

Gauge-Invariant Perturbations in Hybrid Quantum Cosmology

Laura Castelló Gomar,^a Mercedes Martín-Benito^b and Guillermo A. Mena Marugán^a

^aInstituto de Estructura de la Materia, IEM-CSIC,
Serrano 121, 28006 Madrid, Spain

^bRadboud University Nijmegen, Institute for Mathematics, Astrophysics and Particle Physics,
Heyendaalseweg 135, NL-6525 AJ Nijmegen, The Netherlands

E-mail: laura.castello@iem.cfmac.csic.es, m.martin@hef.ru.nl,
mena@iem.cfmac.csic.es

Abstract. We consider cosmological perturbations around homogeneous and isotropic space-times minimally coupled to a scalar field and present a formulation which is designed to preserve covariance. We truncate the action at quadratic perturbative order and particularize our analysis to flat compact spatial sections and a field potential given by a mass term, although the formalism can be extended to other topologies and potentials. The perturbations are described in terms of Mukhanov-Sasaki gauge invariants, linear perturbative constraints, and variables canonically conjugate to them. This set is completed into a canonical one for the entire system, including the homogeneous degrees of freedom. We find the global Hamiltonian constraint of the model, in which the contribution of the homogeneous sector is corrected with a term quadratic in the perturbations, that can be identified as the Mukhanov-Sasaki Hamiltonian in our formulation. We then adopt a hybrid approach to quantize the model, combining a quantum representation of the homogeneous sector with a more standard field quantization of the perturbations. Covariance is guaranteed in this approach inasmuch as no gauge fixing is adopted. Next, we adopt a Born-Oppenheimer ansatz for physical states and show how to obtain a Schrödinger-like equation for the quantum evolution of the perturbations. This evolution is governed by the Mukhanov-Sasaki Hamiltonian, with the dependence on the homogeneous geometry evaluated at quantum expectation values, and with a time parameter defined also in terms of suitable expectation values on that geometry. Finally, we derive effective equations for the dynamics of the Mukhanov-Sasaki gauge invariants, that include quantum contributions, but have the same ultraviolet limit as the classical equations. They provide the master equation to extract predictions about the power spectrum of primordial scalar perturbations.

Keywords: Quantum cosmology, cosmological perturbation theory

Contents

1	Introduction	1
2	The perturbed flat FRW model	5
3	Perturbations in terms of gauge-invariant variables	8
3.1	Canonical transformation for the perturbations	8
3.2	Redefinition of the MS momenta	10
3.3	Inversion of the canonical transformation	11
4	Full canonical set in terms of gauge-invariant variables	12
4.1	Canonical transformation of the zero-modes	12
4.2	Hamiltonian in terms of gauge-invariant variables	13
5	Hybrid quantization and Born-Oppenheimer ansatz	15
5.1	Fock representation for the perturbations	16
5.2	Quantum representation of the constraints	17
5.3	Born-Oppenheimer ansatz	18
5.4	Effective equations for the Mukhanov-Sasaki variables	21
6	LQC representation for the homogeneous sector	23
7	Conclusions	25
A	Spacetime metric and scalar field in the gauge-invariant formulation	28
A.1	Zero-modes	28
A.2	Lapse function and shift vector	29
B	Negligible second derivatives: self-consistency of the approximation	30

1 Introduction

The treatment of linear perturbations has become a prolific field in modern cosmology. The advent of inflationary universe models [1] generated great interest in cosmological perturbations and gave rise to a lot of activity in this area. Beyond its early success, the study of perturbed systems in general relativity requires great care because one must deal with the gauge freedom inherent to the theory, which affects the description of the perturbations. Once this problem was realized, most efforts focused on getting a manifestly gauge-invariant formalism.

Nowadays, cosmology is entering a golden age owing to the recent observational progress, which has opened new windows to test the predictions of theoretical models [2]. The latest observations are providing us with increasingly accurate data of cosmological phenomena and, for the first time, it seems possible for astrophysics to think about finding traces of the quantum geometric structure of the early history of the universe [3]. For this reason, the ultimate hope of the community of physicists working in the quantization of gravity is to develop a quantum theory capable of leading to testable predictions. In order to fully capture

the quantum nature of spacetime, this theory must involve simultaneously both the geometry and the perturbations, with interplay between them.

Over the last decades, the theory of cosmological perturbations [4], combined with the inflationary paradigm [1], has emerged as the framework which conciliates the theoretical models of the early universe with observations, since it provides a good approximation to the anisotropies of the cosmic microwave background (CMB) and explains quite satisfactorily the formation of structures at large scales [5]. The study of the CMB is a powerful tool for understanding the universe in its origins. It supports the approximation that the observed region is homogeneous and isotropic in a suitable average (demonstrated under certain theoretical assumptions [6]). However, this leads to questions about how the anisotropies and cosmological structures formed and developed.

The pioneer work in the analysis of perturbations around classical Friedmann-Robertson-Walker (FRW) cosmologies¹ is due to Lifshitz [7], as a first modelization of the universe considered at a large scale. Nonetheless, it was relatively soon noticed that this analysis had been carried out with a specific gauge choice, and hence it did not address the gauge freedom satisfactorily. This gauge dependence in the description of the perturbations has caused many controversies because keeping track of the gauge modes can get cumbersome, something which makes difficult the extraction of the physically meaningful degrees of freedom. An attempt to provide a covariant treatment of the perturbations was made by Hawking [8], but this work did not resolve totally the gauge ambiguities. It was completed later by Olson [9] for the case of an isentropic perfect fluid in a spatially flat spacetime. However, it was Bardeen who first constructed a truly gauge-invariant formalism (originally for a perfect fluid), which mixes the perturbations of the matter with the perturbations of the four-dimensional metric [10]. This work was followed by many other contributions [11, 12]. Likewise, Mukhanov, based on Sasaki's investigations [12], proposed some gauge-invariant field-like variables for the case of a scalar field on a spatially flat FRW background, directly related to the co-moving curvature perturbations [13]. Mukhanov expanded the action for the gravitational and scalar fields up to second order in the perturbations and introduced a gauge-invariant field that completely characterized those perturbations and allowed one to rewrite their action exclusively in terms of it when the background equations were employed.

In a standard analysis of primordial fluctuations, one studies the perturbations within the scheme of Quantum Field Theory (QFT) in a classical and fixed curved spacetime. From this viewpoint, the nature of these primordial perturbations is in fact quantum and it is rather general to assume that they started out at a very early time as fluctuations with a small amplitude and gradually grew in time as consequence of gravitational instabilities. For doing this, essentially, one represents the perturbations by quantum fields and considers that they were initially propagating on a given geometry, such as a de Sitter spacetime, which describes rather well the inflationary stage of the universe [14]. Actually, we already know from the CMB data that fluctuations with amplitudes of 10^{-5} are sufficient to reproduce the cosmic structures observed today. Despite how well this treatment is able to explain the present observations, the challenge for quantum cosmology is to build a formalism which includes the homogeneous background and the inhomogeneous perturbations and proves to be potentially predictive, in order to elucidate whether the relics of the quantum fluctuations of the early universe may encode information about the quantum character of the spacetime geometry itself.

¹These cosmologies are also called Friedmann-Lemaître-Robertson-Walker cosmologies by many authors.

A canonical approach to cosmological perturbations in which the background and the inhomogeneities (both in the geometry and in a matter scalar field) were treated together quantum mechanically, although fixing the gauge, was developed by Halliwell and Hawking [15], in the context of closed FRW universes. They expanded the action up to second order in the perturbations and then built the corresponding Hamiltonian. Shortly after, Shirai and Wada [16] reformulated this formalism, isolating the quantum dependence on gauge invariants. Actually, they introduced not only a canonical transformation for the perturbations but modified the variables of the background by terms quadratic in those perturbations. Even so, physical states still depended on perturbative variables that were not gauge invariants (although in a way that was completely determined), and the introduced canonical transformations were only defined in a semiclassical limit. Therefore, they depended on the solution to the effective equations that was considered in each case. Later, Langlois [17] tried to refine and clarify this procedure by using techniques inspired by Hamilton-Jacobi theory in order to obtain the gauge-invariant perturbations. Nevertheless, he did not include transformations of the background and, besides, obviated the explicit form in which the Hamiltonian constraint depends on non gauge-invariant variables, because this should be physically irrelevant. More recently, Pinto-Neto and collaborators [18] proposed another approach by means of canonical transformations which involve the perturbations and the background and, in particular, include the Mukhanov-Sasaki (MS) variables. More specifically, they considered the case of a perfect fluid and of a scalar field, and reformulated the system so that the global Hamiltonian constraint depends only on the gauge invariants and the new background variables. In this reformulation, the equations of motion of the background were not used, and the lapse function was redefined. Nonetheless, second-class constraints appeared in this reformulation procedure which had to be eliminated by reduction and the introduction of Dirac brackets, a step that obscures the full gauge invariance of the ultimate system and the role in it of the perturbative constraints.

In the last years, the canonical quantization of cosmological perturbations has received as well a lot of attention within the framework of Loop Quantum Cosmology (LQC). LQC [19] addresses the quantization of cosmological systems using the ideas and techniques of Loop Quantum Gravity (LQG) [20], a non-perturbative and background independent program for the quantization of general relativity that provides one of the most promising approaches for a quantum theory of the gravitational interaction. Initially, LQC was applied successfully to homogeneous models in cosmology, leading to a consistent quantization of the FRW spacetime in which the Big Bang singularity is replaced with a quantum bounce, namely the Big Bounce [21, 22]. Clearly, the limitation of homogeneity is a restriction that must be surpassed when one is interested in studying more realistic scenarios. Therefore, it is natural to try and go beyond this simplification and incorporate small fluctuations in the quantum treatment of the geometry and the matter content.

The loop quantization of the FRW background has been combined with a Fock quantization of the perturbations in a so-called *hybrid* approach in which the whole system that describes the perturbed cosmologies is quantized with canonical methods [23–25]. In particular, in this formalism the perturbed spacetime geometry can be treated as a fully quantum entity. This hybrid approach was originally developed in the linearly polarized Gowdy models [26]. In this case, the inhomogeneities can be dealt with exactly, rather as perturbatively. An alternative proposal, put forward by Ashtekar *et al.* [27], investigates the propagation of the perturbations in the *dressed* FRW geometry obtained by incorporating the quantum effects of LQC on the background. Consequently, in this analysis, known as the *dressed metric* formal-

ism, one loses a full quantum description of the geometry in the perturbed system. Both, in the *hybrid* and the *dressed metric* approaches in LQC, the perturbations have been expressed in terms of gauge invariants, though just after eliminating degrees of freedom in a reduction that casts doubt on the covariance of the system. Finally, a third route to analyze the effects of LQG in cosmological perturbations consists in assuming corrections in the quantum constraints arising from the use of holonomies and the regularization of the inverse of the volume operator [28, 29]. Demanding the closure of the modified algebra of constraints, one can deduce the form of those quantum corrections and the corresponding effective field equations for the propagation of the perturbations. This strategy is intrinsically covariant, although it still rests on a series of hypotheses about, e.g., the form of the possible loop quantum corrections, the use of local expansions for non-local quantities, or the independence of the results on the existence of superselection sectors.

In this work, we will provide a canonical formalism for perturbed flat FRW spacetimes with a matter scalar field which is specially designed to preserve covariance at the considered perturbative level. In particular, the perturbations will be described by the MS gauge invariants, by combinations of the perturbative constraints, and by variables canonically conjugate to them. For completeness and self-consistency in the presentation, we will detail the change to this description from a classical formulation similar to that of Halliwell and Hawking [15] for the case of flat spatial topology. This change will be obtained by means of a canonical transformation which will be completed in order to include as well the background variables. In addition, we will discuss how to quantize the resulting model using generalized hybrid techniques, namely, combining any given quantization of the background with a quantum theory for the perturbations (assuming the compatibility of both quantizations in the whole system). In this framework, we will show how to extract quantum field equations for the MS variables without any gauge fixing. Furthermore, this will be done without appealing to semiclassical approximations, like in Refs. [15, 16], nor a Bohm-De Broglie scheme, like in Ref. [18]. Instead, we will employ a kind of Born-Oppenheimer ansatz and discuss the validity of its application. In this derivation, we will not make use of any background equation of motion or constraint, neither at a classical, effective, nor quantum level. Finally, we will particularize our discussion to the specific case of a hybrid quantization in loop cosmology.

In this sense, our analysis extends other treatments like that of Ref. [25], where the gauge freedom associated with the perturbative constraints was fixed, while we do not make any such classical gauge fixing here. For that, we will need to abelianize the algebra of constraints in our perturbed FRW system. Let us finally mention that, in the context of LQG as well, there have been some attempts to develop a manifestly gauge-invariant perturbation theory in the canonical framework by constructing approximate complete observables [30], adopting an approach different from ours.

The plan of the rest of the paper is as follows. Sec. 2 contains the notation and the description of the classical perturbed system. In Sec. 3 we will introduce the set of transformations for the inhomogeneous modes that leads to gauge invariants, perturbative constraints, and their conjugate variables. This will be extended in Sec. 4 to a transformation for the whole perturbed system that preserves the canonical structure. In doing this, we will have to include quadratic perturbative corrections to the background variables. In addition, in this section we will study how these changes modify the zero-mode of the Hamiltonian constraint. In Sec. 5 we will explain how to proceed to a (generalized) hybrid quantization of the total system formed by the background and the perturbations. Adopting then a Born-Oppenheimer ansatz, we will deduce a quantum equation for the propagation of the

MS variables, discussing its range of validity. Besides, we will derive an effective equation governing the evolution of the corresponding classical variables. This formalism will be applied to LQC in Sec. 6. We will conclude and summarize in Sec. 7. Finally, we will include two appendices with extra details of the calculations.

2 The perturbed flat FRW model

In this section we review the classical description of the FRW model with a minimally coupled scalar field and with inhomogeneous perturbations (see e.g. Ref. [15]). We consider the case of flat spatial topology which (in order to avoid technical problems in the ultraviolet regime, as we will comment later on) we assume to be compact, namely that of a three-torus. This flat case describes the physically relevant situation in cosmology, since it is in agreement with observations, as long as the compactification scales of the three-torus are large compared to the radius of our observable universe. The matter content of the model is provided by a scalar field Φ minimally coupled to the geometry and subject to a potential term. Although the discussion can be carried out for generic potentials, we will particularize our analysis to the case of a mass contribution for simplicity, explaining just how the main formulas of our study generalize for any other potential of the scalar field. On the other hand, as it is well known, scalar, vector, and tensor perturbations decouple from each other at leading order in the perturbative treatment, and may be considered independently. Actually, for scalar fields, vector perturbations are pure gauge and therefore do not contain any physical degree of freedom. On the contrary, tensor perturbations are in fact gauge invariant, and hence their treatment is relatively simple, at least with respect to the issues of covariance of the perturbative formulation that we want to address. Besides, from an observational point of view, scalar perturbations are the most interesting ones, since they are ultimately responsible for the anisotropies measured in the temperature of the CMB. We will thus focus our discussion on scalar perturbations. This model was also considered in Refs. [24, 25] but, unlike in this work, the system was reduced by partially fixing the gauge freedom in those references. Here, we follow conventions and notation similar to the ones of those works, and we refer the reader to them for further details.

Adopting a canonical 3+1 decomposition, we parameterize the four-metric in terms of the three-metric h_{ij} induced on the sections of constant time t , a lapse function N , and a shift vector N^i (or co-vector N_i). Spatial indices i, j run from 1 to 3. For the unperturbed FRW spacetime, these metric functions are completely characterized by a homogeneous lapse $N_0(t)$, and by the scale factor of the spatial metric, its square multiplying a static auxiliary three-metric ${}^0h_{ij}$. Instead of using the (positive) scale factor we will consider its (real) logarithm, denoted by $\alpha(t)$. On the other hand, we take ${}^0h_{ij}$ as the standard Euclidean metric on the three-torus T^3 , appropriate for the considered case of a compact flat universe. We denote the corresponding angular coordinates by θ_i , such that $2\pi\theta_i/l_0 \in S^1$, so that the period is l_0 in each of the orthonormal directions.

Any function defined on the spatial sections, such as those describing the inhomogeneous perturbations, can be expanded in the basis formed by the eigenmodes of the Laplace-Beltrami operator compatible with the metric ${}^0h_{ij}$. In this way, we transform the study of the spatial dependence into a spectral analysis in terms of those modes, whose dynamics decouple (in fixed backgrounds) at leading perturbative order. Then, as in Ref. [25], we adopt a basis of

real Fourier modes, formed by the sine and cosine functions

$$\tilde{Q}_{\vec{n},+}(\vec{\theta}) = \sqrt{2} \cos\left(\frac{2\pi}{l_0} \vec{n} \cdot \vec{\theta}\right), \quad (2.1)$$

$$\tilde{Q}_{\vec{n},-}(\vec{\theta}) = \sqrt{2} \sin\left(\frac{2\pi}{l_0} \vec{n} \cdot \vec{\theta}\right). \quad (2.2)$$

Here, $\vec{n} \cdot \vec{\theta} = \sum_i n_i \theta_i$, and $\vec{n} = (n_1, n_2, n_3) \in \mathbb{Z}^3$ is any tuple of integers such that its first non-vanishing component is strictly positive (this restriction is introduced to avoid repetition of modes). These scalar modes have a norm equal to the square root of the auxiliary volume l_0^3 of the three-torus, and a Laplace-Beltrami eigenvalue equal to $-\omega_n^2 = -4\pi^2 \vec{n} \cdot \vec{n} / l_0^2$.

In the expansion of the inhomogeneities, the vanishing tuple \vec{n} is not included. The zero-modes account for the background homogeneous geometry, parameterized by $N_0(t)$ and $\alpha(t)$, and for the homogeneous part of the matter field Φ , that we denote by $\varphi(t)$. These degrees of freedom are treated exactly at the perturbative order of our truncations in the description of the system.² Using this Fourier expansion, the spacetime metric and the scalar field can be expressed as

$$\begin{aligned} h_{ij}(t, \vec{\theta}) = & \sigma^2 e^{2\alpha(t)} {}^0 h_{ij}(\vec{\theta}) \left[1 + 2 \sum_{\vec{n}, \epsilon} a_{\vec{n}, \epsilon}(t) \tilde{Q}_{\vec{n}, \epsilon}(\vec{\theta}) \right] \\ & + 6\sigma^2 e^{2\alpha(t)} \sum_{\vec{n}, \epsilon} b_{\vec{n}, \epsilon}(t) \left[\frac{1}{\omega_n^2} (\tilde{Q}_{\vec{n}, \epsilon})_{|ij}(\vec{\theta}) + \frac{1}{3} {}^0 h_{ij}(\vec{\theta}) \tilde{Q}_{\vec{n}, \epsilon}(\vec{\theta}) \right], \end{aligned} \quad (2.3a)$$

$$N(t, \vec{\theta}) = \sigma \left[N_0(t) + e^{3\alpha(t)} \sum_{\vec{n}, \epsilon} g_{\vec{n}, \epsilon}(t) \tilde{Q}_{\vec{n}, \epsilon}(\vec{\theta}) \right], \quad (2.3b)$$

$$N_i(t, \vec{\theta}) = \sigma^2 e^{2\alpha(t)} \sum_{\vec{n}, \epsilon} \frac{1}{\omega_n^2} k_{\vec{n}, \epsilon}(t) (\tilde{Q}_{\vec{n}, \epsilon})_{|i}(\vec{\theta}), \quad (2.3c)$$

$$\Phi(t, \vec{\theta}) = \frac{1}{\sigma \sqrt{l_0^3}} \left[\varphi(t) + \sum_{\vec{n}, \epsilon} f_{\vec{n}, \epsilon}(t) \tilde{Q}_{\vec{n}, \epsilon}(\vec{\theta}) \right]. \quad (2.3d)$$

In these equations, we have defined the constant $\sigma^2 = 4\pi G / (3l_0^3)$, G is the Newton constant, the vertical bar stands for the covariant derivative with respect to the auxiliary metric ${}^0 h_{ij}$, and $\epsilon = +, -$ for cosine and sine modes, respectively. Besides, we have scaled for convenience the shift vector and the inhomogeneous part of the lapse function by a power of the scale factor³ (and of the mode frequency in the case of the shift). The time-dependence of the geometry and matter scalar perturbations is parameterized by the functions

$$\{a_{\vec{n}, \epsilon}(t), b_{\vec{n}, \epsilon}(t), g_{\vec{n}, \epsilon}(t), k_{\vec{n}, \epsilon}(t), f_{\vec{n}, \epsilon}(t)\}. \quad (2.4)$$

In what follows we will omit the time-dependence in these quantities to simplify the notation.

Following an approach parallel to that in Ref. [15], we can now substitute these expressions in the Hamiltonian form of the gravitational action with a minimally coupled scalar field

²See e.g. Refs. [15, 16, 18, 31], and the discussion on this issue in Refs. [25, 29], as well as the full treatment beyond perturbation theory adopted in the inhomogeneous Gowdy cosmologies [32], that confirms that when the inhomogeneities are regarded as perturbations, this procedure to deal with the zero-modes is correct.

³We note that $k_{\vec{n}, \epsilon}(t)$ and $g_{\vec{n}, \epsilon}(t)$ are not exactly those of Ref. [25], owing to the commented scaling.

and truncate the result at quadratic order in the inhomogeneous perturbations. In this way, one obtains the total Hamiltonian H of the perturbed system at this order of approximation. As expected, this Hamiltonian is given by a linear combination of constraints, reflecting the freedom inherited from general relativity to perform spatial diffeomorphisms and time reparameterizations. Specifically, we get [15, 24]

$$H = N_0 \left[H_{|0} + \sum_{\vec{n}, \epsilon} H_{|2}^{\vec{n}, \epsilon} \right] + \sum_{\vec{n}, \epsilon} g_{\vec{n}, \epsilon} \tilde{H}_{|1}^{\vec{n}, \epsilon} + \sum_{\vec{n}, \epsilon} k_{\vec{n}, \epsilon} \tilde{H}_{-1}^{\vec{n}, \epsilon}. \quad (2.5)$$

Here, $H_{|0}$ denotes the scalar or Hamiltonian constraint of the unperturbed flat FRW model, that generates homogeneous time reparameterizations in that system:

$$H_{|0} = \frac{e^{-3\alpha}}{2} \left(-\pi_\alpha^2 + \pi_\varphi^2 + e^{6\alpha} \tilde{m}^2 \varphi^2 \right). \quad (2.6)$$

The constant \tilde{m} is the mass m of the scalar field conveniently re-expressed as $\tilde{m} = m\sigma$. Besides, we employ the notation π_q to denote the momentum conjugate to any variable q . Notice that, in the perturbed case under study, the zero-mode of the Hamiltonian constraint (or global Hamiltonian constraint) gets contributions from the inhomogeneities. At our truncation order, these contributions are quadratic in the perturbations. We have included them in the sum of terms $H_{|2}^{\vec{n}, \epsilon}$. For each mode of the perturbations, the contribution is⁴

$$\begin{aligned} H_{|2}^{\vec{n}, \epsilon} = \frac{e^{-3\alpha}}{2} \bigg\{ & -\pi_{a_{\vec{n}, \epsilon}}^2 + \pi_{b_{\vec{n}, \epsilon}}^2 + \pi_{f_{\vec{n}, \epsilon}}^2 + 2\pi_\alpha (a_{\vec{n}, \epsilon} \pi_{a_{\vec{n}, \epsilon}} + 4b_{\vec{n}, \epsilon} \pi_{b_{\vec{n}, \epsilon}}) - 6\pi_\varphi a_{\vec{n}, \epsilon} \pi_{f_{\vec{n}, \epsilon}} \\ & + \pi_\alpha^2 \left(\frac{1}{2} a_{\vec{n}, \epsilon}^2 + 10b_{\vec{n}, \epsilon}^2 \right) + \pi_\varphi^2 \left(\frac{15}{2} a_{\vec{n}, \epsilon}^2 + 6b_{\vec{n}, \epsilon}^2 \right) \\ & - e^{4\alpha} \left(\frac{1}{3} \omega_n^2 a_{\vec{n}, \epsilon}^2 + \frac{1}{3} \omega_n^2 b_{\vec{n}, \epsilon}^2 + \frac{2}{3} \omega_n^2 a_{\vec{n}, \epsilon} b_{\vec{n}, \epsilon} - \omega_n^2 f_{\vec{n}, \epsilon}^2 \right) \\ & \left. + e^{6\alpha} \tilde{m}^2 \left[3\varphi^2 \left(\frac{1}{2} a_{\vec{n}, \epsilon}^2 - 2b_{\vec{n}, \epsilon}^2 \right) + 6\varphi a_{\vec{n}, \epsilon} f_{\vec{n}, \epsilon} + f_{\vec{n}, \epsilon}^2 \right] \right\}. \quad (2.7) \end{aligned}$$

On the other hand, $\tilde{H}_{|1}^{\vec{n}, \epsilon}$ and $\tilde{H}_{-1}^{\vec{n}, \epsilon}$ are linear in the inhomogeneous perturbations. $\tilde{H}_{|1}^{\vec{n}, \epsilon}$ arises from the perturbation of the Hamiltonian constraint around the FRW model in full general relativity, while $\tilde{H}_{-1}^{\vec{n}, \epsilon}$ comes from the perturbation of the momentum (or diffeomorphisms) constraint. These linear perturbative constraints are given by

$$\begin{aligned} \tilde{H}_{|1}^{\vec{n}, \epsilon} = \frac{1}{2} \bigg[& -2\pi_\alpha \pi_{a_{\vec{n}, \epsilon}} + 2\pi_\varphi \pi_{f_{\vec{n}, \epsilon}} - (\pi_\alpha^2 + 3\pi_\varphi^2) a_{\vec{n}, \epsilon} - \frac{2}{3} \omega_n^2 e^{4\alpha} (a_{\vec{n}, \epsilon} + b_{\vec{n}, \epsilon}) \\ & \left. + e^{6\alpha} \tilde{m}^2 \varphi (3\varphi a_{\vec{n}, \epsilon} + 2f_{\vec{n}, \epsilon}) \right], \quad (2.8) \end{aligned}$$

$$\tilde{H}_{-1}^{\vec{n}, \epsilon} = \frac{1}{3} \left[-\pi_{a_{\vec{n}, \epsilon}} + \pi_{b_{\vec{n}, \epsilon}} + \pi_\alpha (a_{\vec{n}, \epsilon} + 4b_{\vec{n}, \epsilon}) + 3\pi_\varphi f_{\vec{n}, \epsilon} \right]. \quad (2.9)$$

We note that $g_{\vec{n}, \epsilon}$ and $k_{\vec{n}, \epsilon}$ do not parameterize physical degrees of freedom, but they are instead the Lagrange multipliers associated with these linear perturbative constraints. Finally, it is worth remarking that, at the order of truncation adopted in the action, the perturbed system is symplectic, with canonical variables given by the zero-modes α and φ , the Fourier coefficients $\{X_{q_l}^{\vec{n}, \epsilon}\} \equiv \{a_{\vec{n}, \epsilon}, b_{\vec{n}, \epsilon}, f_{\vec{n}, \epsilon}\}$ (with $l = 1, 2, 3$), and their corresponding momenta [15, 24].

⁴This formula corrects a misprint in Ref. [24], that was not relevant for the results discussed in that work.

3 Perturbations in terms of gauge-invariant variables

As explained in the introduction, our first goal is to describe our perturbative system in terms of gauge-invariant variables without fixing any gauge freedom, preserving covariance at the level of the perturbative description. With this aim, we first focus our attention on the inhomogeneous sector of the phase space which contains the degrees of freedom of the perturbations. In the previous section, these degrees of freedom were parameterized by the variables $\{X_{q_l}^{\vec{n},\epsilon}\}$, together with their canonically conjugate momenta $\{X_{p_l}^{\vec{n},\epsilon}\} \equiv \{\pi_{a_{\vec{n},\epsilon}}, \pi_{b_{\vec{n},\epsilon}}, \pi_{f_{\vec{n},\epsilon}}\}$, with l running from 1 to 3. We will introduce a canonical transformation to describe these perturbations in terms of the MS variables, two suitable combinations of the linear perturbative constraints, and appropriate conjugate pairs of all of them. Since the intention is to respect the canonical structure of the set of elementary variables used for the perturbations, it is clear that we need to find, in particular, combinations of the linear perturbative constraints which commute, therefore abelianizing the perturbative constraint algebra. Once this abelianization is introduced, the fact that the MS variables are gauge invariant, and hence commute with the perturbative constraints, makes straightforward to find a complete set of compatible elementary variables for the inhomogeneous sector. The wanted canonical transformation is then attained with a convenient choice of conjugate variables. Later on, in Sec. 4, we will complete this transformation into a canonical one in our entire system, that is, considering not only the inhomogeneities but including also the homogeneous sector of the phase space, parameterized by the canonical variables for the zero-modes $\{\alpha, \pi_\alpha, \varphi, \pi_\varphi\}$. In total, the resulting Hamiltonian will be a linear combination of constraints, that are not only first-class (as usual), but furthermore that form an abelian algebra.

So, since in this section we are only interested in the canonical transformations of the inhomogeneous sector of our system and in the symplectic structure induced on it, we will consider momentarily that the homogeneous sector is a fixed background. With this purpose, and only in that sense, we can ignore for the moment the Poisson brackets of the zero-mode variables both between themselves and with the perturbations. Hence, in this section we will operate with a Poisson bracket in the inhomogeneous sector defined accordingly as

$$\{F, G\}_{(P)} \equiv \sum_{l, \vec{n}, \epsilon} \left(\frac{\partial F}{\partial X_{q_l}^{\vec{n},\epsilon}} \frac{\partial G}{\partial X_{p_l}^{\vec{n},\epsilon}} - \frac{\partial F}{\partial X_{p_l}^{\vec{n},\epsilon}} \frac{\partial G}{\partial X_{q_l}^{\vec{n},\epsilon}} \right), \quad (3.1)$$

where F and G are functions of the perturbations, and might also depend on the background variables.

3.1 Canonical transformation for the perturbations

Let us start by introducing the MS variables. In terms of the configuration modes $X_{q_l}^{\vec{n},\epsilon}$, the modes of the MS field are given by [13, 24, 25]

$$v_{\vec{n},\epsilon} = e^\alpha \left[f_{\vec{n},\epsilon} + \frac{\pi_\varphi}{\pi_\alpha} (a_{\vec{n},\epsilon} + b_{\vec{n},\epsilon}) \right]. \quad (3.2)$$

It is straightforward to check that these modes indeed are gauge invariant, as they Poisson commute with the linear perturbative constraints:

$$\{v_{\vec{n},\epsilon}, \tilde{H}_{+1}^{\vec{n}',\epsilon'}\}_{(P)} = 0 = \{v_{\vec{n},\epsilon}, \tilde{H}_{-1}^{\vec{n}',\epsilon'}\}_{(P)}. \quad (3.3)$$

We would like to complete these modes to a set of compatible elementary variables for the description of the perturbations. Since the MS variables are gauge invariant, it is natural to try this completion by considering combinations of the linear perturbative constraints, with which they commute. Besides, in this way, the information about the perturbative constraints will be straightforwardly incorporated in our system in terms of elementary variables. In particular, imposing them quantum mechanically will be a direct task. The fact that we want to extract two compatible variables from the perturbative constraints means, as we have already pointed out, that we have to abelianize the algebra of those constraints (and a posteriori, the entire constraint algebra of the perturbed FRW system). With this aim in mind, we notice that the only non-vanishing Poisson brackets between them is $\{\tilde{H}_{|1}^{\vec{n},\epsilon}, \check{H}_{|1}^{\vec{n},\epsilon}\}_{(P)} = e^{3\alpha} H_{|0}$. It is then easy to abelianize their algebra: We only need to replace $\check{H}_{|1}^{\vec{n},\epsilon}$ with the combination

$$\begin{aligned} \check{H}_{|1}^{\vec{n},\epsilon} &= \tilde{H}_{|1}^{\vec{n},\epsilon} - 3e^{3\alpha} H_{|0} a_{\vec{n},\epsilon} \\ &= -\pi_\alpha \pi_{a_{\vec{n},\epsilon}} + \pi_\varphi \pi_{f_{\vec{n},\epsilon}} + \left(\pi_\alpha^2 - 3\pi_\varphi^2 - \frac{1}{3}\omega_n^2 e^{4\alpha}\right) a_{\vec{n},\epsilon} - \frac{1}{3}\omega_n^2 e^{4\alpha} b_{\vec{n},\epsilon} + e^{6\alpha} \tilde{m}^2 \varphi f_{\vec{n},\epsilon}. \end{aligned} \quad (3.4)$$

Actually, this change amounts to a redefinition of the zero-mode of the lapse function. Indeed, in the action of the system and up to the quadratic order in perturbations that we are working with, we can rewrite the Hamiltonian (2.5) as

$$H = \check{N}_0 \left[H_{|0} + \sum_{\vec{n},\epsilon} H_{|2}^{\vec{n},\epsilon} \right] + \sum_{\vec{n},\epsilon} g_{\vec{n},\epsilon} \check{H}_{|1}^{\vec{n},\epsilon} + \sum_{\vec{n},\epsilon} k_{\vec{n},\epsilon} \tilde{H}_{-1}^{\vec{n},\epsilon}, \quad (3.5)$$

with the new zero-mode of the lapse function acquiring contributions (of quadratic order) from the inhomogeneities:

$$\check{N}_0 = N_0 + 3e^{3\alpha} \sum_{\vec{n},\epsilon} g_{\vec{n},\epsilon} a_{\vec{n},\epsilon}. \quad (3.6)$$

Note that the MS variables $v_{\vec{n},\epsilon}$ remain invariant with respect to the new set of constraints. In summary, the set of variables $\{v_{\vec{n},\epsilon}, \check{H}_{|1}^{\vec{n},\epsilon}, \tilde{H}_{-1}^{\vec{n},\epsilon}\}$ provides a parameterization of the inhomogeneous configuration space (inasmuch as that all the variables are compatible) in terms of constraints and gauge invariants, as we wanted.

We now complete the canonical transformation in the inhomogeneous sector by determining variables canonically conjugate to our new set. It is straightforward to check that

$$C_{-1}^{\vec{n},\epsilon} = 3b_{\vec{n},\epsilon} \quad (3.7)$$

is canonically conjugate to $\tilde{H}_{-1}^{\vec{n},\epsilon}$, namely $\{C_{-1}^{\vec{n},\epsilon}, \tilde{H}_{-1}^{\vec{n},\epsilon}\}_{(P)} = 1$, while it Poisson commutes with $v_{\vec{n}',\epsilon'}$ and $\check{H}_{|1}^{\vec{n}',\epsilon'}$. Finding the other canonical variables requires a bit more of work. We skip the details of the calculation and encourage the reader to check that one can choose

$$C_{|1}^{\vec{n},\epsilon} = -\frac{1}{\pi_\alpha} (a_{\vec{n},\epsilon} + b_{\vec{n},\epsilon}), \quad (3.8)$$

as the variable conjugate to the constraint $\check{H}_{|1}^{\vec{n},\epsilon}$, in the sense that $\{C_{|1}^{\vec{n},\epsilon}, \check{H}_{|1}^{\vec{n},\epsilon}\}_{(P)} = 1$, whereas

$$\pi_{v_{\vec{n},\epsilon}} = e^{-\alpha} \left[\pi_{f_{\vec{n},\epsilon}} + \frac{1}{\pi_\varphi} \left(e^{6\alpha} \tilde{m}^2 \varphi f_{\vec{n},\epsilon} + 3\pi_\varphi^2 b_{\vec{n},\epsilon} \right) \right] \quad (3.9)$$

is a momentum conjugate to the MS variable $v_{\vec{n},\epsilon}$, that is $\{v_{\vec{n},\epsilon}, \pi_{v_{\vec{n},\epsilon}}\}_{(P)} = 1$.

For convenience, we will assign the role of configuration variables to $C_{-1}^{\vec{n},\epsilon}$ and $C_{|1}^{\vec{n},\epsilon}$, and view the constraints $\tilde{H}_{-1}^{\vec{n},\epsilon}$ and $\check{H}_{|1}^{\vec{n},\epsilon}$ as momenta (this will simplify the discussion about the imposition of the perturbative constraints à la Dirac in the quantization of the system). Furthermore, we will use a compact notation for the new set of canonical variables for the perturbations, namely we define

$$\{V_{q_l}^{\vec{n},\epsilon}\} \equiv \{v_{\vec{n},\epsilon}, C_{|1}^{\vec{n},\epsilon}, C_{-1}^{\vec{n},\epsilon}\}, \quad (3.10)$$

$$\{V_{p_l}^{\vec{n},\epsilon}\} \equiv \{\pi_{v_{\vec{n},\epsilon}}, \check{H}_{|1}^{\vec{n},\epsilon}, \tilde{H}_{-1}^{\vec{n},\epsilon}\}. \quad (3.11)$$

In this way, our previous configuration variables $X_{q_l}^{\vec{n},\epsilon}$ and the new ones $V_{q_l}^{\vec{n},\epsilon}$ are related by means of a contact transformation.

3.2 Redefinition of the MS momenta

In Sec. 5 we will carry out a Fock quantization of the perturbations, and in particular of the MS gauge-invariant field. If one reduces the system classically so that it gets described by QFT in a curved background, the resulting MS field verifies the Klein-Gordon equation of a massive scalar field (with time-dependent mass) propagating in an ultrastatic spacetime with compact spatial topology. Remarkably, a series of studies on the uniqueness of the Fock representation for Klein-Gordon fields of this type [33–35] proves that the use of the MS field to describe the perturbations allows for a unitary implementation of the quantum dynamics of the field, and that any other parameterization for the perturbations that includes a time-dependent rescaling of the field (as the one employed in Ref. [27]), prevents the dynamics for being unitarily implementable in the quantum theory [36]. The unitary implementation of the dynamics is possible in a class of (unitarily) equivalent Fock representations with vacua that are invariant under the isometries of the three-torus. Moreover, these results require a specific choice of momentum for the MS variable among all the canonical pairs, namely the evolution can be implemented unitarily as long as the MS modes verify the equation of motion

$$\dot{v}_{\vec{n},\epsilon} = \pi_{v_{\vec{n},\epsilon}} \quad (3.12)$$

at the considered perturbative order. Here the dot denotes, as usual, derivative with respect to the time coordinate t . Since the evolution equations are generated by the Hamiltonian of the system and the MS variables commute with the perturbative constraints, it is not difficult to convince oneself that the above condition on the choice of momenta is achieved only if the zero-mode of the Hamiltonian constraint, that contains quadratic contributions of the perturbations, does not include any linear term in the momenta conjugate to the MS variables.⁵

We want to adhere to the above parameterization of the MS field, and therefore we want to choose its conjugate variable so as to eliminate all terms that are linear in the MS field momentum from the Hamiltonian constraint. With that aim, we need to modify the momentum mode $\pi_{v_{\vec{n},\epsilon}}$ by taking advantage of the freedom in adding to it a linear contribution of the MS configuration variable $v_{\vec{n},\epsilon}$. Thus, we introduce a change of the form

$$\pi_{v_{\vec{n},\epsilon}} \rightarrow \check{\pi}_{v_{\vec{n},\epsilon}} = \pi_{v_{\vec{n},\epsilon}} + F v_{\vec{n},\epsilon}, \quad (3.13)$$

⁵Actually, the condition happens to be not only necessary, but also sufficient.

where F is a function of the homogeneous variables $(\alpha, \pi_\alpha, \varphi, \pi_\varphi)$. We now have to determine this function by appealing to condition (3.12). Notice that, by construction, the new set $\{\check{V}_{p_l}^{\vec{n},\epsilon}\} \equiv \{\check{\pi}_{v_{\vec{n},\epsilon}}, \check{H}_{|1}^{\vec{n},\epsilon}, \check{H}_{-1}^{\vec{n},\epsilon}\}$ still provides a canonical set together with $\{V_{q_l}^{\vec{n},\epsilon}\}$.

In order to find the explicit expression of F we can proceed in a simple way as follows. For a moment, we go to the longitudinal gauge, which is picked out by the pair of conditions $b_{\vec{n},\epsilon} = 0$ and $\pi_{a_{\vec{n},\epsilon}} = \pi_\alpha a_{\vec{n},\epsilon} + 3\pi_\varphi f_{\vec{n},\epsilon}$. In this gauge, we match the resulting MS momentum with the result that was obtained (precisely in that gauge) in Ref. [25]. This procedure turns out to provide a unique and well determined answer, showing the consistency of our calculations. Moreover, the function F is actually mode independent (therefore our notation for it). Its form is

$$F = -e^{-2\alpha} \left(\frac{1}{\pi_\varphi} e^{6\alpha} \tilde{m}^2 \varphi + \pi_\alpha + 3 \frac{\pi_\varphi^2}{\pi_\alpha} \right). \quad (3.14)$$

Thus, the modes of the new MS momentum are

$$\begin{aligned} \check{\pi}_{v_{\vec{n},\epsilon}} &= e^{-\alpha} \left[\pi_{f_{\vec{n},\epsilon}} + \frac{1}{\pi_\varphi} \left(e^{6\alpha} \tilde{m}^2 \varphi f_{\vec{n},\epsilon} + 3\pi_\varphi^2 b_{\vec{n},\epsilon} \right) \right] \\ &- e^{-2\alpha} \left(\frac{1}{\pi_\varphi} e^{6\alpha} \tilde{m}^2 \varphi + \pi_\alpha + 3 \frac{\pi_\varphi^2}{\pi_\alpha} \right) v_{\vec{n},\epsilon}. \end{aligned} \quad (3.15)$$

In order to simplify the notation, in the following we will use again the symbol $\pi_{v_{\vec{n},\epsilon}}$ to denote these redefined momentum modes and $V_{p_l}^{\vec{n},\epsilon}$ for the corresponding set of new momenta.

3.3 Inversion of the canonical transformation

For completeness, we finish this section by giving explicitly the expression of the original perturbative variables $\{X_l^{\vec{n},\epsilon}\} \equiv \{X_{q_l}^{\vec{n},\epsilon}, X_{p_l}^{\vec{n},\epsilon}\}$ in terms of the new ones $\{V_l^{\vec{n},\epsilon}\} \equiv \{V_{q_l}^{\vec{n},\epsilon}, V_{p_l}^{\vec{n},\epsilon}\}$. The result is

$$a_{\vec{n},\epsilon} = -\pi_\alpha V_{q_2}^{\vec{n},\epsilon} - \frac{1}{3} V_{q_3}^{\vec{n},\epsilon}, \quad (3.16a)$$

$$b_{\vec{n},\epsilon} = \frac{1}{3} V_{q_3}^{\vec{n},\epsilon}, \quad (3.16b)$$

$$f_{\vec{n},\epsilon} = e^{-\alpha} V_{q_1}^{\vec{n},\epsilon} + \pi_\varphi V_{q_2}^{\vec{n},\epsilon}, \quad (3.16c)$$

$$\begin{aligned} \pi_{a_{\vec{n},\epsilon}} &= -\frac{1}{\pi_\alpha} V_{p_2}^{\vec{n},\epsilon} + \frac{\pi_\varphi}{\pi_\alpha} e^\alpha V_{p_1}^{\vec{n},\epsilon} + \frac{e^{-\alpha}}{\pi_\alpha} \left(e^{6\alpha} \tilde{m}^2 \varphi + \pi_\varphi \pi_\alpha + 3 \frac{\pi_\varphi^3}{\pi_\alpha} \right) V_{q_1}^{\vec{n},\epsilon} \\ &+ \left(3\pi_\varphi^2 + \frac{1}{3} e^{4\alpha} \omega_n^2 - \pi_\alpha^2 \right) V_{q_2}^{\vec{n},\epsilon} - \frac{1}{3} \pi_\alpha V_{q_3}^{\vec{n},\epsilon}, \end{aligned} \quad (3.16d)$$

$$\begin{aligned} \pi_{b_{\vec{n},\epsilon}} &= 3V_{p_3}^{\vec{n},\epsilon} - \frac{1}{\pi_\alpha} V_{p_2}^{\vec{n},\epsilon} + \frac{\pi_\varphi}{\pi_\alpha} e^\alpha V_{p_1}^{\vec{n},\epsilon} + \frac{e^{-\alpha}}{\pi_\alpha} \left(e^{6\alpha} \tilde{m}^2 \varphi - 2\pi_\varphi \pi_\alpha + 3 \frac{\pi_\varphi^3}{\pi_\alpha} \right) V_{q_1}^{\vec{n},\epsilon} \\ &+ \frac{1}{3} e^{4\alpha} \omega_n^2 V_{q_2}^{\vec{n},\epsilon} - \frac{4}{3} \pi_\alpha V_{q_3}^{\vec{n},\epsilon}, \end{aligned} \quad (3.16e)$$

$$\pi_{f_{\vec{n},\epsilon}} = e^\alpha V_{p_1}^{\vec{n},\epsilon} + e^{-\alpha} \left(\pi_\alpha + 3 \frac{\pi_\varphi^2}{\pi_\alpha} \right) V_{q_1}^{\vec{n},\epsilon} - e^{6\alpha} \tilde{m}^2 \varphi V_{q_2}^{\vec{n},\epsilon} - \pi_\varphi V_{q_3}^{\vec{n},\epsilon}. \quad (3.16f)$$

4 Full canonical set in terms of gauge-invariant variables

Recall that in the previous section we regarded the zero-modes, described by the variables of the unperturbed FRW model $\{w_q^a\} \equiv \{\alpha, \varphi\}$ and $\{w_p^a\} \equiv \{\pi_\alpha, \pi_\varphi\}$ ($a = 1, 2$), as corresponding to a fixed background. Now, we proceed to complete our canonical transformation in the entire system, including these zero-mode variables. In this way we will succeed in describing our perturbed system in terms of gauge invariants, without any gauge fixing, while keeping its full canonical structure.

4.1 Canonical transformation of the zero-modes

The action of our system (truncated at quadratic order in perturbations) in terms of the original parameterization of the inhomogeneous sector has the form

$$S = \int dt \left[\sum_a \dot{w}_q^a w_p^a + \sum_{l, \vec{n}, \epsilon} \dot{X}_{q_l}^{\vec{n}, \epsilon} X_{p_l}^{\vec{n}, \epsilon} + H(w^a, X_l^{\vec{n}, \epsilon}) \right], \quad (4.1)$$

possibly up to surface terms which are not relevant for our discussion. Here, $H(w^a, X_l^{\vec{n}, \epsilon})$ is the total Hamiltonian expressed in terms of the original variables employed in Sec. 2, $\{w^a\} \equiv \{w_q^a, w_p^a\}$ and $\{X_l^{\vec{n}, \epsilon}\} \equiv \{X_{q_l}^{\vec{n}, \epsilon}, X_{p_l}^{\vec{n}, \epsilon}\}$, as in Eq. (3.5). Let us focus on the Legendre term that contains the information about the symplectic structure of the full system, that we call W :

$$W = \int dt \left[\sum_a \dot{w}_q^a w_p^a + \sum_{l, \vec{n}, \epsilon} \dot{X}_{q_l}^{\vec{n}, \epsilon} X_{p_l}^{\vec{n}, \epsilon} \right]. \quad (4.2)$$

Our goal is to find a canonical transformation relating the previous variables with new zero-modes $\tilde{w}^a \equiv (\tilde{w}_q^a, \tilde{w}_p^a)$ such that W retains its canonical form when expressed in terms of the gauge invariants for the perturbations, namely

$$W = \int dt \left[\sum_a \dot{\tilde{w}}_q^a \tilde{w}_p^a + \sum_{l, \vec{n}, \epsilon} \dot{V}_{q_l}^{\vec{n}, \epsilon} V_{p_l}^{\vec{n}, \epsilon} \right]. \quad (4.3)$$

Instead of reproducing all the details of the lengthy calculation that allows one to deduce the form of the desired transformation (and which is essentially based on the consideration of our change of perturbative variables as a canonical transformation in the inhomogeneous sector that depends on a series of time-dependent *external* variables, describing the homogeneous degrees of freedom), we will only sketch the main steps in the derivation. First, we notice that the relations $X_l^{\vec{n}, \epsilon} = X_l^{\vec{n}, \epsilon}(V_m^{\vec{n}, \epsilon})$ given in Eqs. (3.16) do not mix different modes and are linear. Therefore, the original variables $X_l^{\vec{n}, \epsilon}$ can be expressed in the following way

$$X_l^{\vec{n}, \epsilon} = \sum_m \left(\frac{\partial X_l^{\vec{n}, \epsilon}}{\partial V_{q_m}^{\vec{n}, \epsilon}} V_{q_m}^{\vec{n}, \epsilon} + \frac{\partial X_l^{\vec{n}, \epsilon}}{\partial V_{p_m}^{\vec{n}, \epsilon}} V_{p_m}^{\vec{n}, \epsilon} \right), \quad (4.4)$$

where the partial derivatives are functions of the zero-modes w^a and the frequency ω_n only (and $m = 1, 2, 3$). Taking this relation into account in the derivatives with respect to time that appear in W , performing several time integrations by parts, disregarding total time derivatives which contribute at most by surface terms at initial and final times (assuming a well posed variational principle for the system), and truncating up to quadratic order

in perturbations, as well as recalling that $\{V_l^{\vec{n},\epsilon}\}$ is a canonical set as long as one ignores the homogeneous sector, one can obtain the expression of the new canonical homogeneous variables \tilde{w}^a as functions of the old ones w^a and of the new perturbative variables $V_l^{\vec{n},\epsilon}$ (or of $X_l^{\vec{n},\epsilon}$ if preferred). In this way, one gets new zero-mode variables which are given by the original ones plus some additional contributions that are quadratic in the perturbations. This result can be inverted at the considered perturbative order, allowing one to express the original zero-mode variables as functions of the new phase-space variables for the entire system (zero-modes plus inhomogeneities). With this procedure, one arrives at

$$w_q^a = \tilde{w}_q^a - \frac{1}{2} \sum_{l,\vec{n},\epsilon} \left[X_{ql}^{\vec{n},\epsilon} \frac{\partial X_{pl}^{\vec{n},\epsilon}}{\partial \tilde{w}_p^a} - \frac{\partial X_{ql}^{\vec{n},\epsilon}}{\partial \tilde{w}_p^a} X_{pl}^{\vec{n},\epsilon} \right], \quad (4.5a)$$

$$w_p^a = \tilde{w}_p^a + \frac{1}{2} \sum_{l,\vec{n},\epsilon} \left[X_{ql}^{\vec{n},\epsilon} \frac{\partial X_{pl}^{\vec{n},\epsilon}}{\partial \tilde{w}_q^a} - \frac{\partial X_{ql}^{\vec{n},\epsilon}}{\partial \tilde{w}_q^a} X_{pl}^{\vec{n},\epsilon} \right]. \quad (4.5b)$$

In these expressions we have to understand the original variables $X_l^{\vec{n},\epsilon}$ for the perturbations as functions of the new ones $V_l^{\vec{n},\epsilon}$, as given by Eqs. (3.16), with $\{w^a\} = \{\alpha, \pi_\alpha, \varphi, \pi_\varphi\}$ replaced with $\{\tilde{w}^a\} = \{\tilde{\alpha}, \tilde{\pi}_\alpha, \tilde{\varphi}, \tilde{\pi}_\varphi\}$ in those equations. Note that this replacement is consistent within our truncations, since old and new zero-modes differ in terms that are quadratic in perturbations. In this sense, let us also clarify that the partial derivatives in the above identities must be taken keeping $V_l^{\vec{n},\epsilon}$ constant. In Appendix A we give the explicit expressions of the zero-modes w^a in terms of the new phase-space variables \tilde{w}^a and $V_l^{\vec{n},\epsilon}$.

4.2 Hamiltonian in terms of gauge-invariant variables

Once we know the canonical transformation relating the original phase-space parameterization, $\{w^a, X_l^{\vec{n},\epsilon}\}$, with the new one that uses gauge invariants for the perturbations, $\{\tilde{w}^a, V_l^{\vec{n},\epsilon}\}$, we can obtain the form of the Hamiltonian in the new formulation. In particular, let us consider the zero-mode of the Hamiltonian constraint. In the original variables, this zero-mode is given in Eq. (3.5) by the sum

$$H_{|0}(w^a) + \sum_{\vec{n},\epsilon} H_{|2}^{\vec{n},\epsilon}(w^a, X_l^{\vec{n},\epsilon}), \quad (4.6)$$

which contains the contribution of the homogeneous FRW model and the correction quadratic in perturbations.

Since the difference between the original and new variables for the homogeneous sector is precisely of quadratic order in the perturbations, the substitution of the relations between the old and new variables for our system leads to the following new expression for the zero-mode of the Hamiltonian constraint:

$$H_{|0}(\tilde{w}^a) + \sum_b (w^b - \tilde{w}^b) \frac{\partial H_{|0}}{\partial \tilde{w}^b}(\tilde{w}^a) + \sum_{\vec{n},\epsilon} H_{|2}^{\vec{n},\epsilon}[\tilde{w}^a, X_l^{\vec{n},\epsilon}(\tilde{w}^a, V_l^{\vec{n},\epsilon})], \quad (4.7)$$

where the difference $w^a - \tilde{w}^a$ must be regarded as a function of the new phase-space variables, given by the last terms in the two expressions (4.5a) and (4.5b). Note that this difference is a sum of independent contributions of each of the inhomogeneous modes, so that we can write

$$w^a - \tilde{w}^a \equiv \sum_{\vec{n},\epsilon} \Delta \tilde{w}_{\vec{n},\epsilon}^a. \quad (4.8)$$

Besides, evaluation of $H_{|0}$ and $H_{|2}^{\vec{n},\epsilon}$ at the new zero-mode variables \tilde{w}^a is done in Eq. (4.7) by simply replacing the original phase-space coordinates w^a directly with these new ones in the functional dependence of the considered contributions to the constraint. To derive the above expression, it suffices to expand $H_{|0}$ in series around the new zero-mode variables and recall that we are truncating the total Hamiltonian at quadratic order in the perturbations. From these considerations, we see that the quadratic contribution of the perturbations to the zero-mode of the Hamiltonian constraint in our new variables takes the form $\sum_{\vec{n},\epsilon} \tilde{H}_{|2}^{\vec{n},\epsilon}$, with

$$\tilde{H}_{|2}^{\vec{n},\epsilon} = H_{|2}^{\vec{n},\epsilon} + \sum_a \Delta \tilde{w}_{\vec{n},\epsilon}^a \frac{\partial H_{|0}}{\partial w^a}. \quad (4.9)$$

Indeed, this is the form that one would expect if one regarded the change of variables for the perturbations as a canonical transformation that depends on a series of time-dependent *external* variables (namely, the zero-modes) with dynamics governed by the Hamiltonian $H_{|0}$ [17].

If one carries out the calculation explicitly, one gets the following quadratic contribution to the constraint of our system:

$$\tilde{H}_{|2}^{\vec{n},\epsilon} = \check{H}_{|2}^{\vec{n},\epsilon} + F_{|2}^{\vec{n},\epsilon} H_{|0} + F_{|1}^{\vec{n},\epsilon} V_{p_2}^{\vec{n},\epsilon} + \left(F_{-1}^{\vec{n},\epsilon} - 3 \frac{e^{-3\tilde{\alpha}}}{\pi_{\tilde{\alpha}}} V_{p_2}^{\vec{n},\epsilon} + \frac{9}{2} e^{-3\tilde{\alpha}} V_{p_3}^{\vec{n},\epsilon} \right) V_{p_3}^{\vec{n},\epsilon}. \quad (4.10)$$

In this expression we have defined

$$\begin{aligned} \check{H}_{|2}^{\vec{n},\epsilon} &= \frac{e^{-\tilde{\alpha}}}{2} \left[\omega_n^2 + e^{-4\tilde{\alpha}} \pi_{\tilde{\alpha}}^2 + \tilde{m}^2 e^{2\tilde{\alpha}} \left(1 + 15\tilde{\varphi}^2 - 12\tilde{\varphi} \frac{\pi_{\tilde{\varphi}}}{\pi_{\tilde{\alpha}}} - 18e^{6\tilde{\alpha}} \tilde{m}^2 \frac{\tilde{\varphi}^4}{\pi_{\tilde{\alpha}}^2} \right) \right] (V_{q_1}^{\vec{n},\epsilon})^2 \\ &\quad + \frac{e^{-\tilde{\alpha}}}{2} (V_{p_1}^{\vec{n},\epsilon})^2, \end{aligned} \quad (4.11a)$$

$$\begin{aligned} F_{|2}^{\vec{n},\epsilon} &= \frac{3}{2} e^{-2\tilde{\alpha}} \left(1 - 9 \frac{\pi_{\tilde{\varphi}}^2}{\pi_{\tilde{\alpha}}^2} + 12e^{6\tilde{\alpha}} \tilde{m}^2 \frac{\tilde{\varphi}^2}{\pi_{\tilde{\alpha}}^2} \right) (V_{q_1}^{\vec{n},\epsilon})^2 + \frac{1}{2} e^{-2\tilde{\alpha}} (e^{4\tilde{\alpha}} \omega_n^2 - 9\pi_{\tilde{\varphi}}^2) (V_{q_2}^{\vec{n},\epsilon})^2 \\ &\quad - 3 \frac{e^{-2\tilde{\alpha}}}{\pi_{\tilde{\alpha}}^2} \left[(\pi_{\tilde{\alpha}}^2 + 3\pi_{\tilde{\varphi}}^2) \pi_{\tilde{\varphi}} + e^{6\tilde{\alpha}} \tilde{m}^2 \tilde{\varphi} \pi_{\tilde{\alpha}} \right] V_{q_1}^{\vec{n},\epsilon} V_{q_2}^{\vec{n},\epsilon} - 3 \frac{\pi_{\tilde{\varphi}}}{\pi_{\tilde{\alpha}}} V_{p_1}^{\vec{n},\epsilon} V_{q_2}^{\vec{n},\epsilon}, \end{aligned} \quad (4.11b)$$

$$F_{|1}^{\vec{n},\epsilon} = \frac{1}{2} \frac{e^{-4\tilde{\alpha}}}{\pi_{\tilde{\alpha}}} \left[6\pi_{\tilde{\varphi}} V_{q_1}^{\vec{n},\epsilon} + (6e^{3\tilde{\alpha}} H_{|0} + 5\pi_{\tilde{\alpha}}^2) V_{q_2}^{\vec{n},\epsilon} \right], \quad (4.11c)$$

$$\begin{aligned} F_{-1}^{\vec{n},\epsilon} &= 3 \frac{e^{-4\tilde{\alpha}}}{\pi_{\tilde{\alpha}}^2} (e^{6\tilde{\alpha}} \tilde{m}^2 \tilde{\varphi} \pi_{\tilde{\alpha}} + 3\pi_{\tilde{\varphi}}^3 - 2\pi_{\tilde{\alpha}}^2 \pi_{\tilde{\varphi}}) V_{q_1}^{\vec{n},\epsilon} + \omega_n^2 V_{q_2}^{\vec{n},\epsilon} - \frac{3}{2} e^{-6\tilde{\alpha}} \pi_{\tilde{\alpha}} V_{q_3}^{\vec{n},\epsilon} \\ &\quad - 3e^{-2\tilde{\alpha}} \frac{\pi_{\tilde{\varphi}}}{\pi_{\tilde{\alpha}}} V_{p_1}^{\vec{n},\epsilon}. \end{aligned} \quad (4.11d)$$

In the total Hamiltonian, we can account for the term $\sum_{\vec{n},\epsilon} F_{|2}^{\vec{n},\epsilon} H_{|0}$ at the considered perturbative order by means of a new redefinition of the zero-mode of the lapse function, with the new one given by $\bar{N}_0 = \check{N}_0 (1 + \sum_{\vec{n},\epsilon} F_{|2}^{\vec{n},\epsilon})$ [see Eq. (3.6)]. The other terms in Eq. (4.10) different from $\check{H}_{|2}^{\vec{n},\epsilon}$ contribute to the perturbative linear constraints and their presence amounts to a redefinition of the corresponding Lagrange multipliers, that we now denote as $G_{\vec{n},\epsilon}$ and $K_{\vec{n},\epsilon}$. In summary, we obtain at our quadratic order a total Hamiltonian constraint of the form

$$H = \bar{N}_0 \left[H_{|0} + \sum_{\vec{n},\epsilon} \check{H}_{|2}^{\vec{n},\epsilon} \right] + \sum_{\vec{n},\epsilon} G_{\vec{n},\epsilon} V_{p_2}^{\vec{n},\epsilon} + \sum_{\vec{n},\epsilon} K_{\vec{n},\epsilon} V_{p_3}^{\vec{n},\epsilon}. \quad (4.12)$$

In Appendix A we give the explicit expressions of the original lapse function N and shift co-vector N_i in terms of the new phase-space variables $\{\tilde{w}^a, V_l^{\vec{n},\epsilon}\}$ and Lagrange multipliers $\{\bar{N}_0, G_{\vec{n},\epsilon}, K_{\vec{n},\epsilon}\}$.

As we see, the terms $\check{H}_{|2}^{\vec{n},\epsilon}$ provide finally the contributions quadratic in perturbations to the zero-mode of the Hamiltonian constraint in the constructed gauge-invariant formulation. As expected, these terms are precisely those that depend exclusively on the MS variables and their momenta, and hence they are gauge-invariant quantities. For obvious reasons, we call their sum the MS Hamiltonian. Moreover, as we had anticipated, these terms contain no linear contribution of the momenta of the MS variables. In fact, the only contribution from these momenta is quadratic, and its mode-independent coefficient is constant up to a power of the scale factor. We notice as well that the expression given for $\check{H}_{|2}^{\vec{n},\epsilon}$ is linear in the momentum of the zero-mode of the scalar field, $\pi_{\tilde{\varphi}}$. This linear expression has been obtained by taking advantage of the identity $\pi_{\tilde{\varphi}}^2 = 2H_{|0}e^{3\alpha} + \pi_{\tilde{\alpha}}^2 - e^{6\alpha}\tilde{m}^2\varphi^2$, that can be used in Eq. (4.10) at the cost of redefining the zero-mode of the lapse function at the considered order in perturbation theory. Thanks to this, in Sec. 5 we will be able to interpret the zero-mode of the Hamiltonian constraint in a certain approximation as a Schrödinger-like equation that dictates the quantum evolution of the inhomogeneities in some family of states.

5 Hybrid quantization and Born-Oppenheimer ansatz

We can now proceed to quantize the symplectic manifold which describes our cosmological system, and to impose the classical constraints à la Dirac, i.e., as operators that annihilate physical states in the quantum theory. We recall that, in our classical formulation, we have split the phase space in homogeneous and inhomogeneous sectors, starting with the expansion in modes given by the eigenfunctions of the Laplace-Beltrami operator on the spatial sections. The homogeneous sector describes the zero-modes, and its degrees of freedom can be parameterized by the canonical variables $\{\tilde{w}^a\} = \{\tilde{\alpha}, \pi_{\tilde{\alpha}}, \tilde{\varphi}, \pi_{\tilde{\varphi}}\}$. On the other hand, the inhomogeneous sector describes all the non-zero modes of the perturbations, and can be parameterized by $\{V_l^{\vec{n},\epsilon}\} = \{V_{q_i}^{\vec{n},\epsilon}, V_{p_i}^{\vec{n},\epsilon}\}$, that is a set specially selected to make manifest the covariance of the formulation at the considered perturbative level. In this way we have prepared the classical system to now carry out a (generalized) hybrid quantization, in which one adopts specific quantizations for both the homogeneous and inhomogeneous parts. On the tensor product of the corresponding representation spaces, one imposes the constraints quantum mechanically, so that the total system is not at all trivial.

In fact, it seems natural to assume that there exists some regime of the quantum dynamics, in between a fully quantum gravity regime and the regime where QFT in fixed curved backgrounds makes sense, in which the most relevant quantum effects of the geometry are those affecting the zero-modes, and that perturbations, while still having a quantum description, can be represented in more standard ways, like e.g. as in QFT in curved spacetimes. Hence, we can adhere to a hybrid quantization based on this hypothesis, adopting a Fock quantization for the MS gauge invariants (and possibly for the rest of variables for the perturbations) and, in principle, a different type of quantization, of quantum gravity nature, for the homogeneous sector. Note that this hybrid approach is different from treating the perturbations as a test field propagating in a quantum background, which is the viewpoint adopted in the *dressed metric* proposal in Ref. [27]. Rather, the inhomogeneities and the homogeneous sector, even being represented with different quantization methods, form a symplectic manifold that is quantized as a whole. The Hamiltonian constraint of the system affects both

sectors, encoding the backreaction of the perturbations on the homogeneous background inasmuch as this is kept in the quadratic truncation of the action.

In this section we will maintain a general discussion without specifying the concrete quantization of the homogeneous sector. Moreover, most of the details are easily extensible to other quantizations of our MS variables different from the Fock (or QFT-like) quantization that we adopt here. We will simply assume that the quantization of the zero-mode variables provides a representation for the canonical commutation relations such that the operators for the homogeneous FRW geometry commute with those representing the homogeneous sector of the matter field, and in turn all of them commute with the elementary operators representing the variables for the inhomogeneities. We will denote $\mathcal{H}_{\text{kin}}^{\text{grav}} \otimes \mathcal{H}_{\text{kin}}^{\text{matt}}$ the kinematical Hilbert space for the homogeneous sector, such that the operators representing the homogeneous FRW geometry are defined on $\mathcal{H}_{\text{kin}}^{\text{grav}}$, while the operators for the homogeneous matter sector are defined on $\mathcal{H}_{\text{kin}}^{\text{matt}}$. Finally, concerning the quantization of the homogeneous matter sector, we will also assume for simplicity that the operators that represent functions of $\tilde{\varphi}$ act by mere multiplication, while the operator representing $\pi_{\tilde{\varphi}}$ acts as a generalized derivative.

In previous applications of this strategy of a hybrid quantization [23–26], the polymeric representation of LQC was adopted for the background geometry, together with a standard Schrödinger representation for the zero-mode $\tilde{\varphi}$ of the matter field. Although we will briefly explain the particularization of our study to such a representation in the next section, we intend to make here a much more general discussion, by no means restricted to the framework of LQC. Therefore, our analysis generalizes the quantum treatment presented in Ref. [25] in two important ways: a) instead of reducing the system classically, we carry the linear perturbative constraints to the quantum theory, with no gauge fixing; and b) we do not adopt here any specific representation for the homogeneous sector, and in particular for the FRW geometry. In this sense, the formulation can be adapted to any quantization proposal in FRW cosmology. The physics resulting from each specific proposal might be discussed and compared afterwards.

5.1 Fock representation for the perturbations

For the inhomogeneous modes describing the MS gauge invariants, we adopt here a Fock representation similar to that of Refs. [24, 25]. Remarkably, this quantization is selected by the criteria of: a) invariance of the vacuum under the spatial isometries, and b) unitary implementability of the dynamics in the regime in which one recovers a QFT in a curved background for the MS field (in any finite time interval) [33, 34]. As mentioned in Subsec. 3.2, these criteria remove the ambiguity in a possible rescaling of the MS modes by functions of the homogeneous variables, and they assure a unitary implementability of the evolution and a standard quantum mechanical interpretation in the regions where a QFT in a (generally effective) background is recovered. Actually, these results require the spatial sections to be compact; that is why we choose our flat model to have three-torus topology. More concretely, the above criteria pick out not just a Fock representation, but a family of them such that they are all unitarily equivalent [34]. In particular, this family contains the representation in which the annihilation-like variables, $a_{\vec{n},\epsilon}$, and the creation-like variables, $a_{\vec{n},\epsilon}^\dagger$, are those naturally associated with harmonic oscillators of frequency ω_n . A more specific choice of Fock representation among the privileged family selected by those criteria can be done on the basis of further requirements not just on operators that are constructed from exponentials of the MS variables and their momenta (namely, the Weyl algebra), including linear operators, but on other physically relevant operators like e.g. the field Hamiltonian. It seems natural to demand

that the MS Hamiltonian be well defined quantum mechanically (and be essentially self-adjoint). Additional conditions related with regularization may be important, although one would expect that regularization schemes should be a direct consequence of the quantization of the system, since this includes now the background, rather than having to import them from QFT, where one usually appeals to them claiming the absence of a quantization of the curved spacetime.

Assume then that we take a particular Fock representation for the MS modes in the family picked out by the commented criteria of vacuum invariance and unitary dynamics in the standard QFT regime. Let us call $\hat{a}_{\vec{n},\epsilon}$ and $\hat{a}_{\vec{n},\epsilon}^\dagger$ the corresponding annihilation and creation operators, acting on the Fock space \mathcal{F} , such that $[\hat{a}_{\vec{n},\epsilon}, \hat{a}_{\vec{n}',\epsilon'}^\dagger] = \delta_{\vec{n},\epsilon}^{\vec{n}',\epsilon'}$. A basis of states is provided by the occupancy-number states $|\mathcal{N}\rangle$, where \mathcal{N} denotes an array of (positive integer) occupancy-numbers, one for each mode. The creation operator acting on these states excites the mode labeled by \vec{n} and ϵ , increasing the corresponding occupancy-number in one unit, as usual.

5.2 Quantum representation of the constraints

In our gauge-invariant formulation the classical constraints Poisson commute, and therefore they can be imposed quantum mechanically without troubles as far as their quantum counterparts commute as well. We assume that our quantization satisfies this property and impose them independently.

Let us start with the linear perturbative constraints, which classically are $V_{p_i}^{\vec{n},\epsilon} = 0$ for $\tilde{l} = 2, 3$. These constraints are straightforward to impose adopting a quantization, for the part of the inhomogeneous sector parameterized by $\{V_{\tilde{l}}^{\vec{n},\epsilon}\}$ (again with $\tilde{l} = 2, 3$), such that the operators representing the momenta act as derivatives (or as translations, in integrated or discrete versions) with respect to the configuration variables $V_{q_i}^{\vec{n},\epsilon}$. Then, the constraints amount to require that physical states are independent of those variables. Hence, after imposing these constraints, we can just restrict the discussion to a representation space of the form $\mathcal{H} = \mathcal{H}_{\text{kin}}^{\text{grav}} \otimes \mathcal{H}_{\text{kin}}^{\text{matt}} \otimes \mathcal{F}$ to study all physically relevant quantum states. Let us emphasize that the restriction to these states is quantum mechanical, and not a classical reduction. Besides, let us note that these states are not yet fully physical, since we must still impose the zero-mode of the Hamiltonian constraint.

More challenging is the imposition of this global Hamiltonian constraint, for which we will only be able to provide approximate solutions. Let us first focus on its quantum representation. Classically this constraint is given by

$$H_{|0} + \check{H}_{|2} \equiv e^{-3\tilde{\alpha}} \check{H} = 0, \quad (5.1)$$

with $H_{|0}$ defined as in Eq. (2.6) but in terms of the background variables $(\tilde{\alpha}, \pi_{\tilde{\alpha}}, \tilde{\varphi}, \pi_{\tilde{\varphi}})$, and $\check{H}_{|2} = \sum_{\vec{n},\epsilon} \check{H}_{|2}^{\vec{n},\epsilon}$ defined in Eq. (4.11a). In what follows we will focus on the densitized Hamiltonian constraint⁶ $\tilde{H} = 0$.

Before continuing, let us introduce some convenient notation. We first define

$$\mathcal{H}_0^{(2)} \equiv \pi_{\tilde{\alpha}}^2 - e^{6\tilde{\alpha}} \tilde{m}^2 \tilde{\varphi}^2, \quad (5.2)$$

⁶We could also choose to work with the non-densitized constraint and carry out the densitization at the quantum level (see e.g. Ref. [25]).

so that the contribution of the homogeneous sector to \tilde{H} is $e^{3\tilde{\alpha}}H_{|0} = (\pi_{\tilde{\varphi}}^2 - \mathcal{H}_0^{(2)})/2$. In addition, we introduce the following functions of the zero-modes:

$$\vartheta_o \equiv -12e^{4\tilde{\alpha}}\tilde{m}^2\frac{\tilde{\varphi}}{\pi_{\tilde{\alpha}}}, \quad (5.3a)$$

$$\vartheta_e \equiv e^{2\tilde{\alpha}}, \quad (5.3b)$$

$$\begin{aligned} \vartheta_e^q &\equiv e^{-2\tilde{\alpha}}\pi_{\tilde{\alpha}}^2 + \tilde{m}^2e^{4\tilde{\alpha}}\left(1 + 15\tilde{\varphi}^2 - 18e^{6\tilde{\alpha}}\tilde{m}^2\frac{\tilde{\varphi}^4}{\pi_{\tilde{\alpha}}^2}\right) \\ &= e^{-2\tilde{\alpha}}\mathcal{H}_0^{(2)}\left(19 - 18\frac{\mathcal{H}_0^{(2)}}{\pi_{\tilde{\alpha}}^2}\right) + \tilde{m}^2e^{4\tilde{\alpha}}(1 - 2\tilde{\varphi}^2), \end{aligned} \quad (5.3c)$$

and we call

$$\Theta_o^{\vec{n},\epsilon} \equiv -\vartheta_o(V_{q_1}^{\vec{n},\epsilon})^2 \quad (5.4a)$$

$$\Theta_e^{\vec{n},\epsilon} \equiv -\left[(\vartheta_e\omega_n^2 + \vartheta_e^q)(V_{q_1}^{\vec{n},\epsilon})^2 + \vartheta_e(V_{p_1}^{\vec{n},\epsilon})^2\right], \quad (5.4b)$$

$$\Theta_o \equiv \sum_{\vec{n},\epsilon} \Theta_o^{\vec{n},\epsilon}, \quad \Theta_e \equiv \sum_{\vec{n},\epsilon} \Theta_e^{\vec{n},\epsilon}. \quad (5.4c)$$

Then, $2e^{3\tilde{\alpha}}\tilde{H}_{|2} = -(\Theta_e + \Theta_o\pi_{\tilde{\varphi}})$. Let us assume that we can represent the quantities introduced in Eqs. (5.2) and (5.3) as densely defined operators, $\hat{\mathcal{H}}_0^{(2)}$, $\hat{\vartheta}_o$, $\hat{\vartheta}_e$, and $\hat{\vartheta}_e^q$, on $\mathcal{H}_{\text{kin}}^{\text{grav}} \otimes \mathcal{H}_{\text{kin}}^{\text{matt}}$ (while acting as the identity on \mathcal{F}), and the objects given in Eqs. (5.4) as densely defined operators, $\hat{\Theta}_o$ and $\hat{\Theta}_e$, on $\mathcal{H}_{\text{kin}}^{\text{grav}} \otimes \mathcal{H}_{\text{kin}}^{\text{matt}} \otimes \mathcal{F}$. Then, we obtain the following operator representing \tilde{H} :

$$\hat{H} = \frac{1}{2}\left[\hat{\pi}_{\tilde{\varphi}}^2 - \hat{\mathcal{H}}_0^{(2)} - \hat{\Theta}_e - \left(\hat{\Theta}_o\hat{\pi}_{\tilde{\varphi}}\right)_S\right], \quad (5.5)$$

where we have adopted the symmetrization

$$\left(\hat{\Theta}_o\hat{\pi}_{\tilde{\varphi}}\right)_S = \frac{1}{2}\left(\hat{\Theta}_o\hat{\pi}_{\tilde{\varphi}} + \hat{\pi}_{\tilde{\varphi}}\hat{\Theta}_o\right) = \frac{1}{2}[\hat{\pi}_{\tilde{\varphi}}, \hat{\Theta}_o] + \hat{\Theta}_o\hat{\pi}_{\tilde{\varphi}}. \quad (5.6)$$

5.3 Born-Oppenheimer ansatz

We will now analyze solutions to the (zero-mode of the) Hamiltonian constraint of the system by adopting the ansatz

$$\Psi = \Gamma(\tilde{\alpha}, \tilde{\varphi})\psi(\mathcal{N}, \tilde{\varphi}), \quad (5.7)$$

where the dependence on the MS variables is denoted by the label \mathcal{N} of the occupancy-number states for the inhomogeneous subsector \mathcal{F} , and $\tilde{\alpha}$ denotes the dependence on the geometry of the homogeneous FRW sector.⁷ Thus, the ansatz for the quantum states is that their wave functions Ψ can be separated into two factors, one which depends on the homogeneous degrees of freedom of the FRW geometry, and the other on the MS gauge-invariant modes. The ansatz allows for states that present different rates of variation in these two factors with respect to the zero-mode $\tilde{\varphi}$ of the matter scalar field, and when this happens the Hamiltonian constraint can be approximated and simplified. Hence the name of Born-Oppenheimer for

⁷This notation does not imply that we adopt a representation for the homogeneous geometry in which we have an operator representing $\tilde{\alpha}$ that acts by multiplication, but it is rather a symbolic way to indicate functional dependence on the homogeneous geometry sector.

the introduced ansatz. From this perspective, it is convenient to interpret $\tilde{\varphi}$ as an internal time for the system (at least in some intervals of the evolution). We recall that we are assuming a representation of the homogeneous matter sector such that the functions of $\tilde{\varphi}$ act as multiplicative operators on Ψ .

Let us for a moment consider the unperturbed FRW model. Its Hamiltonian constraint operator is proportional to $\hat{\pi}_{\tilde{\varphi}}^2 - \hat{\mathcal{H}}_0^{(2)}$. In the evolution picture that we are employing, where $\tilde{\varphi}$ plays the role of the time, positive-frequency solutions of this constraint are given by

$$\Gamma(\tilde{\alpha}, \tilde{\varphi}) = \hat{U}(\tilde{\alpha}; \tilde{\varphi})\chi(\tilde{\alpha}), \quad (5.8)$$

\hat{U} being the corresponding unitary evolution operator. If $\hat{\mathcal{H}}_0^{(2)}$ is self-adjoint, then one can project on its positive part (P.P.) and take $\sqrt{\text{P.P.}(\hat{\mathcal{H}}_0^{(2)})}$ as the $\tilde{\varphi}$ -dependent self-adjoint operator that generates the time evolution in Eq. (5.8). To maintain the discussion more general, and since the self-adjointness of $\hat{\mathcal{H}}_0^{(2)}$ might not be guaranteed, or a straightforward definition of a square root for it might not be available, we simply assume that the FRW part of the quantum states has the form (5.8), where the family of evolution operators \hat{U} is such that there exists a self-adjoint operator $\hat{\mathcal{H}}_0$ satisfying $[\hat{\pi}_{\tilde{\varphi}}, \hat{U}]\hat{U}^{-1} = \hat{\mathcal{H}}_0$. Note that the concepts of self-adjointness and unitarity that we use here are those corresponding to the FRW-geometry part of our Hilbert space,⁸ $\mathcal{H}_{\text{kin}}^{\text{grav}}$. Actually, within our perturbative framework, it is natural to consider Γ as an approximate solution to the Hamiltonian constraint of the homogeneous sector, and not necessarily an exact one. In fact, as we will see, Γ will supply an approximate solution of the FRW model inasmuch as the square of $\hat{\mathcal{H}}_0$ is *close* to $\hat{\mathcal{H}}_0^{(2)}$ in the sense that $(\hat{\mathcal{H}}_0)^2 - \hat{\mathcal{H}}_0^{(2)} + [\hat{\pi}_{\tilde{\varphi}}, \hat{\mathcal{H}}_0]$ is negligibly small on Γ .

On the other hand, in Eq. (5.8) we normalize the state $\chi(\tilde{\alpha})$ to the unit in $\mathcal{H}_{\text{kin}}^{\text{grav}}$. This part of the quantum state can be understood as the state of the FRW geometry at a fixed initial time $\tilde{\varphi}_0$ (and hence it is independent of $\tilde{\varphi}$). A natural election would be to choose a very semiclassical state χ , so that it is highly peaked on a certain homogeneous geometry (and, if this is possible, that it remains peaked under the evolution dictated by \hat{U}).

Let us plug the ansatz (5.7) in the constraint equation⁹ $\hat{H}\Psi = 0$. Taking into account that

$$\hat{\pi}_{\tilde{\varphi}}\Psi = \Gamma(\hat{\pi}_{\tilde{\varphi}}\psi) + (\hat{\mathcal{H}}_0\Gamma)\psi, \quad (5.9a)$$

$$\hat{\pi}_{\tilde{\varphi}}^2\Psi = \Gamma(\hat{\pi}_{\tilde{\varphi}}^2\psi) + 2(\hat{\mathcal{H}}_0\Gamma)(\hat{\pi}_{\tilde{\varphi}}\psi) + ([\hat{\pi}_{\tilde{\varphi}}, \hat{\mathcal{H}}_0]\Gamma)\psi + \{(\hat{\mathcal{H}}_0)^2\Gamma\}\psi, \quad (5.9b)$$

the constraint can be rewritten as

$$\begin{aligned} & \left\{ \left((\hat{\mathcal{H}}_0)^2 - \hat{\mathcal{H}}_0^{(2)} + [\hat{\pi}_{\tilde{\varphi}}, \hat{\mathcal{H}}_0] \right) \Gamma \right\} \psi + 2(\hat{\mathcal{H}}_0\Gamma)(\hat{\pi}_{\tilde{\varphi}}\psi) + \Gamma(\hat{\pi}_{\tilde{\varphi}}^2\psi) - \frac{1}{2}[\hat{\pi}_{\tilde{\varphi}} - \hat{\mathcal{H}}_0, \hat{\Theta}_o](\Gamma\psi) \\ & - \{ \hat{\Theta}_e + (\hat{\Theta}_o\hat{\mathcal{H}}_0)_S \} (\Gamma\psi) - \hat{\Theta}_o \{ \Gamma(\hat{\pi}_{\tilde{\varphi}}\psi) \} = 0. \end{aligned} \quad (5.10)$$

⁸In a conventional Schrödinger representation in which $\hat{\pi}_{\tilde{\varphi}} = -i\partial_{\tilde{\varphi}}$ we have the standard unitary evolution operator

$$\hat{U}(\tilde{\alpha}; \tilde{\varphi}) = \mathcal{P} \left[\exp \left(i \int_{\tilde{\varphi}_0}^{\tilde{\varphi}} d\tilde{\varphi} \hat{\mathcal{H}}_0(\tilde{\alpha}, \tilde{\varphi}) \right) \right].$$

The symbol \mathcal{P} denotes *time* ordering with respect to $\tilde{\varphi}$. For simplicity, we have fixed the reduced Planck constant \hbar equal to the unit.

⁹Since one expects that solutions do not belong to the kinematical Hilbert space, one should rather impose the constraint on some kind of generalized states $|\psi\rangle$ with the adjoint action, in the form $\langle\psi|\hat{H}^\dagger = 0$. We do not do so because we do not want to complicate the notation even more.

We note that, with our assumptions on the representation of the zero-mode of the scalar field, the operators $[\hat{\pi}_{\tilde{\varphi}}, \hat{\mathcal{H}}_0]$ and $[\hat{\pi}_{\tilde{\varphi}}, \hat{\Theta}_o]$ depend on $\tilde{\varphi}$, that acts by multiplication as an operator, but are independent of $\hat{\pi}_{\tilde{\varphi}}$. Besides, notice that the first term of the above equation is the correction coming from the consideration of wave functions Γ that are not exact solutions to the constraint of the homogeneous sector (to see this, it suffices to make ψ constant and the operators $\hat{\Theta}_o$ and $\hat{\Theta}_e$ of the MS modes equal to zero).

Now, let us consider the approximation that, regarding the homogeneous FRW geometry, only the terms corresponding to the expectation values on Γ are relevant. In other words, when taking the inner product of the left hand side of Eq. (5.10) with Γ with respect to the FRW geometry (namely in $\mathcal{H}_{\text{kin}}^{\text{grav}}$), we disregard possible quantum transitions from Γ to other states mediated by the action of the constraint. From Eq. (5.10) it is not difficult to see that, for this approximation to hold, we only need the following operators to have small relative dispersions on the state Γ for all values of $\tilde{\varphi}$ (or, in other words, the following operators must be highly peaked on Γ along the evolution in $\tilde{\varphi}$): i) $\hat{\mathcal{H}}_0$, ii) $\hat{\vartheta}_e$, iii) $-\frac{i}{2}d_{\tilde{\varphi}}\hat{\vartheta}_o + (\hat{\vartheta}_o\hat{\mathcal{H}}_0)_S + \hat{\vartheta}_e^g$, and iv) $-id_{\tilde{\varphi}}\hat{\mathcal{H}}_0 + (\hat{\mathcal{H}}_0)^2 - \hat{\mathcal{H}}_0^{(2)}$, where we have defined¹⁰

$$-id_{\tilde{\varphi}}\hat{O} \equiv [\hat{\pi}_{\tilde{\varphi}} - \hat{\mathcal{H}}_0, \hat{O}], \quad (5.11)$$

for any operator \hat{O} . In principle, the condition on the operator iv) may be satisfied with a suitable choice of $\hat{\mathcal{H}}_0$, in agreement with our comments above.

Assuming that the considered approximation is valid, we then get

$$\begin{aligned} & \hat{\pi}_{\tilde{\varphi}}^2\psi + \left(2\langle\hat{\mathcal{H}}_0\rangle_{\Gamma} - \langle\hat{\Theta}_o\rangle_{\Gamma}\right)\hat{\pi}_{\tilde{\varphi}}\psi \\ &= \left[\langle\hat{\Theta}_e + (\hat{\Theta}_o\hat{\mathcal{H}}_0)_S\rangle_{\Gamma} + i\langle d_{\tilde{\varphi}}\hat{\mathcal{H}}_0 - \frac{1}{2}d_{\tilde{\varphi}}\hat{\Theta}_o\rangle_{\Gamma} + \langle\hat{\mathcal{H}}_0^{(2)} - (\hat{\mathcal{H}}_0)^2\rangle_{\Gamma}\right]\psi. \end{aligned} \quad (5.12)$$

Here we have introduced the symbol $\langle\hat{O}\rangle_{\Gamma}$ to denote the expectation value of the operator \hat{O} on Γ in the Hilbert space $\mathcal{H}_{\text{kin}}^{\text{grav}}$, namely, only with respect to the inner product of the FRW geometry. Note that, in general, the result is an operator acting on $\mathcal{H}_{\text{kin}}^{\text{matt}} \otimes \mathcal{F}$.

The above equation leads to a generalized Schrödinger equation for the evolution (in $\tilde{\varphi}$) of the inhomogeneities provided that the following conditions are satisfied:

- a) $\langle\hat{\Theta}_o\rangle_{\Gamma}$ has to be negligible as compared to $\langle\hat{\mathcal{H}}_0\rangle_{\Gamma}$. This is valid as long as the quadratic contribution of the MS modes given by $\hat{\Theta}_o$ remains small when compared to the introduced generator of the $\tilde{\varphi}$ -evolution in the FRW case, something that certainly fits in the perturbative scheme that we have adopted for the treatment of the inhomogeneities.
- b) $\hat{\pi}_{\tilde{\varphi}}^2\psi$ has to be negligible as well. The self-consistency of this assumption is checked in Appendix B.
- c) $i\langle d_{\tilde{\varphi}}\hat{\mathcal{H}}_0 - \frac{1}{2}d_{\tilde{\varphi}}\hat{\Theta}_o\rangle_{\Gamma}$ has to be negligible in comparison with $\langle\hat{\Theta}_e + (\hat{\Theta}_o\hat{\mathcal{H}}_0)_S\rangle_{\Gamma}$. Otherwise, the Schrödinger equation would contain a term destroying the unitary evolution of the MS modes and which is not present in the classical field equations of these gauge invariants.

¹⁰In the conventional case $\hat{\pi}_{\tilde{\varphi}} = -i\partial_{\tilde{\varphi}}$, $d_{\tilde{\varphi}}$ is the total derivative in the Heisenberg picture corresponding to a time evolution in $\tilde{\varphi}$ generated by $\hat{\mathcal{H}}_0$.

Indeed, if these three conditions are satisfied, we get the generalized Schrödinger equation

$$\hat{\pi}_{\tilde{\varphi}}\psi = \frac{\langle \hat{\Theta}_e + (\hat{\Theta}_o \hat{\mathcal{H}}_0)_S \rangle_{\Gamma} + \langle \hat{\mathcal{H}}_0^{(2)} - (\hat{\mathcal{H}}_0)^2 \rangle_{\Gamma}}{2\langle \hat{\mathcal{H}}_0 \rangle_{\Gamma}} \psi. \quad (5.13)$$

On top of the previous conditions, let us consider as well the requirement

- d) $\langle \hat{\mathcal{H}}_0^{(2)} - (\hat{\mathcal{H}}_0)^2 \rangle_{\Gamma}$ has to be negligible in comparison with $\langle \hat{\Theta}_e + (\hat{\Theta}_o \hat{\mathcal{H}}_0)_S \rangle_{\Gamma}$. This is natural to assume (once $\langle d_{\tilde{\varphi}} \hat{\mathcal{H}}_0 \rangle_{\Gamma}$ has been ignored) if Γ is an approximate solution of the unperturbed FRW model, as argued before.

If the four conditions a)–d) hold, then the Schrödinger equation simplifies to

$$\hat{\pi}_{\tilde{\varphi}}\psi = \frac{\langle \hat{\Theta}_e + (\hat{\Theta}_o \hat{\mathcal{H}}_0)_S \rangle_{\Gamma}}{2\langle \hat{\mathcal{H}}_0 \rangle_{\Gamma}} \psi. \quad (5.14)$$

Correspondingly, the Hamiltonian generating the dynamics of the perturbations in the internal time $\tilde{\varphi}$ is given by $\langle \hat{\Theta}_e + (\hat{\Theta}_o \hat{\mathcal{H}}_0)_S \rangle_{\Gamma} / 2\langle \hat{\mathcal{H}}_0 \rangle_{\Gamma}$. This Hamiltonian is just the MS Hamiltonian, with its dependence on the variables of the homogeneous geometry evaluated at the expectation values corresponding to the quantum state Γ , and divided by the expectation value of $\hat{\mathcal{H}}_0$ on it.

We see that assuming the Born-Oppenheimer ansatz and introducing a controlled number of approximations allows us to recover standard QFT for the gauge-invariant perturbations, which propagate on an effective homogeneous geometry than can be regarded as *dressed*. The term *dressed* refers here to the fact that this geometry is not a classical one, but it retains the main quantum corrections to the geometry that the MS variables feel [27].

5.4 Effective equations for the Mukhanov-Sasaki variables

Employing the Born-Oppenheimer ansatz and the approximation that the state Γ remains highly peaked on the operators that encode the effect of the FRW geometry in the zero-mode of the Hamiltonian constraint (and hence, strictly speaking, without the need to introduce the rest of approximations that lead to a Schrödinger equation for the wave function of the MS modes), one can further derive effective classical equations for the MS variables. These effective dynamics treat the perturbations as classical, namely they replace the annihilation and creation operators of the gauge-invariant inhomogeneities by their classical counterpart. The assumption that this replacement is acceptable in order to get effective equations does not seem too stringent, because the zero-mode of the Hamiltonian constraint is quadratic in the MS gauge invariants. Since the functions of $\tilde{\varphi}$ act by multiplication, we can see the quantum dynamics of the MS modes ruled by that Hamiltonian constraint as corresponding to time-dependent harmonic oscillators, with $\tilde{\varphi}$ playing the role of internal time.

In the previous subsection we saw that the evolution of the MS variables is governed by Eq. (5.12), which can be interpreted as the result of imposing a constraint of the form:

$$\hat{\mathcal{C}}_{\text{per}} = \hat{\pi}_{\tilde{\varphi}}^2 + D_{\Gamma}(\tilde{\varphi})\hat{\pi}_{\tilde{\varphi}} + E_{\Gamma}(\tilde{\varphi}) - \left\langle \hat{\Theta}_e + (\hat{\Theta}_o \hat{\mathcal{H}}_0)_S - \frac{i}{2} d_{\tilde{\varphi}} \hat{\Theta}_o \right\rangle_{\Gamma}. \quad (5.15)$$

Here, D_{Γ} and E_{Γ} are two functions of $\tilde{\varphi}$ which depend on the state Γ of the homogeneous geometry, and which we do not specify because they are irrelevant for our calculations. As

for the momentum $\hat{\pi}_{\tilde{\varphi}}$, it is supposed to act as a generalized derivative with respect to $\tilde{\varphi}$, and we neither need to specify it at this level of discussion. The constraint operator $\hat{\mathcal{C}}_{\text{per}}$ is imposed on the sector of the model composed by $\mathcal{H}_{\text{kin}}^{\text{matt}} \otimes \mathcal{F}$. The corresponding classical evolution generator, denoted by \mathcal{C}_{per} , is then obtained by replacing $\hat{\pi}_{\tilde{\varphi}}$ by its explicit form as a generalized derivative, and the MS operators $\hat{V}_{q_1}^{\vec{n},\epsilon}$ and $\hat{V}_{p_1}^{\vec{n},\epsilon}$ by their classical counterparts, according to our comments above.

Taking into account the densitization of the constraint [set in Eq. (5.1)] and the definition of the homogeneous part of the lapse function, one can check that $\mathcal{C}_{\text{per}}/2$ generates reparameterizations in a time \bar{T} that, at leading perturbative order, is related with the proper time by $dt = \sigma e^{3\alpha} d\bar{T}$. Moreover, we can change to a conformal time η_Γ , adapted to the *dressed* FRW geometry associated with the state Γ . Observing expressions (5.4), we see that all the dependence of $\mathcal{C}_{\text{per}}/2$ on the MS momenta $V_{p_1}^{\vec{n},\epsilon}$ is given by the term $\langle \hat{\vartheta}_e \rangle_\Gamma (V_{p_1}^{\vec{n},\epsilon})^2/2$ coming from $\langle \hat{\Theta}_e \rangle_\Gamma$. Then, it is natural to define¹¹ $l_0 d\eta_\Gamma = \langle \hat{\vartheta}_e \rangle_\Gamma d\bar{T}$. This change of time is well defined because $\langle \hat{\vartheta}_e \rangle_\Gamma$ is a c-number (depending on $\tilde{\varphi}$ through Γ), and it is monotonous inasmuch as the operator $\hat{\vartheta}_e$ representing $\vartheta_e = e^{2\tilde{\alpha}}$ should be positive. It is worth emphasizing that we would have failed to define a change of time parameter had the time derivative $d\eta_\Gamma/d\bar{T}$ been an operator. Hence, the expectation value on Γ is essential in order to introduce the above change of time. We also point out that the change is state dependent, and hence the properties of the evolution in the times \bar{T} and η_Γ can be quite different when considered in the physical Hilbert space of the system.

In summary, the evolution of $V_1^{\vec{n},\epsilon}$ in the time η_Γ is given (under Poisson brackets) by

$$d_{\eta_\Gamma} V_1^{\vec{n},\epsilon} = \frac{l_0}{2\langle \hat{\vartheta}_e \rangle_\Gamma} \{V_1^{\vec{n},\epsilon}, \mathcal{C}_{\text{per}}\}, \quad (5.16)$$

where d_{η_Γ} denotes the derivative with respect to η_Γ . The evolution of $V_{q_1}^{\vec{n},\epsilon}$ is simply given by $d_{\eta_\Gamma} V_{q_1}^{\vec{n},\epsilon} = l_0 V_{p_1}^{\vec{n},\epsilon}$. Using this result and taking again the time derivative we get the following effective MS equations:

$$d_{\eta_\Gamma}^2 V_{q_1}^{\vec{n},\epsilon} = -V_{q_1}^{\vec{n},\epsilon} \left[\tilde{\omega}_n^2 + \langle \hat{\theta}_e + \hat{\theta}_o \rangle_\Gamma \right], \quad (5.17)$$

where we have defined $\tilde{\omega}_n^2 = l_0^2 \omega_n^2$ and

$$\langle \hat{\theta}_e \rangle_\Gamma \equiv l_0^2 \frac{\langle \hat{\vartheta}_e^q \rangle_\Gamma}{\langle \hat{\vartheta}_e \rangle_\Gamma}, \quad \langle \hat{\theta}_o \rangle_\Gamma \equiv l_0^2 \frac{\langle (\hat{\vartheta}_o \hat{\mathcal{H}}_0)_S - \frac{i}{2} d_{\tilde{\varphi}} \hat{\vartheta}_o \rangle_\Gamma}{\langle \hat{\vartheta}_e \rangle_\Gamma}. \quad (5.18a)$$

We note that the last term in the square brackets of this MS equation is a function of only $\tilde{\varphi}$, and hence of time, when the scalar field is evaluated on the solutions to the effective equations. This factor contains quantum modifications with respect to the standard MS equation, which encode the most relevant quantum geometry effects of the homogeneous background. The derived equations are still of harmonic oscillator type with time-dependent frequencies. Besides, no dissipation term appears and the equations are hyperbolic in the ultraviolet regime, where $\tilde{\omega}_n^2$ dominates in the square brackets. We have included the contribution of $\langle \hat{\theta}_o \rangle_\Gamma$, although, in view of our discussion in Appendix B, we expect it to be negligible in practice as it is proportional to \tilde{m}^2 , in the studied case of a mass term as the potential.

¹¹It is possible to see that, with our choice of numeric factors, this definition of conformal time is not sensitive to the choice of l_0 .

Let us say that, although in our analysis we have assumed for simplicity that the potential of the scalar field is given by a mass contribution, actually we can easily extend the discussion to a general potential $W(\tilde{\varphi})$. For that, we simply need to replace in Eq. (5.3c) the two powers of $\tilde{m}^2\tilde{\varphi}^2$ with $2W(\tilde{\varphi})$, and the term \tilde{m}^2 with $W''(\tilde{\varphi})$, whereas the factor $\tilde{m}^2\tilde{\varphi}$ in Eq. (5.3a) has to be replaced with $W'(\tilde{\varphi})$ (here the prime denotes derivative with respect to $\tilde{\varphi}$). Then, for the approximations related with the Born-Oppenheimer ansatz to hold, instead of requiring a small mass, we need to require small variations of the potential.

6 LQC representation for the homogeneous sector

So far in our discussion we have left unspecified the concrete quantization adopted in the FRW-geometry part of the homogeneous sector. For the sake of illustrating our analysis with an example of particular interest, in this section we will adopt the polymeric quantization employed in LQC. There are plenty of references where the details of the polymeric quantization of the unperturbed FRW model can be found (see e.g. [19, 21, 22, 37]). More specifically, we will adhere to the so-called *improved dynamics* prescription of LQC [22] and, among the possible symmetric orderings for the Hamiltonian constraint operator, we will adopt the prescription put forward in Ref. [37], as it proves to be most convenient both for theoretical and practical purposes [38]. Moreover, in addition, for the homogeneous scalar field, we will use the standard Schrödinger quantization which is usually taken in LQC. Besides, in the perturbed model, when representing the operators of the homogeneous sector that are coupled to perturbations, we will follow Ref. [25], that adopts the factor ordering of Ref. [37] as well.

In the loop formalism, the geometry is described by the Ashtekar-Barbero $su(2)$ connection and by the densitized triad [20], that form a canonical pair. In FRW cosmologies, owing to homogeneity and isotropy, they are respectively determined simply by two dynamical variables, c and p , with Poisson bracket equal to $8\pi G\gamma/3$, where γ is the Immirzi parameter [39]. The mentioned improved dynamics scheme accounts for the existence of a minimum non-vanishing eigenvalue Δ of the area operator in LQC. This scheme involves a transformation to the new variables

$$v = \text{sgn}(p) \frac{|p|^{3/2}}{2\pi G\gamma\sqrt{\Delta}}, \quad b = \sqrt{\frac{\Delta}{|p|}} c \equiv \bar{\mu}c. \quad (6.1)$$

The sign of p determines the orientation of the triad. The new variables verify $\{b, v\} = 2$. The variable v is proportional to the volume of the homogeneous model, which is finite for the three-torus spatial topology under study, and given by $V = 2\pi G\gamma\sqrt{\Delta}|v|$. These variables for the homogeneous geometry are related with those of previous sections, $(\tilde{\alpha}, \pi_{\tilde{\alpha}})$, via the canonical transformation

$$e^{\tilde{\alpha}} = \left(\frac{3\gamma\sqrt{\Delta}}{2\sigma} |v| \right)^{1/3}, \quad \pi_{\tilde{\alpha}} = -\frac{3}{2}vb. \quad (6.2)$$

On the other hand, for the zero-mode of the matter scalar field, it is convenient to introduce the following scaling by a constant, $\phi = \tilde{\varphi}/(l_0^{3/2}\sigma)$ and $\pi_\phi = l_0^{3/2}\sigma\pi_{\tilde{\varphi}}$, since this is the usual parameterization employed in the LQC literature.

In LQC one *polymerizes* the connection, something which means that the connection coefficient c has no well-defined operator in the quantum theory but one represents instead its holonomy elements, given by the exponentials of b . As a consequence, the FRW-geometry

sector of the kinematical Hilbert space, $\mathcal{H}_{\text{kin}}^{\text{grav}}$, is the span of basis states $|v\rangle$, with $v \in \mathbb{R}$, normalized with respect to the discrete inner product $\langle v'|v\rangle = \delta_v^{v'}$. The operator \hat{v} , which acts by multiplication, $\hat{v}|v\rangle = v|v\rangle$, has a discrete spectrum. We are taking the reduced Planck constant \hbar equal to one (as in Sec. 5.3) to simplify the notation. The holonomy operators $N_{\pm\bar{\mu}} \equiv e^{\pm i\bar{\mu}c/2} = e^{\pm ib/2}$ produce a constant shift in the label of these states, $\hat{N}_{\pm\bar{\mu}}|v\rangle = |v \pm 1\rangle$, as one can deduce from the commutation relations $[\hat{N}_{\bar{\mu}}, \hat{v}] = i\{\widehat{N_{\bar{\mu}}}, v\}$. As a result of this loop representation, the classical expression $cp = 2\pi G\gamma vb$ gets promoted in the quantum theory to a symmetric version of $\hat{v} \widehat{\sin(b)}$ multiplied by $2\pi G\gamma$, where $\widehat{\sin(b)} = i(\hat{N}_{-2\bar{\mu}} - \hat{N}_{2\bar{\mu}})/2$. We choose the symmetric ordering proposed in Ref. [37], after which the operator representing cp becomes

$$\hat{\Omega}_0 \equiv \frac{1}{2\sqrt{\Delta}} \hat{V}^{1/2} \left[\widehat{\text{sgn}(v)\sin(b)} + \widehat{\sin(b)\text{sgn}(v)} \right] \hat{V}^{1/2}. \quad (6.3)$$

As we have said, for the zero-mode of the scalar field, we adopt a standard Schrödinger representation, with $\hat{\pi}_\phi = i\partial_\phi$ and $\hat{\phi}$ acting by multiplication, so that $\mathcal{H}_{\text{kin}}^{\text{matt}} = L^2(\mathbb{R}, d\phi)$. In total, this LQC representation yields the following expression for the operator $\hat{\mathcal{H}}_0^{(2)}$ [see Eqs. (5.2) and (5.5)]:

$$\hat{\mathcal{H}}_0^{(2)} = \frac{3}{4\pi G} \left(\frac{3}{4\pi G\gamma^2} \hat{\Omega}_0^2 - \hat{V}^2 m^2 \hat{\phi}^2 \right). \quad (6.4)$$

The operator $\hat{\Omega}_0^2$ annihilates the zero-volume state $|v=0\rangle$ and leaves invariant its orthogonal complement. Moreover, it leaves invariant the subspaces $\mathcal{H}_\varepsilon^\pm$ formed by states with support on the semilattices $\mathcal{L}_\varepsilon^\pm = \{\pm(\varepsilon + 4n)|n \in \mathbb{N}\}$, where $\varepsilon \in (0, 4]$. $\mathcal{H}_\varepsilon^\pm$ are separable, in contrast with the original $\mathcal{H}_{\text{kin}}^{\text{grav}}$. Notice also that, in each of these sectors, the homogeneous volume v has a strictly positive minimum ε (or negative maximum $-\varepsilon$). Below we will represent the quadratic contributions of the inhomogeneities to the (zero-mode of) the Hamiltonian constraint by operators that also leave these semilattices invariant, that therefore get superselected. In the following, we will restrict the discussion, e.g., to $\mathcal{H}_\varepsilon^+$, spanned by states with positive $v \in \mathcal{L}_\varepsilon^+$.

Let us consider now the representation of the homogeneous terms entering the quadratic contribution of the inhomogeneities to the zero-mode of the Hamiltonian constraint, namely the operators $\hat{\vartheta}_o$, $\hat{\vartheta}_e$, and $\hat{\vartheta}_e^q$ representing the ϑ -terms in Eqs. (5.3). These terms are affected in principle by some quantization ambiguities. As it was done in Refs. [24, 25], we will introduce a symmetric factor ordering that tries to respect, as far as possible, the assignments of representation made in the homogeneous sector of the system as follows. i) We represent the inverse of the volume with the standard regularization of LQC, namely $\widehat{[1/V]}$ is the cube of the regularized operator

$$\widehat{\left[\frac{1}{V}\right]}^{1/3} = \frac{3}{4\pi G\gamma\sqrt{\Delta}} \widehat{\text{sgn}(v)} \hat{V}^{1/3} \left[\hat{N}_{-\bar{\mu}} \hat{V}^{1/3} \hat{N}_{\bar{\mu}} - \hat{N}_{\bar{\mu}} \hat{V}^{1/3} \hat{N}_{-\bar{\mu}} \right]. \quad (6.5)$$

This operator annihilates the state $|v=0\rangle$ and commutes with \hat{V} , and hence it is well defined on the subspaces $\mathcal{H}_\varepsilon^\pm$. ii) The products of the form $f(\phi)\pi_\phi$, where f is an arbitrary function, are represented with the symmetric factor ordering $(f(\hat{\phi})\hat{\pi}_\phi + \hat{\pi}_\phi f(\hat{\phi}))/2$. iii) We adopt an algebraic symmetrization in factors of the form $V^r g(cp)$, that are promoted to the operators $\hat{V}^{r/2} \hat{g} \hat{V}^{r/2}$, where g is any function and r a real number. This algebraic symmetric factor ordering is also adopted for powers of the inverse volume. iv) Even powers of $\Omega_0 \equiv -l_0^3 \sigma^2 \gamma \pi_{\bar{\alpha}}$

are represented by the same powers of the operator $\hat{\Omega}_0$, as in FRW. And v) in the case of odd powers of Ω_0 , let us say Ω_0^{2k+1} with k equal to an integer, we adopt the representation $|\hat{\Omega}_0|^k \hat{\Lambda}_0 |\hat{\Omega}_0|^k$, where $|\hat{\Omega}_0|$ is the square root of the positive operator $\hat{\Omega}_0^2$ and $\hat{\Lambda}_0$ is defined exactly as $\hat{\Omega}_0$ in Eq. (6.3), but with holonomies of double length. In this way, its action only shifts v in multiples of four units, and hence preserves the superselection sectors of the homogeneous geometry. Following these prescriptions, we obtain the operators

$$\hat{\vartheta}_o = \frac{4}{l_0} \sqrt{12\pi G \gamma m^2 \hat{\phi}} \hat{V}^{2/3} |\hat{\Omega}_0|^{-1} \hat{\Lambda}_0 |\hat{\Omega}_0|^{-1} \hat{V}^{2/3}, \quad (6.6a)$$

$$\hat{\vartheta}_e = \frac{3l_0}{4\pi G} \hat{V}^{2/3}, \quad (6.6b)$$

$$\begin{aligned} \hat{\vartheta}_e^q &= \frac{4\pi G}{3l_0} \left[\frac{1}{V} \right]^{1/3} \hat{\mathcal{H}}_0^{(2)} \left(19 - 32\pi^2 G^2 \gamma^2 \hat{\Omega}_0^{-2} \hat{\mathcal{H}}_0^{(2)} \right) \left[\frac{1}{V} \right]^{1/3} \\ &+ \frac{3m^2}{4\pi G l_0} \hat{V}^{4/3} \left(1 - \frac{8\pi G}{3} \hat{\phi}^2 \right), \end{aligned} \quad (6.6c)$$

which are densely-defined on $\mathcal{H}_\varepsilon^+ \otimes L^2(\mathbb{R}, d\phi)$. These results coincide with those of Ref. [25] except for some factors of the inverse of the volume operator, which appear now as the inverse powers of the volume operator itself. This difference arises from our different choice of densitization for the zero-mode of the Hamiltonian constraint at the classical level (the densitization was done at the quantum level in that mentioned, previous work).

As we have already said, the validity of the Born-Oppenheimer approximation depends both on the particular representation chosen for the above homogeneous operators, and on the properties of the homogeneous states Γ . We refer the reader to Ref. [25] for more comments at this respect regarding the loop quantization.

7 Conclusions

In this work we have developed a covariant formulation of the gravitational system that describes a perturbed flat FRW spacetime with a compact three-torus topology and minimally coupled to a scalar field. In our perturbative scheme, we truncate the action of the model at quadratic order in the perturbations, and focus our attention on scalar perturbations. We expand the spatial dependence in Fourier modes, using a basis of eigenfunctions of the Laplace-Beltrami operator of the spatial sections. Besides, we treat the degrees of freedom of the zero-modes exactly, up to the order of our truncation. The total Hamiltonian of the resulting model is a sum of constraints, formed by a linear combination of the zero-mode of the scalar (or Hamiltonian) constraint, and of all the inhomogeneous modes of two local constraints which are linear in the perturbations. One of them comes from the perturbations of the Hamiltonian constraint and the other corresponds to the perturbation of the momentum (or diffeomorphisms) constraint. The zero-mode of the Hamiltonian constraint is, in turn, formed by the constraint of the unperturbed model plus contributions that are quadratic in the perturbations.

The first important aspect of our treatment is that we do not fix the gauge freedom associated with the linear perturbative constraints. In order to deal with them, we abelianize their algebra. We achieve this by replacing the linear perturbative Hamiltonian constraint with a suitable linear combination of it and of the constraint of the unperturbed model. At the perturbative order of our truncation of the action, this replacement amounts to a redefinition of

the zero-mode of the lapse function. As a result of this process, we are able to parameterize the perturbations of the model with these Abelian linear perturbative constraints and the modes of the MS field, together with their corresponding canonically conjugate variables. Since the MS modes Poisson commute with the constraints, we indeed attain a parameterization for the perturbations fully adapted to gauge invariance.

Another goal of our work consists in extending the canonical transformation from the perturbations to the entire system, namely considering not only the inhomogeneities, but also the homogeneous sector of the model (the zero-modes describing the degrees of freedom present in the unperturbed FRW model). In this way we get at our level of truncation in the perturbations a fully covariant description of the entire symplectic manifold, and not just of its inhomogeneous sector formed by the perturbations, something that would have required treating the zero-modes as fixed, rather than as genuine degrees of freedom. This is an important step towards the quantization of the system, inasmuch as without this completion of the canonical transformation, the system would have lost its symplectic structure. As a result of this canonical transformation, the zero-mode of the lapse gets redefined again, absorbing a quadratic contribution of the perturbations, but keeping the original freedom as an unspecified Lagrange multiplier. In addition, also the Lagrange multipliers accompanying the perturbative linear constraints get redefined similarly (though in this case with linear contributions of the perturbations), so that the final Hamiltonian is a linear combination of (first-class) commuting constraints. We have derived the explicit expression of the spacetime metric in terms of our gauge-invariant phase-space variables and new Lagrange multipliers.

In our construction, we have started with a form of the linear perturbative constraints that has been suitably scaled by powers of the scale factor. It is worth explaining that we might have started as well without this scaling of the perturbative constraints: The difference would have simply led to another definition of the momentum of the zero-mode of the FRW geometry in the final canonical set of variables for the entire system.

Once the classical description has been completed, we have proceeded to quantize the system. For this, we have proposed a hybrid approach that combines any quantization of the homogeneous FRW sector of the model with a more standard field description of the inhomogeneous degrees of freedom. The utility of this approach rests on the assumption that the most relevant quantum geometry effects are those affecting the zero-modes of the system, and therefore may require a more careful analysis, while perturbations, even being quantum, can be described in a more conventional way. In other words, the philosophy behind the hybrid quantization consists in adopting a representation for the homogeneous sector of the model capable to capture its quantum gravity nature, while the perturbations are treated essentially along the lines of QFT in curved spacetimes.

We have followed the Dirac approach of imposing the constraints as quantum operators in order to find physical states of the system. As the classical constraints Poisson commute, it is natural to pass to quantum counterparts that commute as well, and hence we can impose them independently. The linear perturbative constraints generate translations in their canonically conjugate variables. Their quantum imposition then means that physical states do not depend at all on those gauge degrees of freedom. Therefore, thanks to our abelianization of the constraint algebra, we see that the process of imposing the perturbative linear constraints in the quantum theory leads to quantum states that depend only on the MS modes and the homogeneous sector. Remarkably, this is precisely the kind of states obtained with the approach adopted e.g. in Ref. [25], even if the avenue followed in that case included a gauge fixing of the perturbations, although the description of the remaining physical degrees

of freedom was done only in terms of gauge invariants. In this sense our results, in the present work, support those obtained with gauge fixing and show the covariance of the conclusions attained previously, at the considered perturbative level. After the imposition of the linear perturbative constraints, the only constraint on the model is the zero-mode of the Hamiltonian constraint.

To solve this Hamiltonian constraint, we have adopted an ansatz for physical quantum states of Born-Oppenheimer type, and discussed the validity of such approximation. This Born-Oppenheimer ansatz has a similar motivation as the one proposed in Ref. [40]. It regards the zero-mode of the scalar field as an internal clock and assumes that the MS modes do not affect much the motion of the zero-modes of the geometry. The ansatz assumes a separation of the quantum dependence on the FRW geometry and on the MS gauge invariants. If the state of the FRW geometry remains highly peaked with respect to three operators [the operators listed as i)-iii) in the paragraph above Eq. (5.12)] when it changes in the internal time, and this evolution fits suitably that of the unperturbed system, the zero-mode of the Hamiltonian constraint provides a quantum dynamical equation for the MS variables in terms of this time. Furthermore, under certain conditions, the ansatz allows one to introduce a series of approximations that lead to a Schrödinger equation for the MS modes, characterized by a Hamiltonian that is given by the MS Hamiltonian corrected by the fact that its dependence on the homogeneous geometry gets *dressed* with quantum corrections. In this manner, one recovers a QFT for the gauge-invariant perturbations, that propagate in a fixed (and in general non-classical) spacetime.

Generically, when one is dealing with quantum fields in curved spacetimes, there is ambiguity both in the way that one parameterizes the field (since one can introduce scalings by background functions) and in the Fock representation chosen for that field. Remarkably, for the system under study, there exist uniqueness theorems that pick out a preferred parameterization (in fact, a canonical pair for the field) and a class of unitarily equivalent Fock representations for it [33–36]. These choices are selected by the natural criteria of having a vacuum invariant under the symmetries of the background and a unitary implementation of the dynamics. Our parameterization for the MS modes matches precisely this description.¹² No regime with a unitary QFT would have been accessible if we had chosen a different parameterization for the MS field (by means of a time-depending scaling).

Another important conclusion of our analysis is the robustness of the class of dynamical equations that govern the evolution of the MS gauge invariants. We have shown that the dynamics of these invariants are always given by a harmonic oscillator equation, with a frequency that depends on the internal time, as long as the Born-Oppenheimer states are highly peaked with respect to the operators that feel the FRW geometry in the zero-mode of the Hamiltonian constraint, and irrespectively of whether this constraint equation can be approximated by one of Schrödinger type. To arrive at this conclusion, essentially our only additional assumption has been that the effective dynamics of the MS modes can be derived with the substitution of annihilation and creation operators by their counterparts as classical variables. The MS equation obtained in this manner contains quantum modifications to the time-dependent part of the frequency, but the modification is the same for all modes, and in particular it does not affect the ultraviolet behavior of the classical equations for these perturbations in general relativity.

¹²Nonetheless, we emphasize that our strategy of a hybrid approach to the quantization can be easily adapted to other parameterizations.

Let us also note that, in order to obtain the quantum evolution for the perturbations, we have not employed a semiclassical approximation for the homogeneous geometry. We do not even need to exactly solve the unperturbed model to determine the quantum state of the background geometry. Indeed, within our perturbative scheme, it is enough that we consider approximate solutions, whose difference with respect to the exact ones can be neglected perturbatively.

We finish by emphasizing that, if we believe in QFT in curved spacetimes, it is natural to think that there is a deeper quantum regime, before reaching a complete quantum gravity description, in which our hybrid quantization makes sense. This hybrid approach encodes the main quantum gravity effects and it is suitable to potentially predict whether there exist modifications of quantum gravity nature regarding cosmological observables, for instance in the power spectrum of the CMB. It therefore offers a framework to extract physical consequences of quantum gravity in cosmology.

Acknowledgments

The authors are greatly thankful to J. Olmedo for enlightening conversations, assistance in computations with *Mathematica*, and pointing out some important references. In addition, they are grateful to B. Elizaga Navascués, T. Pereira, and S. Tsujikawa for discussions. This work was partially supported by the Spanish MICINN/MINECO Project No. FIS2011-30145-C03-02 and its continuation FIS2014-54800-C2-2-P. M. M-B acknowledges financial support from the Netherlands Organisation for Scientific Research (NWO) (Project No. 62001772).

A Spacetime metric and scalar field in the gauge-invariant formulation

A.1 Zero-modes

In this appendix, we give the explicit expressions of the original canonical variables for the homogeneous sector of the phase-space of our system, $\{w^a\} = \{\alpha, \pi_\alpha, \varphi, \pi_\varphi\}$, in terms of the new canonical set introduced in the text, formed by the variables \tilde{w}^a and $V_l^{\vec{n},\epsilon}$. Substituting Eqs. (3.16) and their corresponding derivatives into Eqs. (4.5) yields the desired expressions:

$$\begin{aligned} \alpha &= \tilde{\alpha} - \frac{1}{2} \sum_{\vec{n},\epsilon} \left[e^{-2\tilde{\alpha}} \left(1 - 3 \frac{\pi_{\tilde{\varphi}}^2}{\pi_{\tilde{\alpha}}^2} \right) (V_{q_1}^{\vec{n},\epsilon})^2 + 2 \frac{e^{-\tilde{\alpha}}}{\pi_{\tilde{\alpha}}} \left(e^{6\tilde{\alpha}} \tilde{m}^2 \tilde{\varphi} + 3 \frac{\pi_{\tilde{\varphi}}^3}{\pi_{\tilde{\alpha}}} + \pi_{\tilde{\alpha}} \pi_{\tilde{\varphi}} \right) V_{q_1}^{\vec{n},\epsilon} V_{q_2}^{\vec{n},\epsilon} \right] \\ &\quad - \frac{1}{2} \sum_{\vec{n},\epsilon} \left[\left(\pi_{\tilde{\alpha}}^2 + 3\pi_{\tilde{\varphi}}^2 + \frac{1}{3} e^{4\tilde{\alpha}} \omega_{\vec{n}}^2 \right) (V_{q_2}^{\vec{n},\epsilon})^2 + \frac{2}{3} \pi_{\tilde{\alpha}} V_{q_2}^{\vec{n},\epsilon} V_{q_3}^{\vec{n},\epsilon} + 2e^{\tilde{\alpha}} \frac{\pi_{\tilde{\varphi}}}{\pi_{\tilde{\alpha}}} V_{q_2}^{\vec{n},\epsilon} V_{p_1}^{\vec{n},\epsilon} \right] \\ &\quad + \frac{1}{2} \sum_{\vec{n},\epsilon} \left[\frac{2}{\pi_{\tilde{\alpha}}} V_{q_2}^{\vec{n},\epsilon} V_{p_2}^{\vec{n},\epsilon} + \frac{1}{3} (V_{q_3}^{\vec{n},\epsilon})^2 \right], \end{aligned} \tag{A.1a}$$

$$\begin{aligned} \pi_\alpha &= \pi_{\tilde{\alpha}} - \sum_{\vec{n},\epsilon} \left[6e^{5\tilde{\alpha}} \tilde{m}^2 \tilde{\varphi} V_{q_1}^{\vec{n},\epsilon} V_{q_2}^{\vec{n},\epsilon} + \left(3e^{6\tilde{\alpha}} \tilde{m}^2 \tilde{\varphi} \pi_{\tilde{\varphi}} + \frac{2}{3} \pi_{\tilde{\alpha}} e^{4\tilde{\alpha}} \omega_{\vec{n}}^2 \right) (V_{q_2}^{\vec{n},\epsilon})^2 \right] \\ &\quad + \sum_{\vec{n},\epsilon} V_{q_1}^{\vec{n},\epsilon} V_{p_1}^{\vec{n},\epsilon}, \end{aligned} \tag{A.1b}$$

$$\begin{aligned} \varphi = & \tilde{\varphi} + \sum_{\vec{n},\epsilon} \left[e^{\tilde{\alpha}} V_{p_1}^{\vec{n},\epsilon} V_{q_2}^{\vec{n},\epsilon} + e^{-\tilde{\alpha}} \left(\pi_{\tilde{\alpha}} + 3 \frac{\pi_{\tilde{\varphi}}^2}{\pi_{\tilde{\alpha}}} \right) V_{q_1}^{\vec{n},\epsilon} V_{q_2}^{\vec{n},\epsilon} + \left(3\pi_{\tilde{\alpha}}\pi_{\tilde{\varphi}} - \frac{1}{2}e^{6\tilde{\alpha}}\tilde{m}^2\tilde{\varphi} \right) (V_{q_2}^{\vec{n},\epsilon})^2 \right] \\ & + \sum_{\vec{n},\epsilon} \left[e^{-\tilde{\alpha}} V_{q_1}^{\vec{n},\epsilon} V_{q_3}^{\vec{n},\epsilon} + \pi_{\tilde{\varphi}} V_{q_2}^{\vec{n},\epsilon} V_{q_3}^{\vec{n},\epsilon} - 3e^{-2\tilde{\alpha}} \frac{\pi_{\tilde{\varphi}}}{\pi_{\tilde{\alpha}}} (V_{q_1}^{\vec{n},\epsilon})^2 \right], \end{aligned} \quad (\text{A.2a})$$

$$\pi_{\varphi} = \pi_{\tilde{\varphi}} - \sum_{\vec{n},\epsilon} \left[e^{5\tilde{\alpha}} \tilde{m}^2 V_{q_1}^{\vec{n},\epsilon} V_{q_2}^{\vec{n},\epsilon} + \frac{1}{2} e^{6\tilde{\alpha}} \tilde{m}^2 \pi_{\tilde{\varphi}} (V_{q_2}^{\vec{n},\epsilon})^2 \right]. \quad (\text{A.2b})$$

A.2 Lapse function and shift vector

The reformulation of the model in terms of gauge invariants for the perturbations leads to the Hamiltonian given in Eq. (4.12), with Lagrange multipliers $\{\bar{N}_0, G_{\vec{n},\epsilon}, K_{\vec{n},\epsilon}\}$, and constraints expressed in terms of the new zero-mode variables $\{\tilde{\alpha}, \pi_{\tilde{\alpha}}, \tilde{\varphi}, \pi_{\tilde{\varphi}}\}$ and new variables for the inhomogeneities $\{V_{q_i}^{\vec{n},\epsilon}, V_{p_i}^{\vec{n},\epsilon}\}$. Comparing Eq. (4.12) with Eq. (3.5), and taking Eq. (4.10) into account, we get (within our perturbative truncation scheme)

$$G_{\vec{n},\epsilon} = g_{\vec{n},\epsilon} + N_0 F_{|1}^{\vec{n},\epsilon}, \quad (\text{A.3a})$$

$$\bar{N}_0 = N_0 + \sum_{\vec{n},\epsilon} (3e^{3\tilde{\alpha}} g_{\vec{n},\epsilon} a_{\vec{n},\epsilon} + N_0 F_{|2}^{\vec{n},\epsilon}), \quad (\text{A.3b})$$

$$K_{\vec{n},\epsilon} = k_{\vec{n},\epsilon} + N_0 \left(F_{-1}^{\vec{n},\epsilon} - 3 \frac{e^{-3\tilde{\alpha}}}{\pi_{\tilde{\alpha}}} V_{p_2}^{\vec{n},\epsilon} + \frac{9}{2} e^{-3\tilde{\alpha}} V_{p_3}^{\vec{n},\epsilon} \right). \quad (\text{A.3c})$$

From these expressions we deduce the inverse relations (again within the order of our perturbative truncation):

$$N_0 = \bar{N}_0 \left[1 + \sum_{\vec{n},\epsilon} (3e^{3\tilde{\alpha}} F_{|1}^{\vec{n},\epsilon} a_{\vec{n},\epsilon} - F_{|2}^{\vec{n},\epsilon}) \right] - 3e^{3\tilde{\alpha}} \sum_{\vec{n},\epsilon} G_{\vec{n},\epsilon} a_{\vec{n},\epsilon}, \quad (\text{A.4a})$$

$$g_{\vec{n},\epsilon} = G_{\vec{n},\epsilon} - \bar{N}_0 F_{|1}^{\vec{n},\epsilon}, \quad (\text{A.4b})$$

$$k_{\vec{n},\epsilon} = K_{\vec{n},\epsilon} - \bar{N}_0 \left(F_{-1}^{\vec{n},\epsilon} - 3 \frac{e^{-3\tilde{\alpha}}}{\pi_{\tilde{\alpha}}} V_{p_2}^{\vec{n},\epsilon} + \frac{9}{2} e^{-3\tilde{\alpha}} V_{p_3}^{\vec{n},\epsilon} \right), \quad (\text{A.4c})$$

where $a_{\vec{n},\epsilon}$ is given in Eq. (3.16a). These formulas define the modes of the lapse function and the shift vector, given in Eqs. (2.3b) and (2.3c), in terms of the variables and Lagrange multipliers of our gauge-invariant formulation, up to terms that are neglected at our truncation order in the perturbations.

To complete the expressions of the spacetime metric components in our formulation, defined by Eqs. (2.3), as well as of the scalar field (2.3d), we just need to substitute $a_{\vec{n},\epsilon}$, $b_{\vec{n},\epsilon}$, and $f_{\vec{n},\epsilon}$ by their values given in Eq. (3.16), and to replace α and φ in terms of $\tilde{\alpha}$ and $\tilde{\varphi}$ [see Eqs. (A.1a) and (A.2a)]. Recall that the difference between old and new variables for the homogeneous sector is quadratic in the perturbations, and therefore some contributions arising from the above replacement may be disregarded as negligible within our perturbative truncations (namely, if they provide quadratic or higher order corrections to the inhomogeneous modes, or cubic or higher corrections to the zero-modes).

B Negligible second derivatives: self-consistency of the approximation

In this appendix we investigate the self-consistency of neglecting the second derivative of the wave function of the MS modes with respect to the homogeneous part of the scalar field in the constraint equation of the Born-Oppenheimer states.

Suppose that, in Eq. (5.12), we can ignore $\langle \hat{\Theta}_o \rangle_\Gamma$ compared to $\langle \hat{\mathcal{H}}_0 \rangle_\Gamma$. This assumption is justified on the basis of our perturbative scheme. Then, from that equation, we get

$$\hat{\pi}_{\tilde{\varphi}}\psi = \frac{1}{2\langle \hat{\mathcal{H}}_0 \rangle_\Gamma} \left[\langle \hat{\Theta}_e + (\hat{\Theta}_o \hat{\mathcal{H}}_0)_S \rangle_\Gamma + i\langle d_{\tilde{\varphi}} \hat{\mathcal{H}}_0 - \frac{1}{2}d_{\tilde{\varphi}} \hat{\Theta}_o \rangle_\Gamma + \langle \hat{\mathcal{H}}_0^{(2)} - (\hat{\mathcal{H}}_0)^2 \rangle_\Gamma \right] \psi - \frac{1}{2\langle \hat{\mathcal{H}}_0 \rangle_\Gamma} \hat{\pi}_{\tilde{\varphi}}^2 \psi. \quad (\text{B.1})$$

Now we act on this equation with $\hat{\pi}_{\tilde{\varphi}}$ and compute the value of the second derivative of ψ . Taking into account that $[\hat{\pi}_{\tilde{\varphi}}, \langle \hat{O} \rangle_\Gamma] = -i\langle d_{\tilde{\varphi}} \hat{O} \rangle_\Gamma$ and eliminating terms which are negligible perturbatively, we arrive at

$$\begin{aligned} & \left[\frac{3i\langle d_{\tilde{\varphi}} \hat{\mathcal{H}}_0 \rangle_\Gamma + \langle \hat{\mathcal{H}}_0^{(2)} - (\hat{\mathcal{H}}_0)^2 \rangle_\Gamma}{2\langle \hat{\mathcal{H}}_0 \rangle_\Gamma} + 2\langle \hat{\mathcal{H}}_0 \rangle_\Gamma \right] \hat{\pi}_{\tilde{\varphi}}^2 \psi = \frac{2i\langle d_{\tilde{\varphi}} \hat{\mathcal{H}}_0 \rangle_\Gamma + \langle \hat{\mathcal{H}}_0^{(2)} - (\hat{\mathcal{H}}_0)^2 \rangle_\Gamma}{2\langle \hat{\mathcal{H}}_0 \rangle_\Gamma} \\ & \times \left[2\langle \hat{\Theta}_e + (\hat{\Theta}_o \hat{\mathcal{H}}_0)_S \rangle_\Gamma + \frac{i}{2}\langle 3d_{\tilde{\varphi}} \hat{\mathcal{H}}_0 - 2d_{\tilde{\varphi}} \hat{\Theta}_o \rangle_\Gamma + \frac{5}{4}\langle \hat{\mathcal{H}}_0^{(2)} - (\hat{\mathcal{H}}_0)^2 \rangle_\Gamma \right] \psi \\ & - i \left[\langle d_{\tilde{\varphi}} \hat{\Theta}_e + d_{\tilde{\varphi}} (\hat{\Theta}_o \hat{\mathcal{H}}_0)_S \rangle_\Gamma + i\langle d_{\tilde{\varphi}}^2 \hat{\mathcal{H}}_0 - \frac{1}{2}d_{\tilde{\varphi}}^2 \hat{\Theta}_o \rangle_\Gamma + \langle d_{\tilde{\varphi}} \hat{\mathcal{H}}_0^{(2)} - d_{\tilde{\varphi}} (\hat{\mathcal{H}}_0)^2 \rangle_\Gamma \right] \psi \\ & - \frac{1}{8} \frac{(\langle \hat{\mathcal{H}}_0^{(2)} - (\hat{\mathcal{H}}_0)^2 \rangle_\Gamma)^2}{\langle \hat{\mathcal{H}}_0 \rangle_\Gamma} \psi - \hat{\pi}_{\tilde{\varphi}}^3 \psi. \quad (\text{B.2}) \end{aligned}$$

Starting with this equation, it is possible to see by iteration (operating repeatedly with $\hat{\pi}_{\tilde{\varphi}}$) whether it is legitimate to assume that the action of $\hat{\pi}_{\tilde{\varphi}}^{n+1}$ on ψ is negligible compared to the action of $\hat{\pi}_{\tilde{\varphi}}^n$, and in particular if this happens for $n = 1$. For this, one can see that it is sufficient that $\langle \hat{\mathcal{H}}_0^{(2)} - (\hat{\mathcal{H}}_0)^2 \rangle_\Gamma$ be negligible compared with linear terms in the perturbations, and that $\langle d_{\tilde{\varphi}} \hat{\mathcal{H}}_0^{(2)} - d_{\tilde{\varphi}} (\hat{\mathcal{H}}_0)^2 \rangle_\Gamma$ be negligible as well compared to terms of quadratic order, all this assuming that

- i) $\langle d_{\tilde{\varphi}} \hat{\Theta}_e + d_{\tilde{\varphi}} (\hat{\Theta}_o \hat{\mathcal{H}}_0)_S \rangle_\Gamma$ is negligible with respect to terms of quadratic order in the perturbations,
- ii) $\langle d_{\tilde{\varphi}} \hat{\mathcal{H}}_0 \rangle_\Gamma$ and $\langle d_{\tilde{\varphi}} \hat{\Theta}_o \rangle_\Gamma$ are at most of the order of quadratic terms.

Let us recall that $-id_{\tilde{\varphi}} \hat{O} = [\hat{\pi}_{\tilde{\varphi}}, \hat{O}] - [\hat{\mathcal{H}}_0, \hat{O}]$. The first term is a generalized derivative with respect to the explicit dependence on $\tilde{\varphi}$ of the considered operator \hat{O} . For the relevant operators in our discussion, this dependence comes exclusively from the potential of the matter scalar field, that we have considered to be given by a mass term. Then, if the possible values of the mass are considerably small, the term $[\hat{\pi}_{\tilde{\varphi}}, \hat{O}]_\Gamma$ may be treated perturbatively, e.g. by expressing the mass value as a certain power of the amplitude parameter of the inhomogeneous perturbations. On the other hand, when \hat{O} is any of the theta-operators, the second term $[\hat{\mathcal{H}}_0, \hat{O}]_\Gamma$ gives a non-vanishing contribution, whose details will depend on the particular

representation chosen for the homogeneous geometry, and on the properties of the state Γ . In consequence, whether the above conditions are true or not will depend on these particular details and should be carefully checked.¹³

References

- [1] A.A. Starobinsky, *Spectrum of relict gravitational radiation and the early state of the universe*, JETP Lett. **30**, 682 (1979);
A.H. Guth *Inflationary universe: A possible solution to the horizon and flatness problems*, Phys. Rev. D **23**, 347 (1981);
A.D. Linde *A new inflationary universe scenario: A possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems*, Phys. Rev. Lett. B **108**, 389 (1982);
Chaotic inflation, Phys. Rev. Lett. B **129**, 177 (1983).
- [2] G. Hinshaw *et al.*, *Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: Cosmological parameter results*, ApJS **208**, 19 (2013);
P.A.R. Ade *et al.* (Planck Collaboration), *Planck 2015 results. XIII. Cosmological parameters*, arXiv:1502.01589.
- [3] P.A.R. Ade *et al.* (BICEP2 Collaboration), *BICEP2 I: Detection of B-mode polarization at degree angular scales*, Phys. Rev. Lett. **112**, 241101 (2014);
BICEP2/Keck, Planck Collaborations, *A joint analysis of BICEP2/Keck array and Planck data*, Phys. Rev. Lett. **114**, 101301 (2015).
- [4] R. Durrer, *Cosmological perturbation theory*, arXiv:astro-ph/0402129;
D. Langlois, *Lectures on inflation and cosmological perturbations*, Lect. Notes Phys. **800**, 1 (2010);
V.F. Mukhanov, H.A. Feldman, and R.H. Brandenberger, *Theory of cosmological perturbations*, Phys. Rep. **215**, 203 (1992).
- [5] A.R. Liddle and D.H. Lyth, *Cosmological inflation and large-scale structure* (Cambridge University Press, Cambridge, U.K., 2000);
J. Martin, *Inflationary cosmological perturbations of quantum-mechanical origin*, Lect. Notes Phys. **669**, 199 (2005).
- [6] J. Ehlers, P. Geren, and R.K. Sachs, *Isotropic solutions of the Einstein-Liouville equations*, J. Math. Phys. **9**, 1344 (1968);
W.R. Stoeger, R. Maartens, and G.F.R. Ellis, *Proving almost-homogeneity of the universe – An almost Ehlers-Geren-Sachs theorem*, Astrophys. J. **443**, 1 (1995).
- [7] E. Lifshitz, *On the gravitational stability of the expanding universe*, Zh. Eksp. Teor. Fiz. **16**, 587 (1946);
E. Lifshitz and I.M. Khalatnikov, *Investigations in relativistic cosmology*, Adv. Phys. **12**, 185 (1963).
- [8] S.W. Hawking, *Perturbations of an expanding universe*, Astrophys. J. **145**, 544 (1966).
- [9] D.W. Olson, *Density perturbations in cosmological models*, Phys. Rev. D **14**, 327 (1976).
- [10] J.M. Bardeen, *Gauge-invariant cosmological perturbations*, Phys. Rev. D **22**, 1882 (1980).
- [11] G.F.R. Ellis and M. Bruni, *Covariant and gauge-invariant approach to cosmological density fluctuations*, Phys. Rev. D **40**, 1804 (1989).
- [12] M. Sasaki, *Gauge invariant scalar perturbations in the new inflationary universe*, Prog. Theor. Phys. **70**, 394 (1983);

¹³For the case when the representation of the homogeneous geometry is that of LQC, which we apply in Sec. 6, we refer the reader to the discussion of Ref. [25].

- H. Kodama and M. Sasaki, *Cosmological perturbation theory*, Prog. Theor. Phys. Suppl. **78**, 1 (1984).
- [13] V. Mukhanov, *Quantum theory of gauge-invariant cosmological perturbations* Zh. Eksp. Teor. Fiz. **94**, 1 (1988); Sov. Phys. JETP **67**, 1297 (1988); *Physical foundations of cosmology*, (Cambridge University Press, Cambridge, U.K., 2005).
- [14] T.S. Bunch and P.C.W. Davies, *Quantum field theory in de Sitter space: Renormalization by point splitting*, Proc. R. Soc. A **360**, 117 (1978).
- [15] J.J. Halliwell and S.W. Hawking, *Origin of structure in the universe*, Phys. Rev. D **31**, 1777 (1985).
- [16] I. Shirai and S. Wada, *Cosmological perturbations and quantum fields in curved space*, Nucl. Phys. B **303**, 728 (1988).
- [17] D. Langlois, *Hamiltonian formalism and gauge invariance for linear perturbations in inflation*, Class. Quantum Grav. **11**, 389 (1994).
- [18] E.J.C. Pinho and N. Pinto-Neto, *Scalar and vector perturbations in quantum cosmological backgrounds*, Phys. Rev. D **76**, 023506 (2007);
F.T. Falciano and N. Pinto-Neto, *Scalar perturbations in scalar field quantum cosmology*, Phys. Rev. D **79**, 023507 (2009).
- [19] M. Bojowald, *Loop Quantum Cosmology*, Living Rev. Rel. **11**, 4 (2008);
G.A. Mena Marugán, *A brief introduction to Loop Quantum Cosmology*, AIP Conf. Proc. **1130**, 89 (2009); *Loop Quantum Cosmology: A cosmological theory with a view*, J. Phys. Conf. Ser. **314**, 012012 (2011);
K. Banerjee, G. Calcagni, and M. Martín-Benito, *Introduction to Loop Quantum Cosmology*, SIGMA **8**, 016 (2012).
- [20] T. Thiemann, *Modern canonical quantum general relativity* (Cambridge University Press, Cambridge, U.K., 2007);
K. Giesel and H. Sahlmann, *Loop Quantum Gravity*, PoS(QGQGS 2011)002.
- [21] A. Ashtekar, T. Pawłowski, and P. Singh, *Quantum nature of the Big Bang: An analytical and numerical investigation*, Phys. Rev. D **73**, 124038 (2006).
- [22] A. Ashtekar, T. Pawłowski, and P. Singh, *Quantum nature of the Big Bang: Improved dynamics*, Phys. Rev. D **74**, 084003 (2006).
- [23] M. Fernández-Méndez, G.A. Mena Marugán, and J. Olmedo, *Hybrid quantization of an inflationary universe*, Phys. Rev. D **86**, 024003 (2012); *Effective dynamics of scalar perturbations in a flat Friedmann-Robertson-Walker spacetime in Loop Quantum Cosmology*, Phys. Rev. D **89**, 044041 (2014).
- [24] M. Fernández-Méndez, G.A. Mena Marugán, and J. Olmedo, *Hybrid quantization of an inflationary model: The flat case*, Phys. Rev. D **88**, 044013 (2013).
- [25] L. Castelló Gomar, M. Fernández-Méndez, G.A. Mena Marugán, and J. Olmedo, *Cosmological perturbations in Hybrid Loop Quantum Cosmology: Mukhanov-Sasaki variables*, Phys. Rev. D **90**, 064015 (2014).
- [26] M. Martín-Benito, L.J. Garay, and G.A. Mena Marugán, *Hybrid quantum Gowdy cosmology: Combining Loop and Fock quantizations*, Phys. Rev. D **78**, 083516 (2008);
G.A. Mena Marugán and M. Martín-Benito, *Hybrid quantum cosmology: Combining Loop and Fock quantizations*, Int. J. Mod. Phys. A **24**, 2820 (2009);
L.J. Garay, M. Martín-Benito, and G.A. Mena Marugán, *Inhomogeneous Loop Quantum Cosmology: Hybrid quantization of the Gowdy model*, Phys. Rev. D **82**, 044048 (2010);
M. Martín-Benito, G.A. Mena Marugán, and E. Wilson-Ewing, *Hybrid quantization: From Bianchi I to the Gowdy model*, Phys. Rev. D **82**, 084012 (2010);

- D. Brizuela, G.A. Mena Marugán, and T. Pawłowski, *Big Bounce and inhomogeneities*, *Class. Quant. Grav.* **27**, 052001 (2010).
- [27] I. Agullo, A. Ashtekar, and W. Nelson, *Extension of the quantum theory of cosmological perturbations to the Planck era*, *Phys. Rev. D* **87**, 043507 (2013); *The pre-inflationary dynamics of Loop Quantum Cosmology: Confronting quantum gravity with observations*, *Class. Quantum Grav.* **30**, 085014 (2013).
- [28] M. Bojowald, G.M. Hossain, M. Kagan, and S. Shankaranarayanan, *Anomaly freedom in perturbative Loop Quantum Gravity*, *Phys. Rev. D* **78**, 063547 (2008); *Gauge invariant cosmological perturbation equations with corrections from Loop Quantum Gravity*, *Phys. Rev. D* **79**, 043505 (2009); *Phys. Rev. D* **82**, 109903(E) (2010);
M. Bojowald, G. Calcagni, and S. Tsujikawa, *Observational constraints on Loop Quantum Cosmology*, *Phys. Rev. Lett.* **107**, 211302 (2011);
T. Cailleteau, L. Linsefors, and A. Barreau, *Anomaly-free perturbations with inverse-volume and holonomy corrections in Loop Quantum Cosmology*, *Class. Quantum Grav.* **31**, 125011 (2014).
- [29] A. Barrau, M. Bojowald, G. Calcagni, J. Grein, and M. Kagan, *Anomaly-free cosmological perturbations in effective canonical quantum gravity*, [arXiv:1404.1018](https://arxiv.org/abs/1404.1018).
- [30] B. Dittrich and J. Tambornino, *Gauge invariant perturbations around symmetry reduced sectors of general relativity: Applications to cosmology*, *Class. Quantum Grav.* **24**, 4535 (2007).
- [31] C. Kiefer and M. Krämer, *Quantum gravitational contributions to the cosmic microwave background anisotropy spectrum*, *Phys. Rev. Lett.* **108**, 021301 (2012).
- [32] A. Corichi, J. Cortez, and G.A. Mena Marugán, *Quantum Gowdy T3 model: A unitary description*, *Phys. Rev. D* **73**, 084020 (2006).
- [33] J. Cortez, G.A. Mena Marugán, J. Olmedo, and J.M. Velhinho, *Criteria for the determination of time dependent scalings in the Fock quantization of scalar fields with a time dependent mass in ultrastatic spacetimes*, *Phys. Rev. D* **86**, 104003 (2012).
- [34] J. Cortez, G.A. Mena Marugán, J. Olmedo, and J.M. Velhinho, *A uniqueness criterion for the Fock quantization of scalar fields with time-dependent mass*, *Class. Quantum Grav.* **28**, 172001 (2011).
- [35] L. Castelló Gomar, J. Cortez, D. Martín-de Blas, G.A. Mena Marugán, and J.M. Velhinho, *Uniqueness of the Fock quantization of scalar fields in spatially flat cosmological spacetimes*, *JCAP* **11**, 001 (2012);
J. Cortez, L. Fonseca, D. Martín-de Blas, and G.A. Mena Marugán, *Uniqueness of the Fock quantization of scalar fields under mode preserving canonical transformations varying in time*, *Phys. Rev. D* **87**, 044013 (2013);
J. Cortez, D. Martín-de Blas, G.A. Mena Marugán, and J.M. Velhinho, *Massless scalar field in de Sitter spacetime: Unitary quantum time evolution*, *Class. Quantum Grav.* **30**, 075015 (2013);
L. Castelló Gomar and G.A. Mena Marugán, *Uniqueness of the Fock quantization of scalar fields and processes with signature change in cosmology*, *Phys. Rev. D* **89**, 084052 (2014).
- [36] M. Fernández-Méndez, G.A. Mena Marugán, J. Olmedo, and J.M. Velhinho, *Unique Fock quantization of scalar cosmological perturbations*, *Phys. Rev. D* **85**, 103525 (2012).
- [37] M. Martín-Benito, G.A. Mena Marugán, and J. Olmedo, *Further improvements in the understanding of isotropic Loop Quantum Cosmology*, *Phys. Rev. D* **80**, 104015 (2009).
- [38] G.A. Mena Marugán, J. Olmedo, and T. Pawłowski, *Prescriptions in Loop Quantum Cosmology: A comparative analysis*, *Phys. Rev. D* **84**, 064012 (2011).
- [39] G. Immirzi, *Quantum gravity and Regge calculus*, *Nucl. Phys. Proc. Suppl.* **57**, 65 (1997); *Real and complex connections for canonical gravity*, *Class. Quantum Grav.* **14**, L177 (1997).

- [40] C. Rovelli and F. Vidotto, *Stepping out of homogeneity in Loop Quantum Cosmology*, *Class. Quantum Grav.* **25**, 225024 (2008).