, one lensing field, delensed E- and B-mode maps, and a lensing map.

The lensing field $\hat{\phi}$ and delensed E-mode maps were reconstructed in two ways: (1) using a standard quadratic estimator, the output of which we denote $\hat{\phi}^{\rm QE}$ and $\hat{E}^{\rm NILC,QE}$ (Hu & Okamoto 2002; Okamoto & Hu 2003; Maniyar et al. 2021); and (2) using a maximum a-posteriori reconstruction, the output of which is denoted $\hat{\phi}^{\rm MAP}$ and $\hat{E}^{\rm NILC,MAP}$ (Seljak & Hirata 2004). For the quadratic estimator we use `Plancklens`[1] and the same methodology implemented with Planck (Planck Collaboration XV 2016; Planck Collaboration VIII 2020), and the maximum a-posteriori reconstruction is obtained from a curved-sky implementation of the iterative delensing solver described by Carron & Lewis (2017). In the next paragraph we discuss only key elements of this solver and refer the reader to the cited publication for more details.

7.2.1. *Reconstructing the lensing field $\hat{\phi}^{\rm MAP}$ and the delensed E-mode map $\hat{E}^{\rm NILC,MAP}$*

The posterior p for the lensing potential ϕ is

$$\log\left[p(\phi|\hat{E}^{\rm NILC},\hat{B}^{\rm NILC})\right] = \log\left[p(\hat{E}^{\rm NILC},\hat{B}^{\rm NILC}|\phi)\right] - \frac{1}{2}\sum_{LM}\frac{|\phi_{LM}|^2}{C_L^{\phi\phi,{\rm fid}}}\,, \tag{12}$$

where ϕ is assumed to be a Gaussian field, the fiducial lensing field power spectrum $C_L^{\phi\phi,{\rm fid}}$ comes from the Planck FFP10 simulations (Planck Collaboration et al. 2020b), and ϕ_{LM} is the harmonic transform of ϕ. The prior is necessary to handle the large number of poorly-constrained, high-L modes, and it is desirable so as to optimally weigh each lensing multipole when building the lensing map (Sherwin & Schmittfull 2015). We assume that the lensing likelihood $p(\hat{E}^{\rm NILC},\hat{B}^{\rm NILC}|\phi)$ is Gaussian in the maps. It contains a noise covariance matrix on which we apply a Gaussian isotropic beam with 8 arcmin FWHM. The noise is modeled empirically by using smoothed power spectra of $\hat{E}^{\rm NILC}$ and $\hat{B}^{\rm NILC}$ noise maps.

All $\hat{E}^{\rm NILC}$ multipoles $2 \leq \ell \leq 2000$ and $\hat{B}^{\rm NILC}$ multipoles $200 \leq \ell \leq 2000$ are used for the reconstruction of $\hat{\phi}^{\rm MAP}$and the delensed E-mode maps. The $\hat{B}^{\rm NILC}$ multipole range $\ell \leq 200$ is excluded to minimize statistical dependence with the degree-scale information used for r-inference (Teng et al. 2011; Baleato Lizancos et al. 2021a). This exclusion produces a 2% increase in the final residual lensing amplitude, that is, a 2% increase in $A_{\rm lens}$ (Legrand & Carron 2022). We assume that the unlensed CMB and the lensing map are limited to $\ell \leq 2500$ and $L \leq 2000$, respectively, and the reconstruction is done on the full-sky, since the $\hat{E}^{\rm NILC}$ and $\hat{B}^{\rm NILC}$ maps show no obvious spatial features which would have to be masked, see Figure 9.

---

[1] https://github.com/carronj/plancklens

The solver is iterative. It calculates the gradient of the log-posterior with respect to ϕ and combines this information with an estimate of the local posterior curvature, which is built from previous solutions, to produce the next ϕ estimate (Carron & Lewis 2017). Each calculation of the maximum a-posteriori reconstruction is equivalent to obtaining the most probable, or Wiener-filtered (WF), unlensed CMB conditioned on the current estimate of the lensing map and the fiducial cosmology being the truth (Carron & Lewis 2017).

The likelihood model assumes $r = 0$, and we ignore temperature anisotropy information. The iteration of the solver starts from the lensing potential estimated by the QE estimator $\hat{\phi}^{0,\mathrm{MAP}} = \hat{\phi}^{\mathrm{QE}}$. Two maps are produced at each step of the iteration $\hat{E}^{i,\mathrm{WF}}$ and $\hat{\phi}^{i,\mathrm{MAP}}$, and the iteration continues ten times, a number that was verified with a separate set of simulations to produce solver convergence. Construction of $\hat{E}^{i,\mathrm{WF}}$ is the most computationally demanding task. We use the same multigrid-preconditioned conjugate-gradient solver from the Planck analysis (Planck Collaboration VIII 2020; Carron et al. 2022), which we modified to account for the maximum a-posteriori iterative solver. The last iteration produces the best estimate $\hat{\phi}^{\mathrm{MAP}}$, and the final Wiener filtered E-mode map is the best estimate for the delensed E-mode, namely $\hat{E}^{\mathrm{last,WF}}=\hat{E}^{\mathrm{NILC,MAP}}$.

7.2.2. *Estimating the primordial and lensing B-mode maps*

We produce two versions of each pair of the following: (1) a predicted B-mode lensing map, and (2) a delensed CMB B-mode map. One pair is based on $\hat{\phi}^{\mathrm{QE}}$ and $\hat{E}^{\mathrm{NILC,QE}}$ coming from the quadratic estimator, and a second pair is based on $\hat{\phi}^{\mathrm{MAP}}$ and $\hat{E}^{\mathrm{NILC,MAP}}$, which are the result of the maximum a-posteriori solver. We produce both pairs to compare their delensing performance. The B-mode lensing map is estimated by calculating the effect of the lensing field $\hat{\phi}$ on the estimated primordial, delensed E-mode field. The delensed (DL) CMB B-mode map is that obtained by subtracting the lensing B-mode map from $\hat{B}^{\mathrm{NILC}}$, that is,

$$\hat{B}^{\mathrm{QE,DL}} = \hat{B}^{\mathrm{NILC}} - \hat{B}^{\mathrm{QE,L}}\,, \qquad \hat{B}^{\mathrm{MAP,DL}} = \hat{B}^{\mathrm{NILC}} - \hat{B}^{\mathrm{MAP,L}}\,, \tag{13}$$

where $B^{\mathrm{QE,L}}$ and $B^{\mathrm{MAP,L}}$ are the lensing maps, which some authors refer to as 'lensing templates' (Hanson et al. 2013; Ade et al. 2021). When $r \neq 0$ and for an ideal full sky, high sensitivity, and high angular resolution experiment the DL maps would approach the true primordial signal. When $r = 0$ they would only include noise. In practice, delensing is never complete and the DL maps have residual lensing. Because we have access to the simulated input lensing map we assess the residual lensing (RL) for each of the lensing estimators. We make two maps of residuals

$$\hat{B}^{\mathrm{QE,RL}} = B^{\mathrm{L}} - \hat{B}^{\mathrm{QE,L}}\,, \qquad \hat{B}^{\mathrm{MAP,RL}} = B^{\mathrm{L}} - \hat{B}^{\mathrm{MAP,L}}\,. \tag{14}$$

where B^{L} are the simulated input lensing maps. With the RL maps we calculate the residual lensing amplitude A_{lens}. For each delensing estimator we calculate power spectra post-delensing for full sky maps and for maps with the NILC mask. After deconvolution from the mask (following Chon et al. (2004)), the spectra of the ten simulations are averaged to produce an average spectrum.

7.2.3. *Map delensing results*

Figure 9 shows the pre- and post-delensed B-mode maps as produced by the iterative maximum a-posteriori solver. They are compared to the input B-mode lensing B^{L} and the map of residual lensing $\hat{B}^{\mathrm{MAP,RL}}$. The $\hat{B}^{\mathrm{NILC}}$ map pre-delensing (top left) is a high resolution version of the map shown for model Baseline in Figure 3 (left column, third panel). The map variance is $\sigma^2 = (0.086\ \mu\mathrm{K})^2$, where σ is the full-sky pixel standard deviation of the maps band-passed to $2 \leq \ell \leq 200$, with the dominant contribution $\sigma^2 = (0.082\ \mu\mathrm{K})^2$ coming from the input lensing map, shown for a small patch of sky at the bottom left panel; 9% of the total variance comes from residual foregrounds and noise. The post-delensing variance decreases to $\sigma^2 = (0.047\ \mu\mathrm{K})^2$ (top right), of which $(0.039\ \mu\mathrm{K})^2$, or 69%, is due to the residual lensing and the noise introduced in the delensing process.

The reduction of map variance by a factor of 3 to 4 post-delensing is also apparent when comparing the angular power spectra of the pre-delensing map $\hat{B}^{\mathrm{NILC}}$ and the post-delensing map $\hat{B}^{\mathrm{MAP,DL}}$, see Figure 10. The results shown in the figure are equivalent to the first row of Figure 5 except that here the maps are rendered at high resolution (and we are using all frequency bands). The likelihood for r post-delensing, shown in the right panel, is nearly identical to the likelihood calculated with the fiducial 73% that has been used in Section 5.

Figure 11 gives a visualization of a different metric for PICO's delensing efficacy. We show the ten-simulation average spectrum of the input lensing B^{L} and the averages of spectra of residual lensing maps with the QE and maximum a-posteriori iterative solvers. The spectra for the full and masked skies have approximately the same amplitude, with

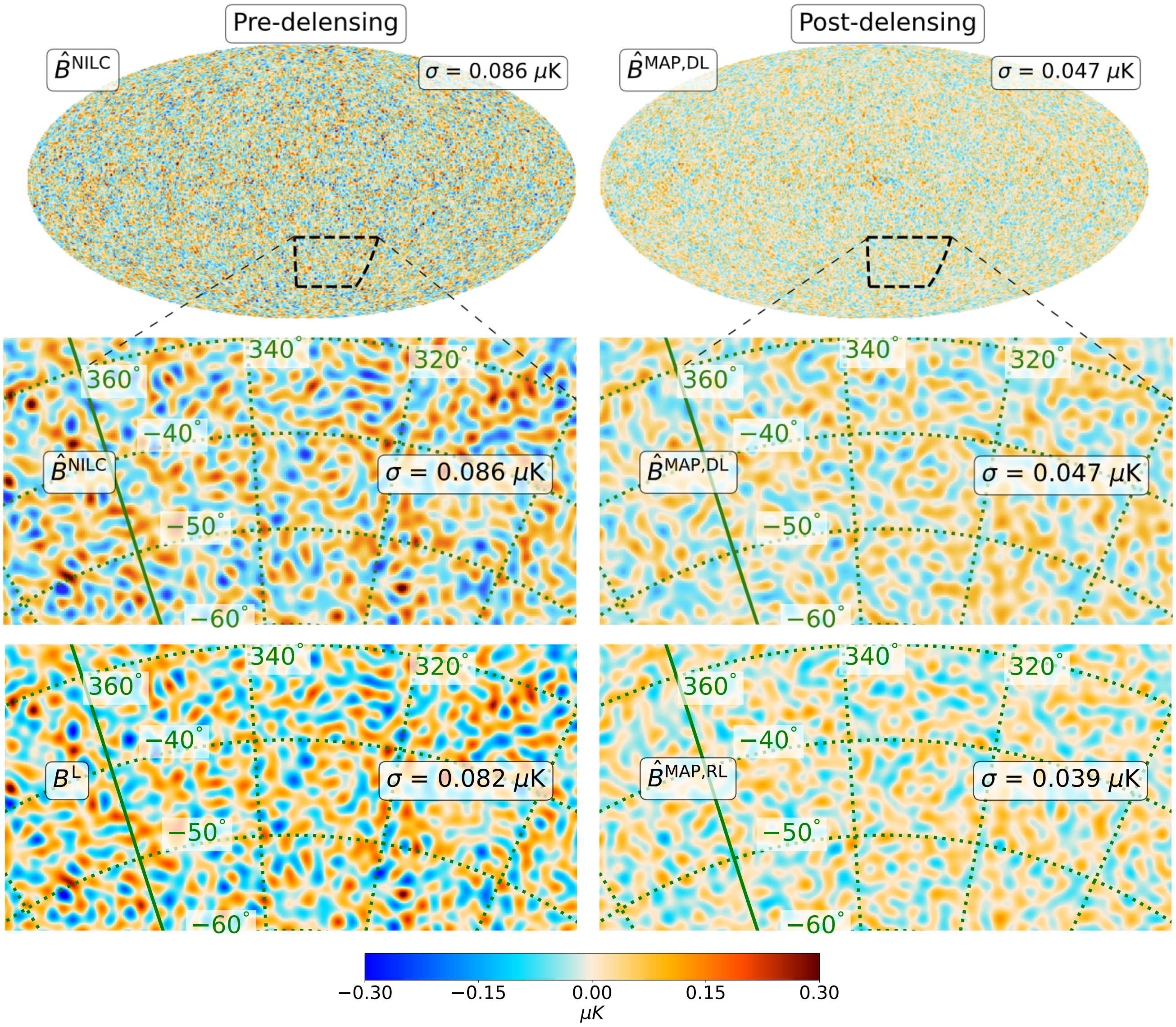

Figure 9. B-mode maps in Galactic coordinates after NILC component separation for model Baseline and $r_{\mathrm{in}} = 0$, before (left) and after (right) iterative delensing including full sky (top panels), smaller ($60^\circ \times 30^\circ$) area centered on a clean patch of the sky (middle panels), the input B-lensing map B^{L} (bottom left), and residual lensing map (bottom right) in the same area. All maps are band-passed to only show $2 \leq \ell \leq 200$, and no masking has been applied to the lensing reconstruction. The quantity σ in all maps is the full-sky pixel standard deviation of the band-passed maps.

the full sky version showing a variance that is on average lower by 48% compared to the masked version, consistent with the f_{sky} difference. Post-delensing there is a factor of $\sim$2 reduction in lensing power with QE and $\sim$4 with the iterative solver. The reduction in power is essentially uniform over $2 \leq \ell \leq 200$.

To quantify the delensing efficiency we calculate the residual delensing A_{lens} from the ratio of the residual lensing power spectrum $C_\ell^{\mathrm{BB,XX,RL}}$ and the input lensing $C_\ell^{\mathrm{BB,L}}$

$$\left\langle A_{\mathrm{lens}}^{\mathrm{XX}} \right\rangle = \left\langle \frac{C_\ell^{\mathrm{BB,XX,RL}}}{C_\ell^{\mathrm{BB,L}}} \right\rangle \tag{15}$$

where XX is for the QE or iterative solvers, and $\langle\rangle$ denotes averaging ten values of A_{lens}. We calculate A_{lens} values after binning the power spectra over the entire range of angular scales $2 \leq \ell \leq 200$, separately for the reionization

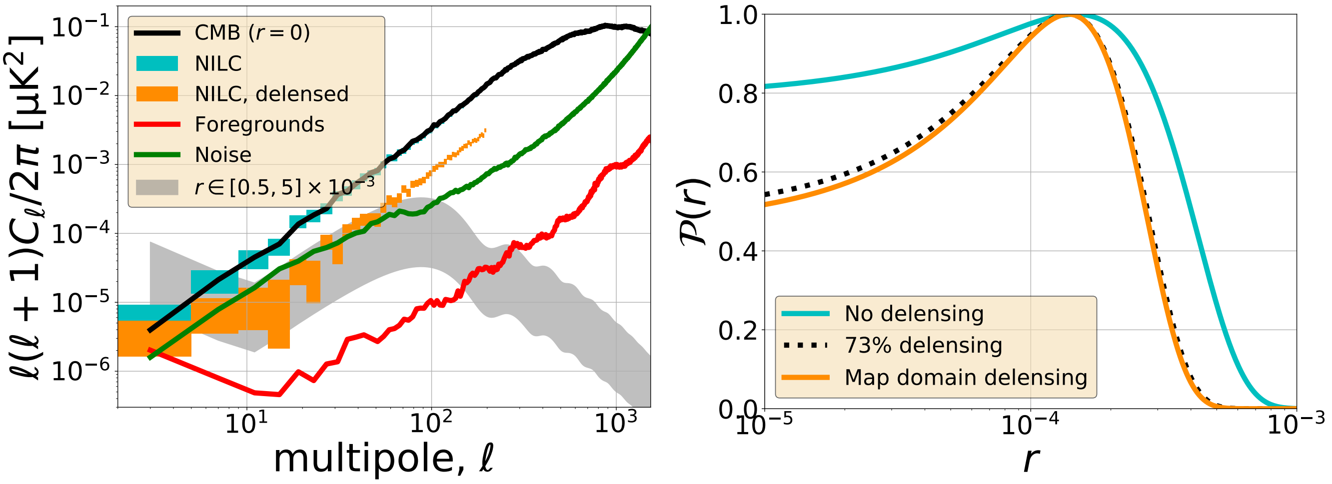

Figure 10. Left: Power spectra similar to those shown in Figure 4 but for high-resolution $\hat{B}^{\mathrm{NILC}}$ map with model Baseline and $r = 0$. Also included is the spectrum of the delensed map (orange). Right: r likelihood when $r_{\mathrm{in}} = 0$ without delensing (cyan), $r = (1.5 \pm 1.7) \times 10^{-4}$, with the approximate 73% delensing (dotted black), $r = (1.4 \pm 1.1) \times 10^{-4}$, and with map-domain delensing (orange), $r = (1.4 \pm 1.1) \times 10^{-4}$.

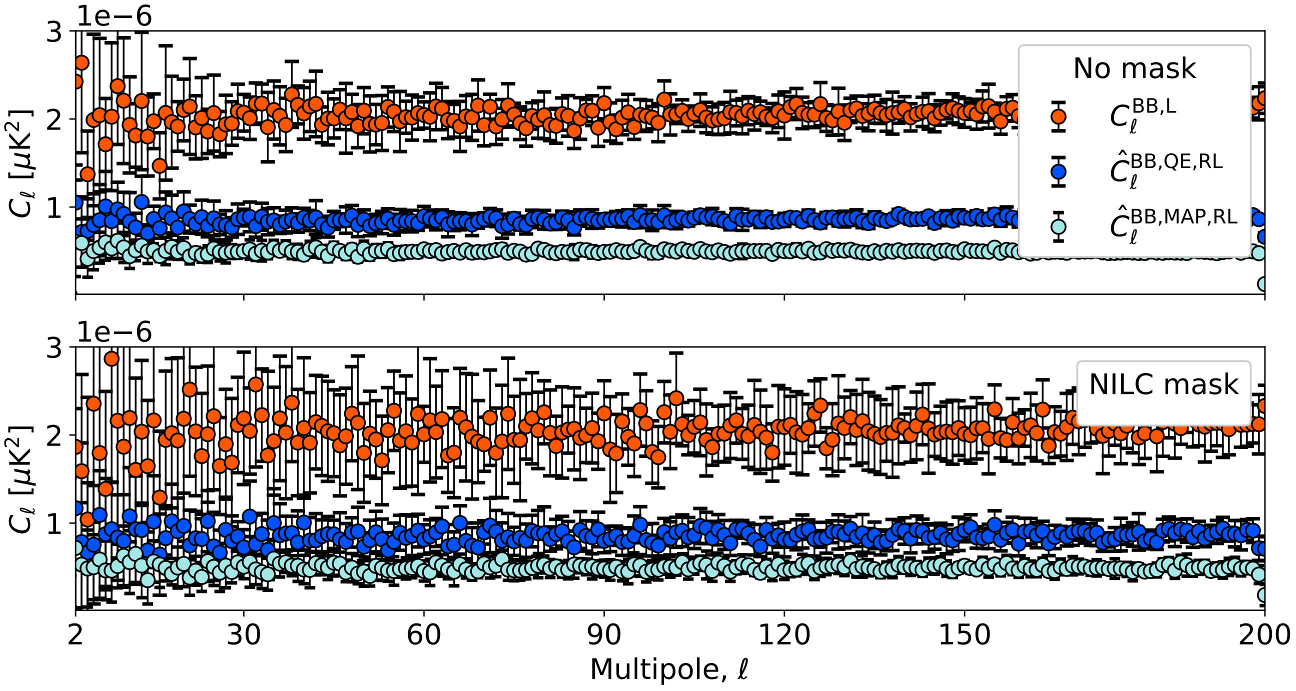

Figure 11. Ten-simulation average spectra of the input lensing B^{L} (red) and residual lensing maps $B^{\mathrm{XX,RL}}$, XX=QE (blue) or MAP (cyan), calculated over the entire sky (top) and with the NILC mask (bottom). Error bars are standard deviation over the ten simulations.

and recombination angular scales encompassing $(\ell_{\mathrm{reio}}, \ell_{\mathrm{reco}}) = ([2, 30), [30, 200])$, and for the full and masked sky. The values are given in Table 6.

The lowest residual lensing A_{lens}= 24% is achieved with the iterative solver when binning the entire range of angular scales, and it is consistent with A_{lens} for the recombination angular scale. Within statistical uncertainty it is also consistent with the value achieved for the reionization scale A_{lens}= 25%, see Table 6. The uncertainty intervals given are standard deviations of the ten simulations. We use two additional methods to validate our $A_{\rm lens}$ estimate. First, it is possible to predict the performance of the iterative solver (Smith et al. 2012; Legrand & Carron 2022), and this prediction gives $A_{\rm lens}$=0.23 and 0.22, for the reionization and recombination scales, respectively, which are close to the values we report. Second, the angular power spectrum domain analysis, described in Section 7.1, gave $A_{\rm lens}$=0.22 with the PICO map noise level we assume in this paper, see Table 5.

The QE solver gives $A_{\rm lens}$ values that are nearly two times larger than obtained with the iterative solver. We find no difference between values derived for the full and masked sky. For completeness, we note that the QE result could have been further optimized by calculating the effect of $\hat{\phi}^{QE}$ on $\hat{E}^{\rm NILC,QE}$ in a perturbative manner (Baleato Lizancos et al. 2021b). This approach would have led to removing at most an additional 3% of the lensing power. There is no analogous further optimization possible for the iterative solver.

Table 6. $A_{\rm lens}$ values obtained with the iterative and QE solvers for the full and masked sky, calculated over different ranges of angular scales. The uncertainties are the standard deviation over 10 simulations.

ℓ range	Iterative		Quadratic Estimator	
	Full	Masked	Full	Masked
$2-200$	0.239 ± 0.002	0.238 ± 0.005	0.416 ± 0.004	0.416 ± 0.005
$2-30$	0.25 ± 0.01	0.25 ± 0.02	0.43 ± 0.03	0.44 ± 0.04
$30-200$	0.238 ± 0.002	0.238 ± 0.005	0.416 ± 0.004	0.416 ± 0.006

8. DISCUSSION AND FUTURE WORK

The analysis that led to the constraints described in Sections 5 and 6 includes several assumptions, some of which could be relaxed in future work. We discuss these below together with other possible extensions:

- We assumed spatially uniform noise. Future work could begin with time-domain simulations that include PICO's sky scan pattern and $1/f$ noise, and that would give a map with a more realistic spatial noise distribution.

- There is limited information about the behavior of foregrounds across the sky. In particular, in diffuse regions observations are limited to $\ell \lesssim 100$. While there are models that include information up to $\ell \sim 1500$, extending simulations into this regime and into even higher ℓ in physically realistic ways is a topic of current research.

- The analysis ignores systematic uncertainties. Preliminary analysis shows that calibration uncertainties on Planck's CMB maps give negligible errors[2]. However, we have not included calibration uncertainties with bands that cannot be calibrated on CMB maps, nor other polarimetric systematic uncertainties.

- With all sky models, the NILC component separation has been carried out over the full sky, and we made no attempts to vary NILC parameters such as the shapes of the needlet window functions. Varying these parameters or masking the Galactic plane prior to component separation might lead to smaller residual foregrounds. With model MultiLayer, masking the Galactic plane prior to component separation might reduce the bias on r, although this is not guaranteed because the NILC algorithm is already fairly local in pixel space. Another strategy to reduce residual foreground biases, when they occur, would be to use a more constrained version of NILC, such as cMILC (Remazeilles et al. 2021) to enforce the minimization of the foreground variance instead of the overall variance. With this approach one deprojects the main statistical moments of the foreground emission (Chluba et al. 2017) out of the reconstructed CMB map.

- We chose to apply the same conservative mask to all sky models so as to compare their forecasts from the same portion of the sky. In a further study, the mask could be optimized separately for each sky model. The choice of the mask could be informed by the small patch analysis carried out for models Baseline and MultiLayer.

---

[2] z.umn.edu/picomission

- Estimates for r after component separation with NILC used a Gaussian approximation for the likelihood, and an approximation to account for delensing. While for $\ell < 12$ we showed that the exact likelihood gives weaker constraints by a factor of 1.3 relative to the Gaussian approximation, the magnitude of this factor for higher ℓ limit is not known. Including the effect of delensing into the r likelihood is also a topic of current active research.

- In this paper we took the focal plane configuration from Hanany et al. (2019). We did not attempt to optimize the focal plane as a function of r limits, neither in terms of the number of frequency bands, nor in terms of the noise level per band.

- Future work could extend to other foreground models, and to assessing the performance of parametric component separation, with Commander or otherwise, with more than the single model we included in this paper and over a larger range in ℓ.

Model MultiLayer stands out as a model that is plausible from a physics point of view, yet posing challenges to the component separation process, even with PICO's low noise and large number of frequency bands. The spectral decorrelation of thermal dust due to multi-layer contributions along the line-of-sight makes such foreground sky more complex because of the larger dimension of the foreground subspace. Model MultiLayer was not included in the analysis of Abazajian et al. (2022), which analyzed the expected performance of CMB-S4 over a small, $\sim 5\%$ area of the sky, and with somewhat looser target constraints on r. There is therefore no information on potential biases with a narrower frequency coverage.

We have shown, however, that PICO's deep integration over the full sky and its high frequency bands that are only sensitive to Galactic dust, enable a powerful cross check by analyzing individual small patches of the sky and correlating the inferred r determinations or limits with the variance in dust polarized intensity; see Section 5.2.1. The analysis methodology we presented – quantifying r constraints as a function of observed local foreground polarized intensity and searching for asymptotic behavior as a function of decreasing foreground levels – only uses data as observed by the mission, and parallels the process that is likely to be pursued by an analysis team. The fact that r determinations or limits converge to the same value does not guarantee the absence of bias, but it is, however, a strong systematic cross-check. The more sky areas that give the same result, the higher the confidence there is no bias.

The small-patch analysis also suggests that constraining observations to the lowest dust polarized intensity variance regions could provide immunity from biased detections. We showed that only having the region away from the Galactic plane is not sufficient. The analysis we presented relies on data from a relatively high frequency band, which is entirely dominated by dust and may not be available for other instruments. We have not repeated the analysis with other instrument noise levels, nor with a different mix of frequency bands.

Parametric component separation with Commander has only been attempted with model Baseline. Both $r = 0$ and $r = 0.003$ are recovered with no bias. No delensing has been applied and therefore the constraints achieved cannot be compared to PICO's requirements. The constraints are nevertheless useful because they illustrate the limitations of using a Gaussian likelihood approximation, specifically when limiting the analysis to low ℓ multipoles. Commander uses a Blackwell-Rao approximation, which, as we show in the Appendix, closely matches the exact likelihood in the case of full sky coverage and uniform noise. Using the Blackwell-Rao approximation, and when $\ell \leq 12$, Commander gives a factor of two higher upper limits compared to NILC. We note that in the NILC analysis we use $\ell \leq 200$. Assessing the differences between the exact, Blackwell-Rao, and Gaussian likelihoods over this extended ℓ range is computationally demanding and there are currently no quantitative estimates.

9. SUMMARY

We used blind and parametric component separation techniques to assess PICO's capability to reach its r determination requirements. The analysis employs realistic noise levels, only the prime mission duration without any extensions (even though the mission has no consumables), sky foreground models that match current data and are broadly accepted as plausible, and analysis techniques that have been implemented with other instruments including Planck. The analysis assumes a spatially uniform noise with white spectrum, and does not include systematic uncertainties. We initially used an angular power spectrum domain analytic calculation to find a predicted level of delensing and took the most conservative value obtained for the residual delensing factor $A_{\rm lens}$= 0.27.

For four out of the five sky models, Baseline, 2MBB, 2PhysDust, and MHD, and with $r_{\rm in} = 0$ the PICO experiment configuration that has 21 frequency bands gives 95% upper limits between 1.3×10^{-4} and 2.7×10^{-4}; see Table 3. These constraints could definitively rule out the simplest models of inflation that predict $r \simeq 0.001$. With the same four models, if the true value of r is $r_{\rm in} = 0.003$ it would be detected with confidence levels between 18σ and 27σ after 5 yrs of the prime mission. The r estimates are based on a Gaussian approximation to the likelihood. They are the strongest upper limits and detections predicted for any instrument in the foreseeable future.

Removal of the lowest or the highest frequency bands weakens the upper limits when $r_{\rm in} = 0$, and gives more biased r likelihoods; see Table 4 and Figure 5. With models Baseline and MHD, nearly 3σ biases occur with removal of either low or high frequency bands. Although removal of the higher frequency bands results in somewhat more pronounced biases, to achieve its requirements a space mission like PICO requires bands over the entire frequency range.

With model MultiLayer the r estimates are significantly biased. However, PICO's deep integration over the full sky and its high frequency bands, enables identification and mitigation of the bias. Model MultiLayer is an example for the limitation of inference using an r determination over a small patch of sky and with limited frequency band data.

Parametric component separation with Commander has only been attempted with model Baseline and without delensing. Both $r = 0$ and $r = 0.003$ are recovered with no bias.

We presented the first application of map-domain iterative delensing, and it is the first application of iterative delensing on maps that have first undergone a component separation exercise. The delensing has been applied over the entire sky because the map for model Baseline after component separation does not indicate the need to mask any region, including along the Galactic plane; see Figure 9.

The map-domain delensing gives a residual delensing factor $A_{\rm lens}$= 0.24, which would have been $A_{\rm lens}$= 0.22 had the range of multipoles $\ell < 200$ been included in the delensing process; see Section 7.2.1. This level reproduces the level we calculated using the analytic approximation, which predicted $A_{\rm lens}$= 0.22 for the noise levels we assumed here; see Section 7.1. The map-domain delensing confirms that the choice we made to use $A_{\rm lens}$= 0.27 is conservative.

APPENDIX

A. COMPARISON OF GAUSSIAN AND BLACKWELL-RAO LIKELIHOODS

In this appendix we compare the Gaussian approximation used in the NILC analysis and the Blackwell-Rao estimator used in the Commander analysis. We consider only ideal primordial Gaussian CMB fluctuations and instrumental noise, and ignore instrumental systematic effects, foreground subtraction, and other potential complications. Explicitly, we assume that the data d may be written as $d = s + n$, where s is a Gaussian distributed true CMB signal with covariance matrix S, and n is Gaussian distributed instrumental noise with covariance matrix N. The sum of the two terms is Gaussian distributed with total covariance $S + N$, and the full exact likelihood is

$$\mathcal{L}_0 = \frac{\exp\left(-\frac{1}{2}d^T(S+N)^{-1}d\right)}{\sqrt{|S+N|}}. \tag{A1}$$

The covariance N is typically a deterministic quantity, given by the noise and scanning strategy of the instrument, and S is typically defined by additionally assuming that the CMB is statistically isotropic. In this case S is block-diagonal in the spherical harmonic domain, and $S_{lm,l'm'} = C_\ell \delta_{ll'}\delta_{mm'}$, where C_l is the angular power spectrum, which is a function of the cosmological parameters. In our case, $S = S(r)$.

For full sky data with uniform noise per pixel Equation A1 may be written in harmonic space

$$\mathcal{L}_1 = \prod_\ell \frac{\exp\left(-\frac{2\ell+1}{2}\frac{\sigma_\ell}{b_\ell^2 p_\ell^2 C_\ell + N_\ell}\right)}{(b_\ell^2 p_\ell^2 C_\ell + N_\ell)^{\frac{2\ell+1}{2}}}, \tag{A2}$$

$$\sigma_\ell \equiv \frac{1}{2\ell+1}\sum_{\ell=-m}^{m} |a_{\ell m}|^2, \tag{A3}$$

$$N_\ell = \sigma_0^2 4\pi/N_{\rm pix}, \tag{A4}$$

where σ_ℓ is the observed realization-specific power spectrum, N_ℓ is the noise power spectrum, σ_0 is the noise per pixel, and b_ℓ and p_ℓ are the beam and pixel window transfer functions, respectively. In order to use this expression to estimate r, one may write (Ade et al. 2019)

$$C_\ell \equiv C_\ell(r, A_{\rm lens}) = rC_\ell^{\rm tens} + A_{\rm lens}C_\ell^{\rm lens}, \tag{A5}$$

where the total power C_ℓ has contributions from the inflationary, tensor term and from the lensing term, which has a delensing factor $1 - A_{\rm lens}$, and a residual lensing factor $A_{\rm lens}$. When $A_{\rm lens}$=1 there is no delensing. The nominal 0.73 delensing factor value used with NILC in Section 5 leaves $A_{\rm lens}$= 0.27 of the original $C_\ell^{\rm lens}$ power in the map. In this expression, we implicitly assume that the net distortion effect from lensing may also be approximated as Gaussian.

The likelihood $\mathcal{L}_1$ is not valid for data with a spatially varying noise distribution or when the data do not include the full sky. Even for uniform noise and full sky, calculating $\mathcal{L}_0$ is too computationally expensive to evaluate by brute-force methods due to expensive matrix inversion and determinant operations. For these reasons many likelihood approaches have been developed that aim to approximate $\mathcal{L}_0$. One common approach is to assume that $\mathcal{L}(C_l)$ is close to a Gaussian distribution. In fact, $\mathcal{L}_1$ is an inverse Gamma distribution in C_l, which converges to a Gaussian for large ℓ. For $\ell \gtrsim 30$ the Gaussian approximation is acceptable (e.g., Hamimeche & Lewis 2008).

Motivated by this observation, the NILC component separation pipeline uses the following expression for the likelihood

$$\mathcal{L}_3 \propto \exp\left(-\frac{1}{2}\sum_\ell \frac{\left(\widehat{C}_\ell^{BB,\rm NILC} - C_\ell\left(r, A_{\rm lens}\right) - \widehat{N}_\ell\right)^2}{\widehat{\Xi}_\ell^2}\right), \tag{A6}$$

$$\widehat{\Xi}_\ell = \sqrt{\frac{2}{(2\ell+1)\, f_{\rm sky}}}\, \widehat{C}_\ell^{BB,\rm NILC}, \tag{A7}$$

where $\hat{N}_\ell$ is an estimate of the NILC B-mode noise power spectrum after component separation, $\widehat{C}_\ell^{BB,\rm NILC}$ is the total measured power spectrum after component separation (see Section 5) and possibly delensing (see Section 7.2), and $\widehat{\Xi}_\ell$ is the corresponding variance, which implicitly includes the cosmic and sample variance contributions of the primordial signal (if any), the residual lensing signal, the residual foreground signal, and the noise. When using NILC to estimate r, we evaluate $\mathcal{L}_3$ as a function of r.

Commander addresses the computational cost associated with evaluating $\mathcal{L}_0$ through Bayesian Monte Carlo sampling. With this method one first draws many Monte Carlo samples from the joint sky signal and power spectrum distribution, i.e., $P(s, C_l|d)$. At first sight, this may appear more complicated than working directly with $P(C_l|d)$. However, joint samples may be drawn through a particularly efficient Gibbs sampling algorithm (Jewell et al. 2004; Wandelt et al. 2004) by first sampling from $P(s|d, C_l)$ and then from $P(C_l|d, s)$, and this two-step approach is in fact computationally far cheaper than evaluating the marginal distribution directly. Each sky map sample s^i represents one possible ideal CMB sky map consistent with the observed data, and the histogram of all these Monte Carlo samples converges to the true likelihood in Equation A1.

A commonly used statistical method for speeding up the convergence rate of a Monte Carlo estimator is the Blackwell-Rao (BR) approximation (Chu et al. 2005). Rather than building up a histogram directly from the sample set, which is subject to a high Monte Carlo uncertainty, this method instead averages the analytical probability distributions for each sample. In our case, this is equivalent to averaging $\mathcal{L}_1$ over all s^i. The advantage of this method is that the Monte Carlo sampler only needs to map out uncertainties due to instrumental noise and masks, but not cosmic variance; that source of uncertainty is instead handled analytically by the expression in Equation A2.

In practice, Monte Carlo convergence can be improved even further through ‘Gaussianization’ (Rudjord et al. 2009). In that approach, one first averages $\mathcal{L}_1$ over a large number of samples, setting $b_\ell = p_\ell = 1$ for each ℓ. We call the resulting distribution $\mathcal{L}_\ell(C_\ell)$. This is then used to define a transformation $x_\ell(C_\ell)$ that maps $\mathcal{L}_l(C_l)$ to a perfect Gaussian by matching their cumulative distributions percentile-by-percentile. One then corrects for correlations between ℓs using

$$\mathcal{L}_2 = P(C_\ell|d) \approx \left(\prod_\ell \frac{\partial C_\ell}{\partial x_\ell}\right)^{-1} \exp\left(-\frac{1}{2}(x-\mu)^T C^{-1}(x-\mu)\right). \tag{A8}$$

Here, $x = \{x_\ell(C_\ell)\}$ is the vector of Gaussianized input power spectrum coefficients, $\frac{\partial C_\ell}{\partial x_\ell}$ is the Jacobian of this Gaussianizing transformation, and the mean μ and covariance matrix $C_{\ell\ell'}$ are estimated from Monte Carlo samples of the full posterior distribution. The advantage of this approach is that only second-order correlations need to be mapped out through Monte Carlo sampling, as opposed to all possible N-point correlations, and this greatly speeds up the overall convergence rate. The disadvantage is that the expression is no longer an exact representation of $\mathcal{L}_1$,

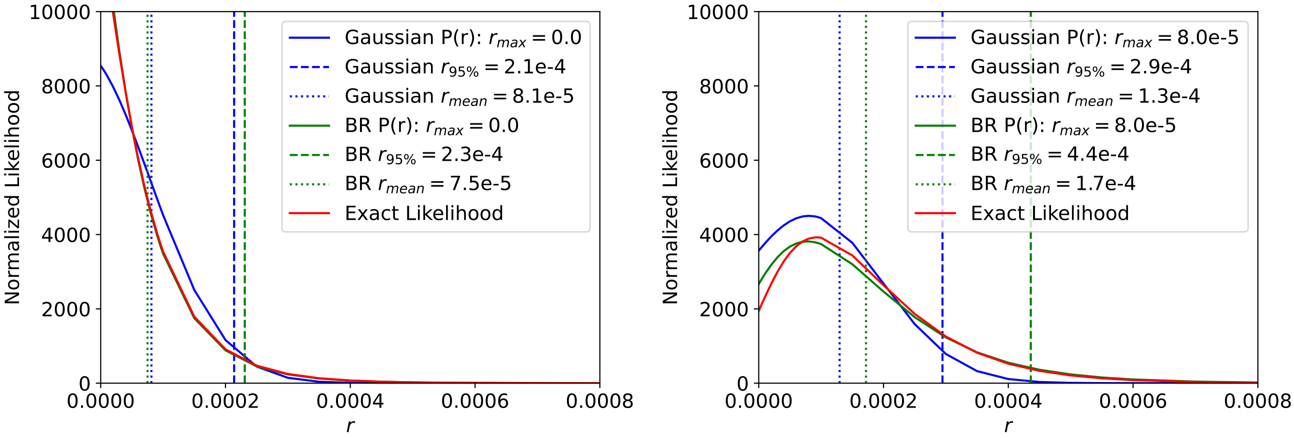

Figure 12. Two simulations with $r = 0$ (left and right) comparing the likelihood curves calculated using the exact (red), Gaussian (blue), and BR (green) likelihoods. In all cases r is estimated for $2 \leq \ell \leq 12$.

and it can break down for data that cover only small sky fractions and that have strong correlations. However, for $f_{\rm sky} \gtrsim 0.70$, as is typical for CMB satellite experiments, it is an excellent approximation (Rudjord et al. 2009).

We quantitatively compare the Gaussian and the BR approximations to the exact likelihood $\mathcal{L}_0$. We generate ten CMB full sky simulations that all have $r = 0$ and a noise level of $\sigma_0 = 0.3\,\mu$K/pixel at $N_{\rm side} = 512$, set by the estimated NILC noise level shown in Figure 4. We omit delensing to simplify the comparison. We compare the $r_{95\%}$ upper limits obtained with $\mathcal{L}_1$, $\mathcal{L}_2$, and $\mathcal{L}_3$ using $2 \leq \ell \leq 12$. The resulting posterior distributions are thus not representative of the full PICO sensitivity, which also includes higher multipoles, but are only useful for comparing the three likelihood approximations. Figure 12 shows the r likelihoods for two of these simulations, one for which the likelihood peaks at zero (left panel) and another for which a noise fluctuation gives a non-zero value for the peak of the likelihood (right panel). In both cases the BR approach gives a larger $r_{95\%}$, although this is more pronounced when the peak has a non-zero value. In both cases the BR likelihood $\mathcal{L}_2$ agrees quite closely with the exact likelihood $\mathcal{L}_0$, and is in better agreement than the Gaussian approximation $\mathcal{L}_3$. This suggests that $\mathcal{L}_2$ might also be closer to $\mathcal{L}_0$ when only part of the sky is used.

The distribution of $r_{95\%}$ for the ten simulations are given in Figure 13. On average the BR likelihood predicts upper limits that are a factor of 1.3 times larger than the Gaussian likelihood. However, this may be split into two classes, those that peak at zero and those that peak at a non-zero value. Examination of the likelihoods indicates that the Gaussian approximation performs well for the former class, while for the latter, the BR $r_{95\%}$ value may be up to nearly two times larger than the Gaussian, as happens in realizations 5 and 8.

This research used resources of the National Energy Research Scientific Computing Center, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. MR would like to thank the Spanish Agencia Estatal de Investigación (AEI, MICIU) for the financial support provided under the project with reference PID2019-110610RB-C21.

REFERENCES

Abazajian, K., Addison, G. E., Adshead, P., et al. 2022, ApJ, 926, 54, doi: 10.3847/1538-4357/ac1596

Abazajian, K. N., et al. 2016. https://arxiv.org/abs/1610.02743

Adak, D., Sen, A., Basak, S., et al. 2021, arXiv e-prints, arXiv:2110.12362. https://arxiv.org/abs/2110.12362

Ade, P., Aguirre, J., Ahmed, Z., et al. 2019, JCAP, 2019, 056, doi: 10.1088/1475-7516/2019/02/056

Ade, P. A. R., Ahmed, Z., Amiri, M., et al. 2021, PhRvL, 127, 151301, doi: 10.1103/PhysRevLett.127.151301

Ade, P. A. R., et al. 2021, Phys. Rev. D, 103, 022004, doi: 10.1103/PhysRevD.103.022004

Baleato Lizancos, A., Challinor, A., & Carron, J. 2021a, JCAP, 03, 016, doi: 10.1088/1475-7516/2021/03/016

—. 2021b, Phys. Rev. D, 103, 023518, doi: 10.1103/PhysRevD.103.023518

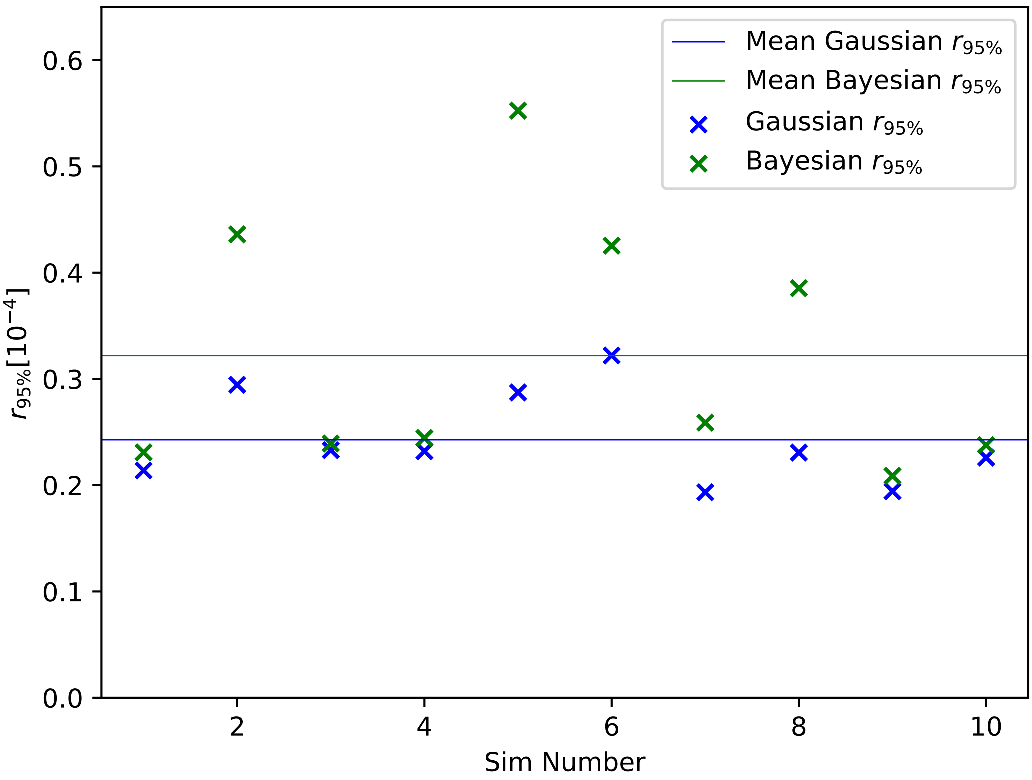

Figure 13. $r_{95\%}$ limits for ten $r = 0$ CMB and noise simulations, for the Gaussian (blue) and BR (green) likelihoods. The averages of the two sets ($r = 2.4 \times 10^{-4}$ for Gaussian and 3.2×10^{-4} for BR) are plotted as the horizontal lines.

Carron, J., & Lewis, A. 2017, Phys. Rev. D, 96, 063510, doi: 10.1103/PhysRevD.96.063510

Carron, J., Mirmelstein, M., & Lewis, A. 2022. https://arxiv.org/abs/2206.07773

Chluba, J., Hill, J. C., & Abitbol, M. H. 2017, MNRAS, 472, 1195, doi: 10.1093/mnras/stx1982

Choi, S. K., & Page, L. A. 2015, JCAP, 2015, 020, doi: 10.1088/1475-7516/2015/12/020

Chon, G., Challinor, A., Prunet, S., Hivon, E., & Szapudi, I. 2004, Mon. Not. Roy. Astron. Soc., 350, 914, doi: 10.1111/j.1365-2966.2004.07737.x

Chu, M., Eriksen, H. K., Knox, L., et al. 2005, PhRvD, 71, 103002, doi: 10.1103/PhysRevD.71.103002

Colombo, L. P. L., Eskilt, J. R., Paradiso, S., et al. 2022, arXiv e-prints, arXiv:2208.14276. https://arxiv.org/abs/2208.14276

Delabrouille, J., & Cardoso, J. F. 2009, in Data Analysis in Cosmology, ed. V. J. Martínez, E. Saar, E. Martínez-González, & M. J. Pons-Bordería, Vol. 665, 159–205, doi: 10.1007/978-3-540-44767-2_6

Delabrouille, J., Cardoso, J. F., Le Jeune, M., et al. 2009, A&A, 493, 835, doi: 10.1051/0004-6361:200810514

Delabrouille, J., Betoule, M., Melin, J. B., et al. 2013, A&A, 553, A96, doi: 10.1051/0004-6361/201220019

Delabrouille, J., de Bernardis, P., Bouchet, F. R., et al. 2018, JCAP, 2018, 014, doi: 10.1088/1475-7516/2018/04/014

Draine, B. T., & Hensley, B. 2013, ApJ, 765, 159, doi: 10.1088/0004-637X/765/2/159

Eriksen, H. K., Jewell, J. B., Dickinson, C., et al. 2008, ApJ, 676, 10, doi: 10.1086/525277

Eriksen, H. K., O'Dwyer, I. J., Jewell, J. B., et al. 2004, ApJS, 155, 227, doi: 10.1086/425219

Galloway, M., Andersen, K. J., Aurlien, R., et al. 2022, arXiv e-prints, arXiv:2201.03509. https://arxiv.org/abs/2201.03509

Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759, doi: 10.1086/427976

Hamimeche, S., & Lewis, A. 2008, PhRvD, 77, 103013, doi: 10.1103/PhysRevD.77.103013

Hanany, S., Alvarez, M., Artis, E., et al. 2019, arXiv e-prints, arXiv:1902.10541. https://arxiv.org/abs/1902.10541

Hanson, D., et al. 2013, Phys. Rev. Lett., 111, 141301, doi: 10.1103/PhysRevLett.111.141301

Haslam, C. G. T., Salter, C. J., Stoffel, H., & Wilson, W. E. 1982, A&AS, 47, 1

Hazumi, M., Ade, P. A. R., Adler, A., et al. 2020, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 11443, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 114432F, doi: 10.1117/12.2563050

Hensley, B. 2015, PhD thesis, Princeton University

Hinshaw, G., Nolta, M. R., Bennett, C. L., et al. 2007, ApJS, 170, 288, doi: 10.1086/513698

Hivon, E., Gorski, K. M., Netterfield, C. B., et al. 2002, The Astrophysical Journal, 567, 2, doi: 10.1086/338126

Hu, W. 2002, Phys. Rev. D, 65, 023003, doi: 10.1103/PhysRevD.65.023003

Hu, W., & Okamoto, T. 2002, Astrophys. J., 574, 566, doi: 10.1086/341110

Jewell, J., Levin, S., & Anderson, C. H. 2004, ApJ, 609, 1, doi: 10.1086/383515

Kim, C.-G., Choi, S. K., & Flauger, R. 2019, ApJ, 880, 106, doi: 10.3847/1538-4357/ab29f2

Kim, C.-G., & Ostriker, E. C. 2017, ApJ, 846, 133, doi: 10.3847/1538-4357/aa8599

Larson, D. L., Eriksen, H. K., Wandelt, B. D., et al. 2007, ApJ, 656, 653, doi: 10.1086/509802

Leach, S. M., Cardoso, J. F., Baccigalupi, C., et al. 2008, A&A, 491, 597, doi: 10.1051/0004-6361:200810116

Legrand, L., & Carron, J. 2022, Phys. Rev. D, 105, 123519, doi: 10.1103/PhysRevD.105.123519

Lewis, A., & Challinor, A. 2006, Phys. Rept., 429, 1, doi: 10.1016/j.physrep.2006.03.002

Linde, A. 2017, JCAP, 02, 006, doi: 10.1088/1475-7516/2017/02/006

LiteBIRD Collaboration, Allys, E., Arnold, K., et al. 2022, arXiv e-prints, arXiv:2202.02773. https://arxiv.org/abs/2202.02773

Maniyar, A. S., Ali-Haïmoud, Y., Carron, J., Lewis, A., & Madhavacheril, M. S. 2021, Phys. Rev. D, 103, 083524, doi: 10.1103/PhysRevD.103.083524

Martínez-Solaeche, G., Karakci, A., & Delabrouille, J. 2018, MNRAS, 476, 1310, doi: 10.1093/mnras/sty204

Meisner, A. M., & Finkbeiner, D. P. 2015, ApJ, 798, 88, doi: 10.1088/0004-637X/798/2/88

Miville-Deschênes, M. A., Ysard, N., Lavabre, A., et al. 2008, A&A, 490, 1093, doi: 10.1051/0004-6361:200809484

Narcowich, F., Petrushev, P., & Ward, J. 2006, SIAM J. Math. Anal., 38, 574

National Academies of Sciences, Engineering, and Medicine. 2021, Pathways to Discovery in Astronomy and Astrophysics for the 2020s

Okamoto, T., & Hu, W. 2003, Phys. Rev. D, 67, 083002, doi: 10.1103/PhysRevD.67.083002

Paradiso, S., Colombo, L. P. L., Andersen, K. J., et al. 2022, arXiv e-prints, arXiv:2205.10104. https://arxiv.org/abs/2205.10104

Pelgrims, V., Clark, S. E., Hensley, B. S., et al. 2021, A&A, 647, A16, doi: 10.1051/0004-6361/202040218

Planck Collaboration, Aghanim, N., Ashdown, M., et al. 2016, A&A, 596, A109, doi: 10.1051/0004-6361/201629022

Planck Collaboration, Akrami, Y., Ashdown, M., et al. 2020a, A&A, 641, A4, doi: 10.1051/0004-6361/201833881

Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020b, A&A, 641, A3, doi: 10.1051/0004-6361/201832909

Planck Collaboration X. 2016, A&A, 594, A10, doi: 10.1051/0004-6361/201525967

Planck Collaboration XV. 2016, A&A, 594, A15, doi: 10.1051/0004-6361/201525941

Planck Collaboration IV. 2020, A&A, 641, A4, doi: 10.1051/0004-6361/201833881

Planck Collaboration VIII. 2020, A&A, 641, A8, doi: 10.1051/0004-6361/201833886

Planck Collaboration XI. 2020, A&A, 641, A11, doi: 10.1051/0004-6361/201832618

Remazeilles, M., Delabrouille, J., & Cardoso, J.-F. 2011a, MNRAS, 418, 467, doi: 10.1111/j.1365-2966.2011.19497.x

—. 2011b, MNRAS, 410, 2481, doi: 10.1111/j.1365-2966.2010.17624.x

Remazeilles, M., Rotti, A., & Chluba, J. 2021, MNRAS, 503, 2478, doi: 10.1093/mnras/stab648

Remazeilles, M., Banday, A. J., Baccigalupi, C., et al. 2018, JCAP, 2018, 023, doi: 10.1088/1475-7516/2018/04/023

Rudjord, Ø., Groeneboom, N. E., Eriksen, H. K., et al. 2009, ApJ, 692, 1669, doi: 10.1088/0004-637X/692/2/1669

Seljak, U., & Hirata, C. M. 2004, Phys. Rev. D, 69, 043005, doi: 10.1103/PhysRevD.69.043005

Seljebotn, D. S., Bærland, T., Eriksen, H. K., Mardal, K. A., & Wehus, I. K. 2017, arXiv e-prints, arXiv:1710.00621. https://arxiv.org/abs/1710.00621

Shandera, S., Adshead, P., Amin, M., et al. 2019, BAAS, 51, 338. https://arxiv.org/abs/1903.04700

Sherwin, B. D., & Schmittfull, M. 2015, Phys. Rev. D, 92, 043005, doi: 10.1103/PhysRevD.92.043005

Smith, K. M., Hanson, D., LoVerde, M., Hirata, C. M., & Zahn, O. 2012, JCAP, 06, 014, doi: 10.1088/1475-7516/2012/06/014

Tassis, K., & Pavlidou, V. 2015, MNRAS, 451, L90, doi: 10.1093/mnrasl/slv077

Teng, W.-H., Kuo, C.-L., & Wu, J.-H. P. 2011. https://arxiv.org/abs/1102.5729

Thorne, B., Dunkley, J., Alonso, D., & Næss, S. 2017, MNRAS, 469, 2821, doi: 10.1093/mnras/stx949

Tristram, M., Banday, A. J., Górski, K. M., et al. 2021, arXiv e-prints, arXiv:2112.07961. https://arxiv.org/abs/2112.07961

Wandelt, B. D., Larson, D. L., & Lakshminarayanan, A. 2004, PhRvD, 70, 083511, doi: 10.1103/PhysRevD.70.083511

Young, K., Alvarez, M., Battaglia, N., et al. 2018, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 10698, Space Telescopes and Instrumentation 2018: Optical, Infrared, and Millimeter Wave, ed. M. Lystrup, H. A. MacEwen, G. G. Fazio, N. Batalha, N. Siegler, & E. C. Tong, 1069846, doi: 10.1117/12.2309421

Zaldarriaga, M., & Seljak, U. 1998, Phys. Rev. D, 58, 023003, doi: 10.1103/PhysRevD.58.023003

4

Studying Non-Gaussian Induced Biases

4.1 Motivation

Beyond the Gaussian approximation of the lensing field lays the uncharted territory of the impact of non-Gaussian imprints onto the iterative reconstruction. Understanding this is a crucial part for future analyses of real CMB data, due to the slight non-Gaussianity of the true deflection field.

It is known that a non-Gaussian deflection field induces a bispectrum in QE. This was a well studied aspect of previous research and was the starting point of this project. This effect had not been quantified in the context of MAP reconstruction. Eventually, it is MAP that will be used to reconstruct the deflection field and support the detection of primordial B-modes, hereby necessitating the quantification of this effect.

Another rather interesting aspect of the analysis is the characterization of the induced $N^{(3/2)}$ bias in terms of higher order correlations: if the majority of the $N^{(3/2)}$ bias is truly induced by a bispectrum, a sign-flip of the convergence field should negate the impact. If however, the bispectrum receives contributions from higher, even-order correlation functions due to MAP, we will see a residual signal as they cannot cancel out by the sign-flip of the κ map.

My contribution to the paper is the lensing of the simulation data with both Gaussian and non-Gaussian deflection fields, calculating the QE and iterative reconstruction for all estimators, calculating the empirical Wiener-filter, calculating the $N^{(3/2)}$ biases, calculating the sign-flipped κ analysis, and supporting writing the section sign-flipped kappa. Further, I calculated the PTE values for this test.

4.2 N32

This section contains the $N^{(3/2)}$ paper as appeared on the ArXiv,[1] and has been approved on Physical Review D.

---

[1] https://arxiv.org/abs/2407.00228

Non-Gaussian deflections in iterative optimal CMB lensing reconstruction

Omar Darwish,[1, *] Sebastian Belkner,[1, †] Louis Legrand,[1, 2] Julien Carron,[1] and Giulio Fabbian[3, 4, 5, 6]
[1] *Université de Genève, Département de Physique Théorique et CAP, 24 Quai Ansermet, CH-1211 Genève 4, Switzerland*
[2] *Instituto de Física Teórica da Universidade Estadual Paulista and ICTP South American Institute for Fundamental Research, R. Dr. Bento Teobaldo Ferraz, 271, Bloco II, Barra-Funda - São Paulo/SP, Brasil*
[3] *Institute of Astronomy, Madingley Road, Cambridge CB3 0HA, UK*
[4] *Kavli Institute for Cosmology Cambridge, Madingley Road, Cambridge CB3 0HA, UK*
[5] *School of Physics and Astronomy, Cardiff University, The Parade, Cardiff, Wales CF24 3AA, United Kingdom*
[6] *Center for Computational Astrophysics, Flatiron Institute, 162 Fifth Avenue, New York, NY, 10010, USA*

(Dated: July 5, 2024)

The gravitational lensing signal from the Cosmic Microwave Background is highly valuable to constrain the growth of the structures in the Universe in a clean and robust manner over a wide range of redshifts. One of the theoretical systematics for lensing reconstruction is the impact of the lensing field non-Gaussianities on its estimators. Non-linear matter clustering and post-Born lensing corrections are known to bias standard quadratic estimators to some extent, most significantly so in temperature. In this work, we explore the impact of non-Gaussian deflections on Maximum a Posteriori lensing estimators, which, in contrast to quadratic estimators, are able to provide optimal measurements of the lensing field. We show that these naturally reduce the induced non-Gaussian bias and lead to unbiased cosmological constraints in ΛCDM at CMB-S4 noise levels without the need for explicit modelling. We also test the impact of assuming a non-Gaussian prior for the reconstruction; this mitigates the effect further slightly, but generally has little impact on the quality of the reconstruction. This shows that higher-order statistics of the lensing deflections are not expected to present a major challenge for optimal CMB lensing reconstruction in the foreseeable future.

I. INTRODUCTION

Gravitational lensing of the cosmic microwave background (CMB) is one of the leading cosmological probes of the next generation of CMB polarization surveys such as Simons Observatory (SO) [1] and CMB-S4 [2]. Maps of the CMB lensing potential and its summary statistics can provide clean and robust probes of the large-scale structure (LSS) in the Universe and allow to constrain cosmological parameters that govern the LSS growth and to which CMB anisotropies alone are weakly sensitive (for example neutrino mass, dark energy properties, curvature) [3–6]. On the other hand, galaxy surveys in different wavelengths probing the same LSS distribution (through galaxy clustering, weak lensing or intensity mapping) are also expected to extract complementary cosmological information on the LSS growth, or on primordial non-Gaussianities. Through galaxy-CMB lensing cross-correlations we will be able to marginalize over some observational systematics (e.g. shear multiplicative bias) [7–9], or modeling uncertainties (e.g. galaxy bias, magnification bias) [10–12]. Masses of high-redshift galaxy cluster samples can also be calibrated with the highest sensitivity through stacking of CMB lensing maps at location of the clusters. and later used to probe cosmology through their mass function [13–15]. Additionally, CMB lensing maps can also be used to predict and subtract the lensing-generated B-mode signal of CMB polarization to enhance constraints on inflationary physics achievable through precise measurements of the primordial B-mode signal on large angular scales [16, 17] Given the importance and diversity of the science case connected to CMB lensing, it is crucial to understand all the properties and shortcomings of statistical estimators employed to reconstruct maps of the CMB lensing potential from the observed maps of the CMB anisotropies. The most commonly employed technique used for this purpose is the so-called quadratic estimator (QE) [18]. This uses the breaking of statistical isotropy introduced by the projected matter potential field into the observed CMB to reconstruct the lensing modes through weighted couplings between pairs of harmonic modes of the observed CMB itself. Such technique has been used to derive the most sensitive measurements in the field so far from the Planck satellite data or ground-based CMB polarization experiments such as ACTpol, SPTpol, Polarbear and BICEP [19–25].

Extensive effort has been recently carried out to evaluate the sensitivity to the QE to noise anisotropies, instrumental systematics, galactic and extragalactic foregrounds, and suitable modification have been proposed to minimize their impact in the final reconstructed lensing map and power spectrum for future experiments [25–31]. But for very deep surveys such as CMB-S4, the QE is expected to be suboptimal since it only accounts for the lensing coupling at linear order. Several additional methods that account for the full lensing information have been proposed. These advocate either sampling techniques [32–34], iterative maximum likelihood [35] or Maximum a-Posteriori (MAP) estimates [36] to reach the lowest possible reconstruction noise for future measurements of the lensing potential and of its power spectrum. MAP estimators in particular offer the advantage of a reduced computational cost compared to sampling or iterative spectrum reconstruction methods and have so far been employed successfully on data covering large sky fraction. Very deep polarization observations of Polarbear were used to validate for the first time the performances of MAP estimates for joint lensing reconstruction and delensing analyses [37]. Similar results have then been achieved by sampling methods on the SPTpol data [38]. Recent work also showed how to achieve unbiased measurements of the CMB lensing spectrum from MAP estimates accounting for residual noise, normalization biases and mean field effects induced by noise anisotropies and masking, even in presence of mischaracterization of the data statistical properties [39, 40], so that it is now possible to start investigating biases induced by foregrounds and theoretical assumptions employed in the design of this new class of lensing estimator. In this work we focus our attention in particular on the assumption of the Gaussianity of the CMB lensing field for iterative MAP CMB lensing reconstruction. As shown in previous work of [41–43], non-Gaussian effects induced by the non-linear evolution of the LSS and post-Born lensing corrections due to multiple photon deflections [44] can bias the QE reconstruction ($N_L^{(3/2)}$ bias) and in turn affect the constraints achievable on cosmological parameters, in particular on the total mass of neutrinos [42]. Such bias becomes more important when correlating CMB lensing with external LSS tracers [45].

In Sec. II we review the properties of MAP estimator and $N_L^{(3/2)}$ for the QE. In Sec. IV we measure the $N_L^{(3/2)}$ bias in the MAP estimator, comparing it to the QE results and assess its importance for future surveys. In Sec. VI B we evaluate the impact $N_L^{(3/2)}$ on cosmological parameters and cross-correlation science and propose mitigation strategies.

In the following we use the standard convention and denote CMB lensing multipoles (L, M) and CMB multipoles with (ℓ, m). We will also mention in the discussion the CMB lensing convergence κ, defined in function of the CMB lensing potential ϕ in real and harmonic domains as

$$\kappa(\hat{n}) = -\frac{\nabla^2}{2}\phi(\hat{n}) \tag{1}$$

$$\kappa_{LM} = \frac{1}{2}L(L+1)\phi_{LM} \tag{2}$$

II. LENSING RECONSTRUCTION AND NON-GAUSSIAN DEFLECTIONS

In this section, we provide background information on the lensing reconstruction methods we employ and offer a brief recap of the origin of the non-Gaussianity induced bias in the CMB lensing estimates.

First, we summarize the Maximum A Posteriori (MAP) CMB lensing estimator. For more details, we refer the reader to [46, 47] for the specific implementation we use.

A. Lensing reconstruction from the MAP estimator

To obtain the MAP estimator [46, 48, 49] we will start by modelling the observed CMB as

$$X^{\rm dat} = \mathcal{B}\mathcal{D}_{\boldsymbol{\alpha}} X + n, \tag{3}$$

where $\mathcal{D}_{\boldsymbol{\alpha}}$ is the lensing operator mapping the primordial CMB X into the lensed CMB through the deflection $\boldsymbol{\alpha}$ [48], $\mathcal{B}$ a linear response matrix that includes the beam, and n the noise, that we assume to be uncorrelated with the CMB; in this work we will ignore foregrounds, and consider only experimental noise.

The goal of the MAP estimator is to derive a maximum a posterior estimate of the CMB lensing field, given the data, by maximising the posterior

$$p(\boldsymbol{\alpha}|X^{\rm dat}) \propto p(X^{\rm dat}|\boldsymbol{\alpha})p_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}) \ , \tag{4}$$

given the likelihood $p(X^{\rm dat}|\boldsymbol{\alpha})$ and a prior $p_{\boldsymbol{\alpha}}(\boldsymbol{\alpha})$. For simplicity, due to the monotony of the natural logarithm, from now own will take the natural logarithm of probabilities, and so maximing it will be equivalent to maximizing the original quantities.

We will assume a Gaussian log-likelihood for the data [46]

$$\ln p(X^{\rm dat}|\boldsymbol{\alpha}) = -\frac{1}{2}X^{\rm dat}\cdot{\rm Cov}_{\boldsymbol{\alpha}}^{-1}X^{\rm dat} - \frac{1}{2}\det {\rm Cov}_{\boldsymbol{\alpha}} + {\rm const} \tag{5}$$

with a covariance

$${\rm Cov}_{\boldsymbol{\alpha}} = \langle X^{\rm dat}X^{{\rm dat},\dagger}\rangle = \mathcal{B}\mathcal{D}_{\boldsymbol{\alpha}}C^{\rm unl}\mathcal{D}_{\boldsymbol{\alpha}}^\dagger\mathcal{B}^\dagger + N \ , \tag{6}$$

where $C^{\rm unl}$ and N are the covariance of the unlensed CMB and that of the noise respectively, and we explicitly specify that the covariance has a structure that depends on the realization of the deflection angle $\boldsymbol{\alpha}$.

For now we will ignore the curl of the deflection field $\boldsymbol{\alpha}$, and we consider only its gradient mode. We can parametrize with κ, such that the final log-posterior for κ can be written as

$$-2\ln p(\kappa|X^{\rm dat}) = X^{\rm dat}\cdot{\rm Cov}_{\boldsymbol{\alpha}}^{-1}X^{\rm dat} + \det{\rm Cov}_{\boldsymbol{\alpha}} + \sum_{LM}\frac{\kappa_{LM}\kappa_{LM}^\dagger}{C_L^{\kappa\kappa}} \tag{7}$$

up to irrelevant constants.

The CMB lensing estimator is then derived by requiring the gradient of the posterior, Eq. (7), to vanish. To solve for this, we will follow [46] and employ an iterative scheme starting from the standard quadratic estimator to finally obtain an estimate of the lensing potential field.

In particular, the calculation of the gradient of the likelihood gives a term that is quadratic in the data, which we will refer to as g^{QD}, and is given by the product of an inverse variance weighted field, and a *deflected* Wiener-filtered one [46]

$$\begin{aligned} g_{\mathrm{QD}}(\hat{n}) = -&\left(\mathcal{B}^{\dagger}\mathrm{Cov}_{\boldsymbol{\alpha}}^{-1}X^{\mathrm{dat}}\right)(\hat{n}) \\ &\cdot\left(\mathcal{D}_{\boldsymbol{\alpha}}\eth C^{\mathrm{unl}}\mathcal{D}_{\boldsymbol{\alpha}}^{\dagger}\mathcal{B}^{\dagger}\mathrm{Cov}_{\boldsymbol{\alpha}}^{-1}X^{\mathrm{dat}}\right)(\hat{n}) \ , \end{aligned} \tag{8}$$

where $\eth$ is the spin-raising operator.

The gradient of the likelihood also introduces a mean field term which, in the absence of other sources of anisotropies, represents the anisotropy introduced delensing the noise map by the estimated $\boldsymbol{\alpha}$ in the quadratic gradient [46]. This term predominantly produces dilations (convergence-like) rather than local anisotropies (shear-like) terms, hence generally very small for polarization-only estimators [47]. However, it is larger for temperature estimators, possibly contributing up to $10-20\%$ to the cross-spectrum with the input field. For further exploration of this impact, a brief discussion is presented in Appendix A, while a more detailed study is deferred to further work.

B. Lensing reconstruction from the QE estimator

The standard quadratic estimator (QE) [18] can be easily obtained from this likelihood perspective. If one writes the gradient ('$g_{\boldsymbol{\alpha}}$') to linear order in $\boldsymbol{\alpha}$, and forces it to vanish, one obtains an estimate $\hat{\boldsymbol{\alpha}}$ defined by:

$$g_{\boldsymbol{\alpha}} \approx g^{\mathrm{QD}}_{\boldsymbol{\alpha}=0} - H_{\boldsymbol{\alpha}=0}\,\hat{\boldsymbol{\alpha}} = 0. \tag{9}$$

In this equation, H (the Hessian) is minus the second derivative of the log-likelihood function. Taking instead of the realization-dependent curvature its average of data realizations, $\hat{\alpha}$ will be quadratic in the data. This average is a Fisher matrix calculated at no deflection, and is identical to the quadratic estimator response function calculated in the standard manner [18].

In this case, the unnormalized estimate $g^{\mathrm{QD}}_{\boldsymbol{\alpha}=0}$ of the CMB lensing field is given by the product of an inverse-variance-weighted field, and an undeflected Wiener-filtered one

$$g^{\mathrm{QE}}(\hat{n}) = \tag{10}$$

$$-\left(\mathcal{B}^{\dagger}\mathrm{Cov}_{\boldsymbol{\alpha}=0}^{-1}X^{\mathrm{dat}}\right)(\hat{n})\left(\eth C^{\mathrm{unl}}\mathcal{B}^{\dagger}\mathrm{Cov}_{\boldsymbol{\alpha}=0}^{-1}X^{\mathrm{dat}}\right)(\hat{n}) \ . \tag{11}$$

This is similar to the quadratic part of Eq. (8), but for the absence of all deflections: the gradient part in (8) is deflected, because the iterative estimate works by capturing the residual lensing at the unlensed position, and then remapping back to give the estimate at the observed locations [50].

C. Non-Gaussian deflections effect on the reconstruction

The autospectrum of an estimated normalized QE CMB lensing map then results in

$$C_L^{\hat{\phi}\hat{\phi}} \approx C_L^{\phi\phi} + N_L^{(0)} + N_L^{(1)} + N_L^{(3/2)} + \dots \ , \tag{12}$$

where the total estimated power spectrum $C_L^{\hat{\phi}\hat{\phi}}$ has contributions from the signal of interest itself $C_L^{\phi\phi}$, from chance Gaussian fluctuations $N^{(0)}$ [3], from additional secondary lensing contractions $N^{(1)}$ [51], and non-Gaussian contributions $N_L^{(3/2)}$. These are induced by non-zero higher order statistics in the CMB lensing potential, arising from non-linear growth of structure and post-Born lensing [44].

The study of these effects on the estimate of the CMB lensing potential has been thoroughly studied for the quadratic estimator [41–43, 45]. They mainly arise due to a non-zero response in the auto-spectrum to the presence of a bispectrum term of the lensing potential.

For next generation surveys, such as CMB-S4 [2], the $N_L^{(3/2)}$ bias, if unaccounted for, can lead to 1-2σ induced biases in cosmological parameters [42] such as the sum of neutrino masses, or have a large impact on cross-correlations with external matter tracers such as galaxy clustering or lensing [45]. [1] While methods to mitigate these effects has been proposed, such as using polarization data only (by excluding temperature information that often leads to large biases), or alternative quadratic estimators, such as bias hardened ones [45, 52], these will not be able to fully harness the statistical power of upcoming CMB data as the MAP estimator does. Indeed, for low noise and high resolution experiments, the MAP will be able to reconstruct the lensing potential to the highest significance [39, 46], and therefore it is important to assess its potential in dealing with known lensing QE biases.

In this paper, we will focus on the quality of reconstruction of the MAP estimator, specifically examining its performance in the presence of non-Gaussian contributions. Our primary focus will be on the lensing auto-spectrum, which serves as the main observable for CMB lensing analyses. We additionally present results for the lensing cross-spectrum with the input lensing potential. We plan to investigate the effect on cross-correlations for MAP with large-scale structure surveys in future studies (as it was done for the QE case [45]).

Finally, we also present simulations results demonstrating the robustness of the bias hardening technique [52] to the non-Gaussian bias.

---

[1] See Section VI B for an updated discussion on the bias induced on the sum of neutrino masses.

III. MEASUREMENT SETUP

A. Experimental setup

We consider acrudely CMB-S4 wide-field-like experimental setup [2]. The assumed observed sky fraction is 40%, and the noise is modelled for simplicity as a homogeneous and isotropic white noise, with a noise level after component separation of $N^{\mathrm{T}}_{\mathrm{lev}}$ ($N^{\mathrm{P}}_{\mathrm{lev}}$) $= 1(\sqrt{2})\mu\mathrm{K}-\mathrm{arcmin}$ for the temperature T (polarization P), and a beam modelled as a Gaussian with full width at half-maximum of 1 arcmin. Our focus is to assess the fundamental performance of the MAP to the non-Gaussianity of CMB lensing, though it is worth noting that, at the small CMB scales considered here, extra-galactic foregrounds will play an important role in temperature data.

These will still play a key role for future CMB lensing measurements like Simons Observatory [1], and are potentially the most sensitive reconstruction channel at even lower noise levels, provided scales deep enough in the damping tail of the CMB spectrum can be used, where the lensing effect is large. Figure 1 shows forecasts[2] for MAP estimators involving different data combinations in our baseline configuration, assuming CMB modes up to $\ell_{\max} = 4000$ are used in the reconstruction. The polarization data on its own can reach a nominal precision of 0.25% on the spectrum using $L \leq 1500$ with the iterative approach, and provided all sources of foregrounds and complications are under control, one can push to slightly less than 0.2%.

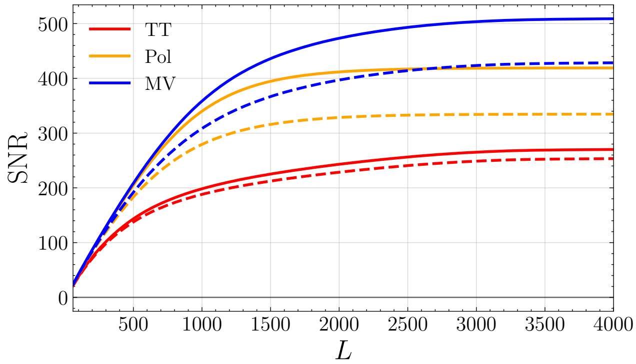

FIG. 1. Predictions of the signal to noise ratio, defined in Equation 50, on the lensing auto spectrum for different lensing estimators, for our baseline CMB-S4 wide-like configuration, with $\ell_{\max} = 4000$. Dashed is the standard QE and solid the MAP estimator. The temperature-only, polarization-only and minimum variance estimators are shown in red, orange and blue respectively.

To extract the $N^{(3/2)}_L$ bias we will consider different kind of simulations, classified based on the nature of the input convergence field. To remove unphysical effects that could come from finite number of particles in the N-body simulation used in this work, we keep input lensing modes up to an $L_{\mathrm{sim}} = 5120$. Given then an input CMB lensing map, we use `lenspyx` [53] to generate lensed CMB maps deflected with $\boldsymbol{\alpha} = \nabla\phi$.[3] Finally, we convolve the CMB simulation with the beam, and add a noise realization. The total map $X^{\mathrm{obs}} = X^{\mathrm{cmb}} + n$ will be our observed map.

We now proceed to describe the sets of deflection fields used to lens the CMB.

B. Simulated data sets

In order to isolate the relevant effects and test the impact of the assumption of the prior on κ for the lensing reconstruction, we made use of different sets of simulations, summarized below:

- $\kappa^{\mathrm{tot}} \equiv \kappa^{\mathrm{LSS+PB}}$, CMB lensing convergence simulations that include the total effect from LSS non-linearity and post-Born (PB) effects.
- κ^{LSS}, CMB lensing convergence simulations that include only the effect from LSS non-linearity.
- $\kappa^{\mathrm{X,R}}$, with $X \in \{\mathrm{tot}, \mathrm{LSS}\}$ where we randomize the phases of the CMB lensing convergence simulations that include the total effect from LSS non-linearity and PB effects, or only LSS if $X = \mathrm{LSS}$.
- κ^{G}, Gaussian CMB lensing simulations that include the exact power spectrum of the raw input convergence field $C_L^{\kappa^{\mathrm{X}}\kappa^{\mathrm{X}}}$, $X \in \{\mathrm{tot}, \mathrm{LSS}\}$. [4]

In addition, in Section III B 2 we will also use log-normal simulations of κ with a custom third moment γ, and they will be indicated by $\kappa^{\log}$, and their randomized version by $\kappa^{\log,\mathrm{R}}$.

Below we provide more details on each type of simulations.

1. *Fully non-linear observables*

To model the realistic effect of the non-linear LSS clustering of matter and the effect of post-Born corrections on

---

[2] The MAP curves were calculated using 5 iterations, with theoretical reconstruction noise calculated using the `plancklens` code.

[3] `https://github.com/carronj/lenspyx`

[4] Though we generate Gaussian simulations with CMB lensing power spectra matching the total and LSS only cases, we note that the impact of post-Born corrections on $C_L^{\kappa\kappa}$ has been proven to be negligible at the level of accuracy considered in this work, at the 0.25% level at $L = 8000$ [44, 54].Therefore one might only use only Gaussian simulations with LSS only spectra.

lensing observables we used the full-sky maps of the CMB lensing and curl potential of [54]. These were constructed using a multiple-lens plane raytracing algorithm [55] with lensing planes constructed from a ΛCDM simulation of the DEMNUni suite. This was designed to study the impact of massive neutrinos on the universe evolution and its interplay with different dark energy models [56, 57]. The ΛCDM simulation used a Planck 2013 cosmology with massless neutrinos

$$\begin{aligned}&\{\Omega_{\rm cdm},\Omega_{\rm b},\Omega_\Lambda,n_{\rm s},\sigma_8,H_0,M_\nu,\tau\} = \\ &\{0.27,0.05,0.68,0.96,0.83,67\,{\rm Km/s/Mpc},0,0.0925\},\end{aligned} \tag{13}$$

and sampled the matter distribution with 2048^3 dark matter particles in a volume of $2{\rm Gpc}/h$ between $z = 99$ and $z = 0$. The mass resolution of the simulation at $z = 0$ is $M_{\rm CDM} = 8.27 \times 10^{10} M_\odot/h$. The details of the full-sky lightcone construction from the finite volume of the N-body simulation are provided in Refs. [58, 59]. The output of the lightcone construction consists in 62 surface mass density planes $\Sigma^{\theta(k)}$ including the mass contained in spherical shells of comoving thickness $\Delta\chi \approx 150$ Mpc$/h$ that are used to construct k-th CMB lensing potential convergence plane as

$$\Delta_\Sigma^{(k)} = \Sigma^{\theta(k)}/\bar{\Sigma}^{\theta(k)} - 1. \tag{14}$$

$$\kappa^{(k)}_{\chi_{CMB}} = 4\pi G \frac{D_a(\chi_{CMB}-\chi_k)}{D_a(\chi_{CMB})}\frac{(1+z_k)}{D_a(\chi_k)}\Delta_\Sigma^{(k)} \tag{15}$$

where D_a is the angular diameter distance and χ_{CMB}, χ_k are comoving distances to the CMB and to the k-th lensing plane respectively. The simulation neglects the matter distribution at $99 < z \leq 1089$ without any effective loss in accuracy. The lensing convergence planes can then be summed together to obtain the lensing potential in Born approximation (κ^{LSS}) following [60] or used to propagate the full lensing distortion tensor beyond the Born approximation as discussed in [54] and obtain also maps of CMB $\kappa^{\rm tot}$ and lensing rotation ω which arises from coupling between subsequent lensing events. The fact that both Born and post-Born lensing maps are derived from the exact same matter distribution allows to disentangle the impact of each of the specific term as anything depending on $\kappa^{\rm tot} - \kappa^{LSS}$ will isolate the effect of post-Born corrections alone. The pipeline is general and can be adopted for lensing planes located at χ_s different from the CMB. We refer the reader to [54, 61] for more technical details of the raytracing procedure used here.

2. *Log-normal simulations*

The main relevant effect for our study is the bispectrum signal introduced in the N-body derived convergence map. For comparison, it will be useful to generate simpler maps that introduce a connected higher than two order correlation function, such as log-normal simulations. These are much cheaper to produce and better understood analytically compared to full N-body simulations, while allowing to tune the skewness of the generated map to match the value found in the full N-body simulation.

We follow the methods of [62] to produce our log-normal simulations.[5] We model the CMB lensing with a shifted lognormal field:

$$\kappa^{\log}(\hat{n}) = e^{Z(\hat{n})} - \lambda \tag{16}$$

where Z is a Gaussian random field with mean μ and variance σ^2, and λ a shift parameter. These values are calculated in such a way to match a desired power spectrum and skewness calculated from a band-limited map. We relate the moments of the desired map $\mu_{\kappa^{\log}}$, $\sigma^2_{\kappa^{\log}}$, and $\gamma_{\kappa^{\log}}$ to the parameters μ, σ, and λ to generate the log-normal simulation (see e.g. [62])

To obtain the desired log-normal field, we first calculate the correlation function of the convergence field, and obtain from it the correlation function of the Gaussian field Z according to [62]

$$\xi_Z = \ln\left(\frac{\xi^\kappa}{\alpha^2} + 1\right) \tag{17}$$

where $\alpha = \mu_{\kappa^{\log}} + \lambda > 0$. We then generate the zero mean Gaussian field $Z - \mu$, from the power spectrum C_l^Z, obtained from ξ_Z,[6] using `Healpy` [63, 64].[7] Finally, we go back to real space and obtain

$$\kappa_{\log}(\hat{n}) = e^{Z(\hat{n})} - \lambda, \tag{18}$$

where $e^\mu = (\mu_{\kappa^{\log}} + \lambda)e^{-\sigma^2/2}$, with $\sigma^2 = \xi_Z(0)$ is the variance of the Gaussian field Z. The generated maps are saved at the same resolution NSIDE = 4096 of the N-body convergence simulation, and are then processed with the same pipeline. In Figure 2 we show on the top panel an histogram of the non-Gaussian input and log-normal maps, and on the lower panel the third moment for the different simulations used in this work. As we will discuss in detail later in the text, the bulk of non-Gaussian biases is related, to leading order, to the bispectrum of the lensing field and recent work has shown that shifted-lognormal simulations like the ones adopted here can provide a moderately accurate approximation to weak lensing fields' higher-order statistics [65].

---

[5] As implemented in `https://github.com/Saladino93/fieldgen`.

[6] We use the package `https://cltools.readthedocs.io/flt/` to perform Legendre transforms to quickly switch from/to correlation functions/angular power spectra.

[7] `healpy.readthedocs.io/`

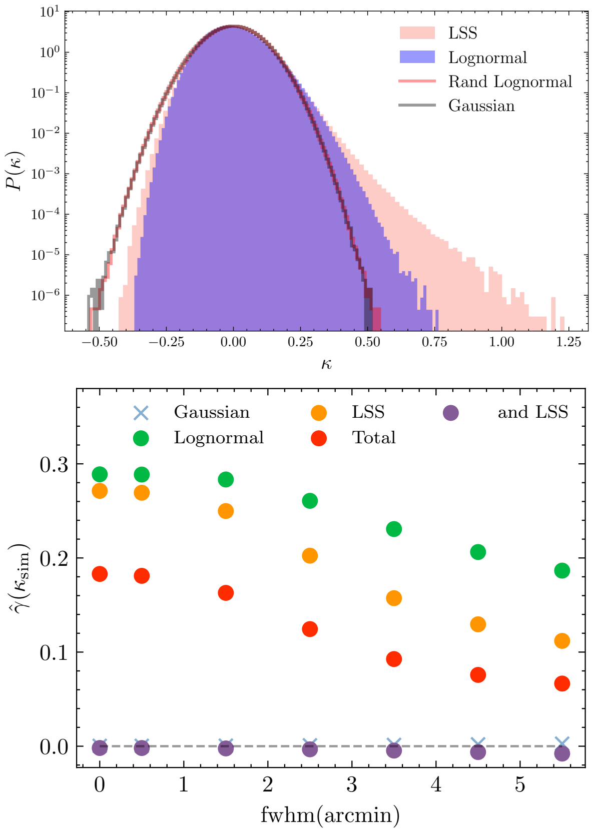

FIG. 2. *Upper panel*: binned histogram of several input convergence maps. We can see that the histograms from the LSS and lognormal convergence maps are slightly skewed compared to the Gaussian and randomized ones. We do not plot an histogram of a convergence map including LSS+Post Born effects, though this will have a reduced skewness compared to the LSS only one [54]. *Lower panel*: The skewness for different input convergence maps in function of the FWHM of a Gaussian smoothing. The lognormal case shown here is built to match the log-normal parameter λ of the LSS map at fwhm $= 0$. The values are calculated considering lensing modes up to $L_{\rm sim} = 5120$ where the CMB lensing modes used to lens the CMB simulations are cut at.

C. Lensing potential reconstruction setup

Unless indicated otherwise, we reconstruct the lensing potentials for $L \in (2, 5120)$ from the generated simulations, using CMB modes in the range $\ell_{\rm T,CMB} \in (10, 4000)$ for T, E and B.

We use `delensalot` [47][8] for the MAP or QE reconstruction of the lensing potentials (the QE reconstruction of `delensalot` follows `plancklens`). Unless stated otherwise, throughout the paper the presented MAP results performed at least 5 iterations, where we find converged results, in the sense of negligible changes in the derived spectra of the reconstructed fields.[9]

For the MAP estimates, reconstruction of lensing potential and the Wiener filtered CMBs are attempted up to an $\ell_{\rm unl,max} = 5120$ at each step. The step size used in the Netwon-Raphson method to updated the iterative estimates is $\lambda = 0.5$, and the tolerance to check for convergence of the conjugate gradient method used for the inverse variance operation is set to `cg_tol` $= 10^{-7}$, running 7-10 iterations per reconstruction.

We renormalize the reconstructed maps as follows: we do not employ analytical expressions, but use reconstructions based on observed CMB simulations lensed with the Gaussian CMB lensing simulations set ($\kappa^{\rm G}$), and we calculate a normalization factor given by

$$\mathcal{W}_L^{\rm emp,XY} = \frac{\langle C_L^{\hat{g}^{XY}\kappa} \rangle_{\rm sims}}{\langle C_L^{\kappa\kappa} \rangle_{\rm sims}} \,,$$

where $\hat{g}^{XY}$ represents an unnormalized CMB lensing potential reconstruction, and κ the corresponding input used to lens the CMB simulation.[10]

We then use this isotropic factor to normalize the lensing potential estimates from all simulation sets (κ^X, $\kappa^{\rm X,R} \mid X \in \{{\rm LSS}, {\rm T}, \log\}$)

$$\hat{\kappa}_{LM}^{XY} = \frac{\hat{g}_{LM}^{XY}}{\mathcal{W}_L^{\rm emp,XY}} \tag{19}$$

from which we calculate the raw power spectrum

$$C_L^{\hat{\kappa}^{XY}\hat{\kappa}^{XY}} = \frac{1}{2L+1} \sum_M \hat{\kappa}_{LM} \hat{\kappa}_{LM}^{\dagger} \,. \tag{20}$$

Finally, to assess the quality of the reconstructions, it is useful to define the cross correlation coefficient ρ_L [11].

$$\rho_L = \frac{C_L^{\hat{\kappa}^{XY}\kappa}}{\sqrt{C_L^{\hat{\kappa}^{XY}\hat{\kappa}^{XY}} C_L^{\kappa\kappa}}} \,,$$

---

[8] `https://github.com/NextGenCMB/delensalot`

[9] In the case of temperature, convergence of the results can be reached earlier though we still keep a larger number of iterations.

[10] We checked that for the QE, we get consistent results both with this method and by using a response function given by CMB gradient spectra, following [45].

[11] By 'quality' of the reconstruction we mean here its information content on the true lensing potential, irrespective of potential biases (the cross-correlation coefficient is invariant under arbitrary multiplicative biases in the reconstructed field (small additive ones can also be cast as multiplicative biases))

IV. $N_L^{(3/2)}$ BIAS

With the notation now established, we are ready to calculate the biases associated with the presence of non-Gaussianity in our CMB lensing fields. The $N_L^{(3/2)}$ biases are determined through simulations, specifically by cross-correlating the estimated CMB lensing field with another tracer, such as the input CMB lensing potential itself, or by examining the CMB lensing auto-correlation. This methodology, as outlined in ([42, 45]) allows us to quantitatively assess and account for the impact of non-Gaussian features in our lensing fields. This is achieved by calculating the difference from the scenario where no non-Gaussianity is present.

The biases, denoted as $N_L^{(3/2)}$, are computed through the following expressions:

$$\begin{aligned}
\text{Total}: N_L^{(3/2)} &= \langle \hat{C}_L^{\hat{\phi}^{XY}\phi^{\rm ext}}[\kappa^{\rm tot}] - \hat{C}_L^{\hat{\phi}^{XY}\phi^{\rm ext}}[\kappa^{\rm tot,R}]\rangle_{\rm CMB} \\
\text{LSS}: N_L^{(3/2)} &= \langle \hat{C}_L^{\hat{\phi}^{XY}\phi^{\rm ext}}[\kappa^{\rm LSS}] - \hat{C}_L^{\hat{\phi}^{XY}\phi^{\rm ext}}[\kappa^{\rm LSS,R}]\rangle_{\rm CMB} \\
\text{PB}: N_L^{(3/2)} &= \langle \hat{C}_L^{\hat{\phi}^{XY}\phi^{\rm ext}}[\kappa^{\rm tot}] - \hat{C}_L^{\hat{\phi}^{XY}\phi^{\rm ext}}[\kappa^{\rm LSS}]\rangle_{\rm CMB} \,.
\end{aligned} \tag{21}$$

We denote $\phi^{\rm ext}$ either the input lensing field or the reconstructed field $\hat{\phi}^{XY}$ itself. The angle brackets signify the average over lensed CMB simulations. To gauge the reliability of our measurements, we will compute the scatter in the biases from these simulations to obtain the uncertainty on the mean measurement. Our subtraction method ensures that simulations which are differentiated in these equations share identical primordial CMB and noise, mitigating realization-dependent biases and cosmic variance effects.

In Figures 3 and 4 we show the $N_L^{(3/2)}$ bias for the CMB lensing autospectrum and cross-spectrum with the input, respectively. The shaded areas are computed from the scatters on the mean of the simulations, while grey area represent statistical error bars as calculated from the diagonal of the covariance matrix

$$\mathbf{C}_{LL'} = \frac{2\delta_{LL'}}{(2L+1)f_{\rm sky}}\left(C_L^{\phi\phi} + N_L^{(0)} + N_L^{(1)}\right)^2 \,. \tag{22}$$

In Figure 3, it is evident that, for the lensing autospectrum, the $N_L^{(3/2)}$ bias of the MAP estimator shows a lower absolute value compared to the QE estimator for $L \leq 1800$. Specifically, at the scales presented, the LSS effect in the QE tends to suppress power in the reconstructed potential. This effect is alleviated in the MAP for all estimators except on the smallest scales. The situation is similar, but with the opposite sign, for the PB bispectrum effect. This results in partial cancellation between the two effects for the total contribution for both the QE and MAP. Similar conclusions are achieved with the cross-spectrum with the input lensing potential as depicted in Figure 4.

Both MAP and QE estimators share the characteristic that the total bias effect depends on the combination of used data. Specifically, in the minimum variance combination, the polarization estimator dominates at lower modes $L \leq 1500$, while the temperature one dominates at higher L's.

Finally, a distinct feature is the noticeable steep rise in the bias of the MAP estimator for the LSS cases, particularly in the TT and MV combinations, evident in both the auto and cross spectra. While the specific origin of this rise remains uncertain, we validated our bias method by assessing Gaussian simulations with no LSS bias and randomized ones. We checked that the difference is consistent with zero. Therefore this rise at small scales in the MAP might be due non-Gaussanities, or from the choice of normalisation of the maps (obtained from Gaussian simulations). To investigate the choice of normalisation, we consider the cross-correlation coefficient of the reconstructions with their respective inputs

$$\rho^X = \frac{\langle C_L^{\hat{\kappa}\kappa}\rangle}{\sqrt{\langle C_L^{\hat{\kappa}\hat{\kappa}}\rangle\langle C_L^{\kappa\kappa}\rangle}} \tag{23}$$

where $\langle . \rangle$ denotes the average over sims. This quantity does not require any normalisation, and it is related to the faithfulness of our reconstruction to the input. We assume that we can directly take the difference of the cross-correlation coefficient with the corresponding input for the LSS non-Gaussian MAP vs the Gaussian MAP estimates, as shown in Figure 5.

As expected, we can see that on large scales the Gaussian case reconstruction is better correlated with the input, for both the QE and MAP cross-correlation differences. While on small scales, $L > 1500$, interestingly we see a small rise in the difference with the cross-correlation coefficient of the MAP of the non-Gaussian case. This might be related to additive terms that we test in Section IV B 2.

A. Simple theory calculations

Given the complexity of beyond-the-QE reconstruction, a comprehensive analytic treatment of the theory curves for the MAP-$N_L^{(3/2)}$ bias seems out of reach. Nevertheless, we can proceed by analogy with quadratic estimator theory to obtain some curves that matches reasonably well our findings.

We will focus on the simpler case of the cross-

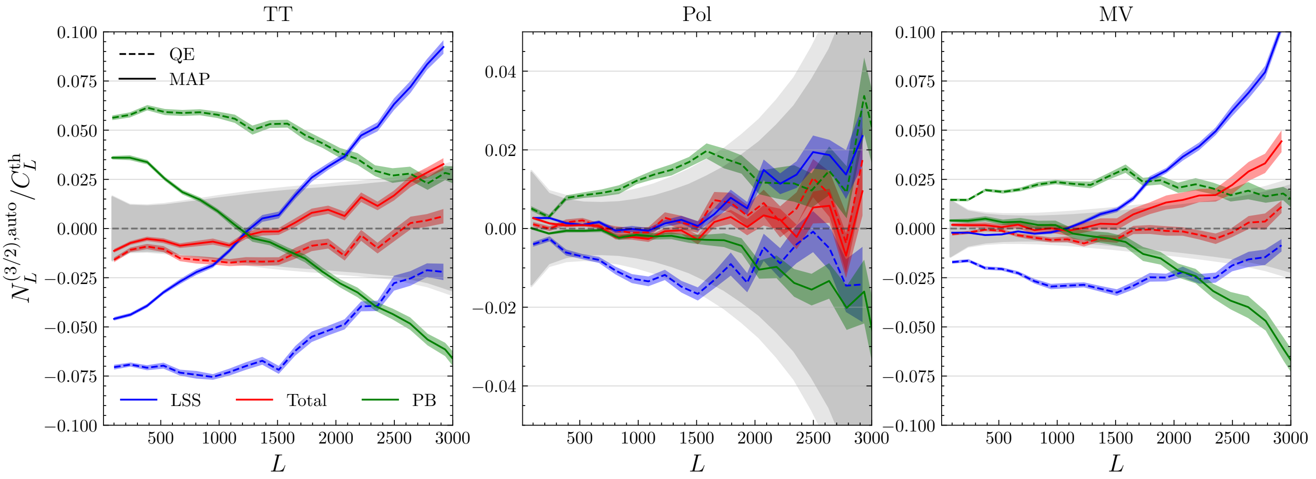

FIG. 3. Fractional $N_L^{(3/2)}$ bias on the CMB lensing auto-correlation for a CMB-S4 like configuration, as described in the text. This is calculated by averaging over 64 realization of the input CMB primordial field. We show results for the LSS only, post-Born (PB) and full cases in blue, green and red, respectively. In grey and light grey we show the statistical error bars for MAP and QE respectively. We bin the spectra from $L_{\min} = 30$ to $L_{\max} = 3000$ with a wide binning of around 140.

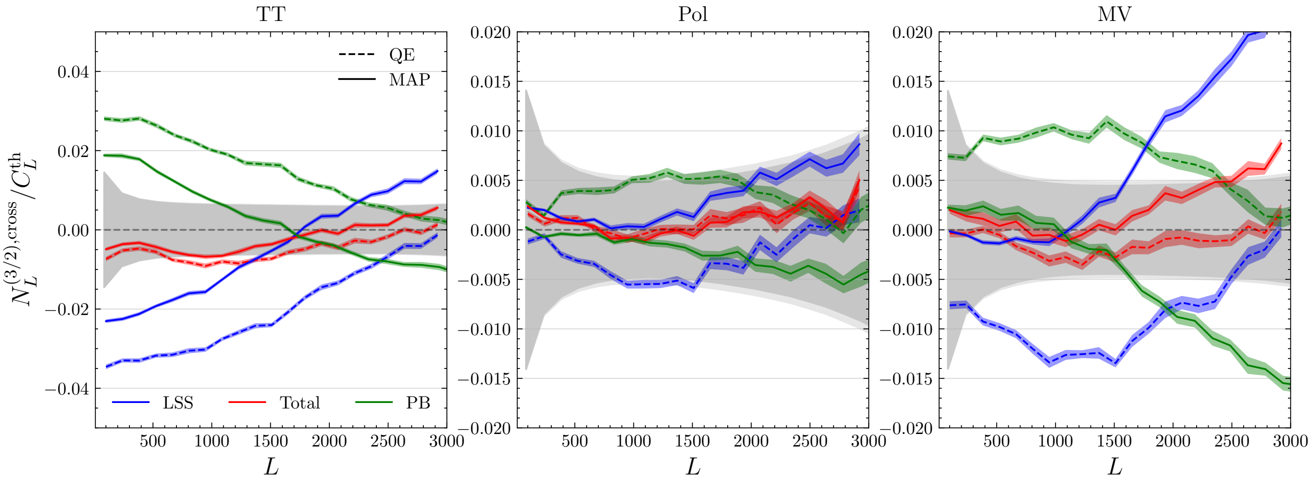

FIG. 4. Same as Figure 3 but for the cross-spectrum between the reconstruction and the input lensing potential.

correlation of the reconstructed CMB lensing potential $\hat{\phi}$ with the input lensing potential ϕ on the flat-sky:[12]

$$N_L^{(3/2),\mathrm{cross}} = \langle \hat{\phi}(\vec{L})\phi^*(\vec{L})\rangle - \langle \phi(\vec{L})\phi^*(\vec{L})\rangle, \tag{24}$$

where we assume the reconstruction $\hat{\phi}(\vec{L})$ is unbiased, in the sense that $\langle\hat{\phi}\rangle = \phi$. The difference between these two terms depends predominantly on the bispectrum of the observed lensing potential $B^{\phi\phi\phi}(L, l_2, l_3)$, where L stands for the external tracer (here ϕ itself) multipole and l's for the ones entering through the CMB modes used for CMB lensing reconstruction.

For the standard QE CMB lensing temperature-only estimator, calculations are given in [41, 45]. We briefly review the rationale here.

A temperature-based quadratic estimator is of the form

$$\hat{\phi}(\vec{L}) = A_L^{TT} \sum_{\vec{l}} g(\vec{l}, \vec{L}) T^{\mathrm{dat}}(\vec{L} - \vec{l}) T^{\mathrm{dat}}(\vec{l}). \tag{25}$$

for some normalization A_L^{TT} and weights $g(\vec{l}, \vec{L})$.

---

[12] We focus on the cross-correlation with the input, and not on the reconstruction auto-spectrum, as the former is easier to predict with respect to the latter. In particular, we perform all of our studies using CMB modes up to $l_{\max} = 4000$. Our baseline perturbative analytical model for prediciting the QE auto-spectrum, based on [41, 45], works well only on the largest scales, where we get the prediction $N^{(3/2),\mathrm{auto}} \sim 2N^{(3/2),\mathrm{cross}}$. This will require a more careful treatment for future studies.

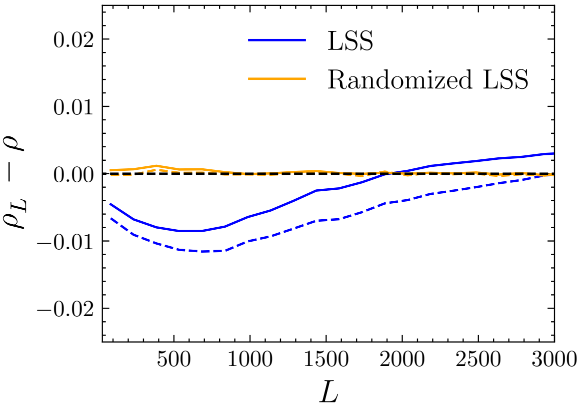

FIG. 5. Difference in the cross-correlation coefficients between the LSS non-Gaussian (blue), randomized LSS non-Gaussian (orange) cases with the Gaussian case, for the TT estimator. In solid (dashed) we have the MAP (QE). We can see that the difference with randomized (orange) is consistent with zero. While as expected, the non-Gaussian case is below zero on large scales. Nevertheless, for the non-Gaussian MAP (blue solid line), we can see a small rise after $L > 1500$.

The lensed CMB may be perturbatively expanded in a series with respect to the lensing potential, leading to

$$T = T^u + \delta T^u + \delta^2 T^u + \mathcal{O}(\phi^3) \ , \tag{26}$$

where T^u is the unlensed CMB, and $\delta^n T^u$ depends on the lensing potential power n [3].

On evaluating the cross-correlation (24),

$$\langle \hat{\phi}(\vec{L})\phi(-\vec{L})\rangle \sim \langle T(\vec{L}-\vec{l})T(\vec{l})\phi(-\vec{L})\rangle \ , \tag{27}$$

the lensing potential bispectrum will give rise to the $N_L^{(3/2),\mathrm{cross}}$,

$$N^{(3/2),\mathrm{cross}} \sim \langle \delta T^u \delta T^u \phi\rangle + 2\langle T^u \delta^2 T^u \phi\rangle \ . \tag{28}$$

The full expression for the temperature case, including to a good approximation non-perturbative lensing remapping effects [45], is

$$\begin{aligned} N_L^{(3/2),\mathrm{cross}} &= A_L^{TT}\int_{\vec{l}_1} B^{\phi\phi\phi}(L, l_1, |\vec{L}-\vec{l}_1|)\int_{\vec{l}_2} g(\vec{l}_2,\vec{L}) \\ &\Big(-C^{T\nabla T}_{|\vec{l}_1-\vec{l}_2|}[(\vec{l}_1-\vec{l}_2)\cdot\vec{l}_1][(\vec{l}_1-\vec{l}_2)\cdot(\vec{L}-\vec{l}_1)] \\ &\qquad + C^{T\nabla T}_{l_2}[\vec{l}_2\cdot\vec{l}_1][\vec{l}_2\cdot(\vec{L}-\vec{l}_1)]\Big) \ . \end{aligned} \tag{29}$$

The spectra $C_l^{T\nabla T}$ are here the same lensed gradient spectra that enters the non-perturbative lensing response functions.

Let's discuss how we connect the MAP bias to this representation. Reference [39] demonstrated through simulations that the converged CMB lensing MAP solution power spectrum (for polarization at least) can be accurately described as a quadratic estimator with partially lensed CMB spectra. These spectra can be obtained using an iterative scheme initially proposed for the EB estimator by [17]. Hence, it is natural to test this recipe for $N_L^{3/2}$ as well.

In this picture, the lensing potential entering the CMB legs are now given by the unresolved, residual lensing map

$$\hat{\phi}^{\mathrm{res}}_{LM} \equiv (1-\mathcal{W}_L)\phi_{LM} \ . \tag{30}$$

We then make use of the same flat-sky prediction Eq (29), but with the substitution

$$B^{\phi\phi\phi}(L, l_1, l_3) \to B^{\phi\phi\phi}(L, l_1, l_3)(1-\mathcal{W}_{l_1})(1-\mathcal{W}_{l_3}) .$$

where $\vec{l}_3 = \vec{L} - \vec{l}_1$, and $\mathcal{W}$ is the Wiener-filter. The normalization A_L^{TT} is the standard quadratic estimator normalization but calculated with partially lensed gradient spectra, and $C_l^{T\nabla T}$ are also calculated with the partially lensed spectra.

In Figure 6 we show the result of this naive calculation, compared to simulations.[13]. For simplicity we consider the case of temperature only. On the upper panel we show the LSS-only case, where $N^{3/2}$ is stronger than with the total bispectrum, shown in the lower panel. It assumes a LSS bispectrum calculated at the tree level in perturbation theory with corrections coming from [66] (though an improved version can be found in [67]). The Post-Born corrections are based on [44]. See the Appendix of [45] for a short review.

We can see that on a wide range of scales this simple predictive scheme is able to recover the estimated biases from simulations, validating the understanding that the bispectrum of the CMB lensing field correctly describe the bias also in the case of the iterative estimator. In the LSS-only case, and on small scales, the prediction deviates significantly from our findings in simulations. In Section IV B 2 we check that its origin is not from higher even CMB lensing connected n-point functions.

B. Consistency and robustness tests

To have a better understanding of our reconstructions and validate our pipeline, we perform a series of consistency checks to explore possible residuals in our estimates, specifically those of higher order than the bispectrum residuals.

1. *Joint reconstruction with lensing field rotation*

In this section we test the impact of post-Born lensing rotation on the MAP reconstructions.

---

[13] The code to calculate these biases is based on `https://github.com/Saladino93/lensbiases`.

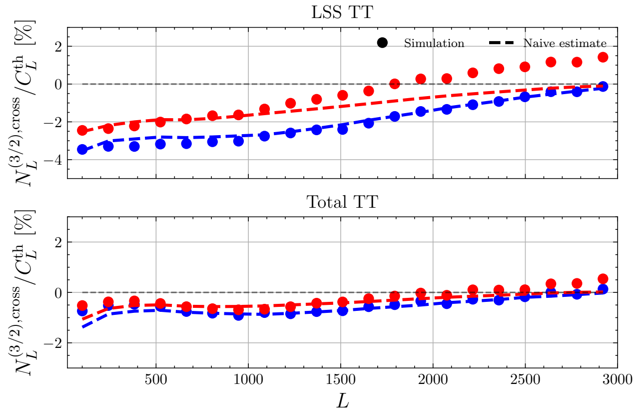

FIG. 6. Approximate theory calculation for the $N_L^{(3/2)}$ bias on the CMB lensing cross-spectrum. In blue, we show results for the quadratic estimator, and in red for the MAP reconstruction. The MAP predictions are obtained from the naive analytic prescription given in the text. This is for our CMB-S4-like reconstruction from temperature.

The CMB lensing deflection vector field $\vec{\alpha}(\hat{n})$ can be written thanks to the Helmholtz decomposition in flat-sky notation as

$$\vec{\alpha}(\hat{n}) = \vec{\nabla}\phi(\hat{n}) + \vec{\nabla} \times \Omega(\hat{n}) \ , \tag{31}$$

On the full-sky, the decomposition is compactly written using the spin-weight formalism,

$${}_1\alpha(\hat{n}) = -\eth\phi(\hat{n}) - i\eth\Omega(\hat{n}), \tag{32}$$

where Ω is the lensing curl potential, and

$$\begin{aligned} \omega(\hat{n}) &= -\frac{\nabla^2\Omega(\hat{n})}{2} \\ \omega_{LM} &= \frac{1}{2}L(L+1)\Omega_{LM} \end{aligned} \tag{33}$$

is the lensing field rotation; the angle by which tiny local images are rotated by lensing. The leading rotation is induced by post-Born lensing, that couples pair of non-aligned shearing lenses at different redshifts, producing in this way a net rotation. For a single deflection or in the Born approximation the rotation is null. Non-zero bispectra of the kind $\kappa\kappa\omega$ or $\kappa\omega\omega$ [44, 54] are also generate at higher perturbative orders. The amplitude of the rotation field power spectrum is about 3 to 4 orders of magnitude smaller than the one of κ [35, 44] on the scales considered here.

The simulations of [45, 54] naturally include the rotation component of CMB lensing. We use this field in combination with the post-Born convergence map to estimate the lensing potentials in the presence of a curl-like displacement using either a joint ϕ-Ω optimal reconstruction (the differences to the standard ϕ-only algorithm are minors and described in [47]), or completely neglecting Ω in the reconstruction.

We can then estimate the effects in the $N_L^{(3/2)}$ bias coming from the mixed-bispectrum terms $\kappa\kappa\omega$, $\kappa\omega\omega$ as :

$$\text{Mixed} : N_L^{(3/2)} = \langle \hat{C}_L^{\hat{\phi}^{XY}\phi^{\rm ext}}[\kappa^{\rm tot},\omega] - \hat{C}_L^{\hat{\phi}^{XY}\phi^{\rm ext}}[\kappa^{\rm tot}]\rangle_{\rm CMB}, \tag{34}$$

where the first term include the lensing rotation in the inputs and reconstruction, and the second does not. As for the biases obtained in previous sections, here we use 64 CMB primordial realizations, and we run the pipeline up to 10 iterations, though outputs appeared to have converged starting from the fifth one.

We show in Fig. 7 the $N_L^{(3/2)}$ biases to the auto and cross-spectra in the presence of a full deflection field (including LSS and post-Born effects in the gradient-like deflection component as well as curl-like deflections), compared to a case without curl-like displacement component in the simulations. We see no statistically significant differences.

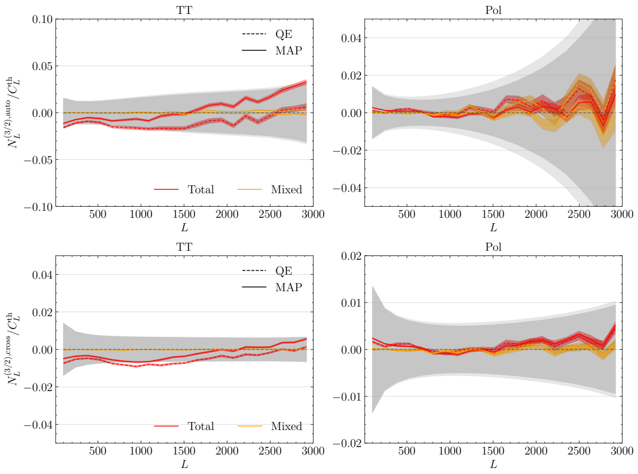

FIG. 7. Non-Gaussian $N_L^{(3/2)}$ lensing biases in the estimated lensing potential auto-spectra (top panels) or cross-spectrum to the true lensing (bottom panels), with and without curl-like deflections in the simulations. In red we show the total bias that does not include post-Born lensing field rotation (same as Figure 3), while in orange we show the difference between the biases obtained including post-Born corrections for both the gradient and curl component of the deflections and the red curve. The difference is consistent with zero in both panels.

2. *Sign-flipped κ*

At the perturbative level, the $N_L^{(3/2)}$ bias is proportional to a projected bispectrum on large scales [41]. We will test this assumption in this section.

Flipping the sign of the lensing potential should flip the sign of the bias, if it depends the bispectrum that is odd in the lensing field. For the QE, this argument breaks down only when using the smallest CMB scales for our reconstructions, where perturbative arguments are less effective (e.g., at $\ell_{\max} \sim 4000$ [41]). On the other hand, the MAP always uses the full likelihood information and, therefore, higher-order point functions to reconstruct the lensing field. This also makes the MAP more difficult to assess analytically compared to the QE. Therefore, we turn to simulations to check if the MAP bias behaves significantly differently with respect to the sign of the input lensing map.

For each of the convergence maps ($\kappa^{\rm G}$, $\kappa^{\rm LSS}$, $\kappa^{\rm tot}$) we calculate their flipped version

$$\kappa^{\rm in,\neg}_{LM} = (-1) \cdot \kappa^{\rm in}_{LM}\,, \tag{35}$$

where $\kappa^{\rm in}_{LM}$ is the reference input CMB convergence lensing field. We generate then 64 lensed CMB simulations with this flipped lensing potential, and calculate the half-sum and half-differences

$$\Delta^{XY}_{\pm} = \left\langle \frac{1}{2}\left(N^{(3/2)}_L[\kappa^{\rm in}_{LM}] \pm N^{(3/2)}_L[\kappa^{\rm in,\neg}_{LM}] \right) \right\rangle_{\rm sims}\,. \tag{36}$$

for the TT, EE, $EE+EB$, and MV estimators, and where the $N^{(3/2)}_L$ are estimated as explained in Section IV.

Any contribution from contractions which are even in the lensing field will appear in the half-sum of the biases, and similarly for odd terms in the half-differences. Figure 8 shows these half-sum in blue and half-differences in orange, for the cross LSS-only reconstruction bias from temperature. We can see that for MAP and QE reconstructions we have a small residual in the half-sum ($\sim 0.2\%$). In particular, the rise at the highest L observed in the previous section is still most likely to an odd n-point function effect.

To quantify the consistency of these results with a signal dominated by the bispectrum, we calculate the probability-to-exceed (PTE) using Welch's test. This test compares the means of two simulation sets (κ, sign-flipped κ) with unequal variances, calculated from simulation scatters. The input data vector is the average over simulation measurements for 21 multipole bins over the range $5 \leq L \leq 3000$. The results are shown in Table I, where the average PTEs over the 21 multipoles are shown. A test is assumed successful if the PTE is greater than 0.05. With this definition, the QE MV, EB cross-spectrum for PB cases fail, though this is acceptable given the number of tests performed.

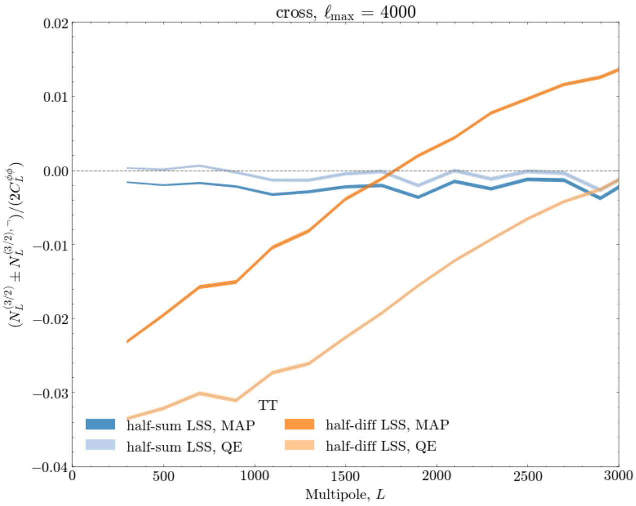

FIG. 8. Half-sum and half-differences of $N^{(3/2)}_L$ biases, built from a baseline simulation, and a second one flipping the sign of the input lensing potential, in order to isolate the contributions of even(blue) and odd(orange) n-point functions on the total bias. The light (dark) colors are the QE (MAP) reconstructions for $\ell_{\max} = 4000$, from temperature-only and LSS bispectrum only.

V. ALTERNATIVE CMB LENSING ESTIMATORS

A. Iterative estimator with a non-Gaussian prior

The conventional iterative estimator for CMB lensing assumes a Gaussian potential field due to its incorporation in the prior distribution. However, the observed CMB lensing field exhibits non-Gaussian characteristics that the standard iterative approach does not directly address. In this context, our objective is to accommodate potential non-Gaussian aspects of the field by proposing an alternative method for estimating ϕ.

To begin with, we introduce a non-Gaussian prior. Specifically, we characterize the lensing convergence field using a log-normal model, as outlined in section III B 2. This choice establishes the foundation for our new approach and is motivated by the fact that the log-normal approximation is a sensible one for moderately non-linear fields, e.g. [68], [69].

The basic idea is that instead of iteratively solving for the CMB lensing potential field directly, we iterate over a posterior that is a function of the Gaussian field whose exponential gives the CMB lensing convergence field.[14]

In this case, we will assume a prior on the Gaussianized field Z, which may be written in harmonic space

$$-2\ln p_Z(Z) = \frac{\left(Z^2_{00} - \mu\sqrt{4\pi}\right)^2}{C^Z_{L=0}} + \sum_{L\geq 0, |M|\leq L} \frac{Z_{LM} Z^{\dagger}_{LM}}{C^Z_L}, \tag{37}$$

up to irrelevant constants. The un-normalized posterior can be written as before

$$p(Z|X^{\rm dat}) \propto p(X^{\rm dat}|Z)p(Z) \tag{38}$$

---

[14] This allows us to use our baseline MAP pipeline with little modifications.

Estimator	QE		MAP	
	auto	cross	auto	cross
	LSS / TOT / PB	LSS / TOT / PB	LSS / TOT / PB	LSS / TOT / PB
		$\ell_{\max} = 3000$		
TT	0.51 / 0.78 / 0.31	0.56 / 0.21 / 0.24	0.89 / 0.51 / 0.98	0.66 / 0.20 / 0.33
EE	0.42 / 0.28 / 0.95	0.44 / 0.63 / 0.56	0.32 / 0.19 / 0.86	0.65 / 0.79 / 0.72
$EE+EB$	0.11 / 0.17 / 0.69	0.07 / 0.61 / 0.04	0.58 / 0.96 / 0.25	0.10 / 0.28 / 0.10
MV	0.90 / 0.47 / 0.13	0.11 / 0.79 / 0.02	0.20 / 0.10 / 0.94	0.75 / 0.74 / 0.45
		$\ell_{\max} = 4000$		
TT	0.97 / 0.87 / 0.88	0.74 / 0.71 / 0.75	0.38 / 0.43 / 0.37	0.25 / 0.09 / 0.36
EE	0.07 / 0.06 / 0.43	0.31 / 0.28 / 0.37	0.52 / 0.29 / 0.74	0.28 / 0.27 / 0.35
$EE+EB$	0.38 / 0.99 / 0.31	0.15 / 0.44 / 0.14	0.82 / 0.88 / 0.65	0.32 / 0.59 / 0.22
MV	0.42 / 0.96 / 0.14	0.25 / 0.25 / 0.27	0.99 / 0.93 / 0.95	0.77 / 0.73 / 0.79

TABLE I. Averaged over L-bins PTEs for the half-sum tests of the $N_L^{(3/2)}$ biases. Shown are the temperature, polarization, their combination, QE and MAP estimators, for $N_L^{(3/2)}$ auto- and cross spectra.

Subsequently, our objective is to determine the optimal estimate of Z based on the available data.[15]

In this case, the total gradient with respect to the Gaussian field is

$$g_Z^{\rm tot}(\hat{n}) = \frac{\delta \ln p(Z|X^{\rm dat})}{\delta Z(\hat{n})} = g_Z^{\rm QD}(\hat{n}) - g_Z^{\rm MF}(\hat{n}) + g_Z^{\rm PR}(\hat{n}) \ , \tag{39}$$

This gradient can be calculated with minimal modifications to our ϕ-based MAP reconstruction code: using the chain rule, we have namely

$$\begin{aligned} g_Z^{\rm QD}(\hat{n}) - g_Z^{\rm MF}(\hat{n}) &= \frac{\delta \ln p(X^{\rm dat}|Z)}{\delta Z(\hat{n})} \\ &= \int d^2\hat{n}' \frac{\delta \kappa(\hat{n}')}{\delta Z(\hat{n})} \frac{\delta \ln p(X^{\rm dat}|\phi)}{\delta \kappa(\hat{n}')} \\ &= e^{Z(\hat{n})} \left(g_\kappa^{\rm QD}(\hat{n}) - g_\kappa^{\rm MF}(\hat{n}) \right) \end{aligned} \tag{40}$$

We have used in the second line

$$\frac{\delta \kappa(\hat{n}')}{\delta Z(\hat{n})} = \delta_D^{(2)}(\hat{n}' - \hat{n}) e^{Z(\hat{n})}, \tag{41}$$

which follows directly from our definition $\kappa(\hat{n}) = e^{Z(\hat{n})} - \lambda$. Finally, the gradients $g_\kappa(\hat{n})$ are easily obtained in harmonic space from those of ϕ which the original MAP reconstruction calculates: from $\kappa_{LM} = \frac{1}{2} L(L+1)\phi_{LM}$ follows for the gradients

$$g_\kappa(\hat{n}) = \int d^2n' \frac{\delta \phi(\hat{n}')}{\delta \kappa(\hat{n})} g_\phi(\hat{n}') \tag{42}$$

$$= \sum_{LM} \frac{2}{L(L+1)} g_{\phi, LM} Y_{LM}(\hat{n}). \tag{43}$$

Starting from the quadratic estimator solution we then iterate over estimates of Z in the same manner than the original code.[16]

Once the MAP point $\hat{Z}$ is reconstructed, we apply the transform to get an unnormalized estimate of the lensing field,

$$\hat{\kappa}_{\log}(\hat{n}) \equiv e^{\hat{Z}(\hat{n})} - \lambda. \tag{44}$$

which we then normalize in the same way as before, by rescaling with the inverse averaged cross-spectrum to input Gaussian fields, Eq. (19). We then reconstruct the biases in the very same way.

The approach we follow (maximizing for Z) does lead to a slightly different map than if we were maximizing directly for κ, even if using the same lognormal prior on Z. This is because the relation between the point-estimates (44) is nonlinear. Our approach is arguably more natural than maximizing for κ, and avoids one issue we encoutered in preliminary work, which is how to enforce the lognormality of κ map across iterations: when maximizing for κ, one must ensure that $\kappa(\hat{n}) + \lambda$ is always positive at each point, since $Z(\hat{n})$ is its logarithm.

In Figures 9 we focus on temperature-only reconstructions, and with the LSS non-linear part of the non-Gaussianity only, which has the strongest signature.

For each set of curves, the dashed lines show the biases found in maps with the N-body LSS kappa convergence map, and the dotted lines the biases found in our simple lognormal simulated convergence maps with the same skewness.

In blue and red we have results for the QE and baseline MAP, respectively, where the reconstruction is done

---

[15] In principle we could also attempt to reconstruct Z jointly with the parameters μ and λ. This is a complication with no relevance for what is being tested here however, and we assume fixed fiducial values of μ and λ, and fixed fiducial spectrum C_L^Z for the reconstructions.

[16] A way to see how the starting point is chosen is the following: if we take the gradient in Z, g_Z, then we require as a first step that: $g_Z \approx g_{Z_0} + F_{Z_0}(Z - Z_0) \approx 0$. By exponentiating the expression that we find for Z we can relate it to $\kappa + \lambda$, and derive that $\hat{Z} = \ln(\hat{\kappa} + \lambda)$.

as in the previous section, with a Gaussian prior on κ. Given the simplicity of the lognormal simulations, it is remarkable how well they reproduce the $N^{(3/2)}$ bias, particularly in the MAP case.

In orange, we show the curves obtained from the reconstructions using the lognormal prior as just discussed. Again, there are essentially no difference found between the N-body input (orange dashed) or lognormal input (orange dotted). Both curves shifts up slightly, reducing slightly the bias on most signal dominated scales, but remain qualitatively very similar.

The fact that the lognormal prior improves only mildly the Gaussian prior is due the fact that we are likelihood dominated, and therefore the MAP reconstruction is mainly data driven.

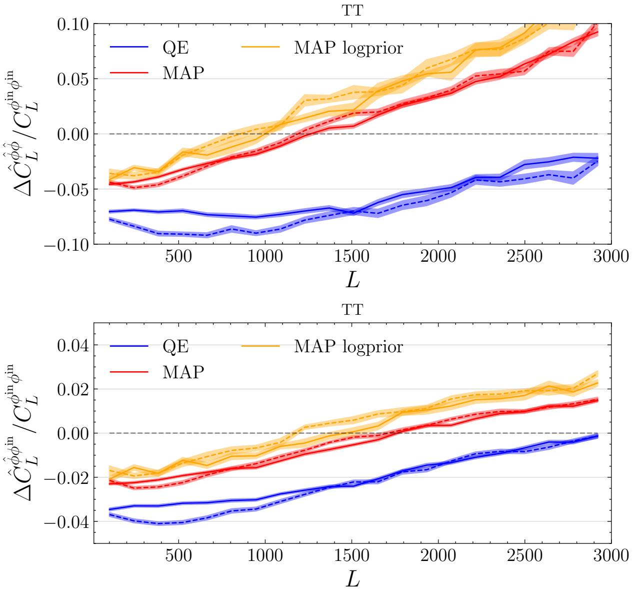

FIG. 9. Comparison of the $N_L^{(3/2)}$ bias for different variants of temperature reconstructions. The solid lines use the N-body LSS map, while the dashed ones a lognormal map with a similar skewness (at the same input resolution). We show QE results in blue, MAP results with Gaussian prior in red, and MAP results with lognormal prior in orange. We see a small shift upwards between these two choices of priors. The biases found on lognormal maps match well the ones found on the LSS input map. The top shows the bias in the auto-spectrum, and bottom for the cross-spectrum to the true lensing.

B. Bias Hardening

The small-scale CMB temperature is strongly contaminated by foregrounds, limiting the use of more modes to perform CMB lensing reconstruction. Recently, foreground-mitigating lensing reconstruction methods have been developed, allowing more robust and powerful CMB lensing measurements.

In particular, bias-hardening [52, 70, 71], deprojecting the response of the CMB lensing QE to a foregrounds QE, have been used in recent data analyses to extract cosmological parameters from observations.

At the likelihood level, we can derive the bias-hardened QE CMB lensing estimator by looking for modulations in the observed map, going beyond the CMB and noise, that could be attributed to point sources.

We imagine to model our data in pixel space as

$$X^{\rm dat} = \mathcal{B}\mathcal{D}_{\boldsymbol{\alpha}} X + \mathcal{B}S + n \ , \tag{45}$$

where S is some source field. From this equation, the pixel-pixel covariance is

$$C \equiv \langle X^d X^{d,\dagger} \rangle = \mathrm{Cov}_{\vec{\alpha},S^2} = \mathcal{B}\mathcal{D}_{\boldsymbol{\alpha}} C^{\rm unl} \mathcal{D}^\dagger_{\boldsymbol{\alpha}} \mathcal{B}^\dagger + \mathcal{B}S^2\mathcal{B}^\dagger + N \tag{46}$$

assuming no cross-terms between CMB, the source, or the noise terms. We can see that at the covariance level the source term induces a variance that is larger than the expected one from just experimental noise.

The log-likelihood becomes then

$$\mathcal{L} \equiv \ln L(X^{\rm dat}|\vec{\alpha}, S^2) = -\frac{1}{2} X^{\rm dat} \cdot \mathrm{Cov}^{-1}_{\vec{\alpha},S^2} X^{\rm dat} - \frac{1}{2} \det \mathrm{Cov}_{\vec{\alpha},S^2} \ . \tag{47}$$

To derive the bias-hardened estimator, we adopt a method akin to the standard quadratic estimator. We start by nulling the gradients for both the lensing potential ϕ and the source term variance S^2, expanding around $(\phi, S^2) = (0, 0)$ using a first-step Newton iteration:

$$\vec{G}_{\phi,S^2} \approx \vec{G}_{0,0} + F_{0,0} \begin{bmatrix} \phi \\ S^2 \end{bmatrix} \approx 0 \rightarrow \begin{bmatrix} \phi \\ S^2 \end{bmatrix} \approx F^{-1}_{0,0} \vec{G}_{0,0} \ , \tag{48}$$

where the total gradient of the log-likelihood $\vec{G} = [\frac{\partial \mathcal{L}}{\partial \phi}, \frac{\partial \mathcal{L}}{\partial S^2}]^T$ is approximated by its value at $(\phi, S^2) = (0, 0)$, $\vec{G}_{0,0}$, and a curvature matrix $F_{0,0}$. The latter, encapsulates the responses R^{ab} of the estimator for a to the presence of b, with $a, b \in \phi, S^2$.

In Figures 10, 11 we show results for the $N_L^{(3/2)}$ bias from our simulations using a bias-hardened estimator against a point source. We can see that, compared to the QE, the bias hardening estimator mitigates the impact of the non-Gaussian bias, confirming analytical expectations of [45]. This suggests a bias-hardened MAP-solution might also be helpful, which just adds a prior to Equation 47, which is work in progress.

VI. IMPACT OF $N_L^{(3/2)}$ FOR COSMOLOGICAL ANALYSES

We now turn to the impact of the bias for parameter inference.

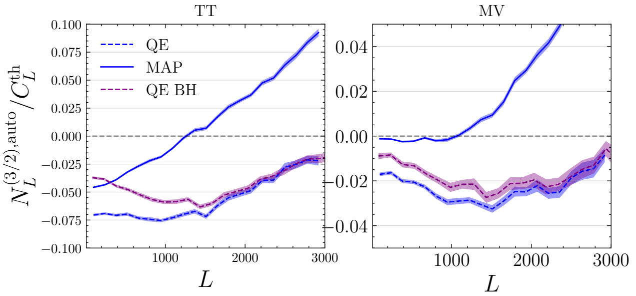

FIG. 10. Auto-spectrum $N_L^{(3/2)}$ biases induced by the non-Gaussian LSS-only map on several lensing reconstructions; in dashed we have the QE, in blue, and the bias hardened QE, in purple. The solid line is the MAP. The QE BH is able to mitigate the bias on large scales on a similar level to the MAP estimator, at a small cost in signal to noise compared to the QE. On the left panel, results with temperature data only, while and on the right including polarization too.

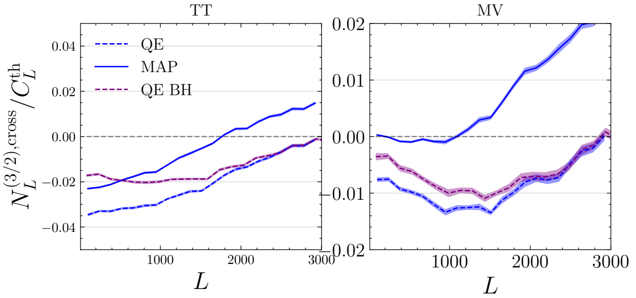

FIG. 11. Similar to Figure 10 but for the bias in cross-correlation to the true lensing. After bias-hardening, the quadratic estimator also shows a somewhat reduced $N_L^{(3/2)}$ bias.

A. Bias in the lensing amplitude

In this subsection we estimate the bias in the amplitude of CMB lensing power spectrum due to $N_L^{(3/2)}$. We assume that the covariance matrix of the CMB lensing power spectrum is diagonal and is given by

$$\mathbf{C}_{LL'} = \frac{2\delta_{LL'}}{(2L+1)f_{\mathrm{sky}}} \left(C_L^{\phi\phi} + N_L^{(0)} + N_L^{(1)}\right)^2 . \tag{49}$$

We report the signal to noise ratio of the CMB lensing power spectrum between multipoles $L_{\min}$ and $L_{\max}$ as

$$\mathrm{SNR} = \sqrt{\sum_{L_{\min}}^{L_{\max}} C_L^{\phi\phi} \mathbf{C}_{LL'}^{-1} C_{L'}^{\phi\phi}} . \tag{50}$$

This signal to noise ratio is exactly the square root of the Fisher information matrix on the amplitude of the CMB lensing spectrum. The inverse of this SNR thus gives the expected constraint on the amplitude of the CMB lensing spectrum if we only vary this parameter.

We show on the upper panel of the Figure 12 the signal to noise ratio as a function of the maximum lensing scale considered. We see that for the noise levels considered, the polarization estimator brings the most information, but the temperature estimator cannot be neglected. The MAP estimator performs particularly better than the QE especially in the polarization channel. We see that for the polarization, the information gain saturates above the scale $L \sim 1500$, while it still grows up to $L \sim 3000$ for the temperature and minimum variance estimators. This is an important point as future lensing analyses, including cross-correlations with large scale structure, will be able to push the signal to noise to high significances, provided that systematic and modelling treatments are accurately handled.

We estimate the bias in the CMB lensing amplitude as in [72]:

$$b(A_L) = \frac{1}{\mathrm{SNR}^2} \sum_{L_{\min}}^{L_{\max}} \Delta A_L \, , \tag{51}$$

with

$$\Delta A_L = N_L^{(3/2)} \mathbf{C}_{LL'}^{-1} C_{L'}^{\phi\phi} . \tag{52}$$

We show in the Table II the bias on the lensing amplitude for the different estimators with $L_{\min} = 0$, $L_{\max} = 4000$, including both the LSS and post-Born effects in $N_L^{(3/2)}$. We see that the MAP seems to be almost unbiased on temperature and polarization, but has an almost 1σ bias on the minimum variance combination.

We show in the lower panel of the Figure 12 the bias ΔA_L as a function of scale. As it can be seen from this figure, the $N_L^{(3/2)}$ introduce a scale dependent bias. In some cases, the $N_L^{(3/2)}$ bias flip signs. This lowers the sum in Eq. 51, and thus gives a low bias despite a large absolute value for some given scales. This is the case for the QE in polarization, which has a negative bias for $L \in [300, 1200]$ and positive outside, and for MAP in temperature, which is negative for $L < 1500$ and positive above. This scale dependent bias, which seems to be more important for the MAP than for the QE, could bias the measurements of cosmological parameters that are sensitive to the shape of the lensing spectrum. However, as the total bias almost cancels-out, the parameters combinations sensitive only to the amplitude of the lensing power spectrum should in principle be immune from shape changing effects.

B. Cosmological parameters

We now estimate the biases on cosmological parameters if one does not take into account the $N_L^{(3/2)}$ bias in the analysis. We focus on the sum of the neutrino masses, as this is a key science goal of CMB-S4. We consider the temperature only (TT), the polarization only (Pol), and the combined (MV) estimators, for both QE

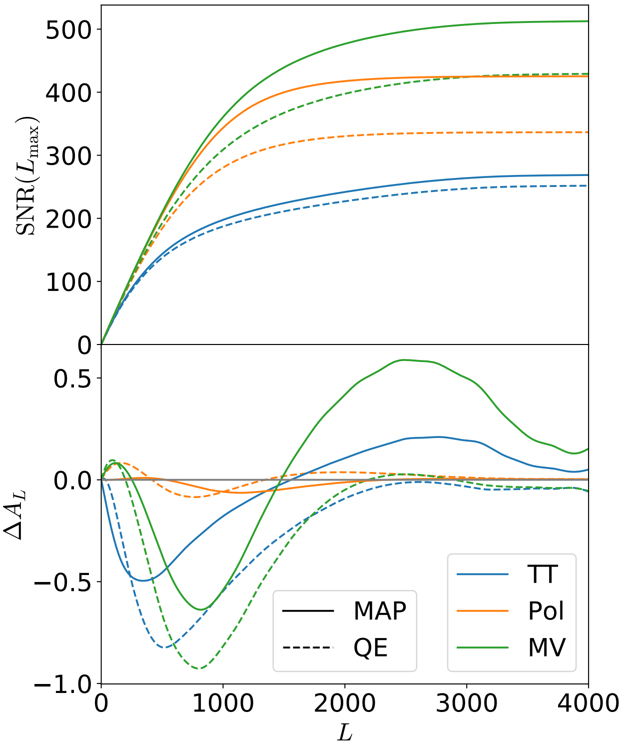

FIG. 12. *Upper panel:* Signal to noise ratio on the lensing power spectrum as a function of the maximum scale considered. We show the QE in dashed lines and the MAP in plain lines. The temperature only, polarization only and minimum variance are respectively in blue, orange and green. *Lower panel:* Bias on the lensing power spectrum amplitude as a function of scale, for the different estimators considered, including both the LSS and post-Born effects in $N_L^{(3/2)}$.

Bias	TT	Pol	MV
QE	-3.7σ	0.07σ	-2.13σ
MAP	-0.36σ	-0.12σ	0.93σ

TABLE II. Bias on the lensing power spectrum amplitude, estimated from the Fisher matrix, given in terms of number of sigmas, including both the LSS and post-Born effects in $N_L^{(3/2)}$.

and MAP. Our fiducial cosmology is the Planck FFP10 cosmology[17], with one massive neutrino of 0.06 eV.

Our analysis combines the CMB lensing spectra, the primary (unlensed) CMB spectra and the BAO. We assume that all three sets of observables are independent, so we can sum their log-likelihood. For the BAO we consider a DESI configuration, following the recipes of [73, 74]. For the CMB likelihood, we consider a CMB-S4 experiment, with a beam $\theta_{\rm FWHM} = 1\,$arcmin, a temperature noise $\Delta T = 1\,\mu$K-arcmin, a polarization noise $\Delta P = \sqrt{2}\,\mu$K-arcmin and a sky fraction $f_{\rm sky} = 0.4$. For the primary CMB likelihood, we consider unlensed spectra C_ℓ^{TT}, C_ℓ^{TE} and C_ℓ^{EE}, between multipoles 30 and 3000, and assume that these spectra follow a Gaussian likelihood, with Gaussian covariance matrix. We discuss in Appendix B the impact of using the unlensed CMB spectra or the lensed CMB spectra. We show there that the bias on the marginalized cosmological parameters are similar as long as we correctly model the non-Gaussian covariance of the lensed CMB and the correlations between the lensed CMB and the reconstructed lensing potential.

The Gaussian CMB lensing power spectrum likelihood is

$$-2\ln\mathcal{L}(\theta) = \left(\hat{C}_L^{\phi\phi} - C_L^{\rm th}(\theta)\right)\mathbf{C}_{LL'}^{-1}\left(\hat{C}_{L'}^{\phi\phi} - C_{L'}^{\rm th}(\theta)\right)\,, \tag{53}$$

where θ is the set of cosmological parameters being sampled. The covariance matrix is assumed to be diagonal

$$\mathbf{C}_{LL'} = \frac{2\delta_{LL'}}{(2L+1)f_{\rm sky}}\left(C_L^{\phi\phi} + N_L^{(0)} + N_L^{(1)}\right)^2\,, \tag{54}$$

where $C_L^{\phi\phi}$, $N_L^{(0)}$ and $N_L^{(1)}$ are evaluated in the fiducial cosmology. We consider lensing multipoles between 10 and 3000, and assume that the lensing field has been reconstructed with CMB multipoles between 10 and 4000 for both temperature and polarization channels. For the different estimators and different configurations of $N_L^{(3/2)}$ bias (LSS, PB or Total), we generate mock data vectors. These mock data vectors assume that the $N_L^{(0)}$ bias can be perfectly subtracted, and that the $N_L^{(1)}$ cosmology dependence is perfectly modelled. Our mock data vectors are given by

$$\hat{C}_L^{\phi\phi} = C_L^{\phi\phi}(\theta^{\rm fid}) + N_L^{(1)}(\theta^{\rm fid}) + \frac{C_L^{\phi\phi}(\theta^{\rm fid})}{C_L^{\phi\phi,\rm sim}}N_L^{(3/2)}\,. \tag{55}$$

where the $N_L^{(3/2)}$ bias has been estimated from simulations, as described in the previous sections, and the CMB lensing power spectrum and $N_L^{(1)}$ bias are taken at the fiducial cosmology. To cancel part of the realization variance of the simulation we divide $N_L^{(3/2)}$ by the lensing power spectrum of the simulation, bin this ratio in 19 multipole bins between 10 and 3810, and fit a spline to these points, weighting by the inverse variance of each bin. We then multiply this spline by our fiducial lensing spectrum. This also allows to partially take into account the difference in the cosmology between the simulation used to estimate the $N_L^{(3/2)}$ (which do not have massive neutrinos, see Equation 13), and the Planck FFP10 cosmology used to generate the data vector. We assume that higher order cosmology dependence of $N_L^{(3/2)}$ can be neglected.

---

[17] The CAMB parameter file used to generate the spectra can be found in `https://github.com/carronj/plancklens/blob/master/plancklens/data/cls/FFP10_wdipole_params.ini`

To accelerate the computation of the theory vector $C_L^{\rm th}(\theta)$, we do not re-estimate the $N_L^{(1)}$ bias for the sampled cosmology. Instead we correct for the variations of the $N_L^{(1)}$ bias around the fiducial, at first order in $C_L^{\phi\phi}(\theta)$, following the procedure of [19]. The theory vector is then

$$C_L^{\rm th}(\theta) = C_L^{\phi\phi}(\theta) + N_L^{(1)}(\theta^{\rm fid}) + \frac{\partial N_L^{(1)}}{\partial C_{L'}^{\phi\phi}}\left[C_{L'}^{\phi\phi}(\theta) - C_{L'}^{\phi\phi}(\theta^{\rm fid})\right] \tag{56}$$

where the $N_L^{(1)}$ correction matrix has been previously evaluated in the fiducial cosmology. We neglect the correction due to the variation of the response around the fiducial. We do not expect that this will impact the results. Indeed in the Gaussian case (i.e. without $N_L^{(3/2)}$ bias) we recover unbiased cosmological parameters.

In this analysis we do not model the reconstruction bias $N_L^{(0)}$ in the data vector, we assume it is perfectly subtracted. In a standard analysis one would use the realization dependent bias estimator RD-$N_L^{(0)}$, which makes the debiasing robust at first order to differences between the fiducial spectra assumed for the reconstruction, and the true CMB spectra of the maps. It was showed in [39, 40] that the realization dependent RD-$N_L^{(0)}$ allows for unbiased cosmological parameter estimates, for both the QE and the MAP estimators. However we note that this analysis was performed with a Gaussian lensing potential, contrary to the maps we are using here, but we assume that this will have a negligible impact on the RD-$N_L^{(0)}$ estimate.

We sample for seven cosmological parameters, namely $\ln(10^{10}A_{\rm s}), n_{\rm s}, \theta_{\rm MC}, \Omega_{\rm c}h^2, \Omega_{\rm b}h^2, \tau$ and $\sum m_\nu$. We include a strong Gaussian prior on τ, the reionisation optical depth, assuming a cosmic variance limit of $\sigma_\tau = 0.002$.

We rely on the CAMB cosmological Boltzmann code [75, 76], and sample the posterior with adaptive, speed-hierarchy-aware MCMC sampler (adapted from CosmoMC) [77, 78]. We explore the posterior using the Cobaya [79] and GetDist packages [80].

The marginalized posterior distribution on the sum of the neutrino masses, for our different lensing estimators and source of $N_L^{(3/2)}$ bias, is shown in Figure 13. We show the impact of each term that contributes to the $N_L^{(3/2)}$ bias: the large scale structures non-Gaussianities (LSS) and the post-Born lensing. We confirm that taking into account only one of the two effects, either LSS or PB, will bias the neutrino mass estimates with the QE. We can see that in general, the MAP estimators are less impacted than the QE. When looking at the total $N_L^{(3/2)}$, we see that the MAP is unbiased, both for temperature and for polarization. The QE shows a $\sim 1\sigma$ bias with the temperature reconstruction, while it is unbiased in the polarization.

We note also that as previously observed in the literature [39, 81], the marginalized constraints on the sum of the neutrino mass from the QE is not improved with the MAP. Indeed, even if, as we showed in the Figure 12, the SNR is increased with the MAP, there are some degeneracies between the cosmological parameters that prevent to reach the full statistical power of the MAP estimator on the marginalized constraints. The marginalized constraints also do not change much if we use the temperature, polarization or the minimum variance estimators, with a one sigma constraint of about 0.016 eV for all cases.

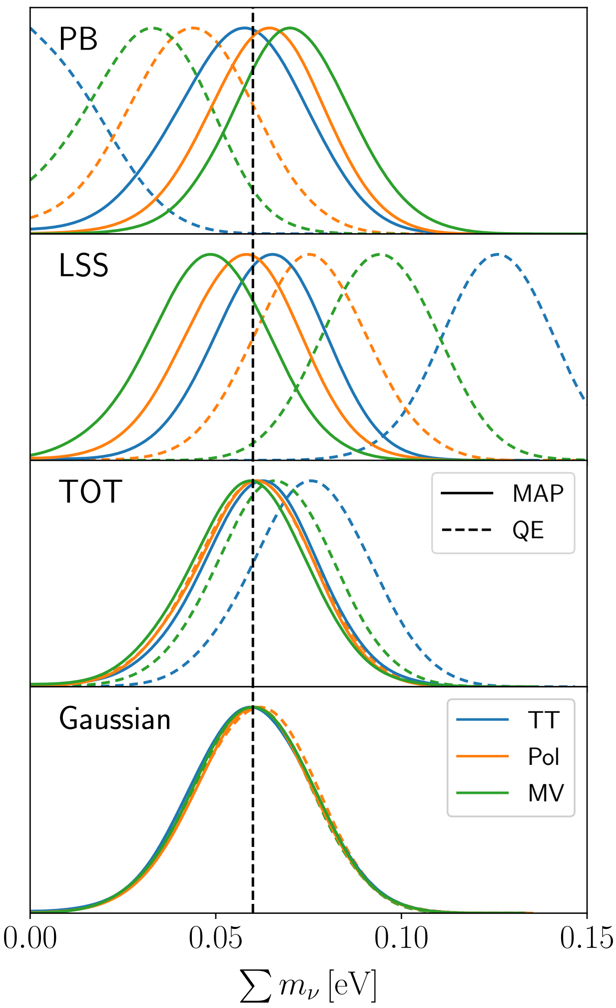

FIG. 13. marginalized posteriors on the sum of the neutrino masses. Panels from top to bottom show respectively the impact of $N_L^{(3/2)}$ when considering the post-Born term only, the large scale structure term only, both terms together, or with no $N_L^{(3/2)}$ bias (fiducial, Gaussian case). The dashed lines are for the QE and the plain lines are for the MAP. Blue, orange and green lines show the temperature only, polarization only or minimum variance estimators. The fiducial cosmology assumes a sum of neutrino masses of 0.06 eV, showed as the vertical dashed line. We combine here the CMB-S4 likelihood with DESI-BAO, and we set a cosmic variance Gaussian prior on the optical depth to reionization.

Figure 14 shows the posterior distribution for the n_s, $\Omega_c h^2$ and $\sum m_\nu$ parameters, for the total $N_L^{(3/2)}$ bias

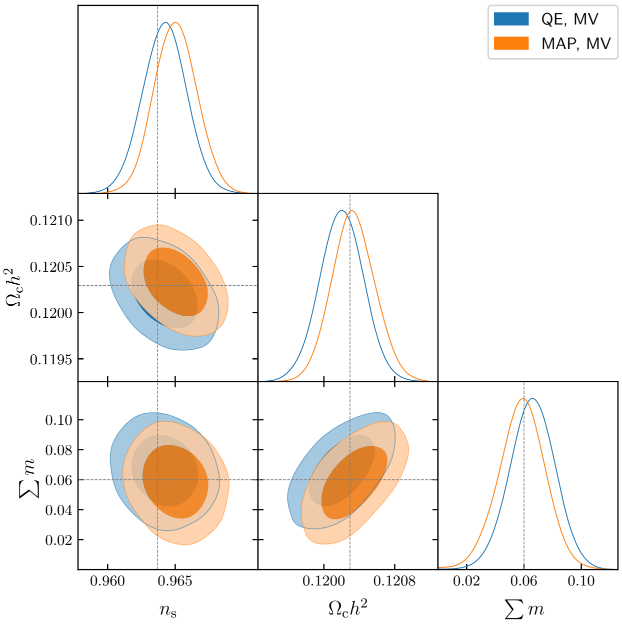

FIG. 14. Posterior distribution for a subset of the sampled cosmological parameters. We show the reconstruction for the minimum variance estimator, for both QE (blue) and MAP (orange), considering the bias due to the total $N_L^{(3/2)}$. The Figure 18 in the Appendix shows the full posterior.

with the minimum variance estimator. We see that even if the MAP is unbiased on the sum of the neutrino mass, it shows a slight 0.5σ bias on the n_s parameter. This bias could be sourced by the shape of the $N_L^{(3/2)}$ bias. Indeed, we showed in Figure 12 that the $N_L^{(3/2)}$ creates a scale dependent bias. In particular, this could tilt the lensing power spectrum, and thus create a bias on the n_s parameter. In comparison, the amplitude bias for the minimum variance QE is mostly negative, so it will not create a strong scale dependent bias, but rather an overall amplitude shift on the CMB lensing power spectrum.

In Appendix B, we compare our results for the QE with the previous results from [42]. They obtained a larger bias on the sum of the neutrino masses. This higher bias comes from the fact that they consider lensed CMB spectra in the likelihood but neglect the non-Gaussian covariance due to lensing and the correlations between CMB and the QE, as described in [82]. In Appendix B we show that using lensed CMB spectra with the correct covariance results in a bias on the sum of the neutrino mass that is similar to the one obtained when considering unlensed spectra and neglecting the non diagonal correlations.

VII. CONCLUSIONS

Future CMB surveys will detect the CMB lensing auto-spectrum well over 100σ significance, and will enable powerful cosmological constraints on the structure formation history and on the sum of neutrino masses. Great care is required before interpreting these measurements at this high level of sensitivity. Standard quadratic estimators are only guaranteed to be unbiased for purely Gaussian fields and under idealized conditions. When the Gaussianity assumptions fail, future CMB surveys such as CMB-S4 can deliver a biased lensing reconstruction induced by the bispectrum of the lensing potential which, in turn will affect constraints on the sum of neutrino masses [42].

In this work we studied the induced non-Gaussian biases in the auto-spectrum of the CMB lensing potential reconstructed through maximum a posteriori estimators. By using state-of-the art simulations that include non-Gaussian lensing deflections induced by nonlinear matter clustering as well as post-Born lensing corrections, we found that for scales relevant for a CMB-S4 like configuration the non-Gaussian induced bias is mitigated compared to the QE one. When including additional effects due to the post-Born rotation we do not find a significant difference.

When performing a ΛCDM cosmological forecast on the sum of neutrino masses through a full MCMC analysis, we found that contrary to the QE, the MAP estimator does not suffer from biases even if $N_L^{(3/2)}$ biases are not specifically accounted in the fitting of the data vector. This shows that cosmological analyses using a MAP reconstructed lensing potential should in principle be more robust than those based on the standard QE. However, if not correcting for the bias, caution should still be taken when investigating extensions to ΛCDM, as the bias has not fully disappeared, but is only to weak to shift the maximum posterior point, and this could change in other models.

We also tested simple modifications to the MAP estimator to take into account the non-Gaussian statistics of the deflection field. Using the prior of a log-normal field instead of a Gaussian field on the input lensing convergence further mitigates slightly the bias. We have also found that the bias itself found on our N-body simulations is reasonably well reproduced by much simpler lognormal simulations. Generally, differences between reconstructions using the Gaussian or lognormal prior are small, and the lensing reconstructions of similar quality, both on N-body and lognormal simulated inputs. Lognormal simulations are effective in replicating non-Gaussian bias effects from large-scale structure non-linearity. They could be leveraged to predict and subtract biases with sufficient accuracy even when applied to more sophisticated models affecting the full probability density function of the projected matter distribution, such as baryonic effects or extended cosmologies. On the other hand we have found that the lognormal distribution is not a good approximation when including post-Born effects, suggesting that the lognormal approximation is insufficient for recovering the shape of the post-Born induced bispectrum (and same would be for highly non-linear fields, necessitating alternative approaches [54, 68, 69]). Nevertheless, using state-of-the art simulations one could learn realistic probability density functions of fields using machine learning techniques to generate new samples or even include alternative priors in our MAP formulation (e.g. [83]).

Our results allow us to conclude that the non-Gaussianity of CMB lensing deflections should not present a major challenge to CMB lensing science in the near future: the bias, small to start with, is mitigated, and the baseline MAP reconstructions with Gaussian prior remain very close to optimal.

In the era of precision-cosmology, additional information will come from the combination and cross-correlation of CMB lensing measurements and multiple large-scale structure observables, such as cosmic shear and galaxy clustering, from DESI, Euclid, and Rubin. It is known that in this case, for QE-based reconstructions, the $N_L^{(3/2)}$ bias in the CMB lensing cross-correlations has a stronger impact depending on the specific tracer and redshift considered in the analysis [45]. This is left for future work.

ACKNOWLEDGMENTS

We thank Antony Lewis for comments. OD, SB, LL and JC acknowledge support from a SNSF Eccellenza Professorial Fellowship (No. 186879). This work was supported by a grant from the Swiss National Supercomputing Centre (CSCS) under project ID s1203. Part of the computations were performed at University of Geneva using Baobab HPC service. GF acknowledges the support of the European Research Council under the Marie Skłodowska Curie actions through the Individual Global Fellowship No. 892401 PiCOGAMBAS.

Appendix A: Mean field impact

The MAP estimator tries to find the best estimate of the CMB lensing potential by nulling the total gradient

$$g^{\rm tot}(\phi) = g^{\rm QD}(\phi) + g^{\rm MF}(\phi) + g^{\rm PR}(\phi) \tag{A1}$$

Ignoring the mean field contribution $g^{\rm MF}$, as we often do for simplicity, one is effectively nulling instead

$$g^{\rm tot,noMF}(\hat\phi) = g^{\rm QD}(\hat\phi) + g^{\rm PR}(\hat\phi) \overset{!}{=} 0 \ . \tag{A2}$$

The neglect of this gradient piece is equivalent to maximizing a slightly different, but still well-defined likelihood function. Hence the iterative procedure converges without problems. The actual total gradient at this point $\hat\phi$ is not zero but

$$g^{\rm tot}(\hat\phi) = g^{\rm MF}(\hat\phi) \neq 0 \ . \tag{A3}$$

Early indications from reconstructions from polarization suggested this term was small [47]. Physically, this is because this term accounts for the anisotropies in the noise maps induced by delensing, and these anisotropies trace for the most part the magnification part of the lensing signal, rather than the shear-like signal that the most powerful EB polarized estimator is sensitive to. Delensing affects the local noise levels by changing areas according to the local magnification, given by $1 - 2\kappa$ to first order, and the mean-field removes the contribution of this noise anisotropy to the quadratic gradient piece.

It turns out that that for reconstructions from temperature, that takes substantial contribution from the magnification-like lensing signals, the mean field contribution is much larger. If the mean-field at the solution $\hat\kappa_{LM}$ were exactly proportional to $\hat\kappa_{LM}$ as suggested by the argument above, the mean-field contamination of $\hat\kappa$ acts simply as a rescaling of the output map. This would have no impact on the cross-correlation coefficient of the reconstruction with the input (the 'quality' of the reconstruction), only on its normalization

$$\mathcal{W}_L^{\rm emp} \equiv \frac{C_L^{\hat\phi\phi}}{C_L^{\phi\phi}} \tag{A4}$$

(the empirical 'Wiener-filter' of the reconstruction). The shift seen on temperature reconstruction for our CMB-S4-like configuration is shown on Fig. 15, and is substantial. Not accounting for the mean field lowers the Wiener-filter curve. This is because the magnification follows κ but with a minus sign. In Fig. 16 we show that after applying the empirical normalization, for the case of mean-field subtraction/no subtraction, we find a power

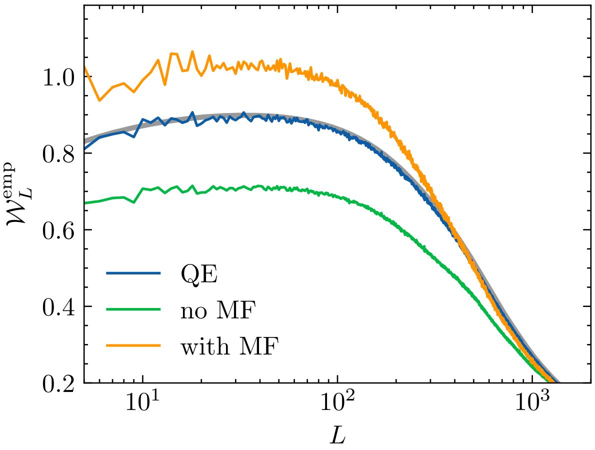

FIG. 15. Comparison of the empirical normalization of output lensing maps, after a few iterations, for our temperature-based CMB-S4 like reconstructions. In blue we show the case of the QE estimator, in green for the MAP estimator if ignoring the mean field contribution, and in orange when including it, using an estimate from a finite number of Monte-Carlo at each step.

spectrum in line with naive expectations,

$$C_L^{\hat{\phi}\hat{\phi}} \sim C_L^{\phi\phi} + N_L^{(0)} + N_L^{(1)} \tag{A5}$$

where the lensing biases are calculated using the partially delensed CMB spectra. The change in cross-correlation coefficient between the two maps is tiny. This gives support to the idea that the bulk of the effect is only a rescaling. Hence, in this paper we proceed using the empirical normalization, and neglecting the delensed-noise mean-field altogether. A more careful study of the mean-field is ongoing.

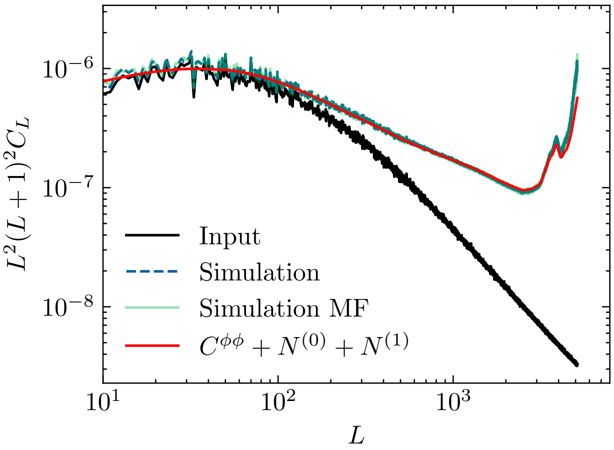

FIG. 16. Lensing spectrum from a MAP temperature reconstruction, after dividing by the empirical normalization of the lensing map (blue). In red the naive theory prediction, obtained from a iterative $N_L^{(0)}$ and $N_L^{(1)}$ calculation. We get a similar result if we account for the mean field, justifying the usage of an empirical normalisation when not account for a mean-field.

Appendix B: About the impact of non-Gaussian lensing correlations on the cosmological parameter biases

In the previous study of [42] (herefater Beck18), the likelihood was neglecting the correlations between the reconstructed lensing field and the CMB spectra. In the results we present in the Section VI B above, assume we can neglect these correlations by using unlensed CMB spectra un the likelihood. We will now make systematic study of this assumption, and evaluate the impact of using the lensed or unlensed CMB spectra, and the impact of the non-Guassian covariance on the final cosmological constraints.

We will consider four different likelihood configurations:

1. The Beck18 likelihood
2. The Beck18 likelihood but with unlensed CMB spectra
3. The likelihood we introduced in Section VI B, which uses unlensed CMB spectra and assumes no correlations between the lensing power spectrum and the CMB fields
4. A complete likelihood including the full non-Gaussian correlations between the lensed CMB spectra and the lensing potential, following [82]

The Beck18 CMB likelihood is

$$-2\log\mathcal{L}(\theta|\hat{\mathbf{C}}) = \sum_\ell (2\ell+1) f_{\rm sky}\left(\ln\frac{|\mathbf{C}_\ell|}{|\hat{\mathbf{C}}_\ell|} + \mathbf{C}_\ell^{-1}\hat{\mathbf{C}}_\ell - 3\right) \tag{B1}$$

where the theoretical covariance matrix $\mathbf{C}$ and the data covariance matrix $\hat{\mathbf{C}}$ are constructed with the lensed CMB spectra and are given by

$$\mathbf{C}_\ell = \begin{pmatrix} C_\ell^{TT} + N_\ell^{TT} & C_\ell^{TE} & C_\ell^{T\phi} \\ C_\ell^{TE} & C_\ell^{EE} + N_\ell^{EE} & 0 \\ C_\ell^{T\phi} & 0 & C_\ell^{\phi\phi} + N_\ell^{\phi\phi} \end{pmatrix} \tag{B2}$$

For the theoretical covariance $\mathbf{C}_\ell$, the CMB lensing noise is $N_\ell^{\phi\phi} = N_\ell^{(0)} + N_\ell^{(1)}$. For the mock data covariance we include the $N_L^{(3/2)}$ bias estimated from the simulations as

$$N_L^{\phi\phi} = N_L^{(0)} + N_L^{(0)} + \frac{C_L^{\phi\phi,\rm fid}}{C_L^{\phi\phi,\rm sim}} N_L^{(3/2),\rm sim} \, . \tag{B3}$$

In this configuration, and in order to reproduce the results of Beck18, we do not take into account the cosmological dependence of the $N_L^{(0)}$ and $N_L^{(1)}$ biases, and we keep them fixed to their fiducial value.

The Beck18 likelihood with unlensed spectra is the same as above but replaces the lensed CMB spectra by their unlensed version.

The likelihood which takes into account the non-Gaussian covariance introduced in [82] has the following data vector:

$$\mathcal{C}_\ell = \left(C_\ell^{TT}, C_\ell^{EE}, C_\ell^{TE}, C_\ell^{\phi\phi} + N_\ell^{(1)}\right) \, . \tag{B4}$$

For the theory vector $\mathcal{C}_\ell(\theta)$ we vary the theoretical $N_\ell^{(1)}$ at first order in the CMB lensing spectra when sampling the cosmology, like in Eq. 56. The mock data vector $\hat{\mathcal{C}}_\ell$ includes the $N_L^{(3/2)}$ bias, which is not modelled by the theory vector. The full likelihood is then

$$-2\log\mathcal{L}(\theta) = \left(\hat{\mathcal{C}}_\ell - \mathcal{C}_\ell(\theta)\right)\mathbf{C}_{\ell\ell'}^{-1}\left(\hat{\mathcal{C}}_{\ell'} - \mathcal{C}_{\ell'}(\theta)\right) \tag{B5}$$

The covariance matrix $\mathbf{C}_{\ell\ell'}$ includes the non-Gaussian terms described in [82], such as the correlations between the CMB spectra due to lensing, and the correlation between the reconstructed lensing field and the CMB.

In all likelihood scenarios we generate the data vector using the same cosmology of Beck18, reproduced in Table B, contrary to the analysis in the Section VI B where

$\Omega_{\rm b}h^2$	0.02225 ± 0.00016
$\Omega_{\rm c}h^2$	0.1198 ± 0.0015
τ	0.058 ± 0.012
$\log(10^{10}A_{\rm s})$	3.094 ± 0.034
$n_{\rm s}$	0.9645 ± 0.0049
$100\theta_{\rm MC}$	1.04077 ± 0.00032
$\sum m_\nu$ [eV]	$[0, 300]$

TABLE III. Fiducial cosmological parameters used in this section, following Beck18, together with their 1σ Gaussian prior, or uniform parameter bound.

we used the FFP10 cosmology. Notably, we now consider massless neutrinos, instead of the minimal mass normal hierarchy considered before. Moreover, contrary to the main analysis, the likelihoods used here do not include external BAO constraints. We follow again Beck18 and assume tight Gaussian priors on the parameters, except for the sum of neutrino mass which assumes a flat prior, given on Table B.

We perform the MCMC samplings like in the main analysis, and we now discuss the results obtained for the QE minimum variance estimator, with the LSS only term in the $N_L^{(3/2)}$ bias. We show in the Figure 17 the constraints for three cosmological parameters $\Omega_c h^2$, $\ln(10^{10}A_s)$ and $\sum m_\nu$. The purple contours are the one obtained with the likelihood of Beck18, with the lensed CMB spectra. We retrieve the same results as in Beck18 (see their Figure 11), with a posterior estimate on the sum of neutrino masses peaking at 0.18 eV. The red contours show the posterior with the Beck18 likelihood with unlensed CMB spectra in the data and theory vectors. The blue contours are for our likelihood, using unlensed CMB spectra, for Beck18 cosmology and priors. In practice, the main difference here with the Beck18 unlensed likelihood is that we take into account the variation of the $N_L^{(1)}$ bias when sampling the cosmology, while it is not the case for the Beck18 likelihood. Finally, the green contours are for the likelihood with lensed CMB spectra and the [82] non-Gaussian covariance.

It appears that the Beck18 likelihood, with the lensed CMB spectra, has the largest bias on the sum of the neutrino masses and on the amplitude of the matter power spectrum. We see that when considering unlensed CMB spectra the Beck18 likelihood obtains similar constraints as ours, whihc has unlensed spectra as well. We show that we obtain almost exactly the same posterior, with a reduced bias, if we use the lensed spectra but correctly taking into account the correlations due to lensing from [82]. This reduction of the bias when using the full non-Gaussian covariance can be understood as lowering the impact of the bias from the lensing power spectrum, by avoiding to "double" count the biased lensing power spectrum in the analysis. If we include as well the lensed C_ℓ^{BB} spectrum in the data vector, even if the correlations are properly taken into account, the bias is not reduced. We interpret that as the fact that the lensed spectrum is another measurement of $C_L^{\phi\phi}$, as we can approximate for $\ell \ll 1000$ [3]

$$C_\ell^{BB} \sim \frac{1}{4\pi} \int \frac{d\ell'}{\ell'} \ell'^4 C_{\ell'}^{\phi\phi} \ell'^2 C_{\ell'}^{EE,\rm unl} . \tag{B6}$$

Thus there is a tension between the lensed BB spectrum, which measures an unbiased $C_L^{\phi\phi}$, and the estimate lensing power spectrum which is biased by $N_L^{(3/2)}$. It appears that the biased $C_L^{\phi\phi}$ dominates and the likelihood is leaning towards the biased measurement of the neutrino mass. In practice, future surveys might not include the lensed BB spectrum in the analysis since the BB spectrum might be more prone to instrumental systematics and polarized foregrounds such as point sources. But comparing the marginalized constraints with or without the BB spectra in the likelihood could serve as additional robustness test to assess the presence of biases coming from non-Gaussian effects.

We do not test the impact of the full non-Gaussian covariance for the MAP estimator. In that case, one would use the partially delensed CMB spectra in the likelihood, reducing the non-diagonal terms in the covariance matrix. Fisher forecasts using the covariance matrix computed with delensed spectra were published in [93, 94]. We leave for a future work a detailed comparison of our MAP reconstructed delensed spectra with this delensed analytical covariance.

We show in the Figure 18 the full posterior of the likelihood for the QE and MAP reconstruction, considering the total $N_L^{(3/2)}$ bias, and seven cosmological parameters as described in the Section VI B.

---

[*] omar.darwish@unige.ch

[†] sebastian.belkner@unige.ch

[1] J. Aguirre *et al.* (Simons Observatory), The Simons Observatory: Science goals and forecasts, JCAP **1902**, 056 (2019), arXiv:1808.07445 [astro-ph.CO].

[2] K. N. Abazajian *et al.*, CMB-S4 Science Book, First Edition, (2016), arXiv:1610.02743.

[3] A. Lewis and A. Challinor, Weak gravitational lensing of the CMB, Physics Reports **429**, 1 (2006), arXiv:astro-ph/0601594.

[4] R. de Putter, O. Zahn, and E. V. Linder, CMB lensing constraints on neutrinos and dark energy, Phys. Rev. D **79**, 065033 (2009), arXiv:0901.0916.

[5] R. Stompor and G. Efstathiou, Gravitational lensing of cosmic microwave background anisotropies and cosmological parameter estimation, MNRAS **302**, 735 (1999), arXiv:astro-ph/9805294 [astro-ph].

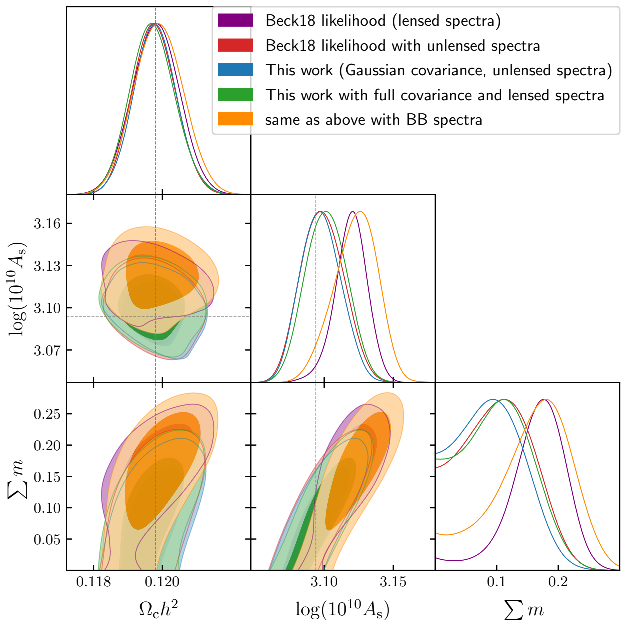

FIG. 17. marginalized contours for three cosmological parameters, for the QE minimum variance, in the presence of the LSS only $N_L^{(3/2)}$ bias. We show the results for the Beck18 likelihood with the lensed CMB spectra (in purple) or with the unlensed CMB spectra (in red). We compare with our likelihood used for the main results in Section VI B (in blue). We also show in green the likelihood which includes the full non-Gaussian covariance of the lensed CMB spectra, and the correlations between the reconstructed lensing spectrum and the CMB. We see that correctly taking into account these correlations gives almost the same posterior as considering unlensed CMB spectra and neglecting correlations. In orange we add the lensed BB spectrum in the data vector, taking into account its correlations with the lensed CMB and the reconstructed lensing potential.

[6] N. Sailer, E. Castorina, S. Ferraro, and M. White, Cosmology at high redshift — a probe of fundamental physics, Journal of Cosmology and Astroparticle Physics **2021** (12), 049.

[7] A. Vallinotto, Using Cosmic Microwave Background Lensing to Constrain the Multiplicative Bias of Cosmic Shear, Astrophys. J. **759**, 32 (2012), arXiv:1110.5339.

[8] E. Schaan, E. Krause, T. Eifler, O. Doré, H. Miyatake, J. Rhodes, and D. N. Spergel, Looking through the same lens: Shear calibration for LSST, Euclid, and WFIRST with stage 4 CMB lensing, Phys. Rev. D **95**, 123512 (2017), arXiv:1607.01761.

[9] R. Cawthon, Effects of redshift uncertainty on cross-correlations of CMB lensing and galaxy surveys, Phys. Rev. D **101**, 063509 (2020), arXiv:1809.09251.

[10] A. Vallinotto, The Synergy between the Dark Energy Survey and the South Pole Telescope, Astrophys. J. **778**, 108 (2013), arXiv:1304.3474.

[11] L. Wenzl, C. Doux, C. Heinrich, R. Bean, B. Jain, O. Doré, T. Eifler, and X. Fang, Cosmology with the Roman Space Telescope - Synergies with CMB lensing, MNRAS **512**, 5311 (2022), arXiv:2112.07681.

[12] Euclid Collaboration, S. Ilić, *et al.*, Euclid preparation. XV. Forecasting cosmological constraints for the Euclid and CMB joint analysis, A&A **657**, A91 (2022), arXiv:2106.08346.

[13] A. Lewis and L. King, Cluster masses from CMB and galaxy weak lensing, Phys. Rev. D **73**, 063006 (2006), arXiv:astro-ph/0512104 [astro-ph].

[14] W. Hu, S. DeDeo, and C. Vale, Cluster mass estimators from CMB temperature and polarization lensing, New Journal of Physics **9**, 441 (2007), arXiv:astro-ph/0701276 [astro-ph].

[15] S. Raghunathan, S. Patil, E. J. Baxter, F. Bianchini, L. E. Bleem, T. M. Crawford, G. P. Holder, A. Manzotti, and C. L. Reichardt, Measuring galaxy cluster masses with CMB lensing using a Maximum Likelihood estimator: statistical and systematic error budgets for future experiments, JCAP **2017**, 030 (2017), arXiv:1705.00411.

[16] U. Seljak and C. M. Hirata, Gravitational lensing as a contaminant of the gravity wave signal in the CMB, Phys. Rev. D

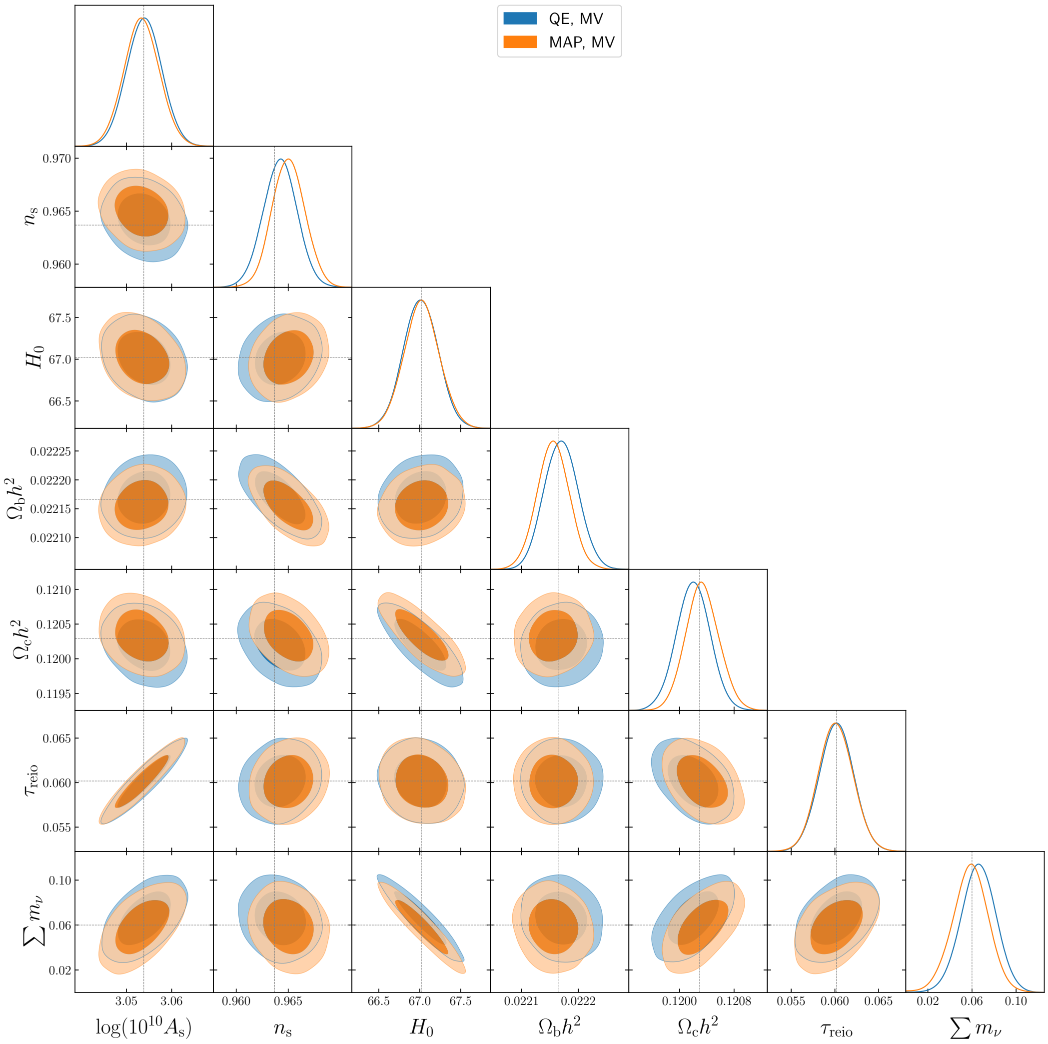

FIG. 18. Full posterior distribution for the seven sampled cosmological parameters. We show the reconstruction for the minimum variance estimator, for both QE (blue) and MAP (orange), considering the bias due to the total $N_L^{(3/2)}$.

69, 043005 (2004), arXiv:astro-ph/0310163 [astro-ph].

[17] K. M. Smith, D. Hanson, M. LoVerde, C. M. Hirata, and O. Zahn, Delensing CMB polarization with external datasets, JCAP **2012**, 014 (2012), arXiv:1010.0048.

[18] W. Hu and T. Okamoto, Mass Reconstruction with Cosmic Microwave Background Polarization, Astrophys. J. **574**, 566 (2002), arXiv:astro-ph/0111606.

[19] N. Aghanim *et al.* (Planck), Planck 2018 results. VIII. Gravitational lensing, A&A **641**, A8 (2020), arXiv:1807.06210 [astro-ph.CO].

[20] J. Carron, M. Mirmelstein, and A. Lewis, CMB lensing from Planck PR4 maps, JCAP **2022**, 039 (2022), arXiv:2206.07773.

[21] M. S. Madhavacheril *et al.*, The Atacama Cosmology Telescope: DR6 Gravitational Lensing Map and Cosmological Parameters, (2023), arXiv:2304.05203.

[22] W. L. K. Wu *et al.*, A Measurement of the Cosmic Microwave Background Lensing Potential and Power Spectrum from 500 deg^2 of SPTpol Temperature and Polarization Data, Astrophys. J. **884**, 70 (2019), arXiv:1905.05777.

[23] M. A. Faúndez *et al.*, Measurement of the Cosmic Microwave Background Polarization Lensing Power Spectrum from Two Years of POLARBEAR Data, Astrophys. J. **893**, 85 (2020), arXiv:1911.10980.

[24] BICEP/Keck Collaboration, :, P. A. R. Ade, *et al.*, BICEP / Keck XVII: Line of Sight Distortion Analysis: Estimates of Gravitational Lensing, Anisotropic Cosmic Birefringence, Patchy Reionization, and Systematic Errors, (2022), arXiv:2210.08038.

[25] N. Sailer, S. Ferraro, and E. Schaan, Foreground-immune CMB lensing reconstruction with polarization, Phys. Rev. Lett. **107**, 023504 (2023), arXiv:2211.03786.

[26] A. S. Maniyar, Y. Ali-Haïmoud, J. Carron, A. Lewis, and M. S. Madhavacheril, Quadratic estimators for CMB weak

lensing, Phys. Rev. D **103**, 083524 (2021), arXiv:2101.12193.
[27] J. Carron, Real-world CMB lensing quadratic estimator power spectrum response, JCAP **02**, 057 (2023).
[28] D. Beck, J. Errard, and R. Stompor, Impact of Polarized Galactic Foreground Emission on CMB Lensing Reconstruction and Delensing of B-Modes, JCAP **06**, 030 (2020), arXiv:2001.02641.
[29] E. Schaan and S. Ferraro, Foreground-Immune Cosmic Microwave Background Lensing with Shear-Only Reconstruction, Phys. Rev. Lett. **122**, 181301 (2019), arXiv:1804.06403.
[30] M. Mirmelstein, G. Fabbian, A. Lewis, and J. Peloton, Instrumental systematics biases in CMB lensing reconstruction: A simulation-based assessment, Phys. Rev. D **103**, 123540 (2021), arXiv:2011.13910.
[31] O. Darwish, B. D. Sherwin, N. Sailer, E. Schaan, and S. Ferraro, Optimizing foreground mitigation for cmb lensing with combined multifrequency and geometric methods, Phys. Rev. D **107**, 043519 (2023), arXiv:2111.00462.
[32] M. Millea, E. Anderes, and B. D. Wandelt, Bayesian delensing of CMB temperature and polarization, Phys. Rev. D **100**, 023509 (2019).
[33] M. Millea, E. Anderes, and B. D. Wandelt, Sampling-based inference of the primordial CMB and gravitational lensing, Phys. Rev. D **102**, 123542 (2020).
[34] M. Millea and U. Seljak, Marginal unbiased score expansion and application to CMB lensing, Phys. Rev. D **105**, 103531 (2022).
[35] C. M. Hirata and U. Seljak, Reconstruction of lensing from the cosmic microwave background polarization, Phys. Rev. D **68**, 083002 (2003).
[36] J. Carron and A. Lewis, Maximum a posteriori CMB lensing reconstruction, Phys. Rev. D **96**, 063510 (2017).
[37] S. Adachi, others, and Polarbear Collaboration, Internal Delensing of Cosmic Microwave Background Polarization B-Modes with the POLARBEAR Experiment, Phys. Rev. Lett. **124**, 131301 (2020), arXiv:1909.13832.
[38] M. Millea *et al.*, Optimal Cosmic Microwave Background Lensing Reconstruction and Parameter Estimation with SPTpol Data, Astrophys. J. **922**, 259 (2021), arXiv:2012.01709.
[39] L. Legrand and J. Carron, Lensing power spectrum of the cosmic microwave background with deep polarization experiments, Phys. Rev. D **105**, 123519 (2022), arXiv:2112.05764.
[40] L. Legrand and J. Carron, Robust and efficient CMB lensing power spectrum from polarization surveys, Phys. Rev. D **108**, 103516 (2023), arXiv:2304.02584 [astro-ph.CO].
[41] V. Böhm, M. Schmittfull, and B. D. Sherwin, A bias to CMB lensing measurements from the bispectrum of large-scale structure, **94**, 043519 (), 1605.01392.
[42] D. Beck, G. Fabbian, and J. Errard, Lensing Reconstruction in Post-Born Cosmic Microwave Background Weak Lensing, **98**, 043512, 1806.01216.
[43] V. Böhm, B. D. Sherwin, J. Liu, J. C. Hill, M. Schmittfull, and T. Namikawa, On the effect of non-Gaussian lensing deflections on CMB lensing measurements, **98**, 123510 (), 1806.01157.
[44] G. Pratten and A. Lewis, Impact of post-Born lensing on the CMB, JCAP **8**, 047 (2016), arXiv:1605.05662.
[45] G. Fabbian, A. Lewis, and D. Beck, CMB lensing reconstruction biases in cross-correlation with large-scale structure probes, **2019**, 057, 1906.08760.
[46] J. Carron and A. Lewis, Maximum a posteriori CMB lensing reconstruction, **96**, 063510, 1704.08230.
[47] S. Belkner, J. Carron, L. Legrand, C. Umiltà, C. Pryke, and C. Bischoff (CMB-S4), CMB-S4: Iterative internal delensing and r constraints, (2023), arXiv:2310.06729 [astro-ph.CO].
[48] C. M. Hirata and U. Seljak, Analyzing weak lensing of the cosmic microwave background using the likelihood function, **67**, 043001 ().
[49] C. M. Hirata and U. Seljak, Reconstruction of lensing from the cosmic microwave background polarization, **68**, 083002 ().
[50] D. Hanson, A. Challinor, and A. Lewis, Weak lensing of the CMB, General Relativity and Gravitation **42**, 2197 (2010).
[51] M. Kesden, A. Cooray, and M. Kamionkowski, Lensing reconstruction with cmb temperature and polarization, Physical Review D **67**, 10.1103/physrevd.67.123507 (2003).
[52] T. Namikawa, D. Hanson, and R. Takahashi, Bias-hardened cmb lensing, Monthly Notices of the Royal Astronomical Society **431**, 609–620 (2013).
[53] M. Reinecke, S. Belkner, and J. Carron, Improved cosmic microwave background (de-)lensing using general spherical harmonic transforms (2023), arXiv:2304.10431 [astro-ph.CO].
[54] G. Fabbian, M. Calabrese, and C. Carbone, CMB weak-lensing beyond the Born approximation: a numerical approach, JCAP **2**, 050 (2018), arXiv:1702.03317.
[55] S. Hilbert, J. Hartlap, S. D. M. White, and P. Schneider, Ray-tracing through the Millennium Simulation: Born corrections and lens-lens coupling in cosmic shear and galaxy-galaxy lensing, A&A **499**, 31 (2009), arXiv:0809.5035.
[56] C. Carbone, M. Petkova, and K. Dolag, DEMNUni: ISW, Rees-Sciama, and weak-lensing in the presence of massive neutrinos, JCAP **7**, 034 (2016), arXiv:1605.02024.
[57] E. Castorina, C. Carbone, J. Bel, E. Sefusatti, and K. Dolag, DEMNUni: the clustering of large-scale structures in the presence of massive neutrinos, JCAP **7**, 043 (2015), arXiv:1505.07148.
[58] C. Carbone, V. Springel, C. Baccigalupi, M. Bartelmann, and S. Matarrese, Full-sky maps for gravitational lensing of the cosmic microwave background, MNRAS **388**, 1618 (2008), arXiv:0711.2655.
[59] M. Calabrese, C. Carbone, G. Fabbian, M. Baldi, and C. Baccigalupi, Multiple lensing of the cosmic microwave background anisotropies, JCAP **3**, 049 (2015), arXiv:1409.7680.
[60] P. Fosalba, E. Gaztañaga, F. J. Castander, and M. Manera, The onion universe: all sky lightcone simulations in spherical shells, MNRAS **391**, 435 (2008), arXiv:0711.1540.
[61] G. Fabbian and R. Stompor, High-precision simulations of the weak lensing effect on cosmic microwave background polar-

ization, A&A **556**, A109 (2013), arXiv:1303.6550 [astro-ph.CO].

[62] H. S. Xavier, F. B. Abdalla, and B. Joachimi, Improving lognormal models for cosmological fields, **459**, 3693, 1602.08503.

[63] K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann, HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere, Astrophys. J. **622**, 759 (2005), arXiv:astro-ph/0409513 [astro-ph].

[64] A. Zonca, L. P. Singer, D. Lenz, M. Reinecke, C. Rosset, E. Hivon, and K. M. Gorski, healpy: equal area pixelization and spherical harmonics transforms for data on the sphere in python, Journal of Open Source Software **4**, 1298 (2019).

[65] A. Hall and A. Taylor, Non-Gaussian likelihood of weak lensing power spectra, Phys. Rev. D **105**, 123527 (2022), arXiv:2202.04095 [astro-ph.CO].

[66] H. Gil-Marín, C. Wagner, F. Fragkoudi, R. Jimenez, and L. Verde, An improved fitting formula for the dark matter bispectrum, **2012**, 047, 1111.4477.

[67] T. Namikawa, B. Bose, F. R. Bouchet, R. Takahashi, and A. Taruya, CMB lensing bi-spectrum: Assessing analytical predictions against full-sky lensing simulations, **99**, 063511, 1812.10635.

[68] S. Das and J. P. Ostriker, Testing a New Analytic Model for Gravitational Lensing Probabilities, Astrophys. J. **645**, 1 (2006), arXiv:astro-ph/0512644 [astro-ph].

[69] A. Barthelemy, S. Codis, and F. Bernardeau, Post-Born corrections to the one-point statistics of (CMB) lensing convergence obtained via large deviation theory, MNRAS **494**, 3368 (2020), arXiv:2002.03625 [astro-ph.CO].

[70] S. J. Osborne, D. Hanson, and O. Doré, Extragalactic foreground contamination in temperature-based cmb lens reconstruction, Journal of Cosmology and Astroparticle Physics **2014** (03), 024–024.

[71] N. Sailer, E. Schaan, and S. Ferraro, Lower bias, lower noise cmb lensing with foreground-hardened estimators, Physical Review D **102**, 10.1103/physrevd.102.063517 (2020).

[72] A. Amara and A. Refregier, Systematic Bias in Cosmic Shear: Beyond the Fisher Matrix, Mon. Not. Roy. Astron. Soc. **391**, 228 (2008), arXiv:0710.5171 [astro-ph].

[73] A. Font-Ribera, P. McDonald, N. Mostek, B. A. Reid, H.-J. Seo, and A. Slosar, DESI and other dark energy experiments in the era of neutrino mass measurements, JCAP **05**, 023, arXiv:1308.4164 [astro-ph.CO].

[74] A. Aghamousa *et al.* (DESI), The DESI Experiment Part I: Science,Targeting, and Survey Design, (2016), arXiv:1611.00036 [astro-ph.IM].

[75] A. Lewis, A. Challinor, and A. Lasenby, Efficient computation of CMB anisotropies in closed FRW models, Astrophys. J. **538**, 473 (2000), arXiv:astro-ph/9911177 [astro-ph].

[76] C. Howlett, A. Lewis, A. Hall, and A. Challinor, CMB power spectrum parameter degeneracies in the era of precision cosmology, JCAP **1204**, 027, arXiv:1201.3654 [astro-ph.CO].

[77] A. Lewis and S. Bridle, Cosmological parameters from CMB and other data: A Monte Carlo approach, Phys. Rev. **D66**, 103511 (2002), arXiv:astro-ph/0205436 [astro-ph].

[78] A. Lewis, Efficient sampling of fast and slow cosmological parameters, Phys. Rev. **D87**, 103529 (2013), arXiv:1304.4473 [astro-ph.CO].

[79] J. Torrado and A. Lewis, Cobaya: Code for Bayesian Analysis of hierarchical physical models, JCAP **05**, 057, arXiv:2005.05290 [astro-ph.IM].

[80] A. Lewis, GetDist: a Python package for analysing Monte Carlo samples, (2019), arXiv:1910.13970 [astro-ph.IM].

[81] R. Allison, P. Caucal, E. Calabrese, J. Dunkley, and T. Louis, Towards a cosmological neutrino mass detection, Phys. Rev. D **92**, 123535 (2015), arXiv:1509.07471 [astro-ph.CO].

[82] J. Peloton, M. Schmittfull, A. Lewis, J. Carron, and O. Zahn, Full covariance of CMB and lensing reconstruction power spectra, Phys. Rev. D **95**, 043508 (2017), arXiv:1611.01446 [astro-ph.CO].

[83] A. Rouhiainen, U. Giri, and M. Münchmeyer, Normalizing flows for random fields in cosmology (2021), arXiv:2105.12024.

[84] R. M. Neal, Taking Bigger Metropolis Steps by Dragging Fast Variables, ArXiv Mathematics e-prints (2005), math/0502099.

[85] V. Böhm, C. Modi, and E. Castorina, Lensing corrections on galaxy-lensing cross correlations and galaxy-galaxy auto correlations, **2020**, 045 (), 1910.06722.

[86] S. S. Boruah, E. Rozo, and P. Fiedorowicz, Map-based cosmology inference with lognormal cosmic shear maps, 2204.13216.

[87] A. Buchalter, M. Kamionkowski, and A. H. Jaffe, The Angular Three-Point Correlation Function in the Quasilinear Regime, **530**, 36, astro-ph/9903486.

[88] P. Coles and B. Jones, A lognormal model for the cosmological mass distribution, **248**, 1.

[89] S. Hilbert, J. Hartlap, and P. Schneider, Cosmic shear covariance: The log-normal approximation, **536**, A85.

[90] A. Lewis and G. Pratten, Effect of lensing non-Gaussianity on the CMB power spectra, JCAP **12**, 003 (2016), arXiv:1608.01263.

[91] C. Carbone, C. Baccigalupi, M. Bartelmann, S. Matarrese, and V. Springel, Lensed CMB temperature and polarization maps from the Millennium Simulation, MNRAS **396**, 668 (2009), arXiv:0810.4145.

[92] M. R. Becker, CALCLENS: weak lensing simulations for large-area sky surveys and second-order effects in cosmic shear power spectra, MNRAS **435**, 115 (2013), arXiv:arXiv:1210.3069.

[93] S. C. Hotinli, J. Meyers, C. Trendafilova, D. Green, and A. van Engelen, The benefits of CMB delensing, JCAP **04** (04), 020, arXiv:2111.15036 [astro-ph.CO].

[94] D. Green, J. Meyers, and A. van Engelen, CMB Delensing Beyond the B Modes, JCAP **12**, 005, arXiv:1609.08143 [astro-ph.CO].

[95] Y. Fantaye, C. Baccigalupi, S. Leach, and A. P. S. Yadav, CMB lensing reconstruction in the presence of diffuse polarized foregrounds, JCAP **12**, 017 (2012), arXiv:1207.0508.

[96] E. Schaan, S. Ferraro, and U. Seljak, Photo-z outlier self-calibration in weak lensing surveys, JCAP **2020**, 001 (2020), arXiv:2007.12795.

5

Summary and Conclusions

This thesis has focused on the optimal lensing reconstruction of CMB lensing, exploring a Bayesian method that goes beyond the traditional quadratic estimator. Central to this work was the optimal lensing reconstruction, map-based delensing, and the development and enhancement of `Delensalot`, a software designed for optimal lensing reconstruction. The discussions covered the limitations of QE, the advantages of iterative lensing reconstruction, and the impact of his novel approach onto the CMB science goals.

This thesis covered five papers that explore the capabilities of this Bayesian method, and we demonstrated successful application of our software in two major collaborations: CMB-S4 and PICO.

In the first and second paper, we developed both CPU and GPU implementations of the deflection operator. The thesis work has pushed the boundaries of computational efficiency by more than one order of magnitude, enabling faster and more precise simulations and analyses that can reach machine precision accuracy.

In the third paper, we discussed the concrete details of the curved sky implementation of this method, and applied it to realistic simulation data of the CMB-S4 collaboration with the aim of measuring r. In the fourth paper, we repeated a similar exercise for the PICO collaboration. In both cases, we could show that the respective experiments' CMB science goals can be reached with our method.

The fifth paper addressed the complexities of non-Gaussian deflection fields and the implications on the lensing reconstruction's accuracy. We found that using a Gaussian likelihood and prior does not prohibit the reconstruction of non-Gaussian features, and that our MAP estimator is less biased compared to the standard QE.

Throughout this research, `Delensalot` was improved in terms of robustness, accuracy, and processing speed, making it not only a powerful tool for researchers but also more accessible and user-friendly. Within the two mentioned collaborations, `Delensalot` was seamlessly integrated into the existing analysis pipelines, demonstrating its utility and effectiveness in real-world research environments and underscores the practical relevance and applicability of the software. The integration of `Delensalot` into the analysis pipelines of major collaborations like CMB-S4 and PICO sets a precedent for broader adoption across other cosmological projects.

Future work could extend beyond CMB studies, incorporating observations from other cosmic phenomena like large-scale structure surveys. Concretely, this could either be done via an additional prior term included in the posterior, or by expanding the likelihood model. Future work could also delve deeper into the statistical properties of correlations beyond the 2-point statistics, exploring their origins and implications for cosmological parameter estimation. This research could enable the development of new statistical tools tailored for impacts from foregrounds that are correlated with the CMB and deflection field itself, enhancing our ability to interpret increasingly complex datasets. `Delensalot` can potentially play a pivotal role in this.

As `Delensalot` becomes more user-friendly and integrated into more research pipelines, there is an excellent opportunity to focus on educational resources and community development. Creating detailed documentation, tutorials, and training programs can empower a new generation of cosmologists to use advanced lensing reconstruction tools effectively. Moreover, open-source contributions and community-driven enhancements to `Delensalot` can foster a more collaborative and innovative research environment.

The outlook for CMB lensing reconstruction is promising. The methodologies developed in this thesis are necessarily needed for unravelling CMB science at current and future experiment noise levels. As we stand on the brink of new astronomical insights with current and upcoming experiments, the continued development of lensing reconstruction techniques will undoubtedly be a key player in discovering new insights in the data analysis.

A

Appendix

A.1 `Delensalot` Language Model

This section shows the complete `Delensalot` language model `DLM`.

```
"""dlensalot_mm.py: Contains classes defining the metamodel of the Dlensalot formalism.
    The metamodel is a structured representation, with the `DLENSALOT_Model` as the main building block.
    We use the attr package. It provides handy ways of validation and defaulting.
"""

class DLENSALOT_Concept:
    """An abstract element base type for the Dlensalot formalism."""
    __metaclass__ = abc.ABCMeta

@attr.s
class DLENSALOT_Chaindescriptor(DLENSALOT_Concept):
    """A root model element type of the Dlensalot formalism.
    This class collects all configurations related to conjugate gradient solver. There are currently not many options for this. Better don't touch it.

    Attributes:
        p0: 0
        p1: type of the conditioner. Can be in ["diag_cl"]
        p2: value of lm_max_ivf[0]
        p3: value of nside of the data
        p4: np.inf
        p5: value of cg_tol
        p6: `tr_cg`: value of cd_solve.tr_cg
        p7: cacher setting
    """
    # TODO change names after testing various chains - can we find better heuristics?
    p0 =  attr.field(validator=chaindescriptor.p0)
    p1 =  attr.field(validator=chaindescriptor.p1)
    p2 =  attr.field(validator=chaindescriptor.p2)
    p3 =  attr.field(validator=chaindescriptor.p3)
    p4 =  attr.field(validator=chaindescriptor.p4)
    p5 =  attr.field(validator=chaindescriptor.p5)
    p6 =  attr.field(validator=chaindescriptor.p6)
    p7 =  attr.field(validator=chaindescriptor.p7)

@attr.s
class DLENSALOT_Stepper(DLENSALOT_Concept):
    """A root model element type of the Dlensalot formalism.
    Defines the stepper function. The stepper function controls how the increment in the likelihood search is added to the current solution.
    Currently, this is pretty much just the harmonicbump class.

    Attributes:
        typ (str): The name of the stepper function
        lmax_qlm (int): maximum `\ell` of the lensing potential reconstruction
        mmax_qlm (int): maximum `m` of the lensing potential reconstruction
        a: TBD
        b: TBD
        xa: TBD
        xb: TBD
    """
    # FIXME this is very 'harmonicbump'-specific.
    typ = attr.field(validator=stepper.typ)
    lmax_qlm =      attr.field(validator=stepper.lmax_qlm) # must match lm_max_qlm -> validator
    mmax_qlm =      attr.field(validator=stepper.mmax_qlm) # must match lm_max_qlm -> validator
    a =   attr.field(validator=stepper.a)
    b =   attr.field(validator=stepper.b)
    xa =  attr.field(validator=stepper.xa)
    xb =  attr.field(validator=stepper.xb)


@attr.s
class DLENSALOT_Job(DLENSALOT_Concept):
    """A root model element type of the Dlensalot formalism.
    delensalot can executte different jobs (QE reconstruction, simulation generation, MAP reconstruction, delensing, ..) which is controlled here.

    Attributes:
        jobs (list[str]): Job identifier(s)
    """
    jobs =attr.field(validator=job.jobs)

@attr.s
class DLENSALOT_Analysis(DLENSALOT_Concept):
    """A root model element type of the Dlensalot formalism.
    This class collects all configurations related to the specific analysis performed on the data.

    Attributes:
        key (str):          reconstruction estimator key
        version (str):      specific configuration for the esimator (e.g. `noMF`, which turns off mean-field subtraction)
        simidxs (np.array[int]):    simulation indices to use for the delensalot job
        simidxs_mf (np.array[int]):simulation indices to use for the calculation of the mean-field
        TEMP_suffix (str):identifier to customize TEMP directory of the analysis
        Lmin (int):         minimum L for reconstructing the lensing potential
        zbounds (tuple[int or str,float]):  latitudinal boundary (-1 to 1), or identifier together with noise level ratio treshold at which lensing reconstruction is perfromed.
        zbounds_len (tuple[int]):  latitudinal extended boundary at which lensing reconstruction is performed, and used for iterative lensing reconstruction
```

```
        lm_max_ivf (tuple[int]):   maximum `\ell` and m for which inverse variance filtering is done
        lm_max_blt (tuple[int]):   maximum `\ell` and m for which B-lensing template is calculated
        mask (list[str]): TBD
        lmin_teb (int):   minimum `\ell` and m of the data which the reconstruction uses, and is set to zero below via the transfer function
        cls_unl (str):    path to the fiducial unlensed CAMB-like CMB data
        cls_len (str):    path to the fiducial lensed CAMB-like CMB data
        cpp (str):        path to the power spectrum of the prior for the iterative reconstruction
        beam (float):     The beam used in the filters
    """
    key = attr.field(validator=analysis.key)
    version =       attr.field(validator=analysis.version) # TODO either make it more useful, or remove
    simidxs =       attr.field(validator=analysis.simidxs)
    simidxs_mf =    attr.field(validator=analysis.simidxs_mf)
    TEMP_suffix =   attr.field(validator=analysis.TEMP_suffix)
    Lmin =attr.field(validator=analysis.Lmin)
    zbounds =       attr.field(validator=analysis.zbounds)
    zbounds_len =   attr.field(validator=analysis.zbounds_len) # TODO rename
    lm_max_ivf =    attr.field(validator=v_filter.lm_max_ivf)
    lm_max_blt =    attr.field(validator=analysis.lm_max_blt)
    mask =attr.field(validator=analysis.mask) # TODO is this used?
    lmin_teb =      attr.field(validator=analysis.lmin_teb)
    cls_unl =       attr.field(validator=analysis.cls_unl)
    cls_len =       attr.field(validator=analysis.cls_len)
    cpp = attr.field(validator=analysis.cpp)
    beam =attr.field(validator=analysis.beam)
    transfunction_desc =    attr.field(validator=analysis.transfunction)


@attr.s
class DLENSALOT_Simulation(DLENSALOT_Concept):
    """A root model element type of the Dlensalot formalism.
    This class collects all configurations related to the input maps, and values can differ from the noise model and analysis.

    Attributes:
        flavour       (str): Can be in ['obs', 'sky', 'unl'] and defines the type of data provided.
        space         (str): Can be in ['map', 'alm', 'cl'] and defines the space of the data provided.
        maps(np.array, optional): These maps will be put into the cacher directly. They are used for settings in which no data is generated or accesed on disk, but directly provided (like in `delensalot.anafast()`) Defaults to DNaV.
        geominfo      (tuple, optional): Lenspyx geominfo descriptor, describes the geominfo of the data provided (e.g. `('healpix', 'nside': 2048)). Defaults to DNaV.
        field         (str, optional): the type of data provided, can be in ['temperature', 'polarization']. Defaults to DNaV.
        libdir        (str, optional): directory of the data provided. Defaults to DNaV.
        libdir_noise  (str, optional): directory of the noise provided. Defaults to DNaV.
        libdir_phi    (str, optional): directory of the lensing potential provided. Defaults to DNaV.
        fns (dict with str with formatter, optional): file names of the data provided. It expects `{'T': <filename{simidx}.something>, 'Q': <filename{simidx}.something>, 'U': <filename{simidx}.something>}`, where `{simidx}` is used by the libraries to format the simulation index into the name. Defaults to DNaV.
        fnsnoise      (dict with str with formatter, optional): file names of the noise provided. It expects `{'T': <filename{simidx}.something>, 'Q': <filename{simidx}.something>, 'U': <filename{simidx}.something>}`, where `{simidx}` is used by the libraries to format the simulation index into the name. Defaults to DNaV.
        fnsP(str with formatter, optional): file names of the lensing potential provided. It expects `<filename{simidx}.something>, where `{simidx}` is used by the libraries to format the simulation index into the name. Defaults to DNaV.
        lmax(int, optional): Maximum l of the data provided. Defaults to DNaV.
        transfunction(np.array, optional): transfer function. Defaults to DNaV.
        nlev(dict, optional): noise level of the individual fields. It expects `{'T': <value>, 'P': <value>}. Defaults to DNaV.
        spin(int, optional): the spin of the data provided. Defaults to 0. Always defaults to 0 for temperature.
        CMB_fn        (str, optional): path+name of the file of the power spectra of the CMB. Defaults to DNaV.
        phi_fn        (str, optional): path+name of the file of the power spectrum of the lensing potential. Defaults to DNaV.
        phi_field     (str, optional): the type of potential provided, can be in ['potential', 'deflection', 'convergence']. This simulation library will automatically rescale the field, if needded. Defaults to DNaV.
        phi_space     (str, optional): can be in ['map', 'alm', 'cl'] and defines the space of the lensing potential provided.. Defaults to DNaV.
        phi_lmax      (_type_, optional): the maximum multipole of the lensing potential. if simulation library perfroms lensing, it is advisable that `phi_lmax` is somewhat larger than `lmax` (+ ~512-1024). Defaults to DNaV.
        epsilon       (float, optional): Lenspyx lensing accuracy. Defaults to 1e-7.
        libdir_suffix(str, optional): defines the directory the simulation data will be stored to, defaults to 'generic'. Helpful if one wants to keep track of different projects.
        CMB_modifier (callable, optional): operation defined in the callable will be applied to each of the input maps/alms/cls
        phi_modifier (callable, optional): operation defined in the callable will be applied to the input phi lms

    """
```

```python
    flavour =        attr.field(validator=data.flavour)
    space =attr.field(validator=data.space)
    maps = attr.field(validator=data.maps)
    geominfo =       attr.field(validator=data.geominfo)
    lenjob_geominfo=attr.field(validator=data.geominfo)
    field =attr.field(validator=data.field)
    libdir =         attr.field(validator=data.libdir)
    libdir_noise =   attr.field(validator=data.libdir_noise)
    libdir_phi =     attr.field(validator=data.libdir_phi)
    fns =  attr.field(validator=data.fns)
    fnsnoise =       attr.field(validator=data.fnsnoise)
    fnsP = attr.field(validator=data.fnsP)
    lmax = attr.field(validator=data.lmax)
    transfunction = attr.field(validator=data.transfunction)
    nlev = attr.field(validator=data.nlev)
    spin = attr.field(validator=data.spin)
    CMB_fn =         attr.field(validator=data.CMB_fn)
    phi_fn =         attr.field(validator=data.phi_fn)
    phi_field =      attr.field(validator=data.phi_field)
    phi_space =      attr.field(validator=data.phi_space)
    phi_lmax =       attr.field(validator=data.phi_lmax)
    epsilon =        attr.field(validator=data.epsilon)
    libdir_suffix = attr.field(default='generic', validator=data.libdir_suffix)
    CMB_modifier =  attr.field(validator=data.modifier)
    phi_modifier =  attr.field(default=lambda x: x)


@attr.s
class DLENSALOT_Noisemodel(DLENSALOT_Concept):
    """A root model element type of the Dlensalot formalism.
    This class collects all configurations related to the noise model used for Wiener-filtering the data.

    Attributes:
        sky_coverage (str):      Can be either 'masked' or 'unmasked'
        spectrum_type (str):     TBD
        OBD (str):      OBD identifier. Can be 'OBD', 'trunc'. Defines how lowest B-modes will be handled.
        nlev_t (float):(central) noise level of temperature data in muK arcmin.
        nlev_p (float):(central) noise level of polarization data in muK arcmin.
        rhits_normalised (str): path to the hits-count map, used to calculate the noise levels, and the mask tracing the noise level. Second entry in tuple is the <inverse hits-count multiplier>.
        geominfo (tuple): geominfo of the noise map
    """
    sky_coverage = attr.field(validator=noisemodel.sky_coverage)
    spectrum_type =attr.field(validator=noisemodel.spectrum_type)
    OBD = attr.field(validator=noisemodel.OBD)
    nlev =attr.field(validator=noisemodel.nlev_t)
    geominfo =      attr.field(validator=noisemodel.ninvjob_geominfo)
    rhits_normalised =       attr.field(validator=noisemodel.rhits_normalised)
    nivt_map =      attr.field(validator=noisemodel.ninvjob_geominfo) # TODO test if it works
    nivp_map =      attr.field(validator=noisemodel.ninvjob_geominfo) # TODO test if it works

@attr.s
class DLENSALOT_Qerec(DLENSALOT_Concept):
    """A root model element type of the Dlensalot formalism.
    This class collects all configurations related to the quadratic estimator reconstruction job.

    Attributes:
        tasks (list[tuple]):         tasks to perfrom. Can be any combination of :code:`calc_phi`, :code:`calc_meanfield`, :code:`calc_blt`
        qlm_type (str):    lensing potential estimator identifier. Can be 'sepTP' or 'jTP'
        cg_tol (float):    tolerance of the conjugate gradient method
        filter_directional (str):   can be either 'isotropic' (unmasked sky) or 'isotropic' (masked sky)
        lm_max_qlm (type): maximum multipole `\ell` and m to reconstruct the lensing potential
        chain (DLENSALOT_Chaindescriptor): configuration of the conjugate gradient method. Configures the chain and preconditioner
        cl_analysis (bool):If tru, performs lensing power spectrum analysis
        blt_pert (bool):   If True, delensing is performed perurbitivly (recommended)

    """

    tasks =        attr.field(validator=qerec.tasks)
    qlm_type =     attr.field(validator=qerec.qlms)
    cg_tol =       attr.field(validator=qerec.cg_tol)
    filter_directional =    attr.field(validator=qerec.filter_directional)
    lm_max_qlm =   attr.field(validator=qerec.lm_max_qlm)
    chain =        attr.field(default=DLENSALOT_Chaindescriptor(), validator=qerec.chain)
    cl_analysis =  attr.field(validator=qerec.cl_analysis) # TODO make this useful or remove
    blt_pert =     attr.field(validator=qerec.btemplate_perturbative_lensremap)

@attr.s
class DLENSALOT_Itrec(DLENSALOT_Concept):
    """A root model element type of the Dlensalot formalism.
    This class collects all configurations related to the iterative reconstruction job.

    Attributes:
        tasks (list[str]): tasks to perfrom. Can be any combination of :code:`calc_phi`, :code:`calc_meanfield`, :code:`calc_blt`
        itmax (int):       maximum number of iterations
```

```
        cg_tol (float):     tolerance of the conjugate gradient method
        iterator_typ (str):mean-field handling identifier. Can be either 'const_mf' or 'pert_mf'
        chain (DLENSALOT_Chaindescriptor): configuration for the conjugate gradient solver
        filter_directional (str):    can be either 'isotropic' (unmasked sky) or 'isotropic' (masked sky)
        lenjob_geominfo (str):       can be 'healpix_geominfo', 'thin_gauss' or 'pbdGeometry'
        lenjob_pbgeominfo (str):     can be 'healpix_geominfo', 'thin_gauss' or 'pbdGeometry'
        lm_max_unl (tuple[int]):     maximum multipoles `\ell` and m for reconstruction the unlensed CMB
        lm_max_qlm (tuple[int]):     maximum multipoles L and m for reconstruction the lensing potential
        mfvar (str):        path to precalculated mean-field, to be used instead
        soltn_cond (type): TBD
        stepper (DLENSALOT_STEPPER):configuration for updating the current likelihood iteration point with the likelihood gradient

    """
    tasks =         attr.field(validator=itrec.tasks)
    itmax =         attr.field(validator=itrec.itmax)
    cg_tol =        attr.field(validator=itrec.cg_tol)
    iterator_typ = attr.field(validator=itrec.iterator_type) # TODO rename
    chain =         attr.field(default=DLENSALOT_Chaindescriptor(), validator=itrec.chain)
    filter_directional =    attr.field(validator=itrec.filter_directional)
    lenjob_geominfo =       attr.field(validator=itrec.lenjob_geominfo)
    lenjob_pbdgeominfo =    attr.field(validator=itrec.lenjob_pbgeominfo)
    lm_max_unl =    attr.field(validator=itrec.lm_max_unl)
    lm_max_qlm =    attr.field(validator=itrec.lm_max_qlm)
    mfvar =         attr.field(validator=itrec.mfvar) # TODO rename and check if it still works
    soltn_cond =    attr.field(validator=itrec.soltn_cond)
    stepper =       attr.field(default=DLENSALOT_Stepper(), validator=itrec.stepper)
    epsilon =       attr.field(validator=data.epsilon)

@attr.s
class DLENSALOT_Mapdelensing(DLENSALOT_Concept):
    """A root model element type of the Dlensalot formalism.
    This class collects all configurations related to the internal map delensing job.

    Attributes:
        data_from_CFS (bool):    if set, use B-lensing templates located at the $CFS directory instead of the $TEMP directory\n
        edges (np.array):        binning to calculate the (delensed) power spectrum on\n
        dlm_mod (bool):if set, modfies the lensing potential before calculating the B-lensing template\n
        iterations (list[int]): which iterations to calculate delensed power spectrum for\n
        nlevels (list[float]):  noiselevel ratio treshold up to which the maps are delensed, uses the rhits_normalized map to generate masks.
        lmax (int):     maximum multipole to calculate the (delensed) power spectrum\n
        Cl_fid (type): fiducial power spectrum, and needed for template calculation of the binned power spectrum package\n
        libdir_it (type):        TBD\n
        binning (type):can be either 'binned' or 'unbinned'. If 'unbinned', overwrites :code:`edges` and calculates power spectrum for each multipole\n
        spectrum_calculator (package): name of the package of the power spectrum calculator. Can be 'healpy' if :code:`binning=unbinned`\n
        masks_fn (list[str]):    the sky patches to calculate the power spectra on. Note that this is different to using `nlevels`. Here, no tresholds are calculated, but masks are used 'as is' for delensing.\n
        basemap (str): the delensed map Bdel is calculated as Bdel = basemap - blt. Basemap can be two things: 'obs' or 'lens', where 'obs' will use the observed sky map, and lens will use the pure B-lensing map.
    """

    data_from_CFS =attr.field(validator=mapdelensing.data_from_CFS)
    edges =         attr.field(validator=mapdelensing.edges)
    dlm_mod =       attr.field(validator=mapdelensing.dlm_mod)
    iterations =    attr.field(validator=mapdelensing.iterations)
    nlevels =       attr.field(validator=mapdelensing.nlevels)
    lmax =attr.field(validator=mapdelensing.lmax)
    Cl_fid =        attr.field(validator=mapdelensing.Cl_fid)
    libdir_it =     attr.field(validator=mapdelensing.libdir_it)
    binning =       attr.field(validator=mapdelensing.binning)
    spectrum_calculator =    attr.field(validator=mapdelensing.spectrum_calculator)
    masks_fn =      attr.field(validator=mapdelensing.masks)
    basemap =       attr.field(validator=mapdelensing.basemap)

@attr.s
class DLENSALOT_Phianalysis(DLENSALOT_Concept):
    """A root model element type of the Dlensalot formalism.
    This class collects all configurations related to the internal map delensing job.

    Attributes:
        custom_WF_TEMP (str):    Path to the dir of an exisiting WF. fn must be 'WFemp_%s_simall%s_itall%s_avg.npy'\n
    """

    custom_WF_TEMP =         attr.field()

@attr.s
class DLENSALOT_OBD(DLENSALOT_Concept):
    """A root model element type of the Dlensalot formalism.
    This class collects all configurations related to the overlapping B-mode deprojection.

    Attributes:
```

```
        libdir (str):       path to the OBD matrix
        rescale (float):    rescaling of OBD matrix amplitude. Useful if matrix already calculated, but noiselevel changed
        tpl (type):function name for calculating OBD matrix
        nlev_dep (float):   deprojection factor, or, strength of B-mode deprojection
    """
    libdir =        attr.field(validator=obd.libdir)
    rescale =       attr.field(validator=obd.rescale)
    tpl = attr.field(validator=obd.tpl)
    nlev_dep =      attr.field(validator=obd.nlev_dep)

@attr.s
class DLENSALOT_Config(DLENSALOT_Concept):
    """A root model element type of the Dlensalot formalism.
    This class collects all configurations related to general behaviour to the operating system.

    Attributes:
        outdir_plot_root (str): root path for the plots to be stored at
        outdir_plot_rel (str):  relative path folder for the plots to be stored at
    """
    outdir_plot_root =      attr.field(default=opj(os.environ['HOME'], 'plots'))
    outdir_plot_rel =       attr.field(default='')

@attr.s
class DLENSALOT_Meta(DLENSALOT_Concept): # TODO do we really need a Meta?
    """A root model element type of the Dlensalot formalism.
    This class collects all configurations related to internal behaviour of delensalot.

    Attributes:
        version (str):  version control of the delensalot model
    """
    version =       attr.field(validator=attr.validators.instance_of(int))


@attr.s
class DLENSALOT_Computing(DLENSALOT_Concept):
    """A root model element type of the Dlensalot formalism.
    This class collects all configurations related to the usage of computing resources.

    Attributes:
        OMP_NUM_THREADS (int):  number of threads used per Job
    """
    OMP_NUM_THREADS =       attr.field(validator=computing.OMP_NUM_THREADS)


@attr.s
class DLENSALOT_Model(DLENSALOT_Concept):
    """A root model element type of the Dlensalot formalism.

    Attributes:
        defaults_to (str):      Identifier for default-dictionary if user hasn't specified value in configuration file
        meta (DLENSALOT_Meta): configurations related to internal behaviour of delensalot
        job (DLENSALOT_Job):    delensalot can executte different jobs (QE reconstruction, simulation generation, MAP reconstruction, delensing, analyse_phi) which is controlled here
        analysis (DLENSALOT_Analysis):  configurations related to the specific analysis performed on the data
        data (DLENSALOT_Data): configurations related to the input CMB maps
        noisemodel (DLENSALOT_Noisemodel):  configurations related to the noise model used for Wiener-filtering the data
        qerec (DLENSALOT_Qerec):        configurations related to the quadratic estimator reconstruction job
        itrec (DLENSALOT_Itrec):        configurations related to the iterative reconstruction job
        madel (DLENSALOT_Mapdelensing): configurations related to the internal map delensing job
        config (DLENSALOT_Config):      configurations related to general behaviour to the operating system
        computing (DLENSALOT_Computing):    configurations related to the usage of computing resources
        obd (DLENSALOT_OBD):    configurations related to the overlapping B-mode deprojection
        phana (DLENSALOT_Phyanalysis):  configurations related to the simple power spectrum analaysis of phi

    """

    defaults_to =   attr.field(default='default_CMBS4_fullsky_polarization')
    validate_model =        attr.field(default=True)
    meta =attr.field(default=DLENSALOT_Meta(), validator=model.meta)
    job = attr.field(default=DLENSALOT_Job(), validator=model.job)
    analysis =      attr.field(default=DLENSALOT_Analysis(), validator=model.analysis)
    simulationdata =        attr.field(default=DLENSALOT_Simulation(), validator=model.data)
    noisemodel =    attr.field(default=DLENSALOT_Noisemodel(), validator=model.noisemodel)
    qerec =         attr.field(default=DLENSALOT_Qerec(), validator=model.qerec)
    itrec =         attr.field(default=DLENSALOT_Itrec(), validator=model.itrec)
    madel =         attr.field(default=DLENSALOT_Mapdelensing(), validator=model.madel)
    config =        attr.field(default=DLENSALOT_Config(), validator=model.config)
    computing =     attr.field(default=DLENSALOT_Computing(), validator=model.computing)
    obd = attr.field(default=DLENSALOT_OBD(), validator=model.obd)
    phana =         attr.field(default=DLENSALOT_Phianalysis())
```

Bibliography

[1] S. Belkner, A. J. Duivenvoorden, J. Carron, N. Schaeffer, and M. Reinecke. "cunuSHT: GPU Accelerated Spherical Harmonic Transforms on Arbitrary Pixelizations". In: (2024). arXiv: 2406.14542 [astro-ph.IM].

[2] S. Belkner et al. "CMB-S4: Iterative Internal Delensing and r Constraints". In: *Astrophys. J.* 964 (2024), p. 148. DOI: 10.3847/1538-4357/ad2351. arXiv: 2310.06729 [astro-ph.CO].

[3] O. Darwish, S. Belkner, L. Legrand, J. Carron, and G. Fabbian. "Non-Gaussian deflections in iterative optimal CMB lensing reconstruction". In: (2024). arXiv: 2407.00228 [astro-ph.CO].

[4] M. Reinecke, S. Belkner, and J. Carron. "Improved cosmic microwave background (de-)lensing using general spherical harmonic transforms". In: *Astron. Astrophys.* 678 (2023), A165. DOI: 10.1051/0004-6361/202346717. arXiv: 2304.10431 [astro-ph.CO].

[5] R. Aurlien, M. Remazeilles, S. Belkner, J. Carron, J. Delabrouille, H. K. Eriksen, R. Flauger, U. Fuskeland, M. Galloway, K. M. Gorski, et al. "Foreground separation and constraints on primordial gravitational waves with the PICO space mission". In: *JCAP* 06 (2023), p. 034. DOI: 10.1088/1475-7516/2023/06/034. arXiv: 2211.14342 [astro-ph.CO].

[6] R. A. Alpher, H. Bethe, and G. Gamow. "The Origin of Chemical Elements". In: *Phys. Rev.* 73 (7 Apr. 1948), pp. 803–804. DOI: 10.1103/PhysRev.73.803. URL: https://link.aps.org/doi/10.1103/PhysRev.73.803.

[7] R. H. Dicke, P. J. E. Peebles, P. G. Roll, and D. T. Wilkinson. "Cosmic Black-Body Radiation." In: ApJ 142 (July 1965), pp. 414–419. DOI: 10.1086/148306.

[8] G. Lemaı tre. "The Beginning of the World from the Point of View of Quantum Theory." In: Nature 127.3210 (May 1931), p. 706. DOI: 10.1038/127706b0.

[9] Alan H. Guth. "The Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems". In: *Phys. Rev. D* 23 (1981). Ed. by Li-Zhi Fang and R. Ruffini, pp. 347–356. DOI: 10.1103/PhysRevD.23.347.

[10] Oliver Heaviside. "A Gravitational and Electromagnetic Analogy". In: *The Electrician, vol.31* (Jan. 1893), pp. 281–282 (part I), p. 359 (part II).

[11] Albert Einstein. “Näherungsweise Integration der Feldgleichungen der Gravitation”. In: *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften* (Jan. 1916), pp. 688–696.

[12] George R. Blumenthal, S. M. Faber, Joel R. Primack, and Martin J. Rees. “Formation of Galaxies and Large Scale Structure with Cold Dark Matter”. In: *Nature* 311 (1984). Ed. by M. A. Srednicki, pp. 517–525. DOI: `10.1038/311517a0`.

[13] Antony Lewis and Anthony Challinor. “Weak gravitational lensing of the CMB”. In: *Phys. Rept.* 429 (2006), pp. 1–65. DOI: `10.1016/j.physrep.2006.03.002`. arXiv: `astro-ph/0601594`.

[14] Mathew Robertson and Antony Lewis. “How to detect lensing rotation”. In: *JCAP* 08 (2023), p. 048. DOI: `10.1088/1475-7516/2023/08/048`. arXiv: `2303.13313 [astro-ph.CO]`.

[15] Antón Baleato Lizancos, Anthony Challinor, and Julien Carron. “Limitations of CMB B-mode template delensing”. In: *Phys. Rev. D* 103.2 (2021), p. 023518. DOI: `10.1103/PhysRevD.103.023518`. arXiv: `2010.14286 [astro-ph.CO]`.

[16] M. Tristram, A. J. Banday, K. M. Górski, R. Keskitalo, C. R. Lawrence, K. J. Andersen, R. B. Barreiro, J. Borrill, L. P. L. Colombo, H. K. Eriksen, R. Fernandez-Cobos, T. S. Kisner, E. Martínez-González, B. Partridge, D. Scott, T. L. Svalheim, and I. K. Wehus. “Improved limits on the tensor-to-scalar ratio using BICEP and *Planck* data”. In: *Phys. Rev. D* 105 (8 Apr. 2022), p. 083524. DOI: `10.1103/PhysRevD.105.083524`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.105.083524`.

[17] Matias Zaldarriaga and Uros Seljak. “Reconstructing projected matter density from cosmic microwave background”. In: *Phys. Rev. D* 59 (1999), p. 123507. DOI: `10.1103/PhysRevD.59.123507`. arXiv: `astro-ph/9810257`.

[18] Takemi Okamoto and Wayne Hu. “CMB lensing reconstruction on the full sky”. In: *Phys. Rev. D* 67 (2003), p. 083002. DOI: `10.1103/PhysRevD.67.083002`. arXiv: `astro-ph/0301031`.

[19] Antony Lewis, Anthony Challinor, and Duncan Hanson. “The shape of the CMB lensing bispectrum”. In: *Journal of Cosmology and Astroparticle Physics* 2011.03 (Mar. 2011), pp. 018–018. ISSN: 1475-7516. DOI: `10.1088/1475-7516/2011/03/018`. URL: `http://dx.doi.org/10.1088/1475-7516/2011/03/018`.

[20] Julien Carron and Antony Lewis. “Maximum a posteriori CMB lensing reconstruction”. In: *Phys. Rev. D* 96.6 (2017), p. 063510. DOI: `10.1103/PhysRevD.96.063510`. arXiv: `1704.08230 [astro-ph.CO]`.

[21] Michael H. Kesden, Asantha Cooray, and Marc Kamionkowski. “Lensing reconstruction with CMB temperature and polarization”. In: *Phys. Rev. D* 67 (2003), p. 123507. DOI: `10.1103/PhysRevD.67.123507`. arXiv: `astro-ph/0302536`.

[22] Elizabeth E. Jenkins, Aneesh V. Manohar, Wouter J. Waalewijn, and Amit P. S. Yadav. "Gravitational Lensing of the CMB: a Feynman Diagram Approach". In: *Phys. Lett. B* 736 (2014), pp. 6–10. DOI: `10.1016/j.physletb.2014.07.002`. arXiv: `1403.2386 [astro-ph.CO]`.

[23] Louis Legrand and Julien Carron. "Robust and efficient CMB lensing power spectrum from polarization surveys". In: *Phys. Rev. D* 108.10 (2023), p. 103516. DOI: `10.1103/PhysRevD.108.103516`. arXiv: `2304.02584 [astro-ph.CO]`.

[24] Giulio Fabbian, Antony Lewis, and Dominic Beck. "CMB lensing reconstruction biases in cross-correlation with large-scale structure probes". In: *JCAP* 10 (2019), p. 057. DOI: `10.1088/1475-7516/2019/10/057`. arXiv: `1906.08760 [astro-ph.CO]`.

[25] Antony Lewis and Geraint Pratten. "Effect of lensing non-Gaussianity on the CMB power spectra". In: *JCAP* 12 (2016), p. 003. DOI: `10.1088/1475-7516/2016/12/003`. arXiv: `1608.01263 [astro-ph.CO]`.

[26] Vanessa Böhm, Marcel Schmittfull, and Blake D. Sherwin. "Bias to CMB lensing measurements from the bispectrum of large-scale structure". In: *Phys. Rev. D* 94.4 (2016), p. 043519. DOI: `10.1103/PhysRevD.94.043519`. arXiv: `1605.01392 [astro-ph.CO]`.

[27] Wayne Hu and Takemi Okamoto. "Mass reconstruction with cmb polarization". In: *Astrophys. J.* 574 (2002), pp. 566–574. DOI: `10.1086/341110`. arXiv: `astro-ph/0111606`.

[28] Ethan Anderes. "Decomposing CMB lensing power with simulation". In: *Physical Review D* 88.8 (Oct. 2013). ISSN: 1550-2368. DOI: `10.1103/physrevd.88.083517`. URL: `http://dx.doi.org/10.1103/PhysRevD.88.083517`.

[29] Emmanuel Schaan and Simone Ferraro. "Foreground-Immune Cosmic Microwave Background Lensing with Shear-Only Reconstruction". In: *Phys. Rev. Lett.* 122.18 (2019), p. 181301. DOI: `10.1103/PhysRevLett.122.181301`. arXiv: `1804.06403 [astro-ph.CO]`.

[30] Omar Darwish, Blake D. Sherwin, Noah Sailer, Emmanuel Schaan, and Simone Ferraro. "Optimizing foreground mitigation for CMB lensing with combined multifrequency and geometric methods". In: *Phys. Rev. D* 107.4 (2023), p. 043519. DOI: `10.1103/PhysRevD.107.043519`. arXiv: `2111.00462 [astro-ph.CO]`.

[31] Toshiya Namikawa, Duncan Hanson, and Ryuichi Takahashi. "Bias-hardened CMB lensing". In: *Monthly Notices of the Royal Astronomical Society* 431.1 (Feb. 2013), pp. 609–620. ISSN: 1365-2966. DOI: `10.1093/mnras/stt195`. URL: `http://dx.doi.org/10.1093/mnras/stt195`.

[32] N. Aghanim et al. "Planck 2018 results. VIII. Gravitational lensing". In: *Astron. Astrophys.* 641 (2020), A8. DOI: `10.1051/0004-6361/201833886`. arXiv: `1807.06210 [astro-ph.CO]`.

[33] Z. Pan et al. "Measurement of gravitational lensing of the cosmic microwave background using SPT-3G 2018 data". In: *Phys. Rev. D* 108.12 (2023), p. 122005. DOI: `10.1103/PhysRevD.108.122005`. arXiv: `2308.11608 [astro-ph.CO]`.

[34] Emmanuel Schaan et al. “Atacama Cosmology Telescope: Combined kinematic and thermal Sunyaev-Zel’dovich measurements from BOSS CMASS and LOWZ halos”. In: *Phys. Rev. D* 103.6 (2021), p. 063513. DOI: `10.1103/PhysRevD.103.063513`. arXiv: `2009.05557 [astro-ph.CO]`.

[35] P. de Bernardis et al. “Mapping the cmb sky: the boomerang experiment”. In: *New Astron. Rev.* 43 (1999), pp. 289–296. DOI: `10.1016/S1387-6473(99)00022-6`. arXiv: `astro-ph/9911461`.

[36] C. Chang et al. “Joint analysis of Dark Energy Survey Year 3 data and CMB lensing from SPT and Planck. II. Cross-correlation measurements and cosmological constraints”. In: *Phys. Rev. D* 107.2 (2023), p. 023530. DOI: `10.1103/PhysRevD.107.023530`. arXiv: `2203.12440 [astro-ph.CO]`.

[37] Zhuoqi (Jackie) Zhang, Chihway Chang, Patricia Larsen, Lucas F. Secco, and Joe Zuntz. “Transitioning from Stage-III to Stage-IV: cosmology from galaxy×CMB lensing and shear×CMB lensing”. In: *Mon. Not. Roy. Astron. Soc.* 514.2 (2022), pp. 2181–2197. DOI: `10.1093/mnras/stac1407`. arXiv: `2111.04917 [astro-ph.CO]`.

[38] Federico Bianchini et al. “Toward a tomographic analysis of the cross-correlation between Planck CMB lensing and H-ATLAS galaxies”. In: *Astrophys. J.* 825.1 (2016), p. 24. DOI: `10.3847/0004-637X/825/1/24`. arXiv: `1511.05116 [astro-ph.CO]`.

[39] Frank J. Qu, Blake D. Sherwin, Omar Darwish, Toshiya Namikawa, and Mathew S. Madhavacheril. “Probing early structure and model-independent neutrino mass with high-redshift CMB lensing mass maps”. In: *Phys. Rev. D* 107.12 (2023), p. 123540. DOI: `10.1103/PhysRevD.107.123540`. arXiv: `2208.04253 [astro-ph.CO]`.

[40] T. M. C. Abbott et al. “Joint analysis of Dark Energy Survey Year 3 data and CMB lensing from SPT and Planck. III. Combined cosmological constraints”. In: *Phys. Rev. D* 107.2 (2023), p. 023531. DOI: `10.1103/PhysRevD.107.023531`. arXiv: `2206.10824 [astro-ph.CO]`.

[41] Abhishek S. Maniyar, Yacine Ali-Haïmoud, Julien Carron, Antony Lewis, and Mathew S. Madhavacheril. “Quadratic estimators for CMB weak lensing”. In: *Phys. Rev. D* 103.8 (2021), p. 083524. DOI: `10.1103/PhysRevD.103.083524`. arXiv: `2101.12193 [astro-ph.CO]`.

[42] Christopher M. Hirata and Uros Seljak. “Reconstruction of lensing from the cosmic microwave background polarization”. In: *Phys. Rev. D* 68 (2003), p. 083002. DOI: `10.1103/PhysRevD.68.083002`. arXiv: `astro-ph/0306354`.

[43] Sebastian Belkner, Julien Carron, Louis Legrand, Caterina Umiltà, Clem Pryke, and Colin Bischoff. “CMB-S4: Iterative Internal Delensing and r Constraints”. In: *Astrophys. J.* 964.2 (2024), p. 148. DOI: `10.3847/1538-4357/ad2351`. arXiv: `2310.06729 [astro-ph.CO]`.

[44] P. A. R. Ade et al. "Planck 2015 results. XV. Gravitational lensing". In: *Astron. Astrophys.* 594 (2016), A15. DOI: `10.1051/0004-6361/201525941`. arXiv: `1502.01591 [astro-ph.CO]`.

[45] Louis Legrand and Julien Carron. "Lensing power spectrum of the cosmic microwave background with deep polarization experiments". In: *Phys. Rev. D* 105.12 (2022), p. 123519. DOI: `10.1103/PhysRevD.105.123519`. arXiv: `2112.05764 [astro-ph.CO]`.

[46] Kevork N. Abazajian et al. "CMB-S4 Science Book, First Edition". In: (Oct. 2016). arXiv: `1610.02743 [astro-ph.CO]`.

[47] Wuhyun Sohn and James Fergusson. "CMB-S4 forecast on the primordial non-Gaussianity parameter of feature models". In: *Phys. Rev. D* 100.6 (2019), p. 063536. DOI: `10.1103/PhysRevD.100.063536`. arXiv: `1902.01142 [astro-ph.CO]`.

[48] J.R Fergusson, M Liguori, and E.P.S Shellard. "The CMB bispectrum". In: *Journal of Cosmology and Astroparticle Physics* 2012.12 (Dec. 2012), pp. 032–032. ISSN: 1475-7516. DOI: `10.1088/1475-7516/2012/12/032`. URL: `http://dx.doi.org/10.1088/1475-7516/2012/12/032`.

[49] Julien Carron and Antony Lewis. "Spherical bispectrum expansion and quadratic estimators". In: *JCAP* 07 (2024), p. 067. DOI: `10.1088/1475-7516/2024/07/067`. arXiv: `2404.16797 [astro-ph.CO]`.

[50] Antony Lewis and Anthony Challinor. *CAMB: Code for Anisotropies in the Microwave Background*. Astrophysics Source Code Library, record ascl:1102.026. Feb. 2011.

[51] Julien Lesgourgues. *The Cosmic Linear Anisotropy Solving System (CLASS) I: Overview*. 2011. arXiv: `1104.2932 [astro-ph.IM]`. URL: `https://arxiv.org/abs/1104.2932`.

[52] Giulio Fabbian, Matteo Calabrese, and Carmelita Carbone. "CMB weak-lensing beyond the Born approximation: a numerical approach". In: *JCAP* 02 (2018), p. 050. DOI: `10.1088/1475-7516/2018/02/050`. arXiv: `1702.03317 [astro-ph.CO]`.

[53] Martin Reinecke, Sebastian Belkner, and Julien Carron. "Improved cosmic microwave background (de-)lensing using general spherical harmonic transforms". In: *Astron. Astrophys.* 678 (2023), A165. DOI: `10.1051/0004-6361/202346717`. arXiv: `2304.10431 [astro-ph.CO]`.

[54] Sigurd K. Næss and Thibaut Louis. "Lensing simulations by Taylor expansion — not so inefficient after all". In: *JCAP* 09 (2013), p. 001. DOI: `10.1088/1475-7516/2013/09/001`. arXiv: `1307.0719 [astro-ph.CO]`.

[55] Antony Lewis. "Lensed CMB simulation and parameter estimation". In: *Phys. Rev. D* 71 (2005), p. 083008. DOI: `10.1103/PhysRevD.71.083008`. arXiv: `astro-ph/0502469`.

[56] R. K. Sachs and A. M. Wolfe. "Perturbations of a Cosmological Model and Angular Variations of the Microwave Background". In: ApJ 147 (Jan. 1967), p. 73. DOI: `10.1086/148982`.

[57] M. J. Rees and D. W. Sciama. “Large-scale Density Inhomogeneities in the Universe”. In: Nature 217.5128 (Feb. 1968), pp. 511–516. DOI: `10.1038/217511a0`.

[58] R. A. Sunyaev and Ya. B. Zeldovich. “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* 4 (Nov. 1972), p. 173.

[59] Ben Thorne, Jo Dunkley, David Alonso, and Sigurd Naess. “The Python Sky Model: software for simulating the Galactic microwave sky”. In: *Mon. Not. Roy. Astron. Soc.* 469.3 (2017), pp. 2821–2833. DOI: `10.1093/mnras/stx949`. arXiv: `1608.02841 [astro-ph.CO]`.

[60] Jean-François Cardoso, Maude Le Jeune, Jacques Delabrouille, Marc Betoule, and Guillaume Patanchon. “Component Separation With Flexible Models—Application to Multichannel Astrophysical Observations”. In: *IEEE Journal of Selected Topics in Signal Processing* 2.5 (Nov. 2008), pp. 735–746. DOI: `10.1109/JSTSP.2008.2005346`.

[61] J. Delabrouille, J. -F. Cardoso, M. Le Jeune, M. Betoule, G. Fay, and F. Guilloux. “A full sky, low foreground, high resolution CMB map from WMAP”. In: A&A 493.3 (Jan. 2009), pp. 835–857. DOI: `10.1051/0004-6361:200810514`. arXiv: `0807.0773 [astro-ph]`.

[62] E. Martínez-González, J. M. Diego, P. Vielva, and J. Silk. “Cosmic microwave background power spectrum estimation and map reconstruction with the expectation-maximization algorithm”. In: MNRAS 345.4 (Nov. 2003), pp. 1101–1109. DOI: `10.1046/j.1365-2966.2003.06885.x`. arXiv: `astro-ph/0302094 [astro-ph]`.

[63] H. K. Eriksen, C. Dickinson, C. R. Lawrence, C. Baccigalupi, A. J. Banday, K. M. Górski, F. K. Hansen, P. B. Lilje, E. Pierpaoli, M. D. Seiffert, K. M. Smith, and K. Vanderlinde. “Cosmic Microwave Background Component Separation by Parameter Estimation”. In: ApJ 641.2 (Apr. 2006), pp. 665–682. DOI: `10.1086/500499`. arXiv: `astro-ph/0508268 [astro-ph]`.

[64] Jesus Torrado and Antony Lewis. “Cobaya: Code for Bayesian Analysis of hierarchical physical models”. In: *JCAP* 05 (2021), p. 057. DOI: `10.1088/1475-7516/2021/05/057`. arXiv: `2005.05290 [astro-ph.IM]`.

[65] Antony Lewis. “Efficient sampling of fast and slow cosmological parameters”. In: *Phys. Rev. D* 87.10 (2013), p. 103529. DOI: `10.1103/PhysRevD.87.103529`. arXiv: `1304.4473 [astro-ph.CO]`.