

This is a repository copy of *Cosmological perturbation theory and quantum gravity*.

White Rose Research Online URL for this paper:  
<https://eprints.whiterose.ac.uk/113541/>

Version: Published Version

---

**Article:**

Brunetti, Romeo, Fredenhagen, Klaus, Hack, Thomas-Paul et al. (2 more authors) (2016)  
Cosmological perturbation theory and quantum gravity. *Journal of High Energy Physics*.  
ISSN 1029-8479

[https://doi.org/10.1007/JHEP08\(2016\)032](https://doi.org/10.1007/JHEP08(2016)032)

---

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:  
<https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.

# Cosmological perturbation theory and quantum gravity

Romeo Brunetti,<sup>a</sup> Klaus Fredenhagen,<sup>b</sup> Thomas-Paul Hack,<sup>c</sup> Nicola Pinamonti<sup>d,e</sup> and Katarzyna Rejzner<sup>f</sup>

<sup>a</sup>*Dipartimento di Matematica, Università di Trento,  
Via Sommarive 14, 38123 Povo TN, Italy*

<sup>b</sup>*II Institute für Theoretische Physik, Universität Hamburg,  
Luruper Chaussee 149, 22761 Hamburg, Germany*

<sup>c</sup>*Institute für Theoretische Physik, Universität Leipzig,  
Brüderstr. 16, 04103 Leipzig, Germany*

<sup>d</sup>*Dipartimento di Matematica, Università di Genova,  
Via Dodecaneso 35, 16146 Genova, Italy*

<sup>e</sup>*INFN, Sezione di Genova,  
Via Dodecaneso 33, 16146 Genova, Italy*

<sup>f</sup>*Department of Mathematics, University of York,  
Heslington, York YO10 5DD, United Kingdom*

*E-mail:* [romeo.brunetti@unitn.it](mailto:romeo.brunetti@unitn.it), [klaus.fredenhagen@desy.de](mailto:klaus.fredenhagen@desy.de),  
[thomas-paul.hack@itp.uni-leipzig.de](mailto:thomas-paul.hack@itp.uni-leipzig.de), [pinamont@dima.unige.it](mailto:pinamont@dima.unige.it),  
[kasia.rejzner@york.ac.uk](mailto:kasia.rejzner@york.ac.uk)

**ABSTRACT:** It is shown how cosmological perturbation theory arises from a fully quantized perturbative theory of quantum gravity. Central for the derivation is a non-perturbative concept of gauge-invariant local observables by means of which perturbative invariant expressions of arbitrary order are generated. In particular, in the linearised theory, first order gauge-invariant observables familiar from cosmological perturbation theory are recovered. Explicit expressions of second order quantities are presented as well.

**KEYWORDS:** Models of Quantum Gravity, Classical Theories of Gravity

**ARXIV EPRINT:** [1605.02573](https://arxiv.org/abs/1605.02573)

---

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Perturbations of the Einstein-Klein-Gordon system on a FLRW space-time</b>	<b>4</b>
<b>3</b>	<b>All-order gauge-invariant observables on FLRW backgrounds</b>	<b>6</b>
3.1	Perturbative expansion up to second order	8
3.2	Non-degenerate covariant coordinates on FLRW backgrounds	8
<b>4</b>	<b>Examples of gauge-invariant observables at second order</b>	<b>10</b>
4.1	The lapse function	11
4.2	The spatial curvature	12
<b>5</b>	<b>Quantization</b>	<b>15</b>
5.1	BRST quantization	15
5.2	Renormalisation	16
<b>6</b>	<b>Conclusions</b>	<b>17</b>

---

## 1 Introduction

The fluctuations of the cosmological microwave background provide a deep insight into the early history of the universe. The most successful theoretical explanation is inflationary cosmology where a scalar field (the inflaton) is coupled to the gravitational field. Usually, the theory is considered in linear order around a highly symmetric background, typically the spatially flat Friedmann-Lemaître-Robertson-Walker spacetime.

Extending the theory to higher orders is accompanied by severe obstacles. Already in a classical analysis the definition of gauge-invariant observables turns out to be rather complicated; moreover, one is immediately confronted with the problem of constructing a theory of quantum gravity. Previous treatments of higher-order cosmological perturbation theory include [4, 6, 17–19, 21, 23, 24]; many further references on the subject can be found e.g. in [17].

In a recent paper [2] three of us reanalysed the field theoretical construction of quantum gravity from the view point of locally covariant quantum field theory. This analysis was based on the methods of perturbative Algebraic Quantum Field Theory (pAQFT), see [13] and references therein, and on an adapted version of the Batalin-Vilkovisky formalism for the treatment of local gauge symmetries [12, 15]. The result was that a consistent theory (in the sense of an expansion into a formal power series) exists and is independent of the background. Due to non-renormalisability, however, in each order of perturbation theory

new dimensionful coupling constants occur, which have to be fixed by experiments; hence the theory should be interpreted as an effective theory that is valid at scales where these new constants are irrelevant. One might hope that non-perturbative effects improve the situation in the sense of Weinberg’s concept of asymptotic safety, since there are encouraging results supporting this perspective; see for example [26, 27]. Furthermore, it is difficult to observe any effects of quantum gravity, so it seems reasonable to start from the hypothesis that at presently accessible scales the influence of these higher order contributions is small.

One of the main questions addressed by [2] in the construction of the theory was the existence of local observables. It was answered, in a way familiar from classical general relativity, by using physical scalar fields, e.g. curvature scalars, as coordinates, and by expressing other fields as functions of these coordinates. Since quantization in the framework of pAQFT relies on a field theoretical version of deformation quantization of classical theories (first introduced in [9]), the classical construction can be transferred to the quantum realm.

The procedure works as follows. One selects 4 scalar fields  $X_\Gamma^a, a = 1, \dots, 4$ , which are functionals of the field configuration  $\Gamma$  which includes the spacetime metric  $g$ , the inflaton field  $\phi$  and possibly other fields. The fields  $X_\Gamma^a$  are supposed to transform under diffeomorphisms  $\chi$  as

$$X_{\chi^*\Gamma}^a = X_\Gamma^a \circ \chi, \tag{1.1}$$

where  $\chi^*$  denotes the pullback (of sections of direct sums of tensor products of the cotangent bundle) via  $\chi$ . We choose a background  $\Gamma_0$  such that the map

$$X_{\Gamma_0} : x \mapsto (X_{\Gamma_0}^1, \dots, X_{\Gamma_0}^4) \tag{1.2}$$

is injective. In order to achieve injectivity on cosmological backgrounds  $\Gamma_0$ , we shall be forced to include the coordinates  $x$  in the construction of  $X_\Gamma$  in a way which is compatible with (1.1). We then consider  $\Gamma$  sufficiently near  $\Gamma_0$  and set

$$\alpha_\Gamma = X_\Gamma^{-1} \circ X_{\Gamma_0}. \tag{1.3}$$

We observe that  $\alpha_\Gamma$  transforms under diffeomorphisms — which leave the background  $\Gamma_0$ , that is by definition fixed, invariant — as

$$\alpha_{\chi^*\Gamma} = \chi^{-1} \circ \alpha_\Gamma. \tag{1.4}$$

Let now  $A_\Gamma$  be any other scalar field which is a local functional of  $\Gamma$  and transforms under diffeomorphisms as in (1.1). Then the field

$$\mathcal{A}_\Gamma := A_\Gamma \circ \alpha_\Gamma \tag{1.5}$$

is invariant under diffeomorphisms and may be considered as a local observable. Note that invariance is obtained by shifting the argument of the field in a way which depends on the configuration.

The physical interpretation of this construction is as follows: the fields  $X_\Gamma^a$  are configuration-dependent coordinates such that  $[A_\Gamma \circ X_\Gamma^{-1}](Y)$  corresponds to the value

of the quantity  $A_\Gamma$  provided that the quantity  $X_\Gamma$  has the value  $X_\Gamma = Y$ . Thus  $A_\Gamma \circ X_\Gamma^{-1}$  is a partial or relational observable [7, 28, 30], and by considering  $\mathcal{A}_\Gamma = A_\Gamma \circ X_\Gamma^{-1} \circ X_{\Gamma_0}$  we can interpret this observable as a field on the background spacetime.

Clearly, to make things precise, one also has to characterise the region in the configuration space where all the maps are well defined and restrict oneself to configurations  $\Gamma$  in the appropriate neighbourhood of the background  $\Gamma_0$ , see [2, 16] for details.

Fortunately, in formal deformation quantization as well as in perturbation theory, only the Taylor expansion of observables around some background configuration enters, hence it is sufficient to establish the injectivity of  $X_{\Gamma_0}$  in order for the expansion of  $\mathcal{A}_{\Gamma_0+\delta\Gamma}$  around  $\Gamma_0$  to be well-defined. As an example we compute this expansion up to the first order. We obtain

$$\mathcal{A}_{\Gamma_0+\delta\Gamma} = A_{\Gamma_0} + \left\langle \frac{\delta A_\Gamma}{\delta\Gamma}(\Gamma_0), \delta\Gamma \right\rangle + \frac{\partial A_{\Gamma_0}}{\partial x^\mu} \left\langle \frac{\delta\alpha_\Gamma^\mu}{\delta\Gamma}(\Gamma_0), \delta\Gamma \right\rangle + O(\delta\Gamma^2). \quad (1.6)$$

The third term on the right hand side is necessary in order to get gauge-invariant fields (up to first order). We calculate

$$\frac{\delta\alpha_\Gamma^\mu}{\delta\Gamma}(\Gamma_0) = - \left( \left( \frac{\partial X_{\Gamma_0}}{\partial x} \right)^{-1} \right)^\mu_a \frac{\delta X_\Gamma^a}{\delta\Gamma}(\Gamma_0). \quad (1.7)$$

In this work we apply this general idea to inflationary cosmology. Unfortunately, due to the high symmetry of a FLRW background one can find only one independent background field which can serve as time coordinate. On the corresponding time slices one can then use three independent coordinates which are harmonic functions for the induced metric on the time slice and are fixed by their asymptotic behavior. Thus they transform correctly under diffeomorphisms which vanish at infinity but depend nonlocally on the dynamical fields. We show that these coordinates can be used for the definition of observables. Moreover, in linear order these observables coincide with the observables used in cosmological perturbation theory. In the quantized theory the nonlocalities induce UV and IR problems. The UV problems seem to be harmless and we indicate a possible solution. For the IR problems one has to investigate the decay of typical correlations, in a similar way as already known from cosmological perturbation theory.

In contrast to other systematic or covariant attempts to define gauge-invariant quantities in higher-order cosmological perturbation theory, see for example [17, 19, 22, 24], our construction works off-shell, is based on a clear and simple concept which is applicable to general backgrounds such that cosmological perturbation theory may be viewed as a particular application of perturbative quantum gravity [2]. Moreover, we construct non-perturbative gauge-invariant quantities whose perturbative expansion to arbitrary orders may be computed algorithmically without the need for additional input at each order. A work which is similar in the latter respect, but approaches the problem from a canonical perspective and with a stronger focus on dynamics, is [8].

This paper is organised as follows: in the second section we recall a few basic facts about perturbation theory of the Einstein-Klein-Gordon system on cosmological backgrounds. In the third section we describe the general method to obtain gauge invariant observables at

all orders on generic backgrounds. We furthermore discuss how to treat the case of a FLRW background where the large symmetry prevents us from using coordinates constructed from the dynamical fields alone. The fourth section contains the analysis of two gauge invariant observables at second order. The steps necessary for the construction of a full all-order quantum theory are briefly sketched in section 5. Finally a number of conclusions are drawn in the last section.

## 2 Perturbations of the Einstein-Klein-Gordon system on a FLRW space-time

We consider the Einstein-Klein-Gordon system, namely a minimally coupled scalar field  $\tilde{\phi}$  with potential  $V(\tilde{\phi})$  propagating on a Lorentzian spacetime  $(M, \tilde{g})$  with field equations

$$R_{ab} - \frac{1}{2}R\tilde{g}_{ab} = T_{ab}, \quad -\square\tilde{\phi} + V^{(1)}(\tilde{\phi}) = 0, \quad (2.1)$$

where  $T_{ab}$  is the stress tensor of  $\tilde{\phi}$ ,  $R_{ab}$  the Ricci tensor and  $R$  the Ricci scalar. We discuss perturbations of this system around a background. A linearised theory is obtained starting from a one-parameter family of solutions  $\lambda \mapsto \Gamma_\lambda := (\tilde{g}_\lambda, \tilde{\phi}_\lambda)$  and considering

$$\delta\Gamma := (\gamma, \varphi) := \left. \frac{d}{d\lambda}(\tilde{g}_\lambda, \tilde{\phi}_\lambda) \right|_{\lambda=0},$$

hence  $\Gamma_0 := (g, \phi) := (\tilde{g}_0, \tilde{\phi}_0)$  is the background configuration while  $\delta\Gamma = (\gamma, \varphi)$  is the linearised perturbation.

The background solution we choose consists of a flat Friedmann-Lemaître-Robertson-Walker (FLRW) spacetime  $(M, g)$  together with a scalar field  $\phi$  which is constant in space. We recall that a flat FLRW spacetime is conformally flat and that

$$M = I \times \mathbb{R}^3, \quad g = a^2(\tau)(-d\tau \otimes d\tau + \sum_i dx^i \otimes dx^i), \quad (2.2)$$

where  $I \subset \mathbb{R}$  is an open interval, the scale factor  $a(\tau)$  is a function of the conformal time  $\tau$  and where  $x^i$  are three-dimensional Cartesian (comoving) coordinates. The background equations of motion of the system are best displayed in terms of the auxiliary function

$$\mathcal{H} := \frac{a'}{a},$$

where  $a'$  indicates the derivative with respect to the conformal time.  $\mathcal{H}$  is related to the Hubble parameter  $H = \mathcal{H}a^{-1}$  and to the Ricci scalar  $R = 6(\mathcal{H}' + \mathcal{H}^2)a^{-2}$ . The background equations of motion are

$$\begin{aligned} 6\mathcal{H}^2 &= (\phi')^2 + 2a^2V(\phi) & 2(2\mathcal{H}' + \mathcal{H}^2) &= -(\phi')^2 + 2a^2V(\phi), \\ \phi'' + 2\mathcal{H}\phi' + a^2V^{(1)}(\phi) &= 0. \end{aligned}$$

A generic perturbation  $\gamma$  of the FLRW metric  $g$  can be decomposed in the following way

$$\gamma = a(\tau)^2 \begin{pmatrix} -2A & (-\partial_i B + V_i)^t \\ -\partial_i B + V_i & 2(\partial_i \partial_j E + \delta_{ij} D + \partial_{(i} W_{j)} + T_{ij}) \end{pmatrix} \quad (2.3)$$

where  $A, B, D, E$  are scalars,  $V, W$  are three dimensional vectors and  $T$  is a tensor on 3-dimensional Euclidean space. The decomposition is unique if all these perturbations vanish at infinity and if

$$T_i^i = 0, \quad \partial_i T^i_j = 0, \quad \partial_i V^i = 0, \quad \partial_i W^i = 0$$

(see e.g. Proposition 3.1 in [14]).

Under an infinitesimal first order gauge transformation the linear perturbations transform in the following way

$$\gamma_{ab} \mapsto \gamma_{ab} + \mathcal{L}_\xi g_{ab} = \gamma_{ab} + 2\nabla_{(a}\xi_{b)}, \quad \varphi \mapsto \varphi + \mathcal{L}_\xi \phi = \varphi + \xi(\phi).$$

In particular

$$\begin{aligned} A &\mapsto A + (\partial_\tau + \mathcal{H})r, & B &\mapsto B + r - s', & D &\mapsto D + \mathcal{H}r, & E &\mapsto E + s, \\ \varphi &\mapsto \varphi + \phi' r, & V_i &\mapsto V_i + v'_i, & W_i &\mapsto W_i + v_i, & T_{ij} &\mapsto T_{ij}, \end{aligned}$$

where the generator  $\xi$  of one-parameter gauge transformations is also decomposed as

$$\xi^0 = r, \quad \xi^i = \partial_i s + v_i, \quad \partial_i v^i = 0. \quad (2.4)$$

Notice that the gauge transformations do not mix scalar, vector or tensor perturbations at linear order.

Furthermore, we observe that tensor perturbations are gauge-invariant and that gauge-invariant vector perturbations can be obtained considering  $X_i := W'_i - V_i$ . Regarding the scalar perturbations we see that the following fields are gauge-invariant

$$\Phi := A - (\partial_t + \mathcal{H})(B + E'), \quad \Psi := D - \mathcal{H}(B + E'), \quad \chi := \varphi - \phi'(B + E'). \quad (2.5)$$

The first two of them are called Bardeen potentials.

Let us recall the form of the linearised equations of motions satisfied by the gauge-invariant perturbations. The first observation is that the equations of motion respect the decomposition in scalar, vector and tensor perturbations. In particular, for the vector and tensor perturbations, it holds that

$$\Delta X_i = 0, \quad (\partial_t + 2\mathcal{H})X_i = 0, \quad \frac{1}{a^2}(\partial_t^2 + 2\mathcal{H}\partial_t - \Delta)T_{ij} = 0. \quad (2.6)$$

For the scalar part the equations of motion are better displayed in terms of the Mukhanov-Sasaki variable

$$\mu := \chi - \frac{\phi'}{\mathcal{H}}\Psi = \varphi - \frac{\phi'}{\mathcal{H}}D. \quad (2.7)$$

The equation of motion for this variable is decoupled also from the other scalars of the theory, in fact

$$\left(-\square + \frac{R}{6} - \frac{z''}{za^2}\right)\mu = 0, \quad z := \frac{a\phi'}{\mathcal{H}}.$$

The other scalar perturbations can be obtained in terms of  $\mu$ . In particular the Bardeen potential  $\Phi$  is the unique solution of

$$\Delta\Phi = \frac{\phi'}{2} \left( \mu' + \left( \frac{\mathcal{H}'}{\mathcal{H}} - \frac{\phi''}{\phi'} \right) \mu \right) \tag{2.8}$$

while the other scalar perturbations are given by

$$\Psi = -\Phi, \quad \chi = \frac{2}{\phi'} (\partial_\tau + \mathcal{H})\Phi. \tag{2.9}$$

We briefly discuss the situation beyond linear order. According to [29], infinitesimal diffeomorphisms may be approximated by so-called knight diffeomorphisms, which are of the form  $\exp \mathcal{L}_\xi$  with  $\xi = \lambda\xi_1 + \frac{1}{2}\lambda^2\xi_2 + O(\lambda^3)$ . Analogously we may expand a configuration  $\Gamma$  as  $\Gamma = \Gamma_0 + \delta\Gamma = \Gamma_0 + \lambda\delta\Gamma_1 + \frac{1}{2}\lambda^2\delta\Gamma_2 + O(\lambda^3)$ , and determine the transformation behaviour of separate orders by considering  $\exp \mathcal{L}_\xi \Gamma$  at fixed order in  $\lambda$ , see for example [4, 6, 19, 21, 23]. Assuming that  $\xi$  and  $\delta\Gamma$  vanish at spatial infinity, each order  $\xi_i$  and  $\delta\Gamma_i$  may be uniquely decomposed as in (2.4) and (2.3). The transformation behaviour of the components of the latter decomposition becomes more complicated than at linear order, since higher-order gauge transformations mix scalar, vector and tensor quantities in a non-local fashion, as do the higher-order equations of motion. We shall not be concerned with the explicit form of higher-order gauge transformations in this work, as our constructions do not rely on these details and the quantities we consider are manifestly all-order gauge-invariant from the outset.

For the remainder of this work we shall use the following notation motivated by the fact that the space of configurations is an affine space. We decompose a general configuration  $\Gamma$  as  $\Gamma := (\tilde{g}, \tilde{\phi}) := \Gamma_0 + \delta\Gamma$ , where  $\tilde{g} := g + \gamma$ ,  $\tilde{\phi} := \phi + \varphi$  and  $\delta\Gamma := (\gamma, \varphi)$  effectively subsumes linear and higher orders of the perturbation of the background  $\Gamma_0 := (g, \phi)$ . This applies analogously to the components of the decomposition (2.3) of  $\gamma$ .

For later use we recall a useful observation regarding Bardeen potentials. The linear Bardeen potentials  $\Phi$ ,  $\Psi$  and the gauge-invariant scalar field perturbation  $\chi$  in (2.5) have the advantage that they coincide with  $A$ ,  $D$ , and  $\varphi$  respectively in the so-called longitudinal or conformal gauge where the components  $B$  and  $E$  of the metric perturbation  $\gamma$  vanish. This gauge and the definition of the gauge-invariant quantities  $\Phi$ ,  $\Psi$  and  $\chi$  may be extended to higher orders, such that also at higher orders  $\Phi = A$ ,  $\Psi = D$ ,  $\chi = \varphi$  if  $B = E = 0$ , see for example [19].

### 3 All-order gauge-invariant observables on FLRW backgrounds

In this section we provide details on the general construction of all-order gauge-invariant quantities on general and FLRW backgrounds before discussing examples in the next section.

In perturbative Algebraic Quantum Field Theory (pAQFT) — the conceptual framework underlying perturbative quantum gravity in [2] — observables of a field theory are described as functionals of smooth field configurations  $\Gamma = (\tilde{g}, \tilde{\phi})$ . For the purpose of



cosmological perturbation theory, we need the additional restriction that configurations vanish at spatial infinity. In order to be able to operate on the functionals, some regularity is required: the functional derivatives to all orders should exist as distributions of compact support.

Moreover, we restrict our attention to local functionals, i.e. those functionals whose  $n$ -th order functional derivatives are supported on the diagonal of  $M^n$  for every  $n$ . Examples of objects of this form are

$$A_\Gamma(f) := \int_M A_\Gamma f \tag{3.1}$$

where  $A_\Gamma$  is a smooth scalar function which is a polynomial in the derivatives of the field configuration  $\Gamma = (\tilde{g}, \tilde{\phi})$  (i.e.  $A_\Gamma(x) = F(j_x(\Gamma))$  with  $F$  a smooth function on the appropriate jet bundle) and where  $f$  is a smooth compactly supported test density. However, later on in this work we are forced to consider also functionals which violate this locality condition as well as the condition of compact support. The diffeomorphisms  $\chi$  of the spacetime act on configurations via pullback  $\Gamma \mapsto \chi^*\Gamma$ , and candidates for gauge-invariant fields are equivariant in the sense that

$$A_{\chi^*\Gamma} = A_\Gamma \circ \chi . \tag{3.2}$$

Thus in order to exhibit gauge-invariant functionals one has to consider test densities  $f_\Gamma$  which depend on the field configuration  $\Gamma$  such that

$$f_{\chi^*\Gamma} = \chi_* f_\Gamma , \tag{3.3}$$

where  $\chi_*$  is the pushforward of test densities via  $\chi$ .

As described in the Introduction, in the general case we solve the problem by choosing four scalar fields  $X_\Gamma^a$  which constitute a coordinate system  $X_\Gamma$  for a given background  $\Gamma_0$ , and define the  $\Gamma$ -dependent diffeomorphism

$$\alpha_\Gamma = X_\Gamma^{-1} \circ X_{\Gamma_0} . \tag{3.4}$$

For arbitrary test densities  $f$ , we may now consider the  $\Gamma$ -dependent test densities  $f_\Gamma = \alpha_{\Gamma*} f$  in order to obtain gauge-invariant observables  $A_\Gamma(f_\Gamma)$  by means of (3.1). Equivalently, we may directly consider the gauge-invariant field

$$\mathcal{A}_\Gamma = A_\Gamma \circ \alpha_\Gamma . \tag{3.5}$$

Scalars that can be used as coordinates on generic backgrounds  $\Gamma_0$  are e.g. traces of powers of the Ricci operator  $\mathbf{R}$

$$X_\Gamma^a := \text{Tr}(\mathbf{R}^a), \quad a \in \{1, 2, 3, 4\} \tag{3.6}$$

(the operator which maps one forms to one forms and whose components are given in terms of the Ricci tensor  $R_a{}^b$ ). When other (matter) fields are present in the considered model, also these can serve as coordinates, e.g., in the case of a Einstein-Klein-Gordon system, the scalar field  $\tilde{\phi}$ .

In view of renormalisation it is advisable to use coordinates  $X_\Gamma$  which are local functionals of the configuration  $\Gamma$ . As we shall discuss in the following, this does not seem to be possible in cosmological perturbation theory on account of the symmetries of FLRW backgrounds  $\Gamma_0$ .

### 3.1 Perturbative expansion up to second order

To illustrate the general procedure we compute the second order expansion of the gauge-invariant field  $\mathcal{A}_\Gamma$  which was to first order described in the Introduction.

We observe that we have to calculate the functional derivatives of the diffeomorphisms  $\alpha_\Gamma$  with respect to  $\Gamma$ . We use the notation

$$\left\langle \frac{\delta^n}{\delta\Gamma^n} X_\Gamma(\Gamma_0), \delta\Gamma^{\otimes n} \right\rangle =: X_n, \quad \left\langle \frac{\delta^n}{\delta\Gamma^n} \alpha_\Gamma(\Gamma_0), \delta\Gamma^{\otimes n} \right\rangle =: x_n \quad (3.7)$$

and find up to second order

$$x_0^\mu(x) = x^\mu, \quad x_1^\mu = -J_a^\mu X_1^a, \quad (3.8)$$

where  $J$  is the inverse of the Jacobian of  $X_{\Gamma_0}$ , and

$$x_2^\mu = -J_a^\mu X_2^a - J_a^\mu J_b^\nu J_c^\rho \frac{\partial^2 X_0^a}{\partial x^\nu \partial x^\rho} X_1^b X_1^c + 2J_a^\mu J_b^\nu \frac{\partial X_1^a}{\partial x^\nu} X_1^b. \quad (3.9)$$

We use an analogous notation for the Taylor expansions of the fields  $A_\Gamma$  and  $\mathcal{A}_\Gamma$  and find

$$\mathcal{A}_0 = A_0, \quad \mathcal{A}_1 = A_1 + \frac{\partial A_0}{\partial x^\mu} x_1^\mu, \quad (3.10)$$

and

$$\mathcal{A}_2 = A_2 + 2 \frac{\partial A_1}{\partial x^\mu} x_1^\mu + \frac{\partial A_0}{\partial x^\mu} x_2^\mu + \frac{\partial^2 A_0}{\partial x^\mu \partial x^\nu} x_1^\mu x_1^\nu. \quad (3.11)$$

### 3.2 Non-degenerate covariant coordinates on FLRW backgrounds

In order to obtain these expansions we need a 4-tuple of equivariant fields which define a non-degenerate coordinate system on the background  $\Gamma_0$ . This is possible in the generic case, e.g. by using the ansatz (3.6), but creates problems, if the background metric possesses non-trivial symmetries. This applies to the case of FLRW backgrounds  $\Gamma_0$  where only time functions can be constructed out of the background metric  $g$  and the background scalar field  $\phi$ . In the following we present a construction of non-degenerate coordinates which solves the above-mentioned problem at the expense of being non-local, albeit in a controlled way. Note that introducing additional external fields as reference coordinates like in the Brown-Kuchař model [5] is not useful in the context of cosmological perturbation theory because these fields would appear in the final gauge-invariant expressions and thus an interpretation of these in terms of only the fundamental dynamical fields is difficult. The construction we present in the following does involve the asymptotic behavior of the comoving spatial coordinates  $x^i$  of the FLRW spacetime as an external input. Therefore our construction is invariant only under diffeomorphisms which decay sufficiently fast at infinity.

The simplest choice of the time coordinate is provided by the inflaton field itself, so we set

$$X_\Gamma^0 = \tilde{\phi} = \phi + \varphi. \quad (3.12)$$

The construction of the spatial coordinates  $X_\Gamma^i$  needs a bit of preparation. To this end, we consider the unit time-like vector

$$n_\phi = \frac{\tilde{g}^{-1}(d\tilde{\phi}, \cdot)}{\sqrt{|\tilde{g}^{-1}(d\tilde{\phi}, d\tilde{\phi})|}} = \frac{1}{a}(1 - A)\partial_\tau + \frac{1}{a}\left(\partial^i B - \frac{\partial^i \varphi}{\phi'}\right)\partial_i + O(\delta\Gamma^2) \quad (3.13)$$

and the tensor

$$h_\phi = \tilde{g} + \tilde{g}(n_\phi, \cdot) \otimes \tilde{g}(n_\phi, \cdot), \quad (3.14)$$

where  $\partial^i := \partial_i := \partial/\partial x^i$  and  $x^i$  for  $i \in \{1, 2, 3\}$  are comoving spatial coordinates on the FLRW spacetime  $(M, g)$ .  $n_\phi$  is a unit normal on the hypersurfaces of constant  $\tilde{\phi}$  and  $h_\phi$  is the induced metric on these hypersurfaces.

Let  $\Delta_\phi$  denote the Laplacian for  $h_\phi$  and  $G_\phi$  its inverse, which we choose by imposing the boundary condition that the background value of  $G_\phi$  is the Coulomb potential  $G_\Delta$  with suitable factors of the scale factor  $a$ . We define and compute

$$\begin{aligned} \Delta_\phi &:= \Delta_0 + \delta\Delta, & \Delta_0 &:= \frac{\Delta}{a^2}, & \Delta &:= \sum_{i=1}^3 \partial_i^2 \\ \delta\Delta &= -\lambda \left( \frac{2(D + \Delta E)\Delta - (\partial^i(D - \Delta E))\partial_i}{a^2} + \frac{(\Delta\varphi)\partial_\tau + (\partial^i\varphi)(2\partial_\tau + \mathcal{H})\partial_i}{a^2\phi'} \right) + O(\delta\Gamma^2) \\ G_\phi &:= G_0 + \delta G, & G_0 &:= a^2 G_\Delta, & G_\Delta \circ \Delta &= \mathbb{1} \quad \text{on functions that vanish at spatial infinity,} \\ \delta G &=, \sum_{n=1}^{\infty} (-1)^n G_0 \circ (\delta\Delta \circ G_0)^{\circ n} = -G_0 \circ \delta\Delta \circ G_0 + O(\delta\Gamma^2). \end{aligned}$$

Using these objects, we obtain

$$Y_\Gamma^i := (1 - G_\phi \circ \Delta_\phi) x^i = x^i + \partial_i(E + G_\Delta \mathfrak{R}) + O(\delta\Gamma^2), \quad \mathfrak{R} := \frac{\mathcal{H}}{\phi'} \mu. \quad (3.15)$$

We observe that  $Y_\Gamma^i$  are harmonic coordinates for  $\Delta_\phi$  that we have constructed by means of  $x^i$ , i.e. harmonic coordinates for  $\Delta_0$ . The construction of  $Y_\Gamma^i$  makes sense for all configurations  $\Gamma$  which vanish at spatial infinity, but not in general. The restriction to this set of configurations from the outset is natural in the context of cosmological perturbation theory — recall that the decomposition (2.5) is unique only in this case — and should not create problems for the pAQFT framework. For consistency, as noted before, we have to restrict the class of infinitesimal diffeomorphisms we consider in the same manner. In fact, a straightforward computation reveals that the functionals  $Y_\Gamma^i$  are equivariant with respect to all diffeomorphisms  $\chi$  that vanish at spatial infinity

$$\chi^* Y_\Gamma^i = Y_{\chi^*\Gamma}^i + (1 - G_{\chi^*\phi} \circ \Delta_{\chi^*\phi})(\chi^* x^i - x^i) = Y_{\chi^*\Gamma}^i,$$

but not with respect to arbitrary diffeomorphisms. Here  $\Delta_{\chi^*\tilde{\phi}}$  denotes the Laplacian constructed analogous to  $\Delta_\phi$  but with  $\chi^*\tilde{\phi}$  instead of  $\tilde{\phi}$  and  $G_{\chi^*\tilde{\phi}}$  denotes its inverse with the discussed boundary condition. Consequently, the observables constructed by means of the equivariant coordinates (3.12) and (3.15) via (3.5) are gauge-invariant with respect to diffeomorphisms which vanish at spatial infinity. As anticipated, the coordinates  $Y_\Gamma^i$  are non-local, but the non-locality of  $G_\phi$  is relatively harmless since its worst singularity has the same wave front set as the  $\delta$ -function, and renormalisation of expressions involving such objects is well under control, cf. section 5.

The coordinates (3.15) are not entirely well-suited for practical computations because of the fact that the rescaled Mukhanov-Sasaki variable  $\mathfrak{R}$  appears convoluted with the Coulomb potential. In order to remedy this we use a different family of spatial hypersurfaces and a corresponding modification of the spatial Laplacian and its inverse. To this end we consider a number of additional quantities related to the slicing induced by the time-function  $\tilde{\phi}$ : the lapse function  $N_\phi$ , the extrinsic curvature  $K_{\phi,ab}$ , and the spatial Ricci scalar  $R_\phi^{(3)}$  which are defined and computed respectively as

$$N_\phi := |\tilde{g}^{-1}(d\tilde{\phi}, d\tilde{\phi})|^{-1/2} = \frac{a}{\phi'} \left( 1 - \frac{\phi'}{\phi'} + A \right) + O(\delta\Gamma^2), \tag{3.16}$$

$$K_{\phi,ab} := h_{\phi,a}{}^c \nabla_c n_{\phi,b}, \quad K_\phi := K_{\phi,a}{}^a = \frac{3\mathcal{H}}{a} + O(\delta\Gamma), \tag{3.17}$$

$$R_\phi^{(3)} := K_{\phi,ab} K_\phi^{ba} - K_\phi^2 + 2 \left( R_{ab} - \frac{1}{2} R \tilde{g}_{ab} \right) n_\phi^a n_\phi^b = \frac{4}{a^2} \Delta \mathfrak{R} + O(\delta\Gamma^2), \tag{3.18}$$

where  $n_\phi$  and  $h_\phi$  are defined respectively in (3.13) and (3.14). Using these quantities, we define a new time function

$$\mathfrak{t} := \tilde{\phi} - \frac{3N_\phi}{4K_\phi} G_\phi R_\phi^{(3)} = \phi + \frac{\phi'}{\mathcal{H}} D + O(\delta\Gamma^2),$$

If we define the spatial metric  $h_\mathfrak{t}$ , the Laplacian  $\Delta_\mathfrak{t}$  and its inverse  $G_\mathfrak{t}$  in analogy to  $h_\phi$ ,  $\Delta_\phi$  and  $G_\phi$  by replacing  $\tilde{\phi}$  with  $\mathfrak{t}$  we obtain

$$X_\Gamma^i := (1 - G_\mathfrak{t} \circ \Delta_\mathfrak{t}) x^i = x^i + \partial_i E + O(\delta\Gamma^2), \tag{3.19}$$

and the spatial coordinates  $X_\Gamma^i$  share the qualitative properties of the initially defined  $Y_\Gamma^i$ .

#### 4 Examples of gauge-invariant observables at second order

In the previous sections we have developed a principle to construct gauge-invariant perturbative observables from non-gauge-invariant ones. In the following we demonstrate this principle at the example of two observables which are relevant in Cosmology. To this end we use the covariant coordinates (3.12) and (3.19).

Despite the mild non-locality inherent in the covariant spatial coordinates (3.19), we are interested in observables  $A_\Gamma$  which are local functionals of the configuration  $\Gamma$ . The non-locality of  $\mathcal{A}_\Gamma = A_\Gamma \circ \alpha_\Gamma$  implied by the non-locality of  $X_\Gamma^i$  in (3.19) appears only because we consider the local functional  $A_\Gamma$  relative to the non-local functional  $X_\Gamma$ . Since

the background  $\Gamma_0$  depends only on time the same applies to the background value of any local functional  $A_\Gamma$ . Consequently, at first order only the field  $X_\Gamma^0$  (3.12) chosen as time coordinate enters the formula for gauge-invariant fields. At second order also the fields used as spatial coordinates  $X_\Gamma^i$  (3.19) enter the expression.

The inverse  $J$  of the Jacobi matrix of the coordinate transform  $X_{\Gamma_0}$  on the background is

$$J = \begin{pmatrix} \frac{1}{\phi'} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The field dependent shifts from section 3.1 with respect to these coordinates up to second order are

$$x_1^0 = -\frac{\varphi}{\phi'}, \quad x_1^i = -\partial_i E,$$

an

$$\begin{aligned} x_2^0 &= -\frac{\phi''\varphi^2}{(\phi')^3} + \frac{2}{\phi'} \left( \frac{\varphi'\varphi}{\phi'} + (\partial_i\varphi)\partial^i E \right), \\ x_2^i &= \frac{2\varphi}{\phi'} \partial_i E' + 2(\partial^i\partial^j E)\partial_j E - (X_\Gamma^i - x^i - \partial_i E). \end{aligned}$$

Thus, for a field  $A_\Gamma$  whose value on the background depends only on time the contributions up to second order for the gauge-invariant modification  $\mathcal{A}_\Gamma = A_\Gamma \circ \alpha_\Gamma$  are

$$\begin{aligned} \mathcal{A}_0 &= A_0, \quad \mathcal{A}_1 = A_1 - \frac{A'_0\varphi}{\phi'}, \\ \mathcal{A}_2 &= A_2 - \frac{2A'_1\varphi}{\phi'} - 2(\partial_i A_1)\partial^i E + A'_0 \left( -\frac{\phi''\varphi^2}{(\phi')^3} + \frac{2}{\phi'} \left( \frac{\varphi'\varphi}{\phi'} + (\partial_i\varphi)\partial^i E \right) \right) + \frac{A''_0\varphi^2}{(\phi')^2}. \end{aligned}$$

If we were to use the fields  $Y_\Gamma^i$  (3.15) as spatial coordinates rather than the fields  $X_\Gamma^i$  (3.19), then the corresponding expression for  $\mathcal{A}_1$  would remain unchanged whereas  $\mathcal{A}_2$  would change by replacing all occurrences of  $\partial_i E$  by  $\partial_i E + G_\Delta \partial_i \mathfrak{R}$ . This demonstrates the dependence of the gauge-invariant constructions on the chosen covariant coordinate system.

#### 4.1 The lapse function

The Sachs-Wolfe effect is one of the main building blocks of the current understanding of the Cosmic Microwave Background (CMB). A rough estimate of this effect can be obtained using the Tolman idea, see e.g. [20]. Given a spacetime with a (conformal) timelike Killing field  $\kappa$  and a state in equilibrium relative to the  $\kappa$ -flow with absolute temperature  $T$ , an observer with four-velocity  $u \propto \kappa$  measures the temperature  $\tilde{T} = T/N$  with  $N$  denoting the lapse function  $N = \sqrt{|g(\kappa, \kappa)|}$ .

In the context of Cosmology we use the Klein-Gordon field  $\tilde{\phi}$  as a time coordinate and consider the vector

$$\kappa_\phi := N_\phi n_\phi = \frac{1}{\phi'} \partial_\tau + O(\delta\Gamma)$$

with  $N_\phi$ ,  $n_\phi$  defined in (3.16) and (3.13) respectively as an approximate conformal Killing vector — in the sense that  $\mathcal{L}_{\kappa_\phi} \tilde{g} - 2\mathcal{H}/\phi' \tilde{g} = O(\phi'', \delta\Gamma)$ . The corresponding lapse function is  $N_\phi = a/\phi' + O(\delta\Gamma)$ . Its background value is not vanishing and thus it is not automatically gauge-invariant at linear order.

As described in section 3, we may obtain a non-perturbatively gauge-invariant version of the lapse function by setting and computing

$$\begin{aligned} \mathcal{N}_\phi &:= N_\phi \circ \alpha_\Gamma = \frac{a}{\phi'} \left( 1 - \left( (\partial_\tau + \mathcal{H}) \frac{\varphi}{\phi'} - A \right) \right) + O(\delta\Gamma^2) \\ &= \frac{a}{\phi'} \left( 1 - \left( (\partial_\tau + \mathcal{H}) \frac{\chi}{\phi'} - \Phi \right) \right) + O(\delta\Gamma^2), \end{aligned} \tag{4.1}$$

where  $\Phi$  and  $\chi$  are the gauge-invariant fields reviewed in section 2. Using the on-shell identities (2.8), (2.9) and the definition of the Mukhanov-Sasaki field  $\mu$  we can rewrite the linear term as

$$\begin{aligned} \mathcal{N}_{\phi,1} &= \lambda \frac{a}{\phi'} \left( (\partial_\tau + \mathcal{H}) \frac{\chi}{\phi'} - \Phi \right) = \frac{a}{(\phi')^2} \left( \mu' + \left( \frac{\mathcal{H}'}{\mathcal{H}} - \frac{\phi''}{\phi'} \right) \mu \right) \\ &= \frac{2a}{(\phi')^3} \Delta\Phi = -\frac{2a}{(\phi')^3} \Delta\Psi. \end{aligned}$$

Using the quantities introduced in section 3.2, we may extract the Bardeen potential on-shell from  $N_\phi$  as

$$\left[ \frac{1}{2N_\phi^3} G_\phi^2 \Delta_\phi N_\phi \right] \circ \alpha_\Gamma = \Phi + O(\delta\Gamma^2).$$

In fact, one could use the above equation as a covariant, gauge-invariant, all-order (and on shell) definition of  $\Phi$ ; however, we shall refrain from doing so.

In order to display second order expressions in a readable form we omit terms containing the metric perturbation components  $V_i$ ,  $W_j$  and  $T_{ij}$  and use once more the Bardeen potentials  $\Phi$ ,  $\Psi$  and the gauge-invariant scalar field perturbation  $\chi$ . We stress that the particular expressions of these fields at linear and higher order are not needed for the actual computations but just for a compact display of the result. Using this, we arrive at the following second order form of the gauge-invariant lapse function

$$\begin{aligned} \mathcal{N}_{\phi,2} &= \frac{a}{\phi'} \left( -\Phi^2 - 2 \left( \frac{\Phi\chi}{\phi'} \right)' - 2\mathcal{H} \frac{\Phi\chi}{\phi'} + 2 \left( \left( \frac{\chi}{\phi'} \right)' \right)^2 + \left( \frac{\phi''}{\phi'} + 2\mathcal{H} \right) \left( \frac{\chi^2}{\phi'^2} \right)' \right. \\ &\quad \left. + \left( \mathcal{H}^2 + \mathcal{H}' + \frac{\phi'''}{\phi'} + \mathcal{H} \frac{\phi''}{\phi'} - \frac{\phi''^2}{\phi'^2} \right) \frac{\chi^2}{\phi'^2} + \sum_{i=1}^3 \left( \partial_i \left( \frac{\chi}{\phi'} \right) \right)^2 + 2 \frac{\chi}{\phi'} \left( \frac{\chi}{\phi'} \right)'' \right), \end{aligned}$$

where, as before, we use the notation that e.g.  $\Phi = \lambda\Phi_1 + \frac{1}{2}\lambda^2\Phi_2 + O(\lambda^3)$  and omit the second order terms linear in  $\Phi$ ,  $\chi$  displayed already in (4.1).

## 4.2 The spatial curvature

A further observable of interest is the scalar curvature of the spatial metric induced by a particular slicing because for a large class of slicings this quantity vanishes in the background and thus is automatically gauge-invariant at linear order. Moreover, for the slicing

defined by the inflation field it is related to the Mukhanov-Sasaki field  $\mu$  which has a very simple dynamical equation.

We have already discussed the spatial curvature relative to the slicing induced by  $\tilde{\phi}$ . It may be computed as (3.18)

$$R_\phi^{(3)} = \frac{4}{a^2} \Delta \mathfrak{R} + O(\delta\Gamma^2), \quad \mathfrak{R} = \frac{\mathcal{H}}{\phi'} \mu = \frac{\mathcal{H}}{\phi'} \varphi - D.$$

In the literature, the quantity  $\mathfrak{R}$  is usually called the *comoving curvature perturbation*. This is due to the fact that the  $\tilde{\phi}$ -slicing may be equivalently characterised by the condition that

$$T(\tilde{\phi})_{ab} n_\phi^a = -\tilde{g}_{ab} n_\phi^a T(\tilde{\phi})_{cd} n_\phi^c n_\phi^d,$$

i.e. that the energy flux of  $\tilde{\phi}$  is parallel to  $n_\phi$ , where  $T(\tilde{\phi})_{ab}$  is the stress tensor of  $\tilde{\phi}$ .

An alternative slicing considered in the literature is the one defined by the energy density  $\tilde{\rho}$  of  $\tilde{\phi}$

$$\tilde{\rho} := T(\tilde{\phi})_{ab} n_\phi^a n_\phi^b = \rho + \varrho,$$

$$\rho := \frac{(\phi')^2}{2a^2}, \quad \varrho := V^{(1)}(\phi)\varphi + \frac{\phi'(\varphi' - \phi'A)}{a^2} + O(\delta\Gamma^2).$$

The spatial curvature  $R_\rho^{(3)}$  with respect to this slicing, defined in analogy to  $R_\phi^{(3)}$ , reads

$$R_\rho^{(3)} = \frac{4}{a^2} \Delta \zeta + O(\delta\Gamma), \quad \zeta := \frac{\mathcal{H}}{\rho'} \varrho - D,$$

where  $\zeta$  is called *uniform density perturbation* because  $\tilde{\rho}$  is by definition constant on the hypersurfaces in the slicing relative to  $\tilde{\rho}$ . The global sign in the definition of  $\zeta$  is conventional.

As anticipated, the background contributions of  $R_\phi^{(3)}$  and  $R_\rho^{(3)}$  vanish and thus

$$\mathcal{R}_\phi^{(3)} := R_\phi^{(3)} \circ \alpha_\Gamma = R_\phi^{(3)} + O(\delta\Gamma^2), \quad \mathcal{R}_\rho^{(3)} := R_\rho^{(3)} \circ \alpha_\Gamma = R_\rho^{(3)} + O(\delta\Gamma^2),$$

cf. (3.10), (3.11). In order to display the second order contribution to  $\mathcal{R}_\phi^{(3)}$ , we make the simplifications discussed for the lapse function in section 4.1. Proceeding like this, we find

$$\mathcal{R}_{\phi,2}^{(3)} = \frac{8}{a^2} \left( \Delta \left( 2\mathfrak{R}^2 - \frac{\chi}{\phi'} (\partial_\tau + 2\mathcal{H})\mathfrak{R} + \frac{1}{2} \left( \mathcal{H}' + 2\mathcal{H}^2 - \frac{\mathcal{H}\phi''}{\phi'} \right) \left( \frac{\chi}{\phi'} \right)^2 \right) - \frac{5(\partial_i \mathfrak{R})\partial^i \mathfrak{R}}{2} \right). \quad (4.2)$$

We omit the result for  $\mathcal{R}_{\rho,2}^{(3)}$  computed with the coordinate system  $X_\Gamma$  defined in (3.12) and (3.19), because it is rather long due to the “mismatch” between the time coordinate  $\tilde{\phi}$  used in  $X_\Gamma^0$  and the time coordinate  $\tilde{\rho}$  used in the definition of  $R_\rho^{(3)}$ . Clearly, using  $\tilde{\rho}$  as a time coordinate in both aspects we would obtain a second order expression  $\mathcal{R}_{\rho,2}^{(3)}$  which is of the form (4.2) up to the replacements

$$\mathfrak{R} \mapsto \zeta, \quad \phi \mapsto \rho, \quad \chi \mapsto \pi := V^{(1)}(\phi)\chi + \frac{\phi'(\chi' - \phi'\Phi)}{a^2}, \quad (4.3)$$

where  $\pi$  is gauge-invariant with  $\pi = \varrho + O(\delta\Gamma^2)$  in the longitudinal gauge.

On shell and at first order,  $\mu$ , and thus  $\mathfrak{R}$ , are preferred observables because they have canonical equal-time Poisson brackets and thus in the quantized theory they commute at spacelike separations, in contrast to  $\Psi$ ,  $\Phi$  and  $\chi$  [10, 14]. Moreover, again on shell and at first order, one may compute

$$\zeta = \mathfrak{R} - \frac{2\Delta\Phi}{3(\phi')^2} = \mathfrak{R} - \frac{\mathfrak{R}'}{3\mathcal{H}}.$$

Consequently,  $\zeta$  shares the causality properties of  $\mu$  and  $\mathfrak{R}$ .

Apart from the phenomenological relevance of an all-order definition of  $\mathfrak{R}$ ,  $\mu$  and  $\zeta$ , it is interesting on conceptual grounds to investigate whether the causality property of these fields persists at higher orders. To this end, we need a fully covariant and gauge-invariant all-order definition of  $\mathfrak{R}$ ,  $\mu$  and  $\zeta$ . Such a definition may be given by means of covariant quantities introduced in section 3.2:

$$\begin{aligned} \left[\frac{1}{4}G_\phi R_\phi^{(3)}\right] \circ \alpha_\Gamma &= \frac{a^2}{4}G_\Delta \mathcal{R}_\phi^{(3)} - a^2 G_\Delta \delta\Delta\mathfrak{R} + O(\delta\Gamma^3) \\ &= \mathcal{H}\frac{\chi}{\phi'} - \Psi + \mathfrak{R}^2 - 2\mathcal{H}\frac{\chi}{\phi'}\mathfrak{R} + \frac{1}{2}\left(\mathcal{H}' + 2\mathcal{H}^2 - \frac{\mathcal{H}\phi''}{\phi'}\right)\left(\frac{\chi}{\phi'}\right)^2 \\ &\quad + G_\Delta\left(\frac{(\partial_i\mathfrak{R})\partial^i\mathfrak{R}}{2}\right) + O(\delta\Gamma^3), \end{aligned} \tag{4.4}$$

$$\left[\frac{3N_\phi}{4K_\phi}G_\phi R_\phi^{(3)}\right] \circ \alpha_\Gamma = \mu + O(\delta\Gamma^2), \quad \left[\frac{1}{4}G_\phi R_\rho^{(3)}\right] \circ \alpha_\Gamma = \zeta + O(\delta\Gamma^2). \tag{4.5}$$

In (4.4) we wrote the  $O(\delta\Gamma)$  term as  $\mathcal{H}\chi/\phi' - \Psi$  instead of  $\mathfrak{R}$  because the fields  $\chi$ ,  $\Psi$  are defined in such a way that they are invariant also with respect to second order gauge transformations (cf. the end of section 2), whereas  $\mathfrak{R} = \mathcal{H}\varphi/\phi' - D$  is only gauge-invariant up to the first order.

In analogy to our discussion of  $R_\rho^{(3)}$ , using  $\tilde{\rho}$  rather than  $\tilde{\phi}$  both as the time coordinate  $X_\Gamma^0$  and as the time function defining a foliation of spacetime, we obtain a higher order definition of  $\zeta$  which is of the form (4.4) up to the replacements in (4.3) (whereby a second order generalisation of  $\pi$ , which can be constructed in analogy to the second order Bardeen potentials, is needed).

In the literature, several possible second order gauge-invariant corrections to  $\mathfrak{R}$  are considered. One often encounters constructions where in a gauge with  $\varphi = 0$  (or  $D = 0$ ), the second order corrections to  $\mathfrak{R}$  vanish — at least in situations where spatial derivatives can be neglected in comparison to temporal ones, see e.g. [18, 19, 25, 31]. In fact  $\mathfrak{R}$  is often defined by the condition  $\mathfrak{R} = -D$  in a gauge where  $\varphi = 0$ . A quick analysis reveals that this is not the case in our construction (4.4). In [31] it is argued that expressions for  $\mathfrak{R}$  valid up to second order that are not of this form, e.g. the one in [1], are potentially physically ill-behaved because they are not conserved on “super-Hubble scales”. Here, conservation of a function  $f(\tau, \vec{x})$  on “super-Hubble scales” means that the Fourier transform  $\hat{f}(\tau, \vec{k})$  of  $f$  with respect to  $\vec{x}$  satisfies  $\partial_\tau \hat{f}(\tau, \vec{k}) = O(|\vec{k}|/\mathcal{H})$ . This property, whose relevance is explained e.g. in [18, 31], usually holds only on-shell. It would be interesting to check whether our result for  $\mathfrak{R}$  as given in (4.4) (and the analogous result for  $\zeta$ ) is conserved in this sense; however, this is beyond the scope of the present work.



## 5 Quantization

In the previous sections we have prepared the ground for an all-order perturbative quantization of the Einstein-Klein-Gordon system on FLRW backgrounds, i.e. for a conceptually clear higher-order generalisation of quantized cosmological perturbation theory. In this section we would like to sketch the steps necessary for a full construction of the quantum theory. A detailed account will be given in a future work [3].

### 5.1 BRST quantization

It is known that a direct quantization of non-linear gauge-invariant observables in a theory with local gauge symmetries is difficult. The standard way out is to perform a gauge fixing in the sense of the BRST method, or more generally, the BV formalism, as treated in [11, 12, 15]. There one adds a Fermionic vector field  $c^\mu$  (the ghost field), which describes the infinitesimal gauge transformations, auxiliary scalar fields  $b_\mu, \bar{c}_\mu$ , where  $b_\mu$  (the Nakanishi-Lautrup field) is Bosonic and  $\bar{c}_\mu$  (antighost) is Fermionic,  $\mu = 0, \dots, 3$ . Infinitesimal coordinate transformations are described by the BRST operator  $s$ , which acts on scalar local functionals  $A$  of the metric, the inflaton and the  $b$  fields by

$$s(A)(x) = c^\mu(x)\partial_\mu A(x),$$

on the components of the ghost field by

$$s(c^\mu)(x) = c^\nu(x)\partial_\nu c^\mu(x),$$

on antighosts by

$$s(\bar{c}_\mu)(x) = ib_\mu(x) - c^\nu(x)\partial_\nu \bar{c}_\mu(x)$$

and satisfies on products the graded Leibniz rule so that  $s^2 = 0$ . One can characterise the classical observables as functionals in the kernel of  $s$  modulo those in the image of  $s$  (i.e. classical observables belong to the 0-th cohomology group of  $s$ ).

The field equations for the extended system are the usual field equation for  $\tilde{\phi}$  as well as

$$\begin{aligned} R_{\mu\nu} &= T(\tilde{\phi})_{\mu\nu} - \frac{1}{2}T(\tilde{\phi})\tilde{g}_{\mu\nu} + s(i\partial_{(\mu}\bar{c}_{\nu)}) \\ \square_{\tilde{g}}c^\mu &= 0 \\ \square_{\tilde{g}}\bar{c}_\mu &= 0 \\ |\det\tilde{g}|^{-\frac{1}{2}}\partial_\mu|\det\tilde{g}|^{\frac{1}{2}}\tilde{g}^{\mu\nu} &= \kappa^{\mu\nu}b_\mu. \end{aligned}$$

Here  $\kappa$  is a non-degenerate fixed tensor.

The quantization of the extended system now proceeds largely analogous to the pure gravity treatment in [2]. The main idea is to use deformation quantization to deform the algebra of functionals as well as the BRST operator  $s$ . The observables we discussed in the previous sections are invariant under diffeomorphisms whose action decays sufficiently fast at spatial infinity. Since the BRST transformation on these fields are just infinitesimal diffeomorphisms the only point to check is whether they have the right decay properties. This can be verified by direct inspection of the correlation functions of the ghost fields. Hence, the quantized versions of the gauge-invariant fields discussed in the previous sections are elements of the cohomology of the quantized BRST operator  $s$ .

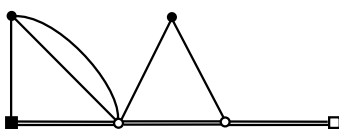
## 5.2 Renormalisation

A conceptual and technical difference to the case of generic backgrounds treated in [2] arises because we were forced to introduce a mild non-locality for observables via the non-local spatial coordinates  $X_\Gamma^i$  (3.19). Despite of this fact, the interaction Lagrangean remains a local field. Therefore only the external vertices in a graph expansion of perturbation theory are affected. In [2] renormalisation was treated in the Epstein-Glaser framework which is initially only suitable for local functionals. As we have to deal with non-local expressions, we need to extend this framework from local quantities to certain non-local ones. The first step in this procedure is to expand the non-local coordinates in terms of its local components. To this end, recall that

$$X_\Gamma^i = (1 - G_\phi \Delta_\phi) x^i = \sum_{k=0}^{\infty} (-G_0 \delta\Delta)^k x^i,$$

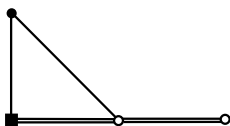
where  $\Delta_\phi = \Delta_0 + \delta\Delta$  is the Laplacian relative to the  $\tilde{\phi}$ -slicing and  $G_\phi = \sum_{k=0}^{\infty} (-G_0 \delta\Delta)^k G_0$  is its Green's function for suitable boundary conditions, cf. section 3.2. Notice that the previous expression has the form of the perturbative expansion of a local field induced by a quadratic interaction Lagrangean and the most severe divergence present in  $G_0$  is local. Our gauge-invariant observables can be expanded as Taylor series in  $X_\Gamma^a$ . Inserting the series written above for the coordinates into the perturbative expansion of the observables one obtains an expansion with only local vertices and two types of propagators. Hence, the renormalisation can be performed as in the standard case, in particular the arising non-localities originate from local counter terms at lower order.

As an example, let us consider the time ordered product of a local observable expanded as described above with several interaction Lagrangeans. In the graphical expansion one finds as a simple but typical contribution the graph



The vertices  $\bullet$  denote the interaction Lagrangean, the vertex  $\blacksquare$  the observable under consideration and the vertices  $\circ$   $\delta\Delta$  terms in the expansion of coordinates. The external vertex  $\square$  represents the action of  $\delta\Delta$  on the background coordinate  $x^i$ .

To see that such graphs can be renormalised, consider the simplest divergent case, namely



The kernel of  $G_0$  considered as a distribution on  $M^2$  is of the form

$$G_0(x, y) = c(\tau_x) \delta(\tau_x, \tau_y) \frac{1}{|\vec{x} - \vec{y}|}$$

with a smooth function  $c$ . The wave front set of  $G_0(x, y)$  at coinciding points is the one of  $\delta(x, y)$  and its scaling degree there is 2. For non-coinciding spatial coordinates no divergences occur. The vertex operators  $\delta\Delta$  are differential operators of at most second order. By direct inspection we thus see that the only singularity of the loop in the above example is at the total diagonal and by power counting we find that the degree of divergence of this loop is at most 2, so that the appropriately renormalised expression is unique up to at most two derivatives of  $\delta$  distributions of the three loop vertices. In general the degree of divergence of a loop containing “internal” propagators may be higher or lower than in the above example depending on the number of Feynman propagators appearing in the loop; the same applies to the renormalisation freedom of general loops.

These arguments indicate that the new types of graphs do not create new problems in the UV regime. We briefly sketch why we do not expect additional IR problems. We have already pointed out that our setup is only meaningful if we restrict the admissible classical configurations to those which vanish at spatial infinity. By consistency we need the same behaviour for the correlation functions of the quantized theory, in particular for the Feynman propagators of the linearised model. Provided quantum states (or more general Hadamard parametrices) with this property exist — this is not obvious and needs to be proven — we expect that the integrals corresponding to the “internal” vertices will converge.

The remaining problem is to deal with the combinatorics of such graphs and ensure that the renormalisation can be performed systematically order by order. This can be done by a slight generalisation of the standard framework and will be discussed in detail in our forthcoming paper [3]. In the same publication we will also prove the validity of Ward identities analogous to the ones proven by Hollands for the Yang-Mills theory [15].

## 6 Conclusions

We described how cosmological perturbation theory may be derived from a full theory of perturbative quantum gravity. This demonstrates that perturbative quantum gravity can already be tested by present observations. Moreover, on a more practical side, our definition of gauge-invariant observables provides a conceptually simple way of extending the observables which are relevant for the interpretation of cosmological observations to arbitrary high orders.

However, even in linear order, our discussion clarifies the choice of good observables, as we have indicated at the example of the lapse function  $N_\phi$  with respect to the spatial hypersurfaces of constant inflaton field. Initially  $N_\phi$  is not gauge-invariant, but our construction yields a gauge-invariant version which at linear order and on shell may be expressed in terms of the Bardeen potential  $\Phi$  that is related to the temperature fluctuations of the CMB via the Sachs-Wolfe effect.

We computed examples of gauge-invariant observables beyond linear order and found a second-order expression for the comoving curvature perturbation which seems to differ from constructions in other works. As in the literature there is some debate about whether some constructions are physically well-behaved, see. e.g. [31], it would be interesting to

investigate the physical properties of our result, even though it is clear from the outset that it has a transparent geometric interpretation.

Finally we have sketched the details of the quantization of the Einstein-Klein-Gordon system on cosmological backgrounds beyond linear order. We believe that the strategy outlined here leads to a full renormalised all-order theory of cosmological perturbations by means of which higher order corrections to standard results in cosmology may be computed.

## Acknowledgments

K.F., N.P. and K.R. would like to thank the Erwin Schrödinger Institute in Vienna, where part of the research reported here was carried out, for the kind hospitality.

**Open Access.** This article is distributed under the terms of the Creative Commons Attribution License ([CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/)), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

## References

- [1] V. Acquaviva, N. Bartolo, S. Matarrese and A. Riotto, *Second order cosmological perturbations from inflation*, *Nucl. Phys. B* **667** (2003) 119 [[astro-ph/0209156](#)] [[INSPIRE](#)].
- [2] R. Brunetti, K. Fredenhagen and K. Rejzner, *Quantum gravity from the point of view of locally covariant quantum field theory*, *Commun. Math. Phys.* **345** (2016) 741 [[arXiv:1306.1058](#)] [[INSPIRE](#)].
- [3] R. Brunetti, K. Fredenhagen, T.-P. Hack, N. Pinamonti and K. Rejzner, in preparation.
- [4] N. Bartolo, S. Matarrese and A. Riotto, *CMB Anisotropies at Second-Order. 2. Analytical Approach*, *JCAP* **01** (2007) 019 [[astro-ph/0610110](#)] [[INSPIRE](#)].
- [5] J.D. Brown and K.V. Kuchar, *Dust as a standard of space and time in canonical quantum gravity*, *Phys. Rev. D* **51** (1995) 5600 [[gr-qc/9409001](#)] [[INSPIRE](#)].
- [6] M. Bruni, S. Matarrese, S. Mollerach and S. Sonego, *Perturbations of space-time: Gauge transformations and gauge invariance at second order and beyond*, *Class. Quant. Grav.* **14** (1997) 2585 [[gr-qc/9609040](#)] [[INSPIRE](#)].
- [7] B. Dittrich, *Partial and complete observables for canonical general relativity*, *Class. Quant. Grav.* **23** (2006) 6155 [[gr-qc/0507106](#)] [[INSPIRE](#)].
- [8] B. Dittrich and J. Tambornino, *Gauge invariant perturbations around symmetry reduced sectors of general relativity: Applications to cosmology*, *Class. Quant. Grav.* **24** (2007) 4543 [[gr-qc/0702093](#)] [[INSPIRE](#)].
- [9] M. Duetsch and K. Fredenhagen, *Perturbative algebraic field theory and deformation quantization*, *Submitted to: Fields Inst. Commun.* (2000) 151 [[hep-th/0101079](#)] [[INSPIRE](#)].
- [10] B. Eltzner, *Quantization of Perturbations in Inflation*, [arXiv:1302.5358](#) [[INSPIRE](#)].
- [11] K. Fredenhagen and K. Rejzner, *Batalin-Vilkovisky formalism in the functional approach to classical field theory*, *Commun. Math. Phys.* **314** (2012) 93 [[arXiv:1101.5112](#)] [[INSPIRE](#)].
- [12] K. Fredenhagen and K. Rejzner, *Batalin-Vilkovisky formalism in perturbative algebraic quantum field theory*, *Commun. Math. Phys.* **317** (2013) 697 [[arXiv:1110.5232](#)] [[INSPIRE](#)].

- [13] K. Fredenhagen, R. Rejzner *Perturbative Construction of Models of Algebraic Quantum Field Theory*, in *Advances in algebraic quantum field theory*, R. Brunetti, C. Dappiaggi, K. Fredenhagen and J. Yngvason eds., Springer (2015), [arXiv:1503.07814](#).
- [14] T.-P. Hack, *Quantization of the linearized Einstein-Klein-Gordon system on arbitrary backgrounds and the special case of perturbations in inflation*, *Class. Quant. Grav.* **31** (2014) 215004 [[arXiv:1403.3957](#)] [[INSPIRE](#)].
- [15] S. Hollands, *Renormalized Quantum Yang-Mills Fields in Curved Spacetime*, *Rev. Math. Phys.* **20** (2008) 1033 [[arXiv:0705.3340](#)] [[INSPIRE](#)].
- [16] I. Khavkine, *Local and gauge invariant observables in gravity*, *Class. Quant. Grav.* **32** (2015) 185019 [[arXiv:1503.03754](#)] [[INSPIRE](#)].
- [17] D. Langlois and F. Vernizzi, *A geometrical approach to nonlinear perturbations in relativistic cosmology*, *Class. Quant. Grav.* **27** (2010) 124007 [[arXiv:1003.3270](#)] [[INSPIRE](#)].
- [18] J.M. Maldacena, *Non-Gaussian features of primordial fluctuations in single field inflationary models*, *JHEP* **05** (2003) 013 [[astro-ph/0210603](#)] [[INSPIRE](#)].
- [19] K.A. Malik and D. Wands, *Cosmological perturbations*, *Phys. Rept.* **475** (2009) 1 [[arXiv:0809.4944](#)] [[INSPIRE](#)].
- [20] V.F. Mukhanov, *CMB-slow, or how to estimate cosmological parameters by hand*, *Int. J. Theor. Phys.* **43** (2004) 623 [[astro-ph/0303072](#)] [[INSPIRE](#)].
- [21] K. Nakamura, *Second-order gauge invariant cosmological perturbation theory: Einstein equations in terms of gauge invariant variables*, *Prog. Theor. Phys.* **117** (2007) 17 [[gr-qc/0605108](#)] [[INSPIRE](#)].
- [22] K. Nakamura, *Recursive structure in the definitions of gauge-invariant variables for any order perturbations*, *Class. Quant. Grav.* **31** (2014) 135013 [[arXiv:1403.1004](#)] [[INSPIRE](#)].
- [23] H. Noh and J.-c. Hwang, *Second-order perturbations of the Friedmann world model*, *Phys. Rev. D* **69** (2004) 104011 [[INSPIRE](#)].
- [24] J.-c. Hwang and H. Noh, *Fully nonlinear and exact perturbations of the Friedmann world model*, *Mon. Not. Roy. Astron. Soc.* **433** (2013) 3472 [[arXiv:1207.0264](#)] [[INSPIRE](#)].
- [25] T. Prokopec and J. Weenink, *Uniqueness of the gauge invariant action for cosmological perturbations*, *JCAP* **12** (2012) 031 [[arXiv:1209.1701](#)] [[INSPIRE](#)].
- [26] M. Reuter, *Nonperturbative evolution equation for quantum gravity*, *Phys. Rev. D* **57** (1998) 971 [[hep-th/9605030](#)] [[INSPIRE](#)].
- [27] M. Reuter and F. Saueressig, *Renormalization group flow of quantum gravity in the Einstein-Hilbert truncation*, *Phys. Rev. D* **65** (2002) 065016 [[hep-th/0110054](#)] [[INSPIRE](#)].
- [28] C. Rovelli, *Partial observables*, *Phys. Rev. D* **65** (2002) 124013 [[gr-qc/0110035](#)] [[INSPIRE](#)].
- [29] S. Sonego and M. Bruni, *Gauge dependence in the theory of nonlinear space-time perturbations*, *Commun. Math. Phys.* **193** (1998) 209 [[gr-qc/9708068](#)] [[INSPIRE](#)].
- [30] T. Thiemann, *Reduced phase space quantization and Dirac observables*, *Class. Quant. Grav.* **23** (2006) 1163 [[gr-qc/0411031](#)] [[INSPIRE](#)].
- [31] F. Vernizzi, *On the conservation of second-order cosmological perturbations in a scalar field dominated Universe*, *Phys. Rev. D* **71** (2005) 061301 [[astro-ph/0411463](#)] [[INSPIRE](#)].