LUDWIGS-MAXIMILIANS-UNIVERSITÄT MÜNCHEN
Fakultät für Physik

Bayesian Inference of Early-Universe Signals

Sebastian Dorn

Dissertation for attainment of the academic degree of

Doctor rerum naturalium (Dr. rer. nat.)

LUDWIGS-MAXIMILIANS-UNIVERSITÄT MÜNCHEN
Fakultät für Physik

Bayesian Inference of Early-Universe Signals

Sebastian Dorn

Dissertation for attainment of the academic degree of

Doctor rerum naturalium (Dr. rer. nat.)

Place & date:	Munich, December 8, 2015
1st assessor:	PD Dr. Torsten A. Enßlin
2nd assessor:	Prof. Dr. Jochen Weller
Oral examination:	February 5, 2016

Begriff ist Summe, Idee Resultat der Erfahrung; jene zu ziehen, wird Verstand, dieses zu erfassen, Vernunft erfordert.

Johann Wolfgang von Goethe
Einzelheiten, Maximen und Reflexionen (1833) [1]

Abstract (GER). Das Kerngebiet dieser Dissertation bildet die Entwicklung und Anwendung von Bayesschen Techniken zur Deduktion von Signalen im frühen Universum sowie die Entwicklung von mathematischen Werkzeugen zur Informationsgewinnung.

Eine wichtige Informationsquelle für das frühe Universum ist das ursprüngliche Skalarpotential sowie seine Statistik. Deren Rekonstruktion aus Messungen der kosmischen Hintergrundstrahlung wird hier präsentiert. Dabei wird das außergewöhnlich große, inverse Problem in eine Vielzahl von Unterproblemen aufgeteilt, von denen jedes durch einen optimalen linearen Filter gelöst wird.

Einmal in Besitz einer Rekonstruktion des ursprünglichen Skalarpotentials und seiner zugehörigen Korrelationsstruktur, ermöglichen diese sodann einen direkten Rückschluss auf die zugrundeliegende Physik des frühen Universums. Zum Beispiel können kleinste Abweichungen des Skalarpotentials von Gauß'scher Statistik benutzt werden, um Parameter von Inflationsmodellen zu deduzieren. Hierzu wird eine Methode entwickelt und verifiziert, die simultan eine Rekonstruktion der spektralen Leistungsdichte liefert. Um die Methode so analytisch wie möglich zu halten und dadurch ressourcenschonend zu gestalten, wird eine Näherung der a posteriori Wahrscheinlichkeit einschließlich der Taylorentwicklung einer Determinante eingeführt.

Die Berechnung einer Matrix-Determinante ist ebenfalls in vielen anderen Bayesschen Methoden, innerhalb und auch außerhalb der Kosmologie, essentiell. Dies kann sehr anspruchsvoll sein, wenn eine Taylorentwicklung fehlschlägt. Insbesondere wenn man mit sehr großen Datenmengen konfrontiert ist, ist das problematisch, weil in diesem Fall Matrizen aufgrund ihrer Größe oft nur indirekt durch Computerroutinen repräsentiert werden. Diese Computerroutinen stellen die Wirkung der Matrizen dar. Um das Determinantenproblem zu lösen wird der Logarithmus der Determinante als Integral reformuliert und durch bekannte Stichprobenziehungstechniken bestimmt.

Für jede hier präsentierte Methode sowie für jede Datenanalyse von wissenschaftlichen Experimenten ist eine korrekte Kalibration des Messinstruments unabdingbare Voraussetzung. Aus diesem Grund steht eine Weiterentwicklung der Theorie der Selbst-Kalibrierung — die Deduktion von Signal und Kalibration aus demselben Datensatz — den anderen Kapiteln voran. Die Weiterentwicklung basiert darauf, sukzessive mehr und mehr Anteile der Kalibrationsunsicherheit in die Gleichungen für die Signalrekonstruktion zu absorbieren. Resultierend daraus erhält man den 'Calibration-Uncertainty Renormalized Estimator' als Lösung einer gekoppelten Differentialgleichung.

Abstract (EN). This thesis focuses on the development and application of Bayesian inference techniques for early-Universe signals and on the advancement of mathematical tools for information retrieval.

A crucial quantity required to gain information from the early Universe is the primordial scalar potential and its statistics. We reconstruct this scalar potential from cosmic microwave background data. Technically, the inference is done by splitting the large inverse problem of such a reconstruction into many, each of them solved by an optimal linear filter.

Once the primordial scalar potential and its correlation structure have been obtained the underlying physics can be directly inferred from it. Small deviations of the scalar potential from Gaussianity, for instance, can be used to study parameters of inflationary models. A method to infer such parameters from non-Gaussianity is presented. To avoid expensive numerical techniques the method is kept analytical as far as possible. This is achieved by introducing an approximation of the desired posterior probability including a Taylor expansion of a matrix determinant.

The calculation of a determinant is also essential in many other Bayesian approaches, both apart from and within cosmology. In cases where a Taylor approximation fails, its evaluation is usually challenging. The evaluation is in particular difficult, when dealing with big data, where matrices are to huge to be accessible directly, but need to be represented indirectly by a computer routine implementing the action of the matrix. To solve this problem, we develop a method to calculate the determinant of a matrix by using well-known sampling techniques and an integral representation of the log-determinant.

The prerequisite for the presented methods as well as for every data analysis of scientific experiments is a proper calibration of the measurement device. Therefore we advance the theory of self-calibration at the beginning of the thesis to infer signal and calibration simultaneously from data. This is achieved by successively absorbing more and more portions of calibration uncertainty into the signal inference equations. The result, the Calibration-Uncertainty Renormalized Estimator, follows from the solution of a coupled differential equation.

Contents

List of Figures

List of Tables

1. Introduction

1.1. The early Universe

This introductory section sketches some milestones of the history of the early[1] and late Universe, which are described in great detail in many of the standard textbooks on cosmology. The argumentation and notation mainly follow references [2–4] as well as references [5–7]. Natural units are used throughout this section, i.e., $c = \hbar = 1$. It is assumed that the reader is familiar with the foundations of particle physics and of the standard model of cosmology, namely the *Lambda Cold Dark Matter model* (ΛCDM).

1.1.1. Chronology of the Universe in a nutshell: the cosmic standard model

The cosmic standard model describes the evolution of the expanding Universe in the framework of quantum field theory and general relativity. The known cosmic history can be broken into three reliably known epochs and the still controversial inflationary period. Each of those was mainly governed by a different form of energy, namely the inflaton field, radiation, matter, and dark energy, see Fig. 1.1. Henceforth some key-events of the thermal evolution of the Universe are recalled.

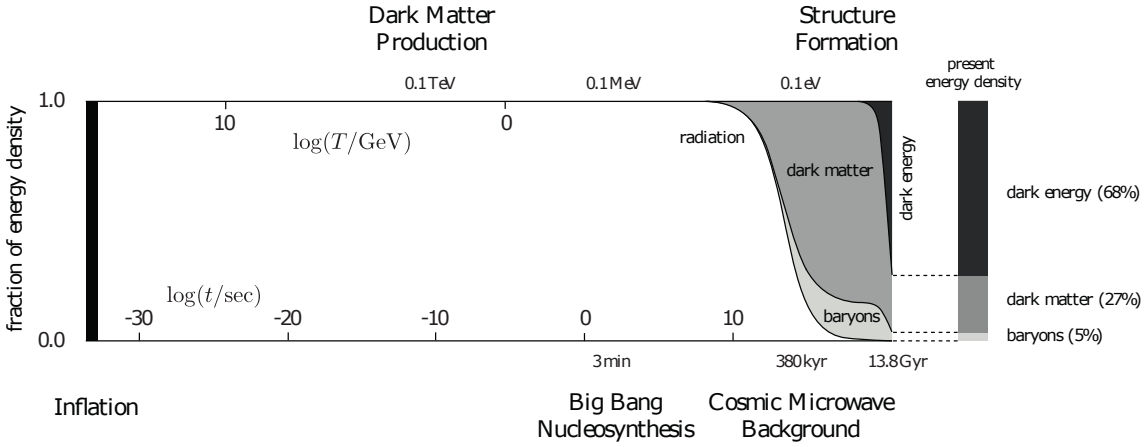


Figure 1.1.: Evolution dependent distribution of the forms of energy in the Universe. After the components – radiation r, matter m, and dark energy Λ – have decoupled they scale differently with time. This can be seen by Eq. (1.17), recalling that $\rho_r \propto a^{-4}$, $\rho_m \propto a^{-3}$, and $\rho_\Lambda \propto$ const. Figure adapted from Ref. [3].

Inflationary period. The inflationary period describes a time in cosmic evolution when the Universe expanded exponentially. This huge expansion was caused by the so-called inflaton field. The period potentially took place between 10^{-36} – 10^{-32} s after the Big Bang

[1]The term *early Universe* refers in this thesis to cosmological times between the key-events of the Big Bang (0 s) and the CMB decoupling ($\approx$ 380 kyr).

and is physically described by inflation theories. The enormous expansion within this short time range generated a dilution of nearly all forms of energy and therefore cooled down the Universe extremely. At the end of inflation the potential energy of the inflaton field was transferred into the generation of particles of the standard model of particle physics. Interactions between the produced particles heated up again the Universe (beginning of the thermal history of the Universe). The inflationary period and the following reheating are described more detailed in Sec. 1.1.2.

Radiation domination. The epoch after inflation was dominated by radiation and lasted 60 kyr. At the very beginning of the epoch, less than 20 ps after the Big Bang, the temperatures of the Universe were extremely high – sufficiently high to unify the electromagnetic and weak interaction into the electroweak force. All particles of the standard model of particle physics were massless at these temperatures, i.e., they solely contributed to the radiation content of the Universe. When the Universe expanded further and hence cooled down the Higgs mechanism provided the masses of the elementary particles. This mechanism triggered the *electroweak (EW) phase transition*, which took place at temperatures around 100 GeV, corresponding to ≈ 20 ps after the Big Bang. At cosmological energies below ≈ 100 GeV and above ≈ 150 MeV the Universe was hot enough for quarks to be asymptotically free and to form a quark-gluon plasma. Below temperatures of ≈ 150 MeV, which corresponds to the time 20 μs, the color confinement prevailed and the quarks combined to baryons and mesons. This transition is known as the *quantum chromodynamics (QCD) phase transition*. Subsequently, some particles with relatively low interaction rates decoupled from the local thermal equilibrium. The first were the weakly interacting neutrinos, which froze out at ≈ 1 MeV (≈ 1 s). Note that the decoupling of potential dark matter particles is assumed to took place around energies of order hundred MeV. At temperatures around 500 keV (6 s) the Universe already cooled down sufficiently to exponentially suppress the production of electron-positron pairs. The next milestone in cosmic evolution is known as the *Big Bang Nucleosynthesis (BBN)* and took place ≈ 3 min (≈ 100 keV) after the Big Bang. During this synthesis the nucleons combined to form bound systems, i.e., the first complex atomic nuclei formed, namely helium ($\approx 25\%$), lithium and other light elements (minor). Eventually, after around 60 kyr (0.75 eV) the energy form of matter became the dominant component.

Matter domination. The process of the first formation of atoms, which is important for studies of the early Universe, took place at temperatures between $\approx 0.23 - 0.33$ eV ($\approx 260 - 380$ kyr) within the epoch of matter domination. It is commonly known as *(re-)combination*. The associated decoupling of photons freed the light of the Cosmic Microwave Background radiation (CMB). Figure 1.2 shows the temperature deviations (anisotropies) ΔT of order $\Delta T/T_{\mathrm{CMB}} \propto \mathcal{O}(10^{-5})$ from a nearly perfectly constant CMB of temperature $T_{\mathrm{CMB}} = 2.7$ K as observed by the Planck collaboration. The correlation structure of these anisotropies contains physical information about the early Universe. Before recombination the Universe was opaque as photons were frequently scattered by free electrons via Thompson scattering, see Fig. 1.3. After the formation of neutral hydrogen atoms the abundance of free electrons decreased and light propagated freely through space for the first time. At this stage of cosmic evolution the epoch of the *dark ages* of the Universe started. It is termed dark because during its period in cosmic chronology no light-emitting objects existed, populated the Universe, with the exception of the 21 cm spin-flip line of hydrogen. Until the beginning of the dark ages the structure (density perturbations) of the

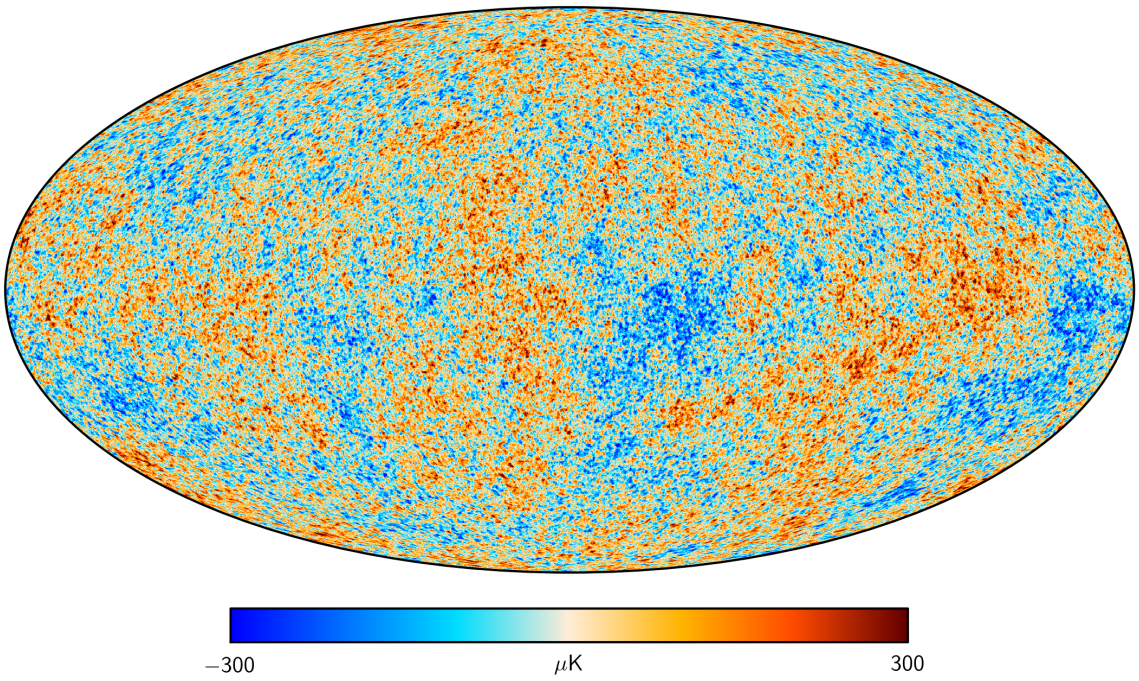


Figure 1.2.: Map of the CMB temperature anisotropies as observed by Planck [8].

Universe can mathematically be described as small density perturbations around a homogeneous background. As structures grew further under the influence of gravity first stars and galaxies formed. These objects ejected high energetic radiation that in turn reionized the neutral hydrogen in the Universe. The epoch related to this process took place around $100 - 400$ Myr ($2.6 - 7.0$ meV) after the Big Bang and is called *reionization*. Eventually more and more matter collapsed gravitationally and galaxies formed significantly. Early representatives of the first galaxies have recently been found around the age of 750 Myr (1.9 meV) after the Big Bang.

Dark energy domination. After the Universe reached the age of around 9 Gyr (0.33 meV) dark energy became the dominant component. Roughly at the same time the solar system formed. The present Universe is 13.8 Gyr old, corresponding to a temperature of 0.24 meV. Extrapolating the current cosmic evolution, i.e. including a cosmological constant Λ in the Friedmann equation, Eq. (1.17), the Universe will expand further and hence will become cooler[2]. Most of the baryonic mass will have been burned within stars and only stellar remnants will be left. Energy will mainly be produced by (hypothetical) proton decays and particle annihilations (estimate: $10^{15} - 10^{37}$ yr). After the epoch of proton decay only black holes will have survived, which will also dissipate by the emission of Beckenstein-Hawking radiation (estimate: $10^{38} - 10^{100}$ yr). When all black holes will have evaporated the Universe will only contain radiation with enormously large wavelengths (due to the spacetime stretch, caused by dark energy) and products of protonic decay. At this stage the Universe will become dark again (estimate: 10^{100} yr).

A summary of the particular epochs of the Universe with corresponding orders of magnitude (energy, redshift, and time) can be found in Tab. 1.1.

[2]Note that we are leaving now the cosmological standard model, describing the idea of the Big Freeze.

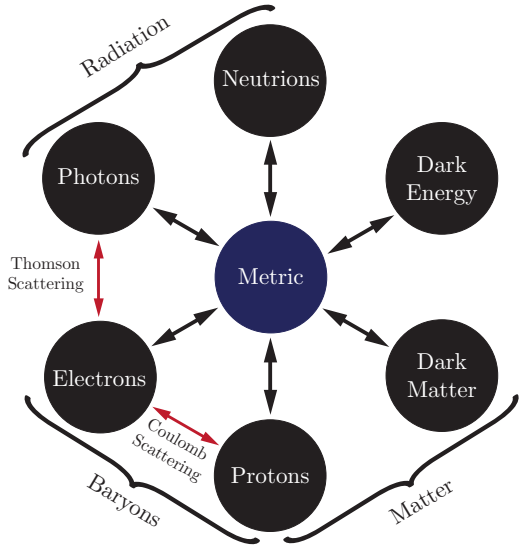


Figure 1.3.: Illustration of the respective interactions of the constituents of the Universe with the background metric and each other. Figure adapted from Ref. [3].

1.1.2. Beyond the standard model – cosmic inflation

In the previous section the thermal chronology of the Universe started with the radiation dominated epoch that includes the EW unification (occurring roughly 20 ps after the Big Bang). Obviously, though, there must be epochs before this, e.g., when the unification of the EW and strong interaction occurred. The epoch of this unification should be described by the so-called Grand Unified Theories (GUT) at $\approx 10^{15} - 10^{16}$ GeV [9]. Theories designed to describe these extremely early epochs, however, are lacking more or less a conclusive experimental confirmation and/or are incomplete and are therefore topics of current research. One of these theories, commonly believed to describe the period of $10^{-36} - 10^{-32}$ s after the Big Bang at energies of order $10^{15} - 10^{16}$ GeV [10], is cosmic inflation [11]. This theory was set up to solve, among others[3], the mystery of the cosmological principle, i.e., the question why the observed Universe is nearly perfectly homogeneous and isotropic. Answering this question led to the well-known horizon problem, which can be easily, qualitatively understood by recalling that Fig. 1.2 shows only the deviations of order $\mathcal{O}\left(10^{-5}\right)$ from a nearly perfectly constant CMB radiation. Detecting only such tiny deviations seems to contradict the fact that formerly causally connected regions should nowadays not be larger than $\approx 2^\circ$ on the sky (if one assumes the thermal cosmic standard model). In the following I will elaborate on the horizon problem and its solution[4] by following Ref. [3].

[3] Inflation was originally set up to solve the magnetic monopole problem [12]. At the same time inflation solves other major problems, e.g. the flatness problem. These problems can be found in many textbooks on cosmology and will not be discussed in this thesis.

[4] Currently, there exists the consensus that the horizon problem has not been solved completely by the theory of inflation given the current data. For several models of inflation one has to require homogeneity and isotropy over more than one Hubble sphere already at the beginning of inflation to be in agreement with current observations [13, 14].

Stage	time t	redshift z	energy E
(Potential) Inflation	10^{-36} s		
EW phase transition	20 ps	10^{15}	100 GeV
QCD phase transition	20 μs	10^{12}	150 MeV
Neutrino decoupling	1 s	6×10^9	1 MeV
Electron - positron annihilation	6 s	2×10^9	500 keV
Big Bang Nucleosynthesis	3 min	4×10^8	100 keV
Matter - radiation equality	60 kyr	3400	0.75 eV
Recombination/ CMB decoupling	260 – 380 kyr	1000 – 1400	0.23 – 0.33 eV
Reionization	100 – 400 Myr	11 – 30	2.6 – 7.0 meV
First galaxy formation	750 Myr	7	1.9 meV
Dark energy - matter equality	9 Gyr	0.4	0.33 meV
Present time	13.8 Gyr	0	0.24 meV
(Potential) Big Freeze	10^{100} yr		

Table 1.1.: Selected key-processes in the history of the Universe. The stated numbers are approximative (or estimated) orders of magnitudes.

Horizon problem. To understand the horizon problem in more detail it is rewarding to consider cosmic length and time scales within the so-called comoving frame, where the Friedmann-Lemaître-Robertson-Walker (FLRW) metric is given by

$$ds^2 = a^2(\tau)\left[d\tau^2 - \left(d\chi^2 + S_\kappa^2(\chi)d\Omega^2\right)\right], \tag{1.1}$$

with a denoting the cosmological scale factor, χ the comoving distance, τ the conformal time, and κ the curvature parameter. The infinitesimal angular element $d\Omega^2$ and the curvature dependent function $S_\kappa(\chi)$ are defined by

$$\begin{aligned} d\Omega^2 &= d\Theta^2 + \sin^2(\Theta)d\phi^2 \text{ and} \\ S_\kappa(\chi) &= \begin{cases} \sinh(\chi) & \kappa = -1 \\ \chi & \kappa = 0 \\ \sin(\chi) & \kappa = +1 \end{cases} . \end{aligned} \tag{1.2}$$

Henceforth, for reasons of clarity and comprehensibility, I focus on photons traveling along radial null geodesics, i.e. $ds^2 = 0$ and $d\Omega^2 = 0$. With these constraints Eq. (1.1) simplifies to

$$d\chi = \pm d\tau, \tag{1.3}$$

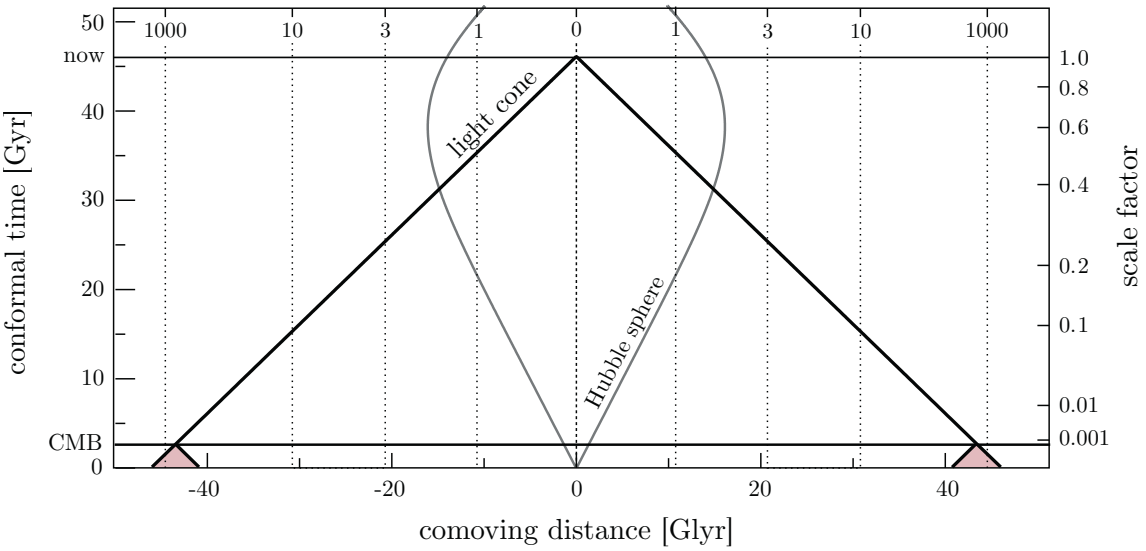


Figure 1.4.: Illustration of the horizon problem: although two spatially widely separated patches of the CMB exhibit the same temperature, they have not been in causal contact before. The scale of causally connected regions is given by the Hubble radius. Figure adapted from Ref. [3].

describing a trivial relation between the conformal time and the comoving distance. It is illustrated by the straight lines of the $\chi - \tau$ diagram in Fig. 1.4. Given Eq. (1.3) the comoving particle horizon $\chi_{\rm ph}$ is defined by

$$\chi_{\rm ph}(\tau) \equiv \tau - \tau_0 = \int_{t_0}^{t} \frac{dt}{a(t)} = \int_{a_0}^{a} d\ln a \; (aH)^{-1} \tag{1.4}$$

with

$$H \equiv \frac{\dot{a}}{a} \equiv \frac{da/dt}{a}. \tag{1.5}$$

$\chi_{\rm ph}$ determines the greatest comoving distance traveled by particles moving with the speed of light between times t_0 and t (e.g. today). The involved term $(aH)^{-1}$ is known as the comoving Hubble radius (often only termed *the horizon*). It describes the maximum distance between particles that are causally connected at a moment of length of the then typical expansion time-scale. Having defined these quantities the horizon problem can be understood by considering Fig. 1.4. The figure illustrates the size of the Hubble sphere, $(aH)^{-3}$, at the time of CMB emission, which is relatively small compared to the distance between two spatially widely separated spots on the CMB sky. The light cones of these two separated CMB spots have not intersected since the beginning of the Universe and thus have not been in causal contact before. The astronomical observation of two such spots having almost exactly the same temperature is therefore not understandable within the cosmic standard model.

A possible solution to the horizon problem would be simply a relatively large particle horizon compared to the Hubble radius, i.e. $\chi_{\rm ph} \gg (aH)^{-1}$, at the last scattering surface. This relation can be achieved by a period, within which the Universe expands accelerated and extremely fast – sufficiently fast to obtain a decreasing (comoving) Hubble radius,

$$\frac{d}{dt}(aH)^{-1} = -\frac{\ddot{a}}{(\dot{a})^2} < 0 \qquad \Rightarrow \qquad \ddot{a} > 0. \tag{1.6}$$

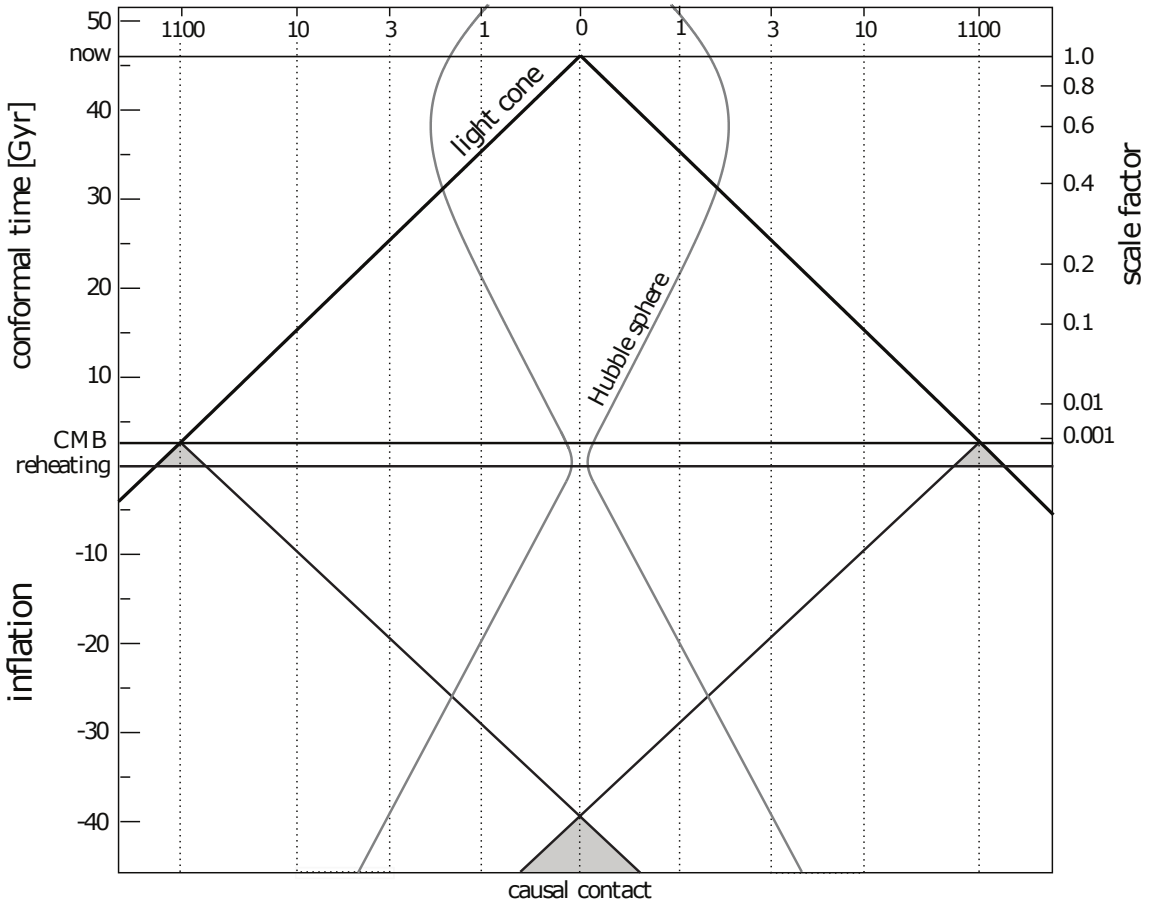


Figure 1.5.: Illustration of the horizon problem solved by cosmic inflation. All regions of the CMB have been in causal contact as their light cones intersect. The causal connection is achieved by a period of a decreasing Hubble radius, referring to a rapid expansion of the Universe. The phase transition between inflation and the thermal history of the Universe is called reheating. Figure adapted from Ref. [3].

This requirement is fulfilled in the theory of cosmic inflation and its effect on the Hubble radius and on the particle horizon is illustrated by Fig. 1.5.

In order to obtain an estimate of the minimal sufficient amount of expansion, caused by inflation, we require that the scale of the observed Universe today, $(a_0 H_0)^{-1}$, is smaller than the Hubble radius before inflation starts, $(a_\mathrm{i} H_\mathrm{i})^{-1}$. This means

$$(a_0 H_0)^{-1} < (a_\mathrm{i} H_\mathrm{i})^{-1}. \tag{1.7}$$

To calculate the inflationary expansion factor, the ratio $(a_\mathrm{i} H_\mathrm{i})^{-1}/(a_\mathrm{f} H_\mathrm{f})^{-1}$, where subscript f denotes the end of inflation, is considered first. For a rough estimate [3] the epochs of matter and dark energy domination are neglected and radiation domination between the end of inflation and today is assumed, where the Hubble parameter is proportional to a^{-2}. Under the assumption that inflation will end before the EW phase transition – at energy scales of GUT, which is roughly of order 10^{15} GeV – the factor between the Hubble radius

at the end of inflation and at its beginning is roughly

$$\frac{(a_\mathrm{i}H_\mathrm{i})^{-1}}{(a_\mathrm{f}H_\mathrm{f})^{-1}} > \frac{(a_0H_0)^{-1}}{(a_\mathrm{f}H_\mathrm{f})^{-1}} \approx \frac{a_\mathrm{f}}{a_0}\left(\frac{a_0}{a_\mathrm{f}}\right)^2 = \frac{T_\mathrm{f}}{T_0} \propto 10^{27} > e^{65}. \tag{1.8}$$

To translate the ratio of scale factors to temperatures the redshift relation $T_\mathrm{f}/T_0 = (1+z) = a_0/a_\mathrm{f}$ was used. As will be shown in the next paragraph, inflation corresponds to an exponential expansion of space,

$$a(t) \propto \exp(tH), \quad \text{with} \quad H \approx \text{const.} \tag{1.9}$$

This means $H_\mathrm{i}/H_\mathrm{f} \approx 1$ and the here estimated amount of inflationary expansion, $a_\mathrm{f}/a_\mathrm{i}$, is larger than 65 e-folds.

Inflationary physics. As estimated above inflation is supposed to stretch the Universe by at least 65 e-folds. But what drives this enormous expansion? In the following I elaborate on the theory of the basic physical mechanism behind inflation to obtain a more concrete idea of it.

As an illustrative representative of the inflationary mechanism we consider a (minimally coupled[5]) single scalar field ϕ – the inflaton – with related potential $V(\phi)$, being located apart from its potential minimum, see Fig. 1.6. Such a situation within curved spacetime can be described by the action [3, 15]

$$S = \int d^4x\sqrt{-g}\left(\frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - V(\phi)\right). \tag{1.10}$$

The symbol $g^{\mu\nu}$ denotes the metric tensor of general relativity and g its determinant with $\mathrm{sgn}(g) < 0$. To obtain the equation of motion of ϕ one has to require $\delta S = 0$, yielding

$$\nabla^\alpha\nabla_\alpha\phi + \frac{dV}{d\phi} = 0, \tag{1.11}$$

which is the covariant[6] form of the Klein-Gordon equation. Next, the scalar field is assumed to consist of a time dependent but spatially constant (due to symmetries of FLRW metric [3]) background field $\phi_0(t)$ and some quantum fluctuations $\delta\phi(t,x)$ with spatial dependence, i.e.,

$$\phi(t,x) = \phi_0(t) + \delta\phi(t,x). \tag{1.12}$$

For an illustration of the inflaton field with related potential see again Fig. 1.6. To keep it simple first, only the constant background field is considered. The evaluation of Eq. (1.11) for a homogeneous and isotropic Universe, described by the FLRW metric, yields the equation of motion of the background field,

$$\ddot{\phi}_0 + 3H\dot{\phi}_0 + dV/d\phi_0 = 0. \tag{1.13}$$

[5] In general, one has the gauge freedom to add a term to Eq. (1.10), which induces a coupling between ϕ and the background gravitational field. Minimally coupled means we chose a gauge, in which this term vanishes [13].

[6] Note that also derivatives are covariant, i.e. metric dependent, and therefore $\nabla^\alpha\nabla_\alpha\ \phi = \frac{1}{\sqrt{-g}}\partial_\mu\left(\sqrt{-g}g^{\mu\nu}\partial_\nu\phi\right)$. They depend in particular on the Christoffel symbols, which again depend on the metric.

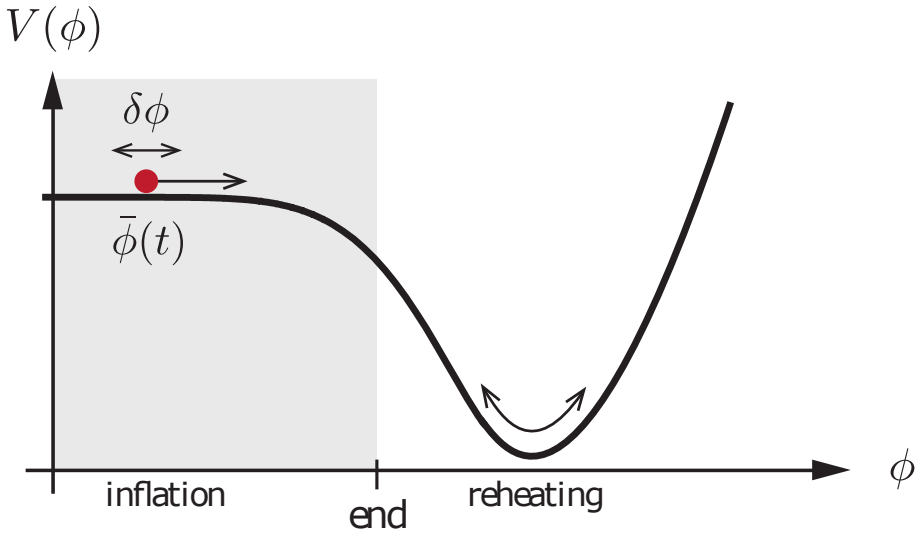


Figure 1.6.: The inflaton field consisting of a constant background field and fluctuations with related potential. Inflation occurs in the shaded region, where the potential is relatively flat. As the inflaton "rolls down" the steeper part of the potential the potential energy is transferred into kinetic energy and the inflaton will oscillate around the potential minimum. It eventually decays into components of the standard model of particle physics. Figure adapted from Ref. [3].

For the evolving scalar field ϕ_0 within the potential the gradient of the potential acts like a force, while the expansion of the Universe acts like a friction [3]. To see in which cases the dynamics of the scalar field lead to an accelerated expansion the energy momentum tensor has to be calculated first. The insertion of Eq. (1.10) into the definition of the energy momentum tensor within curved spacetime yields [15]

$$T_{\mu\nu} \equiv \frac{2}{\sqrt{-g}} \frac{\delta S}{\delta g^{\mu\nu}} = \partial_\mu \phi_0 \, \partial_\nu \phi_0 - g_{\mu\nu} \left(\frac{1}{2} g^{\alpha\beta} \partial_\alpha \phi_0 \, \partial_\beta \phi_0 - V(\phi_0) \right). \tag{1.14}$$

Note in particular that the first term in parenthesis is only different from zero in the time-time component, i.e. $\alpha = \beta = 0$. The same applies for the first term considering $\mu\nu$-indices. By evaluating Eq. (1.14) separately for the time-time component, $T_{00} = \rho_{\phi_0}$, and space-space components, $T_{ij} = -g_{ij} p_{\phi_0}$, one obtains [3]

$$\begin{aligned} \rho_{\phi_0} &= \frac{1}{2} \left(\frac{d\phi_0}{dt} \right)^2 + V(\phi_0), \\ p_{\phi_0} &= \frac{1}{2} \left(\frac{d\phi_0}{dt} \right)^2 - V(\phi_0). \end{aligned} \tag{1.15}$$

Here, ρ_{ϕ_0} denotes the density of the scalar field and p_{ϕ_0} its pressure. If one further requires the *slow-roll condition*[7]

$$\sqrt{V(\phi_0)} \gg \left| \frac{d\phi_0}{dt} \right| \tag{1.16}$$

[7] The two commonly used slow-roll parameters ϵ and η are also directly related to the potential, $\epsilon \propto (dV/d\phi)^2$ and $\eta \propto d^2V/d\phi^2$.

the pressure is given by $p_{\phi_0} \approx -\rho_{\phi_0} < 0$. A negatively valued pressure seems unintuitive in the first place, but is well-known in cosmology by the effect of the cosmological constant Λ on the evolution of the Universe. A positive cosmological constant implies in particular the presence of a component with negative pressure that leads to an exponential expansion of spacetime, $a(t) \propto \exp(Ht)$, with constant H.

The expansion behavior of a Universe, dominated by a cosmological constant, can be studied by considering the evolution equation of the scale factor, which is described by the Friedmann equation,

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G_N}{3}\left(\rho_\kappa + \rho_\mathrm{r} + \rho_\mathrm{m} + \rho_\Lambda\right), \tag{1.17}$$

where G_N denotes Newtons gravitational constant and $\rho_{\kappa,r,m,\Lambda}$ the curvature[8], radiation, matter, or cosmological constant density, respectively. For a Universe dominated by $\rho_\Lambda = \Lambda/(8\pi G_N)$ the Friedmann equation yields $a(t) \propto \exp(\sqrt{\Lambda/3}\ t) = \exp(Ht)$. A space with such an expansion behavior is called de Sitter space.

To understand the negative pressure implication of a positive cosmological constant Λ, we consider Eq. (1.14) and use the slow-roll condition, Eq. (1.16). The energy momentum tensor simplifies to $T_{\mu\nu} \approx V(\phi_0) g_{\mu\nu} \approx -p_{\phi_0} g_{\mu\nu}$ and indeed, a component with negative pressure behaves like a positive Λ within the Einstein equations [3],

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi G_N \left(T_{\mu\nu} + \frac{\Lambda}{8\pi G_N} g_{\mu\nu}\right). \tag{1.18}$$

This means the cosmological constant can be identified with $\Lambda = 8\pi G_N V(\phi_0) = -8\pi G_N p_{\phi_0}$. For the sake of completeness, $R_{\mu\nu}$ represents the Ricci tensor and R the Ricci scalar.

As a consequence of the discussion above, a (minimally coupled) scalar field slowly rolling down a flat potential leads to an approximately exponential expansion of spacetime. The inflationary phase, however, has to stop after roughly 65 e-folds of expansion. Therefore, during this phase, the space coincides not perfectly with a de Sitter space and is commonly called quasi de Sitter space.

The seeds of structure. Up to now I discussed the classical case by considering the inflaton field without fluctuations. To figure out how the fluctuations evolve during inflation they are included in the following. For this purpose variations of the action δS [Eq. (1.10)] have to be calculated to the desired order in cosmological perturbation theory to finally obtain the equation of motion of the inflaton field. Performing this calculation turns out to be a bit involved, which is the reason why this part is skipped here. In lieu thereof a qualitative explanation of the evolution of the inflaton field including fluctuations is provided, further following Ref. [3]. For a fully quantitative description of the cosmological perturbation theory consider Refs. [16, 17].

During inflation there appear (Gaussian) zero point fluctuations of the inflaton, $\delta\phi$. This means statistically there are regions of the Universe where inflation lasted longer and regions where it stopped earlier. This statistical process produces local density perturbations at the end of inflation since some regions get more diluted than others. To be precise, the inflaton fluctuations are produced inside the (comoving) Hubble radius, which decreases during inflation. A specific fluctuation with Fourier mode k exits the Hubble radius at a particular point of time during the inflationary expansion, see Fig. 1.7. By crossing the Hubble radius at $k = aH$ the fluctuation gets conserved. At the time of horizon crossing

[8] Current measurements [7] are consistent with $\rho_\kappa \approx 0$.

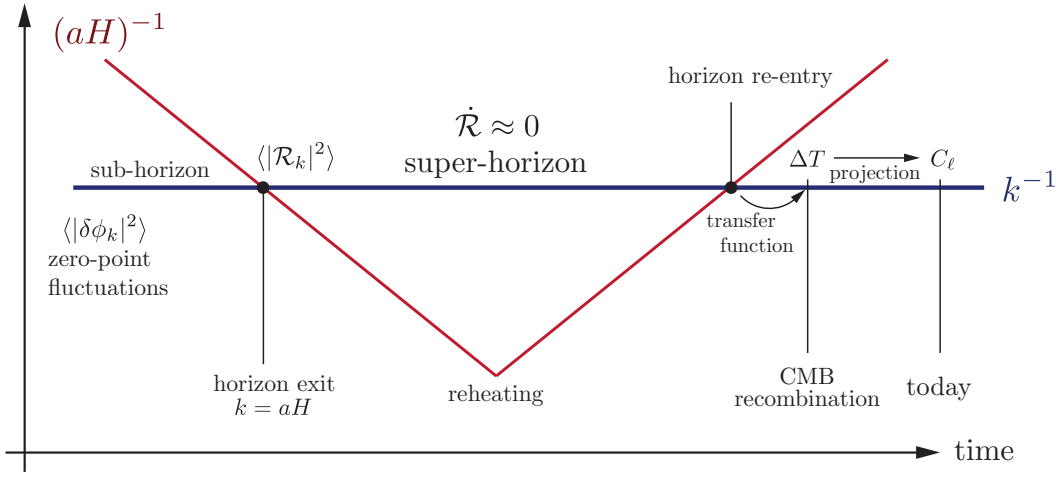


Figure 1.7.: Origin and evolution of the curvature perturbations. Zero-point fluctuations of the inflaton field, generated during inflation, exit *the horizon* (Hubble radius $(aH)^{-1}$) while the Universe expands exponentially. After the end of inflation (reheating) the Hubble radius increases and the curvature perturbations re-enter *the horizon* as real perturbations. These primordial perturbations are imprinted in the CMB anisotropies and initiate the process of cosmic structure formation. Illustration adapted from Ref. [3].

it is beneficial to change from the inflaton fluctuations to the so-called comoving curvature perturbations $\mathcal{R}$, since they are conserved outside the horizon. After inflation, the Hubble radius increases again and the perturbation with Fourier mode k enters the Hubble radius as a real physical curvature perturbation. Perturbations of different scales exit during inflation and re-enter after inflation the Hubble radius at different times. These curvature perturbations, which are related to gravitational potentials [see chapter 3], interact with the composites of the Universe (see again Fig. 1.3). In consequence of this interaction between baryonic matter, radiation, dark matter and gravitational potentials the anisotropies of the CMB are eventually produced. After CMB decoupling, matter clumped more efficiently and cosmic structures grew under the influence of gravity. This means the curvature perturbations finally serve as the seeds of structure formation in the late Universe. Therefore inflation provides also a natural explanation of the origin of cosmic structures.

The seeds of structure – the initial conditions for structure formation – obey almost completely Gaussian statistics when considering simple single field, slow-roll models of inflation as done in the above discussion [18]. Gaussian distributions are completely determined by the two point correlation function, the so-called power spectrum in Fourier space (assuming statistical isotropy and homogeneity). The power spectrum, in turn, describes with how much power/amplitude each Fourier mode contributes to the signal. In the simplest inflationary case the Gaussian fluctuations $\delta\phi$ exit the horizon independently, without any coupling, where they freeze. It can be shown, that the related, conserved, dimensionless, power spectrum $\Delta^2_{\delta\phi}$ at horizon crossing, $k = aH$, is given by

$$\Delta^2_{\delta\phi}(k) \approx \left.\frac{H^2}{4\pi^2}\right|_{k=aH}. \tag{1.19}$$

Recalling that $H \approx$ const. during inflation yields $\Delta^2_{\delta\phi} \approx$ const. and therefore all Fourier

modes contribute equally to the (dimensionless) power spectrum. The scale-invariant power spectrum of Eq. (1.19) is called Harrison-Zel'dovich spectrum. As already mentioned above, it is common to switch the variables from $\delta\phi$ to the comoving curvature perturbation $\mathcal{R}$ at horizon crossing since its statistics are conserved until the perturbations re-enter the horizon (see again Fig. 1.7). Additionally including the freedom for small deviations from the Harrison-Zel'dovich spectrum we arrive at a frequently used form of the power spectrum of $\mathcal{R}$,

$$\Delta^2_{\mathcal{R}}(k) \approx A_s \left(\frac{k}{k_*}\right)^{n_s-1}, \tag{1.20}$$

often called primordial scalar power spectrum. Such deviations from a constant power spectrum are included, because H is a slowly varying function of time and only approximately a constant (the space-time metric is nearly de Sitter, but not exactly). A_s represents the primordial scalar amplitude, n_s the scalar spectral index, and k_* denotes the so-called pivot scale. This relatively simple power-law parametrization enables comparisons with observations. The actual values, derived by the Planck collaboration, are $\ln\left(10^{10}A_s\right) = 3.089 \pm 0.036$ and $n_s = 0.968 \pm 0.006$ for the pivot scale $k_* = 0.05\text{Mpc}^{-1}$ [19]. These results confirm the prediction of the simplest inflation models, stating that $n_s < 1$.

In summary, inflation predicts nearly perfectly Gaussian curvature perturbations with an approximately scale-invariant power spectrum. The curvature perturbations are directly related to density perturbations and serve as the seeds of structure formation in the Universe.

Reheating. The end of inflation is followed by an epoch named reheating. By the inflationary expansion nearly all forms of energy must have been extremely diluted, implying that the Universe cooled down significantly during inflation. Consequently, there must have been a mechanism at the end of inflation that transforms the energy stored in the inflationary potential into the generation of particles of the standard model of particle physics [3], producing a dense thermalized state of the Universe. Otherwise, the thermal evolution of the Universe would not have started and the epochs of radiation and matter domination, described in Sec. 1.1.1, would not have existed. Assume, for simplicity, that the inflaton potential has the shape illustrated by Fig. 1.6. As discussed in the previous paragraphs, inflation takes place in the shaded region, where the potential is flat. As the inflaton "rolls down" the steeper part of the potential, inflation stops and the potential energy gets transformed into kinetic energy of the inflaton field. Around the potential minimum the inflaton starts to oscillate. The local shape of the potential in the vicinity of its minimum can be approximated by $V(\phi) = \frac{1}{2}m^2\phi^2$. Plugging this local parameterization of V into Eq. (1.13) yields the equation of motion of the inflaton around the potential minimum,

$$\ddot{\phi} + 3H\dot{\phi} + m^2\phi = 0, \tag{1.21}$$

which represents a damped harmonic oscillator with undamped frequency m. The oscillating inflaton field, now storing the energy of the potential, is assumed to decay subsequently into the elements of the standard model of particle physics. The interactions between the decay products are heating up the Universe (entropy generation) until thermal equilibrium is reached. After this process the epoch of radiation domination begins and with it the thermal evolution of the Universe. An illustration of the history of the Universe including inflation and reheating can be found in Fig. 1.8.

1.2. Motivation and outline of the thesis

The introduced theory of an inflationary phase in the early Universe seems to solve, among others, the horizon problem and supplies a natural explanation of the origin of cosmic structures. To judge whether inflation theory is true its predictions have to be confronted with observational data. For this purpose one can use the CMB or the large-scale structure of the Universe (LSS). The former carries information of the curvature perturbations at the epoch of recombination, the latter represents the more recent stages of the evolution of the cosmic structures. Although this thesis will mainly focus on the CMB, parts of the here developed methods can be applied to LSS data as well. Whenever this is the case, it will be stated.

Besides the validation of the inflation theory itself, the specific kind of inflation model is of interest. There exist a huge variety of inflationary models, ranging from different inflaton potentials over distinct inflaton couplings [first term in Eq. (1.10)] to the presence of multiple coupled inflaton fields. Many of these models produce unique statistics of the curvature perturbations. Therefore, the respective inflation models can be discriminated by, e.g., the shape of the primordial power spectrum or specific deviations from Gaussianity. By inferring such statistical quantities from data one would be able to constrain the variety of actual still possible inflation models. This opportunity immediately suggests the development of novel statistical methods, which are able to infer statistical properties of early-Universe signals (generated by inflation) from observational data. Such methods are developed in this work.

As stated before, when speaking about observational data this thesis mainly refers to CMB observations, which are measured by terrestrially or extraterrestrially based telescopes. An entity of cosmological observations is that we observe the Universe and in particular the CMB always from the same point of view. Hence the observed statistical realization of the CMB, and in general the realized morphology of the Universe, stays the same for all human observations. Having only one realization of the Universe obviously makes the inference procedures more difficult. In other words, we are only able to infer quantities conditional to the observed data and to the prior assumptions done, since we cannot marginalize over infinite data realizations (as approximately possible in other research fields like particle physics, where an experiment can be repeated as often as desired). For such situations Bayesian statistics are predestinated, because they naturally involve prior assumptions, and are capable of consistently dealing with few data (in the sense of having few realizations of it).

In cosmology, and very often in all other fields of natural science, one is faced with the situation of having a discrete data set and thus a finite number of degrees of freedom, but is interested in inferring a physical field containing an infinite number of degrees of freedom. For such situations, information field theory (IFT) has been developed within the framework of Bayesian statistics [21]. The necessary parts of this theory, used to follow the major derivations of the here presented inference methods, are introduced at the beginning of each chapter of this work.

The thesis is organized as follows. The second chapter, Chap. 2, is named *Signal inference with unknown response: Calibration-uncertainty renormalized estimator* and includes a consistent advancement of the theory of self-calibration in the framework of IFT. We consciously start with this topic since every physical experiment is affected by its calibration. After addressing this generic topic we introduce a method, able to reconstruct

specific parameters of inflation by exploiting higher-order statistics of the CMB. We also present an approach to reconstruct the primordial scalar power spectrum using CMB data. This third chapter, Chap. 3, has the name *Generic inference of inflation models by non-Gaussianity and primordial power spectrum reconstruction.* A required intermediate stage of the methods of Chap. 3 is the reconstruction of the primordial scalar potential[9]. In Chap. 4, *All-sky reconstruction of the primordial scalar potential from WMAP temperature data*, such a reconstruction has been done. Finally, in Chap. 5: *Stochastic determination of matrix determinants*, we develop a novel method to calculate the determinant of an implicitly defined matrix via statistical sampling. Such calculations are required, for instance, in Chap. 3, when considering significant deviations from Gaussianity. A summary of the thesis as well as an outlook can be found in Chap. 6.

Note that throughout the thesis there appear a few double occupied acronyms since I do not want to change the exact wording of the published works, which are reproduced in these chapters. Within each chapter the acronyms are consistent, though.

[9]The primordial scalar potential is related to the comoving curvature perturbations, see Sec. 3.

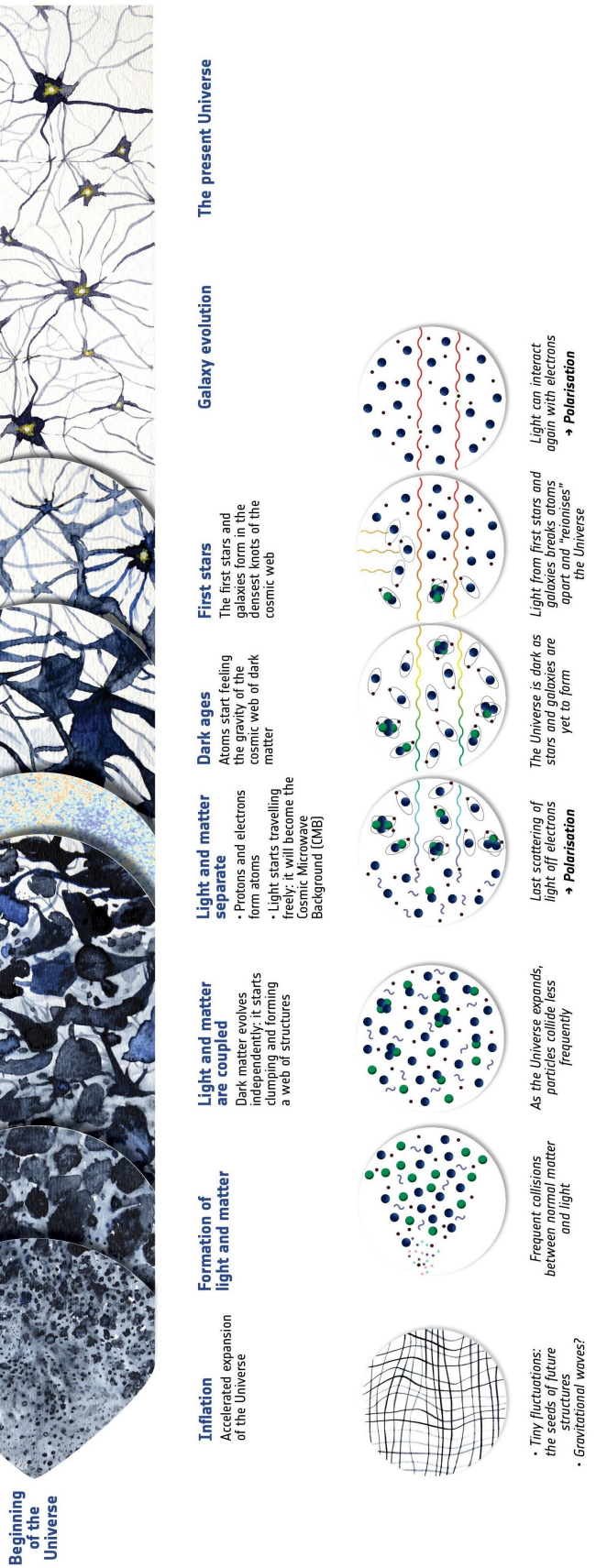


Figure 1.8.: Illustration of the evolution of the Universe. Taken from Ref. [20].

2. Signal inference with unknown response: Calibration-uncertainty renormalized estimator

Note: This chapter has been published in Phys. Rev. E **91**, 013311 (2015) [22].

2.1. Introduction

2.1.1. Motivation

Data analysis is the link between theory and experiment, wherein a signal has to be inferred from measured data. For this purpose the transformation of a signal to data, the measurement response, has to be understood precisely. The reconstruction of this response is called calibration.

In the most simple case of a time independent instrument response, the calibration can be determined by measuring an a priori well known signal in a regime with neglectable noise level. This is commonly called external calibration. However, the assumption of time independency cannot be accepted in the majority of cases. Of course the time dependency caused by, e.g., environmental factors, periodicities and systematics, or the signal itself, can be estimated with utmost effort. The resulting calibration, however, has still to be extrapolated into future time, where the real measurement will be performed and where these influences will not be known exactly. What might be known, however, are their statistics. The resulting uncertainty consequently affects the signal reconstruction and has to be taken into account.

There are methods, which improve the calibration by iteratively calibrating on a signal reconstruction and then improving the reconstruction using the new calibration. Such self-calibration (selfcal) schemes are widely in usage. They can, however, be prone to systematic biases since signal and calibration are partly degenerate, i.e., a feature in the data could be caused by either of them and it is not guaranteed that the selfcal scheme does the correct choice automatically.

An improved selfcal scheme, which takes signal uncertainties in the calibration step into account was presented in Ref. [23]. Since also this new selfcal is an approximative solution to the complex inference problem, we ask if there is room for further improvement using information field theory (IFT) [21]. To this end we develop a calibration uncertainty renormalized estimator (CURE) for a signal, which incorporates calibration uncertainties successively in a so-called renormalization flow equation. In comparison to existent approaches this method is non-iterative. For a review and discussion of previous work on existent calibration methods we point to Refs. [23, 24].

2.1.2. Structure of the work

The remainder of this work is organized as follows. In Sec. 2.2 we review the basics of the free and interacting IFT with focus on the latter. Sec. 2.3 represents the main part of the paper, where the calibration problem is introduced and CURE is derived. The basic ideas as well as the main formulae of alternative selfcal schemes are also presented within this section. In Sec. 5.2.2 the performance of several signal reconstruction methods is studied within a numerical toy example. Results are summarized in Sec. 2.5.

2.2. Information field theory

To follow the derivation of an estimator with renormalized calibration uncertainty in the framework of IFT one has to be familiar with the concepts of interacting IFT (see in particular Secs. 2.2.2, 2.2.3). Thus, a brief review might be helpful, but can be skipped by an advanced reader. For this purpose we basically follow the Refs. [21, 25], where a more detailed description of IFT can be found.

2.2.1. Basic formalism & free theory

Typically, a signal has to be inferred from data with the challenging question, how to do this in an optimal[1] way? To reasonably answer this questions we first have to agree on a particular data model.

Within this work we assume that the data can be expressed by a discrete data tuple, $d = (d_1, \dots, d_m)^T \in \mathbb{R}^m,\ m \in \mathbb{N}$, which is related to a signal s by

$$d = Rs + n, \tag{2.1}$$

where R is a linear response operation acting on the signal, and $n = (n_1, \dots, n_m)^T \in \mathbb{R}^m$ denotes some measurement noise. Contrary to data and noise, the signal $s \equiv s(x),\ x \in \mathcal{U}$ is considered to be a continuous quantity over some Riemannian manifold $\mathcal{U}$, i.e., a physical (scalar) field. The linearity of the signal response, which transforms the continuous signal into data space, is valid for many physical measurements, e.g., observations of the cosmic microwave background and large scale structure in astronomy (cosmology), spectroscopy in different fields of physics, or medical imaging.

We further assume the signal and noise to be uncorrelated, $\mathcal{P}(s,n) = \mathcal{P}(s)\mathcal{P}(n)$, and primarily Gaussian, i.e., $\mathcal{P}(s) = \mathcal{G}(s,S)$ and $\mathcal{P}(n) = \mathcal{G}(n,N)$ with related covariances $S = \left\langle ss^\dagger \right\rangle_{(s|S)}$ and $N = \left\langle nn^\dagger \right\rangle_{(n|N)}$, respectively. Here, we implicitly introduced the notation

$$\begin{aligned} \mathcal{G}(a,A) &\equiv \frac{1}{\sqrt{|2\pi A|}} \exp\left(-\frac{1}{2} a^\dagger A^{-1} a\right), \text{ and} \\ \langle\ .\ \rangle_{(a|A)} &\equiv \int \mathcal{D}a\ .\ \mathcal{P}(a|A), \end{aligned} \tag{2.2}$$

where $\dagger$ denotes a transposition and complex conjugation, $*$. The appropriate inner product of two fields $\{a,\ b\}$ is defined by $a^\dagger b \equiv \int_{\mathcal{U}} d^{\dim \mathcal{U}} x\ a^*(x) b(x)$. If the conditions described above (known linear response, Gaussian signal and noise with known covariances) are met, we term the theory a *free theory*.

[1] Optimal with respect to, e.g., minimizing the $\mathcal{L}^2$-error.

It is often convenient and common to focus on logarithmic probabilities by relating Bayes theorem [26] to statistical physics,

$$\mathcal{P}(s|d) = \frac{\mathcal{P}(s,d)}{\mathcal{P}(d)} \equiv \frac{1}{\mathcal{Z}} \exp\left[-\mathcal{H}(s,d)\right]. \tag{2.3}$$

Here, we introduced the information Hamiltonian

$$\mathcal{H}(s,d) \equiv -\ln\left[\mathcal{P}(s,d)\right], \tag{2.4}$$

and the partition function

$$\mathcal{Z}(d) \equiv \mathcal{P}(d) = \int \mathcal{D}s \exp\left[-\mathcal{H}(s,d)\right]. \tag{2.5}$$

Still considering the above free theory we find

$$\begin{aligned} \mathcal{H}(s,d) &= \mathcal{H}_0 - j^\dagger s + \frac{1}{2} s^\dagger D^{-1} s, \text{ and} \\ \mathcal{Z}(d) &= \sqrt{|2\pi D|} \exp\left(\frac{1}{2} j^\dagger D j - \mathcal{H}_0\right), \end{aligned} \tag{2.6}$$

with the abbreviations

$$\begin{aligned} \mathcal{H}_0 &= \frac{1}{2}\ln|2\pi N| + \frac{1}{2}\ln|2\pi S| + \frac{1}{2} d^\dagger N^{-1} d, \\ D^{-1} &= S^{-1} + R^\dagger N^{-1} R, \text{ and} \\ j^\dagger &= d^\dagger N^{-1} R, \end{aligned} \tag{2.7}$$

where the so-called information propagator, D, and the information source, j, have been introduced. $|\,.\,|$ denotes the determinant.

To exploit the whole machinery of statistical physics we additionally include a moment generating term, $J^\dagger s$, into the partition function,

$$\mathcal{Z}(d,J) = \int \mathcal{D}s \exp\Big[-\mathcal{H}(s,d) + J^\dagger s\Big]. \tag{2.8}$$

The last definition permits to express the connected correlation functions (= cumulants) of a probability density function (PDF) via functional derivatives [21],

$$\langle s(x_1)\ldots s(x_n)\rangle^c_{(s|d)} \equiv \left.\frac{\delta^n \ln\left[\mathcal{Z}(d,J)\right]}{\delta J(x_1)\ldots\delta J(x_n)}\right|_{J=0}. \tag{2.9}$$

Since we consider a Gaussian signal, its mean is equivalent to the well known Wiener filter [27] solution,

$$\langle s\rangle_{(s|d)} = Dj \equiv m_{\mathrm{w}}. \tag{2.10}$$

Its two point correlation function describes the uncertainty of the reconstruction, $\left\langle ss^\dagger\right\rangle^c_{(s|d)} = \left\langle (s-m_{\mathrm{w}})(s-m_{\mathrm{w}})^\dagger\right\rangle_{(s|d)} = D$, and all cumulants with $n > 2$ vanish. Therefore, the posterior is Gaussian and given by

$$\mathcal{P}(s|d) = \mathcal{G}(s-m_{\mathrm{w}}, D). \tag{2.11}$$

2.2.2. n-th order perturbation theory

Within the free theory we required the noise and in particular the signal to be Gaussian. However, this requirement cannot be met in some cases, e.g., in case noise or response are signal dependent, or simply a non-linear signal field. In the framework of IFT these scenarios can often[2] be described by a Taylor-expanded Hamiltonian [21] composed of a free part, $\mathcal{H}_{\text{free}}$ (Eq. (2.6)), and a so-called interacting part, $\mathcal{H}_{\text{int}}$,

$$\mathcal{H} = \mathcal{H}_{\text{free}} + \underbrace{\sum_{n=0}^{\infty} \frac{1}{n!} \Lambda^{(n)} \left[s^{(n)}\right]}_{\equiv \mathcal{H}_{\text{int}}}, \tag{2.12}$$

where the deviation from Gaussianity is encoded in the anharmonic terms, $n > 2$. The term $\Lambda^{(n)} \left[s^{(n)}\right]$ denotes a complete, fully symmetric[3], contraction between the rank-n tensor $\Lambda^{(n)}$ and the n fields $s^{(n)} = (s^1, \dots, s^n)$. If a decent estimate m_0 is known, one should Taylor-expand the Hamiltonian around this reference field m_0 in terms of residuals $\phi \equiv s - m_0$. A well working estimate is, for instance, the Wiener filter solution of the free theory, Eq. (2.10). Using this reference field expansion often permits to truncate the Taylor-expansion earlier, since the anharmonic terms become smaller.

Analogously to the free theory we define the partition function,

$$\begin{aligned}
\mathcal{Z}(d, J) &= \int \mathcal{D}s \exp\left[-\mathcal{H}(s, d) + J^\dagger s\right] \\
&= \int \mathcal{D}s \exp\left[-\mathcal{H}_{\text{int}}\right] \exp\left[-\mathcal{H}_{\text{free}} + J^\dagger s\right] \\
&= \exp\left(-\mathcal{H}_{\text{int}}\left[\frac{\delta}{\delta J}\right]\right) \int \mathcal{D}s \exp\left[-\mathcal{H}_{\text{free}} + J^\dagger s\right] \\
&\equiv \exp\left(-\mathcal{H}_{\text{int}}\left[\frac{\delta}{\delta J}\right]\right) \mathcal{Z}_{\text{free}} \\
&= \left(1 - \mathcal{H}_{\text{int}}\left[\frac{\delta}{\delta J}\right] + \frac{1}{2!}\mathcal{H}^2_{\text{int}}\left[\frac{\delta}{\delta J}\right] - \dots\right) \mathcal{Z}_{\text{free}}.
\end{aligned} \tag{2.13}$$

In principle, Eqs. (2.9) and (2.13) enable to calculate all correlation functions of a PDF perturbatively. These calculations, however, are very uncomfortable and lengthy. Fortunately, there exists a well known diagrammatic treatment in analogy to quantum field theory and thermal field theory [21]. E.g., including the first two correction terms, the signal mean m is given by

$$\begin{aligned}
m_x &= \quad + \quad + \quad + \dots \\
&= D_{xy}\left(j_y - \frac{1}{2}\Lambda^{(3)}_{yzv} D_{zv} - \frac{1}{2}\Lambda^{(3)}_{yzv}(Dj)_z (Dj)_v\right) + \dots,
\end{aligned} \tag{2.14}$$

where the ordering of diagrams corresponds to those of the equations and dots (...) representing the residual Feynman-series of correction terms. The external dots (•) represent

[2] See Sec. 2.2.3 for cases in which such a treatment is not sufficient.

[3] $\Lambda^{(n)} \equiv \frac{1}{n!} \sum_\pi \Lambda^{(n)}_{\pi(x_1,\dots,x_n)}$, with π representing every permutation of $\{1, \dots, n\}$.

source terms, internal dots vertices (the tensors $\Lambda^{(n)}$), and lines (——) propagator terms, respectively. Repeated indices are to be integrated over.

The Feynman rules used in this work, which are necessary to switch between the mathematical expressions and the corresponding diagrams, can be found in App. A.1.

2.2.3. Uncertainty renormalization

2.2.3.1. Motivation

The approach of perturbative diagrammatic expansion is supposed to work well if the Hamiltonian is dominated by linear and quadratic terms. That in turn means that the the tensors $\Lambda^{(n)}$ describing the deviation from Gaussianity are sufficiently small for the Feynman-series to converge. This is, however, not always the case, e.g., within the calibration problem where the signal response cannot be known exactly due to some potential time-dependencies or uncontrolled systematics. This calibration uncertainty can lead to large, non-vanishing terms $\Lambda^{(n)}$ as we show in Sec. 2.3.1 of this paper.

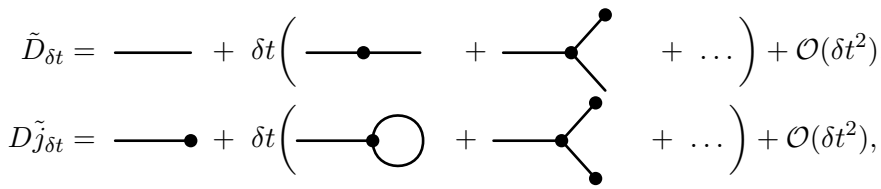
Following the concept of Ref. [25], we can circumvent such a problem by including successively more and more small portions of, e.g., calibration uncertainty into a signal inference equation. The basic idea is to include only a sufficiently small amount of uncertainty per step to ensure the non-Gaussian (interaction) terms to be weak. Finally, this process results in a renormalized propagator, $\tilde{D}$, and information source, $\tilde{j}$. This process is called uncertainty renormalization [25].

2.2.3.2. Concept

For reasons of clarity and comprehensibility we skip the most general derivation and justification of uncertainty renormalization, which can be found in Ref. [25], and focus more on the pragmatic procedure thereof. In the following we consider the Taylor-expanded, effective Hamiltonian to be of the form of Eq. (2.12). To suppress the strength of the non-Gaussian contributions we include a so-called expansion parameter, $\delta t \ll 1$, into the Hamiltonian,

$$\mathcal{H} = \mathcal{H}_{\text{free}} + \delta t \sum_{n=0}^{\infty} \frac{1}{n!} \Lambda^{(n)} \left[s^{(n)} \right], \tag{2.15}$$

and concentrate on this new Hamiltonian for a moment. For an appropriately small δt the interaction terms become sufficiently small and the diagrammatic expansion of Sec. 2.2.2 is justified again. Hence, by including the first correction terms into the propagator, $D \to \tilde{D}_{\delta t}$, and into the information source, $j \to \tilde{j}_{\delta t}$, we obtain



(2.16)

where the dots (...) represent all diagrams of order $\mathcal{O}(\delta t)$, i.e., all possible one-vertex diagrams. This way, $t \in [0, 1]$ can be identified with a pseudo-time, which measures the accumulated uncertainty correction to the information propagator and source, and the expansion parameter δt represents the time step in which D and j are increased from their

intermediate values, D_t and j_t, to their one-step-renormalized (but not final!) values $D_{t+\delta t}$ and $j_{t+\delta t}$, i.e.

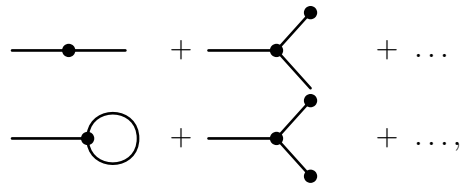
$$\begin{aligned} D_t &\to D_{t+\delta t}, \text{ and} \\ j_t &\to j_{t+\delta t}. \end{aligned} \tag{2.17}$$

We want to emphasize that δt cannot simply be set to unity to obtain the fully renormalized propagator, $\tilde{D}$, because this step would violate the justification of our perturbative expansion (see Sec. 2.2.3.1). However, a single step of this analytical *resummation* can be infinitesimally small, permitting for the formal definition of the derivatives [25]

$$\begin{aligned} \frac{dD_t}{dt} &\equiv \lim_{\delta t \to 0} \frac{D_{t+\delta t} - D_t}{\delta t} \quad \text{and} \\ \frac{dj_t}{dt} &\equiv \lim_{\delta t \to 0} \frac{j_{t+\delta t} - j_t}{\delta t}, \end{aligned} \tag{2.18}$$

whereby the *renormalization flow equations* can be formulated,

$$\begin{aligned} \frac{dD_t}{dt} &= \quad + \quad + \ \ldots \\ D\frac{dj_t}{dt} &= \quad + \quad + \ \ldots, \end{aligned} \tag{2.19}$$



which is a system of coupled differential equations for operators with boundary values $D_{t=0} = D$ and $j_{t=0} = j$. By solving these equations one obtains the fully renormalized quantities $\tilde{D} = D_{t=1}$, $\tilde{j} = j_{t=1}$, and the renormalized Wiener filter formula

$$\tilde{m} = \tilde{D}\tilde{j}. \tag{2.20}$$

This means, by solving Eq. (2.19), we finally calculate a Gaussian approximation to the correct posterior mean of s, $P(s|d) \approx \mathcal{G}(s - \tilde{m}, \tilde{D})$.

2.3. Self-calibration

Now we address the calibration problem, i.e., how to infer a physical signal field given a data set without precise knowledge of the signal response. We consider the case in which an external calibration is not possible (see Sec. 2.1). Thus, the instrument has to be self-calibrated during the measurement process. If we had absolutely no information about the signal response (how a measurement device transforms the signal into data) there would be absolutely no chance to infer the signal appropriately. However, if we have some information about the statistics of the response, e.g., the two point correlation function, this task becomes solvable. For this purpose we introduce the CURE method in the framework of IFT (Sec. 2.3.1) and review already existing methods (Sec. 2.3.2) to compare it against.

The aim is to calculate an optimal[4] estimator for the signal (or in general the moments $\langle s \ldots s \rangle_{(s|d)}$) given the data without exact information of the calibration. A way to approach

[4]Optimal in the sense of minimizing the $\mathcal{L}^2$-error.

this challenge is to consider the unknown calibration as a nuisance parameter, i.e., to marginalize over the calibration when calculating the signal posterior,

$$\mathcal{P}(s|d) = \int \mathcal{D}\gamma\ \mathcal{P}(s,\gamma|d) = \underbrace{\int \mathcal{D}\gamma\ \mathcal{P}(d,\gamma|s)}_{\mathcal{P}(d|s)}\ \frac{\mathcal{P}(s)}{\mathcal{P}(d)}, \tag{2.21}$$

which involves the calculation of the calibration marginalized likelihood. To do so, we assume the response to be a linear function in the calibration coefficients γ_a with Gaussian statistics, i.e. $R^\gamma \approx R^0 + \sum_a \gamma_a R^a$. The assumption of Gaussianity is appropriate as long as we have a priori no information about higher moments of γ, $\langle \gamma_1 \dots \gamma_n \rangle_{(\gamma)}$ with $n > 2$. The linearity can be considered as a first order approximation around $\gamma_0 = 0$ in γ,

$$\begin{aligned} R^\gamma &= R(\gamma_0) + \left.\frac{\partial R(\gamma)}{\partial \gamma_a}\right|_{\gamma=\gamma_0} (\gamma - \gamma_0) + \mathcal{O}(\gamma^2) \\ &= R^0 + \sum_a \gamma_a R^a + \mathcal{O}(\gamma^2). \end{aligned} \tag{2.22}$$

Under these assumptions one obtains [23, 28]

$$\begin{aligned} \mathcal{P}(d|s) &= \int \mathcal{D}\gamma\ \mathcal{P}(d|s,\gamma)\mathcal{P}(\gamma) \\ &= \int \mathcal{D}\gamma\ \mathcal{G}\left(d - \left(R^0 + \sum_a \gamma_a R^a\right) s, N\right) \mathcal{G}\left(\gamma, \Gamma\right) \\ &= \mathcal{G}\left(d - R^0 s, N + \sum_{ab} \Gamma_{ab} R^a s s^\dagger R^{b\dagger}\right). \end{aligned} \tag{2.23}$$

The data variance of this Gaussian likelihood, Eq. (2.23), depends on the correlation structure of the calibration, $\Gamma = \left\langle \gamma\gamma^\dagger \right\rangle_{(\gamma|\Gamma)}$, as well as on the signal s. This, in turn, results in a non-Gaussian posterior, $\mathcal{P}(s|d) \propto \mathcal{P}(d|s)\mathcal{P}(s)$, such that calculations of moments cannot be done analytically anymore. In principle one can adapt posterior sampling techniques like Markov Chain Monte Carlo (MCMC) methods to calculate, e.g., the posterior mean, m_{MCMC}. These approaches, however, are usually very expensive, which increases the attractivity of developing (semi-)analytical methods.

2.3.1. Calibration uncertainty renormalized estimator

Now, we apply the concept of uncertainty renormalization to the selfcal problem. According to Sec. 2.2.3 we introduce an expansion parameter $\delta t \ll 1$ in the ansatz:

$$\mathcal{P}(s|d) \propto \mathcal{G}\left(d - R^0 s, N + \delta t \sum_{ab} \Gamma_{ab} R^a s s^\dagger R^{b\dagger}\right) \mathcal{P}(s). \tag{2.24}$$

To simplify the notation we define an auxiliary para-meter $\Xi \equiv \sum_{ab} \Gamma_{ab} R^a s s^\dagger R^{b\dagger}$ and assume a Gaussian signal prior, $\mathcal{P}(s) = \mathcal{G}\left(s - s_0, S\right)$, with the a priori mean $s_0 \equiv \langle s \rangle_{(s)}$.

The Hamiltonian becomes

$$\begin{aligned}\mathcal{H}(d,s) =& -\ln \mathcal{P}(d,s)\\ =& -\ln\left[\mathcal{G}\left(d-R^0 s, N+\delta t\ \Xi\right)\mathcal{G}\left(s-s_0,S\right)\right]\\ =& \frac{1}{2}\ln|2\pi S| + \frac{1}{2}\ln|2\pi(N+\delta t\ \Xi)|\\ &+\frac{1}{2}\left(d-R^0 s\right)^\dagger (N+\delta t\ \Xi)^{-1}\left(d-R^0 s\right) + \frac{1}{2}(s-s_0)^\dagger S^{-1}(s-s_0).\end{aligned} \tag{2.25}$$

We can use that the expansion parameter δt is small, i.e. $\delta t\ \Xi \ll N$ (spectrally[5]), whereby the approximations

$$\begin{aligned}\ln|2\pi(N+\delta t\ \Xi)| \approx&\ \ln|2\pi N| + \mathrm{tr}\left(\delta t\ \Xi\ N^{-1}\right), \text{ and}\\ (N+\delta t\ \Xi)^{-1} \approx& \ N^{-1} - N^{-1}\delta t\ \Xi\ N^{-1}\end{aligned} \tag{2.26}$$

can be made. Using Eqs. (2.25), (2.26) yields

$$\mathcal{H}(d,s) = \mathcal{H}^{\text{free}} + \delta t \sum_{n=2}^{4}\frac{1}{n!}\lambda^{(n)}\left[s^{(n)}\right] \tag{2.27}$$

with

$$\begin{aligned}\mathcal{H}^{\text{free}} =& \ \mathcal{H}_0 + \frac{1}{2}s^\dagger D^{-1}s - j^\dagger s,\\ \lambda^{(2)}[s,s] =& \sum_{ab}\Gamma^{ab}\left(s^\dagger M^{ba}s - j^{a\dagger}ss^\dagger j^b\right) + 1 \text{ perm.},\\ \lambda^{(3)}[s,s,s] =& \sum_{ab}\Gamma^{ab}\left(\frac{1}{2}j^{a\dagger}ss^\dagger M^{b0}s + \text{cc.}\right) + 5 \text{ perm.},\\ \lambda^{(4)}[s,s,s,s] =& \sum_{ab}\Gamma^{ab}\left(-\frac{1}{2}s^\dagger M^{0a}ss^\dagger M^{b0}s\right) + 23 \text{ perm.},\end{aligned} \tag{2.28}$$

with permutations (perm.) with respect to s and the abbreviations

$$\begin{aligned}\mathcal{H}_0 =& \frac{1}{2}\ln|2\pi N| + \frac{1}{2}\ln|2\pi S| + \frac{1}{2}d^\dagger N^{-1}d + \frac{1}{2}s_0^\dagger S^{-1}s_0\\ D^{-1} =& \left(S^{-1} + R^{0\dagger}N^{-1}R^0\right),\\ j^\dagger =& \ d^\dagger N^{-1}R^0 + s_0^\dagger S^{-1},\\ M^{ab} =& \ R^{a\dagger}N^{-1}R^b,\\ j^{a\dagger} =& \ d^\dagger N^{-1}R^a.\end{aligned} \tag{2.29}$$

Terms higher than fourth order in the signal are dropped by making the approximation of Eq. (2.26).

[5]Means that $\xi^\dagger\delta t\ \Xi\ \xi \ll \xi^\dagger N\xi\ \forall\xi \in \mathbb{R}^m\backslash 0$.

2.3.1.1. Zero point expansion

Since the information Hamiltonian, Eqs. (2.27), (2.28), and (2.29), has the structure of Eq. (2.15), we can start to derive the renormalization flow equations. First, we consider (also for pedagogical reasons) the special case, in which the a priori signal mean is zero but the signal two point statistic is known, i.e., we use a zero centered, Gaussian prior, $\mathcal{P}(s) = \mathcal{G}(s, S)$.

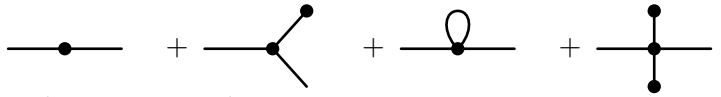
Following Sec. 2.2.3, the interaction terms of Eq. (2.27) (Eq. (2.28)) can be absorbed in a so-called renormalized information propagator $\tilde{D}_{\delta t}$ and information source $\tilde{j}_{\delta t}$ of order δt. Including this (first) correction these quantities read

$$
\begin{aligned}
\left(\tilde{D}_{\delta t}\right)_{xy} &= D_{xy} + \delta t\Big(- D_{xz}\lambda^{(2)}_{zv} D_{vy} - D_{xz}\lambda^{(3)}_{zvu}(Dj)_v D_{uy} - \frac{1}{2} D_{xz}\lambda^{(4)}_{zvur} D_{vu} D_{ry} \\
&\quad - \frac{1}{2} D_{xz}\lambda^{(4)}_{zvur}(Dj)_v (Dj)_u D_{ry}\Big) + \mathcal{O}\left(\delta t^2\right) \\
&= \quad + \ \delta t\Big(\quad + \quad + \quad \\
&\quad + \quad \Big) + \mathcal{O}\left(\delta t^2\right), \\
D_{xy}\left(\tilde{j}_{\delta t}\right)_y &= D_{xy}\Big[j_y + \delta t\Big(- \frac{1}{2}\lambda^{(3)}_{yzv} D_{zv} - \lambda^{(2)}_{yz}(Dj)_z - \frac{1}{2}\lambda^{(3)}_{yzv}(Dj)_z (Dj)_v \\
&\quad - \frac{1}{2}\lambda^{(4)}_{yzvu} D_{zv}(Dj)_u - \frac{1}{3!}\lambda^{(4)}_{yzvu}(Dj)_z (Dj)_v (Dj)_u\Big)\Big] + \mathcal{O}\left(\delta t^2\right) \\
&= \quad + \ \delta t\Big(\quad + \quad + \quad \\
&\quad + \quad + \quad \Big) + \mathcal{O}\left(\delta t^2\right).
\end{aligned}
\tag{2.30}
$$

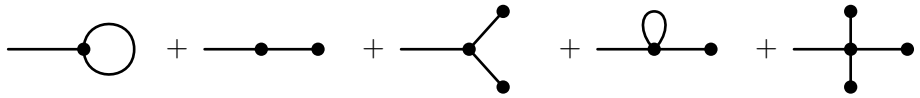
Just as a reminder, the vertices (internal dots) are multiplied by δt while the source terms (external dots) are independent of δt. In the diagrammatic expansions, Eq. (2.30), we place δt outside the brackets to underline this dependency. Therefore, to include all corrections up to order δt, we have to include all possible one-vertex diagrams. It is crucial to realize that δt cannot simply be set to one in order to obtain the fully renormalized propagator, $\tilde{D}$, because this step would violate Eq. (2.26). Appart from this it might also break down the perturbative expansion. However, instead of setting $\delta t = 1$ we can interpret $t \in [0, 1]$ as a pseudo-time, which measures the accumulated correction to the information propagator and source (see Sec. 2.2.3), $D_{t+\delta t}$ and $j_{t+\delta t}$. Thereby we can formulate the

renormalization flow equations,

$$\frac{dD_t}{dt} = \lim_{\delta t\to 0}\frac{D_{t+\delta t}-D_t}{\delta t}$$



$$D\frac{dj_t}{dt} = D\left(\lim_{\delta t\to 0}\frac{j_{t+\delta t}-j_t}{\delta t}\right)$$



(2.31)

which is a system of coupled differential equations for operators with boundary values $D_{t=0} = D$ and $j_{t=0} = j$. By solving these equations one obtains the fully renormalized quantities $\tilde{D} = D_{t=1}$, $\tilde{j} = j_{t=1}$, and the renormalized Wiener filter formula

$$\tilde{m} = \tilde{D}\tilde{j}. \tag{2.32}$$

However, instead of solving the coupled differential equations of Eq. (2.31) we could also solve the system where dD_t/dt is replaced by an equivalently valid equation for dD_t^{-1}/dt leading to the new differential system

$$\begin{aligned}
\frac{dD^{-1}_{t,xy}}{dt} &= \lambda^{(2)}_{xy} + \lambda^{(3)}_{xyz}(D_t j_t)_z + \frac{1}{2}\lambda^{(4)}_{xury}D_{t,ur} + \frac{1}{2}\lambda^{(4)}_{xvuy}(D_t j_t)_v(D_t j_t)_u,\\
\frac{dj_{t,y}}{dt} &= -\frac{1}{2}\lambda^{(3)}_{yzv}D_{t,zv} - \frac{1}{2}\lambda^{(3)}_{yzv}(D_t j_t)_z(D_t j_t)_v\\
&\quad - \lambda^{(2)}_{yz}(D_t j_t)_z - \frac{1}{2}\lambda^{(4)}_{yzvu}D_{t,zv}(D_t j_t)_u - \frac{1}{3!}\lambda^{(4)}_{yzvu}(D_t j_t)_z(D_t j_t)_v(D_t j_t)_u.
\end{aligned} \tag{2.33}$$

Solving these equations might simplify the numerical effort in some cases. Afterwards we invert $D^{-1}_{t=1} \equiv \tilde{D}^{-1}$ to finally solve Eq. (2.32).

2.3.1.2. Reference field expansion

There is also the option to introduce a residual field $\phi = s - m_0$ with respect to a reference field, e.g., $m_0 = Dj^0$ the Wiener filter solution without information of the proper calibration, Eq. (2.10). By deriving a Hamiltonian of ϕ the perturbative expansion gets more exact while the non-Gaussian terms become smaller. The Hamiltonian then reads

$$\mathcal{H}(d,\phi) = \mathcal{H}_0' + \frac{1}{2}\phi^\dagger D^{-1}\phi + \delta t\sum_{n=1}^{4}\frac{1}{n!}\Lambda^{(n)}\left[\phi^{(n)}\right], \tag{2.34}$$

where $\mathcal{H}_0'$ includes all ϕ-independent terms[6] and $\Lambda^{(n)}$ denotes the new (vertex-)tensor. Due to the fact that now already the source term is of $\mathcal{O}(\delta t)$ the diagrammatic expansion up to

[6]Note that among the ϕ-independent terms of $\mathcal{H}_0'$ are terms, collected in $\Lambda^{(0)}$, that depend on δt. These terms, however, only shift the Hamiltonian by a constant value but they do not influence its shape/structure.

order δt reduces to a sum of Feynman diagrams containing only a single source and single vertex term, given by

$$\begin{aligned}
\left(\tilde{D}_{\delta t}\right)_{xy} &= D_{xy} + \delta t \left(-D_{xz}\Lambda^{(2)}_{zv} D_{vy} - \frac{1}{2} D_{xz}\Lambda^{(4)}_{zvur} D_{vu} D_{ry} \right) \\
&= \text{——} \; + \; \delta t \Big(\text{——●——} \quad + \quad \text{—●◯} \quad \Big), \\
D_{xy}\left(\tilde{j}_{\delta t}\right)_y &= D_{xy}\left[\delta t \Big(- \Lambda^{(1)}_y - \frac{1}{2}\Lambda^{(3)}_{yzv} D_{zv} \Big) \right] \\
&= \delta t \Big(\text{——●} \; + \; \text{—●◯} \Big).
\end{aligned} \tag{2.35}$$

After restoring the original signal s by replacing the source term, $j_{\delta t} \to j_{\delta t} + D^{-1}_{\delta t} m_t$, $m_t \equiv D_t j_t$ [23], this leads in analogy to the previous section to the renormalization flow equations,

$$\begin{aligned}
\frac{dD_{t,xy}}{dt} &= -D_{xz}\Lambda^{(2)}_{zv} D_{vy} - \frac{1}{2} D_{xz}\Lambda^{(4)}_{zvur} D_{vu} D_{ry}, \\
&\quad \text{or alternatively} \\
\frac{dD^{-1}_{t,xy}}{dt} &= \Lambda^{(2)}_{xy} + \frac{1}{2}\Lambda^{(4)}_{xury} D_{t,ur}, \text{ and} \\
\frac{dj_{t,y}}{dt} &= -\Lambda^{(1)}_y + \Lambda^{(2)}_{yz}(D_t j_t)_z - \frac{1}{2}\Lambda^{(3)}_{yzv} D_{t,zv} + \frac{1}{2}\Lambda^{(4)}_{yzvu} D_{t,zv}(D_t j_t)_u,
\end{aligned} \tag{2.36}$$

with boundaries $j_{t=0} = j^0$ and $D^{-1}_{t=0} = D^{-1}$. Note that the positive terms in the differential equation of j_t arise from the restoration of the original signal.

Further note that the gained simplicity in the diagrammatic expansion has turned into a higher complexity of the vertex structure. The explicit structure of these vertices can be found in App. A.2. These are also implemented for our numerical example, see Sec. 5.2.2 and Fig. 2.1. The effect of the resummation process (involving absolute calibration measurements, see Sec. 5.2.2) on the information propagator is illustrated by Fig. 2.2.

2.3.1.3. Approach optimization

Until now, the vertex-tensors $\Lambda^{(n)}$ were pseudo-time independent (see, for instance, Eq. (A.2)). However, the CURE approach can in principle be improved if we replace the residual field $\phi = s - m_0$ after every timestep by $\tilde{\phi}_t = s - m_t = s - D_t j_t$. This way the support point of the expansion is always chosen optimally so that the first term of $\Lambda^{(1)}$ does still vanish (see Eq. (A.2)) and the definition of time derivatives, Eq. (2.31), remains valid.

2.3.2. Self-calibration schemes

To compare the derived CURE method not only against the Wiener filter solution without information of calibration (which is the starting value of CURE) but also to two other iterative self-calibration (selfcal) schemes we review the basic ideas of the latter briefly. A full description of the following methods can be found in Ref. [23]. The response is still considered to be linear, $R^\gamma = R^0 + \sum_a \gamma_a R^a$.

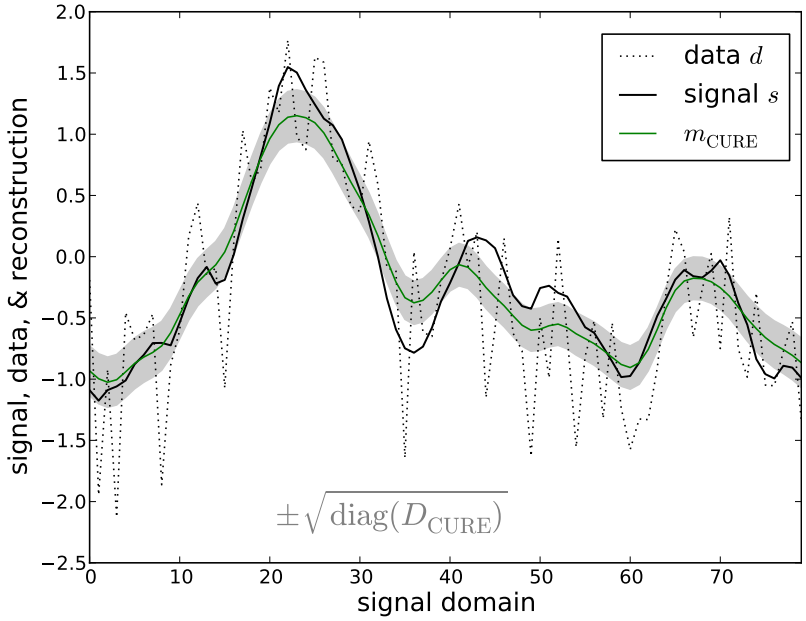


Figure 2.1.: Signal, data, and signal reconstruction (considering unknown calibration) with related 1σ-uncertainty according to Eq. (2.36) for the numerical example described in Sec. 5.2.2. For comparison to other methods see Fig. 2.3. Calibration and its reconstructions are shown in Fig. 2.4.

2.3.2.1. Classical selfcal

Classical selfcal is an iterative method, alternately inferring the signal while assuming the calibration to be known and vice versa until a fix-point is reached. The respectively inferred quantities $s^\star$ and $\gamma^\star$ are often maximum a posterori (MAP) estimators. This procedure of simultaneously estimating s and γ can be identified with searching for the maximum of the joint posterior, $\mathcal{P}(\gamma, s|d)$, or equally for the minimum of the joint information Hamiltonian [23], $\mathcal{H}(d, \gamma, s) = -\ln[\mathcal{P}(d, \gamma, s)]$, given by

$$\left.\frac{\partial \mathcal{H}(d,\gamma,s)}{\partial \gamma_a}\right|_{\gamma=\gamma^\star} = 0 \text{ and } \left.\frac{\partial \mathcal{H}(d,\gamma,s)}{\partial s}\right|_{s=s^\star} = 0. \tag{2.37}$$

The resulting equations (Eq. (2.39) with $T = 0$) must be iterated until convergence.

2.3.2.2. New selfcal

The new selfcal method is based on the above described idea of classical selfcal. However, in marked contrast to the latter new selfcal uses the signal marginalized posterior to infer the calibration, and determines a signal estimate under the usage of the resulting calibration estimate and its uncertainty afterwards. Therefore, the gradient and Hessian of the Hamiltonian, $\mathcal{H}(d, \gamma) = -\ln \int \mathcal{D}s\ \mathcal{P}(d, \gamma, s)$, have to be calculated to find the MAP esti-

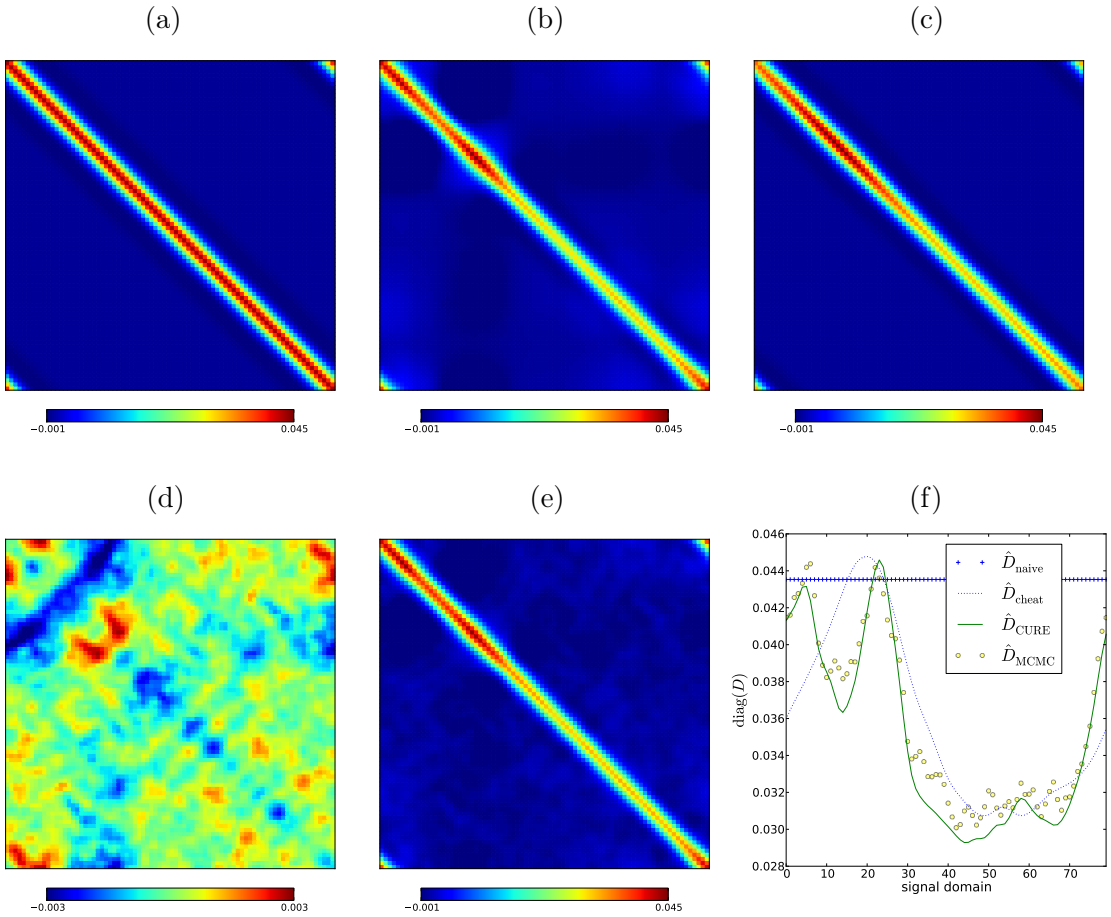


Figure 2.2.: Explicit structure of propagator operators for the realization shown in Fig. 2.1, 2.3, and 2.4. The figures (a)-(c), (e) refer to the propagators D_{method} with method = naive (a), CURE (b), cheat (c), and MCMC (e) according to Eq. (2.10) with unknown calibration set to zero, Eq. (2.36), and Eq. (2.10) with known calibration, respectively. Upper panels & lower middle panel: The renormalized propagator exhibits the same diagonal structure as the MCMC propagator. Lower panels (left, right): Comparison of the renormalized propagator to the MCMC result [CURE - MCMC, (d)] and explicit structures of the propagator diagonals (denoted by $\hat{D}_{\mathrm{method}}$, (f)). Emerging from the process of resummation (involving absolute calibration measurements, see Sec. 5.2.2), Eq. (2.36), the renormalized propagator obtains a non-diagonal structure due to the complex, non-local vertex structure of the non-Gaussian contributions to the Hamiltonian, see in particular Eq. (A.2).

mate $\gamma^\star$ and its uncertainty Δ, given by

$$\left.\frac{\partial \mathcal{H}(d,\gamma)}{\partial \gamma_a}\right|_{\gamma=\gamma^\star} = 0 \text{ and } \left.\frac{\partial^2 \mathcal{H}(d,\gamma)}{\partial \gamma_a \partial \gamma_b}\right|_{\gamma=\gamma^\star} \equiv \Delta^{-1}_{ab}. \tag{2.38}$$

By following Ref. [23], but skipping here the full derivation, we obtain the resulting calibration formula,

$$\begin{aligned}
\gamma^\star &= \Delta h, \\
\Delta^{-1}_{ab} &= \Gamma^{-1}_{ab} + \mathrm{tr}\left[\left(mm^\dagger + TD\right) M^{ab}\right], \text{ and} \\
h_b &= m^\dagger j^b - \mathrm{tr}\left[\left(mm^\dagger + TD\right) M^{ab}\right], \text{ with} \\
T &= \begin{cases} 1 & \text{for new selfcal} \\ 0 & \text{for classic selfcal} \end{cases}.
\end{aligned} \tag{2.39}$$

Note that the Wiener filter signal estimate $m = m(\gamma^\star)$ and its uncertainty $D = D(\gamma^\star)$ still depend on the calibration and thus Δ of Eq. (2.39) is not exactly the one of Eq. (2.38) [23]. For further details, as well as an extensive discussion of the selfcal methods we want to point to Ref. [23].

2.4. Numerical Example

2.4.1. Setup & results

To demonstrate the efficiency of the derived CURE approach we address the illustrative, simplistic, one-dimensional example used in Ref. [23] and perform a direct comparison to the selfcal schemes and MCMC sampling (see, e.g., Ref. [29]). There, a measurement device with a perfect point-like response scans a signal field s over a (periodic) domain $\Omega = \{x\}_x = [0, 1) \subset \mathbb{R}$ within a time $t \in [0, 1) \subset \mathbb{R}$, but with a time dependent calibration uncertainty, given by the calibration coefficients γ_t. This instrument exhibits a sampling rate[7] of $1/\tau = 80$ so that the ith data point, measured at time $t = i\tau$, is related to the signal at position $x_t = i\tau$. During the measurement process spatial and temporal coordinates are aligned and the data are given by

$$d_t = R_{tx} s_x + n_t = (1 + \gamma_t)\delta(x - x_t)s_x + n_t, \tag{2.40}$$

where the signal, measurement noise n_t, and calibration coefficients γ_t are Gaussian with $\mathcal{G}(s, S)$, $\mathcal{G}(n, N)$, and $\mathcal{G}(\gamma, \Gamma)$ the corresponding PDF's with related covariance matrices S, N, and Γ. These are assumed to be known[8] and might be described by their respective power spectra in Fourier space. Following Ref. [23] we use

$$\begin{aligned}
\mathcal{P}_s(k) &= \frac{a_s}{(1 + (k/k_s)^2)^2}, \\
\mathcal{P}_\gamma(w) &= \frac{a_\gamma}{(1 + (w/w_\gamma)^2)^2}, \text{ and} \\
\mathcal{P}_n(w) &= a_n.
\end{aligned} \tag{2.41}$$

By Eq. (2.41) the amplitudes $a_s = \sigma_s^2 \lambda_s$, $a_\gamma = \sigma_\gamma^2 \tau_\gamma$, and $a_n = \sigma_n^2 \tau_n$ with related variances $\sigma^2_{s,\gamma,n}$ and correlation lengths $\lambda_s = 4/k_s$, $\tau_\gamma = 4/\omega_\gamma$, and $\tau_n = \tau$ have been introduced.

[7] Since this work is supposed to be a proof of concept we work with explicit matrices and tensors, whereby we have to limit the size of the problem for computational reasons. Further investigations are needed on how to transform this into a method using implicit tensors, and therefore suitable for "big data" problems.

[8] In case they are unknown there exist well known methods which are able to extract the correlation structure simultaneously from data, see, e.g., Ref. [25]

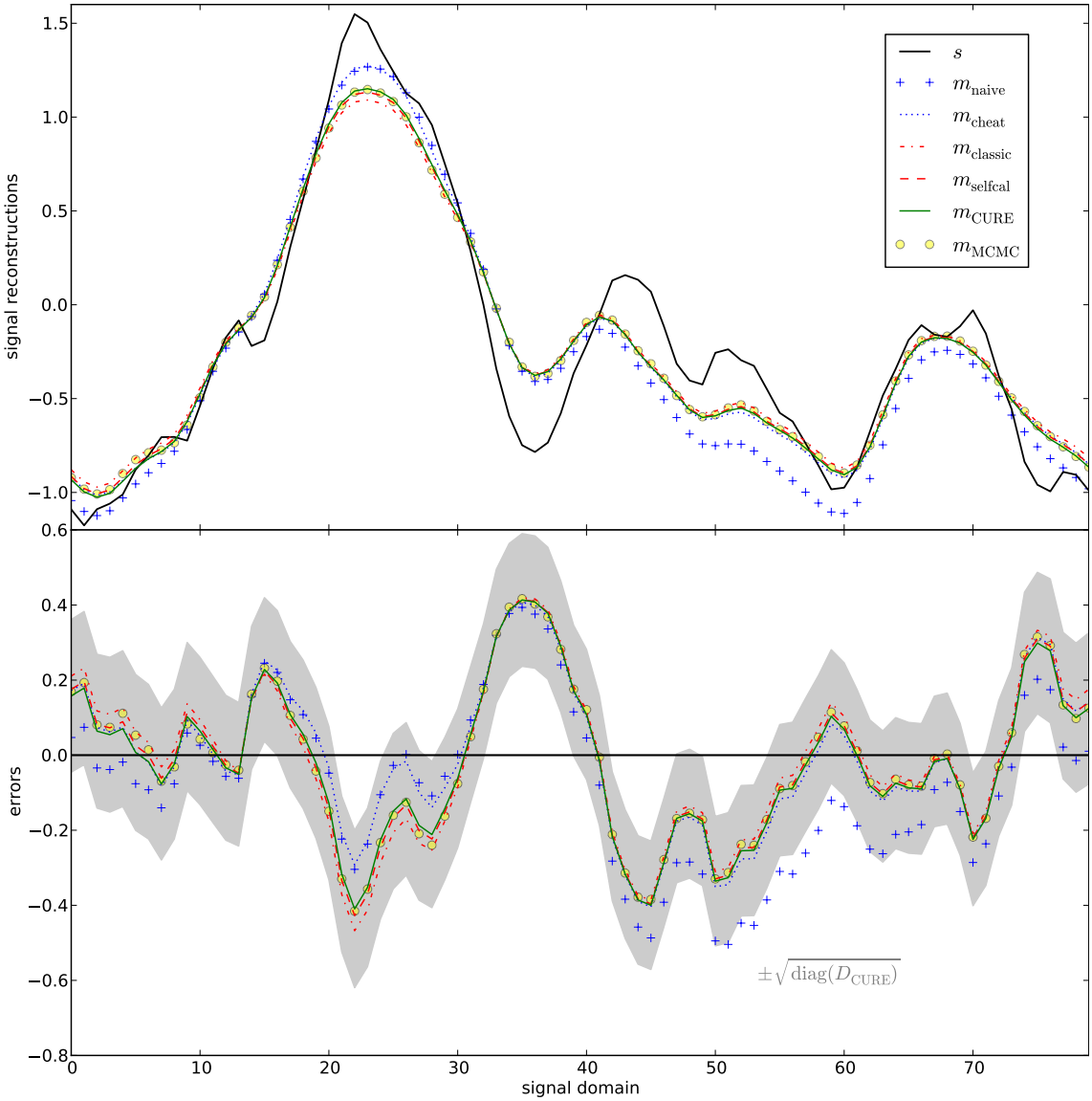


Figure 2.3.: Signal reconstructions and related errors of different approaches. The following terminology is used: naive: Wiener filter with unknown calibration set to zero; classic: classical selfcal (Eq. (2.43), $T = 0$); CURE (Eqs. (2.36) and (A.2)); selfcal: new selfcal (Eq. (2.43), $T = 1$); cheat: Wiener filter with known calibration; MCMC: Markov Chain Monte Carlo sampling. The gray shaded region represents the 1σ uncertainty of the CURE method.

Within the numerical implementation we use the values $\sigma_s = 1$, $\sigma_\gamma = 0.3$, $\sigma_n = 0.5$, $\lambda_s = 0.3$ and $\tau_\gamma = 1.5$. This means we get an unit variance signal with calibration uncertainty of 30% and noise of 50%, which is still white (percentage values with respect to the typical signal strength).

Relating to Ref. [23] we also introduce so-called absolute calibration measurements to have additional information about the calibration that is beneficial to break the global degeneracy of the data with respect to signal and calibration variations. This means, we switch off the signal for four particular times $t_i \in \{0, 0.25, 0.5, 0.75\}$, where the calibration has the strength $c = 4$. Here, the data d' is given by

$$d'_{t_i} = (1 + \gamma_{t_i})c + n'_{t_i}. \tag{2.42}$$

During these measurements we assume the same noise statistics as before, $n' \hookleftarrow \mathcal{G}(n', N)$.

Including the absolute calibration measurements the iterative selfcal equations, Eq. (2.39), become [23]

$$\begin{aligned}
\gamma^\star &= \Delta h, \\
\Delta^{-1}_{tt'} &= \Gamma^{-1}_{tt'} + \sigma_n^{-2}\delta_{tt'}\left(q_t + c^2\sum_i \delta_{tt_i}\right), \\
h_t &= \sigma_n^{-2}\left(d_t m_{x_t} - q_t + c^2\sum_i \delta_{tt_i} d'_i\right), \text{ and} \\
q_t &= m^2_{x_t} + TD_{x_t x_t} \text{ with} \\
T &= \begin{cases} 1 & \text{for new selfcal} \\ 0 & \text{for classic selfcal} \end{cases}.
\end{aligned} \tag{2.43}$$

To apply the CURE approach including the absolute calibration measurements we have to solve the ordinary differential equation of first order, according to Eq. (2.33) or Eq. (2.36), depending on whether the *zero-point* or *reference field* expansion is used. We present here the more general, but more complex version of the *reference field* expansion, Eqs. (2.35) and (A.2), because this version is constructed to deal with a larger uncertainty of the calibration than the *zero-point* expansion. To solve Eq. (2.36) we use the ordinary differential equation solver of `scipy`[9] with integrator settings: `vode`, method = `adams`. All numerical calculations have been performed using NIFTy[10] [30].

Figs. 2.3 and 2.4 show a typical result for signal and calibration reconstruction, respectively. Fig. 2.5 and Tabs. 5.1, 2.2, and 2.3 show the squared error averages of the different calibration methods according to Eq. (2.44) at a given number of realizations[11] for signal and calibration, where the following terminology is used,

$$\begin{aligned}
\Delta^s_i &\equiv \left\langle (s - m_i)^\dagger (s - m_i) \right\rangle_{(d,s,\gamma)}, \\
\Delta^\gamma_i &\equiv \left\langle (\gamma - \gamma_i)^\dagger (\gamma - \gamma_i) \right\rangle_{(d,s,\gamma)} \text{ with} \\
i &= \text{naive, cheat, classic, selfcal,} \\
&\quad \text{CURE, and MCMC,}
\end{aligned} \tag{2.44}$$

[9] `http://docs.scipy.org/doc/scipy/reference/generated/scipy.integrate.ode.html`
[10] `http://www.mpa-garching.mpg.de/ift/nifty`
[11] Note that for the statistics of 500 realizations we use a four times coarser sampling rate.

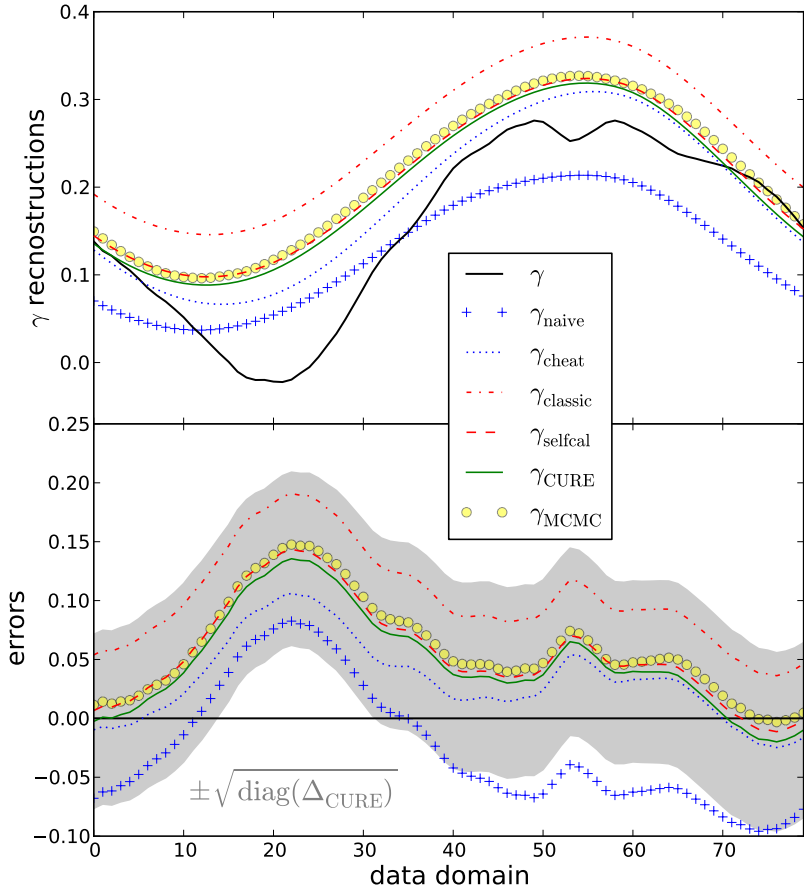


Figure 2.4.: Calibration reconstructions and related errors of different approaches using Eq. (2.43). The terminology is used following Fig. 2.3. The gray shaded region represents the 1σ uncertainty of the CURE method. The reconstruction of cheat is not perfect, because Eq. (2.43) uses the Wiener filtered data (assuming the correct calibration, but non-neglectable noise). The relatively good result of naive is a pure coincidence.

referring to the Wiener filter methods without and with information of calibration, the classic and new selfcal scheme, the CURE scheme, and MCMC sampling, respectively.

2.4.2. Discussion

As Figs. 2.3, 2.4, and in particular Fig. 2.5 with related Tabs. 5.1, 2.2, and 2.3 illustrate, the CURE and new selfcal (selfcal) approach prevail against classical selfcal (classic) and Wiener filtering with unknown calibration (naive) and perform similar to the MCMC method. The latter represents in principle the best method by avoiding any approximations, but also the most expensive one. Its small underperformance in comparison to CURE and selfcal has its origin in using still not sufficient samples for the MCMC-chains[12] to converge fully. Increasing their number, however, would increase the numerical effort significantly.

[12]For each signal realization we have to run a separate chain. In the numerical example used in this work, a single chain consists of 2×10^3 independent samples.

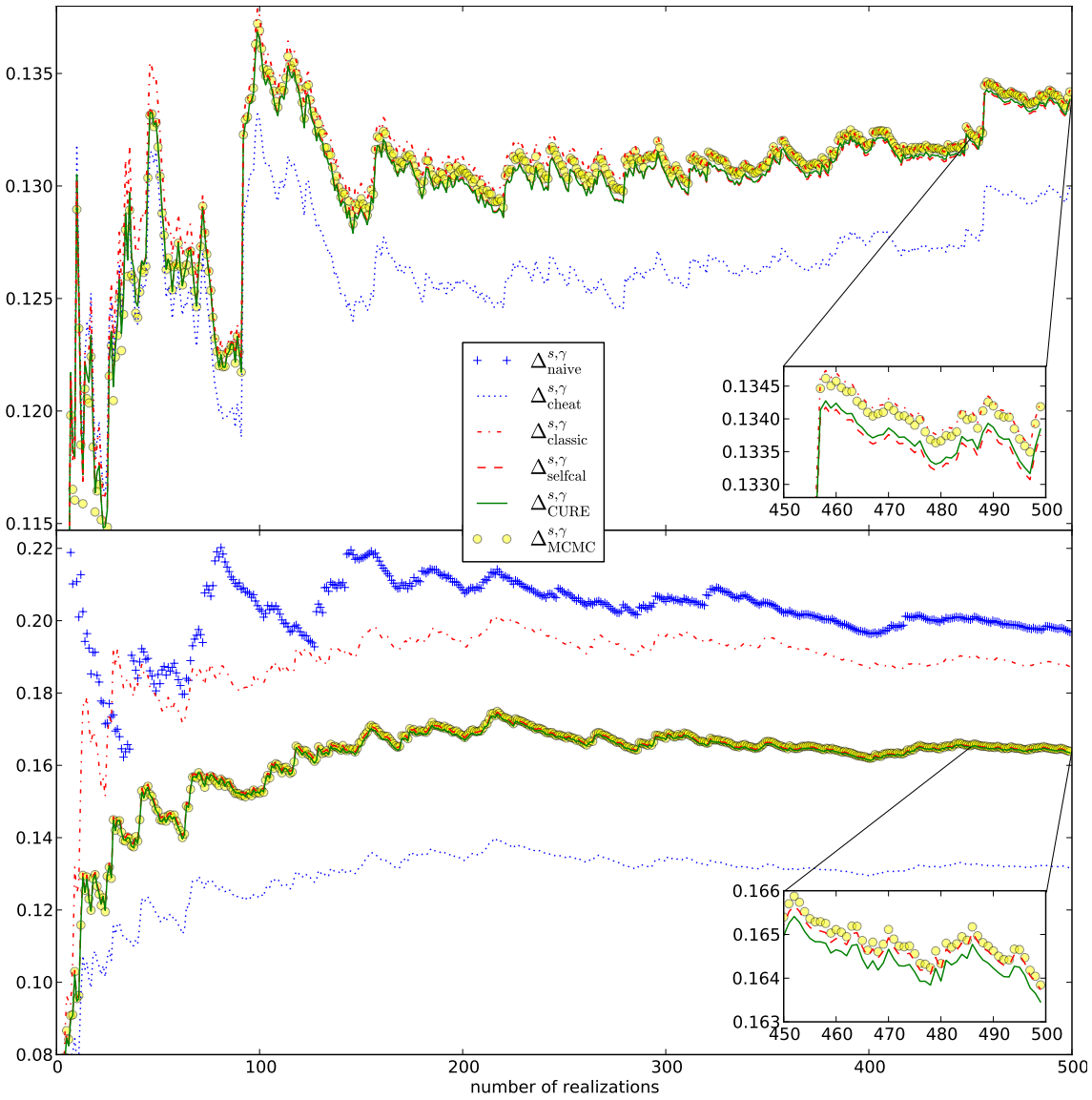


Figure 2.5.: Squared error averages according to Eq. (2.44) at a given number of realizations for signal (upper panel) and calibration (lower panel). The best and worst result for signal and calibration yields the cheat and naive method, respectively. In the signal domain (upper panel) all three advanced methods are very close to each other, although there is a slight preference for the CURE and selfcal method followed by MCMC and classic. The results of the naive method are beyond the range of the upper panel. For the inference of calibration (lower panel) CURE and selfcal perform clearly better than classic and very similar to MCMC (see Tab. 5.1, 2.2, and 2.3).

Table 2.1.: Squared errors of signal and calibration for all methods, averaged over 500 realizations, see Fig. 2.5.

i	Δ_i^s	Δ_i^γ
naive	0.1635	0.1968
cheat	0.1300	0.1316
classic	0.1343	0.1873
selfcal	0.1338	0.1637
CURE	0.1338	0.1635
MCMC	0.1342	0.1638

Table 2.2.: Improvements of the methods' signal squared errors with respect to the naive method, averaged over 500 realizations.

i	$\Delta^s_{\text{naive}} - \Delta^s_i$	improvement
naive	0.0000	0.00%
cheat	0.0335	100.00%
classic	0.0292	87.16%
selfcal	0.0297	88.86%
CURE	0.0297	88.66%
MCMC	0.0293	87.46%

The upside of CURE is that it is not iterative since it only involves the solution of a single system of coupled ordinary differential equations (ODE's). For ODE's, in turn, exist a number of well working numerical solvers with adaptive stepsize control, which might save significant amounts of computational time[13]. This is, however, only true if one finds a clever implementation or sparse representation of $\Lambda^{(3)}$, because the term $\Lambda^{(3)}_{yzv} D_{zv}$, required in Eq. (2.35) might become a bottleneck within a calculation due to its complex correlation structure (contrary to the $\Lambda^{(4)}$ term). Another downside is the higher level of complexity in comparison to new selfcal that naturally arises with a renormalization calculation.

[13]We want to mention that we also found realizations for certain levels of signal, noise, and calibration where the `scipy` solver did not converge. The reason for this might be a initial guess too far away from the correct solution. In many cases one could cope with this problem by significantly reducing the stepsize, δt, or using the optimization described in Sec. 2.3.1.3. We, however, did not elaborate on this since this work is supposed to be a proof of concepts.

Table 2.3.: Improvements of the methods' calibration squared errors with respect to the naive method, averaged over 500 realizations.

i	$\Delta^\gamma_{\text{naive}} - \Delta^\gamma_i$	improvement
naive	0.0000	0.00%
cheat	0.0652	100.00%
classic	0.0095	14.57%
selfcal	0.0331	50.77%
CURE	0.0333	51.07%
MCMC	0.0330	50.61%

2.5. Concluding remarks

We derived the calibration uncertainty renormalized estimator (CURE) method to infer a signal and consequently the calibration without knowledge of the calibration but its covariance. The basic idea of CURE is to perform a perturbation calculation around a reference field, an a priori determined reconstruction of the signal without knowledge of calibration. Perturbatively means that we successively take into account higher-order terms of calibration uncertainty. This way, the problem of signal reconstruction without knowledge of the calibration, which is often solved by iterative or brute-force sampling methods, turns into a single system of ordinary differential equations.

We applied the method to a mock example and compared it against other existent calibration methods. For this example we found that CURE performs extremely similar to new selfcal and MCMC sampling, and clearly beats the Wiener filter without calibration as well as the classic selfcal method in terms of reconstruction accuracy. Although it obviously performs well, a recommendation to favor this method over new and classical selfcal depends on the particular problem at hand as well as on the numerical implementation, as discussed in Sec. 2.4.2. Therefore it serves as an alternative to them.

3. Generic inference of inflation models by non-Gaussianity and primordial power spectrum reconstruction

Note: This chapter has been published in JCAP **6**, 048 (2014) [31].

3.1. Introduction

3.1.1. Motivation

By precision measurements of the cosmic microwave background (CMB) [32, 33] it has become possible to determine the exact statistics of its temperature anisotropies. These anisotropies are strongly connected to the curvature perturbations on uniform density hypersurfaces ζ, predicted by inflationary models, with the result that the zoo of models can be constrained by exploiting observational data, e.g., by the usage of Gaussian statistics [34–38]. The viable models that are compatible with current *Planck* constraints on primordial non-Gaussianity, often represented by the $f_{\rm NL}$ parameter, are given by $-3.1 \leq f_{\rm NL} \leq 8.5$ (68%C.L. statistical) [39] for the local type of non-Gaussianity. In particular, a value of $|f_{\rm NL}| \propto \mathcal{O}(1)$ is in agreement with the data. Such a low value of non-Gaussianity opens the possibility to include the effect of primordial non-Gaussianity when performing routine cosmological parameter estimates in order to maximally exploit the data, since it permits approximations which prevent the computations from becoming numerically too expensive. The contributions from higher-order statistics can in many cases (see Sec. 3.3) be parametrized by the local non-Gaussianity parameter $f_{\rm NL}$ and $g_{\rm NL}$ [40],

$$\zeta = \zeta_1 + \frac{3}{5} f_{\rm NL} \zeta_1^2 + \frac{9}{25} g_{\rm NL} \zeta_1^3 + \mathcal{O}(\zeta_1^4), \tag{3.1}$$

where ζ_1 is the Gaussian curvature perturbation. $f_{\rm NL}$ contributes to the bi- and trispectrum, while $g_{\rm NL}$ contributes only to the trispectrum of the curvature perturbation.

As things turned out, there are inflation models among the ones, which are favored by current data, e.g., stated in Ref. [37] (AI, BI, ESI, HI, LI, MHI, RGI, SBI, SFI)[1] or Ref. [38], predicting values of $|f_{\rm NL}| \propto \mathcal{O}(1)$ and distinctly deviate from $g_{\rm NL} = 0$ if the possibility of non-Gaussianity is taken into account. It is crucial to realize that it is less likely for (at least) two disjunct inflation models to predict the same combination $(f_{\rm NL}, g_{\rm NL})$ than only the same value of $f_{\rm NL}$ or $g_{\rm NL}$. In other words, if we would be able to infer these two non-Gaussianity parameters simultaneously from CMB or large scale structure (LSS) data, we had a powerful tool to distinguish between the remaining inflation models. This requires to derive a posterior probability density function (pdf) for $(f_{\rm NL}, g_{\rm NL})$ within a

[1]Terminology according to Ref. [34]: AI = Arctan Inflation, BI = Brane Inflation, ESI = Exponential SUSY Inflation, HI = Higgs Inflation, LI = Loop Inflation, MHI = Mutated Hilltop Inflation, RGI = Radiation Gauge Inflation, SBI = Supergravity Brane Inflation, SFI = Small Field Inflation.

Bayesian framework. How this can be done analytically is presented in the first part of this paper. Additionally, we show how this method can be recast to infer parameters specific to inflationary models, e.g., shape parameters of inflationary potentials, or the presence of an additional bosonic field, directly from data. We also provide a validation of our approach to show its precision despite using an approximation.

The second quantity of interest here is the primordial power spectrum, $P_\zeta(k)$ or $P_{\zeta_1}(k)$, in particular due to its constraining character with respect to inflationary scenarios. The *Planck* collaboration might have seen some features within the primordial power spectrum which in turn would indicate non-linear physics and thus could point to inflation models beyond single-field slow-roll inflation [38]. Additionally, these types of deviations are well motivated by, e.g., implications of the recent BICEP2 data [41–44], or special features of the inflaton potential [45, 46]. However, for the detection of such features one has to appropriately reconstruct the power spectrum from observational data. For this purpose we suggest two non-parametric spectral inference methods in Sec. 3.5.

3.1.2. Previous Bayesian work

The majority of publications [21, 47–52], which are dealing with Bayesian reconstructions of non-Gaussian quantities from CMB have their focus only on estimators or the pdf of the $f_{\rm NL}$ or $g_{\rm NL}$ parameter. They usually require computationally expensive calculations like Monte Carlo sampling except for some, e.g. Refs. [53, 54], which derive analytic expressions by performing approximations.

High precision CMB measurements of the *WMAP* and *Planck* satellites have opened a new window to the physics of the early Universe and have thus improved the constraints on some parameters of non-Gaussianity [39] and on many inflation models [38] based on the two-point function, but have not connected the inflationary parameters directly to higher-order statistics. A way of direct inference of single-field slow-roll inflation models from CMB data of the *Planck* satellite was recently presented by Refs. [35–37]. Here, the CMB power spectrum was analyzed already ruling out a huge amount of inflation models. We, however, go beyond Gaussian and three-point statistics to achieve tighter constraints on reasonable, not necessarily single-field slow-roll inflation models given the *Planck* and future data.

An independent cross-check of CMB results is the analysis of the LSS data. Current results for non-Gaussianity values, e.g. Refs. [55–61] and forwarding references thereof, are consistent with CMB constraints. Thus, the LSS provides also a natural data set to infer inflation models. The inference approach presented in this paper is in principle able to deal with this type of data sets as well (see Sec. 3.4).

According to the reconstruction of the primordial power spectrum, there exist a huge amount of approaches and an overview of the literature can be found in section 7 of Ref. [38] and in Ref. [62]. Within this work we exclusively focus on the approach of Refs. [25] and [63], which developed approximative, but inexpensive Bayesian inference schemes for spectra within the framework of information field theory.

For a brief review on inferring primordial non-Gaussianities in the CMB beyond Bayesian techniques (e.g., bi- and trispectrum estimators, Minkowski Functionals, wavelets, needlets, etc.) we want to point to Ref. [64] and forwarding references thereof.

3.1.3. Structure of the work

The remainder of this work is organized as follows. In Sec. 3.2 we describe the considered data model and introduce the generic method of inferring inflation models postulating $f_{\rm NL}$, $g_{\rm NL}$. In Sec. 3.3 we review a few inflation models that are not ruled out by current *Planck* data and quote corresponding expressions for $f_{\rm NL}$ and $g_{\rm NL}$. Additionally, we show where the specific models are localized in the $f_{\rm NL}$-$g_{\rm NL}$-plane. The Bayesian posterior for special inflationary parameters is shown in Sec. 3.4 as well as a numerical implementation (toy case) of the also pedagogically important curvaton scenario in the Sachs-Wolfe limit and its validation by the Diagnostics of Insufficiencies of Posterior distribution (DIP) test [65]. In Sec. 3.5 we introduce a method to reconstruct the primordial power spectrum of ζ and ζ_1. We summarize our findings in Sec. 3.6.

Being at the interface of statistical analysis and physical cosmology, it seems appropriate to guide the reader by giving some reading instructions. For a reader who is rather interested in the statistical analysis, i.e. how to infer (inflationary) parameters from CMB data and how to reconstruct a power spectrum in a non-parametric way in general, paragraphs starting with symbol ▶ and ending with symbol ◀ might be skipped. For a reader rather interested in physical cosmology these symbols might mark paragraphs of special interest.

3.2. Generic inference of inflation models postulating $f_{\rm NL}$, $g_{\rm NL}$

In order to decide which inflation model is favored by current CMB or LSS data one should use as much information as possible during the inference process without becoming numerically too expensive. This implies, in particular, to aim for information sensitive to non-Gaussian statistics. Usually, this leads to non-trivial phase space integrals which cannot be performed analytically and require numerically expensive techniques like Monte Carlo sampling. Within this section, however, we show how to set up a fully analytic posterior for the scalar, local non-Gaussianity parameters $f_{\rm NL}(p)$ and $g_{\rm NL}(p)$, which in general depend on inflation or reheating model specific parameters, $p = (p_1, \ldots, p_u)^T \in \mathbb{R}^u,\ u \in \mathbb{N}$. In turn, this also enables to calculate analytically the posterior pdf for the model specific parameters p, which encode, e.g., the particular shape of an inflation model or the density fraction of an additional bosonic field (see Sec. 3.3). To keep this analyticity and simultaneously avoid numerically expensive sampling techniques we introduce a saddle-point approximation in the actual section, whose sufficiency is validated in Sec. 3.4.2.3.

For reasons of clarity and comprehensibility we drop the p-dependency in our notation within this section. For the same reason we focus on global values of $f_{\rm NL}$ and $g_{\rm NL}$ although the formalism described below is generic and can deal with spatially varying non-Gaussianity parameters as shown in Ref. [21].

3.2.1. Data model

To infer physical quantities from data we have to agree on a particular data model. Following the logic of information field theory [21], a CMB observation is represented by a discrete data tuple $d = (d_1, \ldots, d_m)^T \in \mathbb{R}^m,\ m \in \mathbb{N}$, composed of uncorrelated Gaussian noise $n = (n_1, \ldots, n_m)^T \in \mathbb{R}^m$ and a linear response operation R acting on the, in general, non-Gaussian comoving curvature perturbations ζ, a continuous physical field over the

Riemannian manifold $\mathcal{U}$,

$$d = \frac{\delta T}{T_{\rm CMB}} = R\zeta + n = R\left(\zeta_1 + \frac{3}{5} f_{\rm NL}\zeta_1^2 + \frac{9}{25} g_{\rm NL}\zeta_1^3 + \mathcal{O}\left(\zeta_1^4\right)\right) + n, \tag{3.2}$$

with the Gaussian curvature perturbations ζ_1. The pdfs of ζ_1 and n are given by $P(\zeta_1) = \mathcal{G}(\zeta_1, \Xi)$ with covariance $\Xi = \left\langle \zeta_1 \zeta_1^\dagger \right\rangle_{(\zeta_1|\Xi)}$, and $P(n) = \mathcal{G}(n, N)$ with covariance N, respectively. Here, we use the notation

$$\langle\ .\ \rangle_{P(a)} = \langle\ .\ \rangle_{(a|A)} \equiv \int \mathcal{D}a\ .\ P(a|A), \tag{3.3}$$

and

$$\mathcal{G}(a, A) \equiv |2\pi A|^{-1/2}\ \exp\left(-\frac{1}{2} a^\dagger A^{-1} a\right), \tag{3.4}$$

where $\dagger$ denotes a transposition and complex conjugation, $*$, and $a^\dagger b \equiv \int_{\mathcal{U}} d^{\rm d}x\ a^*(x) b(x)$ with $\mathrm{d} \equiv \dim\mathcal{U}$ defining the inner product on the fields a, b. The comoving curvature perturbation ζ on uniform density hypersurfaces, which is a conserved quantity outside the horizon[2] [68], is the seed of the structure growth during the evolution of the Universe and its statistics are precisely predicted by inflation models. Therefore, ζ is directly related to inflationary parameters, p. If the statistics of ζ, predicted by inflation scenarios, are non-Gaussian, the dependence on p can often be absorbed in the non-Gaussianity parameters $f_{\rm NL}(p)$ and $g_{\rm NL}(p)$. The linear response R in Eq. (3.2) transfers the curvature perturbations into temperature deviations, δT, and contains all instrumental and measurement effects, i.e. R represents the radiation transfer function. In this way the data is directly related to the initial Gaussian curvature perturbation ζ_1 or to inflationary parameters p and we can set up the inference scheme.

3.2.2. Posterior derivation

We derive the posterior by following Ref. [53], i.e. we first calculate the pdf for the Gaussian curvature perturbation ζ_1 given the non-Gaussianity parameters and data via Bayes theorem [26],

$$\begin{aligned} P(\zeta_1|d, f_{\rm NL}, g_{\rm NL}) &= \frac{P(\zeta_1, d, f_{\rm NL}, g_{\rm NL})}{P(d, f_{\rm NL}, g_{\rm NL})} \\ &= \frac{P(d|\zeta_1, f_{\rm NL}, g_{\rm NL}) P(\zeta_1|f_{\rm NL}, g_{\rm NL})}{P(d|f_{\rm NL}, g_{\rm NL})} \equiv \frac{1}{\mathcal{Z}} e^{-H(\zeta_1, d|f_{\rm NL}, g_{\rm NL})}, \end{aligned} \tag{3.5}$$

where $H(\zeta_1, d|f_{\rm NL}, g_{\rm NL}) \equiv -\ln[P(d|\zeta_1, f_{\rm NL}, g_{\rm NL}) P(\zeta_1|f_{\rm NL}, g_{\rm NL})]$ defines the information Hamiltonian and $\mathcal{Z} \equiv P(d|f_{\rm NL}, g_{\rm NL})$ the partition function. Assuming Gaussian noise, $\mathcal{G}(n, N)$, with $N = \left\langle n n^\dagger \right\rangle_{(n|N)}$ denoting the noise covariance matrix and that $f_{\rm NL}$ and $g_{\rm NL}$ are constant scalars and that all quantities are real, the information Hamiltonian is given

[2]Note that this is true in the standard ΛCDM model. However, if there are sources of anisotropic stress before neutrino decoupling, such as, e.g., in the case of primordial magnetic fields, then ζ is no longer a constant on superhorizon scales [66, 67].

by

$$\begin{aligned} H(\zeta_1, d|f_{\rm NL}, g_{\rm NL}) =& -\ln\left[P(d|\zeta_1, f_{\rm NL}, g_{\rm NL})P(\zeta_1|f_{\rm NL}, g_{\rm NL})\right] \\ =& -\ln\left[\mathcal{G}\left(d - R\zeta, N\right)\mathcal{G}\left(\zeta_1, \Xi\right)\right] \\ =& H_0 + \frac{1}{2}\zeta_1^\dagger D^{-1}\zeta_1 - j^\dagger\zeta_1 - \frac{3}{5}f_{\rm NL}j^\dagger\zeta_1^2 - \frac{9}{25}g_{\rm NL}j^\dagger\zeta_1^3 + \frac{3}{5}f_{\rm NL}\zeta_1^\dagger M\zeta_1^2 \\ &+ \frac{9}{25}g_{\rm NL}\zeta_1^\dagger M\zeta_1^3 + \frac{9}{50}f_{\rm NL}^2\left(\zeta_1^\dagger\right)^2 M\zeta_1^2 + \frac{27}{125}f_{\rm NL}g_{\rm NL}\left(\zeta_1^\dagger\right)^2 M\zeta_1^3 \\ &+ \frac{81}{1250}g_{\rm NL}^2\left(\zeta_1^\dagger\right)^3 M\zeta_1^3. \end{aligned} \tag{3.6}$$

Note that some terms of order $\mathcal{O}(\zeta_1^5)$ have already been neglected because we did not state the exact expression for the term proportional to $\mathcal{O}(\zeta_1^4)$ in Eq. (3.2). Eq. (3.6) contains the abbreviations

$$\begin{aligned} D^{-1} = \Xi^{-1} + M, \quad M = R^\dagger N^{-1}R, \quad j = R^\dagger N^{-1}d, \\ \text{and } H_0 = \frac{1}{2}\ln|2\pi\Xi| + \frac{1}{2}\ln|2\pi N| + \frac{1}{2}d^\dagger N^{-1}d. \end{aligned} \tag{3.7}$$

Now, we are able to determine the posterior for the non-Gaussianity parameters $f_{\rm NL}$ and $g_{\rm NL}$, which can be calculated by combining Eqs. (3.5) and (3.6),

$$\begin{aligned} P(f_{\rm NL}, g_{\rm NL}|d) =& \frac{P(d|f_{\rm NL}, g_{\rm NL})P(f_{\rm NL}, g_{\rm NL})}{P(d)} \propto P(f_{\rm NL}, g_{\rm NL})\int \mathcal{D}\zeta_1 P(\zeta_1, d|f_{\rm NL}, g_{\rm NL}) \\ =& P(f_{\rm NL}, g_{\rm NL})\int \mathcal{D}\zeta_1 \exp\left[-H(\zeta_1, d|f_{\rm NL}, g_{\rm NL})\right]. \end{aligned} \tag{3.8}$$

Due to the fact that the Hamiltonian contains higher orders than ζ_1^2 we cannot perform the path-integration analytically. To circumvent this obstacle we conduct a saddle-point approximation in ζ_1 around $\bar{\zeta_1} \equiv \arg\min\left[H(\zeta_1, d|f_{\rm NL}, g_{\rm NL})\right]$ up to the second order in ζ_1 to be still able to perform the path-integration analytically, cf. [53]. This Taylor approximation is justified by $|\zeta_1| \propto \mathcal{O}\left(10^{-5}\right)$. For the expansion of the Hamiltonian we need the first and second derivative with respect to ζ_1, given by

$$\begin{aligned} 0 =& \left.\frac{\delta H(\zeta_1, d|f_{\rm NL}, g_{\rm NL})}{\delta\zeta_1}\right|_{\zeta_1=\bar{\zeta_1}} \\ =& \left(D^{-1} - \frac{6}{5}f_{\rm NL}\,\hat{j}\right)\bar{\zeta_1} - j - \frac{27}{25}g_{\rm NL}\hat{j}\bar{\zeta_1}^2 + \frac{3}{5}f_{\rm NL}\left(M\bar{\zeta_1}^2 + 2\bar{\zeta_1}\star M\bar{\zeta_1}\right) \\ &+ \frac{9}{25}g_{\rm NL}\left(M\bar{\zeta_1}^3 + 3\bar{\zeta_1}^2\star M\bar{\zeta_1}\right) + \left(\frac{18}{25}f_{\rm NL}^2\bar{\zeta_1}\star M\bar{\zeta_1}^2\right) \\ &+ \frac{27}{125}f_{\rm NL}g_{\rm NL}\left(2\bar{\zeta_1}\star M\bar{\zeta_1}^3 + 3\bar{\zeta_1}^2\star M\bar{\zeta_1}^2\right) + \left(\frac{243}{625}g_{\rm NL}^2\bar{\zeta_1}^2\star M\bar{\zeta_1}^3\right), \end{aligned} \tag{3.9}$$

and

$$
\begin{aligned}
D^{-1}_{d,f_{\rm NL},g_{\rm NL}} &\equiv \left.\frac{\delta^2 H(\zeta_1, d|f_{\rm NL}, g_{\rm NL})}{\delta\zeta_1^2}\right|_{\zeta_1=\bar{\zeta_1}} \\
&= D^{-1} - \frac{6}{5} f_{\rm NL}\hat{j} - \frac{54}{25} g_{\rm NL}\widehat{\bar{\zeta_1}\star j} + \frac{6}{5} f_{\rm NL}\left(2\bar{\zeta_1}\star M + \widehat{M\bar{\zeta_1}}\right) \\
&\quad + \frac{27}{25} g_{\rm NL}\left(\widehat{M\bar{\zeta_1}^2} + \left(\bar{\zeta_1}^2\star M + 2\widehat{\bar{\zeta_1}}\widehat{M\bar{\zeta_1}}\right)\right) + \frac{18}{25} f^2_{\rm NL}\left(\widehat{M\bar{\zeta_1}^2} + 2\bar{\zeta_1}\star M\star\bar{\zeta_1}\right) \\
&\quad + \frac{54}{125} f_{\rm NL} g_{\rm NL}\left(3\widehat{\bar{\zeta_1}}\widehat{M\bar{\zeta_1}^2} + \widehat{M\bar{\zeta_1}^3} + 6\bar{\zeta_1}\star M\star\bar{\zeta_1}^2\right) \\
&\quad + \frac{243}{625} g^2_{\rm NL}\left(2\widehat{\bar{\zeta_1}}\widehat{M\bar{\zeta_1}^3} + 3\bar{\zeta_1}^2\star M\star\bar{\zeta_1}^2\right),
\end{aligned}
\tag{3.10}
$$

where $\star$ denotes a pixel-by-pixel multiplication, e.g., $\zeta_x^2 = (\zeta\star\zeta)_x \equiv \zeta_x\zeta_x$ and a hat over fields denotes the transformation of a field to a diagonal matrix, $\zeta_x \mapsto \zeta_x\delta_{xy} \equiv \hat{\zeta}_{xy}$.

With Eqs. (3.9) and (3.10) we are able to perform the saddle-point approximation of the posterior yielding

$$
\begin{aligned}
P(f_{\rm NL}, g_{\rm NL}|d) &\propto P(f_{\rm NL}, g_{\rm NL}) \int \mathcal{D}\zeta_1 \exp\left[-H(d, \zeta_1|f_{\rm NL}, g_{\rm NL})\right] \\
&\approx P(f_{\rm NL}, g_{\rm NL}) \int \mathcal{D}(\zeta_1 - \bar{\zeta_1}) \left|\frac{\delta(\zeta_1 - \bar{\zeta_1})}{\delta\zeta_1}\right|^{-1} \\
&\quad \times \exp\left[-H(d, \bar{\zeta_1}|f_{\rm NL}, g_{\rm NL}) - \frac{1}{2}(\zeta_1 - \bar{\zeta_1})^\dagger D^{-1}_{d,f_{\rm NL},g_{\rm NL}}(\zeta_1 - \bar{\zeta_1})\right] \\
&= |2\pi D_{d,f_{\rm NL},g_{\rm NL}}|^{\frac{1}{2}} \exp\left[-H(d, \bar{\zeta_1}|f_{\rm NL}, g_{\rm NL})\right] P(f_{\rm NL}, g_{\rm NL}).
\end{aligned}
\tag{3.11}
$$

Considering Eq. (3.11), we are able to calculate analytically the full posterior pdf of the $f_{\rm NL}$ and $g_{\rm NL}$ parameter without using expensive Monte Carlo sampling techniques. These techniques have been avoided by replacing the joint pdf for data and curvature perturbation, $P(\zeta_1, d|f_{\rm NL}, g_{\rm NL})$, by the Gaussian distribution $\mathcal{G}(\zeta_1 - \bar{\zeta_1}, D_{d,f_{\rm NL},g_{\rm NL}})$, whose precision is validated for particular inflation models in Sec. 3.4 as well as in Ref. [53] by applying the DIP test [65].

Note that the evaluation of Eq. (3.11) requires a priori knowledge about the primordial power spectrum, Ξ (see Sec. 3.5 for a more detailed description). In the realistic case of small non-Gaussianity one might try, for instance, to see what consequences the power-law power spectrum of the *Planck* cosmology [38] yields, Eq. (3.48), as long as $\Xi = \left\langle \zeta_1\zeta_1^\dagger \right\rangle_{(\zeta_1|\Xi)} \approx \left\langle \zeta\zeta^\dagger \right\rangle_{(\zeta|\Xi)}$ holds. In regimes of larger non-Gaussianity, where the last approximation is violated, the primordial power spectrum and the reconstruction of ζ_1 (wherefore we need a priori Ξ) have to be inferred simultaneously from the data. For this purpose we introduce an Empirical Bayes method in Sec. 3.5.

3.3. Special models of inflation

► There is a large number of different inflationary models, so what particular type should one focus on? Fortunately, recently published papers given by Refs. [38] and [37] address

this question. The first by mainly pointing out parameter constraints to many representative inflation models as well as a Bayesian model comparison thereof, the second by suggesting to concentrate on nine specific types of single-field slow-roll inflation, which are favored by current *Planck* data. To be more precise, the favors of Ref. [37] have been determined by calculating the Bayesian evidence and complexity of the models.

The remaining nine models are all single-field slow-roll models and thus are characterized by, e.g., two slow-roll parameters[3], ϵ and η, which are given by

$$\epsilon \equiv \frac{1}{2}\left(\frac{M_{\mathrm{Pl}}V_\phi}{V}\right)^2 \quad \text{and} \quad \eta \equiv \frac{M_{\mathrm{Pl}}^2 V_{\phi\phi}}{V}, \tag{3.12}$$

where $V \equiv V(\phi)$ denotes the potential of the inflaton ϕ, M_{Pl} the *Planck* mass, subscript letters represent derivatives, and ϵ and η fulfill the bounds $\epsilon, |\eta| \ll 1$. For these specific models the non-Gaussianity parameters are usually much smaller than one and can often be written as a function of ϵ and η, e.g., for single-field slow-roll inflation models with standard kinetic term and Bunch-Davis vacuum as initial vacuum state the parameter f_{NL} is proportional to $\mathcal{O}(\epsilon, \eta)$ (for details see [39, 69, 70]). In particular, the quantitative dependence of the non-Gaussianity parameters on inflationary parameters p can be worked out for every inflation model by conducting cosmological perturbation theory [71, 72] to desired order or by applying the so-called δN-formalism [68, 73–75]. This means, by replacing the non-Gaussianity parameters by (ϵ, η)-dependent functions, which again depend on inflation model specific parameters p as clarified in Eq. (3.12), we are able to infer the slow-roll parameters as well as p directly from data according to Eq. (3.11). Unfortunately, such a tiny amount of non-Gaussianity is currently expected not to be observable due to other general relativistic effects (for details see [39]).

The other case of inflation models with Lagrangians including non-standard kinetic terms leads to non-Gaussianity of equilateral type depending on the so-called sound speed of the inflaton, c_s ($= 1$ for standard kinetic terms) [38]. This type of non-Gaussianity can approximately be described by the parameter $f_{\mathrm{NL}}^{\mathrm{eq}} \propto (1 - c_s^{-2})$. The sound speed of the inflaton, again, depends on the particular inflation model and its parameters. Unfortunately, this type of non-Gaussianity cannot be expressed in a form similar to Eq. (3.1) and one has to go back, for instance, to templates.

There are also multi-field inflation models that are not ruled out yet [38]. In particular models where initially isocurvature perturbations are (not necessarily completely) transformed to adiabatic perturbations. Such a transformation requires at least two fields, e.g., as it happens in the curvaton, axion, higgs inflation scenario, and indeed a slight favor to a non-vanishing amount of isocurvature modes might have been seen by *Planck* [38]. In this section we focus on such scenarios and review the calculations of the non-Gaussianity parameters for a selection of realistic models without claiming that this selection is the most favored one. We are picking only one representative single-parameter mechanism per inflation model for minimal complexity and simplicity. Note, that the approach of Sec. 3.2 would also allow to focus on the other inflation scenarios mentioned above. We choose the following scenarios for illustration only. ◀

[3] Analogously one can use the HFF slow-roll parameter definition as done in Ref. [37].

3.3.1. Simplest curvaton model

► Here, the simplest curvaton model, taking into account radiation and the curvaton, is considered. Without interactions perfect fluids have conserved curvature perturbations [40, 76],

$$\zeta_i = \delta N + \frac{1}{3}\int_{\bar{\rho}_i}^{\rho_i} \frac{d\tilde{\rho}_i}{\tilde{\rho}_i + P_i(\tilde{\rho}_i)}, \quad i \in \{r, \chi\}, \tag{3.13}$$

with r denoting radiation, χ the curvaton, δN the perturbation of the number of e-folds N during inflation, ρ the particle specific density, and P the respective pressure. Barred quantities refer to homogeneous background values. Assuming the curvaton decays on a uniform total density hypersurface determined by $H = \Gamma$, where $H = \frac{\dot{a}}{a}$ is the Hubble parameter, a the cosmological scale factor, and Γ the decay rate of the curvaton (assumed to be constant). Then on this hypersurface

$$\rho_r(t_{\rm decay}, \vec{x}) + \rho_\chi(t_{\rm decay}, \vec{x}) = \bar{\rho}(t_{\rm decay}). \tag{3.14}$$

However, the local curvaton and radiation densities on this decay surface will be inhomogeneous, with $\zeta = \delta N$,

$$\zeta_r = \zeta + \frac{1}{4}\ln\left(\frac{\rho_r}{\bar{\rho}_r}\right) \Rightarrow \rho_r = \bar{\rho}_r\, e^{4(\zeta_r - \zeta)}, \tag{3.15}$$

$$\zeta_\chi = \zeta + \frac{1}{3}\ln\left(\frac{\rho_\chi}{\bar{\rho}_\chi}\right) \Rightarrow \rho_\chi = \bar{\rho}_\chi\, e^{3(\zeta_\chi - \zeta)}. \tag{3.16}$$

Here it was used that once the curvaton starts oscillating it effectively behaves as a non-relativistic perfect fluid, $\rho_\chi \propto a^{-3}$. Using Eqs. (3.15) and (3.16) in Eq. (3.14) leads to [40]

$$\Omega_{\chi,\rm decay}\, e^{3(\zeta_\chi - \zeta)} + (1 - \Omega_{\chi,\rm decay})\, e^{4(\zeta_r - \zeta)} = 1, \tag{3.17}$$

where

$$\Omega_{\chi,\rm decay} \equiv \frac{\bar{\rho}_\chi}{\bar{\rho}_\chi + \bar{\rho}_r}. \tag{3.18}$$

This equation will now be expanded order by order in ζ. Following Ref. [40] the simplest case is considered in which any perturbation in the radiation fluid is neglected, due to, say an inflationary curvature perturbation. Hence $\zeta_r = 0$.

To first order Eq. (3.17) reads [40]

$$4\,(1 - \Omega_{\chi,\rm decay})\,\zeta_1 = 3\,\Omega_{\chi,\rm decay}\,(\zeta_{\chi_1} - \zeta_1)\,, \tag{3.19}$$

where subscript 1 denotes the first order expansion so that

$$\zeta_1 = r\zeta_{\chi_1}, \tag{3.20}$$

where

$$r \equiv \frac{3\Omega_{\chi,\rm decay}}{4 - \Omega_{\chi,\rm decay}} = \left.\frac{3\bar{\rho}_\chi}{3\bar{\rho}_\chi + 4\bar{\rho}_r}\right|_{t_{\rm decay}} \in [0, 1]. \tag{3.21}$$

For $r \approx 1$, i.e. the curvaton is highly dominant, the curvature perturbations are purely adiabatic whereas for $r \ll 1$ not all isocurvature modes have converted to adiabatic ones. Assuming that the curvaton energy density is determined by a simple quadratic potential,

$$\rho_\chi = \frac{1}{2} m^2 \chi^2, \tag{3.22}$$

and assuming it is a weakly coupled field during inflation so that its quantum fluctuations induce a classical Gaussian random field after horizon exit on superhorizon scales, then

$$\chi_* = \bar{\chi}_* + \delta_1\chi_*, \tag{3.23}$$

where the * indicates the time of horizon exit and 1 in the perturbation emphasizes the linear perturbation. Moreover, $\delta_1\chi_*$ is a Gaussian random field with 2-point correlation function in k-space,

$$\langle \delta_1\chi_{*,\vec{k}}\, \delta_1\chi^\dagger_{*,\vec{k}'} \rangle = \frac{2\pi^2}{k^3} \left(\frac{H}{2\pi}\right)^2 \delta_{\vec{k}\vec{k}'}, \tag{3.24}$$

where $H \simeq const.$ during inflation. Now there could be a nonlinear evolution of χ on superhorizon scales after horizon exit up to the beginning of the curvaton oscillations and subsequent decay during the radiation dominated era. In Ref. [40] this is taken into account by introducing a function $g(\chi)$ such that during the curvaton oscillations the value of the curvaton field is given by

$$\chi = g(\chi_*). \tag{3.25}$$

Hence $\bar{\rho}_\chi = \frac{1}{2} m^2 \bar{g}^2$ and [40]

$$\zeta_{\chi_1} = \frac{2}{3}\frac{\delta_1\chi}{\bar{\chi}} = \frac{2}{3}\left.\frac{g'}{g}\right|_{\chi=\chi_*} \delta_1\chi_*. \tag{3.26}$$

In real space the nonlinearity parameters $f_{\rm NL}$ and $g_{\rm NL}$ are defined by (e.g. [40])

$$\zeta(t,\vec{x}) = \zeta_1(t,\vec{x}) + \frac{3}{5} f_{\rm NL} \zeta_1(t,\vec{x})^2 + \frac{9}{25} g_{\rm NL} \zeta_1(t,\vec{x})^3 + \mathcal{O}(\zeta_1^4). \tag{3.27}$$

The Bardeen potential $\Phi_{\rm md}$ on large scales in the matter dominated era is related to ζ_1 by $\Phi_{\rm md} = \frac{3}{5}\zeta_1$ so that

$$\frac{3}{5}\zeta = \Phi_{\rm md} + f_{\rm NL}\Phi_{\rm md}^2 + g_{\rm NL}\Phi_{\rm md}^3. \tag{3.28}$$

In Ref. [40] $f_{\rm NL}$ and $g_{\rm NL}$ are calculated by expanding Eq. (3.19) respectively up to second and third order, e.g.,

$$f_{\rm NL} = \frac{5}{4r}\left(1 + \frac{g g''}{g'^2}\right) - \frac{5}{3} - \frac{5r}{6}, \tag{3.29}$$

and $g_{\rm NL}$ can be found in Ref. [40].

Now considering the simplest model and neglecting any nonlinear evolution of χ between Hubble exit and the start of curvaton oscillations, so that $g'' = 0 = g'''$. In this case we obtain

$$\begin{aligned} f_{\rm NL} =& \frac{5}{4}\kappa - \frac{5}{3} - \frac{5}{6\kappa}, \\ g_{\rm NL} =& \frac{25}{54}\left(-9\kappa + \frac{1}{2} + \frac{10}{\kappa} + \frac{3}{\kappa^2}\right), \end{aligned} \tag{3.30}$$

with

$$\kappa \equiv \frac{1}{r} = \frac{4\bar{\rho}_r}{3\bar{\rho}_\chi} + 1 \in [1, \infty), \tag{3.31}$$

where the parameter κ was introduced for reasons that become clear in Sec. 3.5. Eq. (3.30) is illustrated in Fig. 3.1. Currently, an upper bound on the isocurvature contribution was given by the *Planck* collaboration corresponding to $f_{\rm NL} = -1.23 \pm 0.02$. Note that this constraining interval for $f_{\rm NL}$ was found by a power spectrum fit including adiabatic and isocurvature modes [38] and thus is independent of the limit ($f_{\rm NL} = 2.7 \pm 5.8$) found in Ref. [39] and is only valid for the here considered curvaton scenario. On the other hand this corresponds to the interval $g_{\rm NL} = 1.97 \pm 0.11$. A $g_{\rm NL}$ outside this interval would put some pressure on this simplest curvaton model. ◀

3.3.2. Modulated Higgs inflation

▶ Next, we consider the Standard Model Higgs field h in addition[4] to the inflaton field ϕ with related potential $V(\phi)$ as pointed out in Ref. [77] and as representative mechanism of the Higgs inflation (HI) class. The Higgs field is responsible for modulating the efficiency of reheating, whereby primordial curvature perturbations are generated by converting isocurvature perturbations (produced by h during inflation) to adiabatic ones [79]. In particular we assume a simple Higgs potential during the energy scale of inflation μ,

$$V(h) = \frac{\lambda}{4}h^4, \tag{3.32}$$

with $\lambda \equiv \lambda(\mu) \approx \mathcal{O}(10^{-2})$ the Higgs self coupling with logarithmic dependence on the energy scale.

Within this model we can write the total decay rate of the inflaton, $\Gamma(h)$, as a sum of a Higgs dependent and independent term,

$$\Gamma(h) = \Gamma^I + \Gamma^D(h). \tag{3.33}$$

Then the curvature perturbation is given by [77, 80]

$$\zeta = \frac{1}{M_{\rm Pl}^2}\frac{V(\phi)}{V_\phi(\phi)}\delta\phi_* + Q_h\delta h_* + \frac{1}{2}Q_{hh}\delta h_*^2 + \frac{1}{6}Q_{hhh}\delta h_*^3 + \mathcal{O}(\delta h_*^4), \tag{3.34}$$

with $Q \approx a_0 \log\left(\frac{\Gamma}{H_c}\right)$. $*$ denotes the horizon exit, H_c is the Hubble constant at t_c (a time before the decay of the inflaton, for details cf. [77]), and subscript letters represent derivatives. a_0 is a model dependent constant of the order of $\mathcal{O}(10^{-1})$.

[4]It is well known that the Standard Model Higgs field cannot serve as an inflaton field [77, 78].

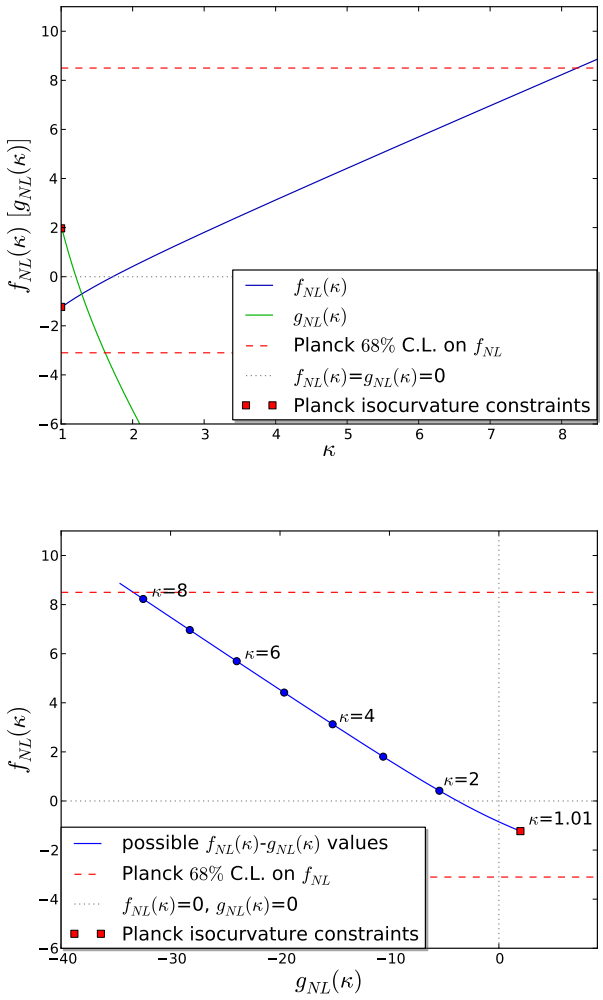


Figure 3.1.: Simplest curvaton model: Possible values of $f_{\rm NL}$ and $g_{\rm NL}$ within current *Planck* constraints parametrized by the curvaton parameter $\kappa = 4\bar{\rho}_r/(3\bar{\rho}_\chi) + 1$. The current constraints on isocurvature modes (red squares) narrow down the allowed region significantly.

Assuming the Higgs dependent decay rate to be of polynomial form in h, $\Gamma^D(h) \propto h^n$, the non-Gaussianity parameters can be calculated from the statistics of ζ, which yields [77]

$$\begin{aligned} f_{\rm NL} &= -\frac{5}{6}\frac{\beta^2}{a_0}\left(1 - \frac{\Gamma\Gamma_{hh}}{\Gamma_h^2}\right) \approx -\frac{5}{6}\frac{\beta^2}{a_0}\left(1 - \frac{1}{B_h}\frac{n-1}{n}\right), \\ g_{\rm NL} &= -\frac{25\beta^3}{54a_0^2}\left(2 - 3\frac{\Gamma\Gamma_{hh}}{\Gamma_h^2} + \frac{\Gamma^2\Gamma_{hhh}}{\Gamma_h^3}\right) \approx \frac{2(n-2)}{3(n-1)\beta} f_{\rm NL}^2 - \frac{5}{3}\frac{\beta}{a_0} f_{\rm NL}, \end{aligned} \tag{3.35}$$

with $B_h = \Gamma^D/\Gamma \leq \mathcal{O}(10^{-3} - 10^{-2})$ and $\beta \approx \mathcal{O}(10^{-2} - 1)$. Eq. (3.35) is illustrated in Fig. 3.2. ◀

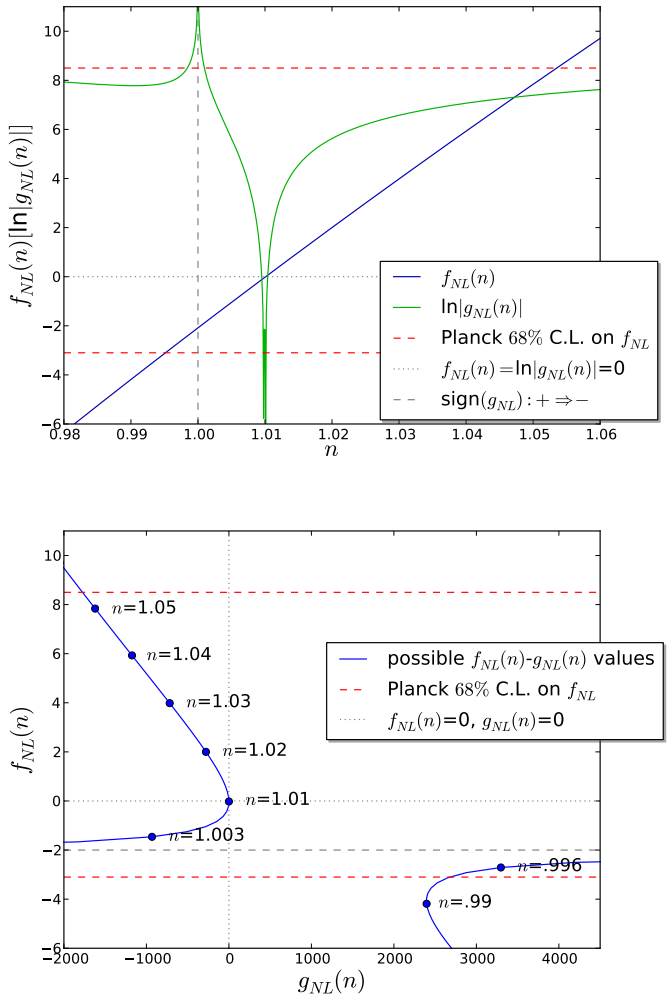


Figure 3.2.: Modulated Higgs inflation: Possible values of $f_{\rm NL}$ and $g_{\rm NL}$ for some model-typical values of $a_0 = 0.1$, $\beta = 0.5$, and $B_h = 0.01$ within the *Planck* constraints parametrized by the decay rate index n from $\Gamma^D(h) \propto h^n$.

3.4. Posterior for special inflationary parameters

3.4.1. Generic procedure

For all models discussed in Sec. 3.3 an expression for the posterior of a model specific quantity can be derived by replacing $f_{\rm NL}$ and $g_{\rm NL}$ by their corresponding model dependent parameters p, which are pointed out in the stated section. This is, of course, also true for all other inflation models postulating these two non-Gaussianity parameters. Thereby one obtains p-dependent equations for $\left.\frac{\delta H(\zeta_1,d|p)}{\delta\zeta_1}\right|_{\zeta_1=\bar{\zeta_1}} = 0$ and $D^{-1}_{d,p} \equiv \left.\frac{\delta^2 H(\zeta_1,d|p)}{\delta\zeta_1^2}\right|_{\zeta_1=\bar{\zeta_1}}$, which allow to derive the posterior analytically, Eq. (3.11).

Eventually, the response R has to be replaced by its respective, corresponding expression, depending on whether one uses the CMB data, LSS data, or something else. The resulting

posterior can then be implemented and evaluated numerically. With its help one might obtain a single point (e.g. mean of p or maximum of the posterior) within the $f_{\rm NL}(p) - g_{\rm NL}(p)-$plane with corresponding error interval being a sub-area of the plane or it just maps out the parameter posterior.

3.4.2. Simplest curvaton model

3.4.2.1. Posterior derivation

To demonstrate the applicability of the inference approach we study the simplest curvaton inflation scenario. We follow Secs. 3.2 and 3.4.1 to derive the posterior. The replacement of $f_{\rm NL}$ and $g_{\rm NL}$ by their corresponding κ expressions leads to the information Hamiltonian,

$$\begin{aligned} H(\zeta_1, d|\kappa) = H_0 &+ \frac{1}{2}\zeta_1^\dagger D^{-1}\zeta_1 - j^\dagger\zeta_1 \\ &- \left(\frac{3}{4}\kappa - 1 - \frac{1}{2\kappa}\right) j^\dagger\zeta_1^2 - \left(-\frac{3}{2}\kappa + \frac{1}{12} + \frac{5}{3\kappa} + \frac{1}{2\kappa^2}\right) j^\dagger\zeta_1^3 \\ &+ \left(\frac{3}{4}\kappa - 1 - \frac{1}{2\kappa}\right) \zeta_1^\dagger M\zeta_1^2 + \left(-\frac{3}{2}\kappa + \frac{1}{12} + \frac{5}{3\kappa} + \frac{1}{2\kappa^2}\right) \zeta_1^\dagger M\zeta_1^3 \\ &+ \frac{1}{2}\left(\frac{3}{4}\kappa - 1 - \frac{1}{2\kappa}\right)^2 \left(\zeta_1^2\right)^\dagger M\zeta_1^2 + \left(-\frac{3}{2}\kappa + \frac{1}{12} + \frac{5}{3\kappa} + \frac{1}{2\kappa^2}\right) \\ &\times \left(\frac{3}{4}\kappa - 1 - \frac{1}{2\kappa}\right) \left(\zeta_1^2\right)^\dagger M\zeta_1^3 + \frac{1}{2}\left(-\frac{3}{2}\kappa + \frac{1}{12} + \frac{5}{3\kappa} + \frac{1}{2\kappa^2}\right)^2 \left(\zeta_1^3\right)^\dagger M\zeta_1^3. \end{aligned} \tag{3.36}$$

Analogously, by replacing $f_{\rm NL}$ and $g_{\rm NL}$ by their corresponding κ expressions in Eqs. (3.9) and (3.10), one obtains expressions for $\left.\frac{\delta H(\zeta_1,d|\kappa)}{\delta\zeta_1}\right|_{\zeta_1=\bar{\zeta_1}} = 0$ and $D^{-1}_{d,\kappa} \equiv \left.\frac{\delta^2 H(\zeta_1,d|\kappa)}{\delta\zeta_1^2}\right|_{\zeta_1=\bar{\zeta_1}}$, whereby we are able to perform the saddle-point approximation of the posterior, which yields (see Eq. (3.11))

$$P(\kappa|d) \propto P(\kappa) \int \mathcal{D}\zeta_1 \exp\left[-H(d,\zeta_1|\kappa)\right] \approx |2\pi D_{d,\kappa}|^{\frac{1}{2}} \exp\left[-H(d,\bar{\zeta_1}|\kappa)\right] P(\kappa). \tag{3.37}$$

For numerical reasons $D^{-1}_{d,\kappa}$ is split into a diagonal part, $D^{-1}_{d,\kappa,\rm diag}$, and a non-diagonal one, $D^{-1}_{d,\kappa,\text{non-diag}}$, which leads to (cf. [53])

$$\begin{aligned} \ln\left[P(\kappa|d)\right] = &- H(\kappa|d) \\ \approx &- \frac{1}{2}\mathrm{tr}\left[\ln\left(\frac{1}{2\pi}D^{-1}_{d,\kappa,\rm diag}\right)\right] + \frac{1}{2}\mathrm{tr}\left[\sum_{n=1}^{\infty}\frac{(-1)^n}{n}\left(D_{d,\kappa,\rm diag}D^{-1}_{d,\kappa,\text{non-diag}}\right)^n\right] \\ &- H(d,\bar{\zeta_1}|\kappa) + \ln\left[P(\kappa)\right] + \text{const.}. \end{aligned} \tag{3.38}$$

The series expansion of the logarithm in Eq. (3.38) can be truncated if the terms become sufficiently small.

3.4.2.2. Numerical implementation

For a numerical implementation of Eq. (3.38) we have to choose a prior pdf for κ, $P(\kappa)$. A naive choice, for instance, according to Eq. (3.31) would be

$$P(\kappa) = \frac{1}{\kappa_0 - 1}\Theta(\kappa - 1)\Theta(\kappa_0 - \kappa), \tag{3.39}$$

with Θ the Heaviside step function and κ_0 a large but finite number to normalize the distribution. However, at this point we do not want to open a discussion of how to choose an appropriate prior pdf for the curvaton parameter. Thus we focus on the likelihood pdf, $P(d|\kappa) \propto \frac{P(\kappa|d)}{P(\kappa)}$, given by subtracting $\ln[P(\kappa)]$ from Eq. (3.38).

In order to implement the likelihood we study the inference from CMB data in the Sachs-Wolfe limit [81] within a toy example in two and three dimensions. Here the data are given by

$$\begin{aligned} d &= R\zeta + n = R_{\mathrm{CMB}}\Phi_{\mathrm{md}} + n = R_{\mathrm{CMB}}\frac{3}{5}\zeta + n \\ &= R_{\mathrm{CMB}}\frac{3}{5}\left(\zeta_1 + \frac{3}{5}f_{\mathrm{NL}}\zeta_1^2 + \frac{9}{25}g_{\mathrm{NL}}\zeta_1^3 + \mathcal{O}(\zeta_1^4)\right) + n, \end{aligned} \tag{3.40}$$

with Φ_{md} the Bardeen potential in the matter dominated era. In this limit the response becomes local[5] [82],

$$R(x, y) = -\frac{3}{5}\frac{1}{3}\delta(x - y), \tag{3.41}$$

with x, y two positions on the two-(three-) dimensional sky. Additionally, we assume white noise, $N_{xy} = \sigma_n^2\delta_{xy}$. For the inference process the so-called free information propagator D is required, which depends in particular on the power spectrum of ζ_1. For this we assume

$$\Xi = \left\langle \zeta_1\zeta_1^\dagger \right\rangle_{(\zeta_1|\Xi)} \equiv P_{\zeta_1}(k)\delta_{kk'} = A_s\left(\frac{k}{k_*}\right)^{n_s - 1}\delta_{kk'}, \tag{3.42}$$

which is diagonal in Fourier space with related modes k, k'. This power spectrum is parametrized by the scalar amplitude A_s, the spectral index n_s, and the pivot scale k_*. A detailed discussion about this power spectrum and its parameters can be found in Sec. 3.5 and Ref. [38].

The numerical implementation is done in NIFTy [30], where possible calculations of traces of operators are determined by operator probing. The NIFTy package uses implicit operators and therefore avoids to store matrices explicitely. Thus, highly resolved data sets like the CMB map of the *Planck* satellite should be treatable in principle. However, to show the efficiency of the derived inference method we use a data set in two (three) flat dimensions with best fit parameters (for scalar amplitude and spectral index from *Planck* [38]) $A_s = 2.2 \times 10^{-9}$, $n_s = 0.9603$, $k_* = 1$, $N_{\mathrm{pix}} = 10000$ (10648), and $\sigma_n^2 = 10^{-14}$, where ζ_1 and n are sampled from Ξ and N, respectively.

An implementation of the likelihood (posterior[6] with constant prior) in two (three) dimensions with a true underlying value of $\kappa_{\mathrm{gen}} = 5$ ($\kappa_{\mathrm{gen}} = 19.8$) is shown in Fig. 3.3. The numerical result coincides [perfectly, (a)] with a Gaussian fit and would deviate from this shape only for unrealistically high values of κ. A slight deviation from Gaussianity, however,

[5] The treatment of non-local responses was shown in [53] for a similar case.
[6] For a prior choice according to Eq. (3.39), Fig. 3.3 shows the posterior of κ.

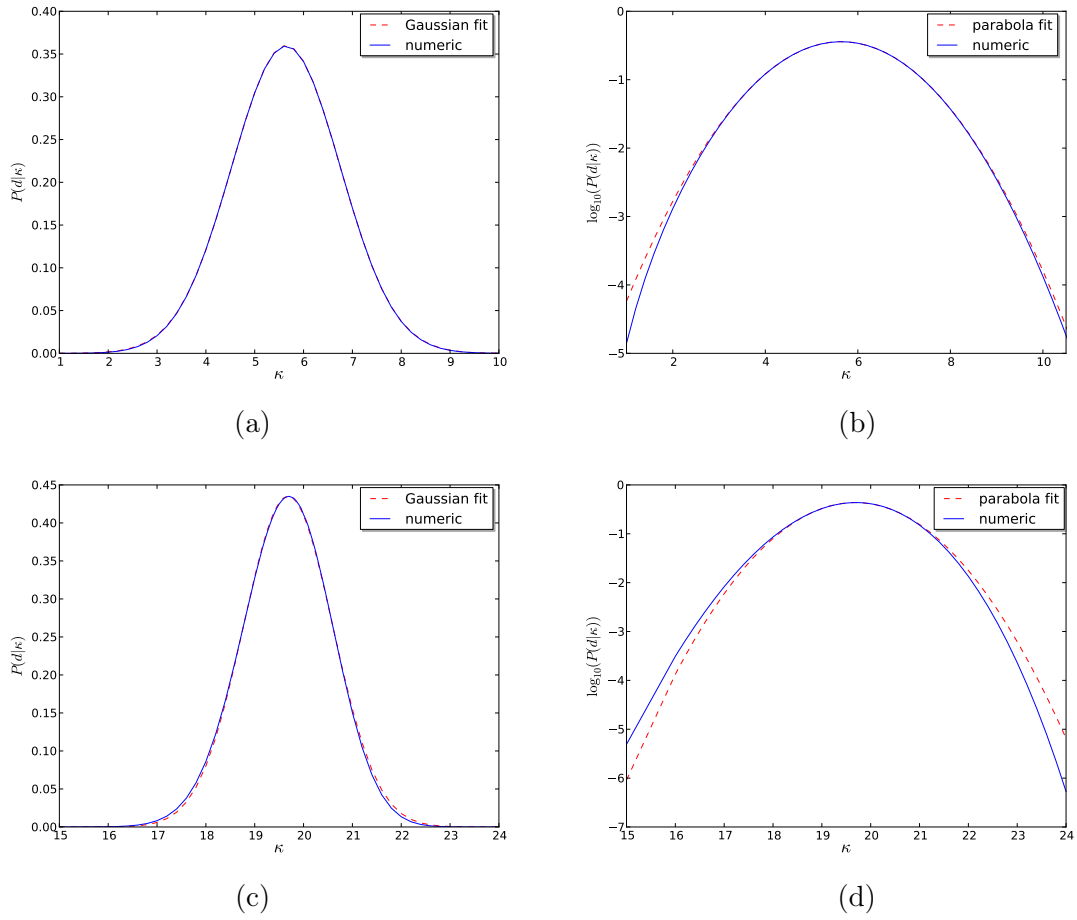


Figure 3.3.: Normalized likelihood distributions for κ in a two-(three-)dimensional test case with data generated from $\kappa_{\mathrm{gen}} = 5$ [(a), (b)] ($\kappa_{\mathrm{gen}} = 19.8$ [(c), (d)]).

can even be observed in the three-dimensional realization with $\kappa = 19.8$ (likelihood is negatively skewed) [21, 47–52]. As we will show in Sec. 3.5.3, even this slight deviation (and even smaller ones) can affect the reconstruction of the primordial power spectrum, $P_\zeta(k)$, significantly. Note, however, that this statement is only true for the likelihood of κ and has to be taken into account case by case. Significant non-Gaussianity in the likelihood and posterior pdf might be induced by the p dependent determinant of Eq. (3.11) or by a specific prior choice. A brief discussion about the comparison of a skewed posterior of p can be found in App. B.1.

3.4.2.3. Posterior validation

To validate the implementation of Eq. (3.38) we consider the two-dimensional test case of Sec. 3.4.2.2 and apply the DIP test [65] by following Refs. [53, 83], i.e. conducting the following steps:

1. Sample uniformly a value of κ_{gen} from an interval $I = [\kappa_{\mathrm{ini}}, \kappa_{\mathrm{fin}}]$, i.e. from a prior[7]

[7]Note that for validating both, the sufficiency of numerical implementation of the posterior distribution

$$P(\kappa) = \begin{cases} \frac{1}{\kappa_{\rm fin} - \kappa_{\rm ini}} & \text{if } \kappa \in I \\ 0 & \text{else} \end{cases} . \tag{3.43}$$

2. Generate data d for $\kappa_{\rm gen}$ according to Eq. (3.2).

3. Calculate a posterior curve for given data by determining $P(\kappa|d)$ for $\kappa \in I$ according to Eq. (3.38).

4. Calculate the posterior probability for $\kappa \leq \kappa_{\rm gen}$ according to

$$x \equiv \int_{\kappa_{\rm ini}}^{\kappa_{\rm gen}} {\rm d}\kappa \; P(\kappa|d) \;\; \in [0,1] \, . \tag{3.44}$$

5. If the calculation of the posterior was correct, the distribution for x, $P(x)$, should be uniform between 0 and 1.

The result of this posterior validation test is shown by Fig. 3.4. Here, the histogram represents the distribution of 500 x-values within eight bins. The uniformity of the distribution verifies the numeric and analytic (due to the saddle-point approximation) sufficiency of the posterior. In particular, this means that the shape of the posterior (and therefore the error-bars around the posterior mean) are calculated correctly. Otherwise the histogram of the DIP test would have exhibited a characteristic deviation from uniformity, e.g., a dip in the case of an underestimation of the variance.

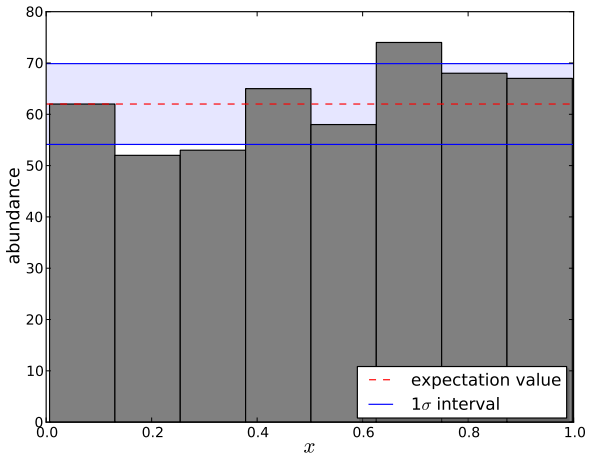


Figure 3.4.: DIP distribution of calculated x values for the two-dimensional test case. The histogram shows the unnormalized distribution of 500 x values within eight bins. The standard deviation interval (1σ, blue solid line) around the expectation value (red dashed line) as calculated from Poissonian statistics is also shown.

and analytic approximations including its derivation, it is not necessary to choose a physical prior. Thus this kind of prior with appropriate values $\kappa_{\rm ini,\ fin}$ has been chosen for simplicity only. Here, appropriate means that the interval I is sufficiently large to take care of the shape of the posterior.

3.4.3. Modulated Higgs inflation

Next, we consider the scenario of modulated Higgs inflation discussed in Sec. 3.3.2. By following again Secs. 3.2 and 3.4.1, i.e. by replacing $f_{\rm NL}$ and $g_{\rm NL}$ by their corresponding n dependent expressions[8], we obtain

$$\begin{aligned} H(\zeta_1, d|f_{\rm NL}(n)) = & -\ln\left[P(d|\zeta_1, f_{\rm NL}(n))P(\zeta_1|f_{\rm NL}(n))\right] = -\ln\left[\mathcal{G}\left(d - R\zeta, N\right)\mathcal{G}(\zeta_1, \Xi)\right] \\ = & H_0 + \frac{1}{2}\zeta_1^\dagger D^{-1}\zeta_1 - j^\dagger\zeta_1 - \frac{3}{5}f_{\rm NL}j^\dagger\zeta_1^2 \\ & - \left(\frac{6(n-2)}{25(n-1)\beta}f_{\rm NL}^2 - \frac{3}{5}\frac{\beta}{a_0}f_{\rm NL}\right)j^\dagger\zeta_1^3 + \frac{3}{5}f_{\rm NL}\zeta_1^\dagger M\zeta_1^2 \\ & + \left(\frac{6(n-2)}{25(n-1)\beta}f_{\rm NL}^2 - \frac{3}{5}\frac{\beta}{a_0}f_{\rm NL}\right)\zeta_1^\dagger M\zeta_1^3 + \frac{9}{50}f_{\rm NL}^2\left(\zeta_1^\dagger\right)^2 M\zeta_1^2 \\ & + \left(\frac{18(n-2)}{125(n-1)\beta}f_{\rm NL}^3 - \frac{81}{625}\frac{\beta}{a_0}f_{\rm NL}^2\right)\left(\zeta_1^\dagger\right)^2 M\zeta_1^3 \\ & + \left(\frac{27(n-2)}{625(n-1)\beta}f_{\rm NL}^2 - \frac{27}{250}\frac{\beta}{a_0}f_{\rm NL}\right)^2\left(\zeta_1^\dagger\right)^3 M\zeta_1^3, \end{aligned} \tag{3.45}$$

with $f_{\rm NL}(n)$ given by

$$f_{\rm NL}(n) \approx -\frac{5}{6}\frac{\beta^2}{a_0}\left(1 - \frac{1}{B_h}\frac{n-1}{n}\right). \tag{3.46}$$

In Eq. (3.45) we dropped the explicit dependency of $f_{\rm NL}$ on n in our notation for reasons of clarity. To derive the posterior pdf for the n parameter of the Higgs inflation model, we again conduct the saddle-point approximation introduced in Sec. 3.2. This yields

$$\begin{aligned} \ln[P(n|d)] = & -H(n|d) \\ \approx & -\frac{1}{2}\mathrm{tr}\left[\ln\left(\frac{1}{2\pi}D_{d,n,\rm diag}^{-1}\right)\right] + \frac{1}{2}\mathrm{tr}\left[\sum_{m=1}^{\infty}\frac{(-1)^m}{m}\left(D_{d,n,\rm diag}D_{d,n,\text{non-diag}}^{-1}\right)^m\right] \\ & - H(d, \bar{\zeta_1}|n) + \ln[P(n)] + \text{const.}, \end{aligned} \tag{3.47}$$

where $\bar{\zeta_1}$ and $D_{d,n}$ are defined in Eqs. (3.9) and (3.10) and the labels *diag* and *non-diag* refer to the diagonal and non-diagonal part of $D_{d,n}^{-1}$. As before, the series expansion of the logarithm can be truncated if the terms become sufficiently small.

The numerical implementation of Eq. (3.47) and its validation is completely analogous to the one of the curvaton scenario. Therefore we do not present it here.

[8] As a reminder: the model is parametrized by the decay rate index n, $\Gamma^D(h) \propto h^n$ (see Sec. 3.3.2 for details).

3.5. Primordial power spectrum reconstruction

3.5.1. Motivation

► An essential quantity that allows to discriminate between inflationary scenarios is the primordial power spectrum. It is commonly parametrized by [38]

$$\begin{aligned} \ln\left[P_{\mathcal{R}}(k)\right] &= \ln(A_s) + \left(n_s - 1 + \frac{1}{2}\frac{dn_s}{d(\ln k)}\ln\left(\frac{k}{k_*}\right) + \dots\right)\ln\left(\frac{k}{k_*}\right), \\ \ln\left[P_t(k)\right] &= \ln(A_t) + \left(n_t + \frac{1}{2}\frac{dn_t}{d(\ln k)}\ln\left(\frac{k}{k_*}\right) + \dots\right)\ln\left(\frac{k}{k_*}\right), \end{aligned} \tag{3.48}$$

where A_s (A_t) denotes the scalar (tensor) amplitude, n_s (n_t) the scalar (tensor) spectral index, k_* the mode k crossing the Hubble radius, and $\mathcal{R}$ the comoving curvature perturbation, which is approximately equal to the comoving curvature perturbation on uniform density hypersurfaces, ζ, on large scales. Henceforth we will focus on the scalar part of the power spectrum. For considering the pure power-law form of Eq. (3.48), $P_{\mathcal{R}}(k) = A_s\,(k/k_*)^{n_s-1}$, the *Planck* collaboration [38] recently found the best fit values $A_s = 2.2 \times 10^{-9}$ and $n_s = 0.9603$ (± 0.0073) for $k_* = 0.05$ Mpc^{-1}, which constrain all inflationary scenarios. However, also an extension to this simple power-law shape is currently investigated taking into account bumps, sharp features, or wiggles. These types of deviations are well motivated by, e.g., implications of the recent BICEP2 data [41–44], or special features of the inflaton potential [45, 46]. Such features, in turn, might indicate non-linear physics and thus correspond to non-vanishing non-Gaussianity parameters [38]. These features would therefore be an indicator for inflation models beyond single-field slow-roll scenarios. ◄

The primordial power spectrum is a valuable quantity since it depends on the physics of the early Universe. Its inference process is highly non-trivial. Therefore we would like to present two Bayesian, non-parametric reconstruction schemes in the framework of information field theory, which, however, are related to each other. To reconstruct the primordial power spectrum we have to know the inferred field ζ and its variance. This means we are interested in the posterior pdf of ζ. Following Sec. 3.2 the Hamiltonian is given by

$$H(d,\zeta) = -\ln\left[\mathcal{G}(d - R\zeta, N)P(\zeta)\right]. \tag{3.49}$$

In general, Eq. (3.49) cannot be evaluated further because the shape of $P(\zeta)$ is unknown. Fortunately, a Gaussian is a very good approximation for $P(\zeta)$ as argued in Sec. 3.2.2 and motivated by the actual constraints on $f_{\rm NL}$ [39] that becomes exact if $f_{\rm NL} = g_{\rm NL} = 0$. This enormously simplifies the derivation and under this approximation the posterior is given by

$$P(\zeta|d) = \mathcal{G}(\zeta - m_{\rm w}, D), \tag{3.50}$$

with $m_{\rm w} \equiv Dj = \left(\Xi^{-1} + M\right)^{-1} R^\dagger N^{-1} d$ the Wiener filter solution.

One may also be interested in the case[9] where $|f_{\rm NL}| \propto \mathcal{O}(1)$ and $g_{\rm NL} \neq 0$. Here, the quantity of interest is the power spectrum of the primordial Gaussian perturbation ζ_1, $P_{\zeta_1}(k)$. The approach of reconstructing ζ_1 for fixed non-Gaussianity parameters is already

[9] We declare the case of $|f_{\rm NL}| \propto \mathcal{O}(1)$ to be interesting due to the current constraints on $f_{\rm NL}$. The discussion that follows, however, is generic and therefore valid for arbitrary values of $f_{\rm NL}$ and $g_{\rm NL}$ still satisfying the saddle-point approximation, Eq. (3.11).

described in Sec. 3.2.2 and determined by Eqs. (3.9) and (3.10). Therefore the posterior is given by $\mathcal{G}(\zeta_1 - \bar{\zeta_1}, D_{d,f_{\rm NL},g_{\rm NL}})$, which implicitly depends on parameters of non-Gaussianity or alternatively on parameters of inflation.

Given a reconstructed map of ζ or ζ_1 and its uncertainty D or $D_{d,f_{\rm NL},g_{\rm NL}}$, the challenge is now to appropriately infer the power spectrum $P_\zeta(k)$ or $P_{\zeta_1}(k)$ under consideration of the uncertainty. For this purpose we suggest the two following approaches (Eqs. (3.55) and (3.58)), which have already successfull applications in cosmology and astrophysics, e.g., Refs. [84–87]. We will show that these methods are able to reconstruct the spectrum of the primordial curvature perturbations even in case of significant non-Gaussianity and partial sky coverage.

3.5.2. Filter formulae

Critical filter. The first filter captures the concepts of the well known Karhunen-Loève [88, 89] and Feldman-Kaiser-Peacock [90] estimators and has been derived in Ref. [25]. The aim here is to reconstruct the power spectrum for Gaussian signals, which determines the statistics completely under the cosmological assumption of translationally and rotationally invariance. This implies the existence of an orthonormal basis O in which Ξ become diagonal, e.g., the Fourier space with elements $\vec{k} = (k_1, \ldots, k_3) \in \mathbb{R}^3$ of length $k \equiv |\vec{k}|$ (Fourier mode) for signals defined in Euclidean space, or the spherical harmonics space for signals defined on the sphere. Following Ref. [25], the signal covariance, here $\Xi_{kk'}$, and its inverse are linearly parametrized by non-overlapping basis functions $f_i(k)$, commonly denoted as spectral bands, and coefficients $\tilde{p}$,

$$P_\zeta(k) = \sum_i \tilde{p}_i f_i(k), \tag{3.51}$$

where $(\Xi_i)_{xy} = O^*_{xk} f_i(k) O_{ky}$, and therefore

$$\Xi_{\tilde{p}} = \sum_i \tilde{p}_i \; \Xi_i \quad \text{and} \quad \Xi^{-1}_{\tilde{p}} = \sum_i \tilde{p}_i^{-1} \; \Xi_i^{-1}. \tag{3.52}$$

Ξ_i^{-1} denotes the pseudo-inverse of the band-variances, given by $\left(\Xi_i^{-1}\right)_{xy} = O^*_{xk} g_i(k) O_{ky}$ with $g_i(k) = 1/f_i(k)$ if $f_i(k) > 0$ and $g_i(k) = 0$ if $f_i(k) = 0$ [25]. In all cases addressed in this paper the non-overlapping basis functions $f_i(k)$ are projections from the vectors $\vec{k}$ onto the Fourier modes $k = |\vec{k}| = \sqrt{k_1^2 + k_2^2 + k_3^2}$. This way, the primordial power spectrum can be parametrized as in Eq. (3.42).

The priors of $\tilde{p}$ are assumed to be mutually independent, $P(\tilde{p}) = \prod_i P(\tilde{p}_i)$, and obey an inverse Gamma distribution,

$$P(\tilde{p}_i) = \mathcal{I}(\tilde{p}_i; \alpha_i, q_i) \equiv \frac{1}{q_i \Gamma(\alpha_i - 1)} \left(\frac{\tilde{p}_i}{\alpha_i}\right)^{-\alpha_i} \exp\left(-\frac{q_i}{\tilde{p}_i}\right), \tag{3.53}$$

with Γ the Gamma function. Constructed in this way one obtains an informative prior by $\alpha_i \gg 1$ and a non-informative prior, e.g. Jeffreys prior, by $\alpha_i = 1, q_i = 0$.

To derive the critical filter formula we calculate the minimum of the ζ marginalized Hamiltonian,

$$H(d, \tilde{p}) = \frac{1}{2}\mathrm{tr}\left(\ln \Xi\right) - \frac{1}{2}\mathrm{tr}\left(\ln D\right) - \frac{1}{2} j^\dagger D j + \sum_i (\alpha_i - 1)\tau_i + q_i e^{-\tau_i}, \tag{3.54}$$

with respect to $\tau \equiv \ln(\tilde{p})$, thereby maximizing the posterior probability for the logarithmic power spectrum yielding the coupled system of equations [25]

$$\begin{aligned} m_{\tilde{p}_{\min}} &= D_{\tilde{p}_{\min}} j, \\ \tilde{p}_{i,\min} &= \frac{q_i + \frac{1}{2}\mathrm{tr}\left(m_{\tilde{p}_{\min}} m^\dagger_{\tilde{p}_{\min}} + D_{\tilde{p}_{\min}}\right)\Xi_i^{-1}}{\alpha_i - 1 + \frac{1}{2}\mathrm{tr}\left(\Xi_i^{-1}\Xi_i\right)}. \end{aligned} \tag{3.55}$$

For the parameter choice according to Jefferys prior, $\alpha_i = 1, q_i = 0$, Eq. (3.55) is called critical filter. To solve this coupled system iteratively, we need a boundary condition, e.g., for the power spectrum (remember that $D_{\tilde{p}_{\min}}$ depends on the spectral coefficients $\tilde{p}_{i,\min}$). A well motivated initial guess might be the primordial power spectrum from Planck [38], which is a pure power law, Eq. (3.48).

Critical filter with smoothness prior. For some physical reasons [38], e.g., that physics do not change suddenly during inflation, one may want to enforce the reconstructed power spectrum to be smooth. This can be incorporated by an extension of the prior [63], given by

$$P(\tau) = P_{\mathrm{sm}}(\tau) \prod_i P(\tau_i), \tag{3.56}$$

with the smoothness prior

$$P_{\mathrm{sm}}(\tau) \propto \exp\left(-\frac{1}{2\sigma_\tau^2}\int d(\ln k)\left(\frac{\partial^2 \ln \tilde{p}(\tau_k)}{\partial(\ln k)^2}\right)^2\right) \equiv \exp\left(-\frac{1}{2}\tau^\dagger T \tau\right), \tag{3.57}$$

which punishes any deviation from a power-law power spectrum with a strength σ_τ. This means in the limit of $\sigma_\tau \to \infty$ we recover Eq. (3.55) whereas for a finite decreasing, especially small value of σ_τ the smoothness increases. Here we introduced the linear operator T whose explicit form can be found in Ref. [63]. T includes the integral as well as the scaling constant σ_τ. An analogous derivation to the critical filter case then yields

$$\tilde{p}_{i,\min} = \frac{q_i + \frac{1}{2}\mathrm{tr}\left(m_{\tilde{p}_{\min}} m^\dagger_{\tilde{p}_{\min}} + D_{\tilde{p}_{\min}}\right)\Xi_i^{-1}}{\alpha_i - 1 + \frac{1}{2}\mathrm{tr}\left(\Xi_i^{-1}\Xi_i\right) + (T\tau)_i}. \tag{3.58}$$

In comparison to Eq. (3.55) the result exhibits the additional term $(T\tau)_i$ in the denominator that enforces smoothness. By appropriately choosing the scale parameter σ_τ one is able to permit the reconstruction of features on specific scales. However, for a detailed discussion of the critical filter with smoothness prior including the choice of σ_τ see Ref. [63].

3.5.3. Numerical toy example

3.5.3.1. Inferring a power spectrum of approximately Gaussian curvature perturbations

To demonstrate the performance of the filter formulae, Eqs. (3.55) and (3.58), according to the inference of a power spectrum of approximately Gaussian curvature perturbations ζ we use the two dimensional test case of Sec. 3.4.2.2, but with $N_{\mathrm{pix}} = 10^6$, and $f_{\mathrm{NL}}, g_{\mathrm{NL}} \approx 0$ to satisfy the condition of negligible non-Gaussianity. Additionally, we adopt the parameter choice according to Jefferys prior, i.e. $\alpha_i = 1, q_i = 0$, and a scaling constant of $\sigma_\tau^2 = 0.1$.

Figure 3.5 (a) shows the result and confirms the properness of the reconstruction, which exhibits typical deviations from the true underlying spectrum for low k modes due to the effect of cosmic variance. Reconstruction errors are not included in the figure but could be incorporated by evaluating the inverse Hessian of the Hamiltonian $H(d, \tilde{p})$ for real scenarios, as done in Ref. [63].

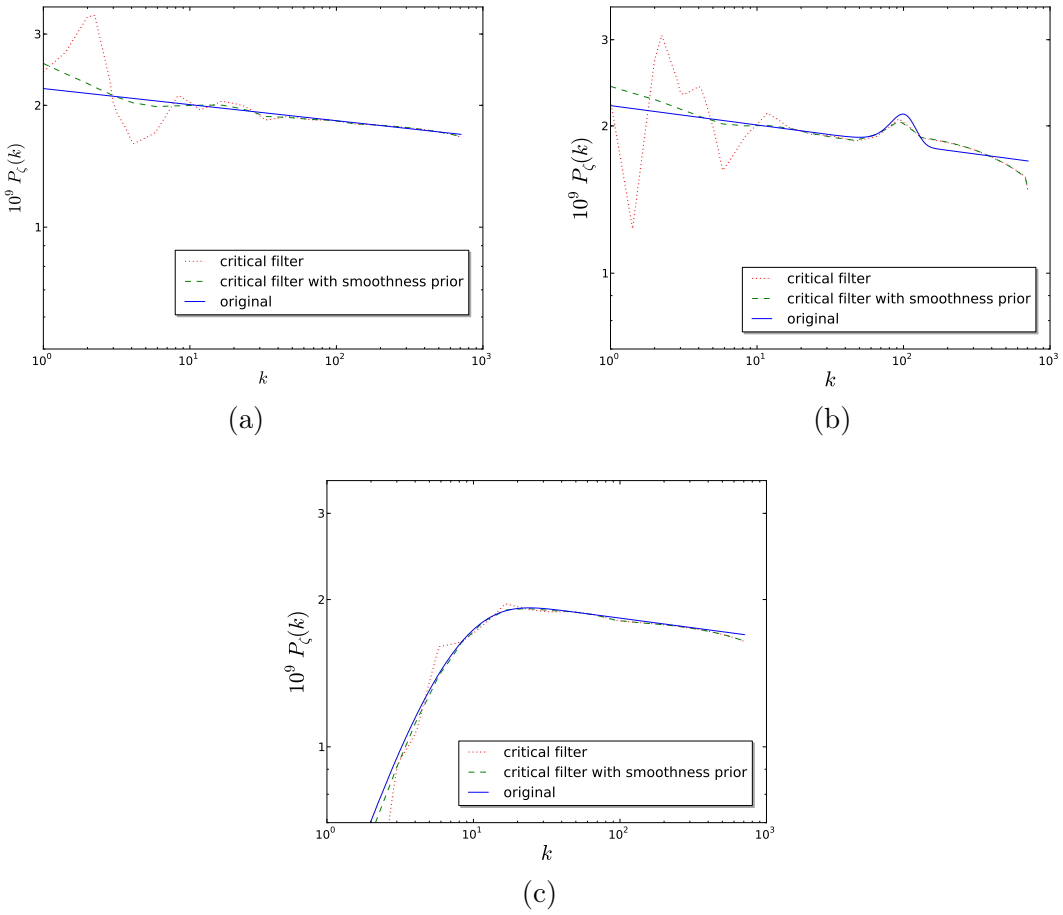
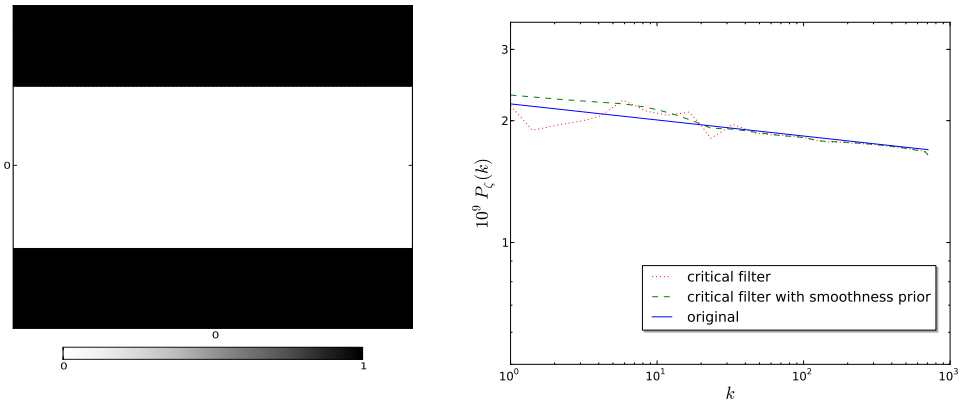


Figure 3.5.: Primordial power spectrum reconstruction of approximately Gaussian curvature perturbations without (a) and with [(b), (c)] features by a critical filter solely (red dotted line) and including a smoothness prior (green dashed line) according to Eqs. (3.55) and (3.58) and compared to the original power spectrum (blue solid line).

The ability of a non-parametric reconstruction of features on the power spectrum like bumps and cutoffs as well as dealing with a partial sky coverage is also illustrated by Figs. 3.5 [(b),(c)] and 3.6. For the latter case we consider a mask in addition to the response, R_{mask}, so that the sky is observed by 50% only. In all cases the critical filter with and without smoothness prior works well, i.e., it is able to reconstruct the primordial power spectrum including possible features.

We want to emphasize, however, that for a proper reconstruction the condition of Gaussianity has to be fulfilled because the approach introduced is very sensitive to deviations from this restriction. To illustrate this we consider again the curvaton scenario where



(a) (b)

Figure 3.6.: Primordial power spectrum reconstruction of approximately Gaussian curvature perturbations [see (b)] at a sky coverage of 50% [see (a)] by a critical filter solely (red dotted line) and including a smoothness prior (green dashed line) according to Eqs. (3.55) and (3.58) and compared to the original power spectrum (blue solid line).

$f_{\rm NL}$ and $g_{\rm NL}$ are parametrized as a function of κ, cf. Eq. (3.30). This means that for large values of κ the curvature perturbation ζ is now falsely assumed to be Gaussian with deviations according to Eq. (3.30). Fig. 3.7 shows the performance of the critical filter with smoothness prior for this case. Panel (a) shows how the reconstructed power spectrum deviates from the true underlying one as a function of the level of non-Gaussianity, parametrized by κ. For a small level of non-Gaussianity there is no observable effect on the spectral index n_s whereas the scalar amplitude A_s depends strongly on κ as shown by panel (c) and (d). The quadratic fits appearing within this panels obey the formula $10^9 A_s(\kappa) = a_q\kappa^2 + b_q\kappa + c_q$ with $a_q = 0.0036$, $b_q = -0.044$ and $c_q = 2.3$ for $\kappa \geq 1$, which can be reformulated analytically into a $f_{\rm NL}$ dependency, if we neglect contributions of the trispectrum (see App. B.2).

For a higher level of non-Gaussianity also n_s becomes affected [Fig. 3.7 (b)]. The two linear fits in panel (b) are generated independently, because the region of $\kappa < 1$ is unphysical. The physically relevant fit for $\kappa \geq 1$ obeys the formula $n_s(\kappa) - 1 = a_{\rm l}\kappa + b_{\rm l}$ with $a_{\rm l} = 8 \times 10^{-4}$ and $b_{\rm l} = -0.04$. The latter formula can also be reformulated into a dependency on $f_{\rm NL}$. Additionally one can derive a quadratic relation between the spectral index and the scalar amplitude, cf. App. B.2.

Note that these particular dependencies on the level of non-Gaussianity (or alternatively on inflationary parameters) is not a generic statement, but valid for the critical filter with and without smoothness prior, which were not informed here about the presence of non-linearities, and that the magnitude of the deviations depends additionally on the number of pixels used.

3.5.3.2. Inferring a power spectrum of non-Gaussian curvature perturbations

Now we show that the critical filter with and without smoothness prior in combination with results of Sec. 3.2 is able to reconstruct the spectrum of the Gaussian, primordial curvature perturbations ζ_1 even in case of significant non-Gaussianity. For this case (of inferring a

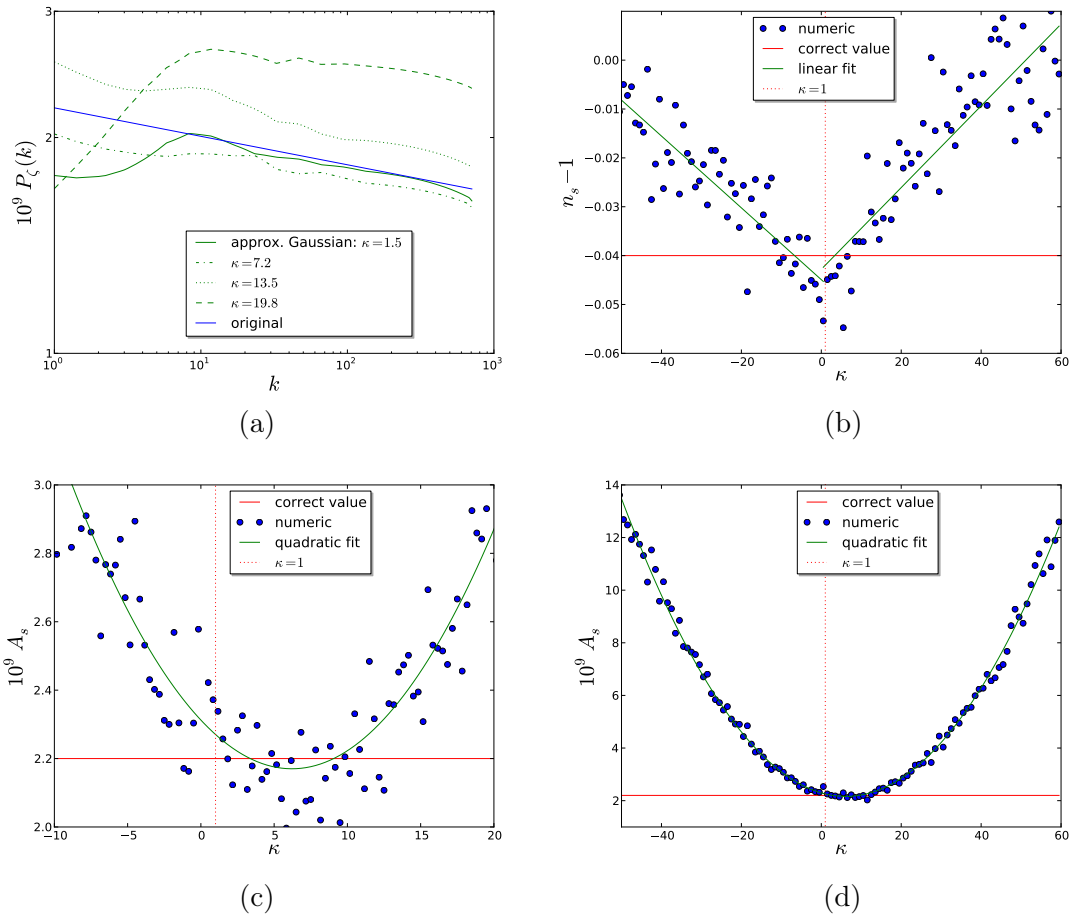


Figure 3.7.: Primordial power spectrum reconstruction of non-Gaussian curvature perturbations, which are falsely assumed to be Gaussian, in the curvaton scenario for various choices of κ by a critical filter solely and including a smoothness prior according to Eqs. (3.55) and (3.58). Panel (a) shows the comparison to the original power spectrum. The dependence of A_s [(c), (d)] and $n_s - 1$ [(b)] on κ has been determined by conducting a linear fit to the reconstructed power spectrum for $k \gtrsim 100$.

power spectrum of non-Gaussian curvature perturbations ζ) we leave all numerical specifications in place, but use $N_{\mathrm{pix}} = 1.6 \times 10^5$, $P_{\zeta_1}(k)\delta_{kk'}$ according to Eq. (3.48), and use $\kappa = 7.2$ ($\widehat{=} f_{\mathrm{NL}}, g_{\mathrm{NL}} = 7.2, -29$), 13.5 ($\widehat{=} f_{\mathrm{NL}}, g_{\mathrm{NL}} = 15, -56$) to seed a non-vanishing level of non-Gaussianity. The reconstructed map of ζ_1, from which we infer the power spectrum, is calculated according to Eq. (3.9) and its uncertainty according to Eq. (3.10). Note that these quantities now depend on $f_{\mathrm{NL}}(p)$ and $g_{\mathrm{NL}}(p)$ and thus on a specific inflation model. Figure 3.8 shows the result both for critical filter and its extension including a smoothness prior. Reconstruction errors are also not included. Compared with the improper reconstructed power spectra of Fig. 3.7 (a) (compare in particular the case of $\kappa = 13.5$), the advanced method used here yields adequate results. This comparison also suggests that one could infer the level of non-Gaussianity by measuring both, non-Gaussian and Gaussian power spectrum. ▶ These spectra might be inferred from, e.g., $T-$ and $B-$modes due to the fact that $B-$modes might be less non-Gaussian than $T-$ modes [91]. Afterwards

the level of non-Gaussianity could be determined by the difference between the respective spectral amplitudes. ◀

Once the power spectrum of the Gaussian curvature perturbation ζ_1 is determined the power spectrum of the non-Gaussian curvature perturbation, $\langle \zeta\zeta^\dagger \rangle_{P(\zeta)}$, can also be calculated from $P_{\zeta_1}(k)$. The distribution $P(\zeta)$ required for this calculation can be calculated approximately and is pointed out, e.g., in Ref. [92].

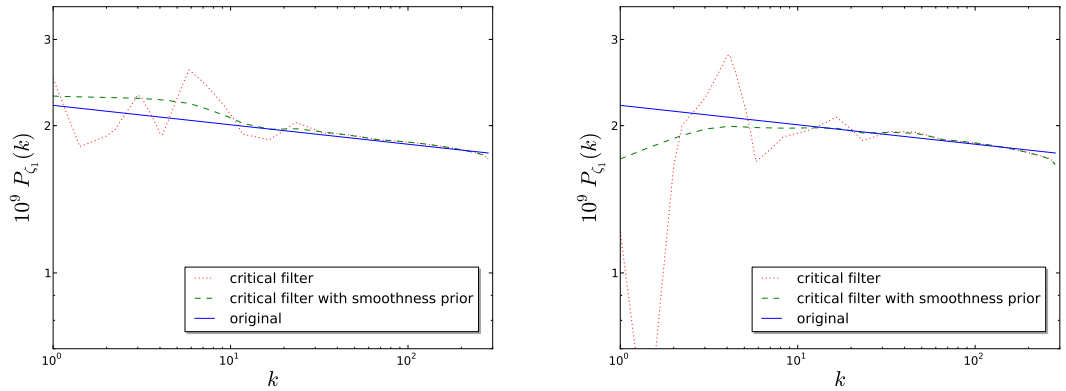


Figure 3.8.: Primordial power spectrum reconstruction of non-Gaussian curvature perturbations with non-vanishing [(a) $\kappa_{\text{gen}} = 7.2$, (b) $\kappa_{\text{gen}} = 13.5$] non-Gaussianity by a critical filter solely (red dotted line) and including a smoothness prior (green dashed line) according to Eqs. (3.55) and (3.58) and compared to the original power spectrum (blue solid line).

3.6. Conclusion

We have presented a novel and generic method to infer inflation models from observations by the non-Gaussianity parameters f_{NL} and g_{NL} and how to reformulate this method to infer specific parameters of inflation models, p, directly (see especially Secs. 3.2 and 3.3). This approach, i.e. the analytical derivation of a posterior for f_{NL} and g_{NL} as well as for p can be used to further distinguish between the already restricted amount of inflation models. It is formulated in a generic manner in the framework of information field theory, so that it is applicable to CMB data as well as to LSS data (see especially the three dimensional example of Sec. 3.4.2) by tuning the response appropriately. The analyticity of the method, achieved by a saddle-point approximation, allows to dispense with numerically expensive sampling techniques like the commonly used Monte Carlo method. The analytic approximation we introduced has been validated successfully by the DIP test [65].

The second quantity of interest here is the primordial power spectrum due to its far-reaching implications for inflationary cosmology. We have presented two computationally inexpensive, approximative Bayesian methods to infer the primordial power spectrum from CMB data, the so called critical filter, Eq. (3.55), and an extension thereof with smoothness prior, Eq. (3.58). Both methods allow a non-parametric reconstruction of the power spectrum including the reconstruction of possible features on specific scales. Additionally, both methods are able to perform this inference process even in the case of partial sky

coverage and non-Gaussianity. We have argued that this property would allow to infer the level of non-Gaussianity of a field if one could measure both, the power spectrum of the non-Gaussian and Gaussian curvature perturbations. These spectra might be inferred from, e.g., $T-$ and $B-$modes due to the fact that $B-$ modes might be less non-Gaussian than $T-$ modes [91]. A fully quantitative analysis thereof, however, is left for future work.

4. All-sky reconstruction of the primordial scalar potential from WMAP temperature data

Note: This chapter has been published in JCAP **2**, 041 (2015) [93].

4.1. Introduction & motivation

The cosmic microwave background radiation (CMB) is presently one of the most informative data sets for cosmologists to study the physics of the early Universe. Of actual interest is in particular the verification of the existence of an inflationary phase of the Universe and investigations of the physical properties of the involved inflaton field(s). An essential quantity is thereby the primordial adiabatic scalar potential Φ. Its statistic, especially the two-point function, was determined during inflation, when the quantum fluctuations of the inflationary field were frozen during their exit of the Hubble horizon. This statistic is conserved on super-horizon scales during the epoch of reheating until the individual perturbed modes re-enter the horizon. Therefore, significant information on the inflationary phase is encoded in the observable quantity Φ. The processes translating the initial modes after their horizon re-entry into the observed CMB fluctuations are described by the so-called radiation transfer functions, see Refs. [4, 64]. As a consequence, many inference methods aim at constraining parameters of the early Universe involve Φ or their statistics. Therefore the CMB fluctuations provide a highly processed view on the primordial scalar potential. In this work, we attempt, however, their direct reconstruction and visualization via Bayesian inference. Once they are reconstructed a direct investigation of their statistics is possible, e.g., the inference of the primordial power spectrum, their connection to large scale structure [94], or primordial magnetic fields [95, 96].

The Planck observation, Ref. [39], of the almost homogeneous and isotropic CMB have shown that the statistical deviations from Gaussianity of the primordial modes/perturbations are still consistent with zero. Therefore, the two-point correlation function of Φ seems to describe nearly fully the statistics of the early Universe up to high accuracy. This fact simplifies the inference of these modes significantly (see, e.g., Ref. [31, 53]), and enables a well justified all-sky reconstruction of the primordial scalar potential from real data.

This work is organized as follows. In Sec. 4.2 we present a Bayesian inference approach to reconstruct the primordial scalar potential. This method, initially proposed by Ref. [64], requires the knowledge of the primordial power spectrum. We show further how Φ and its spectrum can be inferred (unparametrized) even without such an a priori knowledge or assumption. In Sec. 4.3, we reconstruct the primordial scalar potential with corresponding 1σ-uncertainty from WMAP temperature data [97] and partially its initial power spectrum. In Sec. 4.4, we summarize our findings. Exact derivations of all used reconstruction methods can be found in appendices C.1-C.3.

4.2. Inference approach

We derive the inference methods within the framework of information field theory (IFT) [21], where Φ is considered to be a physical scalar field, defined over the Riemannian manifold $\mathbb{R}^3$. Since there is no solid evidence that Φ is non-Gaussian, we assume its statistics to be Gaussian with a covariance matrix determined by its power spectrum[1], i.e.,

$$\Phi \hookleftarrow \mathcal{G}(\Phi, P^\Phi) \quad \text{with} \quad P^\Phi(k,q) \equiv \left\langle \Phi\Phi^\dagger \right\rangle_{(\Phi)} = (2\pi)^3 \delta(k-q) P^\Phi(k). \tag{4.1}$$

Thereby we introduced the notation

$$\mathcal{G}(a,A) \equiv \frac{1}{\sqrt{|2\pi A|}} \exp\left(-\frac{1}{2} a^\dagger A^{-1} a\right) \quad \text{and} \quad \langle\ .\ \rangle_{(a)} \equiv \int \mathcal{D}a\ .\ \mathcal{G}(a,A), \tag{4.2}$$

with corresponding inner product

$$a^\dagger b \equiv \int_{\mathbb{R}^3} d^3x\ a^*(x) b(x) \tag{4.3}$$

for the fields $a,\ b$. Here, $\dagger$ denotes a transposition, t, and complex conjugation, $*$. The CMB data, on the other hand, are of discrete nature, i.e., $d \equiv (d_1, \ldots, d_n)^t \in \mathbb{R}^n,\ n \in \mathbb{N}$.

4.2.1. Temperature only

To set up a Bayesian inference scheme for the primordial scalar potential Φ we have to know how the data d are related to Φ. In the case of the data being the WMAP CMB temperature map this relation is well known, given by [98]

$$\begin{aligned} d_{\ell m} &\equiv (R\Phi)_{\ell m} + n_{\ell m} \\ &= M_{\substack{\ell m \\ \ell' m'}} B_{\ell'} \frac{2}{\pi} \int dk\ k^2 \int dr\ r^2 \Phi_{\ell' m'}(r) g^T_{\ell'}(k) j_{\ell'}(kr) + n_{\ell m}, \end{aligned} \tag{4.4}$$

where $g^T_\ell(k)$ denotes the adiabatic radiation transfer function of temperature, $j_\ell(kr)$ the spherical Bessel function, $n \in \mathbb{R}^n$ the additive Gaussian noise, and B_ℓ the beam transfer function of the WMAP satellite. Repeated indices are implicitly summed over unless they are free on both sides of the equation. We assume the noise to be uncorrelated to Φ. The operator R, which transforms Φ into the CMB temperature map, is assumed to be linear consisting of an integration in Fourier space as well as over the radial (comoving distance) coordinate plus the instrument's beam convolution and a foreground mask, M. Since there is currently no hint for isocurvature modes we exclude them from all calculations.

The next logical step, the construction of an optimal[2] linear filter within the framework of IFT, e.g. the Wiener filter [27] (see, e.g., Ref. [21]), is straightforward. Given the actual, very high resolution of current CMB data sets this, however, turns out to be extremely expensive.

Fortunately, there is a way to split this single computation of reconstructing the primordial scalar potential into multiple. Instead of reconstructing the three-dimensional Φ in a single blow, one can reconstruct it spherically slice by slice, each slice corresponding to a

[1] Here we assume that Φ is also statistically homogeneous and isotropic.
[2] Optimal with respect to the $\mathcal{L}^2-$error norm.

specific radial coordinate starting from $r = 0$ to beyond the surface of last scattering (LSS), r_{LSS}. To understand this procedure we want to recall the definition of the response stated in Ref. [21], where R is the part of the data which correlates with the signal, $R\Phi = \langle d \rangle_{(d|\Phi)}$. It is straightforward to show that this is equivalent to

$$R \equiv \left\langle d\Phi^\dagger \right\rangle_{(\Phi,d)} \left\langle \Phi\Phi^\dagger \right\rangle^{-1}_{(\Phi,d)} . \tag{4.5}$$

To obtain the response acting on a sphere with corresponding comoving distance r it can now also be defined as the expectation value of the data given Φ restricted to a sphere instead of over the three-dimensional regular space, i.e., $R^{(2)}\Phi(r = \text{const.}) = \langle d \rangle_{(d|\Phi(r=\text{const.}))}$. The exact derivation of this modification can be found in App. C.1 and yields

$$\begin{aligned} R^{(2)}_{\substack{\ell m \\ \ell' m'}}(r) &= M_{\substack{\ell m \\ \ell'' m''}} B_{\ell''} \frac{\int dk\ k^2 P^\Phi(k) j_{\ell''}(kr) g^T_{\ell''}(k)}{\int dk\ k^2 P^\Phi(k) j^2_{\ell''}(kr)} \delta_{\ell''\ell'} \delta_{m''m'} \\ &\equiv M_{\substack{\ell m \\ \ell'' m''}} B_{\ell''} R_{\ell''} \delta_{\ell''\ell'} \delta_{m''m'}, \end{aligned} \tag{4.6}$$

with superscript "(2)" indicating that this response acts on the (two-dimensional) sphere $\Phi_{\ell m}(r = \text{const.})$. Initially, we assume P^Φ to be known (see Sec. 4.2.3 if not), i.e. that it is determined via the primordial power spectrum of comoving curvature perturbations $\mathcal{R}$, given by

$$P^{\mathcal{R}}(k) \equiv \frac{2\pi^2}{k^3} A^s_* \left(\frac{k}{k_*} \right)^{n^s_* - 1}, \tag{4.7}$$

with k_* the pivot scale with related primordial scalar amplitude A^s_* and scalar spectral index n^s_*. During matter domination, the relation

$$\mathcal{R} = -\frac{5}{3}\Phi \tag{4.8}$$

is valid. Hence, the primordial power spectrum of Φ is given by

$$P^\Phi(k) = \frac{9}{25} \frac{2\pi^2}{k^3} A^s_* \left(\frac{k}{k_*} \right)^{n^s_* - 1} . \tag{4.9}$$

Figure 4.3 shows the predicted data power spectrum using $R^{(2)}(r = \text{const.})$ without instrumental beam, noise, or mask. Having this response, we are able to construct the (data-space version of the) Wiener filter formula (see App. C.13 for details),

$$m^{(2)}(r) = P^\Phi_\ell(r) R^{(2)\dagger}(r) \left[\tilde{C}^{TT} + N \right]^{-1} d, \tag{4.10}$$

with $P^\Phi_\ell(r)$ the primordial power spectrum projected onto the sphere at comoving distance r and $\tilde{C}^{TT} = RP^\Phi R^\dagger = MBC^{TT}B^\dagger M^\dagger$ where

$$C^{XY}_\ell = \frac{2}{\pi} \int dk\ k^2 P^\Phi(k) g^X{}_\ell(k) g^Y{}_\ell(k). \tag{4.11}$$

X, Y can denote temperature T or polarization $E-$mode. Equation (4.10) provides an optimal estimator of $\Phi_{\ell m}(r)$ and was stated first[3] in Ref. [99]. The huge advantage of this

[3]For a detailed derivation see App. C.1 and C.2.

method is the reduction of computational time, by separating the whole inverse problem into many independent distance-dependent ones. This method permits an easy parallelization of the Wiener filter[4] in the three-dimensional space. The 1σ uncertainty of this estimate, $\Delta m^{(2)}(r)$, is given by [21]

$$\begin{aligned}\Delta m^{(2)}(r) &\equiv \pm\sqrt{\mathrm{diag}\,[D]} \\ &= \pm\sqrt{\mathrm{diag}\left[P_\ell^\Phi - P_\ell^\Phi R^{(2)\dagger}\left(\tilde{C}^{TT}+N\right)^{-1}R^{(2)}P_\ell^\Phi\right]},\end{aligned} \tag{4.12}$$

where we have introduced the posterior covariance D in data space. A proxy of this formula, used in our numerical calculations, can be found in App. C.2.

4.2.2. Temperature and polarization

With future data releases of current experiments like Planck [33], it should be possible to include polarization data (P) with acceptable signal-to-noise level into considerations. Including polarization measurements, parametrized by the Stokes parameters Q,and U, the data is given by

$$d = \begin{pmatrix} d^T \\ d^Q \\ d^U \end{pmatrix} = R\Phi + \begin{pmatrix} n^T \\ n^Q \\ n^U \end{pmatrix} \tag{4.13}$$

with corresponding response

$$R = \underbrace{\begin{pmatrix} M_T B & 0 & 0 \\ 0 & M_P B & 0 \\ 0 & 0 & M_P B \end{pmatrix} W^{T,E}_{T,Q,U}}_{\equiv R^{T,E}_{T,Q,U}} \begin{pmatrix} R^T \\ R^E \\ 0 \end{pmatrix}, \tag{4.14}$$

where $R^{T,E}$ captures the radiation transfer, i.e.,

$$\left(R^{T,E}\Phi\right)_{\ell m} \equiv \frac{2}{\pi}\int dk\; k^2 \int dr\; r^2 \Phi_{\ell m}(r) g_\ell^{T,E}(k) j_\ell(kr). \tag{4.15}$$

The adiabatic radiation transfer functions are $g^{T,E}$ for temperature and E-mode polarization, respectively. For the formal definition of $g^{T,E}$ see, e.g., Refs. [100, 101]. The operator $W^{T,E}_{T,Q,U}$ transforms a vector, containing temperature and E-mode polarization, into Stokes I, Q, U parameters, which are directly measured by experiments like WMAP or Planck. Therefore the generalized data-space version of the Wiener filter equation reads

$$\begin{aligned} m^{(2)}(r) = P_\ell^\Phi(r) &\left(R_T^{(2)\dagger}(r)\; R_E^{(2)\dagger}(r)\; 0\right)\left(R^{T,E}_{T,Q,U}\right)^\dagger \\ &\times\left[R^{T,E}_{T,Q,U}\begin{pmatrix} C_\ell^{TT} & C_\ell^{TE} & 0 \\ C_\ell^{TE} & C_\ell^{EE} & 0 \\ 0 & 0 & 0 \end{pmatrix}\left(R^{T,E}_{T,Q,U}\right)^\dagger + N\right]^{-1}\begin{pmatrix} d^T \\ d^Q \\ d^U \end{pmatrix},\end{aligned} \tag{4.16}$$

[4]The matrix inversion within Eq. (4.10), often solved by Krylov subspace methods like the conjugate gradient method, is often computationally (very) expensive.

where $R^{(2)}_{X=T,E}$ denotes the two-dimensional version of R^X, analogous to Eq. (4.6). The uncertainty is given analogously to Eq. (4.12).

The inclusion of polarization data will result in a significant improvement of reconstruction quality not least because $g^T_\ell(k)$ and $g^E_\ell(k)$ are out of phase and thus compensating the $\ell-$blind spots of each other, which was also noticed by Ref. [99] and can be observed in their Fig. 1. This is, however, only correct if the polarization data is not highly dominated by noise.

4.2.3. Primordial power spectrum reconstruction

Once a signal estimate (optimally with uncertainty) is available the power spectrum of the stochastic process underlying the signal generation might be inferred. Usually, however, an initial guess of the signal power spectrum is required to obtain a Wiener filter signal in the first place. This initial guess spectrum can affect the spectrum estimate and therefore might act as a hidden prior. In order to forget the initial guess, the procedure of signal and spectrum inference should be iterated until it has converged onto a spectrum that is then independent of the initial starting value. Fortunately, the primordial power spectrum is constrained well by the existing[5] CMB data-sets so that this process should converge rapidly. This iterative, unparametrized method was derived in Refs. [25, 103] and named critical filter. It can be regarded as a maximum a posteriori estimate of the logarithmic power spectrum and the assumption of a scale invariant Jeffreys prior of its amplitudes. The power spectrum on the sphere is written as

$$P^\Phi_{\vec{\ell}\vec{\ell}'} = \delta_{\vec{\ell}\vec{\ell}'} P^\Phi_\ell \quad \text{with} \quad \vec{\ell} \equiv (\ell, m). \tag{4.17}$$

The iterative critical filter formula including a spectral smoothness prior is then given by Eq. (4.10) and

$$P^\Phi_\ell = \frac{\sum_{\{\vec{\ell}'|\ell'=\ell\}} \left(m^{(2)}_{\vec{\ell}} m^{(2)\dagger}_{\vec{\ell}} + D_{\vec{\ell}\vec{\ell}'} \right)}{\rho_\ell + 2(S \ln P^\Phi)_\ell}, \tag{4.18}$$

where $\rho_\ell = \sum_{\{\vec{\ell}'|\ell'=\ell\}} 1$ is the number of degrees of freedom on the multipole ℓ and S an operator that enforces smoothness (for details see Ref. [103]).

4.3. Temperature-only reconstruction of the primordial scalar potential

4.3.1. Input values and settings

We analyze the full resolution (nside = 512) coadded nine-year WMAP (foreground-cleaned) V-band frequency temperature map, masked with the primary temperature analysis mask (KQ85: 74.8% of the sky). The data as well as the corresponding beam transfer function and noise properties (see App. C.3) we used can be found at `http://lambda.gsfc.nasa.gov/product/map/dr5/m_products.cfm` [97, 104]. We did not take polarization data into considerations due to the suboptimal signal-to-noise levels. To be consistent with the WMAP team's measurements we use the cosmological parameters obtained by their data

[5] See section 7 of Ref. [10] and Ref. [102] for an overview of the literature on such methods.

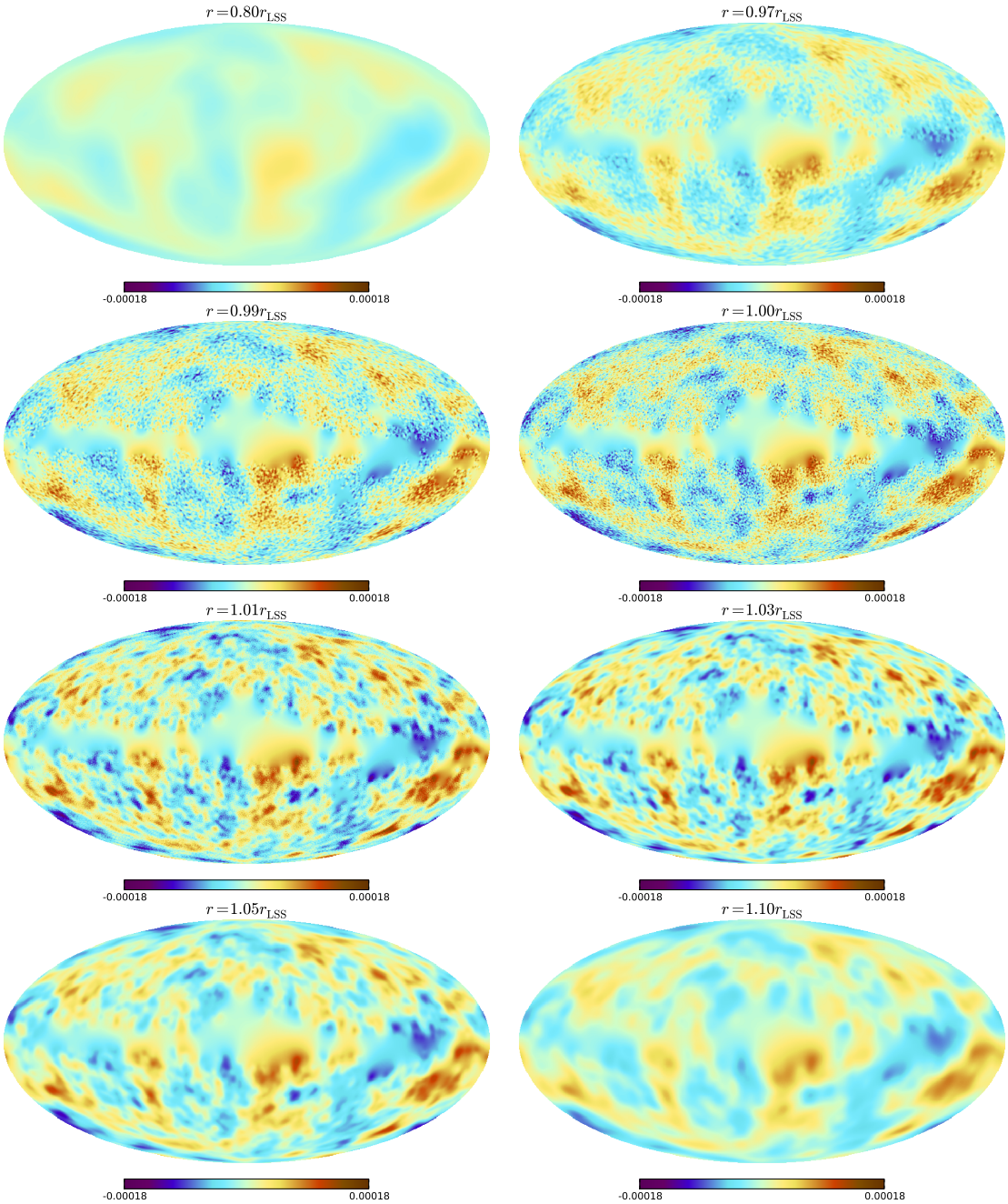


Figure 4.1.: All-sky maps of the reconstructed primordial scalar potential at different comoving distances according to Eq. (4.10) in the vicinity of the recombination sphere with $r = r_{\rm LSS}$. A Mollweide projection is used.

analysis to compute the radiation transfer function as well as the primordial power spectrum. In particular this has been done by using `gTfast`[6], which is based on `CMBFAST`[7] [105]. We used the following settings: pivot scale $k_* = 0.002$ Mpc^{-1}, spectral index $n_*^s = 0.962$, spectral amplitude $A_*^s = 2.46 \times 10^{-9}$, noise level $\sigma_0^{\rm V-band} = 3.131 \times 10^{-3}$ K, CMB temper-

[6] `http://www.mpa-garching.mpg.de/~komatsu/CRL/nongaussianity/radiationtransferfunction/`
[7] `http://lambda.gsfc.nasa.gov/toolbox/tb_cmbfast_ov.cfm`

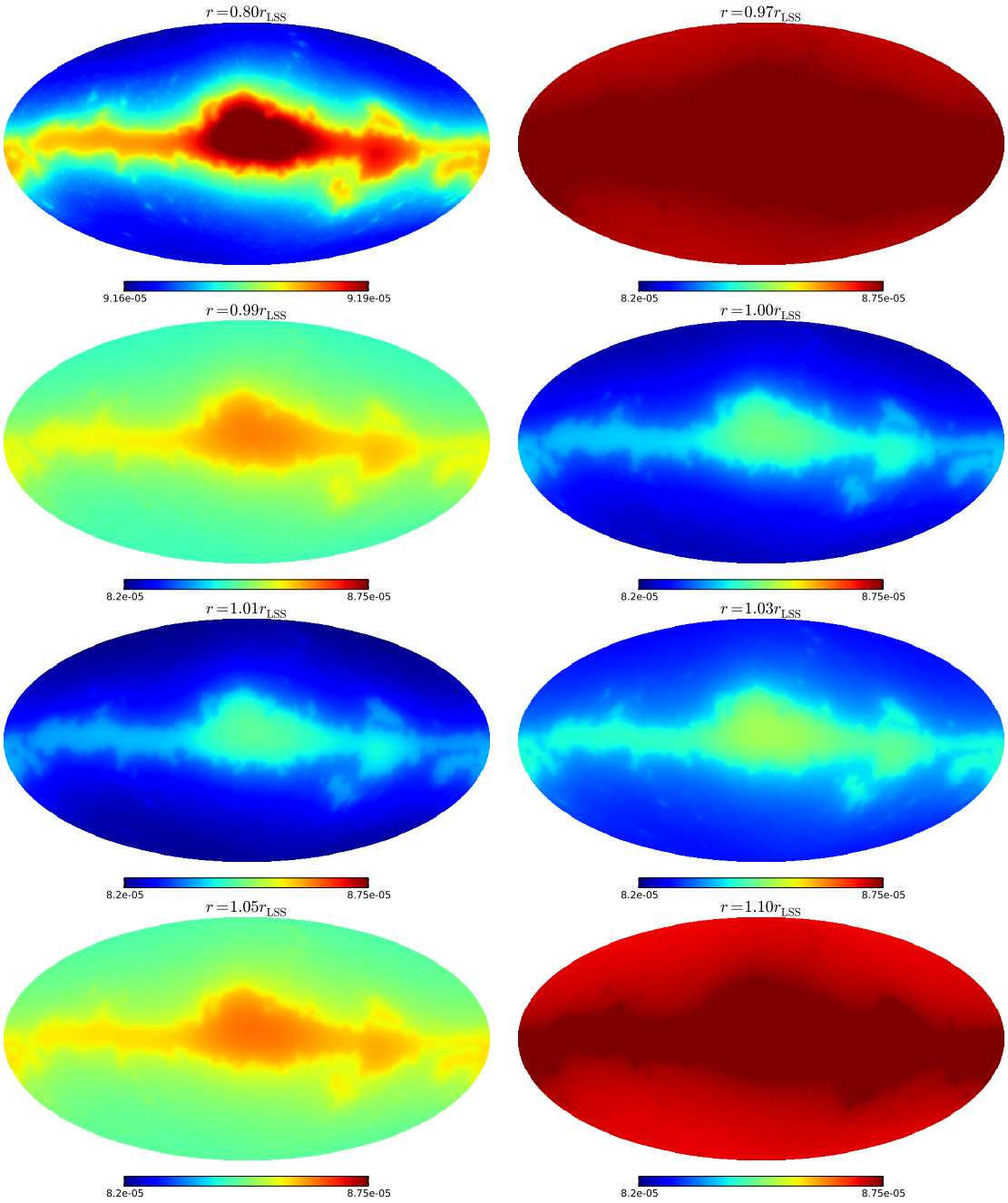


Figure 4.2.: 1σ uncertainty maps of the corresponding all-sky maps of of Fig. 4.1 according to Eq. (C.19) in the vicinity of the recombination sphere with $r = r_{\rm LSS}$. A Mollweide projection is used. Note that the color bar for $r = 0.80 r_{\rm LSS}$ is a different one, showing the natrual bounds of the uncertainty map. All uncertainty maps share this morphology.

ature $T_{\rm CMB} = 2.726$ K, optical depth $\tau = 0.088$, density parameters $\Omega_b = 0.046$, $\Omega_c = 0.0231$, $\Omega_\Lambda = 0.723$, Hubble constant $H_0 = 70.2$ km/s/Mpc, helium abundance $Y_{\rm He} = 0.24$, and the effective number of massless neutrino species $N_\nu^{\rm eff} = 3.04$. The resulting distance to the LSS amounts 1.40147×10^4 Mpc.

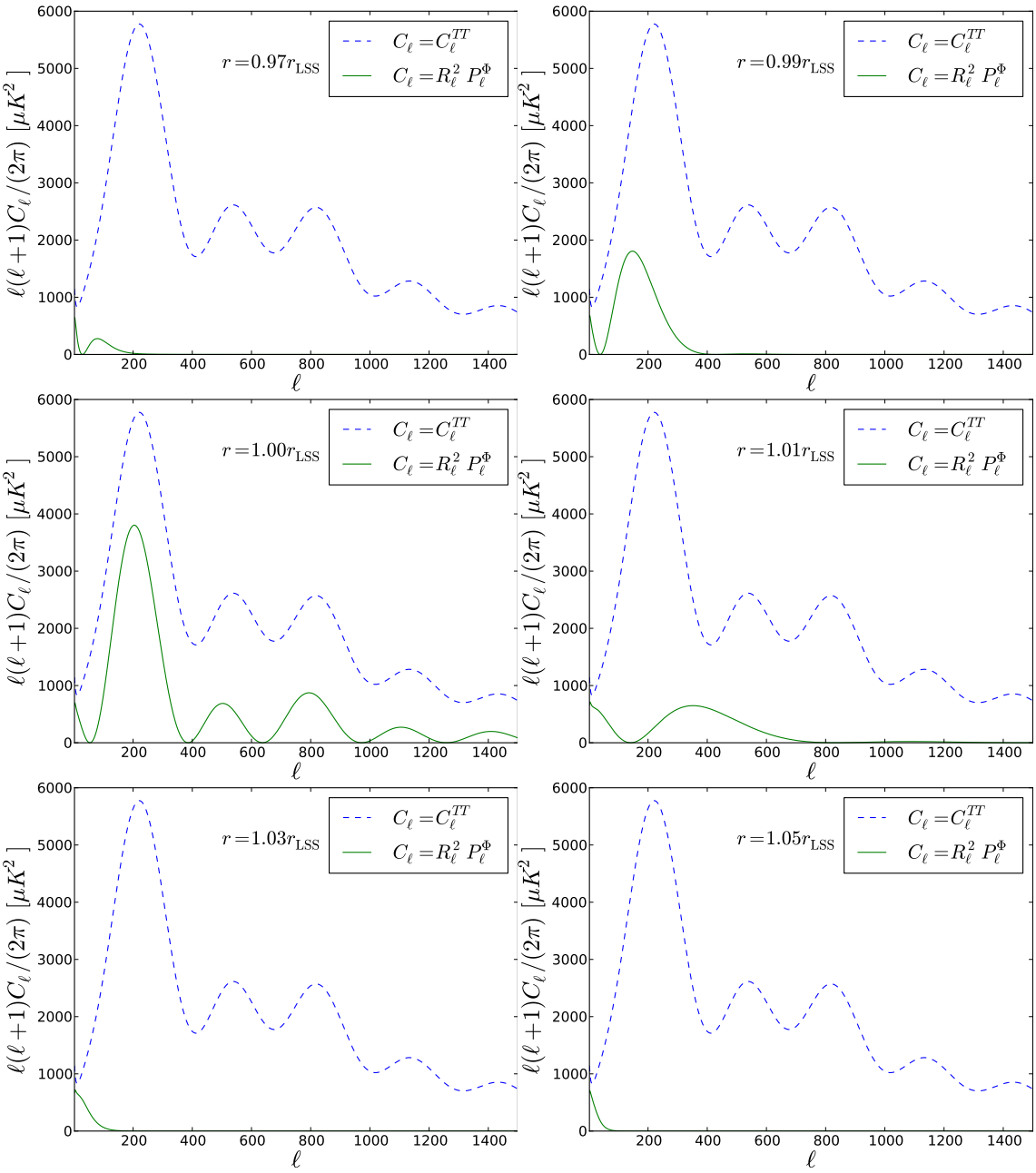


Figure 4.3.: Predicted power spectra of data simulated with the estimator response $R^{(2)}$ compared to the CMB data power spectrum. The $\ell-$blind spots move from large scales at distances $r < r_{\rm LSS}$ to small scales at $r > r_{\rm LSS}$. The amplitude of the predicted power spectra gets maximal at $r = r_{\rm LSS}$. For clarity and comprehensibility we exclude the instrumental beam, noise, and observational mask.

4.3.2. Results

With the parameters defined in the previous paragraph, we have reconstructed a shell around the last scattering surface ($0.8 \times r_{\rm LSS}$ to $1.1 \times r_{\rm LSS}$) in 151 slices as well as additional 6 slices within the range $(50\% - 80\%) \times r_{\rm LSS}$ from real data, see Fig. 4.1. For all reconstructions 1σ-uncertainty maps are provided, see Fig. 4.2 as well as the relative

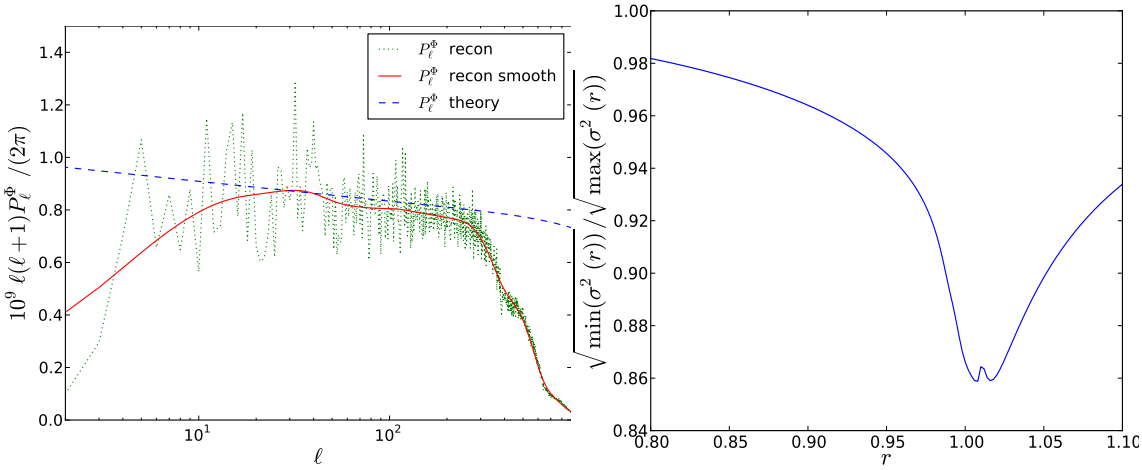


Figure 4.4.: Left: Estimated primordial power spectrum of $\Phi(r = r_{\rm LSS})$ according to Eq. (4.18). Masking effects as well as the estimators power loss are compensated. At scales smaller than $\ell \approx 300$ the reconstruction fails due to sub-horizon physics [99] and noise-dominance. Right: Relative 1σ-uncertainty along the radial coordinate. Minimal values of σ correspond to Eq. (C.18) with "no mask", maximal values to the same equation with "all mask".

1σ-error along the radial coordinate, see Fig. 4.4 (Right). A detailed description of the calculation of these uncertainty maps can be found in App. C.2. The respective data files of the reconstruction can be found at `http://www.mpa-garching.mpg.de/ift/primordial/`. For the most interesting sphere at $r = r_{\rm LSS}$ we also provide a power spectrum estimate, see Fig. 4.4 (Left). This power spectrum estimate has been obtained with the critical filter formula with smoothness prior but without iterations[8] and D set to zero (defined in Eq. (4.12)).

We also phenomenologically[9] corrected for the effect of masking and power-loss in the predicted power spectra of data simulated with the estimator response $R^{(2)}$ in comparison to the power spectrum of Eq. (4.11). Therefore our spectrum estimate should rather be regarded as providing a consistency check of the algorithm than to necessarily provide precisely the cosmological power spectrum. Having stated these caveats, we like to note that a deviation from the power-law primordial power spectrum is not apparent over roughly one order of magnitude in Fourier space.

Some of the reconstructed slices of the primordial scalar potential might look suspiciously crumby at first. The reason for this property are the $\ell-$blind spots in the response R_ℓ. Figure 4.3 shows the noiseless data power spectrum, $C_\ell^{TT} = RP^\Phi R^\dagger$, as well as the power spectrum $R^{(2)}P_\ell^\Phi R^{(2)\dagger}$ expected from noiseless, distance dependent data obtained with the estimator response, $d^{(2)} = R^{(2)}\Phi$. The $\ell-$blind spots are clearly recognizable, which move

[8] With the correct application of the critical filter (iterative) one might be able to detect features in the primordial power spectrum [31]. This, however, would require a highly resolved data set including polarization to compensate for the $\ell-$bind spots (one cannot get rid of with temperature data only) with a high signal-to-noise level in T-, Q-, and U-data maps. Perfect candidates for such data sets are future CMB experiments and Planck polarization data releases.

[9] The power-loss is corrected by convolving the reconstructed Φ with $\alpha_\ell \equiv \sqrt{C_\ell^{TT}/(R_\ell^2 P_\ell^\Phi)}\ \ \forall \ell : R_\ell^2 P_\ell^\Phi \neq 0$ before performing the power spectrum estimation. We also investigated how the mask affects the power spectrum of Φ, by calculating $\beta_l \equiv \left\langle \text{power}\left[R^{\rm mask}(\Phi)\right]/P_\ell^\Phi \right\rangle$ where power[.] denotes the application of the critical filter formula with smoothness prior. We re-scaled the inferred power spectrum with $1/\beta_\ell$.

from large scales at distances $r < r_{\mathrm{LSS}}$ to small scales at $r > r_{\mathrm{LSS}}$, where the amplitude of this power spectrum gets maximal at $r = r_{\mathrm{LSS}}$.

The numerical and computational effort to reconstruct one slice by one CPU amounts to roughly 45 minutes, which simultaneously represents the time for reconstructing the whole three-dimensional primordial scalar potential at full parallelization. In our numerical implementation we used the `conjugate gradient` method to solve Eq. (4.10).

4.4. Conclusion & outlook

We have presented a reconstruction of the primordial scalar potential Φ with corresponding 1σ-uncertainty from WMAP temperature data. This has been achieved by setting up an inference approach that separates the whole inverse problem of reconstructing Φ into many independent ones, each corresponding to the primordial scalar potential projected onto a sphere with specific comoving distance. This way the reconstruction is done sphere by sphere until one obtains a thick shell of nested spheres around the surface of last scattering. This results in a significant reduction of computational costs (since the reconstruction equation (Wiener filter) parallelizes fully), if only the small region around the last scattering surface is reconstructed, which is accessible through CMB data.

We did not include polarization information yet due to the suboptimal signal-to-noise ratios of the WMAP polarization data. Hence we do not expect a huge improvement when additionally including WMAP Stokes Q and U parameters into the Wiener filter equation. This, however, will definitely change when the polarization data of Planck will be available in the near future. Once one uses simultaneously temperature and polarization data, the $\ell-$blind spots in the reconstructions will disappear and with it the crumbliness of the maps. At this point it also might be more rewarding to apply the critical filer equations to simultaneously obtain the power spectrum of the primordial scalar potential.

5. Stochastic determination of matrix determinants

Note: This chapter has been published in Phys. Rev. E **92**, 013302 (2015) [106].

5.1. Motivation

Current and future physical observations generate huge data streams to be analyzed. Particle physics, biophysics, astronomy, and cosmology are representatives of current scientific fields of interest that are undergoing a revolution driven by increasing data volume. Typical large data sets in cosmology are, for instance, observations of the cosmic microwave background [33, 97] as well as of the large-scale structure [107–109] as they are often wide- or all-sky observations carried out by telescopes with remarkable resolution. In order to extract information about the universe or physics in general, Bayesian inference methods becomes more and more frequently used as their large computational demands become more feasible thanks to technology developments. The signal of interest to be extracted from data could be almost everything, ranging from just a single parameter (e.g., the level of local non-Gaussianity of the cosmic microwave background [53, 98]) to a full four-dimensional reconstruction of the structure growth in the universe [94, 110]. Such ambitious Bayesian analyses often invoke linear transformations of the data or of estimated signal vectors.

The size of the involved data and signal spaces often bans the explicit representation of matrices acting on these spaces by their individual matrix elements. A prominent example appearing in many analyses is, for instance, the covariance matrix of a multivariate Gaussian distribution of a vector valued quantity, which describes the two-point correlation structure of the said quantity. Due to their large dimensions such matrices are often only representable by a computer routine, which implements the application of the matrix to a vector without storing or even calculating the individual matrix elements. Such routines often invoke fast Fourier transformations and other efficient operations, which in combination render nonsparse matrices into easily computable basis systems. We refer to such a matrix as an *implicit matrix*. For instance, calculating the model evidence often requires calculating determinants of such matrices. This work provides an efficient way to numerically calculate determinants given only by an implicit matrix representation.

The remainder of this work is organized as follows: In Sec. 5.2 we introduce the formalism of the stochastic estimation of an implicit matrix and present two numerical examples. Section 5.3 provides a perspective of possible applications in science. Results are summarized in Sec. 5.4.

5.2. Probing the log-determinant of an implicit matrix

5.2.1. Formalism

Let $A = (a_{ij}) \in \mathbb{C}^{n\times n}$ be an implicitly defined, complex-valued, square matrix of order n. Implicitly means that the particular entries of the matrix are not accessible, for instance, if dealing with large data sets, where an explicit storage of A might exceed the memory of the computer. However, the action of the matrix as a linear operator is assumed to be known and given by a computer routine implementing the mapping $x \mapsto Ax$.

Motivated by applications in science and statistics (Secs. 5.1 and 5.3), in particular by signal reconstruction techniques and model comparison in astronomy and cosmology, where the determinant of a covariance matrix is required (Sec. 5.3), we constrain the variety of different types of matrices by requesting that the matrix A of interest is either *weak diagonal dominant* or *Hermitian positive definite*. The term *weak diagonal dominant* is defined by

$$|a_{ii}| \geq \sum_{i\neq j} |a_{ij}| \quad \forall i, \tag{5.1}$$

while *Hermitian positive definite* means

$$A^\dagger = A \quad \text{and} \quad x^\dagger A x > 0 \quad \forall x \in \mathbb{C}^n \backslash \{0\} \tag{5.2}$$

with † denoting the adjoint.

The diagonal and the trace of an implicit matrix can be obtained by exploiting common probing routines [111–114]. A stochastic estimate of the diagonal of the linear operator A is given by

$$\mathrm{diag}(A) = \langle \xi \star A\xi \rangle_{\{\xi\}} \approx \frac{1}{M} \sum_{i=1}^{M} \xi_i \star A\xi_i, \tag{5.3}$$

where $\star$ denotes a componentwise product, $M = |\{\xi\}|$ the sample size, and $\langle\cdot\rangle_{\{\xi\}}$ the arithmetic mean over ξ with $M \to \infty$. The probing vectors $\xi \in \mathbb{C}^n$ are random variables, whose components x (x') fulfill the condition $\langle \xi_x \xi_{x'} \rangle_{\{\xi\}} = \delta_{xx'}$. Analogously to the diagonal of an operator its trace can be probed by, e.g.,

$$\mathrm{tr}(A) = \left\langle \xi^\dagger A \xi \right\rangle_{\{\xi\}}. \tag{5.4}$$

Recently, there have been investigations to improve these straightforward probing methods by exploiting Bayesian inference [111]. This has been achieved by reformulating the process of stochastic probing of an operator's diagonal (trace) as a signal inference problem. As a result, it requires fewer probes than the purely stochastic methods and thus can decrease the computational costs. With the phrase *operator probing*, be it trace or diagonal probing, we subsequently refer to the entirety of probing methods in general.

The linear operator A can be split into a diagonal matrix $D \in \mathbb{C}^{n\times n}$ and a matrix $N \in \mathbb{C}^{n\times n}$, which contains the off-diagonal part of A, i.e.,

$$A = D + N. \tag{5.5}$$

We are now interested in the value of its determinant or of its log-determinant, $\Delta \equiv \ln[\det(A)]$. In case A is mainly dominated by its diagonal (i.e. $ND^{-1} \ll \mathbb{1}$ spectrally), a

Taylor expansion of the log-determinant might be a reasonable approximation,

$$\begin{aligned}\Delta =& \ln[\det(D+N)] \\ =& \ln[\det(D)] + \operatorname{tr}\left[ND^{-1}\right] + \mathcal{O}\left(\operatorname{tr}\left[\left(ND^{-1}\right)^2\right]\right),\end{aligned} \tag{5.6}$$

which is sometimes feasible dealing with implicit operators, e.g., see Refs. [31, 53] for recent applications in cosmic microwave background physics. This approximation, however, breaks down when the relation $ND^{-1} \ll \mathbb{1}$ (spectrally) is violated. In order to circumvent this problem we introduce the quantity

$$\Delta(t) \equiv \ln[\det(D+tN)] \tag{5.7}$$

with the pseudotime parameter $t \in [0,1]$. For a sufficiently small t the approximation of Eq. (5.6) becomes valid. This property can be used together with a few mathematical manipulations (for details see Appendix D) to obtain the formula

$$\begin{aligned}\Delta =& \int_0^1 dt\ \operatorname{tr}\left[N\left(D+tN\right)^{-1}\right] + \Delta(0) \\ =& \int_0^1 dt\ \left\langle \xi^\dagger N \left(D+tN\right)^{-1} \xi \right\rangle_{\{\xi\}} + \Delta(0)\end{aligned} \tag{5.8}$$

that represents a stochastic estimate of the log-determinant of A using operator probing. In particular, the following steps are required to evaluate Eq. (5.8):

1. Diagonal (operator-) probing to split A into

$$A = \underbrace{\operatorname{diag}(A)}_{\equiv D} + \underbrace{A - \operatorname{diag}(A)}_{\equiv N},$$

2. an approach to invert $D+tN$ in Eq. (5.8), e.g., the `conjugate gradient` method [115],

3. trace (operator-) probing to evaluate the integrand,

4. a numerical integration method, e.g., applying `Simpson's` rule.

It might immediately strike the eye of the reader that one recaptures the simple first-order Taylor-expanded version of the log-determinant, Eq. (5.6), when dropping the pseudotime dependency of the integrand in Eq. (5.8) by requesting $t = 0$. This means that in case of dealing with diagonal dominant operators the value of the correct log-determinant might be received by a coarse numerical integration since the integrand close to $t = 0$ already yields the main correction, which might decrease the computational costs, see Sec. 5.2.2.

Equation (5.8) further represents the main result of this paper and can be regarded as a special case of calculating partition functions (see Sec. 5.3 and Refs. [116, 117]). Although the first line of it, the integral representation of the log-determinant, was also, independently of our work, found by mathematicians 10 years ago [118], it is (to our knowledge) not known in the community of physics or signal inference. The connection to stochastic estimators, however, is a novel way to evaluate the log-determinant of implicitly defined matrices that enables previously impossible calculations, see Sec. 5.3.

5.2.2. Numerical example

We address here a simple and also exactly solvable numerical example referring to (Bayesian) signal inference problems or, in general, statistical problems in physics (see Secs. 5.3.1 and 5.3.2), where the log-determinant of a covariance matrix A is of interest. If we assume statistical isotropy and homogeneity of a physical field, its covariance matrix can be parametrized by a so-called power spectrum. This is often a reasonable assumption[1], e.g., in astronomy and physical cosmology, when applying the cosmological principle. In this case the covariance matrix becomes diagonal in Fourier space,

$$A_{kk'} = c_k \delta_{kk'}, \tag{5.9}$$

with respective Fourier modes k, k' and power spectrum c_k. It is straightforward to show that the position space representation of $A_{kk'}$, given by $A_{xx'} = \mathcal{F}^\dagger_{xk} A_{kk'} \mathcal{F}_{k'x'}$ with Fourier transformation $\mathcal{F}$, is nondiagonal if and only if $c_k \neq \text{const}\ \forall k$. In order to apply the stochastic estimator of the log-determinant we use two special forms of the power spectrum, given by

$$c_k = \frac{1}{(1+k)^\alpha} \tag{5.10}$$

with α set to 2 or 4. A value of $\alpha = 2$ describes a mostly diagonal dominant matrix, whereas $\alpha = 4$ exhibits a significant nondiagonal structure in position space. To be precise, in the following we use a regular, two-dimensional, real-valued grid (over $\mathcal{T}^2$) of $n = 20 \times 20$ pixels to represent our position space, resulting in a matrix A consisting of $n \times n = 1.6 \times 10^5$ real numbers. See Fig. 5.1 for an illustration thereof.

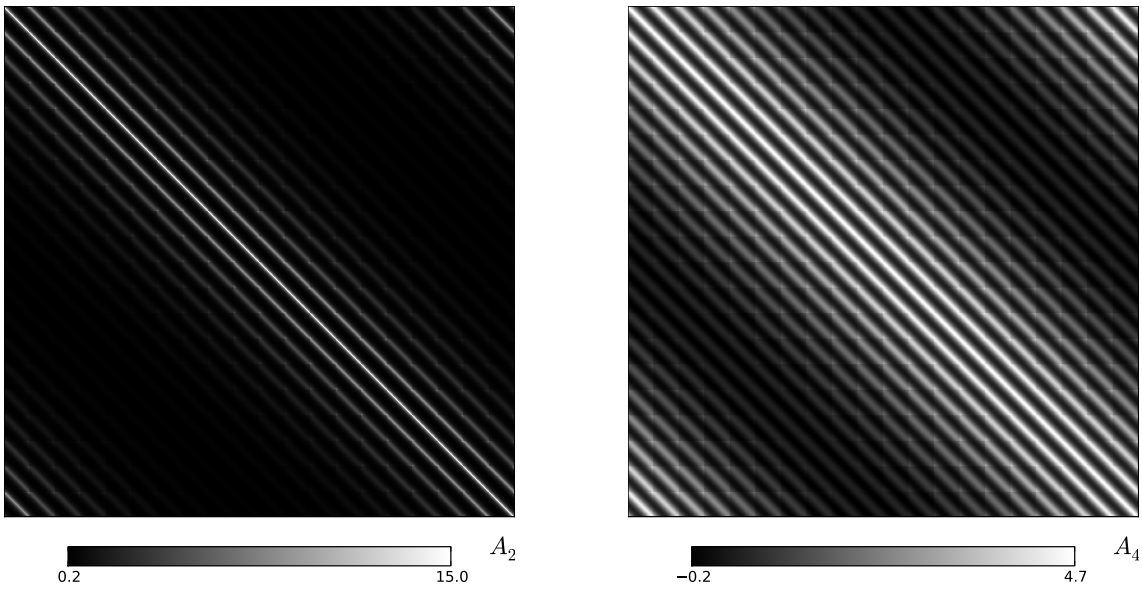


Figure 5.1.: Illustration of the matrices A_2 (left) and A_4 (right) in position space with linear color bars.

For both matrices, which we refer to as A_2 and A_4, we apply Eq. (5.8) given an explicit and implicit numerical implementation. For the explicit variant there also exist

[1] Referring to Bayesian evidence calculations such a matrix might be the prior or posterior covariance, see Sec. 5.3 for details.

Table 5.1.: Results of the numerical determinant calculations with and without probing. The absolute errors of the probing method are defined by $\epsilon_1 = |\Delta_{\text{explicit}}(1) - \Delta_{\text{implicit}}(1)|$ and $\epsilon_2 = |\Delta_{\text{correct}} - \Delta_{\text{implicit}}(1)|$. Differences between ϵ_1 and ϵ_2 arise from the discretized, numerical integration.

	A_2	A_4
$\Delta(0)$	-1308.05	-1771.57
Δ_{correct}	-1566.99	-3107.28
$\Delta_{\text{explicit}}(1)$	-1566.81	-3107.29
$\Delta_{\text{implicit}}(1)$	-1565.33	-3108.41
m	10	1000
M	8	8
ϵ_1	1.48	1.12
ϵ_2	1.66	1.13

well-understood, precise numerical methods[2] to calculate the determinant. Therefore, the numerical results of such a method can be regarded as our gold standard and hence serve as a reference for the probing results. Henceforth we will refer to it using the subscript "correct". Both variants, the explicit and implicit implementation, are realized using the tools of NIFTy [30].

After the separation of A_2 and A_4 into diagonal and off-diagonal parts by applying diagonal probing we calculate the integrands of Eq. (5.8) for the m-part-discretized interval of $t \in [0, 1]$ by using the `conjugate gradient` method as well as trace probing and perform the numerical integration afterwards by using `Simpson's` rule. The operator probing as well as the `conjugate gradient` method have also been realized using NIFTy. Furthermore we introduce the quantities

$$\Delta(x) \equiv \int_0^x dt \ \mathrm{tr}\left[N\,(D + tN)^{-1}\right] + \Delta(0), \quad x \in [0, 1] \tag{5.11}$$

to study the convergence to the final value and $\Delta(m)$ to investigate the dependency on the discretization of the integration interval, see Figs. 5.2, 5.3, and 5.4.

We used a rather low sample size of $M = 8$ for trace and diagonal probing [see Eqs. (5.4) and (5.3)] to demonstrate the applicability of the method to large data sets. The discretization of the pseudotime interval into m parts was chosen to be $m = 10^3$ for A_4 and only $m = 10$ for A_2, see in particular Fig. 5.4, which illustrates the dependence of the probing result on m.

[2]See, for instance, the method described at `http://docs.scipy.org/doc/numpy/reference/generated/numpy.linalg.slogdet.html`, which is based on LU-factorization.

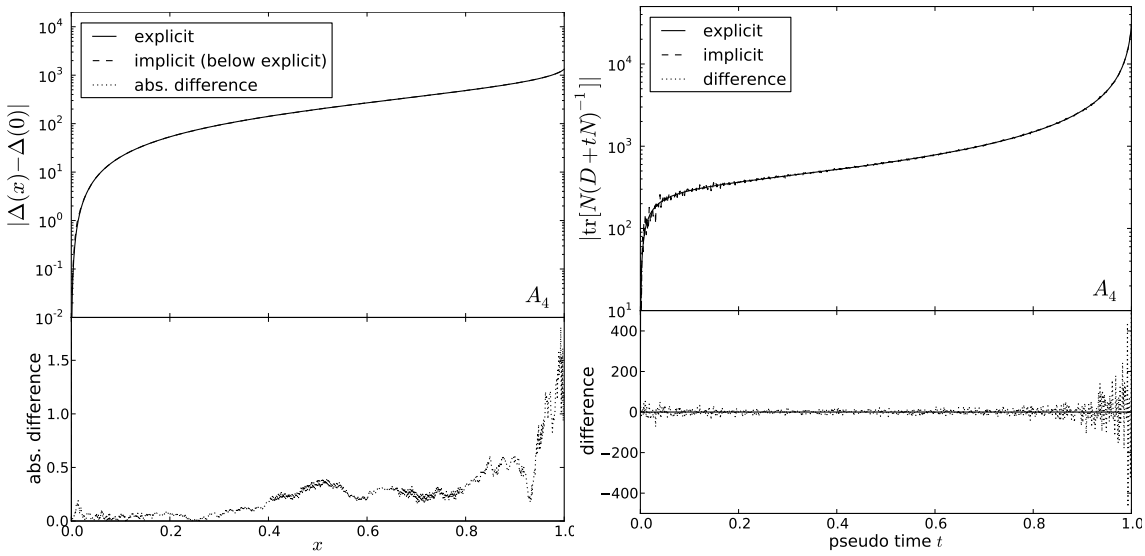


Figure 5.2.: The integrand of Eq. 5.8 (right panel) and $\Delta(x)$ (left panel) for explicit and implicit representations of A_4.

5.2.3. Discussion

The exact numerical values of the determinant calculation using explicit and implicit representations of A_4 and A_2 can be found in Table 5.1. The results of the probing method (implicit) compared to the correct and the explicit method, where Eq. (5.8) can be evaluated without using a `conjugate gradient` or probing techniques, are accurate for both matrices. It is remarkable that despite using a relatively small sample size of $M = 8$ the absolute errors remain relatively small. The reason for this is that the pseudotime integration over all probed integrands averages the probing error. This is of particular importance when applying the log-determinant probing to large data sets, where a large sampling size should be avoided to save computational time. These errors can be decreased further, of course, by an increase of the sampling size and a refinement of the numerical integration.

The results of the trace (integrand) probing and the determinant's convergence behavior as well as their respective errors with respect to the explicit representation can be found in Figs. 5.2 and 5.3. Note that the scaling of the ordinate is logarithmic. For both matrices, but especially for A_4, the largest contribution to the integral of Eq. (5.8) comes from late t-values. Therefore, if dealing with big data sets, one could divide the integration interval not into m equal parts but by starting with a rather coarse discretization for small t-values and subsequently refining it for larger values, e.g., by substituting dt by $d\ln(t')$ and thereby saving computational costs. This, however, might depend on the particular shape of the matrix and has to be studied case by case.

The dependency of the numerical value of the determinant of A_4 on the discretization (in m equal parts) of the integration interval can be found in Fig. 5.4 and shows that even a small value of m adds significant corrections to the result. The result for $m \propto \mathcal{O}(10)$ is, for instance, better than just using the determinant of the diagonal, $\Delta(0)$. This might be used in practice to investigate cheaply whether the nondiagonal structure of a matrix influences the determinant significantly.

A huge advantage of the probing method discussed here is the possibility to parallelize the numerical calculation almost completely. To be precise, the diagonal probing beforehand,

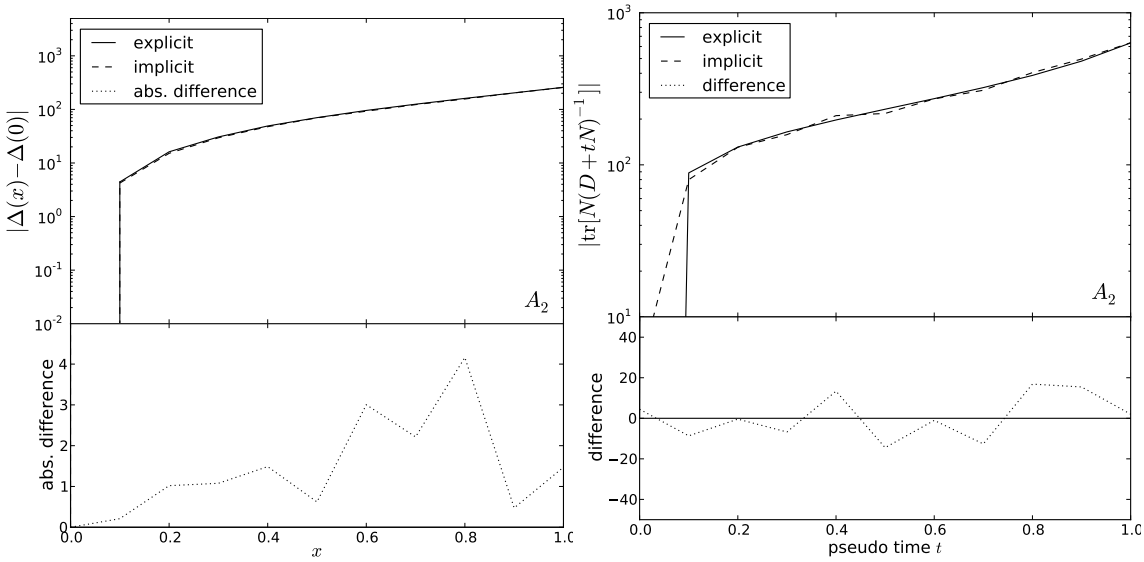


Figure 5.3.: The integrand of Eq. 5.8 (right panel) and $\Delta(x)$ (left panel) for explicit and implicit representations of A_2 with only $m = 10$ steps in pseudotime.

the pseudotime integral, as well as every single trace probing can be parallelized fully. The only operation that cannot be parallelized is the `conjugate gradient` method as it is a potential minimizer, using at least the previous step to calculate the next one.

The determination of a suitable choice of the involved parameters m and M as well as the precision parameters for the used `conjugate gradient` approach and numerical integration method depend highly on the matrix to be studied. The computational costs and precision of the introduced determinant calculation thus depend on the combination of the chosen methods for diagonal and trace probing, numerical integration, the method to numerically invert the matrix $D + tN$, and the matrix A itself. Since it is therefore not possible to make general statements we consciously avoid here such a discussion of computational costs and precision with respect to m and M. A more pragmatic way to estimate these parameters would be to downscale the problem of interest until the matrix of interest fits into the memory of the computer and to subsequently perform mock tests to obtain a suitable choice for the parameters discussed above. Afterwards these values can be extrapolated to the size of the real problem.

5.3. Applications in science

Within this section we present a selection of possible applications in science. Although there are a vast number of research fields and topics which might benefit from the stochastic estimation of a log-determinant we focus henceforth on a selection of usages in Bayesian signal inference, in particular in physics and only present simple examples. Exact, more complicated examples can be found in the cited works within this section.

5.3.1. Evidence calculations & model selection

The Bayesian evidence $\mathcal{P}(d)$ is a measure for the quality of the model and hence for all assumed model parameters for the data d [119]. To keep it short and simple we assume

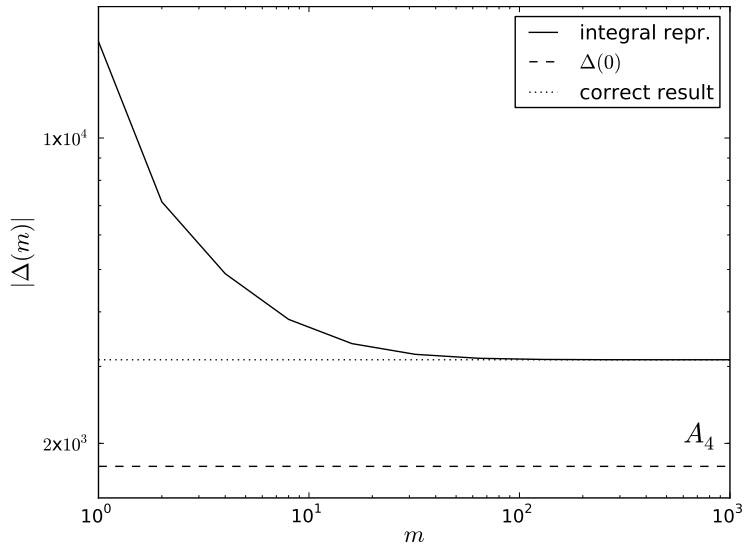


Figure 5.4.: Dependency of the determinant's result on the discretization of the integration interval into m parts, using A_4.

a model that describes a linear measurement of a Gaussian signal s with additive, signal-independent, Gaussian noise n, i.e.,

$$d = Rs + n, \tag{5.12}$$

where R represents a linear operator. A Gaussian distribution of a variable x is defined by

$$\mathcal{P}(x) = \mathcal{G}(x, X) \equiv \frac{1}{\sqrt{|2\pi X|}} \exp\left\{-\frac{1}{2} x^\dagger X^{-1} x\right\} \tag{5.13}$$

with related covariance matrix X and mean

$$\langle x \rangle_{\mathcal{P}(x)} \equiv \int \mathcal{D}x \; x \; \mathcal{P}(x). \tag{5.14}$$

$\int \mathcal{D}[\cdot]$ denotes a phase space integral and $|\cdot|$ the determinant. Under these circumstances the evidence can be calculated as

$$\begin{aligned} \mathcal{P}(d) &= \int \mathcal{D}s \int \mathcal{D}n \; \mathcal{P}(d, s, n) \\ &= \int \mathcal{D}s \int \mathcal{D}n \; \delta(d - Rs - n) \mathcal{P}(n|s) \mathcal{P}(s) \\ &= \sqrt{\frac{|2\pi C_{s|d}|}{|2\pi C_s||2\pi C_n|}} \exp\left\{-\frac{1}{2}\left(d^\dagger C_n^{-1} d - j^\dagger C_{s|d} j\right)\right\}, \end{aligned} \tag{5.15}$$

with

$$\begin{aligned} j &= R^\dagger C_n^{-1} d, \\ C_{s|d}^{-1} &= R^\dagger C_n^{-1} R + C_s^{-1}, \end{aligned} \tag{5.16}$$

and the signal and noise covariances C_s and C_n, respectively. Therefore, to calculate the Bayesian model evidence, one often[3] has to calculate determinants of covariance matri-

[3]By the word "often" we refer to cases, in which at least one marginalization [see Eq. (5.17)] can be performed analytically (approximated with high precision) to obtain a model-dependent determinant.

ces. This might be done by probing [Eq. (5.8)] if dealing with implicit matrices [last line of Eq. (5.15)] instead of performing the multidimensional integral [second last line in Eq. (5.15)] numerically. The latter has been done, for instance, in the field of inflationary cosmology [38, 120] by the method of nested sampling [121, 122].

This is especially of importance in the field of model selection or comparison [119], where from an observation – the data – one wants to infer which theory reproduces the observation best. Switching from one model to another means, for instance[4], to exchange R in Eq. (5.15), which directly affects the determinant containing $C_{s|d}$. Thus, the calculation of the determinant is mandatory here.

5.3.2. Posterior distribution including marginalizations

In the field of signal inference one is typically interested in reconstructing a set of i parameters p^i with uncertainty from some observation, the data d. This information is delivered by the posterior, given by[5] [26] $\mathcal{P}(p^i|d) \propto \mathcal{P}(d|p^i)\mathcal{P}(p^i)$. Often, however, this inference problem is degenerate, caused by a so-called nuisance parameter. For example, consider the calibration of an instrument is of interest and not the signal. In this case the signal s represents the nuisance parameter. The common procedure to circumvent this problem is to marginalize over these parameters,

$$\mathcal{P}(p^i|d) \propto \int \mathcal{D}s \int \mathcal{D}n \; \mathcal{P}(d, s, n|p^i)\mathcal{P}(p^i). \tag{5.17}$$

To continue with the simple example of Sec. 5.3.1 we assume again Gaussian distributions for s and n and a linear measurement but with explicit dependency on p^i, i.e., $d = (Rs)[p^i] + n$. If we further follow the example of calibration, the parameter p^i might be a calibration coefficient, thus affecting only R. This yields $(Rs)[p^i] = R[p^i]s$ and therefore

$$\mathcal{P}(p^i|d) \propto \left\{ \int \mathcal{D}s \; \mathcal{G}\left(d - R[p^i]s, C_n\right) \mathcal{G}(s, C_s) \right\} \mathcal{P}(p^i). \tag{5.18}$$

This integration can be performed analytically, producing an in general non-Gaussian probability distribution with p^i-dependent normalization (and exponent) similar to Eq. (5.15),

$$\mathcal{P}(p^i|d) \propto \sqrt{\left|2\pi C_{s|d}[p^i]\right|} \; \mathcal{P}(p^i) \exp\left\{ \frac{1}{2} j^\dagger[p^i] C_{s|d}[p^i] j[p^i] \right\} \tag{5.19}$$

with $C_{s|d}[p^i]$ and $j[p^i]$ now containing $R[p^i]$ instead of R. In case the covariance matrices or $R[p^i]$ are only given by a computer routine (implicit representation of a matrix) one could use Eq. (5.8) to probe the determinant.

A variety of scientific fields are affected by this problem. For example, the extraction of the level of non-Gaussianity of the cosmic microwave background [31, 53] in cosmology, the problem of self-calibration [22, 28, 123] in general, or lensing in astronomy [124].

5.3.3. Realistic astronomical example

In order to study a more realistic example we consider a measurement device with spatially constant but unknown calibration amplitude, parametrized by $1+\gamma \in \mathbb{R}$, scanning a specific

[4] We focus here on R for simplicity only. One could also, additionally, exchange the prior covariances C_n and C_s, the assumed prior statistics, the parametrization of the data, and so on.

[5] Note that in this case the evidence is just a scalar which normalizes the posterior, therefore we merely state proportionalities.

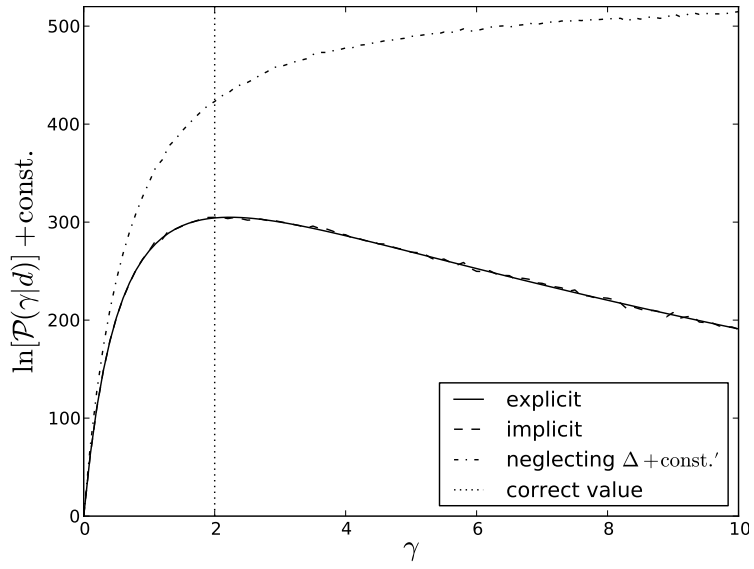


Figure 5.5.: Logarithmic posterior of the calibration amplitude parameter γ using implicit and explicit representations of the involved operators, see Eq. (5.21) and Eq. (5.22) for details. The abbreviation Δ denotes the logarithm of the term given by Eq. (5.22).

patch of the sky. The measured and assumed to be Gaussian sky signal s is affected by the instrument via a convolution $\mathcal{C}$ with a Gaussian kernel of standard deviation $\sigma = 0.05$. Additionally, the observation might be disturbed by fore- and backgrounds. For this reason we include an observational mask M_o, which cuts out 20% of the sky. The noise n is still assumed to be Gaussian and uncorrelated with the signal. Hence, the measurement equation is given by

$$d = R\,[\gamma]\,s + n = (1+\gamma) M_o \mathcal{C} s + n. \tag{5.20}$$

To calibrate the measurement device the calibration posterior $\mathcal{P}(\gamma|d)$ has to be determined. The resulting calibration mean $\langle\gamma\rangle_{\mathcal{P}(\gamma|d)}$ can be regarded as an external calibration if the *a priori* knowledge on the signal is sufficiently strong. Otherwise one could infer the signal and calibration amplitude γ simultaneously from data using iterative approaches [123]. Using Eq. (5.19) as well as a flat prior on γ we obtain

$$\ln \mathcal{P}(\gamma|d) = -\frac{1}{2} \ln \left| C^{-1}_{s|d}\,[\gamma] \right| + \frac{1}{2} j^\dagger\,[\gamma]\, C_{s|d}\,[\gamma]\, j\,[\gamma] + \text{const.}, \tag{5.21}$$

which exhibits in particular the γ-dependent determinant

$$\left| C^{-1}_{s|d}\,[\gamma] \right| = \left| (1+\gamma)\; \mathcal{C}^\dagger M_o^\dagger C_n^{-1} M_o \mathcal{C} (1+\gamma) + C_s^{-1} \right|. \tag{5.22}$$

For the numerical evaluation of Eq. (5.21) we use the settings of Sec. 5.2.2 with $C_s(k,k') = (1+k)^{-3}\,\delta_{kk'}$, a calibration amplitude parameter of $\gamma = 2$, and a noise covariance of $(C_n)_{x,x'} = 10^{-1}\delta_{xx'}$ to generate a data realization. The pseudotime interval has been discretized into 10^2 parts. The numerically determined calibration posterior for a given data realization can be found in Fig. 5.5, which demonstrates again the efficiency of the stochastic method using only eight probes for a single trace probing operation. The figure also illustrates the impact of the determinant on the log-posterior, which would not peak in the shown interval without it.

5.4. Summary

Motivated by the problem of finding a way to efficiently determine the determinant of an implicitly defined matrix or operator, we derived a formula, Eq. (5.8), representing a stochastic estimate of its log-determinant. This has been achieved by reformulating the log-determinant as an integral representation and transforming the involved terms into stochastic expressions, which includes a numerical integration and a trace probing. Numerical examples have shown that the discretization of the integration interval may be very coarse in case the probed operator is sufficiently diagonal. In case it exhibits a significant nondiagonal structure one has to fine-grain the discretization of this interval. The number of probes necessary for the trace probing, however, remains very low in the studied examples. These facts combined with the almost complete parallelizability of this approach might keep the computational costs within reasonable limits in many situations.

This method clearly has more general applications but might in particular be useful for Bayesian signal inference and model comparison when dealing with large data sets as often given, for instance, in astronomy and cosmology. To be precise, it might be beneficial in all fields where the numerical calculation of a determinant of an operator is mandatory.

6. Conclusion

6.1. Summary

This thesis bridges cosmology and information theory. For studying the early Universe we have developed a method to reconstruct inflationary model parameters by detecting deviations from primordial Gaussianiy and reconstructed the primordial scalar potential including uncertainty from CMB data. In order to advance the mathematical tools of information retrieval, needed for our cosmological but also for other studies, we have developed further the theory of self-calibration and have presented a novel technique to sample a determinant of an implicitly defined matrix.

The first part of this thesis deals with the advancement of the theory of self-calibration, since it involves a detailed introduction to information field theory. Based on the latter we developed the Calibration-Uncertainty Renormalized Estimator (CURE), to infer a signal and its unknown calibration from the same data. For this purpose we assumed the signal and calibration covariance structure to be known and successively included more and more portions of calibration uncertainty into the signal inference equations. The final inference equations for the renormalized signal and calibration, which are coupled differential equations, absorbed these uncertainty corrections. CURE is able to keep up in accuracy, but not yet in performance, with the best self-calibration methods currently available.

In the following, we introduced a novel approach to reconstruct inflationary model parameters by exploiting higher-order statistics of the curvature perturbations. This reconstruction has been achieved by a saddle-point approximation of the posterior probability. In this approximative way we preserved the analyticity of the problem as far as possible and therefore circumvented expensive sampling techniques. Furthermore, we described how to infer the primordial power spectrum from the curvature perturbations.

Subsequently, a successful reconstruction of the primordial scalar potential has been presented. The development of a method that splits this large inverse problem of the reconstruction into many, each of them solved by a Wiener filter, enabled us to infer the primordial scalar potential in the vicinity of the last scattering surface. This method has been applied to the CMB data of the WMAP satellite and has led to a couple of future research opportunities, see Sec. 6.2.

Finally, we developed a method to sample the determinant of an implicitly defined matrix. This sampling method requires relatively few probes, is based on an integral representation of the log-determinant of a matrix, and can be used, e.g., in the field of Bayesian model comparison or posterior calculations.

6.2. Outlook

The respective inference methods presented in this thesis enable subsequent research possibilities, which are listed in the following:

- **Calibration-uncertainty renormalized estimator.** Having developed the basic methodology of the CURE, the next logical step should be its preparation for realistic measurement situations. This would require the usage of a sparse representation of the occurring tensors of higher order, since these tensors are computationally expensive to handle. As a consequence CURE could be applied to larger data sets. Finally one could try to extend the CURE-method to non-Gaussian signals, e.g., log-normal fields. The latter development would be especially interesting for the ground-based radio astronomy, where sky intensity can be modeled as a log-normal field and where a reliable self-calibration is needed (due to, e.g., effects of the Earth's atmosphere).

- **Inference of inflation models by non-Gaussianity.** The introduction of the posterior for inflationary parameters, mapping constraints on non-Gaussianity onto constraints on inflation, has been done in a generic manner. It would be straightforward to extend the approach to many other inflationary models that have not been ruled out yet by observational data. Unfortunately, as it turned out, the level of local non-Gaussianity of the primordial curvature perturbations is tiny. Therefore the real application of the developed method to CMB data seems to be unnecessarily complex in comparison with other approximative approaches like the KSW estimator [52]. A further development for the application to large-scale structure observations might be promising, though.

- **Primordial scalar potential and power spectrum**. The next logical step is the reconstruction of the primordial scalar potential and its power spectrum from Planck data including polarization whilst taking into account all frequency bands. This should be done by using the critical filter formula, Eq. (4.18), for the reasons already pointed out in Sec. 4.2.3. Afterwards, there are a couple of possible directions for future research, e.g.,
  - cross-checks with reconstructions of the initial conditions from large-scale structure (see, e.g., Ref. [94]),
  - inference of inflation or reheating parameters,
  - symmetry/morphology investigations,
  - cross-correlation with other results, e.g., the CMB lensing potential,
  - inference of primordial magnetic fields generated during inflation/reheating [95],
  - etc.

  Furthermore, there are several routes in which the method can be improved further. One could think of incorporating the delensing operator into the response operation, the possibility of incorporating the isocurvature modes, or to infer the primordial scalar potential directly from time-ordered data. The latter could be done, for instance, by using a modified version of the already existing deconvolution code ArtDeco [125].

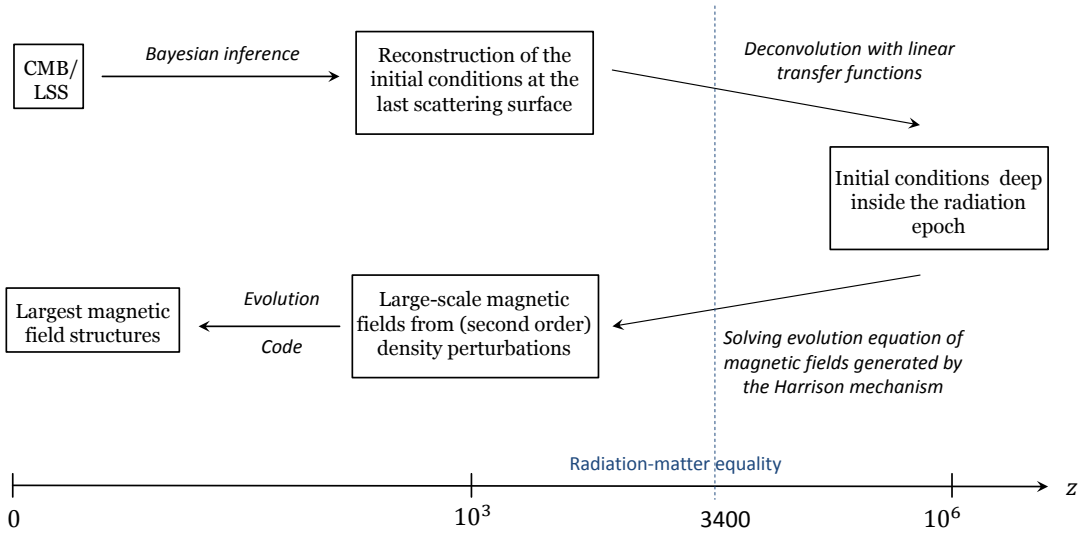


Figure 6.1.: Sketch of a method to calculate the primordial magnetic field generated by the Harrison mechanism.

- **Large-scale magnetic fields from cosmological initial conditions**. The inferred primordial scalar potential can also be used to calculate the minimal amount of large-scale magnetic fields generated mainly during the radiation dominated epoch by the so-called Harrison mechanism [126].

  The Harrison mechanism, summarized in Ref. [96], describes the production of magnetic fields in rotating regions (i.e. second order cosmological perturbation theory; angular velocity ω) within the expanding Universe. It considers the period in radiation domination, where the constituents[1] of the Universe, i.e. protons p, electrons e, and photons γ are tightly coupled by Thompson and Coulomb scattering as well as interactions via lepton-pair production [96]. During this period the Universe can be described by a single fluid. As the Universe expanded and cooled down the strength of Coulomb scattering was not able to keep up with the one of Thomson scattering. Therefore the angular velocity of the fluid within rotating regions splits up into ω_p and ω_e with $\omega_p < \omega_e$. During the expansion of the Universe (scale factor a) the angular velocities of the components scale with $\omega_p \propto 1/a^2$ and $\omega_e \propto 1/a$, respectively. The relative motion between these two rotating components produced non-vanishing currents and thus magnetic fields. This process, however, ended when electrons and protons combined to neutral hydrogen and photons decoupled. A mathematical description of the magnetic field generation using a simplified version[2] (mainly neglecting the anisotropic stress tensor) of this process can be found in Ref. [96] and

[1] We neglect here, for simplicity only, interactions with further components of the Universe like neutrinos, dark matter particles, etc.

[2] For a strict derivation of the magnetic fields produced by the Harrison mechanism see Ref. [127].

yields the spatial ($\vec{x}$) and redshift (z) dependent magnetic field

$$\vec{B}(\vec{x},\tau) = -\frac{m_p(1+z)}{e\mathcal{H}}\left(\nabla\Phi'\times\nabla\Phi - \frac{1}{12\mathcal{H}}\nabla\Phi\times\nabla(\Delta\Phi) - \frac{1}{12\mathcal{H}^2}\nabla\Phi'\times\nabla(\Delta\Phi)\right), \tag{6.1}$$

with $\mathcal{H}$ the Hubble constant in conformal time τ. The primordial scalar potential Φ and its derivative with respect to conformal time, denoted by apostrophe ($'$), are defined by

$$\Phi(\vec{k},\tau) = \frac{3j_1(k\tau/\sqrt{3})}{k\tau/\sqrt{3}}\Phi_0(\vec{k}), \tag{6.2}$$

where the function $j_1(\cdot)$ denotes the spherical bessel function of the first kind. The initial $\Phi_0(\vec{k})$ represents the primordial scalar potential at an early point in the radiation epoch, where all relevant scales are well outside the horizon. It can be calculated from a given reconstruction of the primordial scalar potential or matter density contrast at the surface of last scattering. Applying Eq. (6.1) subsequently yields the early-time large-scale magnetic fields, whose late-time evolution (between the last scattering and today) might be simulated in the following. The logical steps are illustrated in Fig. 6.1.

The implementation of this research has already started. However, due to the cumulative nature of this thesis, which requires accepted reviewed publications as its basis, it is excluded from its content.

A. Appendix Chapter 2

Note: This appendix has been published in Phys. Rev. E **91**, 013311 (2015) [22].

A.1. Feynman rules

The Feynman rules originally stated in and inherited from Ref. [25] read as follows:

1. Open ends of lines in diagrams correspond to external coordinates and are labeled by such. Since the partition sum in particular does not depend on any external coordinate, it is calculated only from summing up closed diagrams. However, the field expectation value $m(x) = \langle s(x) \rangle_{(s|d)} = \delta \ln[\mathcal{Z}(d, J)]/\delta J(x)|_{J=0}$ and higher order correlation functions depend on coordinates and therefore are calculated from diagrams with one or more open ends, respectively.

2. A line with coordinates x' and y' at its end represents the propagator $D_{x'y'}$ connecting these locations.

3. Vertices with one leg get an individual internal, integrated coordinate x' and represent the term $j_{x'} + J_{x'} - \Lambda^{(1)}_{x'}$.

4. Vertices with n legs represent the term $-\Lambda^{(n)}_{x'_1 \dots x'_n}$, where each individual leg is labeled by one of the internal coordinates $x'_1 \dots x'_n$. This more complex vertex-structure, as compared to QFT, is a consequence of non-locality in IFT.

5. All internal (and therefore repeatedly occurring) coordinates are integrated over, whereas external coordinates are not.

6. Every diagram is divided by its symmetry factor, the number of permutations of vertex legs leaving the topology invariant, as described in any book on field theory.

A.2. Renormalization flow equations including absolute calibration measurements

This section derives the generalization of the renormalization flow equations in presence of absolute calibration measurements. These measurements can be included in the prior knowledge of the calibration coefficients, $\mathcal{P}(\gamma) = \mathcal{G}(\gamma - m_\gamma, D^\gamma)$, with $(m_\gamma)_a$ the Wiener filter solution for γ with uncertainty D^γ using the absolute calibration measurements only. Hence, the likelihood becomes

$$\begin{aligned} \mathcal{P}(d|s) &= \int \mathcal{D}\gamma \; \mathcal{P}(d|s,\gamma)\mathcal{G}(\gamma - m_\gamma, D^\gamma) = \mathcal{G}\left(d - \check{R}s, N + \sum_{ab} D^\gamma_{ab} R^a s s^\dagger R^{b\dagger}\right), \\ \check{R} &\equiv R^0 + \sum_a (m_\gamma)_a R^a. \end{aligned} \tag{A.1}$$

Compared to the result without measurements of absolute calibration, Eq. (2.23), the response R and the calibration covariance have been replaced by $\check{R}$ and D^γ, respectively. This means the response became modified by new, additional, information from the absolute calibration measurements and associated the uncertainty D^γ, which is not diagonal anymore. The resulting reference field expansion of the Hamiltonian, Eq. (2.34), yields the following assignments:

$$
\begin{aligned}
\check{D} =& \left(S^{-1} + \check{R}^\dagger N^{-1} \check{R}\right)^{-1}, \quad \check{j} = \check{R}^\dagger N^{-1} d, \quad \check{m} = \check{D}\check{j}, \quad M^{\check{}x} \equiv \check{R}^\dagger N^{-1} R^x, \\
\Lambda^{(1)}\phi =& \frac{1}{\delta t} \underbrace{\left(\check{m}^\dagger \check{D}^{-1} - \check{j}^\dagger\right)}_{=0} \phi + \sum_{ab} D^\gamma_{ab} \Big\{ \check{m}^\dagger M^{ab} \phi - \frac{1}{2} j^{a\dagger} \left(\phi \check{m}^\dagger + \check{m}\phi^\dagger\right) j^b \\
& - \frac{1}{2} \check{m}^\dagger M^{\check{}a} \left(\phi \check{m}^\dagger + \check{m}\phi^\dagger\right) M^{b\check{}} \check{m} - \frac{1}{2} \check{m}^\dagger M^{\check{}a} \check{m} \check{m}^\dagger M^{b\check{}} \phi + j^{a\dagger} \check{m} \check{m}^\dagger M^{b\check{}} \phi \\
& + \frac{1}{2} j^{a\dagger} \left(\phi \check{m}^\dagger + \check{m}\phi^\dagger\right) M^{b\check{}} \check{m} + \frac{1}{2} \check{m}^\dagger M^{\check{}a} \left(\phi \check{m}^\dagger + \check{m}\phi^\dagger\right) j^b \Big\}, \\
\Lambda^{(2)}[\phi,\phi] =& \frac{1}{2} \sum_{ab} D^\gamma_{ab} \Big\{ \phi^\dagger M^{ab} \phi - j^{a\dagger} \phi\phi^\dagger j^b - \phi^\dagger M^{\check{}a} \check{m} \check{m}^\dagger M^{b\check{}} \phi - \check{m}^\dagger M^{\check{}a} \phi\phi^\dagger M^{b\check{}} \check{m} \\
& - \phi^\dagger M^{\check{}a} \left(\phi \check{m}^\dagger + \check{m}\phi^\dagger\right) M^{b\check{}} \check{m} - \check{m}^\dagger M^{\check{}a} \left(\phi \check{m}^\dagger + \check{m}\phi^\dagger\right) M^{b\check{}} \phi \\
& + j^{a\dagger} \left(\phi \check{m}^\dagger + \check{m}\phi^\dagger\right) M^{b\check{}} \phi + \phi^\dagger M^{\check{}a} \left(\phi \check{m}^\dagger + \check{m}\phi^\dagger\right) j^b + j^{a\dagger} \phi\phi^\dagger M^{b\check{}} \check{m} \\
& + \check{m}^\dagger M^{\check{}a} \phi\phi^\dagger j^b \Big\} + 1 \text{ perm.}, \\
\Lambda^{(3)}[\phi,\phi,\phi] =& -\sum_{ab} D^\gamma_{ab} \Big\{ \frac{1}{2} \phi^\dagger M^{\check{}a} \left(\phi \check{m}^\dagger + \check{m}\phi^\dagger\right) M^{b\check{}} \phi + \frac{1}{2} \check{m}^\dagger M^{\check{}a} \phi\phi^\dagger M^{b\check{}} \phi \\
& + \frac{1}{2} \phi^\dagger M^{\check{}a} \phi\phi^\dagger M^{b\check{}} \check{m} - \frac{1}{2} j^{a\dagger} \phi\phi^\dagger M^{b\check{}} \phi - \frac{1}{2} \phi^\dagger M^{\check{}a} \phi\phi^\dagger j^b \Big\} + 5 \text{ perm.}, \\
\Lambda^{(4)}[\phi,\phi,\phi,\phi] =& -\frac{1}{2} \sum_{ab} D^\gamma_{ab} \phi^\dagger M^{\check{}a} \phi\phi^\dagger M^{b\check{}} \phi + 23 \text{ perm.}.
\end{aligned}
\tag{A.2}
$$

B. Appendix Chapter 3

Note: This appendix has been published in JCAP **6**, 048 (2014) [31].

B.1. Shape of posterior and estimator of inflationary parameters p

In general, the posterior distribution for p does not have to be Gaussian. If so, one should be very careful if one compares the posterior pdf for p with an estimator pdf, $P(\hat{p})$, because they can exhibit different types of deviations from Gaussianity. This means in particular that in some cases an unbiased constructed estimator might exhibit a skewness behavior different from the posteriors one. For instance, compare the posterior pdf in Ref. [53] with the estimator pdf in Ref. [48], where the pdf is negatively skewed in one case and positively in another. The reason for this apparent contradiction is illustrated in Fig. B.1, where the joint probability of p and d is shown. To determine the posterior pdf we consider a varying p given d, which is a one-dimensional hypersurface, parallel to the horizontal axis. If one wants to obtain the pdf for the estimator, one has to vary d given p corresponding to a one-dimensional hypersurface parallel to the vertical axis. If the probability distribution is symmetric, e.g. Gaussian, the shapes of estimator pdf and posterior pdf coincide. However, if the distribution is asymmetric the shapes do not have to coincide as sketched in Fig. B.1 for a one-dimensional parameter p. In turn, this means that the shapes of posterior and the distribution of an estimator of a quantity do not have to exhibit the same skewness. Hence there is no real contradiction. A discussion about other advantages and disadvantages of the usage of posterior distributions can be found, e.g., in Ref. [54].

B.2. Dependency of the scalar amplitude and spectral index on f_{NL} under usage of the critical filter

In Sec. 3.5.3 a relation between the scalar amplitude A_s (spectral index n_s) of the primordial power spectrum and the non-Gaussianity parameter f_{NL} was mentioned, if one uses the critical filter with or without smoothness prior to reconstruct the power spectrum from a field that is (in some cases) falsely assumed to be Gaussian. In turn, that means by applying these filter formulae, Eqs. (3.55) and (3.58), one could infer the level of non-Gaussianity of a field.

Within Sec. 3.5.3 we did this calculation for the curvaton scenario. However, we can transform the dependency of A_s, n_s on κ into a dependency on f_{NL} by neglecting contributions of the trispectrum. Solving Eq. (3.30) for κ and substituting it within the quadratic and linear fitting formula, pointed out in Sec. 3.5.3, yields

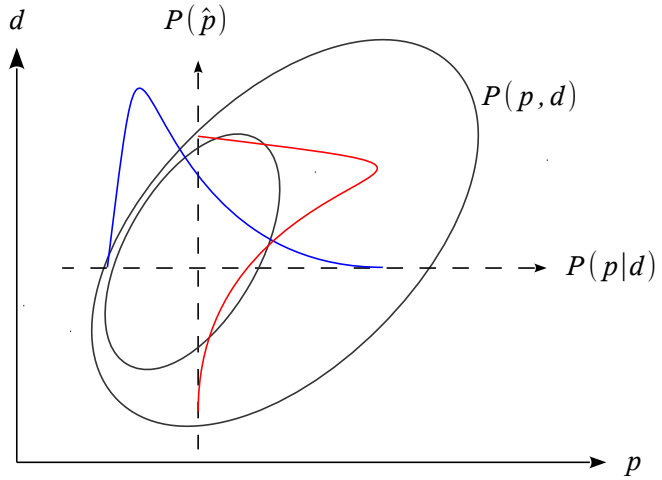


Figure B.1.: Sketch of posterior (blue) and estimator (red) pdf with different skewness behaviors drawn from an asymmetric joint probability (black).

$$10^9 A_s(f_{\rm NL}) = \frac{a_q}{125}\left(10 + 6f_{\rm NL} + \sqrt{2}\sqrt{18f_{\rm NL}^2 + 60f_{\rm NL} + 125}\right)^2 + \frac{b_q}{15}\left(10 + 6f_{\rm NL} + \sqrt{2}\sqrt{18f_{\rm NL}^2 + 60f_{\rm NL} + 125}\right) + c_q, \tag{B.1}$$

and

$$n_s(f_{\rm NL}) - 1 = \frac{a_{\rm l}}{15}\left(10 + 6f_{\rm NL} + \sqrt{2}\sqrt{18f_{\rm NL}^2 + 60f_{\rm NL} + 125}\right) + b_{\rm l}, \tag{B.2}$$

for

$$f_{\rm NL} \geq -\frac{5}{4}, \tag{B.3}$$

with a_q, b_q, c_q, $a_{\rm l}$, $b_{\rm l}$ fitting parameters (see Sec. 3.5.3) that might depend on the number of pixels used. Note that relations between the spectral index and the scalar amplitude of the primordial power spectrum can also be derived for other inflation models, e.g., the modulated Higgs inflation scenario of Sec. 3.3.2.

Analogously one could solve the linear fitting formula of the spectral index for κ and substitute the latter within the quadratic fitting formula to derive a relation between A_s and n_s. This yields

$$10^9 A_s(n_s) = \frac{a_q}{a_{\rm l}^2}\left(n_s - 1 - b_{\rm l}\right)^2 + \frac{b_q}{a_{\rm l}}\left(n_s - 1 - b_{\rm l}\right) + c_q, \tag{B.4}$$

whereby we have not neglected the contributions of the trispectrum.

C. Appendix Chapter 4

Note: This appendix has been published in JCAP **2**, 041 (2015) [93].

C.1. Response projected onto the sphere of LSS

The data is given by

$$\begin{aligned} d_{\ell m} &\equiv M_{\substack{\ell m \\ \ell' m'}} a^{\mathrm{CMB}}_{\ell' m'} + n_{\ell m} = (R\Phi)_{\ell m} + n_{\ell m} \\ &= M_{\substack{\ell m \\ \ell' m'}} B_{\ell'} \frac{2}{\pi} \int dk\ k^2 \int dr\ r^2 \Phi_{\ell' m'}(r) g^T_{\ell'}(k) j_{\ell'}(kr) + n_{\ell m}. \end{aligned} \tag{C.1}$$

Considering Gaussian statistics for the primordial curvature perturbations, Φ, the response is defined by

$$R \equiv \left\langle d\Phi^\dagger \right\rangle_{(\Phi,d)} \left\langle \Phi\Phi^\dagger \right\rangle^{-1}_{(\Phi,d)}. \tag{C.2}$$

Instead of using the full three-dimensional response R, we introduce a two-dimensional response, $R^{(2)}$, which acts on the primordial potential projected onto the last scattering surface (LSS), $\Phi^{(2)} \equiv \Phi\left(r = r_{\mathrm{LSS}}\right) = \tilde{T}\Phi$, where $\tilde{T}$ denotes the projection operator:

$$R^{(2)} = \left\langle R\Phi(\tilde{T}\Phi)^\dagger \right\rangle_{(\Phi,d)} \left\langle \tilde{T}\Phi(\tilde{T}\Phi)^\dagger \right\rangle^{-1}_{(\Phi,d)} = \left(RP^\Phi\tilde{T}^\dagger\right)\left(\tilde{T}P^\Phi\tilde{T}^\dagger\right)^{-1}. \tag{C.3}$$

To derive the denominator at the distance of the LSS, we first transform it into position-space,

$$\begin{aligned} \left(\tilde{T}P^\Phi\tilde{T}^\dagger\right)_{\hat{\mathbf{n}},\hat{\mathbf{n}}'} &= \int d^3\mathbf{x} \int d^3\mathbf{y} \delta\left(\mathbf{x} - r_{\mathrm{LSS}}\hat{\mathbf{n}}\right) \delta\left(\mathbf{y} - r_{\mathrm{LSS}}\hat{\mathbf{n}}'\right) \\ &\quad \times \int \frac{d^3\mathbf{k}}{(2\pi)^3} \int \frac{d^3\mathbf{q}}{(2\pi)^3} (2\pi)^3 \delta(\mathbf{k}-\mathbf{q}) P^\Phi(k) e^{-i\mathbf{k}\cdot\mathbf{x}} e^{i\mathbf{q}\cdot\mathbf{y}} \\ &= \int \frac{d^3\mathbf{k}}{(2\pi)^3} P^\Phi(k) e^{-ir_{\mathrm{LSS}}\mathbf{k}\cdot\hat{\mathbf{n}}} e^{ir_{\mathrm{LSS}}\mathbf{k}\cdot\hat{\mathbf{n}}'}. \end{aligned} \tag{C.4}$$

Vectors are printed in bold for reasons of clarity and comprehensibility; unit vectors are denoted by $\hat{}$. Subsequently we use the Rayleigh expansion,

$$e^{i\mathbf{k}\cdot\mathbf{r}} = 4\pi \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} i^\ell j_\ell(kr) Y_\ell^{m*}(\hat{\mathbf{k}}) Y_\ell^m(\hat{\mathbf{r}}), \tag{C.5}$$

as well as the transformation rules

$$\begin{aligned} f_{\ell m} &\equiv \oint d\hat{\mathbf{n}}\ Y_\ell^{m*}(\hat{\mathbf{n}}) f(\hat{\mathbf{n}}), \\ f(\hat{\mathbf{n}}) &= \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} f_{\ell m} Y_\ell^m(\hat{\mathbf{n}}), \\ \text{and } &\oint d\hat{\mathbf{n}}\ Y_\ell^m(\hat{\mathbf{n}}) Y_{\ell'}^{m'*}(\hat{\mathbf{n}}) = \delta_{\ell\ell'}\delta_{mm'}, \end{aligned} \tag{C.6}$$

to obtain the final corresponding expression in the spherical harmonic space,

$$\begin{aligned}\left(\tilde{T}P^{\Phi}\tilde{T}^{\dagger}\right)_{\substack{\ell m\\ \ell' m'}} &= \oint d\hat{\mathbf{n}} \oint d\hat{\mathbf{n}}'\; Y_{\ell}^{m}(\hat{\mathbf{n}}) Y_{\ell'}^{m'*}(\hat{\mathbf{n}}') \int \frac{d^3\mathbf{k}}{(2\pi)^3} P^{\Phi}(k) \\ &\quad \times \sum_{\substack{\ell''\ell'''\\ m''m'''}} (4\pi)^2 i^{\ell'''-\ell''} j_{\ell''}(kr_{\mathrm{LSS}}) j_{\ell'''}(kr_{\mathrm{LSS}}) \\ &\quad \times Y_{\ell''}^{m''}(\hat{\mathbf{k}}) Y_{\ell'''}^{m'''*}(\hat{\mathbf{k}}) Y_{\ell''}^{m''*}(\hat{\mathbf{n}}) Y_{\ell'''}^{m'''}(\hat{\mathbf{n}}') \\ &= \frac{2}{\pi} \int dk\; k^2 \oint d\hat{\mathbf{k}}\; P^{\Phi}(k) i^{\ell'-\ell} j_{\ell}(kr_{\mathrm{LSS}}) j_{\ell'}(kr_{\mathrm{LSS}}) Y_{\ell'}^{m'*}(\hat{\mathbf{k}}) Y_{\ell}^{m}(\hat{\mathbf{k}}) \\ &= \frac{2}{\pi} \int dk\; k^2 P^{\Phi}(k) j_{\ell}^2(kr_{\mathrm{LSS}}) \delta_{\ell\ell'}\delta_{mm'} \equiv P_{\ell}^{\Phi}\delta_{\ell\ell'}\delta_{mm'}.\end{aligned} \tag{C.7}$$

P_{ℓ}^{Φ} denotes the primordial power spectrum projected onto the sphere of LSS.

To determine the numerator we fist have to transform $P^{\Phi}\tilde{T}^{\dagger}$ into the basis of spherical harmonics. Analogous to the calculation above we obtain

$$\left(P^{\Phi}\tilde{T}^{\dagger}\right)_{\substack{\ell m\\ \ell' m'}}(r) = \frac{2}{\pi} \int dk\; k^2 P^{\Phi}(k) j_{\ell}(kr_{\mathrm{LSS}}) j_{\ell}(kr) \delta_{\ell\ell'}\delta_{mm'}, \tag{C.8}$$

and thus

$$\begin{aligned}\left(RP^{\Phi}\tilde{T}^{\dagger}\right)_{\substack{\ell m\\ \ell' m'}} &= M_{\substack{\ell m\\ \ell'' m''}} B_{\ell''} \frac{2}{\pi} \\ &\quad \times \int dk\; k^2 \int dr\; r^2 \left\{ \frac{2}{\pi} \int dk' k'^2 P^{\Phi}(k') j_{\ell''}(k' r_{\mathrm{LSS}}) j_{\ell''}(k'r) \right\} \\ &\quad \times g_{\ell''}^{T}(k) j_{\ell''}(kr) \delta_{\ell''\ell'}\delta_{m''m'}.\end{aligned} \tag{C.9}$$

Using the identity

$$\int_0^{\infty} dr\; r^2 j_{\ell}(kr) j_{\ell}(k'r) = \frac{\pi}{2}\frac{1}{k^2}\delta(k-k') \tag{C.10}$$

finally yields

$$\left(RP^{\Phi}\tilde{T}^{\dagger}\right)_{\substack{\ell m\\ \ell' m'}} = M_{\substack{\ell m\\ \ell'' m''}} B_{\ell''} \frac{2}{\pi} \int dk\; k^2 P^{\Phi}(k) j_{\ell''}(kr_{\mathrm{LSS}}) g_{\ell''}^{T}(k) \delta_{\ell''\ell'}\delta_{m''m'}. \tag{C.11}$$

Putting the results together, the two-dimensional response is given by

$$R^{(2)}_{\substack{\ell m\\ \ell' m'}} = M_{\substack{\ell m\\ \ell'' m''}} B_{\ell''} \frac{\int dk\; k^2 P^{\Phi}(k) j_{\ell''}(kr_{\mathrm{LSS}}) g_{\ell''}^{T}(k)}{\int dk\; k^2 P^{\Phi}(k) j_{\ell''}^2(kr_{\mathrm{LSS}})} \delta_{\ell''\ell'}\delta_{m''m'}. \tag{C.12}$$

The response for arbitrary comoving distances r' can be obtained by replacing r_{LSS} by r'.

C.2. Wiener filter formula and uncertainty estimate in data space

The Wiener filter in data space is defined by

$$\begin{aligned} m_{\mathrm{w}}^{(2)} &\equiv \left\langle \Phi^{(2)} \right\rangle_{(\Phi|d)} = \tilde{T} \left\langle \Phi \right\rangle_{(\Phi|d)} = \tilde{T} \left\langle \Phi d^{\dagger} \right\rangle_{(\Phi,n)} \left\langle dd^{\dagger} \right\rangle_{(\Phi,n)}^{-1} d \\ &= \tilde{T} P^{\Phi} R^{\dagger} \left[RP^{\Phi}R^{\dagger} + N \right]^{-1} d = \tilde{T} P^{\Phi} R^{\dagger} \left[\tilde{C}^{TT} + N \right]^{-1} d \\ &\overset{\text{Eq. (C.12)}}{=} P_{\ell}^{\Phi} R^{(2)\dagger} \left[\tilde{C}^{TT} + N \right]^{-1} d. \end{aligned} \tag{C.13}$$

Formally, the corresponding posterior covariance matrix is constructed as

$$D = P_\ell^\Phi - P_\ell^\Phi R^{(2)\dagger} \left(\tilde{C}^{TT} + N\right)^{-1} R^{(2)} P_\ell^\Phi. \tag{C.14}$$

The square root of its position space diagonal would give us the 1σ uncertainty map. However, as the operator is not directly accessible to us, but is only defined as a sequence of linear functions, calculating the diagonal requires very expensive probing routines which need to evaluate the covariance matrix several thousand times before converging.

However, the covariance matrix becomes diagonal in spherical harmonic space under two conditions[1]: We assume that there is no masking in the data and the noise covariance N is a multiple of the identity. The noise covariance matrix for TT data is already diagonal and dominated by white uncorrelated noise. So this approximation seems appropriate given the benefits in computational costs. The assumption that there is no masking is more drastic of course. We therefore construct our uncertainty map out of the limiting cases of having no masking and masking the whole sky. Both scenarios make the posterior covariance matrix diagonal in spherical harmonic space.

The constant approximation to the noise covariance is constructed as

$$\tilde{N}_{\hat{n}\hat{n}'} = \frac{\operatorname{tr} N}{\operatorname{tr} \mathbb{1}} \delta(\hat{n} - \hat{n}'). \tag{C.15}$$

The response with no mask is diagonal in spherical harmonic space,

$$\tilde{R}_{\substack{\ell m \\ \ell' m'}} = B_\ell R_\ell \, \delta_{\ell\ell'} \, \delta_{mm'}, \tag{C.16}$$

and the response with an all-sky mask is zero. Therefore the covariance matrix is diagonal in either case. Since a diagonal matrix in spherical harmonic space results in a constant diagonal in position space, we can exploit the invariance of the trace to get the position space diagonal of the covariance matrix,

$$D_{\hat{n}\hat{n}} = \frac{\operatorname{tr} D}{4\pi}, \tag{C.17}$$

where the trace is easily calculated in spherical harmonic space, where D is diagonal.

In a region that is fully masked and where the edges of the mask are further away than the correlation length of Φ the uncertainty approaches the limiting case of an all-sky mask. In a region that is fully exposed and more than a correlation length away from a masked region the uncertainty approaches the limiting case of no mask. We therefore combine the two cases into one map by setting the uncertainty to the "all-sky masked" value in regions which are masked and to the "no mask" value in regions which are not masked, i.e.

$$\sigma_{\hat{n}}^2 = \begin{cases} D_{\hat{n}\hat{n}}^{\text{all mask}} & \text{if } M_{\hat{n}\hat{n}'} = 0 \\ D_{\hat{n}\hat{n}}^{\text{no mask}} & \text{otherwise.} \end{cases} \tag{C.18}$$

The interpolation between these two regions is dictated by the prior covariance. It describes precisely how information is correlated between masked and unmasked regions. Our final uncertainty map is therefore the result of a smoothing of σ with the normalized square root of the prior covariance,

$$\sigma_{\text{smooth}} = \frac{1}{\mathcal{N}} \sqrt{P_\ell^\Phi} \sigma, \tag{C.19}$$

[1]Note that this procedure is only valid for a temperature-only analysis. Once polarization data is included the 1σ uncertainty must be calculated by the square root of the diagonal of Eq. (C.14).

where

$$\mathcal{N} = \oint d\hat{n} d\hat{n}' \left(\sqrt{P_\ell^\Phi}\right)_{\hat{n}\hat{n}'} \delta(\hat{n}'). \tag{C.20}$$

C.3. WMAP noise characterization

The pixel noise level (in units mK) of a single map can be determined by $\sigma = \sigma_0/\sqrt{N_{\rm obs}}$, where σ_0 can be found at `http://lambda.gsfc.nasa.gov/product/map/dr5/skymap_info.cfm` and the effective number of observations $N_{\rm obs}$, which can vary from pixel to pixel, is stored in the FITS file of a map, see `http://lambda.gsfc.nasa.gov/product/map/dr4/skymap_file_format_info.cfm`. Thus, the noise covariance matrix of a single map is given by

$$N_{\hat{\mathbf{n}},\hat{\mathbf{n}}'} = \frac{\sigma_0^2}{N_{\rm obs}(\hat{\mathbf{n}}')} \delta_{\hat{\mathbf{n}}\hat{\mathbf{n}}'}. \tag{C.21}$$

Including polarization data, the noise covariance matrix in position space has to be generalized by

$$N^{-1} = \begin{pmatrix} N_{\rm obs}^{TT}/\sigma_T^2 & 0 & 0 \\ 0 & N_{\rm obs}^{QQ}/\sigma_P^2 & N_{\rm obs}^{QU}/\sigma_P^2 \\ 0 & N_{\rm obs}^{QU}/\sigma_P^2 & N_{\rm obs}^{UU}/\sigma_P^2 \end{pmatrix}, \tag{C.22}$$

where $\sigma_{T,P}$ is the respective noise level of temperature and polarization as given by WMAP.

D. Appendix Chapter 5

Note: This appendix has been published in Phys. Rev. E **92**, 013302 (2015) [106].

Here Eq. (5.8) is derived. Following Sec. 5.2 the log-determinant Δ of an operator A can be parametrized by $\Delta = \ln[\det(D+N)]$ with D being the diagonal and N the off-diagonal part of A. Since Δ can be Taylor-expanded for small N (spectrally compared to D) only, we employ a method from the field of renormalization theory [22, 25]. Accordingly, we introduce an expansion parameter $\delta t \ll 1$ to suppress the influence of N. In particular, we replace Δ by $\ln[\det(D + \delta t N)]$ for a moment. For sufficiently small values of δt, in the following interpreted as tiny pseudotime steps, we can approximate Δ by Eq. (5.6). Theoretically, a single pseudotime step could be infinitesimal small, enabling the formal definition of the derivative

$$\begin{aligned}\frac{d\Delta(t)}{dt} &\equiv \lim_{\delta t\to 0} \frac{\ln[\det(D+(t+\delta t)N)] - \ln[\det(D+tN)]}{\delta t} \\ &= \lim_{\delta t\to 0} \frac{1}{\delta t} \ln\left[\det\left(\mathbb{1} + \delta t N[D+tN]^{-1}\right)\right] \\ &= \lim_{\delta t\to 0} \frac{1}{\delta t} \mathrm{tr}\left[\ln\left(\mathbb{1} + \delta t N[D+tN]^{-1}\right)\right] \\ &= \mathrm{tr}\left[N[D+tN]^{-1}\right],\end{aligned} \tag{D.1}$$

with the definition

$$\Delta(t) \equiv \ln[\det(D+tN)]. \tag{D.2}$$

Integrating the pseudotime derivative of $\Delta(t)$ yields the integral representation of the log-determinant,

$$\Delta = \int_0^1 dt\ \mathrm{tr}\left[N\left(D+tN\right)^{-1}\right] + \Delta(0). \tag{D.3}$$

This integral representation has also been found by Ref. [118], where its validity has been proven for weak diagonal dominant and Hermitian positive definite matrices. In particular one has to ensure the existence of the inverse matrix of the integrand of Eq. (D.3).

Finally, we replace the trace by stochastic trace probing and perform the pseudotime integral by an numeric integration method. This yields

$$\Delta = \int_0^1 dt\ \left\langle \xi^\dagger N\left(D+tN\right)^{-1}\xi \right\rangle_{\{\xi\}} + \Delta(0). \tag{D.4}$$

Bibliography

[1] J. W. Goethe, *Einzelheiten, Maximen und Reflexionen* (J. G. Cotta'schen Buchhandlung, 1833).

[2] S. Dodelson, *Modern Cosmology* (Academic Press, 2003) iSBN: 9780122191411.

[3] D. Baumann, *Cosmology* (`http://www.damtp.cam.ac.uk/user/db275/Cosmology/cosmology.pdf`, 2014) [online version date: 19 Nov. 2014].

[4] D. Baumann, ArXiv e-prints (2009), arXiv:0907.5424 [hep-th] .

[5] R. McMahon, Nature **443**, 151 (2006).

[6] F. C. Adams and G. Laughlin, Reviews of Modern Physics **69**, 337 (1997), astro-ph/9701131 .

[7] Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett, and et al., ArXiv e-prints (2015), arXiv:1502.01589 .

[8] Planck Collaboration, R. Adam, P. A. R. Ade, N. Aghanim, Y. Akrami, M. I. R. Alves, M. Arnaud, F. Arroja, J. Aumont, C. Baccigalupi, and et al., ArXiv e-prints (2015), arXiv:1502.01582 .

[9] J. Beringer, J. F. Arguin, R. M. Barnett, K. Copic, O. Dahl, D. E. Groom, C. J. Lin, J. Lys, H. Murayama, C. G. Wohl, W. M. Yao, P. A. Zyla, C. Amsler, M. Antonelli, D. M. Asner, H. Baer, H. R. Band, T. Basaglia, C. W. Bauer, J. J. Beatty, V. I. Belousov, E. Bergren, G. Bernardi, W. Bertl, S. Bethke, H. Bichsel, O. Biebel, E. Blucher, S. Blusk, G. Brooijmans, O. Buchmueller, R. N. Cahn, M. Carena, A. Ceccucci, D. Chakraborty, M. C. Chen, R. S. Chivukula, G. Cowan, G. D'Ambrosio, T. Damour, D. de Florian, A. de Gouvêa, T. DeGrand, P. de Jong, G. Dissertori, B. Dobrescu, M. Doser, M. Drees, D. A. Edwards, S. Eidelman, J. Erler, V. V. Ezhela, W. Fetscher, B. D. Fields, B. Foster, T. K. Gaisser, L. Garren, H. J. Gerber, G. Gerbier, T. Gherghetta, S. Golwala, M. Goodman, C. Grab, A. V. Gritsan, J. F. Grivaz, M. Grünewald, A. Gurtu, T. Gutsche, H. E. Haber, K. Hagiwara, C. Hagmann, C. Hanhart, S. Hashimoto, K. G. Hayes, M. Heffner, B. Heltsley, J. J. Hernández-Rey, K. Hikasa, A. Höcker, J. Holder, A. Holtkamp, J. Huston, J. D. Jackson, K. F. Johnson, T. Junk, D. Karlen, D. Kirkby, S. R. Klein, E. Klempt, R. V. Kowalewski, F. Krauss, M. Kreps, B. Krusche, Y. V. Kuyanov, Y. Kwon, O. Lahav, J. Laiho, P. Langacker, A. Liddle, Z. Ligeti, T. M. Liss, L. Littenberg, K. S. Lugovsky, S. B. Lugovsky, T. Mannel, A. V. Manohar, W. J. Marciano, A. D. Martin, A. Masoni, J. Matthews, D. Milstead, R. Miquel, K. Mönig, F. Moortgat, K. Nakamura, M. Narain, P. Nason, S. Navas, M. Neubert, P. Nevski, Y. Nir, K. A. Olive, L. Pape, J. Parsons, C. Patrignani, J. A. Peacock, S. T. Petcov, A. Piepke,

A. Pomarol, G. Punzi, A. Quadt, S. Raby, G. Raffelt, B. N. Ratcliff, P. Richardson, S. Roesler, S. Rolli, A. Romaniouk, L. J. Rosenberg, J. L. Rosner, C. T. Sachrajda, Y. Sakai, G. P. Salam, S. Sarkar, F. Sauli, O. Schneider, K. Scholberg, D. Scott, W. G. Seligman, M. H. Shaevitz, S. R. Sharpe, M. Silari, T. Sjöstrand, P. Skands, J. G. Smith, G. F. Smoot, S. Spanier, H. Spieler, A. Stahl, T. Stanev, S. L. Stone, T. Sumiyoshi, M. J. Syphers, F. Takahashi, M. Tanabashi, J. Terning, M. Titov, N. P. Tkachenko, N. A. Törnqvist, D. Tovey, G. Valencia, K. van Bibber, G. Venanzoni, M. G. Vincter, P. Vogel, A. Vogt, W. Walkowiak, C. W. Walter, D. R. Ward, T. Watari, G. Weiglein, E. J. Weinberg, L. R. Wiencke, L. Wolfenstein, J. Womersley, C. L. Woody, R. L. Workman, A. Yamamoto, G. P. Zeller, O. V. Zenin, J. Zhang, R. Y. Zhu, G. Harper, V. S. Lugovsky, and P. Schaffner, Phys. Rev. D **86**, 010001 (2012).

[10] Planck Collaboration, P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. J. Banday, and et al., Astronomy and Astrophysics **571**, A22 (2014), arXiv:1303.5082 .

[11] A. H. Guth, Phys. Rev. D **23**, 347 (1981).

[12] A. H. Guth and S.-H. H. Tye, Physical Review Letters **44**, 631 (1980).

[13] A. Linde, ArXiv High Energy Physics - Theory e-prints (2005), hep-th/0503203 .

[14] A. Linde, JCAP **10**, 004 (2004), hep-th/0408164 .

[15] V. Mukhanov and S. Winitzki, *Introduction to Quantum Effects in Gravity* (Cambridge University Press, 2007).

[16] C.-P. Ma and E. Bertschinger, The Astrophysical Journal **455**, 7 (1995), astro-ph/9506072 .

[17] S. Weinberg, *Cosmology*, Cosmology (OUP Oxford, 2008).

[18] E. Komatsu, ArXiv Astrophysics e-prints (2002), astro-ph/0206039 .

[19] Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, F. Arroja, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, and et al., ArXiv e-prints (2015), arXiv:1502.02114 .

[20] `http://www.esa.int/spaceinimages/Images/2015/02/The_history_of_the_Universe`, [online version date: 05 Feb. 2015].

[21] T. A. Enßlin, M. Frommert, and F. S. Kitaura, Phys. Rev. D **80**, 105005 (2009), arXiv:0806.3474 .

[22] S. Dorn, T. A. Enßlin, M. Greiner, M. Selig, and V. Boehm, Phys. Rev. E **91**, 013311 (2015), arXiv:1410.6289 [physics.data-an] .

[23] T. A. Enßlin, H. Junklewitz, L. Winderling, M. Greiner, and M. Selig, ArXiv e-prints (2013), arXiv:1312.1349 [astro-ph.IM] .

[24] C. Osborne, International Statistical Review / Revue Internationale de Statistique **59**, pp. 309 (1991).

[25] T. A. Enßlin and M. Frommert, Phys. Rev. D **83**, 105014 (2011), arXiv:1002.2928 [astro-ph.IM] .

[26] T. Bayes, Phil. Trans. of the Roy. Soc. **53**, 370 (1763).

[27] N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series* (New York: Wiley, 1949).

[28] S. L. Bridle, R. Crittenden, A. Melchiorri, M. P. Hobson, R. Kneissl, and A. N. Lasenby, Mon. Not. Roy. Astron. Soc. **335**, 1193 (2002), astro-ph/0112114 .

[29] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).

[30] M. Selig, M. R. Bell, H. Junklewitz, N. Oppermann, M. Reinecke, M. Greiner, C. Pachajoa, and T. A. Enßlin, Astronomy and Astrophysics **554**, A26 (2013), arXiv:1301.4499 [astro-ph.IM] .

[31] S. Dorn, E. Ramirez, K. E. Kunze, S. Hofmann, and T. A. Enßlin, JCAP **6**, 048 (2014), arXiv:1403.5067 .

[32] C. L. Bennett, D. Larson, J. L. Weiland, N. Jarosik, G. Hinshaw, N. Odegard, K. M. Smith, R. S. Hill, B. Gold, M. Halpern, E. Komatsu, M. R. Nolta, L. Page, D. N. Spergel, E. Wollack, J. Dunkley, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, and E. L. Wright, Astrophys. J. Suppl. Ser. **208**, 20 (2013), arXiv:1212.5225 [astro-ph.CO] .

[33] Planck Collaboration, P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, C. Aumont, A. J. Banday, and et al., Astronomy and Astrophysics **571**, A16 (2014), arXiv:1303.5076 [astro-ph.CO] .

[34] J. Martin and C. Ringeval, JCAP **8**, 009 (2006), astro-ph/0605367 .

[35] J. Martin, C. Ringeval, and V. Vennin, ArXiv e-prints (2013), arXiv:1303.3787 [astro-ph.CO] .

[36] C. Ringeval, ArXiv e-prints (2013), arXiv:1312.2347 [astro-ph.CO] .

[37] J. Martin, C. Ringeval, R. Trotta, and V. Vennin, ArXiv e-prints (2013), arXiv:1312.3529 [astro-ph.CO] .

[38] Planck Collaboration, P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. J. Banday, and et al., ArXiv e-prints (2013), arXiv:1303.5082 [astro-ph.CO] .

[39] Planck Collaboration, P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. J. Banday, and et al., ArXiv e-prints (2013), arXiv:1303.5084 [astro-ph.CO] .

[40] M. Sasaki, J. Väliviita, and D. Wands, Phys. Rev. D **74**, 103003 (2006), astro-ph/0607627 .

[41] BICEP2 Collaboration, P. A. R. Ade, R. W. Aikin, D. Barkats, S. J. Benton, C. A. Bischoff, J. J. Bock, J. A. Brevik, I. Buder, E. Bullock, C. D. Dowell, L. Duband, J. P. Filippini, S. Fliescher, S. R. Golwala, M. Halpern, M. Hasselfield, S. R. Hildebrandt, G. C. Hilton, V. V. Hristov, K. D. Irwin, K. S. Karkare, J. P. Kaufman, B. G. Keating, S. A. Kernasovskiy, J. M. Kovac, C. L. Kuo, E. M. Leitch, M. Lueker, P. Mason, C. B. Netterfield, H. T. Nguyen, R. O'Brient, R. W. Ogburn, IV, A. Orlando, C. Pryke, C. D. Reintsema, S. Richter, R. Schwarz, C. D. Sheehy, Z. K. Staniszewski, R. V. Sudiwala, G. P. Teply, J. E. Tolan, A. D. Turner, A. G. Vieregg, C. L. Wong, and K. W. Yoon, ArXiv e-prints (2014), arXiv:1403.3985 [astro-ph.CO] .

[42] D. K. Hazra, A. Shafieloo, G. F. Smoot, and A. A. Starobinsky, ArXiv e-prints (2014), arXiv:1403.7786 .

[43] B. Hu, J.-W. Hu, Z.-K. Guo, and R.-G. Cai, ArXiv e-prints (2014), arXiv:1404.3690 .

[44] D. K. Hazra, A. Shafieloo, G. F. Smoot, and A. A. Starobinsky, ArXiv e-prints (2014), arXiv:1405.2012 .

[45] A. Enea Romano and A. Gallego Cadavid, ArXiv e-prints (2014), arXiv:1404.2985 .

[46] J. Martin, L. Sriramkumar, and D. K. Hazra, ArXiv e-prints (2014), arXiv:1404.6093 .

[47] P. Creminelli, L. Senatore, and M. Zaldarriaga, JCAP **3**, 019 (2007), arXiv:astro-ph/0606001 .

[48] T. L. Smith, M. Kamionkowski, and B. D. Wandelt, Phys. Rev. D **84**, 063013 (2011), arXiv:1104.0930 [astro-ph.CO] .

[49] F. Elsner and B. D. Wandelt, The Astrophysical Journal **724**, 1262 (2010), arXiv:1010.1254 [astro-ph.CO] .

[50] F. Elsner, B. D. Wandelt, and M. D. Schneider, Astronomy and Astrophysics **513**, A59 (2010), arXiv:1002.1713 [astro-ph.CO] .

[51] T. L. Smith, D. Grin, and M. Kamionkowski, Phys. Rev. D **87**, 063003 (2013).

[52] E. Komatsu, A. Kogut, M. R. Nolta, C. L. Bennett, M. Halpern, G. Hinshaw, N. Jarosik, M. Limon, S. S. Meyer, L. Page, D. N. Spergel, G. S. Tucker, L. Verde, E. Wollack, and E. L. Wright, Astrophys. J. Suppl. Ser. **148**, 119 (2003), arXiv:astro-ph/0302223 .

[53] S. Dorn, N. Oppermann, R. Khatri, M. Selig, and T. A. Enßlin, Phys. Rev. D **88**, 103516 (2013), arXiv:1307.3884 [astro-ph.CO] .

[54] L. Verde, R. Jimenez, L. Alvarez-Gaume, A. F. Heavens, and S. Matarrese, JCAP **6**, 023 (2013), arXiv:1301.6017 [astro-ph.CO] .

[55] T. Giannantonio, A. J. Ross, W. J. Percival, R. Crittenden, D. Bacher, M. Kilbinger, R. Nichol, and J. Weller, ArXiv e-prints (2013), arXiv:1303.1349 [astro-ph.CO] .

[56] K. N. Abazajian, K. Arnold, J. Austermann, B. A. Benson, C. Bischoff, J. Bock, J. R. Bond, J. Borrill, I. Buder, D. L. Burke, E. Calabrese, J. E. Carlstrom, C. S. Carvalho, C. L. Chang, H. C. Chiang, S. Church, A. Cooray, T. M. Crawford, B. P. Crill, K. S. Dawson, S. Das, M. J. Devlin, M. Dobbs, S. Dodelson, O. Doré, J. Dunkley, J. L. Feng, A. Fraisse, J. Gallicchio, S. B. Giddings, D. Green, N. W. Halverson, S. Hanany, D. Hanson, S. R. Hildebrandt, A. Hincks, R. Hlozek, G. Holder, W. L. Holzapfel, K. Honscheid, G. Horowitz, W. Hu, J. Hubmayr, K. Irwin, M. Jackson, W. C. Jones, R. Kallosh, M. Kamionkowski, B. Keating, R. Keisler, W. Kinney, L. Knox, E. Komatsu, J. Kovac, C.-L. Kuo, A. Kusaka, C. Lawrence, A. T. Lee, E. Leitch, A. Linde, E. Linder, P. Lubin, J. Maldacena, E. Martinec, J. McMahon, A. Miller, L. Newburgh, M. D. Niemack, H. Nguyen, H. T. Nguyen, L. Page, C. Pryke, C. L. Reichardt, J. E. Ruhl, N. Sehgal, U. Seljak, L. Senatore, J. Sievers, E. Silverstein, A. Slosar, K. M. Smith, D. Spergel, S. T. Staggs, A. Stark, R. Stompor, A. G. Vieregg, G. Wang, S. Watson, E. J. Wollack, W. L. K. Wu, K. W. Yoon, O. Zahn, and M. Zaldarriaga, ArXiv e-prints (2013), arXiv:1309.5381 [astro-ph.CO] .

[57] N. Dalal, O. Doré, D. Huterer, and A. Shirokov, Phys. Rev. D **77**, 123514 (2008), arXiv:0710.4560 .

[58] A. Slosar, C. Hirata, U. Seljak, S. Ho, and N. Padmanabhan, JCAP **8**, 031 (2008), arXiv:0805.3580 .

[59] T. Giannantonio, C. Porciani, J. Carron, A. Amara, and A. Pillepich, Mon. Not. Roy. Astron. Soc. **422**, 2854 (2012), arXiv:1109.0958 [astro-ph.CO] .

[60] J.-Q. Xia, C. Baccigalupi, S. Matarrese, L. Verde, and M. Viel, JCAP **8**, 033 (2011), arXiv:1104.5015 [astro-ph.CO] .

[61] A. J. Ross, W. J. Percival, A. Carnero, G.-b. Zhao, M. Manera, A. Raccanelli, E. Aubourg, D. Bizyaev, H. Brewington, J. Brinkmann, J. R. Brownstein, A. J. Cuesta, L. A. N. da Costa, D. J. Eisenstein, G. Ebelke, H. Guo, J.-C. Hamilton, M. V. Magaña, E. Malanushenko, V. Malanushenko, C. Maraston, F. Montesano, R. C. Nichol, D. Oravetz, K. Pan, F. Prada, A. G. Sánchez, L. Samushia, D. J. Schlegel, D. P. Schneider, H.-J. Seo, A. Sheldon, A. Simmons, S. Snedden, M. E. C. Swanson, D. Thomas, J. L. Tinker, R. Tojeiro, and I. Zehavi, Mon. Not. Roy. Astron. Soc. **428**, 1116 (2013), arXiv:1208.1491 [astro-ph.CO] .

[62] P. Paykari, F. Lanusse, J.-L. Starck, F. Sureau, and J. Bobin, ArXiv e-prints (2014), arXiv:1402.1983 [astro-ph.CO] .

[63] N. Oppermann, M. Selig, M. R. Bell, and T. A. Enßlin, Phys. Rev. E **87**, 032136 (2013), arXiv:1210.6866 [astro-ph.IM] .

[64] A. P. S. Yadav and B. D. Wandelt, Advances in Astronomy **2010**, 565248 (2010), 10.1155/2010/565248, arXiv:1006.0275 [astro-ph.CO] .

[65] S. Dorn, N. Oppermann, and T. A. Enßlin, Phys. Rev. E **88**, 053303 (2013), arXiv:1307.3889 [astro-ph.IM] .

[66] K. Kojima, T. Kajino, and G. J. Mathews, JCAP **2**, 018 (2010), arXiv:0910.1976 [astro-ph.CO] .

[67] J. R. Shaw and A. Lewis, Phys. Rev. D **81**, 043517 (2010), arXiv:0911.2714 [astro-ph.CO] .

[68] N. S. Sugiyama, E. Komatsu, and T. Futamase, Phys. Rev. D **87**, 023530 (2013), arXiv:1208.1073 [gr-qc] .

[69] V. Acquaviva, N. Bartolo, S. Matarrese, and A. Riotto, Nuclear Physics B **667**, 119 (2003), astro-ph/0209156 .

[70] J. Maldacena, Journal of High Energy Physics **5**, 013 (2003), astro-ph/0210603 .

[71] H. Kodama and M. Sasaki, Progress of Theoretical Physics Supplement **78**, 1 (1984).

[72] V. F. Mukhanov, H. A. Feldman, and R. H. Brandenberger, Physical Reports **215**, 203 (1992).

[73] A. A. Starobinskiĭ, Soviet Journal of Experimental and Theoretical Physics Letters **42**, 152 (1985).

[74] M. Sasaki and E. D. Stewart, Progress of Theoretical Physics **95**, 71 (1996), astro-ph/9507001 .

[75] M. Sasaki and T. Tanaka, Progress of Theoretical Physics **99**, 763 (1998), gr-qc/9801017 .

[76] N. Bartolo and A. R. Liddle, Phys. Rev. D **65**, 121301 (2002), astro-ph/0203076 .

[77] K.-Y. Choi and Q.-G. Huang, Phys. Rev. D **87**, 043501 (2013), arXiv:1209.2277 [hep-ph] .

[78] G. Isidori, V. S. Rychkov, A. Strumia, and N. Tetradis, Phys. Rev. D **77**, 025034 (2008), arXiv:0712.0242 [hep-ph] .

[79] L. Alabidi, K. Malik, C. T. Byrnes, and K.-Y. Choi, JCAP **11**, 037 (2010), arXiv:1002.1700 [astro-ph.CO] .

[80] K. Ichikawa, T. Suyama, T. Takahashi, and M. Yamaguchi, Phys. Rev. D **78**, 063545 (2008), arXiv:0807.3988 .

[81] R. K. Sachs and A. M. Wolfe, The Astrophysical Journal **147**, 73 (1967).

[82] M. White and W. Hu, Astronomy and Astrophysics **321**, 8 (1997), arXiv:astro-ph/9609105 .

[83] S. R. Cook, A. Gelman, and D. B. Rubin, Journal of Computational and Graphical Statistics **15**, 675 (2006).

[84] T. A. Enßlin and C. Weig, Phys. Rev. E **82**, 051112 (2010), arXiv:1004.2868 [astro-ph.IM] .

[85] N. Oppermann, H. Junklewitz, G. Robbers, M. R. Bell, T. A. Enßlin, A. Bonafede, R. Braun, J. C. Brown, T. E. Clarke, I. J. Feain, B. M. Gaensler, A. Hammond, L. Harvey-Smith, G. Heald, M. Johnston-Hollitt, U. Klein, P. P. Kronberg, S. A. Mao, N. M. McClure-Griffiths, S. P. O'Sullivan, L. Pratley, T. Robishaw, S. Roy,

D. H. F. M. Schnitzeler, C. Sotomayor-Beltran, J. Stevens, J. M. Stil, C. Sunstrum, A. Tanna, A. R. Taylor, and C. L. Van Eck, Astronomy and Astrophysics **542**, A93 (2012), arXiv:1111.6186 [astro-ph.GA] .

[86] H. Junklewitz, M. R. Bell, M. Selig, and T. A. Enßlin, ArXiv e-prints (2013), arXiv:1311.5282 [astro-ph.IM] .

[87] M. Selig and T. Enßlin, ArXiv e-prints (2013), arXiv:1311.1888 [astro-ph.IM] .

[88] K. Karhunen, Ann. Acad. Sci. Fennicae. Ser. A. **37**, 1 (1947).

[89] M. Loève, *Probability Theory II* -, 4th ed. (Springer, Berlin, Heidelberg, 1978).

[90] H. A. Feldman, N. Kaiser, and J. A. Peacock, The Astrophysical Journal **426**, 23 (1994), astro-ph/9304022 .

[91] M. Shiraishi, S. Yokoyama, D. Nitta, K. Ichiki, and K. Takahashi, Phys. Rev. D **82**, 103505 (2010), arXiv:1003.2096 [astro-ph.CO] .

[92] D. Babich, Phys. Rev. D **72**, 043003 (2005), astro-ph/0503375 .

[93] S. Dorn, M. Greiner, and T. A. Enßlin, JCAP **2**, 041 (2015), arXiv:1412.8315 .

[94] J. Jasche and B. D. Wandelt, Mon. Not. Roy. Astron. Soc. **432**, 894 (2013), arXiv:1203.3639 [astro-ph.CO] .

[95] S. Nurmi and M. S. Sloth, JCAP **7**, 012 (2014), arXiv:1312.4946 .

[96] S. Matarrese, S. Mollerach, A. Notari, and A. Riotto, Phys. Rev. D **71**, 043502 (2005), astro-ph/0410687 .

[97] C. L. Bennett, D. Larson, J. L. Weiland, N. Jarosik, G. Hinshaw, N. Odegard, K. M. Smith, R. S. Hill, B. Gold, M. Halpern, E. Komatsu, M. R. Nolta, L. Page, D. N. Spergel, E. Wollack, J. Dunkley, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, and E. L. Wright, Astrophys. J. Suppl. Ser. **208**, 20 (2013), arXiv:1212.5225 [astro-ph.CO] .

[98] E. Komatsu, D. N. Spergel, and B. D. Wandelt, The Astrophysical Journal **634**, 14 (2005), arXiv:astro-ph/0305189 .

[99] A. P. Yadav and B. D. Wandelt, Phys. Rev. D **71**, 123004 (2005), astro-ph/0505386 .

[100] M. Zaldarriaga and U. Seljak, Phys. Rev. D **55**, 1830 (1997), astro-ph/9609170 .

[101] E. Komatsu and D. N. Spergel, Phys. Rev. D **63**, 063002 (2001), arXiv:astro-ph/0005036 .

[102] P. Paykari, F. Lanusse, J.-L. Starck, F. Sureau, and J. Bobin, Astronomy and Astrophysics **566**, A77 (2014), arXiv:1402.1983 .

[103] N. Oppermann, M. Selig, M. R. Bell, and T. A. Enßlin, ArXiv e-prints (2012), arXiv:1210.6866 [astro-ph.IM] .

[104] G. Hinshaw, D. Larson, E. Komatsu, D. N. Spergel, C. L. Bennett, J. Dunkley, M. R. Nolta, M. Halpern, R. S. Hill, N. Odegard, L. Page, K. M. Smith, J. L. Weiland, B. Gold, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, E. Wollack, and E. L. Wright, Astrophys. J. Suppl. Ser. **208**, 19 (2013), arXiv:1212.5226 [astro-ph.CO] .

[105] U. Seljak and M. Zaldarriaga, The Astrophysical Journal **469**, 437 (1996), arXiv:astro-ph/9603033 .

[106] S. Dorn and T. A. Enßlin, Phys. Rev. E **92**, 013302 (2015), arXiv:1504.02661 .

[107] D. G. York, J. Adelman, J. E. Anderson, Jr., S. F. Anderson, J. Annis, N. A. Bahcall, J. A. Bakken, R. Barkhouser, S. Bastian, E. Berman, W. N. Boroski, S. Bracker, C. Briegel, J. W. Briggs, J. Brinkmann, R. Brunner, S. Burles, L. Carey, M. A. Carr, F. J. Castander, B. Chen, P. L. Colestock, A. J. Connolly, J. H. Crocker, I. Csabai, P. C. Czarapata, J. E. Davis, M. Doi, T. Dombeck, D. Eisenstein, N. Ellman, B. R. Elms, M. L. Evans, X. Fan, G. R. Federwitz, L. Fiscelli, S. Friedman, J. A. Frieman, M. Fukugita, B. Gillespie, J. E. Gunn, V. K. Gurbani, E. de Haas, M. Haldeman, F. H. Harris, J. Hayes, T. M. Heckman, G. S. Hennessy, R. B. Hindsley, S. Holm, D. J. Holmgren, C.-h. Huang, C. Hull, D. Husby, S.-I. Ichikawa, T. Ichikawa, Ž. Ivezić, S. Kent, R. S. J. Kim, E. Kinney, M. Klaene, A. N. Kleinman, S. Kleinman, G. R. Knapp, J. Korienek, R. G. Kron, P. Z. Kunszt, D. Q. Lamb, B. Lee, R. F. Leger, S. Limmongkol, C. Lindenmeyer, D. C. Long, C. Loomis, J. Loveday, R. Lucinio, R. H. Lupton, B. MacKinnon, E. J. Mannery, P. M. Mantsch, B. Margon, P. McGehee, T. A. McKay, A. Meiksin, A. Merelli, D. G. Monet, J. A. Munn, V. K. Narayanan, T. Nash, E. Neilsen, R. Neswold, H. J. Newberg, R. C. Nichol, T. Nicinski, M. Nonino, N. Okada, S. Okamura, J. P. Ostriker, R. Owen, A. G. Pauls, J. Peoples, R. L. Peterson, D. Petravick, J. R. Pier, A. Pope, R. Pordes, A. Prosapio, R. Rechenmacher, T. R. Quinn, G. T. Richards, M. W. Richmond, C. H. Rivetta, C. M. Rockosi, K. Ruthmansdorfer, D. Sandford, D. J. Schlegel, D. P. Schneider, M. Sekiguchi, G. Sergey, K. Shimasaku, W. A. Siegmund, S. Smee, J. A. Smith, S. Snedden, R. Stone, C. Stoughton, M. A. Strauss, C. Stubbs, M. SubbaRao, A. S. Szalay, I. Szapudi, G. P. Szokoly, A. R. Thakar, C. Tremonti, D. L. Tucker, A. Uomoto, D. Vanden Berk, M. S. Vogeley, P. Waddell, S.-i. Wang, M. Watanabe, D. H. Weinberg, B. Yanny, N. Yasuda, and SDSS Collaboration, The Astronomical Journal **120**, 1579 (2000), astro-ph/0006396 .

[108] D. N. Spergel, R. Bean, O. Doré, M. R. Nolta, C. L. Bennett, J. Dunkley, G. Hinshaw, N. Jarosik, E. Komatsu, L. Page, H. V. Peiris, L. Verde, M. Halpern, R. S. Hill, A. Kogut, M. Limon, S. S. Meyer, N. Odegard, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright, Astrophys. J. Suppl. Ser. **170**, 377 (2007), astro-ph/0603449 .

[109] G. J. Hill, K. Gebhardt, E. Komatsu, N. Drory, P. J. MacQueen, J. Adams, G. A. Blanc, R. Koehler, M. Rafal, M. M. Roth, A. Kelz, C. Gronwall, R. Ciardullo, and D. P. Schneider, *Panoramic Views of Galaxy Formation and Evolution*, Astronomical Society of the Pacific Conference Series **399**, 115 (2008), arXiv:0806.0183 .

[110] F. S. Kitaura and T. A. Enßlin, Mon. Not. Roy. Astron. Soc. **389**, 497 (2008), arXiv:0705.0429 .

[111] M. Selig, N. Oppermann, and T. A. Enßlin, Phys. Rev. E **85**, 021134 (2012), arXiv:1108.0600 [astro-ph.IM] .

[112] E. Aune, D. Simpson, and J. Eidsvik, Statistics and Computing **24**, 247 (2014).

[113] M. Hutchinson, Communications in Statistics - Simulation and Computation **18**, 1059 (1989).

[114] C. Bekas, E. Kokiopoulou, and Y. Saad, Applied Numerical Mathematics **57**, 1214 (2007).

[115] M. R. Hestenes and E. Stiefel, Journal of Research of the National Bureau of Standards **49**, 409 (1952).

[116] M.-D. Wu and W. Fitzgerald, *Maximum Entropy and Bayesian Methods*, edited by K. Hanson and R. Silver, Fundamental Theories of Physics, Vol. 79 (Springer Netherlands, 1996) pp. 153–162.

[117] D. C. M. Dickson, M. R. Hardy, and H. R. Waters, *Physics and Probability - Essays in Honor of Edwin T. Jaynes*, 1st ed. (Cambridge University Press, Cambridge, 2004).

[118] J. Du and J. Ji, Dynamical Systems **2005**, 225 (2005).

[119] E. Jaynes and G. Bretthorst, *Probability Theory: The Logic of Science, Chapter 12.4* (Cambridge University Press, 2003).

[120] J. Martin, C. Ringeval, R. Trotta, and V. Vennin, JCAP **3**, 039 (2014), arXiv:1312.3529 .

[121] J. Skilling, AIP Conference Proceedings **735**, 395 (2004).

[122] F. Feroz, M. P. Hobson, and M. Bridges, Mon. Not. Roy. Astron. Soc. **398**, 1601 (2009), arXiv:0809.3437 .

[123] T. A. Enßlin, H. Junklewitz, L. Winderling, M. Greiner, and M. Selig, Phys. Rev. E **90**, 043301 (2014), arXiv:1312.1349 [astro-ph.IM] .

[124] C. M. Hirata and U. Seljak, Phys. Rev. D **67**, 043001 (2003), astro-ph/0209489 .

[125] E. Keihänen and M. Reinecke, Astronomy and Astrophysics **548**, A110 (2012), arXiv:1208.1399 [astro-ph.IM] .

[126] E. R. Harrison, Mon. Not. Roy. Astron. Soc. **147**, 279 (1970).

[127] S. Saga, K. Ichiki, K. Takahashi, and N. Sugiyama, Phys. Rev. D **91**, 123510 (2015), arXiv:1504.03790 .

[128] K. M. Gòrski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann, The Astrophysical Journal **622**, 759 (2005).

Acknowledgments

Acknowledgements according to Chapters

Note: The acknowledgments of each chapter have been published in Phys. Rev. E **91** (2015), 013311 [22], JCAP **6**, 048 (2014) [31], JCAP **2**, 041 (2015) [93], and Phys. Rev. E **92**, 013302 (2015) [106], respectively.

Chapter 2

Calculations were realized using the NIFTy [30] package to be found at `https://github.com/mselig/nifty` as well as `scipy` to be found at `http://www.scipy.org`.

Chapter 3

We gratefully acknowledge Eiichiro Komatsu and Vanessa Boehm for useful discussions and comments on the manuscript. We also want to thank Marco Selig and Maksim Greiner for numerical support and discussions hereof. KEK would like to thank the Max-Planck-Institute for Astrophysics for hospitality where this work was initiated and acknowledges financial support by Spanish Science Ministry grants FIS2012- 30926 and CSD2007-00042. The work of SH was supported by the DFG cluster of excellence *Origin and Structure of the Universe* and by TRR 33 *The Dark Universe*. Calculations were realized using the NIFTy [30] package to be found at `http://www.mpa-garching.mpg.de/ift/nifty/`.

Chapter 4

We gratefully acknowledge Vanessa Boehm and Marco Selig for useful discussions and comments on the manuscript, as well as Eiichiro Komatsu for numerical support concerning `gTfast`, to be found at `http://www.mpa-garching.mpg.de/~komatsu/CRL/nongaussianity/radiationtransferfunction/`. All calculations have been done using NIFTy [30] to be found at `http://www.mpa-garching.mpg.de/ift/nifty/`, in particular involving `HEALPix` [128] to be found at `http://healpix.sourceforge.net/`. We also acknowledge the support by the DFG Cluster of Excellence "Origin and Structure of the Universe". The calculations have been carried out on the computing facilities of the Computational Center for Particle and Astrophysics (C2PAP).

Chapter 5

We gratefully acknowledge Maksim Greiner for discussions and David Butler for useful comments on the manuscript. All calculations were realized using NIFTy [30] to be found at `http://www.mpa-garching.mpg.de/ift/nifty`.

General acknowledgments

A lot of people have contributed to the success of this thesis, but first of all, I would like to thank the whole Max-Planck-Institute for Astrophysics for a very friendly, informal, and emancipated working atmosphere that formed day by day an enjoyable period of residence and enabled the travel to conferences and meetings. A couple of people, however, deserve a mentioning by name:

- The IFT group – in particular Niels, Henrik, Daniel, Valentina, Vanessa, Theo, Mahsa, Jens, Jörg, Martin, David, Chris, and Natalia for many fruitful and interesting discussions about science, politics, and much else. Thank you also for proofreading a couple of papers and an introduction to radio astronomy (thanks Valentina & Henrik).

- Marco and Maksim – who assisted & supported me in solving many numerical and theoretical problems. Thank you for the friendly team play.

- Torsten – as my Ph.D. supervisor he never failed in inspiring, guiding, and supporting me. He had a sympathetic ear at all times for all kind of problems and invested significant parts of his time into this project – not least by giving uncountable advises to solve scientific problems or by the promotion of my work. Thank you for your engagement.

Finally, I would like to give a cordial thanks to my whole family and closest friends for their support – above all to my wife Marina, who deserves a thousand thanks for unconditionally supporting me during my doctoral studies.