The Pennsylvania State University
The Graduate School

EFFECTIVE METHODS IN COSMOLOGY AND GRAVITY

A Dissertation in
Physics
by
Ding Ding

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2022

The dissertation of Ding Ding was reviewed and approved by the following:

Martin Bojowald
Professor of Physics
Dissertation Advisor
Chair of Committee

Eugenio Bianchi
Associate Professor of Physics

Radu Roiban
Professor of Physics

Ping Xu
Professor of Mathematics

Nitin Samarth
Professor of Physics
Head of the Department of Physics

Abstract

In this dissertation, I summarize how canonical effective methods can be applied to describe quantum corrected gauge symmetries, non-adiabatic corrections in inflation, and backreactions that seed correlations and inhomogeneities.

Special attention will be paid to the canonical structure and dynamics of quantum fluctuations (which can be interpreted as the moments of a quantum state). I will use canonical methods to highlight the role of constraints in the quantum corrected low energy effective theory of gravity; deformed covariance is shown to be a generic consequence. Reconciliation between seemingly different effective Hamiltonians is also discussed in this context. I then explore the effects of dynamical quantum fluctuations on effective potentials in inflation. I show how these fluctuations affect observables and help us characterize the quantum state, thus establishing a state-to-observable correspondence. Finally, I discuss the role of backreactions on background correlations and inhomogeneity growth. I find that distinct oscillation behaviors arise when correlations are present. Furthermore, the trickling of infrared inhomogeneities to smaller length scales is demonstrated to exist with the inclusion of non-local effects from quantum moments.

Table of Contents

List of Figures

Acknowledgments

I am first and foremost grateful to my advisor Prof. Martin Bojowald for his guidance and encouragement, as well as for providing me with the opportunity to develop independence in my research. From him, I came to appreciate the power and beauty of canonical methods in physics, much beyond what one is typically exposed to in a graduate course. In research, I learned from him how to use my intuition to navigate through the tough waters of technical details. Most of all, in the course of working with Martin, he helped me learn how to maintain my resolve in the face of seemingly insurmountable obstacles.

I would also like to thank my committee members, many of whom gave me spectacular lectures on advanced topics in gravity, cosmology, and quantum field theory. Additionally, I have benefited from many stimulating discussions with colleagues and collaborators, particularly with Suddhasattwa Brahma, Sean Crowe, and Bekir Baytas.

Last but not least, I would like to thank my family for their unconditional love and support; they have fostered for me an environment conducive to science and critical thinking.

The dissertation is based on work that appeared in [3–6]. The studies presented in this dissertation have been in part supported by NSF grant PHY-1607414 and PHY-1912168, along with the Verne M. Willaman Distinguished Graduate Fellowship. Any findings and conclusions appearing in the dissertation are those of the author and do not necessarily reflect the view of the funding agency.

Chapter 1 Introduction

1.1 Motivations

Both quantum field theory and general relativity have had countless triumphs in the past century—from the precise calculation of the anomalous magnetic moment and explaining the perihelion precession of Mercury, to predicting the observation of the Higgs Boson and gravitational waves. All indications seem to suggest the two theories are extremely reliable in their respective domains. Few theorists would dispute this. However, despite countless laborious attempts by some of the greatest minds in modern history, we have yet to unite the two theories. Even more conservative expeditions aimed at only quantizing gravity, without attempting to unify it with the other known forces, have been met with considerable challenges.

One might ask, why bother? If quantum gravity only matters around the Planck scale (roughly 15 orders of magnitude away from what is accessible in near-future colliders), does quantum gravity truly concern 21st-century humans? Surprisingly, the answer is yes. In fact, a more poetic (and perhaps a slightly overdramatic) theorist might argue that it is quite literally a matter of life and death.

The existence of rich chemical and biological processes on Earth owes itself to the structure formation of our early universe. If everything was perfectly homogeneous, it would result in a lifeless world. Structures of our universe originate from the inhomogeneities that are believed to be seeded by quantum processes in our early universe—during a period of fast exponential expansion. This expansion process, known as the paradigm of inflation, happens around a few orders of magnitude below the Planck scale and magnifies the otherwise elusive features of the extremely-early universe. Therefore, we have good reasons to believe that, in the initial states of inflation, there might be footprints left behind by quantum gravity. These footprints could very well still exist in our skies today, offering us clues on how life arose on Earth.

On the other hand, it is surprising and worrisome that the Higgs boson—which concerns the stability of the laws of nature—can potentially be sensitive to Planck scale physics. Quantum corrections from the top quark, the heaviest known fundamental particle, can cause instabilities for the Higgs particle. In an unstable or metastable scenario, the Higgs particle will tunnel to a true stable vacuum, altering fundamental constants of nature and destroying the chemical laws that life depends on. Traditional wisdom assumes that this process is slow and our universe is safe. But recently, more in-depth analysis show that Planck scale physics—quantum gravity included—can change this conclusion. The threat of an electroweak vacuum decay in our universe thus becomes very real. Consequently, Planck scale physics become important for our little 10^4 GeV civilization to understand if (and why) one day we might simply cease to exist. (The catastrophe that is vacuum decay can propagate at the speed of light, meaning that we are unlikely to have any empirical evidence of its materialization until it is too late.)

In both of the aforementioned cases, the best we can do for understanding the quantum physics whose full detail is obscure to us is to adopt the philosophy of effective methods. In this paradigm, a quantum theory reduces to an effective system with new interactions induced by quantum corrections. These interactions, from a Lagrangian perspective, can contain any higher power operator as well as non-local terms that respects the symmetries of the theory. The resulting effective theory is generally difficult to deal with due to these new terms. To make such highly non-linear and non-local theories manageable, one can try to postulate a state where these difficult terms are subdominant or simply vanish. Unfortunately, in such a prescription, the role of non-adiabatic contributions (such as large temporal derivative terms) as well as backreactions becomes obscured and sometimes largely neglected. At the Planck scale or scales close to the inflation era, backreactions and non-adiabatic quantum corrections can become important. Assumptions regarding the background state tend to be too restrictive and may exclude important quantum corrections by fiat. This calls for new methods that are more inclusive toward time-dependent quantum corrections.

Canonical analysis is a powerful tool that can be adapted to account for non-adiabatic evolutions and backreactions. It is also a natural framework within which gauge symmetries, including the diffeomorphism symmetry of general relativity, can be described with algebras of first-class constraints. These topics will be the main focus of this dissertation.

1.2 Outline of the dissertation

In this dissertation, we will focus on the application of canonical effective methods to discuss modified covariance in pure gravity, homogeneous quantum corrections to inflation models, and backreactions of inhomogeneous field theory. Emphasis will be laid upon the description of quantum corrections with effective or semi-classical degrees-of-freedom (DoFs).

We will start our analysis with a spherically symmetric model of gravity. When a quantization is carried out, we expect classical covariance to be modified. While there is little consensus on the specific quantization procedure of gravity, from a purely effective point of view, we expect these modifications to be a generic result of the additional terms induced by integrating out degrees of freedom. These modifications, from the canonical perspective, manifest in the form of modified algebras between first-class constraints. In an effective or semi-classical regime, where the notion of space-time is still expected to hold, space-time covariance is thus modified. While the above conclusions are expected to be independent of quantization schemes, subtleties can arise when different basic variables are used. Gravity is the most constrained theory among all known fundamental interactions—it also has by far the largest gauge symmetry. When the search for generic gauge-invariant variables is still an open issue, we are often forced to pick a set of gauge-dependent variables as a starting point of quantization. The resulting effective corrections to space-time can appear to be variable dependent. Factoring in the fact that first-class constraints can muddy the hierarchy of energy scales, the task of checking the genericness of quantum corrections becomes non-trivial. In chapter 2, we take a small step forward by showing how one can use the closure of constraint algebra as a guideline to find the allowed quantum corrections (up to some fixed order of derivative terms in the action). The role that derivative terms play in deforming covariance is highlighted. We then discuss how first-class constraints can transmute the order and type of derivatives while maintaining the closure of the gauge algebra—this can be used to show equivalence between seemingly different effective theories.

In chapter 3, we will discuss how quantum fluctuations in an inflationary model can induce new DoFs that behave as an additional field driving the expansion. We will mainly focus on quantum corrections induced by the background homogeneous field. Starting from a simple scalar model, we show how the resulting effective system essentially becomes a multi-field model. We will also show how the uncertainty principle and non-adiabatic contributions—which are often neglected or hidden when using the traditional assumptions behind an effective potential calculation—set the initial and final stages of inflation. Standard tools in multi-field inflation will be used to show how the effective model conforms with observations without excess fine-tuning. We find that observables are sensitive to quantities that parameterize the quantum state of the background field.

In chapter 4, we move on to an inhomogeneous model inspired by the cosmological perturbation theory. We will be mainly interested in the interplay between the background field and the inhomogeneous field (which plays the role of perturbations). We show, using canonical effective methods, how correlations between the background and the degree of freedom reflecting backreactions can arise. We also make the surprising discovery of how backreactions allow large-scale modes to trickle down to smaller scales, thus generating inhomogeneity buildup for the background field. The analysis will highlight how canonical effective methods can connect observables with the parameters that reflect the properties of the quantum state.

In chapter 5, we summarize the results of the dissertation and discuss the outlooks of several future directions.

Chapter 2
Gravity

2.1 Brief review of canonical gravity

In this section, we briefly review how gravity can be described using canonical methods. We will first go over the standard 3+1 splitting along with the ADM formulation. The significance of constraints is highlighted. We will see that it is possible to enlarge our gauge symmetries and rewrite our gravitational theory in a form that bears resemblance to Yang-Mills type gauge theories. A key component is the introduction of the Gauss constraint. After a brief review of the Gauss constraint in Yang-Mills theories, we will introduce analogous connection variables to be used in gravity. This will serve as a starting point for our later discussions regarding potential modifications to space-time symmetries arising from quantization. We will mainly be following [7].

2.1.1 Summary of basic results: geometry

Below we list basic results of the kinematics behind a 3+1 splitting of space-time

1. Staring point: a manifold $\mathcal{M} = \mathbb{R} \times \Sigma_t$. Here Σ_t represents a spatial slicing.

2. The slicing Σ_t induces *time arrow* t^μ such that $t^\mu \nabla_\mu t = 1$. (A typical choice of coordinates requires $t^\mu \nabla_\mu = \partial/\partial t$ with $t = x^0$.)

3. The normal to the slicing $n_\mu \propto \partial_\mu t$ [1] helps introduce the *induced metric* $h_{\mu\nu} = g_{\mu\nu} + n_\mu n_\nu$. (We stress that $h_{\mu\nu} \equiv g_{\mu\rho} g_{\nu\sigma} h^{\rho\sigma}$ acts not as the inverse of $h^{\mu\nu}$ on space-time tensors but only on tensors tangent to Σ_t.)

[1] We always pick orientation $n^\mu n_\mu = -1$.

4. *Lapse function* and *shift vector*: $N \equiv -n_\mu t^\mu = 1/(n^\mu \nabla_\mu t)$ (last equality uses $n_\mu \propto \partial_\mu t$ and its normalization) and $N^\mu = h^{\mu\nu} t_\nu$. As a corollary, we may expand $n_\mu = N^{-1}(t_\mu - g_{\mu\nu} N^\nu)$.

5. The full metric can be written as $g_{\mu\nu} = h_{\mu\nu} - N^{-2}(t_\mu - g_{\mu\rho}N^\rho)(t_\nu - g_{\nu\sigma}N^\sigma)$.

6. Coordinates adapted to Σ_t: choose coordinate *functions* $x^\mu = (x^0, x^i)$ such that $t^\mu \nabla_\mu x^0 = 1$ and $t^\mu \nabla_\mu x^i = 0$. (The first requirement ensures $x^0 = t$.)

7. The determinant of the metric, which is a weight $w = 1$ density, is $\sqrt{-g} = N\sqrt{h}$. (Here we are assuming the "natural" coordinate choice, to be explained later. Also, h refers to the determinant of the spatial part of $h_{\mu\nu}$ as is obvious since the full $h_{\mu\nu}$ is not invertible due to $h_{ab}n^b = 0$—it has a zero eigenvalue.)

Coordinate expressions. We will now summarize how the above definition and results lead to coordinate expressions for basic quantities, assuming the choice for coordinate functions $x^\mu = (t, x^i)$.

- $t^\mu = (1, 0, 0, 0)$ and $n_\mu = (-N, 0, 0, 0)$, from the definition of our coordinate choice and lapse. (Note that while the normal is normalized, the time vector t^a is not normalized to 1. Instead $t^\mu t_\mu = g_{00}$ as we will see later.)

- $N^\mu = (0, \mathbf{N})$ and $N_\mu = (N^i N^j h_{ij}, h_{ij} N^j)$.

As for the full metric and induced metric, we have

$$g_{\mu\nu} = \begin{pmatrix} -N^2 + h_{ij}N^iN^j & h_{ji}N^i \\ h_{ij}N^j & h_{ij} \end{pmatrix} \quad \text{and} \quad g^{\mu\nu} = \begin{pmatrix} -\frac{1}{N^2} & \frac{N^j}{N^2} \\ \frac{N^i}{N^2} & h^{ij} - \frac{N^iN^j}{N^2} \end{pmatrix}, \tag{2.1}$$

where $h^{ij} = (h_{ij})^{-1}$. And for the induced metric

$$h_{\mu\nu} = \begin{pmatrix} h_{ij}N^iN^j & h_{ji}N^i \\ h_{ij}N^j & h_{ij} \end{pmatrix} \quad \text{and} \quad h^{\mu\nu} = \begin{pmatrix} 0 & 0^j \\ 0^i & h^{ij} \end{pmatrix}. \tag{2.2}$$

The derivation of the above component expressions requires us to compute different parts of the normal, full metric and induced metric in parallel. A good starting point is to find n^μ using $n^\mu n_\mu = -1$ while taking advantage of the fact that t^μ and n_μ are only non-vanishing in their 0th component.

So far we are using the space-time indexed description of 3+1 decomposition. While this description may seem more covariant, it can be mentally taxing for practical component calculations. Alternatively, one might prefer to use coordinate functions adapted to the hypersurface. In the next section, we will briefly summarize an equivalent and more calculation-friendly description based on embeddings of hypersurfaces.

2.1.2 Embedding and hypersurfaces

In this section, we will briefly review the ideas behind embeddings. We can embed a (hyper)surface Σ into a larger manifold $\mathcal{M}$ with the map

$$X : \Sigma \to \mathcal{M}, \quad y^a \mapsto X^\mu(y^a)\,, \tag{2.3}$$

where y^a are coordinates on the hypersurface Σ and X^μ are coordinates on the manifold $\mathcal{M}$. (For indices of this section, we will follow the convention that $\alpha, \beta \dots \mu, \nu, \rho$ are reserved for space-time coordinates X^μ and a, b, c index the Σ-intrinsic coordinates y^a that appears in $X^\mu(y^a)$. The lowercase indices i, j, k, similar to the previous section, denoted the spatial components of a space-time tensor (which is in general different from the a, b-indexed coordinates unless we pick the special coordinate embedding. However, we will show that objects like h_{ij} are equivalent to the pulled-back metric using the "natural" embedding map which we will soon define).

2.1.2.1 Expressions of components

We will sketch the key steps in deriving the previous stated results for component expressions. We will also show two ways, one geometric and one component-flavored, that allow us to get the expression for the metric.

Starting point: We will use the *natural coordinates* adapted from the intrinsic coordinates on Σ_t. Namely, we will use (2.10) on the constraint surface $\Phi(x^\mu) = t$. First, it is easy to show that

$$t^\mu \nabla_\mu t = 1 \quad \text{and} \quad t^\mu \nabla_\mu x^i = 0 \tag{2.4}$$

are satisfied from (2.10). (The spatial superscript i should be interpreted as a label denoting the coordinate function x^i. Namely, we have three is to specify three independent spatial coordinate functions.)

The coordinate choice (2.10) along with the definition of normal to a constraint surface trivially ensures

$$t^\mu = (1, \mathbf{0}), \quad n_\nu = (n_0, \mathbf{0})\,. \tag{2.5}$$

The definition of the lapse function fixes the normalization for the normal $n_0 = -N$.

The definition of the lapse and the shift leads to $n^\mu = 1/N(t^\mu - N^\mu)$. Additionally, natural coordinates and $N^\mu n_\mu = 0$ ensures $N^\mu = (0, \mathbf{N})$. Combine these two with (2.5) and we get

$$n^\mu = (\frac{1}{N}, -\frac{\mathbf{N}}{N}) , \tag{2.6}$$

which agrees with previous stated results.

The metric: covariant derivation. Now we will utilize $g_{\mu\nu} = h_{\mu\nu} - n_\mu n_\nu$ to obtain coordinate expressions $g_{\mu\nu}$. Recalling that t^μ, n_μ selects out the 0-th component while N^μ selects out the spatial components we have

$$g_{ij} = - n_i n_j + h_{ij} = h_{ij} \tag{2.7}$$

$$g_{0j} = - n_0 n_j + h_{0j} = t^\mu h_{\mu j} = N_j = h_{j\mu} N^\mu = h_{ji} N^j \tag{2.8}$$

$$g_{00} = - t^\mu t^\nu n_\mu n_\nu + t^\mu t^\nu h_{\mu\nu} = -N^2 + t^\mu N_\nu = -N^2 + N_0 = -N^2 + g_{0\mu} N^\mu = -N^2 + h_{ij} N^i N^j \, , \tag{2.9}$$

where the last equality of the last line uses the expression for g_{0j} and the fact that N^μ selects out spatial components when being contracted. Therefore, utilizing t^μ, n_μ and N^ν as projectors, we have derived the components for the full space-time metric in natural coordinates using a relatively geometric perspective. However, we will also show a embedding/coordinate-flavored derivation later. We first need to introduce how we define our embedding.

Natural coordinates and embedding

Physical systems in general relativity allows a canonical decomposition $\mathcal{M} = \mathbb{R} \times \Sigma_t$. This allows for a natural coordinate system which induces an embedding. The natural coordinate system can be thought of as choosing the time coordinate to be $x^0 = t$ (where t parameterizes Σ_t) and the intrinsic coordinates to Σ_t as the spatial coordinates for $\mathcal{M} = \mathbb{R} \times \Sigma_t$. Specifically, this means for the space-time we pick

$$x^0 \equiv X^0(y) = t \quad \text{and} \quad x^{\mu=a} \equiv X^{\mu=a}(y) = y^a \quad \text{with} \quad a = 1, 2, 3 \tag{2.10}$$

and y^a as the intrinsic coordinates to Σ_t. Note that the relations (2.10) defines an embedding (2.3). While the embedding map appears to be defined with coordinates, it is *not* a coordinate dependent map! An embedding map is a map between points of two manifolds Σ and $\mathcal{M}$, in which coordinates $x^\mu \equiv X^\mu$ and y^a merely serve as labels for these points.

A mixed tensorial object like $E^\mu_a = \partial x^\mu/\partial y^a$ is a tangent vector (to the hypersurface Σ_t) with a label "a" as seen by $\mathcal{M}$. This is why the object E^μ_a is actually also a projector from the full space-time to a spatial slice. That is, it eliminates any normal component to the constraint surface

$$E^\mu_a n_\mu(X)|_{\text{on }\Phi} \propto \frac{\partial X^\mu}{\partial y^a}\partial_\mu \Phi \approx 0\,, \tag{2.11}$$

where the weak equivalence $\approx$ means the restriction to the constraint surface $\Phi = C$ that defines the hypersurface Σ_t in $\mathcal{M}$. The above equivalence follows trivially from the fact that $\partial_a \Phi = 0$ by definition of the constraint surface and the intrinsic coordinates y^a on that surface.

The "Jacobian" projector

Given an embedding $X^\mu : \Sigma \to \mathcal{M}$, call the following object the Jacobian projector

$$E^\mu_a \equiv \frac{\partial X^\mu}{\partial y^a}\,, \tag{2.12}$$

where y^a are the intrinsic coordinates to Σ. This Jacobian projector is what appears in the Jacobian. It is also responsible for the pull-backs

$$\Phi^*(\mathbf{g})_{ab} = E^\mu_a E^\nu_b g_{\mu\nu}\,. \tag{2.13}$$

Now suppose we pick the natural coordinates (2.10), clearly $E^\mu_a = \delta^\mu_a$ and vanishes for E^0_a. This means that $h_{ij} \equiv g_{ij}$ is the same as the pull-back of $g_{\mu\nu}$

$$\Phi^*(\mathbf{g})_{ij} = \delta^\mu_i \delta^\nu_j g_{\mu\nu} = h_{ij}\,. \tag{2.14}$$

So from now on we will simply use h_{ij} to also represent the pull-back or the induced metric of $g_{\mu\nu}$ under the natural embedding (2.10).

Following the same line of logic as above, we arrive at a very useful conclusion regarding the natural embedding: $E^\mu_a = \delta^\mu_a$ represents the pull-back to Σ_t and, in natural coordinates (2.10), just trivially selects out the spatial components of the space-time tensors. That is

$$\Phi^*(T)_{ijk\ldots} = E^\mu_i E^\nu_j E^\rho_k \ldots T_{\mu\nu\rho\ldots} = T_{ijk\ldots}\,. \tag{2.15}$$

This conclusion means that even if we start off with the covariant definition of 3+1 splitting (using μ, ν indices), we can calculate intrinsic quantities to Σ_t by only referring to g_{ij}.

The metric: embedding derivation. The previous covariant-indexed derivation does not rely on an explicit embedding. All our derivations refer only to spatial indices of the full space-time. We only had to assume that the slicing of $\mathcal{M}$ into Σ_t is allowed (hence allowing for a normal n_μ to some surface).

An embedding perspective starts with a hypersurface (Σ_t) with its own inherent coordinate functions y^a. It is not until we embed it into the full space-time, are we required to refer to the full space-time (typically through n_μ-related quantities like the extrinsic curvature). The natural coordinates give us a natural embedding that maps all results derived in the embedding perspective to the results of our previous section, which are expressed in i, j indices.

We start with the embedding-induced expansion of the full space-time element $dx^\mu = t^\mu dt + E^\mu_a dy^a$. Here we are not yet assuming x^μ are natural coordinates. Then we can express our invariant distance element as

$$g_{\mu\nu} dx^\mu dx^\nu = g_{\mu\nu}(t^\mu dt + E^\mu_a dy^a)(t^\nu dt + E^\nu_b dy^b) \tag{2.16}$$

$$= t^\mu t^\nu g_{\mu\nu} dt dt + 2t^\mu g_{\mu\nu} E^\nu_a dy^a dt + g_{\mu\nu} E^\mu_a E^\nu_b dy^a dy^b \,. \tag{2.17}$$

To bring in the lapse and shift, we recall that the normal to Σ_t obeys $n_\mu E^\mu_a = 0$ which makes E^μ_a a tangent vector to the hypersurface. We can show that the following are consistent with the geometric definitions used previously

$$t^\mu \nabla_\mu t = 1, \quad N = -t^\mu n_\mu \tag{2.18}$$

$$n_\mu = -N \partial_\mu t, \quad t^\mu = N n^\mu + N^a E^\mu_a \,, \tag{2.19}$$

where we introduced the *3-shift* N^a, which is related to the 4-shift by $N^\mu = E^\mu_a N^a$. (Now it's obvious why the 4-shift has no 0-th component if we use natural coordinates.) We also hope to point out that the above four equalities also ensure n_μ is normalized.

Plugging the expressions for t^μ back into (2.17), we get

$$g_{\mu\nu} = (-N^2 + h_{ab} N^a N^b) dt dt + 2h_{ab} N^a dy^b dt + h_{ab} dy^a dy^b \tag{2.20}$$

$$= -N^2 dt^2 + h_{ab}(dy^a + N^a dt)(dy^b + N^b dt) \,, \tag{2.21}$$

which resembles the more commonly found expression for the metric in an ADM exposition.

The induced metric $h_{ab} = E_a^\mu E_b^\nu g_{\mu\nu}$ here is the pull-back of the full space-time metric onto the hypersurface. It is the same as the spatial components of the space-time metric, "h_{ab}" = "h_{ij}" ≡ "g_{ij}", if we chose natural coordinates (2.10).

2.1.3 Dynamics of hypersurfaces: the ADM formalism

Here we list the main results of the Arnowitt-Deser-Misner (ADM) formalism, which describe how hypersurfaces evolve in time. A reminder on notation: The space-time tensors will be labeled by Greek indices μ, ν, ρ while tensors tangent to the embedding surface (i.e. the ones orthogonal to the hypersurface normal) will be labelled by indices a, b, c.

The 3+1 decomposition of the space-time allows us to express the (space-time) Ricci scalar R into quantities that are intrinsic and extrinsic to the hypersurface. Intuitively, the extrinsic quantities tell us how the space-time is actually decomposed into slices using n_μ, while the intrinsic quantities describe how the resulting slices Σ_t look geometrically with the help of spatial covariant derivative $\mathcal{D}_a$. The dynamics of the theory require a time direction, which we can supply with the previously defined time-flow vector (field) $t^\mu = t^\mu(N, N^\mu, n_\mu)$. Specifically, the vector field t^μ identifies points on different Σ_t slices as the same spatial point—points P and P' that have the same spatial coordinate, e.g. $y^a(P) = y^a(P')$. (Geometrically, the vector uses the data on an initial slice Σ_{t_0} and generates a congruence of curves. For a given curve in the congruence, the different intersections between this curve and different future slices are identified as the same spatial point.) What Einstein's equations need to do, in this context, is to make sure that this identification is physically correct—the evolution of physical distance between two spatially separated points are in line with observations.

The geometry intrinsic to a spatial slice is described by h_{ab} while the extrinsic properties are described by how n_μ "moves" along the spatial hypersurface. Similar to the space-time case, to describe how a tensor moves along a direction we can use the covariant derivative. In our case we define the *spatial covariant derivative* as the pull-back of space-time covariant derivative onto the spatial slice. Using the projector we previously described, we have

$$\mathcal{D}_a T_c^b \equiv E_a^\mu E_\nu^b E_c^\rho \nabla_\mu T_\rho^\nu \, , \tag{2.22}$$

For the purpose of ADM formulation, the projector E_a^μ is chosen to be the one associated with natural embedding (2.10).

Things get easier if we are acting on a **spatial tensor** S_c^b. Because the projector annihilates n_μ after contraction, one can prove that

$$\mathcal{D}_a S^b_c = \partial_a S^b_c + \Gamma^b_{ad} S^d_c - \Gamma^d_{ac} S^b_d \,, \tag{2.23}$$

meaning that when acting on a spatial tensor, $\mathcal{D}_a$ is compatible with the spatial metric and is simply the covariant derivative associated with the "spatial" Christoffel symbol produced by h_{ab}.

Now we can introduce the extrinsic quantity of interest—the extrinsic curvature K_{ab}

$$K_{ab} \equiv E^\mu_a E^\nu_b \nabla_\mu n_\nu \,. \tag{2.24}$$

Using orthogonality between the normal and the projector, along with $E^\mu_{[a} \nabla_{|\mu|} E^\nu_{b]} = 0$ (easily proven with (2.10)), we can show that

$$K_{ab} = K_{ba} \,. \tag{2.25}$$

Another important identity is

$$K_{ab} = \frac{1}{2N} (\dot{h}_{ab} - 2\mathcal{D}_{(a} N_{b)}) \,, \tag{2.26}$$

where $(\dot{\ldots}) = \mathcal{L}_t$, and is the same as $\partial_t h_{ab}$ since h_{ab} is spatial (however $\dot{h}_{ab}$ is not spatial). This identity strongly hints that the extrinsic curvature is related to the conjugate momentum of h_{ab}.

To obtain the ADM action, we use the Gauss-Codazzi relation to rewrite

$$R =^{(3)} R + K_{ab} K^{ab} - (K^a_a)^2 + 4\nabla_a (n^{[a} \nabla_c n^{c]}) \tag{2.27}$$

$$^{(3)}R_{abc}{}^e \omega_e \equiv [\mathcal{D}_a, \mathcal{D}_b] \omega_c \,, \tag{2.28}$$

where we have defined the spatial Ricci scalar $^{(3)}R$ associated with h_{ab}. Plugging (2.27) into the Einstein-Hilbert action we obtain

$$S_{\mathrm{E-H}}[h_{ab}, N, N^c] = \int \mathrm{d}t L_{\mathrm{E-H}} = \frac{1}{16\pi G} \int \mathrm{d}^4 x N \sqrt{h} (^{(3)}R + K_{ab} K^{ab} - (K^a_a)^2) \,, \tag{2.29}$$

where we have ignored boundary terms. (The boundary terms are important for general relativity. Indeed, generic space-times do not always admit a good fall-off condition. In fact, the Einstein-Hilbert action requires an additional term, called the *Gibbons-Hawking-York* boundary term, to admit a good variational problem. However, boundary terms do not change any of our following results, so we will simply set them aside for now and assume the variational principle works.)

2.1.3.1 The ADM Hamiltonian and constraints

Now we wish to obtain the Hamiltonian for the ADM action (2.29). To obtain the symplectic structure, we only need to look at the terms with the extrinsic curvature because the spatial Ricci scalar does not contain any time derivatives. Following this strategy we find the momenta to be

$$
\begin{aligned}
p_N =& \frac{\delta L_{\text{E-H}}}{\delta \dot{N}} = 0 \\
p_{N^a} =& \frac{\delta L_{\text{E-H}}}{\delta \dot{N}^a} = 0 \\
p^{ab} =& \frac{\delta L_{\text{E-H}}}{\delta \dot{h}_{ab}} = \frac{\sqrt{h}}{16\pi G}(K^{ab} - K^c_c h^{ab}) ,
\end{aligned}
$$

where we used $K_{ab} \sim 1/2N\dot{h}_{ab}$ in the last line. After a Legendre transform (with the help of multipliers) we obtain the Hamiltonian

$$
\begin{aligned}
H_{\text{E-H}} =& \int \mathrm{d}^3x \left(\frac{16\pi G N}{\sqrt{h}}(p_{ab}p^{ab} - \frac{1}{2}(p^c_c)^2) - \frac{N\sqrt{h}}{16\pi G}{}^{(3)}R + 2p^{ab}D_a N_b \right) \\
&+ \int \mathrm{d}^3x (\lambda p_N + \mu^a p_{N^a}) .
\end{aligned} \tag{2.30}
$$

The fact that momenta of N and N^a vanish is a primary constraint. The restriction that these constraints are preserved throughout evolution gives rise to secondary constraints called the *Hamiltonian (scalar)* and *diffeomorphism (vector)* constraints

$$
C_{\text{Ham}} \equiv - \dot{p}_N = -\{p_N, H_{\text{E-H}}\} = 0 \tag{2.31}
$$

$$
C^{\text{Diff}}_a \equiv - \dot{p}_{N^a} = -\{p_{N^a}, H_{\text{E-H}}\} = 0 \tag{2.32}
$$

Because the shift and lapse appear in the Hamiltonian (2.30) linearly (after integration by parts on the $2p^{ab}\mathcal{D}_a N^b$ term), a trivial rearrangement gives

$$
H_{\text{E-H}} = \int \mathrm{d}^3x (N C_{\text{Ham}} + N^a C^{\text{Diff}}_a + \lambda p_N + \mu^a p_{N^a}) . \tag{2.33}
$$

The Hamiltonian for gravity is a fully constrained one!

2.1.3.2 Interpreting the constraints

If one computes the bracket between the smeared secondary constraints in (2.33) one would arrive at the *hypersurface deformation algebra* (HDA)

$$\{D[N^b], D[M^a]\} = D[\mathcal{L}_{N^b} M^a] \tag{2.34}$$

$$\{D[N^a], H[N]\} = H[\mathcal{L}_{N^a} N] \tag{2.35}$$

$$\{H[N], H[M]\} = - D[h^{ab}(N\partial_b M - M\partial_b N)] \tag{2.36}$$

$$D[N^a] \equiv \int \mathrm{d}^3x N^a C_a^{\mathrm{Diff}} \tag{2.37}$$

$$H[N] \equiv \int \mathrm{d}^3x N C_{\mathrm{Ham}} \,. \tag{2.38}$$

The fact that brackets between constraints equal to constraints means that the Hamiltonian and diffeomorphism constraints are first-class. Physically, this means that they generate gauge transformations—in our case these are the space-time diffeomorphisms that make gravity so unique. (Also, notice that the resulting smearing function of the $H - H$ bracket now is a function of phase-space variables. A consequence of this is that brackets between constraints like $\int N(x)C_A$ and $\int M(h_{ab}(y))C_B$ can contain a $(\dots)\{C_A, h^{ab}\}C_B$ term. Ultimately, it means that the gauge transformations are only generated after imposing the on-shell condition in the final step of the bracket computation.)

We will show more concrete examples of how first-class constraints generate gauge transformations when we talk about Yang-Mills theory in later sections.

2.1.3.3 Example calculation: inflationary space-times

In this section, we will illustrate how constraints and gauge redundancies are dealt with in the ADM formalism using the example of inflationary (quasi de Sitter) space-times. These space-times are essentially the perturbed versions of the familiar FLRW space-time. The perturbed metric, in this case, is not unique due to gauge redundancies induced by the first-class constraints introduce in the previous section. The logic of our calculation is straightforward (following [8]): We chose a gauge which then specifies a form of the perturbed metric, solve the constraints perturbatively and plug the solution back into the original action to obtain a new action that contains only the true degrees of freedom.

We start with a general action that contains a scalar field minimally coupled to gravity (in units where reduced Planck mass is set to 1)

$$S = \frac{1}{2}\int \mathrm{d}^4x\sqrt{|g|}(R - g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - 2V(\phi))\,. \tag{2.39}$$

Using (2.1), (2.2) and $\sqrt{|g|} = N\sqrt{h}$ (with $h \equiv \det(h_{ab})$) we have

$$S =\frac{1}{2}\int \mathrm{d}^4xN\sqrt{h}(R + N^{-2}(\dot{\phi} - N^a\partial_a\phi)^2 - h^{ab}\partial_a\phi\partial_b\phi - 2NV) \tag{2.40}$$

$$\int \mathrm{d}^4x\sqrt{|g|}R = \int \mathrm{d}^4xN\sqrt{h}({}^{(3)}R + K_{ab}K^{ab} - (K^a_a)^2) + \text{boundary terms}\,, \tag{2.41}$$

where we used the Gauss-Codazzi relations to decompose R. (Note that in the Lagrangian formalism K_{ab} is a function of N as well as derivatives of N^a and h_{ab}, while ${}^{(3)}R$ only depends on derivatives of h_{ab}.)

We want to expand the action (2.39) perturbatively around the FLWR metric to obtain a quadratic action. So up to what order do we need to solve the constraints? One might naively think we need everything up to second order. But fortunately, solving the constraints up to first order in lapse and shift is sufficient! Here is the proof: The Lagrangian is a functional of lapse, shift, and the spatial metric (which we denote with the shorthand notation h in the following). We now expand the Lagrangian in perturbations of only lapse and shift (where δ^n represents the n-th order of perturbation)

$$\begin{aligned}
N =&N_0 + \delta^1N + \delta^2N + \dots\\
N^a =&N^a_0 + \delta^1N^a + \delta^2N^a + \dots\\
L[N, N^a, h] \approx&L[N_0, N^a_0, h]\\
&+ \frac{\delta L}{\delta N}[N_0, N^a_0, h](\delta^1N + \delta^2N + \dots) + \frac{\partial L}{\delta N^a}[N_0, N^a_0, h](\delta^1N^a + \delta^2N^a + \dots)\\
&+ \frac{1}{2}\frac{\delta^2L}{\delta N\delta N}L[N_0, N^a_0, h](\delta^1N + \delta^2N + \dots)^2\\
&+ \frac{1}{2}\frac{\delta^2L}{\delta N^a\delta N^b}[N_0, N^a_0, h](\delta^1N^a + \delta^2N^a + \dots)(\delta^1N^b + \delta^2N^b + \dots)\\
&+ \frac{\delta^2L}{\delta N\delta N^a}[N_0, N^a_0, h](\delta^1N + \delta^2N + \dots)(\delta^1N^a + \delta^2N^a + \dots)\,.
\end{aligned} \tag{2.42}$$

Clearly, if we are keeping up to second order, we only need to examine up to quadratic $\delta^1(\dots)$ terms and linear $\delta^2(\dots)$ terms. This means that $\delta^2(\dots)$ terms can only come from the second line of the $L[N, N^a, h]$ expansion, which is multiplied with the constraint $\delta L/\delta N$ and $\delta L/\delta N^a$ evaluated with zeroth-order lapse, shift, and metric (metric too because we are only keeping up to second order). They vanish. In fact, the constraints evaluated at zeroth order is simply one of the background equations of motion (the Hubble-parameter squared equation). The story is similar for h_{ab} expansion: We need not worry about $\delta^2 h_{ab}$ when restricting the Lagrangian up to second order because these terms are multiplied with the background equations of motion for h_{ab}.

We fix our gauge to be the comoving gauge defined by

$$\delta\phi = 0, \quad h_{ab} = a^2((1+2\zeta(x))\delta_{ab} + \gamma_{ab}), \quad N = 1+\alpha, \quad N_a = \partial_a\psi + \tilde{N}_a \tag{2.43}$$

$$\gamma^a_a = 0, \quad \partial^a\gamma_{ab} = 0, \quad \partial_a\tilde{N}^a = 0\,. \tag{2.44}$$

Note that the indices of perturbations (or derivatives thereof) can simply be raised with $a^2\delta^{ij}$ since we only need to keep up to the first order. Due to similar reasons, the spatial covariant derivatives acting on shift vectors can be replaced with spatial partial derivatives because the Christoffel symbols contain only spatial derivatives of the metric—they are proportional to perturbations.

After a lengthy calculation one can show that the diffeomorphism constraint gives

$$\mathcal{D}_b(K^b_a - K^c_c h^b_a) \approx 2\partial_a(\alpha H - \dot{\zeta}) - \frac{1}{2}\partial^b\partial_b\tilde{N}_a = 0\,, \tag{2.45}$$

which can be solved by

$$\alpha = \frac{\dot{\zeta}}{H}, \quad \tilde{N}_a = 0\,. \tag{2.46}$$

Using these two solutions we can show that Hamiltonian constraint becomes

$$^{(3)}R - (2V + \frac{1}{1+2\alpha}\dot{\phi}^2) - (K^{ab}K_{ab} - (K^a_a)^2) \approx -\frac{4}{a^2}\delta^{ab}\partial_a\partial_b\zeta + 2\frac{\dot{\zeta}}{H}\dot{\phi}^2 - 4\frac{H}{a^2}\delta^{ab}\partial_a\partial_b\psi = 0\,, \tag{2.47}$$

which can be solved by

$$\delta^{ab}\frac{\partial_a\partial_b\psi}{a^2} = -\delta^{ab}\frac{\partial_a\partial_b\zeta}{a^2 H} + \frac{1}{2}\frac{\dot{\phi}^2}{H^2}\dot{\zeta}\,. \tag{2.48}$$

(There is an apparent a^{-2} factor difference before the term $\partial_a\partial_b\psi$ compared to Maldacena's original result. This is because Maldacena defines ψ as $N^i = \partial_i\psi + N^i_T$, which differs to our ψ by a factor of a^{-2}. We also wish to point out that it is common in literature for $(\partial)^2$ to be defined as $\delta^{ab}\partial_a\partial_b$ rather than being raised with h^{ab}. In our results, all $a, b, c\ldots$ indices are raised with the inverse spatial metric h^{ab}.)

After plugging the above solutions back into the action, using equations of motion and integration by parts we obtain the quadratic part of the perturbations

$$S_2 = \frac{1}{2}\int \mathrm{d}^4 x a^3 \frac{\dot{\phi}^2}{H^2}(\dot{\zeta}^2 - a^{-2}\delta^{ab}\partial_a\zeta\partial_b\zeta)\,. \tag{2.49}$$

This action can be directly used for the quantization of perturbations.

2.1.4 Gravity as a gauge theory

We have shown previously how the gravitational Hamiltonian can be described with canonical variables h_{ab} and its conjugate p^{ab}. However, these are not the only variables allowed. In fact, it is possible to use tetrad and connection variables to make the gravitational Hamiltonian look more similar to the one in a Yang-Mills theory. In the process, we will however have to pay a price by introducing additional redundancies. Fortunately, these redundancies can be interpreted as the familiar ones generated by the Gauss constraint in a standard Yang-Mills theory.

2.1.4.1 Lightning review of Yang-Mills

In this section, we aim to show how classical pure Yang-Mills theory exhibit gauge freedom due to constraints. The calculation in gravity is (non-trivially) analogous but much more tedious. In gravity, we also have more complicated constraints and gauge group(oid)s making the corresponding algebra contain phase-space dependent structure functions (as opposed to constants). However, by understanding the process for the Yang-Mills case, it will become clear how we can use triads and connections as basic variables for gravity.

Definition and conventions: we make the following representation-independent conventions

$$[T^a, T^b] = if^{abc}T^c, \quad \mathrm{Tr}[T^aT^b] = T_R\delta^{ab}\,, \tag{2.50}$$

where the structure constants f^{abc} are *totally anti-symmetric* and T_R is the *index* of a representation—for $SU(N)$, they are $T(\text{def.}) \equiv T_F = 1/2$ and $T(\text{adj}) \equiv T_A = N$. (In the context of $SU(N)$, the name defining rep. and fundamental rep. will be used interchangeably.) The adjoint representation is defined by its generator

$$T_A^{abc} = -if^{abc}\,. \tag{2.51}$$

A note on representation: There are two important representations used in the context of Yang-Mills: the fundamental/defining rep. and the adjoint rep. Matter fields transform naturally in the fundamental rep.

$$\psi_i \to (e^{i\alpha^a T^a})_{ij}\psi_j \,. \tag{2.52}$$

Preservation of local gauge symmetry requires our gauge/connection field to transform (infinitesimally) as

$$A_\mu^a \to A_\mu^a + \frac{1}{g}\partial_\mu \alpha^a + f^{abc} A_\mu^b \alpha^c \,, \tag{2.53}$$

where the structure constant refers naturally to an adjoint action $\mathrm{ad}_A(\alpha) = [A, \alpha]$. This results in the field strength to transform as

$$F_{\mu\nu}^a \to F_{\mu\nu}^a - f^{abc}\alpha^b F_{\mu\nu}^c \,, \tag{2.54}$$

which again means that the field strength transforms in the adjoint representation. Also note that when acting with $\mathcal{D}$ on gauge fields or functions thereof we should use the adjoint action

$$\begin{aligned}(\mathcal{D}_\rho F_{\mu\nu})^a =& \partial_\rho F_{\mu\nu}^a - igA_\rho^c T_A^{cab} F_{\mu\nu}^b && (2.55)\\ =& \partial_\rho F_{\mu\nu}^a - gA_\rho^c f^{cab} F_{\mu\nu}^b \\ =& \partial_\rho F_{\mu\nu}^a - gi[T^c, T^b]^a A_\rho^c F_{\mu\nu}^b \\ =& \partial_\rho F_{\mu\nu}^a - ig[A, F_{\mu\nu}]^a \,, && (2.56)\end{aligned}$$

where we have used the definition of $\mathcal{D}$ and the total anti-symmetry of indices for the structure constants.

The pure Yang-Mills action [2] is written as

$$S_{YM} = -\frac{1}{2}\int \mathrm{Tr}[F_{\mu\nu}F^{\mu\nu}]_{\text{fund.}} = -\frac{1}{4}\int F_{\mu\nu}^a F^{\mu\nu a}|_{\text{adj.}} \equiv \int \mathcal{L}_{YM} \tag{2.57}$$

$$F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a - ig[A_\mu, A_\nu]^a = 2\partial_{[\mu}A_{\nu]}^a + f^{abc}A_\mu^b A_\nu^c \,. \tag{2.58}$$

To get the momentum we first compute

$$\frac{\partial \mathcal{L}_{YM}}{\partial \partial_\mu A_\nu^f} = -\frac{1}{4}2\frac{\partial}{\partial \partial_\mu A_\nu^f}(2\partial_{[\alpha}A_{\beta]})^a F_{\rho\sigma}^a \eta^{\alpha\rho}\eta^{\beta\sigma} \tag{2.59}$$

[2] We are ignoring the topological θ term for now.

$$= -\,(\delta^{[\mu}_{\alpha}\delta^{\nu]}_{\beta})\eta^{\alpha\rho}\eta^{\beta\sigma}F^f_{\rho\sigma} \tag{2.60}$$

$$= -\,F^f_{\rho\sigma}\eta^{\rho[\mu}\eta^{\nu]\sigma}\,, \tag{2.61}$$

where we have used the fact that $\partial/(\partial\partial_\mu A^f_\nu)$ does not see the non-derivative terms in $F^a_{\alpha\beta}$ and $\eta^{\mu\nu}$ is the Minkowski metric. The momentum is therefore obtained if we restric to $\mu = 0$

$$\pi^{\nu f} = -\frac{1}{2}(F^{0\nu f} - F^{\nu 0 f}) = F^{\nu 0 f}\,. \tag{2.62}$$

Recall that the "electric" and "magnetic" fields can be defined as ($A^\mu = (\phi, \mathbf{A})$ in E&M)

$$E^i_a = -\pi^i_a = F^{0i}_a \xrightarrow{U(1)} -\partial_0 A_i - \partial_i\phi \tag{2.63}$$

$$B^i_a = \frac{1}{2}\varepsilon^{ijk}F^{jk}_a\,. \tag{2.64}$$

Notice that the electric field is the negative of the conjugate momentum for A_i. (Depending on the context, the above definitions can differ by an overall sign.) This is why, in gravity, triad variables are sometimes called electric fields as they are conjugate to the connection variable.

Now it is clear that, due to anti-symmetry of the field strength, we have our primary and thus also secondary constraint

$$\pi^\nu \approx 0 \tag{2.65}$$

$$\dot{\pi}^\nu \approx 0\,. \tag{2.66}$$

The secondary constraint is expected to give us the non-Abelian version of Gauss's law

$$\mathcal{D}_i E^i \approx 0\,, \tag{2.67}$$

which reduces to the $U(1)$ Gauss's Law when the group is Abelian (meaning that the adjoint action vanishes).

To show explicitly that the secondary constraint implies Gauss's law, we use the Euler-Lagrange equation for fields (denoting $\mathcal{L} \equiv \mathcal{L}_{YM}$)

$$\partial_\alpha \frac{\partial\mathcal{L}}{\partial\partial_\alpha A^f_0} - \frac{\partial\mathcal{L}}{\partial A^f_0} = 0\,. \tag{2.68}$$

Isolating the $\dot{\pi}^0_f = \partial_0(\partial\mathcal{L}/\partial\dot{A}^f_0)$ part of the EoM we obtain

$$\dot{\pi}^0_f = \frac{\partial\mathcal{L}}{\partial A^f_0} - \partial_j \frac{\partial\mathcal{L}}{\partial\partial_j A^f_0} \,. \tag{2.69}$$

Using (2.58), the first term on the RHS is

$$\begin{aligned}
\frac{\partial\mathcal{L}}{\partial A^f_0} =& -\frac{1}{4}\frac{\partial}{\partial A^f_0}(F^a_{\mu\nu}F^{\mu\nu a}) && (2.70)\\
=& -\frac{1}{2}g(f^{afc}A^c_\nu F^{0\nu a} + f^{abf}A^b_\mu F^{\mu 0a}) && (2.71)\\
=& -f^{afc}A^c_\nu F^{0\nu a} && (2.72)\\
=& ig[A^c_i T^c, F^{0ia}T^a]^f \,, && (2.73)
\end{aligned}$$

where we are allowed to use $\partial(F^a_{\mu\nu}F^{\mu\nu a}) = 2\partial F^a_{\mu\nu}F^{\mu\nu a}$ in the second line because the non-commutative matrix part already has been traced into a δ^{ab}, so every quantity in the calculation commutes.

Using again (2.58), the second term on the RHS of (2.69) is

$$\begin{aligned}
-\partial_j \frac{\partial\mathcal{L}}{\partial\partial_j A^f_0} =& \partial_j\left(\frac{1}{4}2\frac{\partial}{\partial\partial_j A^f_0}(2\partial_{[\mu}A^a_{\nu]})F^{\mu\nu a}\right)\\
=& \partial_j\left(\delta^j_{[\mu}\delta^0_{\nu]}F^{\mu\nu f}\right)\\
=& \partial_j F^{j0f}\\
=& -\partial_j E^{jf} \,, && (2.74)
\end{aligned}$$

where have we used the definition for the electric field, $F^{0j} = E^j = E_j$. Finally, plugging (2.73) and (2.74) into (2.69) we get

$$\dot{\pi}^0 = -(\partial_i F^{0if} - ig[A_i, F^{0i}]^f) = (\mathcal{D}_i\pi^i)^f = -(\mathcal{D}_i E^i)^f \approx 0 \,. \tag{2.75}$$

This is Gauss's law/constraint for Yang-Mills.

2.1.4.2 Constraint algebra and gauge transformations

Given a Lagrangian, students are typically taught to look for gauge transformations by using the naked eye. In simple Lagrangians, this simple trick is good enough. However, the canonical (Hamiltonian) formalism offers a much more systematic way of finding and describing gauge redundancies, which is especially useful for non-linear systems such as gravity.

In the context of a Yang-Mills theory, a canonical analysis reveals the Gauss constraint to be a non-trivial constraint—categorized as a *first-class constraint*, meaning the algebra of constraints is closed. Indeed, the Gauss constraint is expected to generate the original gauge transformation of the theory. This is not a trivial fact and needs to be checked. Specifically, we should check whether: (1) the constraint algebra is indeed isomorphic to the original algebra of the gauge group; and (2) the constraint does indeed recover the correct expression for gauge transformations.

The gauge algebra. In the phase space of Yang-Mills, the generators of gauge transformations (not to be confused with generators of the Lie group) are the first-class constraints—in our case it is the smeared version of the Gauss constraint (2.75)

$$G[N^a(x)] = -\int \mathrm{d}^3x N^a(x)(\mathcal{D}_i E^i)^a \,, \tag{2.76}$$

where we only smear over the spatial slice as we are in the canonical formalism. (The -1 factor can be absorbed if we express Gauss constraint using the canonical momentum to A_i, as is often done in the context of gravity.)

Using the fact that $\{A_i^a(x), E^{jb}(y)\} = -\delta_i^j \delta^{ab}\delta(x-y)$ (up to some absorbable constants of π) we have

$$\begin{aligned}
\{G[N^a(x)], G[M^d(y)]\} = &\int\int N^a(x)M^d(y) \\
&\times \Big(\{\partial_i E^{ia}(x) + gf^{abc}A_i^b(x)E^{ic}(x), \partial_j E^{jd}(y) + gf^{def}A_j^e(y)E^{jf}(y)\}\Big) \\
= &-\int\int N^a(x)M^d(y)\Big(- gf^{daf}E^{jf}(y)\partial_j\delta_{xy} + gf^{adc}E^{ic}(x)\partial_i\delta_{yx} \\
&+ g^2(f^{afc}E^{jc}(x)f^{def}A_j^e(x) - f^{abe}A_j^b(x)f^{def}E^{jf}(x))\delta_{xy}\Big) \,,
\end{aligned} \tag{2.77}$$

where we have introduced the notation $\partial_i\delta_{xy} = \partial_{x^i}\delta(x-y) = -\partial_i\delta_{yx}$ and that ∂_i always means derivative with respect to x^i or y^i depending on whether it acts on δ_{xy} or δ_{yx}.

We can simplify the Poisson bracket further by using the trick

$$E^{jf}(y)\partial_i\delta_{yx} = E^{jf}(x)\partial_i\delta_{yx} - \partial_i E^{jf}(x)\delta_{xy} \,, \tag{2.78}$$

where its validity requires the existence (but not the use) of an overall $\int \mathrm{d}^3x$. The trick allows us to simplify the first line of the last equality in (2.77)

$$-gf^{daf}E^{jf}(y)\partial_j\delta_{xy} + gf^{adc}E^{ic}(x)\partial_i\delta_{yx} = -\,gf^{daf}\partial_j E^{jf}(x)\delta_{xy}\,, \tag{2.79}$$

where we used $\partial_i\delta_{xy} = -\partial_i\delta_{yx}$ and the anti-symmetry of structure constants.

The second line can be simplified with some index gymnastics and the Jacobi identity

$$\begin{aligned}
&g^2(f^{afc}E^{jc}f^{def}A^e_j - f^{abe}A^b_j f^{def}E^{jf})\delta_{xy} \qquad (2.80)\\
&= g^2A^e_jE^{jc}(f^{afc}f^{def} - f^{aef}f^{dfc})\\
&= g^2A^e_jE^{jc}(-f^{acf}f^{fde} - f^{fae}f^{cdf})\\
&= g^2A^e_jE^{jc}f^{daf}f^{fce}\\
&= g^2f^{daf}(-1)(-i)[A_j, E^j]^f\\
&= ig^2f^{daf}[A_j, E^j]^f\,. \qquad (2.81)
\end{aligned}$$

Combining (2.79) with (2.81) we get back the factor $\mathcal{D}_jE^j$ inside a smeared Gauss constraint. Therefore, we ultimately obtain

$$\{G[N^a(x)], G[M^d(y)]\} = G[f^{adf}gN^a(x)M^d(y)] \tag{2.82}$$

and thus recover the algebra of the gauge group through structure constants f^{adf}. (The factor g can be gotten rid of if we define our action with a re-scaled A_μ.) In the case of E&M the structure constant vanishes and the Gauss constraints form the algebra of the $U(1)$ gauge group.

Gauge transformations. From the bracket between the constraints, it is clear that the Gauss constraints are first-class. This means that they generate gauge transformations. We will show that we can generate the familiar Yang-Mills gauge transformations with $\{\ ,G[N^a]\}$, making the Gauss constraints bona fide generators of gauge transformations.

A straightforward computation reveals

$$\begin{aligned}
\{A^a_i(x), G[N^b(y)]\} &= -\int N(y)^b\{A^a_i, \partial_jE^{jb} + gf^{bce}A^c_jE^{ej}\}\\
&= -\int N(y)^b(\delta^{ab}\partial_i\delta_{yx} + gf^{bca}A^c_i\delta_{yx})\\
&= \int \partial_iN^a + gf^{acb}A^c_iN^b\\
&= (\mathcal{D}_iN)^a\,. \qquad (2.83)
\end{aligned}$$

This agrees with the standard result (2.53) for infinitesimal gauge transformations

$$A_i \to A_i + \mathcal{D}_i N \xrightarrow{U(1)} A_i + \partial_i N \,. \tag{2.84}$$

We stress that $N = N^a T^a$ is a Lie-algebra valued object and should be acted on using the adjoint action; otherwise, we cannot obtain the correct $U(1)$ limit. (The transformation law for A_0 is not obtained as the variable is not dynamical and instead can be seen as a Lagrangian multiplier associated with the secondary constraint.)

2.1.4.3 Gauge fields for gravity: tetrads and connections

While the metric represents an intuitive distance-measuring object, using it to express the action for gravity leads to a highly non-linear Hamiltonian. The equations of motion, in this case, are also second-order partial differential equations. In this sense, both classical calculations and quantization can be difficult when using the metric as a fundamental field. Sometimes it is worthwhile to look for first-order formulations of a system by viewing certain derivative terms in the Lagrangian as independent variables. When this is done, we typically pay the price of having more equations of motion (albeit of lower order) while also introducing redundancies in the form of new gauge symmetries. A familiar example of introducing redundancies is when we introduce 4-vector potentials in classical electrodynamics instead of using the electric and magnetic fields. In gravity, we can do something similar by using tetrads and connections as our basic fields.

On our space-time manifold, we can find four independent vectors to describe a basis. It is typically convenient to combine these independent vectors into a set of four (Lorentzian)-orthogonal basis vectors, called tetrads e^μ_I, obeying

$$g_{\mu\nu} e^\mu_I e^\nu_J = \eta_{IJ} \,, \tag{2.85}$$

where I labels which basis vector it is and μ denotes the usual space-time components. We can think of index I as one referring to an internal Lorentzian space (an $SO(1,3)$ index). An example of a set of tetrads is the four orthonormal vectors that describe the frame of a time-like observer, called frame fields.

Tetrads offer a two-way map between I-indexed and μ-indexed objects thanks to its invertability in both I and μ indices

$$\begin{aligned} e^{\mu K} e_{\mu J} = g_{\mu\nu} e^\mu_I e^\nu_J \eta^{IK} = \eta_{IJ} \eta^{IK} &\Rightarrow e^{\mu K} e_{\mu J} = \delta^K_J \\ e_{\mu I}(e^\mu_J e^J_\nu) = g_{\mu\rho} e^\rho_I e^\mu_J e^J_\nu = \eta_{IJ} e^J_\nu = e_{\nu I} &\Rightarrow e^\mu_J e^J_\nu = \delta^\mu_\nu \,, \end{aligned} \tag{2.86}$$

where the last line used the fact that tetrads $e_{\mu J}$ make up a basis (in μ index), so if $e^\mu_J e^J_\nu$ acts like a δ^μ_ν on $e_{\mu J}$, it does so on all μ-indexed objects.

Due to (2.85), if we know the tetrads we can effectively reconstruct the local geometry by using its invertibility

$$\eta^{IJ} e^\mu_I e^\nu_J = g^{\mu\nu} \, . \tag{2.87}$$

However, we immediately notice that a Lorentz transformation of the form $e^a_I \to \Lambda^J_I e^a_J$ reproduces the same metric. So we have introduced a new gauge symmetry on top of the already existent diffeomorphism symmetries. This new gauge symmetry is very similar to the gauge symmetry of $SU(N)$ in Yang-Mills theory. Indeed, we will show later that it is closely related to the Gauss constraint.

We can understand our construction intuitively by imagining gluing an internal space to every point in space-time. Naturally, we would hope to have a way of comparing internal vectors v^I at different points. Just like on a curved space-time manifold or Yang-Mills, we would need a connection (space-time-)one-form $\omega_{\mu IJ}$ for the construction of a covariant derivative $\mathcal{D}_\mu$ with the following properties

$$\begin{aligned}
(\mathcal{D}_\mu v)^I =& \nabla_\mu v^I + \omega_{\mu KJ} \eta^{KI} v^J \\
\mathcal{D}_\mu(\Lambda v)|_{\omega'} =& \Lambda \mathcal{D}_\mu(v)|_{\Lambda^{-1}\omega'\Lambda + \Lambda^{-1}\nabla\Lambda} \\
\mathcal{D}_\mu(\eta_{IJ}) =& 0 \Rightarrow \omega_{\mu IJ} = -\omega_{\mu JI} \, ,
\end{aligned} \tag{2.88}$$

where the subscript on the second line indicates what connection is being used in the covariant derivative $\mathcal{D}_\mu$.

Geometry with tetrads. From a kinematical point of view, we hope that by knowing tetrads we would know about the geometry. This can be done by using the first and second structure equations (assuming there is no torsion)

$$\mathrm{d}\mathbf{e} = \mathbf{e}_J \wedge \boldsymbol{\omega}_I{}^J \tag{2.89}$$

$$\mathbf{R}_I{}^J = \mathrm{d}\boldsymbol{\omega}_I{}^J + \boldsymbol{\omega}_I{}^K \wedge \boldsymbol{\omega}_K{}^J \, , \tag{2.90}$$

where we suppressed the space-time indices on which the differential-form operators act. (The first equation can be obtained by demanding that the tetrads are constants as seen by the covariant derivative. Using the same assumption, we can obtain the second equation by fully contracting the tetrads with the space-time Riemann tensor and using the definition $R \sim [\nabla, \nabla]$.)

If one starts off with the known tetrads, then using (2.89) one can solve for $\boldsymbol{\omega}_I{}^J$. Then plugging into the second structure equation (2.90) we would find the (mixed-)Riemann tensor. The original space-time Riemann tensor can be obtained by mapping the mixed Riemann tensor $\mathbf{R}_{IJ} \equiv R_{\mu\nu IJ} dx^\mu \wedge dx^\nu$ back using tetrads.

Now we know what to expect from a candidate dynamical theory built from tetrads and connections. Whatever action we come up with, the resulting equations of motion should be consistent with kinematical equations (2.89) and (2.90). It should also tell about the dynamics, namely how the tetrads evolve, and it better be consistent with Einstein's equations!

2.1.4.4 Action for gravity

We hope to write the theory of gravity akin to a gauge theory. We know that a Yang-Mills action is written in terms of a 2-form that is the field strength. In gravity, the mixed curvature tensor seems like a good candidate. However, the Yang-Mills action is quadratic in field strength while we expect the gravitational action to be linear in the (contracted) curvature tensor. So we need objects, like the tetrads, to contract with the indices of a linear curvature 2-form. Furthermore, the integration over space-time would also require that the measure d^4x is multiplied with some density of weight 1, with the convention that $\sqrt{|g|}$ is a weight-1 scalar density. (This fact is important in dynamical theories of gravity because the excess weights are neutralized by determinants of the metric, which is itself a function of phase space variables.) The above observations lead us to try the following action

$$S[e,\omega] \sim \int \mathrm{d}^4x |e| e^\mu_I e^\nu_J F^{IJ}_{\mu\nu}(\omega) \sim \int \epsilon_{IJKL} \mathbf{e}^K \wedge \mathbf{e}^L \wedge \mathbf{F}^{IJ} \tag{2.91}$$

$$\mathbf{F}^{IJ} = \mathrm{d}\boldsymbol{\omega}^{IJ} + \boldsymbol{\omega}^{IK} \wedge \boldsymbol{\omega}_K{}^J , \tag{2.92}$$

where using the definition (2.85) and invertibility of the tetrad, the determinant of tetrads is shown to be the same as the determinant of the metric $|e| \equiv |\det(e^I_a)| = \sqrt{-\det(g_{\mu\nu})} = \sqrt{|g|}$.

We emphasize that while the definition of the curvature tensor (2.92) resembles the second structure equation (2.90), the connections in the curvature tensor $\mathbf{F}$ need not obey the first structure equation. It is only after the equations of motion are imposed that the curvature tensors agree. (To distinguish, we will call the $\mathbf{F}$ the curvature tensor, while referring to $\mathbf{R}$ as the Ricci curvature tensor.) Indeed, the equations of motion can be obtained by varying the action with respect to $\mathbf{e}^I$ and $\boldsymbol{\omega}^{IJ}$

$$\varepsilon^{\mu\nu\rho\sigma}\epsilon_{IJKL}\mathcal{D}_\mu(e^K_\mu e^L_\nu) = 0 \quad (\text{from } \delta\boldsymbol{\omega}^{IJ}) \tag{2.93}$$

$$\varepsilon^{\mu\nu\rho\sigma}\epsilon_{IJKL}e^I_\mu F^{JK}_{\nu\rho} = 0 \quad (\text{from } \delta e^I_\mu)\,, \tag{2.94}$$

where the first equation gives us the compatibility condition between the connection and the tetrad. Therefore it leads us to the first structure equation (2.89). The second equation can be shown, after some Levi-Civita index gymnastics, to be equivalent to Einstein's equation in vacuum.

So we have shown that (2.91) is a good action that gives rise to the correct local equations of motion. However, similar to the theta term in Yang-Mills theory, we can add a topological term to the our gravitational Lagrangian, parameterized by the so called *Barbero-Immiriz parameter*

$$\mathcal{L} = \mathcal{L}_{\text{original}} + \mathcal{L}_\gamma \tag{2.95}$$

$$\mathcal{L}_\gamma = \frac{1}{2\gamma}|e|e^\mu_I e^\nu_J \epsilon^{IJ}{}_{KL} F^{KL}_{\mu\nu}(\omega)\,. \tag{2.96}$$

From now on, we will assume that our full Lagrangian contains the $\mathcal{L}_\gamma$ term. (Our guess for the action (2.91) has an overall constant ambiguity. It is expected to contain Newton's constant G. The exact expression for the constant can be fixed by coupling gravity to matter and demanding that we regain Einstein's equations with matter. The result is an overall constant of $1/(16\pi G)$ for the first $\sim$ relation in (2.91).)

2.1.4.5 ADM for a first-order formulation

Even though we are now using tetrad variables e^μ_I instead of the space-time metric $g_{\mu\nu}$, we still need to perform an ADM splitting (in space-time, i.e. in index μ) in order to carry out a canonical analysis. This splitting will allow us to use triads instead of *tetrads* and deal with *spin connections* Γ and extrinsic curvature instead of the original connection 1-form ω. A canonical analysis in this context will reveal 4 types of constraints: the original Hamiltonian (scalar) constraint and diffeomorphism (vector) constraints, the Gauss constraint similar to Yang-Mills, and lastly a constraint which we will call the *compatibility constraint* for lack of a better name. The compatibility constraint combined with the Gauss constraint will give us the compatibility condition between the spatial covariant derivative (defined using Γ) and the triads, as well as relate components of the connection 1-form ω to the extrinsic curvature. The ADM splitting and canonical analysis are tedious but straightforward. So instead of deriving them in detail, we will simply list important results while only sketching how they are obtained.

First, let us clarify some notations. Regarding indices we will use the following convention

	Original space-time ($g_{\mu\nu}$)	
Greek letters like	μ, ν, ρ	for space-time indices
Coordinate names like	t, x, y, z	for specific space-time components
Lowercase letters like	a, b, c	for spatial indices of space-time
	Internal space-time (η_{IJ})	
Capital letters like	I, J, K	for internal space-time indices
Numbering of components	$0, 1, 2, 3$	for specific internal space-time components
Lower case letters like	i, j, k	for internal spatial indices.

Additionally, while we will not a priori assume the use of natural embedding coordinates (2.10), it will be worthwhile to use the intuition they provide; the most immediate one being the observation that contractions with n_μ and t^μ select out the "temporal" components of the contracted indices. This intuition will help us keep track of which objects are considered spatial (or on the slice Σ_t) and which are temporal based on their orthogonality with n_μ and t^μ.

The ADM split. With now two spaces (the original space-time and the internal space), we ought to do two splittings. The original space-time can be split up using the original method. The key relation we will be using is simply $n^\mu = 1/N(t^\mu - N^\mu)$.

Moving on to the internal space. We first construct, using the tetrads, a space-time vector $\mathcal{E}_I^\mu$ that lies on Σ_t

$$\mathcal{E}_I^\mu = e_I^\mu + n^\mu n_\nu e_I^\nu \,.$$

We can check that $\mathcal{E}_I^\mu n_\mu = \mathcal{E}_I^\mu n^I = 0$ (where $n^I = e_\nu^I n^\nu$) is satisfied. So this $\mathcal{E}_I^\mu$ is a spatial vector on Σ_t. Therefore, when the I index is restricted to the spatial ones, we will call objects that are proportional to $\mathcal{E}_I^\mu$ a *triad*.

Because $g_{\mu\nu} = \eta_{IJ} e_\mu^I e_\nu^J$, we have an additional $SO(1,3)$ redundancy when using tetrads. It will be convenient to reduce this redundancy with partial gauge fixing

$$e_0^\mu = n^\mu \,,$$

which can be done by demanding $n^I = (1,0,0,0)$ in the usual Minkowski coordinates. This partial gauge fixing allows us to select out the temporal components of internal space tensors, that are later shown to be related to the extrinsic quantities that make up canonical conjugates and multipliers for constraints. (This is similar to what happens in the original ADM construction using metric variables. A rough analogy is: n^I is to internal space what t^μ is to space-time in the original ADM construction.) Kinematically, we now have everything we need for a canonical analysis. With the introduction of $\mathcal{L}_\gamma$ we will see that we can find at least two sets of canonical variables to use.

Finding a canonical pair. While we could look for the canonical pairs with brute force calculation, it is tedious and not very telling. The structure of the theory will become much clearer if we instead try to finesse our way through by utilizing analogies with the Yang-Mills theory.

Starting with the action

$$S[e,\omega] = \frac{1}{16\pi G}\int \mathrm{d}^4x|e|e^\mu_I e^\nu_J(\delta^{[I}_K\delta^{J]}_L - \frac{1}{2\gamma}\epsilon^{IJ}{}_{KL})F^{KL}_{\mu\nu}(\omega)\,, \tag{2.97}$$

we look for a canonical variable as starting point. While our action is very different from a Yang-Mills action, the expression of the curvature tensor (analogous to the field strength) along with the role the connection 1-form plays in the covariant derivative hints that the connection should be the fundamental field, while the tetrad plays the role of the canonical momentum. So we can try using the tetrad as a canonical momentum. However, there is one more subtlety. In curved space-time, the Legendre transform of a covariant action typically contains a term like

$$\int \mathrm{d}^3x\sqrt{h}p_\varphi\dot{\varphi}\,.$$

This indicates that densitized quantity $\sqrt{h}p_\varphi$ should be the conjugate to $\dot{\varphi}$ instead of p_φ. In the context of tetrads and connections, this observation suggests that we should define a candidate canonical momentum as the tetrad with a $\sqrt{h}$ factor

$$P^\mu_i = \frac{\sqrt{h}}{8\pi\gamma G}\mathcal{E}^\mu_i\,. \tag{2.98}$$

We will call these vectors *densitized triads*. So far we have not proven anything, but are only defining objects. Our goal is to show that starting with (2.97), P^μ_i is indeed a canonical momentum to the canonical coordinate that is *some* combination of the connection ω.

We first re-express all the triads e_I^μ in terms of P_I^μ, n^μ and n_I. Then combined with the expansion of $n^\mu = 1/N(t^\mu - N^\mu)$, we can identify the time-derivative terms associated with P_i^μ by picking out, in the action, any term that contains a factor of the form

$$P_i^\mu t^\nu \partial_\nu (\dots)^i_\mu \,. \tag{2.99}$$

(This trick relies on the assumption that we are using the temporal coordinate natural to t^μ. A more coordinate-independent way of finding "time-derivative" terms is by looking for Lie derivative terms $\mathcal{L}_t$ along t^μ. However, this more generic method requires integration by parts as an additional step.) Our previous ADM split (including the partial gauge fixing $n^I = (1,0,0,0)$) and anti-symmetry of indices ensures that many terms in the expanded action vanish after contraction. Consequently, the terms that fit our criteria turn out to be

$$\int \mathrm{d}^4x P_j^\nu t^\mu \partial_\mu (\gamma \omega_\nu^{0j} + \frac{1}{2}\epsilon^j{}_{kl}\omega_\nu^{kl}) \subset S[e,\omega] \,. \tag{2.100}$$

Immediately, we see that the canonical conjugate to P_j^ν is

$$A_\nu^j \equiv \frac{1}{2}\epsilon^j{}_{kl}\omega_\nu^{kl} + \gamma\omega_\nu^{0j} \,. \tag{2.101}$$

This quantity is called the *Ashtekar-Barbero connection* (AB connection). Note that with coordinates (2.10), the triad is "spatial", $P_j^\nu = P_j^a \delta_a^\nu$, due to its orthogonality with $n_\nu = (1,0,0,0)$. So only the pull-back of A_ν^j (to Σ_t) can be considered as the canonical momentum to P_j^ν. Therefore, whenever we write A_μ^j and P_j^ν (upper and lower indices matter), we will implicitly assume that the pull-back is already performed, namely, their Greek letter indices will only be spatial. (In fact, we could define a quantity by naively forcing $\nu = t$ in (2.101). This term does exist in the action but its time derivative does not. Eventually, the term A_t^j will turn out to be a multiplier for the Gauss constraint, similar to A_0 in electrodynamics.)

With a bit of foreshadowing, we can rewrite the AB connection as

$$\begin{aligned}
A_\mu^i =& \Gamma_\mu^i + \gamma K_\mu^i \\
\Gamma_\mu^i =& \frac{1}{2}\epsilon^i{}_{kl}\omega_\mu^{kl} \\
K_\mu^i =& \omega_\mu^{0i} \,,
\end{aligned}$$

where after imposing the Gauss and compatibility constraint, K_μ^i will become the (mixed-index) extrinsic curvature. Here, Γ^i_μ is the *spin connection* that appears in the spatial covariant derivative defined as

$$\mathcal{D}_\mu v^i \equiv \nabla_\mu v^i + h^\nu_\mu \omega_\nu{}^i{}_j v^j = \nabla_\mu v^i - \epsilon^i{}_{jk}\Gamma^j_\mu v^k \,. \tag{2.102}$$

Finding a second canonical pair. One might wonder what happens if we did not introduce $\mathcal{L}_\gamma$ in our action. Obviously, one should not try to find out by naively taking $\gamma \to 0$ as $P^\mu_i \propto 1/\gamma$ will blow up. What ought to be done is to take $\gamma \to \infty$ after the following canonical transformation and obtain

$$(A^i_\mu, P^\nu_j) \to (\gamma^{-1}A^i_\mu, \gamma P^\nu_j) \xrightarrow{\gamma\to\infty} (K^i_\mu, \frac{1}{8\pi G}E^\nu_j)\,, \tag{2.103}$$

where we have introduced the densitized triad $E^\mu_i \equiv \sqrt{h}\mathcal{E}^\mu_i$. After imposing the some of the previously mentioned constraints, we are left with a second canonical pair that allows us to describe the geometry with extrinsic curvature and densitized triads. This second canonical pair has the advantage of being independent of γ whose value can be imaginary, leading us to a complex phase space. In quantum theory, the value of γ can lead to complications regarding both the quantization procedure as well as the allowed semi-classical corrections.

The Hamiltonian and constraints. The (primary) constraints in the first-order formulation with triads are found, again, by looking for independent combinations of connections that have no time derivatives in the Lagrangian. The original Hamiltonian (scalar) and diffeomorphism (vector) constraints are easy to find since they are expected to be proportional to N and N^μ just like the case in the original ADM formulation (because the old constraints are obtained from the Poisson bracket between the conjugate to the lapse or shift and the Hamiltonian). Indeed, after our previous substitution of $e^\mu_I = e^\mu_I(\mathcal{E}^\mu_I, n)$ and $n^\mu = n^\mu(t, N, N^\mu)$ into the action (2.97), the new constraints resulting from the use of triads and connections must be proportional to t^μ. This makes our search easier, especially since we have already made sense of some of the $t^\mu\partial_\mu(\dots)$ terms in the previous section. Employing this strategy, it is straightforward to find the terms of interest are contained in the action as

$$\int \mathrm{d}^4x \left(A^j_t \mathcal{D}^{(A)}_\mu P^\mu_j + (1+\gamma^2)\epsilon_{jm}{}^n \omega_t{}^{0j}\omega_\mu{}^{0m} P^\mu_n \right) \subset S[e,\omega]\,, \tag{2.104}$$

where we had to integrate by parts and do some gymnastics with $\epsilon - \epsilon$ contractions. The superscript on the covariant derivative $\mathcal{D}^{(A)}$ means that the connection associated with it is the AB connection (instead of ω) and A_t^j is (2.101) but with $\nu = t$. These t-subscripts are results of contractions with t^μ.

Just like we hinted before, the term A_t^j indeed does not have any time derivatives, and neither does ${\omega_t}^{0j}$. In an extended Hamiltonian H_{ext} where they are considered canonical variables, their momenta is 0, which is a primary constraint. The secondary constraints are obtained by demanding their conjugate momenta

$$\dot{\pi}_{A_t^j \text{ or } {\omega_t}^{0j}} = \{\pi_{A_t^j \text{ or } {\omega_t}^{0j}}, H_{\text{ext}}\} = 0\,. \tag{2.105}$$

These conditions give us the remaining 2 of the 4 types of constraints we mentioned previously

$$G_j \equiv \mathcal{D}_a^{(A)} P_j^a = 0 \tag{2.106}$$

$$S_j \equiv \epsilon_{jm}^{\ \ n} {\omega_a}^{0m} P_n^a = \epsilon_{jm} K_a^m P_n^a = 0\,, \tag{2.107}$$

where we recall that a, b, c indicate spatial indices of space-time objects. Here, G_j is the Gauss constraint and S_j is the compatibility constraint. The two combined together ensures $\mathcal{D}$ (the spatial covariant derivative) is compatible with P_j^a along with the justification for K_a^i being the mixed-index extrinsic curvature.

There is one more remaining issue we need to address before we can write out the Hamiltonian: Our search for time derivatives (or absence thereof) is not yet exhaustive. So far we have only looked for the conjugates to the triads, which have accounted for specific combinations of connections ω. As a simple analogy, in a 2-dimensional plane, $x+y$ is not a faithful representation of independent coordinates x and y. However, $x+y$ along with x *can* faithfully represent the two independent coordinates. In our context, we have only talked about $A = \Gamma + \gamma K$ (indices suppressed). To cover all independent components of connection ω, we also check whether Γ has any canonical conjugate. One can check that it doesn't and we obtain more constraints. These constraints, with the help of the Gauss and compatibility constraint, can be partially solved to give us an expression for the spin connection Γ in terms of the triads. In this sense, we say that the spin connections are compatible with the triads.

Now we can write down the Hamiltonian using only first-class constraints. This is because by using compatible spin connections in the action, we can now simply ignore the secondary (and second class) constraint $S_i = 0$ due to the fact that

$$\mathcal{D}_a^{(A)} P_j^a - \mathcal{D}_a P_j^a \propto \gamma S_j = 0$$

is already imposed as a result of the Gauss constraint and spin connection compatibility. The Hamiltonian for triads and AB connection is

$$H[A,P] = \int \mathrm{d}^3x(-\Lambda^i G_i + NC_{\mathrm{Ham}} + N^a C_a^{\mathrm{Diff}}) \tag{2.108}$$

where C_{Ham} and C_a^{Diff} are the Hamiltonian (scalar) and diffeomorphism (vector) constraints. The AB connection here is understood to contain the triad-compatible spin connection Γ.

2.2 Modified symmetries: motivations

Canonical effective methods offer a systematic way to discuss potential corrections to the symmetries of a theory. As we have previously alluded to, classical gravity is ultimately a gauge theory, albeit one with a complicated algebraic structure. Therefore, symmetries are a key element for constructing models of quantum gravity. At the effective level, a given quantum gravity theory will induce corrections that generate new terms in a semi-classical Lagrangian. These terms must not break the gauge symmetry for the theory to be consistent. When there are no experiments to help us construct a quantum theory of gravity, these consistency conditions make up one of the few guiding principles available to us.

In this chapter, we apply canonical effective methods to analyze quantum corrections arising from quantization that uses different canonical variables. We will exemplify how symmetries of gravity guide us in clarifying ambiguities in effective theories. However, unlike the usual analysis of effective field theories which takes place in Lagrangian formalism, we will adopt the canonical (Hamiltonian) perspective. Special emphasis will be laid on the closure of Poisson brackets between first-class constraints—more commonly known as the Dirac hypersurface deformation algebras (HDA) in the context of gravity. We will chose to consider effective theories coming from loop quantum gravity as this theory follows more directly from canonical quantization. However, our analysis can be easily generalized to any theory of quantum gravity so long as an effective Hamiltonian (or Lagrangian) is obtainable.

2.3 Canonical quantum gravity corrections

Several independent studies have shown that holonomy and inverse-triad corrections from loop quantum gravity (LQG) modify hypersurface-deformation brackets for spherically symmetric gravity and related midisuperspace models [9–18], thereby realizing a deformation of general covariance [19–22]. These modifications are closely related [23] to anomaly-free models of perturbative cosmological inhomogeneity constructed within the same framework [24–28], suggesting that modified space-time structures may be a generic consequence of quantum-geometry effects in loop quantum gravity. In [29] (see also [30]), however, it has been shown that such modifications may be avoided if one uses self-dual connections and a densitized lapse function, as in [31–33], instead of real variables [34]. These models, valid for self-dual Lorentzian gravity with Barbero–Immirzi parameter $\gamma = \pm i$ or Euclidean gravity with Barbero–Immirzi parameter $\gamma = \pm 1$, are rather special because the Hamiltonian constraint simplifies considerably compared with general γ. It is therefore of interest to compare the structures encountered in various models in order to determine whether undeformed space-time structures could be realized more broadly.

Such a comparison is not obvious, for instance, because the modifications considered in [29] are different from those found in anomaly-free models using real variables. In particular, those modifications cannot be implemented in an anomaly-free manner for arbitrary choices of the Barbero–Immirzi parameter: We will show that the classical form of the constraint brackets can be retained only with a specific class of holonomy modifications for $\gamma = \pm i$ (self-dual Lorentzian gravity) or $\gamma = \pm 1$ (a special version of Euclidean gravity). More general treatments of the self-dual or Euclidean case, implemented in close analogy with the real connection formulation, lead to either anomalies or deformations of the space-time structure. This result then allows us to draw conclusions about properties of the Hamiltonian constraint required for certain types of modifications to be consistent.

At a technical level, an analysis of the Hamiltonian constraint and its Poisson brackets indicates a formal relationship between modifications of space-time structures and the appearance of spatial derivatives of the densitized triads (canonically conjugate to the connection). Spatial derivatives of the triad generically appear in the Hamiltonian constraints of gravitational theories because they are required for curvature components. But for $\gamma^2 = \pm 1$, and *only in this case*, they are completely absorbed in the connection components through the spin connection which, in combination with extrinsic-curvature components, forms the Ashtekar connection in the self-dual case [31], or the Ashtekar–Barbero connection in the real case [34].

This structural statement allows us to draw a first conclusion about the genericness of modified space-time structures. Using standard arguments from effective field theory (generalized here to a canonical setting), modified brackets should be considered generic, unless one can show that the full quantum theory has a symmetry that protects the derivative structure of terms in the Hamiltonian constraint as encountered for self-dual variables, or more generally for $\gamma^2 = \pm 1$. No such symmetry is known. Although it has been shown that the real Ashtekar–Barbero connection, unlike the self-dual one, cannot be identified with the pull-back of a space-time connection, this result is of an "aesthetic nature" [35] and does not characterize the case of $\gamma^2 = \pm 1$ via a physical symmetry that could restrict possible quantum corrections. Moreover, applying this result in the present context would amount to pre-supposing the classical space-time structure in a model of quantum gravity. In canonical quantum gravity, the structure of space-time is determined intrinsically, based on the observation that space-time symmetries of a gravitational theory are gauge transformations, generated in Hamiltonian form by the constraints that are to be quantized in order to define canonical quantum gravity. Poisson brackets of these constraints, or commutators of their operator versions, then encode the structure of space-time. An analysis of possible consistent modifications of these brackets, such that they remain closed but possibly with non-classical structure functions, show whether the symmetries remain unviolated after quantization. As we will see, such modifications with intact (but possibly deformed) symmetry exist for any value of γ. Therefore, no value of γ is distinguished by the presence of a symmetry.

In this work, we will mainly focus on an interpretation of the constraints as representing Euclidean gravity. We will then be exempt from having to consider a possible role of reality conditions, the implementation of which remains poorly understood in a quantum theory of self-dual variables. However, as the constraints are formally identical in Euclidean gravity with $\gamma = \pm 1$ and self-dual Lorentzian gravity, our results can formally be used also in the latter case.

2.4 Unsolved Gauss constraint

The model considered in [29], following [32], consists of three canonical pairs of fields — $A_i(x)$ and $E^i(x)$ for $i = 1, 2, 3$ depending on the radial coordinate x of a spherically symmetric manifold — subject to three constraints. Two of the constraints function as generators of hypersurface deformations in space-time and therefore encode the structure of space-time. The third one, a Gauss constraint, implements an internal symmetry of SO(2)-rotations of two of the canonical pairs.

While the form of the Gauss constraint and the spatial generator of hypersurface deformations (the diffeomorphism constraint) are strictly determined by the canonical structure together with the corresponding Lie algebras of infinitesimal rotations and 1-dimensional diffeomorphisms, respectively, there is much freedom in specifying the normal generator of hypersurface deformations, or the Hamiltonian constraint, even if the physical dynamics is fixed. The version used in [29, 32] is rather special in that it is quadratic in the canonical fields and does not contain spatial derivatives of E^i (while first-order spatial derivatives of A_i do appear). In the first part of this section, we will strengthen the result of [29] by showing that the consistent deformation found in this paper is unique within a family of models that preserve the quadratic nature and derivative structure of the Hamiltonian constraint. In the second part of this section, however, we will show that this rigidity is not stable within a larger class of models that determine the same classical dynamics but do not respect the restricted derivative structure (parameterized by the so-called Barbero–Immirzi parameter γ [34, 36]). The following sections will then place our discussion in a setting of effective field theory, and highlight the role played by the Gauss constraint.

2.4.1 Regaining the quadratic Hamiltonian constraint

In order to derive our rigidity result, we start from the condition that the Poisson brackets of constraints are closed and see what kind of restrictions it imposes on the form of constraints. The specific procedure follows the classical (and classic) result [37] that the full Hamiltonian constraint, up to second order in derivatives, can be regained uniquely from the classical hypersurface-deformation brackets, as specified in [38]. This procedure has already been applied to spherically symmetric models in [19], but only for modifications of the dependence of the Hamiltonian constraint on the triad variables E^i. Our calculations here differ from [19] in that we use connection variables A_i, and take into account potential modifications of the dependence on these variables.

As already indicated, we assume for now that the Hamiltonian constraint is quadratic in the canonical fields without spatial derivatives of the triad E^i. This version of the constraint is realized in spherically symmetric gravity if one uses self-dual connection variables [31] in Lorentzian signature, or real Barbero-type variables [34] in Euclidean signature such that the Barbero–Immirzi parameter is equal to $\gamma = \pm 1$. (One should also smear the Hamiltonian constraint with a lapse function of density weight minus one to guarantee the quadratic nature. We will come back to the subtleties arising from densities later.) This parameter is therefore fixed and does not appear in the remainder of this subsection. Working with

$$\{A_1(x), E^1(y)\} = 2G\delta(x,y) \tag{2.109}$$

and

$$\{A_2(x), E^2(y)\} = G\delta(x,y) \quad , \quad \{A_3(x), E^3(y)\} = G\delta(x,y) \tag{2.110}$$

while all other brackets of basic variables vanish. (Note the missing factor of 2 in the last two brackets, compared with (2.109), which is a consequence of the fact that (A_2, E^2) and (A_3, E^3) encode the same degree of freedom after the Gauss constraint is implemented.)

$$\{A_1(x), E^1(y)\} = 2\{A_{2/3}(x), E^{2/3}(y)\}\,. \tag{2.111}$$

This canonical structure completely determines the Gauss constraint

$$G[\Lambda] = \frac{1}{2G}\int \mathrm{d}x\Lambda\left((E^1)' - 2E^2A_3 + 2E^3A_2\right) \tag{2.112}$$

and the diffeomorphism constraint

$$D[M] = \frac{1}{2G}\int \mathrm{d}xM\left(2A_3'E^3 + 2A_2'E^2 - A_1(E^1)'\right) \tag{2.113}$$

but not the Hamiltonian constraint. Sometimes, it is convenient to combine the diffeomorphism constraint $D[M]$ and the Gauss constraint $G[\Lambda]$ to form the vector constraint

$$\begin{aligned} V[M] &= D[M] + G[A_1M] \\ &= \frac{1}{G}\int \mathrm{d}xM\left((A_3' + A_1A_2)E^3 + (A_2' - A_1A_3)E^2\right)\,. \end{aligned} \tag{2.114}$$

2.4.1.1 Constraints algebra and densities

We will now use these constraints and attempt to derive the most general form of the Hamiltonian constraint, purely quadratic in the canonical fields and with up to first derivatives of A_i but no derivatives of E^i, such that all constraints have closed Poisson brackets. With this assumption, we can write the local (unsmeared) constraint as

$$\mathcal{H} = H^{110}E^1E^2 + H^{101}E^1E^3 + H^{011}E^2E^3$$

$$+H^{200}(E^1)^2 + H^{020}(E^2)^2 + H^{002}(E^3)^2\,, \tag{2.115}$$

where we use the convention that $H[\underset{\sim}{N}] = (2G)^{-1}\int \mathrm{d}x\underset{\sim}{N}(x)\mathcal{H}$, H^{ijk} may be functions of A_1, A_2, A_3 and their spatial derivatives up to first order.

The appearance of densitized smearing function stems from the way we organize our integrand inside the full Hamiltonian constraint H. For instance, if we chose to write our Hamiltonian in terms of the metric, we would typically not use a densitized smearing function but instead write it as $H \sim \int \mathrm{d}x N\mathcal{H}_{\mathrm{metric}}$ with a smearing function N of density weight $\omega = 0$. However, if we wish to obtain a Hamiltonian polynomial in canonical variables then we should use densitized triads and connections; but the price we pay is the introduction of densitized smearing function $\underset{\sim}{N}$.

As shown in the introductory sections, rewriting the Hamiltonian with densitized triads is straightforward albeit tedious in the classical theory. Among the steps is an absorption of $1/\sqrt{|h|}$ into the smearing function. We may subsequently calculate brackets concerning H in two ways: We can be rigorous and treat the densitized smearing function as a function of phase space variables (metric or triads), or we can pretend it is a generic (-1)-weighted scalar function. The differences between these two operations vanish on-shell

$$\begin{aligned}\{H[N(x)], D[M(y)]\} =& \int \mathrm{d}x\mathrm{d}y\{\underset{\sim}{N}(A,E)\mathcal{H}, M(x)\mathcal{D}\}\\ =& \int \mathrm{d}x\mathrm{d}y\Big(\{\mathcal{H},\mathcal{D}\}\underset{\sim}{N}(A,E)M(x) + \{\underset{\sim}{N}(A,E),\mathcal{D}\}\mathcal{H}M(x)\Big)\\ \approx& \{H[\underset{\sim}{N}(x)], D[M(y)]\}_{\text{pretending } \underset{\sim}{N}=\underset{\sim}{N}(x)}\,,\end{aligned} \tag{2.116}$$

where the second term in the second line vanishes on-shell. Note that so far the conclusion applies to any constraint as opposed to just first-class ones. It is essentially the bracket $\{\mathcal{H},\mathcal{D}\}$ that informs us of the first-class nature of the constraints and thus the gauge structure.

The choice to ignore brackets containing densitized smearing functions is not so innocent anymore in a quantum theory; path integral quantization tells us that off-shell paths contribute to the dynamics of a quantum system. Therefore, the quantum version of the constraints algebra (whatever that means in quantum gravity) should be analyzed by treating the smearing function's dependence on phase space variables rigorously. However, as we have argued, the space-time structure is mainly encoded in the first-class nature of the constraints. At least in a semi-classical regime, the first-class information is contained in the bracket $\{\mathcal{H},\mathcal{D}\}\sim\mathcal{H}$ as opposed to $\{\underset{\sim}{N},\mathcal{D}\}$, which exists even if $\mathcal{H}$ and $\mathcal{D}$ are not first-class. Therefore, we will still be able to obtain valuable information on the allowed modifications to the bracket algebras by just by focusing on the $\mathcal{H}$-$\mathcal{D}$ bracket (and similarly the $\mathcal{H}$-$\mathcal{H}$ bracket).

2.4.1.2 Diffeomorphism constraint

We first consider the bracket of the Hamiltonian and diffeomorphism constraints, writing it in local form as

$$\begin{aligned}\{\mathcal{H}(x),\mathcal{D}(y)\} &= G\int \mathrm{d}z\left(2\frac{\delta\mathcal{H}(x)}{\delta A_1(z)}\frac{\delta\mathcal{D}(y)}{\delta E^1(z)}-2\frac{\delta\mathcal{H}(x)}{\delta E^1(z)}\frac{\delta\mathcal{D}(y)}{\delta A_1(z)}\right.\\ &\quad+\frac{\delta\mathcal{H}(x)}{\delta A_2(z)}\frac{\delta\mathcal{D}(y)}{\delta E^2(z)}-\frac{\delta\mathcal{H}(x)}{\delta E^2(z)}\frac{\delta\mathcal{D}(y)}{\delta A_2(z)}\\ &\quad\left.+\frac{\delta\mathcal{H}(x)}{\delta A_3(z)}\frac{\delta\mathcal{D}(y)}{\delta E^3(z)}-\frac{\delta\mathcal{H}(x)}{\delta E^3(z)}\frac{\delta\mathcal{D}(y)}{\delta A_3(z)}\right)\end{aligned} \tag{2.117}$$

where $D[M]=(2G)^{-1}\int \mathrm{d}xM(x)\mathcal{D}(x)$. If this bracket is to correspond to classical hypersurface deformations, it should be equal to

$$\{\mathcal{H}(x),\mathcal{D}(y)\}=2G\left(\mathcal{H}'(x)\delta(x,y)+2\mathcal{H}(x)\delta'(x,y)\right)\ , \tag{2.118}$$

using the convention that a prime on a delta function always indicates a derivative with respect to the first argument. Therefore,

$$\delta'(x,y)=-\delta'(y,x)\,. \tag{2.119}$$

If the bracket is of the given form, the smeared constraints have the bracket

$$\begin{aligned}\{H[\underset{\sim}{N}],D[M]\} &= \frac{1}{4G^2}\int \mathrm{d}x\mathrm{d}y\underset{\sim}{N}(x)M(y)\{\mathcal{H}(x),\mathcal{D}(y)\}\\ &= \frac{1}{2G}\int \mathrm{d}x\mathrm{d}y\underset{\sim}{N}(x)M(y)\left((\partial_x\mathcal{H}(x))\delta(x,y)-2\mathcal{H}(x)\partial_y\delta(x,y)\right)\\ &= -H[(\underset{\sim}{N}M)']+2H[\underset{\sim}{N}M']=-H[M\underset{\sim}{N}'-M'\underset{\sim}{N}]\end{aligned} \tag{2.120}$$

as required if $\underset{\sim}{N}$ has density weight minus one for the purpose of having a quadratic Hamiltonian constraint.

We proceed by evaluating the Poisson bracket. Considering the assumed dependence

(2.115) of $\mathcal{H}$ on the canonical variables, we have

$$
\begin{aligned}
\{\mathcal{H}(x), \mathcal{D}(y)\} &= 2G \int \mathrm{d}z \left(\left(\frac{\partial \mathcal{H}(x)}{\partial A_1(z)} \delta(x,z) + \frac{\partial \mathcal{H}(x)}{A_1'(z)} \delta'(x,z) \right) (-A_1(y)\delta'(y,z)) \right. \\
&\quad - \frac{\partial \mathcal{H}(x)}{\partial E^1(z)} \delta(x,z) \left(-(E^1)'(y)\delta(y,z) \right) \\
&\quad + \left(\frac{\partial \mathcal{H}(x)}{\partial A_2(z)} \delta(x,z) + \frac{\partial \mathcal{H}(x)}{A_2'(z)} \delta'(x,z) \right) A_2'(y)\delta(y,z) \\
&\quad - \frac{\partial \mathcal{H}(x)}{\partial E^2(z)} \delta(x,z) E^2(y) \delta'(y,z) \\
&\quad + \left(\frac{\partial \mathcal{H}(x)}{\partial A_3(z)} \delta(x,z) + \frac{\partial \mathcal{H}(x)}{A_3'(z)} \delta'(x,z) \right) A_3'(y)\delta(y,z) \\
&\quad \left. - \frac{\partial \mathcal{H}(x)}{\partial E^3(z)} \delta(x,z) E^3(y) \delta'(y,z) \right) \\
&= 2G \left(\frac{\partial \mathcal{H}(x)}{\partial A_2(x)} A_2'(x) + \frac{\partial \mathcal{H}(x)}{\partial A_3(x)} A_3'(x) + \frac{\partial \mathcal{H}(x)}{\partial E^1(x)} (E^1)'(x) \right) \delta(x,y) \\
&\quad - \left(\frac{\partial \mathcal{H}(x)}{\partial A_1(x)} A_1(y) + \frac{\partial \mathcal{H}(x)}{\partial E^2(x)} E^2(y) + \frac{\partial \mathcal{H}(x)}{\partial E^3(x)} E^3(y) \right. \\
&\quad \left. + \frac{\partial \mathcal{H}(x)}{\partial A_2'(x)} A_2'(y) + \frac{\partial \mathcal{H}(x)}{\partial A_3'(x)} A_3'(y) \right) \delta'(y,x) \\
&\quad - \int \mathrm{d}z \frac{\partial \mathcal{H}(x)}{\partial A_1'(z)} A_1(y) \delta'(x,z) \delta'(y,z) \,, \qquad (2.121)
\end{aligned}
$$

where we used (2.119).

The last term has a product of two derivatives of delta functions, which does not occur in (2.118). Integrating by parts can remove one of the derivatives, but it also gives a second-order derivative of a delta function which does not appear either in (2.118). The term, therefore, must be zero, so that we already know that $\mathcal{H}$ cannot depend on A_1'. In order to bring the remaining terms to a form close to (2.118), we use the identity

$$
\begin{aligned}
A(x)B(y)\delta'(y,x) &= A(x)\partial_y \left(B(y)\delta(y,x) \right) - A(x)B'(y)\delta(x,y) \\
&= A(x)\partial_y \left(B(x)\delta(y,x) \right) - A(x)B'(x)\delta(x,y) \\
&= A(x)B(x)\delta'(y,x) - A(x)B'(x)\delta(x,y) \qquad (2.122)
\end{aligned}
$$

and write

$$
\{\mathcal{H}(x), \mathcal{D}(y)\} = 2G \left(\frac{\partial \mathcal{H}(x)}{\partial A_1(x)} A_1'(x) + \frac{\partial \mathcal{H}(x)}{\partial A_2(x)} A_2'(x) + \frac{\partial \mathcal{H}(x)}{\partial A_3(x)} A_3'(x) \right.
$$

$$
\begin{aligned}
&+\frac{\partial\mathcal{H}(x)}{\partial A_2'(x)}A_2''(x)+\frac{\partial\mathcal{H}(x)}{\partial A_3'(x)}A_3''(x)\\
&+\frac{\partial\mathcal{H}(x)}{\partial E^1(x)}(E^1)'(x)+\frac{\partial\mathcal{H}(x)}{\partial E^2(x)}(E^2)'(x)+\frac{\partial\mathcal{H}(x)}{\partial E^3(x)}(E^3)'(x)\Bigg)\,\delta(x,y)\\
&+2G\left(\frac{\partial\mathcal{H}(x)}{\partial A_1(x)}A_1(x)+\frac{\partial\mathcal{H}(x)}{\partial A_2'(x)}A_2'(x)+\frac{\partial\mathcal{H}(x)}{\partial A_3'(x)}A_3'(x)\right.\\
&\left.+\frac{\partial\mathcal{H}(x)}{\partial E^2(x)}E^2(x)+\frac{\partial\mathcal{H}(x)}{\partial E^3(x)}E^3(x)\right)\delta'(x,y)\,. \qquad (2.123)
\end{aligned}
$$

Since $\mathcal{H}$ does not depend on A_1', the first parenthesis (multiplied by a delta function) is equal to $\mathcal{H}'$ without any further restriction on the dependence on other canonical variables. In order to evaluate the second parenthesis, which according to (2.118) should equal $4G\mathcal{H}$, we use the quadratic form (2.115) and obtain the condition

$$
\begin{aligned}
&\frac{\partial\mathcal{H}(x)}{\partial A_1(x)}A_1(x)+\frac{\partial\mathcal{H}(x)}{\partial A_2'(x)}A_2'(x)+\frac{\partial\mathcal{H}(x)}{\partial A_3'(x)}A_3'(x)\\
&+H^{110}E^1E^2+H^{101}E^1E^3+2H^{011}E^2E^3+2H^{020}(E^2)^2+2H^{002}(E^3)^2\\
=\;&2\left(H^{110}E^1E^2+H^{101}E^1E^3+H^{011}E^2E^3+H^{020}(E^2)^2+H^{002}(E^3)^2\right) \qquad (2.124)
\end{aligned}
$$

or

$$
\frac{\partial\mathcal{H}(x)}{\partial A_1(x)}A_1(x)+\frac{\partial\mathcal{H}(x)}{\partial A_2'(x)}A_2'(x)+\frac{\partial\mathcal{H}(x)}{\partial A_3'(x)}A_3'(x)=H^{110}E^1E^2+H^{101}E^1E^3+2H^{200}(E^1)^2
$$

after some cancellations. Comparing coefficients of E^iE^j in this equation, we obtain

$$
\begin{aligned}
\frac{\partial H^{110}}{\partial A_1}A_1+\frac{\partial H^{110}}{\partial A_2'}A_2'+\frac{\partial H^{110}}{\partial A_3'}A_3' &= H^{110} &(2.125)\\
\frac{\partial H^{101}}{\partial A_1}A_1+\frac{\partial H^{101}}{\partial A_2'}A_2'+\frac{\partial H^{101}}{\partial A_3'}A_3' &= H^{101} &(2.126)\\
\frac{\partial H^{011}}{\partial A_1}A_1+\frac{\partial H^{011}}{\partial A_2'}A_2'+\frac{\partial H^{011}}{\partial A_3'}A_3' &= 0 &(2.127)\\
\frac{\partial H^{200}}{\partial A_1}A_1+\frac{\partial H^{200}}{\partial A_2'}A_2'+\frac{\partial H^{200}}{\partial A_3'}A_3' &= 2H^{200} &(2.128)\\
\frac{\partial H^{020}}{\partial A_1}A_1+\frac{\partial H^{020}}{\partial A_2'}A_2'+\frac{\partial H^{020}}{\partial A_3'}A_3' &= 0 &(2.129)\\
\frac{\partial H^{002}}{\partial A_1}A_1+\frac{\partial H^{002}}{\partial A_2'}A_2'+\frac{\partial H^{002}}{\partial A_3'}A_3' &= 0\,. &(2.130)
\end{aligned}
$$

If we assume polynomial dependence of $\mathcal{H}$ on the connection variables, we can conclude that the coefficients H^{110} and H^{101} must be linear in A_1, A_2' and A_3', while H^{200} must be quadratic in these variables. The coefficients H^{011}, H^{020} and H^{002} cannot depend on A_1, A_2' or A_3'.

2.4.1.3 Bracket of Hamiltonian constraints

The Poisson bracket of two Hamiltonian constraints can be computed in a similar way. Classically, we expect

$$\begin{aligned}\{\mathcal{H}(x),\mathcal{H}(y)\} &= 2G\Big(E^1(x)^2\mathcal{V}(x)\delta'(y,x)\\ &\qquad -E^1(y)^2\mathcal{V}(y)\delta'(x,y)\Big) \end{aligned} \tag{2.131}$$

with the local vector constraint $\mathcal{V}(x)$ such that $V[M]=(2G)^{-1}\int \mathrm{d}x M(x)\mathcal{V}(x)$. If the space-time structure is deformed, the bracket is multiplied by a non-constant function β which, for a comparison with [29], we assume to depend only on the A_i. (This function should approach $\beta=1$ in some classical limit, usually for small A_i.) After using (2.115) and comparing coefficients of E^iE^j, we obtain the equations

$$\begin{aligned} &2\left(-2\frac{\partial H^{110}}{\partial A_1'}H^{200}-\frac{\partial H^{200}}{\partial A_1'}H^{110}\right)-\frac{\partial H^{110}}{\partial A_2'}H^{110}-2\frac{\partial H^{200}}{\partial A_2'}H^{020}-\frac{\partial H^{110}}{\partial A_3'}H^{101}-\frac{\partial H^{200}}{\partial A_3'}H^{011}\\ &=4\beta(A_2'-A_1A_3) \qquad (2.132)\\ &2\left(-2\frac{\partial H^{101}}{\partial A_1'}H^{200}-\frac{\partial H^{200}}{\partial A_1'}H^{101}\right)-\frac{\partial H^{101}}{\partial A_2'}H^{110}-2\frac{\partial H^{200}}{\partial A_2'}H^{011}-\frac{\partial H^{101}}{\partial A_3'}H^{101}-\frac{\partial H^{200}}{\partial A_3'}H^{002}\\ &=4\beta(A_3'+A_1A_2)\,, \qquad (2.133)\end{aligned}$$

which are sensitive to the modification the function β, as well as several β-independent equations:

$$\begin{aligned} &4\frac{\partial H^{200}}{\partial A_1'}H^{200}+\frac{\partial H^{200}}{\partial A_2'}H^{110}+\frac{\partial H^{200}}{\partial A_3'}H^{101}=0 \qquad (2.134)\\ &2\left(\frac{\partial H^{110}}{\partial A_1'}H^{110}+2\frac{\partial H^{020}}{\partial A_1'}H^{200}\right)+2\frac{\partial H^{110}}{\partial A_2'}H^{020}+\frac{\partial H^{020}}{\partial A_2'}H^{110}+\frac{\partial H^{110}}{\partial A_3'}H^{011}+\frac{\partial H^{020}}{\partial A_3'}H^{101}\\ &=0 \qquad (2.135)\\ &2\left(\frac{\partial H^{101}}{\partial A_1'}H^{101}+2\frac{\partial H^{002}}{\partial A_1'}H^{200}\right)+\frac{\partial H^{101}}{\partial A_2'}H^{011}+\frac{\partial H^{002}}{\partial A_2'}H^{110}+2\frac{\partial H^{101}}{\partial A_3'}H^{002}+\frac{\partial H^{002}}{\partial A_3'}H^{101}\\ &=0 \qquad (2.136)\\ &2\left(2\frac{\partial H^{011}}{\partial A_1'}H^{200}+\frac{\partial H^{101}}{\partial A_1'}H^{110}+\frac{\partial H^{110}}{\partial A_1'}H^{101}\right)\end{aligned}$$

$$+\frac{\partial H^{011}}{\partial A'_2}H^{110}+2\frac{\partial H^{101}}{\partial A'_2}H^{020}+\frac{\partial H^{110}}{\partial A'_2}H^{011}+\frac{\partial H^{011}}{\partial A'_3}H^{101}+\frac{\partial H^{101}}{\partial A'_3}H^{011}+2\frac{\partial H^{110}}{\partial A'_3}H^{002}$$
$$=0\,. \tag{2.137}$$

Four additional equations,

$$2\frac{\partial H^{020}}{\partial A'_1}H^{110}+2\frac{\partial H^{020}}{\partial A'_2}H^{020}+\frac{\partial H^{020}}{\partial A'_3}H^{011}$$
$$=0 \tag{2.138}$$
$$2\frac{\partial H^{002}}{\partial A'_1}H^{101}+\frac{\partial H^{002}}{\partial A'_2}H^{011}+2\frac{\partial H^{002}}{\partial A'_3}H^{002}$$
$$=0 \tag{2.139}$$
$$2\left(\frac{\partial H^{011}}{\partial A'_1}H^{110}+\frac{\partial H^{020}}{\partial A'_1}H^{101}\right)+2\frac{\partial H^{011}}{\partial A'_2}H^{020}+\frac{\partial H^{020}}{\partial A'_2}H^{011}+\frac{\partial H^{011}}{\partial A'_3}H^{011}+2\frac{\partial H^{020}}{\partial A'_3}H^{002}$$
$$=0 \tag{2.140}$$
$$2\left(\frac{\partial H^{011}}{\partial A'_1}H^{101}+\frac{\partial H^{002}}{\partial A'_1}H^{110}\right)+\frac{\partial H^{011}}{\partial A'_2}H^{011}+2\frac{\partial H^{002}}{\partial A'_2}H^{020}+2\frac{\partial H^{011}}{\partial A'_3}H^{002}+\frac{\partial H^{002}}{\partial A'_3}H^{011}$$
$$=0 \tag{2.141}$$

are identically satisfied, given that H^{011}, H^{020} and H^{002} cannot depend on A'_i. Because $\mathcal{H}$ cannot depend on A'_1, we may simplify the set of equations to

$$-\frac{\partial H^{110}}{\partial A'_2}H^{110}-2\frac{\partial H^{200}}{\partial A'_2}H^{020}-\frac{\partial H^{110}}{\partial A'_3}H^{101}-\frac{\partial H^{200}}{\partial A'_3}H^{011}=4\beta(A'_2-A_1A_3) \tag{2.142}$$
$$-\frac{\partial H^{101}}{\partial A'_2}H^{110}-2\frac{\partial H^{200}}{\partial A'_2}H^{011}-\frac{\partial H^{101}}{\partial A'_3}H^{101}-\frac{\partial H^{200}}{\partial A'_3}H^{002}=4\beta(A'_3+A_1A_2) \tag{2.143}$$
$$\frac{\partial H^{200}}{\partial A'_2}H^{110}+\frac{\partial H^{200}}{\partial A'_3}H^{101}=0 \tag{2.144}$$
$$2\frac{\partial H^{110}}{\partial A'_2}H^{020}+\frac{\partial H^{110}}{\partial A'_3}H^{011}=0 \tag{2.145}$$
$$\frac{\partial H^{101}}{\partial A'_2}H^{011}+2\frac{\partial H^{101}}{\partial A'_3}H^{002}=0 \tag{2.146}$$
$$2\frac{\partial H^{101}}{\partial A'_2}H^{020}+\frac{\partial H^{110}}{\partial A'_2}H^{011}+\frac{\partial H^{101}}{\partial A'_3}H^{011}+2\frac{\partial H^{110}}{\partial A'_3}H^{002}=0 \tag{2.147}$$

2.4.1.4 Gauss constraint

The Gauss constraint further restricts the combinations of basic variables which can appear in the Hamiltonian constraint. The gauge-invariant combinations that contribute to the classical constraint are E^1, $(E^2)^2+(E^3)^2$, $A_2E^2+A_3E^3$, $A_2^2+A_3^2$ and

$A_1(A_2E^2 + A_3E^3) - (A_2'E^3 - A_2'E^2)$. (The identity (2.122) is useful for seeing that the last combination has a vanishing Poisson bracket with the unsmeared Gauss constraint.) These expressions show that A_1, A_2' and A_3' can appear in gauge-invariant form only in combination with E^2 and E^3. It is therefore impossible to fulfill the condition that H^{200} be quadratic in A_1, A_2' and A_3' because H^{200} is defined as the E-independent coefficient of $(E^1)^2$ in the Hamiltonian constraint. For Hamiltonian constraints quadratic in E^i, we have $H^{200} = 0$.

Equations (2.142) and (2.143) then simplify to

$$-\frac{\partial H^{110}}{\partial A_2'}H^{110} - \frac{\partial H^{110}}{\partial A_3'}H^{101} = 4\beta(A_2' - A_1A_3) \tag{2.148}$$

$$-\frac{\partial H^{101}}{\partial A_2'}H^{110} - \frac{\partial H^{101}}{\partial A_3'}H^{101} = 4\beta(A_3' + A_1A_2)\,. \tag{2.149}$$

For $\beta = 1$, these equations are obeyed by the classical $H^{110}_{\rm cl} = 2(A_1A_2 + A_3')$ and $H^{101}_{\rm cl} = 2(A_1A_3 - A_2')$, as they should. For $\beta \neq 1$, we can solve these two equations by $H^{110} = \beta_1 H^{110}_{\rm cl}$ and $H^{101} = \beta_2 H^{101}_{\rm cl}$, provided that β_1 and β_2 do not depend on spatial derivatives of A_i and are such that $\beta_1\beta_2 = \beta$. Invariance under transformations generated by the Gauss constraint, which mix the terms of $H^{110}_{\rm cl}$ and $H^{101}_{\rm cl}$, implies that $\beta_1 = \beta_2$, and therefore $\beta > 0$ and $\beta_1 = \beta_2 = \sqrt{\beta}$. This modification function can be eliminated from the contributions of H^{110} and H^{101} to the constraint by absorbing it in the lapse function, thus moving the modification to the remaining contributions from $H^{020} = \beta^{-1/2}H^{020}_{\rm cl}$ and $H^{002} = \beta^{-1/2}H^{002}_{\rm cl}$. Therefore, the only non-trivial modification of the dynamics is in the contributions from H^{020} and H^{002} which, as already shown, can only depend on A_2 and A_3. Again invoking transformations generated by the Gauss constraint, the modified term $\beta^{-1/2}(H^{020}_{\rm cl} + H^{002}_{\rm cl})$ is an arbitrary (positive) function of $A_2^2 + A_3^2$, which is equivalent to the modification found in [29] and therefore strengthens their result.

If we relax the condition that the Hamiltonian constraint does not depend on spatial derivatives of the densitized triad, additional gauge-invariant combinations are possible. For instance, the extrinsic-curvature component

$$K_1 = A_1 - \frac{(E^2)'E^3 - E^2(E^3)'}{(E^2)^2 + (E^3)^2} \tag{2.150}$$

is gauge invariant. Moreover, if spatial derivatives of the densitized triad are allowed, the Gauss constraint can be used to rewrite the Hamiltonian constraint without changing the on-shell behavior. For instance, the identity

$$\begin{aligned} & A_1(A_2E^2+A_3E^3)+2E^2A_3'-2E^3A_2' \\ = \quad & (E^1)''+A_2(A_1E^2+2(E^3)')+A_3(A_1E^3-2(E^2)')-\mathcal{G}' \end{aligned} \tag{2.151}$$

eliminates spatial derivatives of A_2 and A_3 from the Hamiltonian constraint, in favor of a second-order spatial derivative of E^1. This new form is much closer to the expression of the Hamiltonian constraint in extrinsic-curvature variables [39], and may allow different modified brackets than the quadratic version (2.115) even if one works with the reduced Ashtekar connections A_i.

The possibility of rewriting the Hamiltonian constraint by using the Gauss constraint explains why different formulations of the same classical theory may give rise to different modified brackets: The Gauss constraint depends on A_2 and A_3, and therefore, depending on how it is used in writing the Hamiltonian constraint, restricts possible modifications. In extrinsic-curvature variables, this ambiguity does not appear because the Gauss constraint is solved explicitly.

From the perspective of effective field theory, applied here to the classical structure of up to second-order derivatives, restricting the dependence of the Hamiltonian constraint on spatial derivatives of E^i leads to non-generic models. The classical constraint is quadratic in A_i, which, according to the field equations implied by the theory, amounts to terms with up to two derivatives. Any term that is consistent with the symmetries of the theory (generated by the constraints) and has up to two derivatives (temporal or spatial) should then be allowed for a generic model. Such theories should include terms with up to second-order spatial derivatives of E^i, in addition to the quadratic terms in A_i which contribute two time derivatives. (A higher-derivative theory beyond second order would be obtained by including quantum back-reaction effects, which is not the purpose of this paper.)

2.4.2 Arbitrary Barbero–Immirzi parameter

We will now show that the preceding rigidity result is not stable within a class of models in which spatial derivatives of the densitized triad are allowed to appear. A suitable set of constraints that describes the same classical physics as, depending on the signature, Euclidean or self-dual gravity is obtained by letting the Barbero–Immirzi parameter vary, instead of fixing it to a specific value such that $\gamma^2 = \pm 1$. The modification found in [29] is therefore not generic. To this end, we will now switch to a general setting of spherically symmetric gravity in which the Barbero–Immirzi parameter and other numerical factors (as well as the gravitational constant G) are included.

Spherically symmetric gravity can be formulated as a Hamiltonian theory with phase space given by the canonical pairs, subject to three constraints. This setting has been formulated in [32] for self-dual variables and in [39] for real variables. In order to avoid having to impose reality conditions, we follow the latter notation, in which the canonical pairs (A_1, E^1), (A_2, E^2) and (A_3, E^3) are such that

$$\{A_1(x), E^1(y)\} = 2\gamma G\delta(x,y) \tag{2.152}$$

and

$$\begin{aligned}\{A_2(x), E^2(y)\} &= \gamma G\delta(x,y) &(2.153)\\ \{A_3(x), E^3(y)\} &= \gamma G\delta(x,y) &(2.154)\end{aligned}$$

(a version of (2.109) and (2.110) for arbitrary real γ). They are subject to the Gauss constraint

$$G[\Lambda] = \frac{1}{2\gamma G}\int \mathrm{d}x\Lambda\left((E^1)' + 2A_2E^3 - 2A_3E^2\right) \tag{2.155}$$

smeared with a multiplier Λ, the diffeomorphism constraint

$$D[N^x] = \frac{1}{2\gamma G}\int \mathrm{d}x\, N^x\, \left(-A_1(E^1)' + 2A_3'E^3 + 2A_2'E^2\right) \tag{2.156}$$

smeared with the shift vector N^x, and the Hamiltonian constraint

$$\begin{aligned} H[\underset{\sim}{N}] &= \frac{1}{2G}\int \mathrm{d}x\, \underset{\sim}{N}\, \Big(2A_1E^1(A_2E^2 + A_3E^3) \\ &\quad +(A_2^2 + A_3^2 - 1)\left((E^2)^2 + (E^3)^2\right) + 2E^1\left(E^2A_3' - E^3A_2'\right) \\ &\quad + (\epsilon - \gamma^2)\left(2K_1E^1(K_2E^2 + K_3E^3) + ((K_2)^2 + (K_3)^2)((E^2)^2 + (E^3)^2)\right)\Big) \\ &= H^{\mathrm{E}}[\underset{\sim}{N}] + H^{\mathrm{L}}[\underset{\sim}{N}] \end{aligned} \tag{2.157}$$

smeared with the lapse function $\underset{\sim}{N}$ of density weight -1. The non-polynomial relationship between the extrinsic-curvature components K_1, K_2 and K_3 with the basic variables is given below.

In all three constraints, the prime represents a derivative with respect to the radial coordinate x. Moreover, γ in (2.157) is the Barbero–Immirzi parameter [34, 36] and $\epsilon = \pm 1$ the space-time signature, such that $\epsilon = 1$ in the Euclidean case and $\epsilon = -1$ in the Lorentzian case. As usual, it is convenient to split the Hamiltonian constraint into the Euclidean part

$$\begin{aligned} H^{\rm E}[\underset{\sim}{N}] &= \frac{1}{2G}\int {\rm d}x\, \underset{\sim}{N}\, \Big(2A_1E^1(A_2E^2+A_3E^3) \\ &\quad + (A_2^2+A_3^2-1)\left((E^2)^2+(E^3)^2\right) + 2E^1\left(E^2A_3'-E^3A_2'\right)\Big) \end{aligned} \quad (2.158)$$

and the "Lorentzian" contribution

$$\begin{aligned} H^{\rm L}[\underset{\sim}{N}] = -\frac{\gamma^2-\epsilon}{2G}\int {\rm d}x\, \underset{\sim}{N}\, \Bigg(& 2K_1E^1(K_2E^2+K_3E^3) \\ & +((K_2)^2+(K_3)^2)((E^2)^2+(E^3)^2)\Bigg)\,. \end{aligned} \quad (2.159)$$

Thus, $H[\underset{\sim}{N}] = H^{\rm E}[\underset{\sim}{N}]$ for $\gamma = \pm 1$ in Euclidean signature ($\epsilon = 1$), while the "Lorentzian" contribution (a slight misnomer) also contributes in Euclidean signature if $\gamma \neq \pm 1$. (The Lorentzian contribution is always required in Lorentzian signature if one works with real γ such that the Poisson brackets are real.) The canonical variables A_1, E^2 and E^3 have density weight one.

The geometrical meaning of the phase-space variables is determined as follows: The fields E^1, E^2 and E^3, as the components of a spherically symmetric densitized triad, describe a spatial metric q_{ab} according to the line element

$$\begin{aligned} {\rm d}s^2 &= q_{ab}{\rm d}x^a{\rm d}x^b \quad (2.160) \\ &= \frac{(E^2)^2+(E^3)^2}{|E^1|}{\rm d}x^2 + |E^1|({\rm d}\vartheta^2+\sin^2\vartheta{\rm d}\varphi^2)\,. \end{aligned}$$

The densitized triad also determines a spin connection such that it is constant with respect to the resulting covariant derivative. The components of this spin connection are functions of the densitized triad and its first spatial derivatives:

$$\Gamma_1 = \frac{E^3(E^2)'-E^2(E^3)'}{(E^2)^2+(E^3)^2} \quad (2.161)$$

$$\Gamma_2 = -\frac{1}{2}\frac{(E^1)'E^3}{(E^2)^2+(E^3)^2} \quad (2.162)$$

$$\Gamma_3 = \frac{1}{2}\frac{(E^1)'E^2}{(E^2)^2+(E^3)^2}\,. \quad (2.163)$$

The densitized triad is canonically conjugate to components of extrinsic curvature, K_i,

$i = 1, 2, 3$. Since the Γ_i depend only on E^i, one can add them to K_i without changing the latter's canonical relationships with E^i. In this way, the canonical connection components $A_i = \Gamma_i + \gamma K_i$ are obtained, using the Barbero–Immirzi parameter γ.

The constrained system is first-class, with brackets of the constraints $D[N^x]$ and $H[\underset{\sim}{N}]$ according to Dirac's hypersurface deformations [38] (taking into account the density weight of $\underset{\sim}{N}$ in the Hamiltonian constraint used here). In particular, the bracket $\{H[\underset{\sim}{N}], H[\underset{\sim}{M}]\}$ should be proportional to the diffeomorphism constraint, up to possible contributions from the Gauss constraint. We display the relevant derivations in a more general setting, following the observation [29] that, for $\gamma^2 = \epsilon$, the constraint brackets remain closed in the presence of a "magnetic-field" modification, replacing $B_1 := A_2^2 + A_3^2 - 1$ in the Euclidean part of the Hamiltonian constraint with an arbitrary function $f(A_2^2 + A_3^2 - 1)$. Our aim is to determine whether this modification can be carried over to the Lorentzian contribution.

We begin with the bracket of two modified Euclidean parts, $\{H^{\rm E}[\underset{\sim}{N}], H^{\rm E}[\underset{\sim}{M}]\}$. Thanks to antisymmetry of the bracket in $\underset{\sim}{N}$ and $\underset{\sim}{M}$, we need consider only those brackets of terms that lead to derivatives of delta functions. There are two such contributions,

$$\begin{aligned} &\{2A_1(x)E^2(x)(A_2(x)E^2(x) + A_3(x)E^3(x)), 2E^1(y)(E^2(y)A_3(y)' - E^3(y)A_2(y)')\} \\ =\ & (\cdots)\delta(x,y) - 4\gamma G A_1(x)E^1(x)E^1(y)\left(A_3(x)E^2(y) - A_2(x)E^3(y)\right)\partial_y\delta(x,y) \end{aligned} \tag{2.164}$$

and

$$\begin{aligned} &\{2E^1(x)(E^2(x)A_3(x)' - E^3(x)A_2(x)'), 2E^1(y)(E^2(y)A_3(y)' - E^3(y)A_2(y)')\} \\ =\ & (\cdots)\delta(x,y) - 4\gamma G E^1(x)E^1(y)\left(\left(E^2(x)A_2(y)' + E^3(x)A_3(y)'\right)\partial_x\delta(x,y) \right.\\ & \left. - \left(E^2(y)A_2(x)' + E^3(y)A_3(x)'\right)\partial_y\delta(x,y)\right)\,. \end{aligned} \tag{2.165}$$

With these two ingredients, we obtain

$$\begin{aligned} \{H^{\rm E}[\underset{\sim}{N}], H^{\rm E}[\underset{\sim}{M}]\} &= \frac{\gamma}{G}\int {\rm d}x \left(\underset{\sim}{N}'\underset{\sim}{M} - \underset{\sim}{N}\underset{\sim}{M}'\right)(E^1)^2\left(A_1(A_2E^3 - A_3E^2) + E^2A_2' + E^3A_3'\right) \\ &= \gamma^2 V[(E^1)^2(\underset{\sim}{N}'\underset{\sim}{M} - \underset{\sim}{M}'\underset{\sim}{N})] \end{aligned} \tag{2.166}$$

where

$$V[\Lambda] = \frac{1}{\gamma G}\int {\rm d}x\,\Lambda\,\left(A_1(E^2A_3 - E^3A_2) + A_3'E^3 + A_2'E^2\right) \tag{2.167}$$

is the vector constraint constraint (2.114), $V[\Lambda] = D[\Lambda] + G[A_1\Lambda]$, related to the diffeomorphism constraint D through a contribution from the Gauss constraint (2.155).

Using $\sqrt{\det q} = \sqrt{|E^1|((E^2)^2 + (E^3)^2)}$ from (2.160), we can write the smearing function in (2.166) as

$$(E^1)^2 \left(\underset{\sim}{N}' \underset{\sim}{M} - \underset{\sim}{M}' \underset{\sim}{N} \right) = \frac{|E^1|}{(E^2)^2 + (E^3)^2} (N'M - M'N) \tag{2.168}$$

where $N = \sqrt{|E^1|((E^2)^2 + (E^3)^2)}\underset{\sim}{N}$ and $M = \sqrt{|E^1|((E^2)^2 + (E^3)^2)}\underset{\sim}{M}$ are lapse functions without density weight. The coefficient $|E^1|/\left((E^2)^2 + (E^3)^2\right)$ in (2.168) is, according to (2.160), the radial component of the inverse spatial metric, in agreement with the classical form of hypersurface-deformation brackets. The system is therefore anomaly-free for any modification f in (2.157) without any modification of the constraint brackets and the space-time structure — provided the Lorentzian part does not contribute to the Hamiltonian constraint, that is in Euclidean gravity with $\gamma = \pm 1$ or in Lorentzian gravity with $\gamma = \pm i$. This is consistent with the results reported in [29].

It is easy to see that any function $f(A_2^2 + A_3^2 - 1)$ can be used in the modified Euclidean part because this term does not produce derivatives of delta functions in the Poisson bracket of two Euclidean constraints. Moreover, because A_2 and A_3 are scalars without density weight, any such term has the correct Poisson bracket with the diffeomorphism constraint. However, if $\gamma^2 \neq \epsilon$, the cross-term $\{H^{\rm E}[\underset{\sim}{N}], H^{\rm L}[\underset{\sim}{M}]\}$ in the Poisson bracket of two Hamiltonian constraints does receive a contribution from $f(A_2^2 + A_3^2 - 1)$ in $H^{\rm E}[\underset{\sim}{N}]$ because $H^{\rm L}[\underset{\sim}{M}]$, written in the canonical variables A_i and E^i, contains spatial derivatives of E^i through Γ_i. An explicit calculation is therefore required to check whether the bracket can still be closed for $f(A_2^2 + A_3^2 - 1) \neq A_2^2 + A_3^2 - 1$.

We first compute The Poisson brackets of each individual term in $H^{\rm E}[\underset{\sim}{N}]$ with the full $H^{\rm L}[\underset{\sim}{M}]$: We obtain

$$\begin{aligned}
& \frac{1}{G}\{\textstyle\int {\rm d}x N(x) A_1(x) E^1(x) (A_2(x)E^2(x) + A_3(x)E^3(x)), H^{\rm L}[\underset{\sim}{M}]\} \\
= \; & \frac{\gamma^2 - \epsilon}{2\gamma^2 G^2} \int {\rm d}x{\rm d}y N(x) M(y) \left((\cdots)\delta(x,y) \right. \\
& \quad -2A_1(x)E^1(x)E^1(y)(A_2(y)E^2(y) + A_3(y)E^3(y))\{A_2(x)E^2(x) + A_3(x)E^3(x), \Gamma_1(y)\} \\
& \quad +E^1(x)(A_2(x)E^2(x) + A_3(x)E^3(x)) \left(E^2(y)^2 + E^3(y)^2\right) \\
& \qquad \times \left. \{A_1(x), -2(A_2(y)\Gamma_2(y) + A_3(y)\Gamma_3(y)) + \Gamma_2(y)^2 + \Gamma_3(y)^2\} \right) \\
= \; & \frac{\gamma^2 - \epsilon}{2\gamma G} \int {\rm d}x{\rm d}y N(x) M(y) \Big(-2A_1(x)E^1(x)E^1(y)(A_2(y)E^2(y) \\
& \quad +A_3(y)E^3(y)) \frac{E^2(x)E^3(y) - E^2(y)E^3(x)}{E^2(y)^2 + E^3(y)^2}
\end{aligned}$$

$$
\begin{aligned}
&+2E^1(x)(E^2(y)^2+E^3(y)^2)(A_2(x)E^2(x)+A_3(x)E^3(x))\\
&\qquad\times\frac{A_2(y)E^3(y)-A_3(y)E^2(y)-E^3(y)\Gamma_2(y)+E^2(y)\Gamma_3(y)}{E^2(y)^2+E^3(y)^2}\Big)\partial_y\delta(x,y)\\
=&\ -\frac{\gamma^2-\epsilon}{2\gamma G}\int \mathrm{d}x N(x)M'(x)E^1(A_2E^2+A_3E^3)\left((E^1)'+2A_2E^3-2A_3E^2\right)\\
=&\ -(\gamma^2-\epsilon)G[NM'E^1(A_2E^2+A_3E^3)] \qquad (2.169)
\end{aligned}
$$

up to terms that cancel out when inserted in the antisymmetric $\{H^{\mathrm{E}}[\underset{\sim}{N}],H^{\mathrm{L}}[\underset{\sim}{M}]\}+\{H^{\mathrm{L}}[\underset{\sim}{N}],H^{\mathrm{E}}[\underset{\sim}{M}]\}$. In the detailed calculations, we have used the explicit expressions for the Γ_i, from which we also obtain the useful identity

$$
\gamma(K_2E^2+K_3E^3)=A_2E^2+A_3E^3 \qquad (2.170)
$$

because $\Gamma_2E^2+\Gamma_3E^3$ is identically zero.

The second term,

$$
\begin{aligned}
&\frac{1}{2G}\{\textstyle\int \mathrm{d}x N(x)f(A_2(x)^2+A_3(x)^2-1)(E^2(x)^2+E^3(x)^2),H^{\mathrm{L}}[\underset{\sim}{M}]\}\\
=&\ \frac{\gamma^2-\epsilon}{2\gamma^2G^2}\int \mathrm{d}x\mathrm{d}y N(x)M(y)\Big((\cdots)\delta(x,y)\\
&-2\dot{f}(x)(E^2(x)^2+E^3(x)^2)E^1(y)(A_2(y)E^2(y)+A_3(y)E^3(y))\{A_2(x)^2+A_3(x)^2,\Gamma_1(y)\}\Big)\\
=&\ \frac{\gamma^2-\epsilon}{2\gamma G}\int \mathrm{d}x\mathrm{d}y N(x)M(y)\Big((\cdots)\delta(x,y)\\
&-2\dot{f}(x)(E^2(x)^2+E^3(x)^2)E^1(y)(A_2(y)E^2(y)+A_3(y)E^3(y))\\
&\qquad\times\frac{2A_2(x)E^3(y)-A_3(x)E^2(y)}{E^2(y)^2+E^3(y)^2}\partial_y\delta(x,y)\Big)\\
=&\ 2(\gamma^2-\epsilon)G[NM'\dot{f}E^1(A_2E^2+A_3E^3)]-\frac{\gamma^2-\epsilon}{2\gamma G}\int \mathrm{d}x NM'\dot{f}E^1(E^1)'(A_2E^2+A_3E^3)\,, \qquad (2.171)
\end{aligned}
$$

does not vanish on the constraint surface. Therefore, the function f, whose derivative by its argument we have denoted by $\dot{f}$, is now relevant for closed brackets. In particular, the last contribution containing $(E^1)'$ must be canceled by a corresponding term in the remaining bracket.

In this last bracket,

$$
\begin{aligned}
B\ :=&\ \frac{1}{G}\{\textstyle\int \mathrm{d}x N(x)E^1(x)(E^2(x)A_3(x)'-E^3(x)A_2(x)'),H^{\mathrm{L}}[\underset{\sim}{M}]\}\\
=&\ \frac{\gamma^2-\epsilon}{2\gamma^2G^2}\int \mathrm{d}x\mathrm{d}y N(x)M(y)\Big((\cdots)\delta(x,y)
\end{aligned}
$$

$$
\begin{aligned}
&+2E^1(x)E^1(y)(A_2(y)E^2(y)+A_3(y)E^3(y))\{E^2(x)A_3(x)'-E^3(x)A_2(x)',-\Gamma_1(y)\}\\
&+2E^1(x)E^1(y)(A_1(y)-\Gamma_1(y))\{E^2(x)A_3(x)'-E^3(x)A_2(x)',A_2(y)E^2(y)+A_3(y)E^3(y)\}\\
&-2E^1(x)(E^2(y)^2+E^3(y)^2)\Big((A_2(y)-\Gamma_2(y))\{E^2(x)A_3(x)'-E^3(x)A_2(x)',\Gamma_2(y)\}\\
&\qquad+(A_3(y)-\Gamma_3(y))\{E^2(x)A_3(x)'-E^3(x)A_2(x)',\Gamma_3(y)\}\Big)\Big)
\end{aligned}
$$

$$
\begin{aligned}
=\ &\frac{\gamma^2-\epsilon}{2\gamma G}\int \mathrm{d}x\mathrm{d}yN(x)M(y)\Big((\cdots)\delta(x,y)\\
&-2E^1(x)E^1(y)(A_2(y)E^2(y)+A_3(y)E^3(y))\frac{E^2(x)E^2(y)'+E^3(x)E^3(y)'}{E^2(y)^2+E^3(y)^2}\partial_x\delta(x,y)\\
&+2E^1(x)E^1(y)(A_2(y)E^2(y)+A_3(y)E^3(y))\frac{E^2(x)E^3(y)+E^3(x)E^2(y)}{E^2(y)^2+E^3(y)^2}\partial_x\partial_y\delta(x,y)\\
&+2(A_1(y)-\Gamma_1(y))E^1(x)E^1(y)(E^2(x)A_3(y)-E^3(x)A_2(y))\partial_x\delta(x,y)\\
&+E^1(x)E^1(y)\left((A_2(y)-\Gamma_2(y))E^2(x)+(A_3(y)-\Gamma_3(y))E^3(y)\right)\partial_x\delta(x,y)\Big),
\end{aligned}
\tag{2.172}
$$

we have a contribution from a second-order derivative of the delta function. Integrating by parts once in this term and taking into account its contributions to NM' and $N'M$, respectively, (noting that terms with $N'M'$ cancel out in the final antisymmetric bracket) we write

$$
\begin{aligned}
B\ =\ &\frac{\gamma^2-\epsilon}{2\gamma G}\int \mathrm{d}x\mathrm{d}yN(x)M(y)\Bigg((\cdots)\delta(x,y)\\
&-2\frac{E^1(x)E^1(y)}{E^2(y)^2+E^3(y)^2}\Bigg((E^2(x)E^2(y)'+E^3(x)E^3(y)')(A_2(y)E^2(y)+A_3(y)E^3(y))\\
&\qquad+(E^3(y)E^2(y)'-E^2(y)E^3(y)')(E^2(x)A_3(y)-E^3(x)A_2(y))\Big)\partial_x\delta(x,y)\\
&+E^1(x)E^1(y)\Big(2A_1(y)(E^2(x)A_3(y)-E^3(x)A_2(y))\\
&\qquad+E^1(x)E^1(y)'(A_2(y)E^2(x)+A_3(y)E^3(x))\Big)\partial_x\delta(x,y)\\
&-2E^1(x)E^1(y)\Big(A_2(y)E^2(y)'+A_3(y)E^3(y)'+A_2(y)'E^2(y)+A_3(y)'E^3(y)\\
&\qquad-2(A_2(y)E^2(y)+A_3(y)E^3(y))\frac{E^2(x)E^2(y)'+E^3(x)E^3(y)'}{E^2(y)^2+E^3(y)^2}\Bigg)\partial_x\delta(x,y)\Bigg)\\
=\ &\frac{\gamma^2-\epsilon}{2\gamma G}\int \mathrm{d}x\mathrm{d}yN(x)M(y)\Bigg((\cdots)\delta(x,y)+2E^1(x)E^1(y)\\
&\qquad\times\Big(A_1(y)(E^2(x)A_3(y)-E^3(x)A_2(y))-(A_2(y)'E^2(y)+A_3(y)'E^3(y))\Big)\partial_x\delta(x,y)\\
=\ &(\gamma^2-\epsilon)\left(D[(E^1)^2N'M]+G[A_1(E^1)^2N'M]\right)
\end{aligned}
$$

$$-\frac{\gamma^2-\epsilon}{2\gamma G}\int \mathrm{d}x N' M E^1 (E^1)'(A_2E^2+A_3E^3)\,. \tag{2.173}$$

This result provides the diffeomorphism constraint as well as a term which cancels the previous non-constraint contribution in (2.171), but only if $\dot{f}=1$. Therefore, if the Lorentzian contribution is included, no modification of the classical $A_2^2+A_3^2-1$ is allowed. The final bracket now equals

$$\begin{aligned}
\{H[\underset{\sim}{N}],H[\underset{\sim}{M}]\} &= \{H^{\mathrm{E}}[\underset{\sim}{N}],H^{\mathrm{E}}[\underset{\sim}{M}]\}+\{H^{\mathrm{E}}[\underset{\sim}{N}],H^{\mathrm{L}}[\underset{\sim}{M}]\}-\{H^{\mathrm{E}}[\underset{\sim}{M}],H^{\mathrm{L}}[\underset{\sim}{N}]\}\\
&= \gamma^2 D[(E^1)^2(\underset{\sim}{N}'\underset{\sim}{M}-\underset{\sim}{N}\underset{\sim}{M}')]+\gamma^2 G[A_1(E^1)^2(\underset{\sim}{N}'\underset{\sim}{M}-\underset{\sim}{N}\underset{\sim}{M}')]\\
&\quad -(\gamma^2-\epsilon)G[E^1(A_2E^2+A_3E^3)(1-2\dot{f})(\underset{\sim}{N}'\underset{\sim}{M}-\underset{\sim}{N}\underset{\sim}{M}')]\\
&\quad -(\gamma^2-\epsilon)\left(D[(E^1)^2(\underset{\sim}{N}'\underset{\sim}{M}-\underset{\sim}{N}\underset{\sim}{M}')]+G[A_1(E^1)^2(\underset{\sim}{N}'\underset{\sim}{M}-\underset{\sim}{N}\underset{\sim}{M}')]\right)\\
&= \epsilon\left(D[(E^1)^2(\underset{\sim}{N}'\underset{\sim}{M}-\underset{\sim}{N}\underset{\sim}{M}')]+G[A_1(E^1)^2(\underset{\sim}{N}'\underset{\sim}{M}-\underset{\sim}{N}\underset{\sim}{M}')]\right)\\
&\quad +(\gamma^2-\epsilon)G[E^1(A_2E^2+A_3E^3)(\underset{\sim}{N}'\underset{\sim}{M}-\underset{\sim}{N}\underset{\sim}{M}')]\\
&\approx -\epsilon D[(E^1)^2(\underset{\sim}{N}\underset{\sim}{M}'-\underset{\sim}{N}'\underset{\sim}{M})]\,,
\end{aligned} \tag{2.174}$$

using $\dot{f}=1$ in the last step because the bracket would not be closed otherwise. (Note that $\{H^{\mathrm{L}}[\underset{\sim}{N}],H^{\mathrm{L}}[\underset{\sim}{M}]\}=0$, which can most easily be seen if one uses the canonical variables K_i and E^i, of which no spatial derivatives appear in the Lorentzian contribution.)

2.5 Connection variables in a canonical effective field theory

We have seen a crucial difference between gravitational theories governed by the Euclidean Hamiltonian constraint H^{E} and the full $H^{\mathrm{E}}+H^{\mathrm{L}}$, respectively. Formally, the reason is the difference in derivative structures implied by the spin-connection terms in H^{L}: While H^{E} contains derivatives only of the spatial connection, H^{L} also contributes spatial derivatives of the triad. As a consequence, the two versions allow different modifications while maintaining closed brackets.

Derivative structures are best dealt with in a setting of effective field theory, in which one formulates generic theories by selecting the basic fields and the maximum order of derivatives to which they contribute, as well as relevant symmetries. For our purposes, we need an adaptation of the usual arguments to a canonical formulation, in which some derivatives may not be explicit because they appear only if some of the canonical equations are used, mainly in the relationship between momenta and "velocities."

In order to determine the correct derivative orders in a canonical theory, we must first choose which of the basic fields should play the role of configuration variables and therefore are considered free of time derivatives. We are looking for a canonical theory of triads, which will correspond to a space-time metric or triad theory, and therefore choose as our basic fields a densitized spatial triad with momenta. The latter may be given in terms of a connection or extrinsic curvature. The derivative order depends on the quantum effects we wish to include. For now, we will analyze the classical setting and therefore consider up to second-order derivatives of the fields. Symmetries are implemented by the requirement that the constraint brackets be closed, and in the classical case amount to hypersurface-deformation brackets.

2.5.1 Basic strategy

In our explicit calculations of generic terms, we again follow the conventions of section 2.2 and set $\gamma = 1$ for simplicity. For our effective Hamiltonian, we choose to allow up to second-order in derivatives of densitized triads. Since the conjugate momenta are of the form $A \sim \partial E$, using the equations of motion for $\dot{E}$, we have the following general form of the Hamiltonian constraint $H[\underset{\sim}{N}] = (2G)^{-1}\int \mathrm{d}x \underset{\sim}{N}(x)\mathcal{H}(x)$ with

$$\mathcal{H} = \alpha^i(E^j, \partial E^j)A_i + \beta^{ij}(E^k)A_{ij} + \gamma^i(E)\partial A_i + Q(E, \partial E, \partial^2 E)\,, \tag{2.175}$$

where we have introduced the notation $\partial \equiv \partial/\partial x$, $A_{ij\cdots k} = A_i A_j \cdots A_k$ and $E^{ij\cdots k} = E^i E^j \cdots E^k$. We can already observe some preliminary restrictions on the coefficients $\alpha^i(E, \partial E)$ and $Q(E, \partial E, \partial E\partial E, \partial^2 E)$. Both coefficients are initially allowed to depend on ∂E^i and $\partial^2 E^i$. But since we only allow up to second-order derivatives in the Hamiltonian constraint, the dependence cannot be arbitrary. Specifically, we have

$$\begin{cases} \alpha^i = \bar{\alpha}^i(E) + \alpha^i_j(E)\partial E^j \\ Q = \bar{Q}(E) + a_i(E)\partial E^i + b_{ij}(E)\partial E^i \partial E^j + c_i(E)\partial^2 E^i\,. \end{cases}$$

We want the Hamiltonian density $\mathcal{H}$ to respect the classical symmetries,

$$\begin{cases} \{\mathcal{H}(x), \mathcal{G}(y)\} & = 0 \\ \{\mathcal{H}(x), \mathcal{D}(y)\} & = 2G(\partial\mathcal{H}(x)\delta_{xy} + 2\mathcal{H}(x)\delta'_{xy}) \\ \{\mathcal{H}(x), \mathcal{H}(y)\} & \approx -2G(\partial(E^{11}\mathcal{D}(x))\delta_{xy} + 2E^{11}\mathcal{D}(x)\delta'_{xy})\,, \end{cases} \tag{2.176}$$

where $G[\Lambda] = (2G)^{-1}\int \mathrm{d}x\Lambda(x)\mathcal{G}(x)$ and $D[N] = (2G)^{-1}\int \mathrm{d}xN(x)\mathcal{D}(x)$ are the diffeomorphism and Gauss constraints, respectively. We have introduced the shorthand notation $\delta'_{xy} := \partial_x\delta(x-y)$, and $\approx$ means "equal" when setting $\mathcal{G} = 0$ in the final step of the calculation. These symmetries will impose restrictions on the coefficients $\alpha_i, \beta^{ij}, \gamma^i, Q$ in (2.175), telling us what a generic Hamiltonian constraint looks like.

2.5.2 Brackets

The first bracket, $\{\mathcal{H}, \mathcal{G}\}$, represents the restriction to gauge-invariant terms for any allowed $\mathcal{H}$. Inserting (2.175), we have

$$
\begin{aligned}
\{\mathcal{H}(x), \mathcal{G}(y)\} &= 2G\int \mathrm{d}z[(\alpha^1 + 2\beta^{1j}A_j)\delta_{xz} + \gamma^1\delta'_{xz}](x)\delta'_{yz} \\
&\quad +[(\alpha^2 + 2\beta^{2j}A_j)\delta_{xz} + \gamma^2\delta'_{xz}](x)(-A_3(y)\delta_{yz}) \\
&\quad -[(\delta_{xz}\partial_2 + \delta'_{xz}\partial_{2'})(\alpha^i)A_i + (\delta_{xz}\partial_2 + \delta'_{xz}\partial_{2'} + \delta''_{xz}\partial_{2''})Q \\
&\quad +\delta_{xz}\partial_2\beta^{ij}A_{ij} + \delta_{xz}\partial_2\gamma^i\partial A_i](x)E^3(y)\delta_{yz} \\
&\quad +[(\alpha^3 + 2\beta^{3j}A_j)\delta_{xz} + \gamma^3\delta'_{xz}](x)(A_2(y)\delta_{yz}) \\
&\quad -[(\delta_{xz}\partial_3 + \delta'_{xz}\partial_{3'})(\alpha^i)A_i + (\delta_{xz}\partial_3 + \delta'_{xz}\partial_{3'} + \delta''_{xz}\partial_{3''})Q \\
&\quad +\delta_{xz}\partial_3\beta^{ij}A_{ij} + \delta_{xz}\partial_3\gamma^i\partial A_i](x)(-E^2(y)\delta_{yz}) \\
&= 0\,,
\end{aligned}
$$

where we have introduced further shorthand notation $\partial_i := \partial/\partial E^i$ and $\partial_{i'} := \partial/\partial(\partial_x E^i)$. To make the right-hand side of the equation vanish, we need several cancellations. We can do this by first making all functions depend on x using delta functions and integrating over z. Then we group terms with the same dependence on A_i and derivatives of δ_{xy} together and demand that each grouping vanish by itself. (Different order of derivatives on δ may be dependent, for instance in $\delta'_{yx}A(x) = A(y)\delta'_{yx} + \partial_y A(y)\delta_{yx}$. Therefore, some δ' can produce terms that group with a δ.) This procedure produces several dozens of partial differential equations which we will list later along with those from the $\{\mathcal{H}, \mathcal{D}\}$ bracket.

Inserting our form of $\mathcal{H}$ into the $\mathcal{H}$-$\mathcal{D}$ bracket, we obtain

$$
\begin{aligned}
\{\mathcal{H}(x), \mathcal{D}(y)\} &= 2G\int \mathrm{d}z[\delta_{xz}(\alpha^1 + 2\beta^{1j}A_j) + \gamma^1\delta'_{xz}](x)(-A_1(y)\delta'_{yz}) \\
&\quad -[(\delta_{xz}\partial_1 + \delta'_{xz}\partial_{1'})(\alpha^i)A_i + \delta_{xz}\partial_1\beta^{ij}A_{ij} \\
&\quad +\delta_{xz}\partial_1\gamma^i\partial A_i + (\delta_{xz}\partial_1 + \delta'_{xz}\partial_{1'} + \delta''_{xz}\partial_{1''})(Q)](x)(-\partial E^1(y)\delta_{yz})
\end{aligned}
$$

$$
\begin{aligned}
&+[\delta_{xz}(\alpha^2+2\beta^{2j}A_j)+\gamma^2\delta'_{xz}](x)(\partial A_2(y)\delta_{yz})\\
&-[(\delta_{xz}\partial_2+\delta'_{xz}\partial_{2'})(\alpha^i)A_i+\delta_{xz}\partial_2\beta^{ij}A_{ij}\\
&+\delta_{xz}\partial_2\gamma^i\partial A_i+(\delta_{xz}\partial_2+\delta'_{xz}\partial_{2'}+\delta''_{xz}\partial_{2''})(Q)](x)(E^2(y)\delta'_{yz})\\
&+[\delta_{xz}(\alpha^3+2\beta^{3j}A_j)+\gamma^3\delta'_{xz}](x)(\partial A_3(y)\delta_{yz})\\
&-[(\delta_{xz}\partial_3+\delta'_{xz}\partial_{3'})(\alpha^i)A_i+\delta_{xz}\partial_3\beta^{ij}A_{ij}\\
&+\delta_{xz}\partial_3\gamma^i\partial A_i+(\delta_{xz}\partial_3+\delta'_{xz}\partial_{3'}+\delta''_{xz}\partial_{3''})(Q)](x)(E^3(y)\delta'_{yz})\\
= \quad & 2G(\partial_x\mathcal{H}(x)\delta_{xy}+2\mathcal{H}(x)\delta'_{xy})\,.
\end{aligned}
$$

Similarly to how we dealt with the condition of gauge invariance, we first integrate over z to make all functions depend on x, then match term by term with the right-hand side, expanded in A_i and derivatives of δ_{xy}. Again, we obtain a few dozen partial differential equations.

We next list the partial differential equations that the coefficients of terms in $\mathcal{H}$ have to obey. These equations will completely determine the dependence on E^2 and E^3, leaving free functions of E^1 which the $\mathcal{H}$-$\mathcal{H}$ bracket will further restrict. These conditions then determine possible modifications of the classical $\mathcal{H}_{\rm cl}$. In the following equations, we use the differential operators $\hat{D} := E^2\partial_2 + E^3\partial_3$ and $\hat{C} := E^2\partial_3 - E^3\partial_2$.

2.5.2.1 The $\mathcal{H}$-$\mathcal{G}$ bracket

For β^{ij} and γ^i we have

$$
\begin{cases}\hat{C}\beta^{11}=0\\ \hat{C}\beta^{12}=-\beta^{13}\\ \hat{C}\beta^{13}=\beta^{12}\end{cases}
\begin{cases}\hat{C}\beta^{22}=-2\beta^{23}\\ \hat{C}\beta^{33}=2\beta^{23}\\ \hat{C}\beta^{23}=\beta^{22}-\beta^{33}\end{cases}
\begin{cases}\hat{C}\gamma^{1}=0\\ \hat{C}\gamma^{2}=-\gamma^3\\ \hat{C}\gamma^{3}=\gamma^2\end{cases}
\tag{2.177}
$$

For α^i we have

$$
\begin{cases}\hat{C}\bar{\alpha}^1=0\\ \hat{C}\bar{\alpha}^2=-\bar{\alpha}^3\\ \hat{C}\bar{\alpha}^3=\bar{\alpha}^2\end{cases}
\begin{cases}\hat{C}\alpha^1_1=0\\ \hat{C}\alpha^2_1=-\alpha^3_1\\ \hat{C}\alpha^3_1=\alpha^2_1\end{cases}
\begin{cases}\hat{C}\alpha^1_2=-\alpha^1_3\\ \hat{C}\alpha^2_2=-\alpha^3_2-\alpha^2_3\\ \hat{C}\alpha^3_2=\alpha^2_2-\alpha^3_3\end{cases}
\begin{cases}\hat{C}\alpha^1_3=\alpha^1_2\\ \hat{C}\alpha^2_3=\alpha^2_2-\alpha^3_3\\ \hat{C}\alpha^3_3=\alpha^3_2+\alpha^2_3\end{cases}
\tag{2.178}
$$

For Q we have

$$
\hat{C}\bar{Q}=0 \tag{2.179}
$$

$$\begin{cases}\hat{C}a_1 = 0\\ \hat{C}a_2 = -a_3\\ \hat{C}a_3 = a_2\end{cases}\begin{cases}\hat{C}b_{11} = 0\\ \hat{C}b_{12} = -b_{13}\\ \hat{C}b_{13} = b_{12}\end{cases}\begin{cases}\hat{C}b_{22} = -2b_{32}\\ \hat{C}b_{33} = 2b_{32}\\ \hat{C}b_{23} = b_{22} - b_{33}\end{cases}\begin{cases}\hat{C}c_1 = 0\\ \hat{C}c_2 = -c_3\\ \hat{C}c_3 = c_2\end{cases} \tag{2.180}$$

The remaining equations mix different coefficients:

$$\begin{cases}E^2a_3 - E^3a_2 = \bar{\alpha}^1\\ (-\alpha^1_j + 2E^2b_{3j} - 2E^3b_{2j})\partial E^j = -2(\partial E^2c_3 - \partial E^3c_2)\\ E^2c_3 - E^3c_2 = \gamma^1\end{cases}\begin{cases}E^2\alpha^1_3 - E^3\alpha^1_2 = 2\beta^{11}\\ E^2\alpha^2_3 - E^3\alpha^2_2 = 2\beta^{12} - \gamma^3\\ E^2\alpha^3_3 - E^3\alpha^3_2 = 2\beta^{13} + \gamma^2\end{cases} \tag{2.181}$$

2.5.2.2 The $\mathcal{H}$-$\mathcal{D}$ bracket

For β^{ij} and γ^i we have

$$\begin{cases}\hat{D}\beta^{11} = 0\\ \hat{D}\beta^{12} = \beta^{12}\\ \hat{D}\beta^{13} = \beta^{13}\end{cases}\begin{cases}\hat{D}\beta^{22} = 2\beta^{22}\\ \hat{D}\beta^{33} = 2\beta^{33}\\ \hat{D}\beta^{23} = 2\beta^{23}\end{cases}\begin{cases}\hat{D}\gamma^1 = 0\\ \hat{D}\gamma^2 = \gamma^2\\ \hat{D}\gamma^3 = \gamma^3\end{cases} \tag{2.182}$$

For α^i we have

$$\begin{cases}\hat{D}\bar{\alpha}^1 = \bar{\alpha}^1\\ \hat{D}\bar{\alpha}^2 = 2\bar{\alpha}^2\\ \hat{D}\bar{\alpha}^3 = 2\bar{\alpha}^3\end{cases}\begin{cases}\hat{D}\alpha^1_1 = 0\\ \hat{D}\alpha^2_1 = \alpha^2_1\\ \hat{D}\alpha^3_1 = \alpha^3_1\end{cases}\begin{cases}\hat{D}\alpha^1_2 = -\alpha^1_2\\ \hat{D}\alpha^2_2 = 0\\ \hat{D}\alpha^3_2 = 0\end{cases}\begin{cases}\hat{D}\alpha^1_3 = -\alpha^1_3\\ \hat{D}\alpha^2_3 = 0\\ \hat{D}\alpha^3_3 = 0\end{cases}\begin{cases}E^2\alpha^2_2 + E^2\alpha^2_3 = 0\\ E^2\alpha^3_2 + E^3\alpha^3_3 = 0\end{cases} \tag{2.183}$$

For Q we have

$$\begin{cases}\hat{D}\bar{Q} = 2\bar{Q}\\ E^2c_2 + E^3c_3 = 0\\ E^2a_2 + E^3a_3 = 0\end{cases}\begin{cases}c_1 + 2(b_{12}E^2 + b_{13}E^3) = 0\\ 3c_2 + 2(b_{22}E^2 + b_{23}E^3) = 0\\ 3c_3 + 2(b_{32}E^2 + b_{33}E^3) = 0\end{cases} \tag{2.184}$$

$$\begin{cases}\hat{D}c_1 = 0\\ \hat{D}c_2 = -c_2\\ \hat{D}c_3 = -c_3\end{cases}\begin{cases}\hat{D}a_1 = a_1\\ \hat{D}a_2 = 0\\ \hat{D}a_3 = 0\end{cases}\begin{cases}\hat{D}b_{11} = 0\\ \hat{D}b_{12} = -b_{12}\\ \hat{D}b_{13} = -b_{13}\end{cases}\begin{cases}\hat{D}b_{22} = -2b_{22}\\ \hat{D}b_{33} = -2b_{33}\\ \hat{D}b_{23} = -2b_{23}\end{cases} \tag{2.185}$$

One equation mixes different coefficients:

$$E^2\alpha^1_2 + E^3\alpha^1_3 = -\gamma^1\,. \tag{2.186}$$

2.5.2.3 The $\mathcal{H}$-$\mathcal{H}$ bracket

Matching term by term for $\mathcal{H}$-$\mathcal{H}$ is quite tedious, mainly because the classical bracket $\{\mathcal{H},\mathcal{H}\}$ is fully determined only after setting $\mathcal{G}=0$. For example, if there is a term $f(\alpha,\beta,\gamma,Q)\partial E^1$ on the left-hand side of $\{\mathcal{H}(x),\mathcal{H}(y)\}\approx -2G(E^{11}\partial_x\mathcal{D}(x)\delta_{xy}+2E^{11}\mathcal{D}(x)\delta'_{xy})$ which is not on the right hand side, do we demand $f(\alpha,\beta,\gamma,Q)=0$ or do we demand $f(\alpha,\beta,\gamma,Q)\propto\mathcal{G}$ or $\partial\mathcal{G}$, or does $f(\alpha,\beta,\gamma,Q)\partial E^1$ combine with possible $f(\alpha,\beta,\alpha,Q)(-E^2A_3+E^3E_2)$ terms to become something proportional to $\mathcal{G}$? There are about 10^2 terms on the left-hand side of the $\mathcal{H}$-$\mathcal{H}$ bracket, each of which has several possibilities of respecting the symmetry (in the form of second-order polynomial equations of α,β,γ,Q). It is therefore necessary to check whether these $(10^2)^n, n\sim 10^0$ possibilities are consistent with one another, rendering our current strategy impractical. Luckily, we can use an alternative strategy to find a subset of the most generic Hamiltonian by adding "semi-symmetric Gaussian" terms to the classical Hamiltonian constraint.

2.5.3 Real vs. self-dual variables

We define a *semi-symmetric* term to be any term in a generic Hamiltonian constraint that is allowed by the $\{H,D\}$ and $\{H,G\}$ brackets. These terms are solutions to our previous partial differential equations (2.177)-(2.186). We define a *Gaussian* term to be any term that is a polynomial of $\mathcal{G}$ and $\partial^n\mathcal{G}$, with coefficients denoted collectively as $C(E)$, which may depend on densitized triads and its derivatives. Namely, for a semi-symmetric Gaussian term $g(x):=g[\mathcal{G}(x),\partial^n\mathcal{G}(x),C(E(x))]$ we demand

$$\begin{cases}\{g(x),\mathcal{G}(y)\} &= 0\\ \{g(x),\mathcal{D}(y)\} &= 2G(\partial g(x)\delta_{xy}+2g(x)\delta'_{xy})\,,\end{cases} \tag{2.187}$$

Any semi-symmetric Gaussian term, $g[\mathcal{G},\partial^n\mathcal{G},C(E)]$, that we add to the classical Hamiltonian constraint $\mathcal{H}_{cl}$ is guaranteed to respect all our symmetries as shown below.

Suppose we add one semi-symmetric Gaussian term $g[\mathcal{G},\partial^n\mathcal{G},C(E)]$ to the classical Hamiltonian constraint $\mathcal{H}_{\rm cl}$

$$H[\underset{\sim}{N}]=\frac{1}{2G}\int \mathrm{d}x\underset{\sim}{N}(x)(\mathcal{H}_{\rm cl}+g)\,. \tag{2.188}$$

Since $\mathcal{H}_{\rm cl}$ respects all symmetries by definition and g is built out of semi-symmetric Gaussian terms,

$$\{H[\underset{\sim}{N}],G[M]\}=0 \tag{2.189}$$

is trivial. Similarly, the H-D bracket is satisfied:

$$
\begin{aligned}
\{H[\underset{\sim}{N}], D[M]\} &= \frac{1}{4G^2}\int \mathrm{d}x\mathrm{d}y\underset{\sim}{N}(x)M(y)(\{\mathcal{H}_{\rm cl}, \mathcal{D}\} + \{g, \mathcal{D}\}) \\
&= \frac{1}{2G}\int \mathrm{d}x\mathrm{d}y\underset{\sim}{N}(x)M(y)(\partial_x\mathcal{H}_{\rm cl}(x)\delta_{xy} + 2\mathcal{H}_{\rm cl}(x)\delta'_{xy} + \partial_x g(x)\delta_{xy} + 2g(x)\delta'_{xy}) \\
&= \frac{1}{2G}\int \mathrm{d}x\mathrm{d}y\underset{\sim}{N}(x)M(y)(\partial_x\mathcal{H}(x)\delta_{xy} + 2\mathcal{H}(x)\delta'_{xy}) = -H[M\underset{\sim}{N}' - M'\underset{\sim}{N}]
\end{aligned}
\tag{2.190}
$$

because g is built out of semi-symmetric Gaussian terms. The $H[\underset{\sim}{N}]$-$H[\underset{\sim}{M}]$ bracket then has additional terms compared with the classical case, given by $\{\mathcal{H}_{\rm cl}, g\}$ and $\{g, g\}$. Both terms are of the form $\{f, g\}$ with some semi-symmetric f, and share the property that $\int \mathrm{d}x\mathrm{d}yN(x)M(y)\{f(x), g(y)\}$ vanishes when $\mathcal{G} = 0$: In

$$
\begin{aligned}
&\int \mathrm{d}x\mathrm{d}yN(x)M(y)\{f(x), g[\mathcal{G}(y), \partial^n\mathcal{G}(y), C(E)]\} \\
= &\int \mathrm{d}x\mathrm{d}yN(x)M(y)\left(\{f(x), \mathcal{G}(y)\}\frac{\partial g}{\partial\mathcal{G}}(y) + \{f(x), \partial_y^n\mathcal{G}(y)\}\frac{\partial g}{\partial(\partial_y^n\mathcal{G})}(y)\right. \\
&\left. + \{f(x), C(E)\}\frac{\partial g}{\partial C(E)}\right) \\
= &\int \mathrm{d}x\mathrm{d}yN(x)M(y)\left(\{f(x), \mathcal{G}(y)\}\frac{\partial g}{\partial\mathcal{G}}(y) + \{f(x), C(E)\}\frac{\partial g}{\partial C(E)}\right) \\
&+ \int \mathrm{d}x\mathrm{d}yN(x)(-\partial_y)^n\left(M(y)\frac{\partial g}{\partial(\partial_y^n\mathcal{G})}(y)\right)\{f(x), \mathcal{G}(y)\},
\end{aligned}
\tag{2.191}
$$

the first and last term vanish because f is semi-symmetric, while $\partial g/\partial C(E) \approx 0$ because $C(E)$, by definition, represents coefficients in g of the Gauss constraint or its spatial derivatives.

With this result, we confirm that

$$
\begin{aligned}
\{H[\underset{\sim}{N}], H[\underset{\sim}{M}]\} &= \frac{1}{4G^2}\int \mathrm{d}x\mathrm{d}y\underset{\sim}{N}(x)\underset{\sim}{M}(y)(\{\mathcal{H}_{\rm cl}(x), \mathcal{H}_{\rm cl}(y)\} \\
&\quad +\{g[\mathcal{G}(x), \partial^n\mathcal{G}(x), C(E)], g[\mathcal{G}(y), \partial^n\mathcal{G}(y), C(E)]\} \\
&\quad +\{\mathcal{H}_{\rm cl}(x), g[\mathcal{G}(y), \partial^n\mathcal{G}(y), C(E)]\} \\
&\quad +\{g[\mathcal{G}(x), \partial^n\mathcal{G}(x), C(E)], \mathcal{H}_{\rm cl}(y)\}) \\
&\approx \frac{1}{4G^2}\int \mathrm{d}x\mathrm{d}y\underset{\sim}{N}(x)\underset{\sim}{M}(y)\{\mathcal{H}_{\rm cl}(x), \mathcal{H}_{\rm cl}(y)\}
\end{aligned}
\tag{2.192}
$$

obeys the classical brackets for any semi-symmetric g. Thus, semi-symmetric Gaussian terms indeed preserve all symmetries.

When written in real variables, the classical Hamiltonian constraint contains a term with second-order derivative of $E^1 \sim E^x$, given by $2\partial\Gamma_\phi E^x = -\partial(\partial E^x/(E^\varphi))E^x$. But when using self-dual variables, there are no second-order derivative of triads. As already mentioned, this discrepancy is caused by the fact that $\mathcal{G} \approx 0$ is already solved in the real variable case. Indeed, using semi-symmetric terms (see appendix 1) for constructing modifications we have the following allowed terms when using self-dual variables

$$\begin{aligned}\mathcal{H}_2(A,E) &= \mathcal{H}_{\rm cl}(A,E) + c_1(E^1)\left(\partial\mathcal{G} - \frac{1}{2}\frac{\partial((E^\varphi)^2)}{(E^\varphi)^2}\mathcal{G}\right) \\ &\quad + \partial E^1[b_{11}(E^1)\partial E^1 + \tilde{C}_{\alpha_1^2}(E^1)(E^3A_2 - E^2A_3)]\,, \qquad (2.193)\end{aligned}$$

where $\partial\mathcal{G} \sim \partial^2 E^1$ provides the second-order derivative. Note that the second semi-symmetric term (proportional to ∂E^1) becomes a semi-symmetric Gaussian term if we pick $b_{11} = \frac{1}{2}\tilde{C}_{\alpha_1^2}$.

Substituting $A_i = \gamma K_i + \Gamma_i$, $c_1 = E^1, b_{11} = \frac{1}{2}\tilde{C}_{\alpha_1^2} = 1/2$ in the classical Hamiltonian constraint and de-densitizing, we obtain

$$\mathcal{H}_2(K,E) = |E^x|^{-1/2}\left(K_\varphi^2 E^\varphi + 2K_\varphi K_x E^x - \left(1 - \left(\frac{\partial E^x}{2E^\varphi}\right)^2\right)E^\varphi + \frac{E^x\partial^2 E^x}{E^\varphi} - \frac{E^x\partial E^x\partial E^\varphi}{(E^\varphi)^2}\right)\,, \qquad (2.194)$$

where we used the Gauss constraint in real variables. This result matches the standard classical Hamiltonian constraint in real variables. Thus, including semi-symmetric Gaussian terms in the quadratic constraint, it is equivalent to the classical one written in real variables.

Revisiting the setting of the previous section, it follows that a further restriction of our $\mathcal{H}$ to be only quadratic in densitized triads implies that all allowed modifications to the classical $\mathcal{H}_{\rm cl}$ are in the form of semi-symmetric Gaussian terms:

$$\begin{aligned}\mathcal{H}_{\rm quad} &= C_1(\partial A_3 E^{21} - \partial A_2 E^{31} + A_{12}E^{12} + A_{13}E^{13}) + C_2\left(A_{22} + A_{33} + \frac{C_3}{C_2}\right)(E^{22} + E^{33}) \\ &\quad + C_4\partial E^1\mathcal{G} + C_5(A_2E^2 + A_3E^3)\mathcal{G}\,. \qquad (2.195)\end{aligned}$$

The first two terms are present in $\mathcal{H}_{\rm cl}$ while the last two are new semi-symmetric Gaussian terms and all C_i are constants. However, the complexity of the general equations makes it difficult to show that all possible modifications to the Hamiltonian constraint up to second order in derivatives can be constructed from semi-symmetric Gaussian terms.

2.6 Eliminating the Gauss constraint

Our analysis of gravitational theories in a setting of effective field theory has highlighted the role of the Gauss constraint, which implies that the hypersurface-deformation generators are not uniquely defined. Since the Gauss constraint contains a spatial derivative, and spatial derivatives of this constraint can also be added to the hypersurface-deformation generators, the derivative structure and therefore the possibility of modifications is ambiguous as long as the Gauss constraint remains unsolved. We will therefore now solve the Gauss constraint explicitly and analyze the resulting hypersurface-deformation generators and their brackets.

2.6.1 Gauge-invariant variables

We begin with the classical constraint

$$\begin{aligned} H[N] &= \frac{1}{2G}\int \mathrm{d}x \frac{N}{\sqrt{E^1((E^2)^2+(E^3)^2)}}\Big(2E^1(E^2A_3'-E^3A_2') \\ &\quad +2A_1E^1(A_2E^2+A_3E^3)+(A_2^2+A_3^2-1)((E^2)^2+(E^3)^2) \\ &\quad +(\epsilon-\gamma^2)(2K_1E^1(K_2E^2+K_3E^3)+(K_2^2+K_3^2)((E^2)^2+(E^3)^2)\Big) \end{aligned} \tag{2.196}$$

in which the lapse function no longer has a density weight. The next few transformations closely follow the derivations given in [39], but are presented here in a different form using vector notation.

The pairs (E^2, E^3) and (A_2, A_3) (as well as (K_2, K_3)) transform under the defining representation of SO(2) with respect to the Gauss constraint. It will be convenient to arrange them in 3-vectors, such that

$$\vec{E} = E^2\vec{e}_2 + E^3\vec{e}_3 \tag{2.197}$$

$$\vec{A} = A_2\vec{e}_2 + A_3\vec{e}_3 \tag{2.198}$$

$$\vec{K} = K_2\vec{e}_2 + K_3\vec{e}_3 \tag{2.199}$$

with standard basis vectors $\vec{e}_i$. Obvious invariant variables are therefore

$$E^\varphi = |\vec{E}| = \sqrt{(E^2)^2+(E^3)^2} \tag{2.200}$$

$$A_\varphi = |\vec{A}| = \sqrt{A_2^2+A_3^2} \tag{2.201}$$

$$K_\varphi = |\vec{K}| = \sqrt{K_2^2+K_3^2}\,. \tag{2.202}$$

Moreover, we obtain another invariant α from the angle between $\vec{E}$ and $\vec{A}$,

$$\cos\alpha = \frac{\vec{E}\cdot\vec{A}}{E^\varphi A_\varphi}\,. \tag{2.203}$$

While E^1 and K_1 are also invariant, A_1 has a non-trivial transformation. A final gauge-invariant expression can be written as $A_1 + \beta'$, where

$$\cos\beta = \frac{\vec{e}_2\cdot\vec{A}}{A_\varphi}\,. \tag{2.204}$$

Using our definitions of α and β, we can write the unit vectors

$$\vec{e}_A = \frac{\vec{A}}{A_\varphi} = \vec{e}_2\cos(\beta) + \vec{e}_3\sin(\beta) \tag{2.205}$$

$$\vec{e}_E = \frac{\vec{E}}{E^\varphi} = \vec{e}_2\cos(\alpha+\beta) + \vec{e}_3\sin(\alpha+\beta)\,. \tag{2.206}$$

From the last relation one can derive the spin-connection component $\Gamma_1 = -(\alpha+\beta)'$ [39]. Therefore, $\gamma^{-1}(A_1+\alpha'+\beta') = K_1$ is nothing but an extrinsic-curvature component. Since α and K_1 are gauge invariant, $A_1 + \beta'$ must be gauge invariant, as claimed above.

Moreover, computing the extrinsic curvature and spin connection for a spherically symmetric triad [39] shows that the angular part $\vec{K}$ points in the same internal direction as the triad,

$$\vec{e}_K = \vec{e}_E\,, \tag{2.207}$$

while the angular part of the spin connection, $\vec{\Gamma}$, is orthogonal,

$$\vec{e}_\Gamma = -\vec{e}_1\times\vec{e}_E\,, \tag{2.208}$$

with coefficient

$$\Gamma_\varphi = -\frac{(E^1)'}{2E^\varphi}\,; \tag{2.209}$$

see (2.161). Therefore,

$$A_\varphi^2 = |\vec{A}|^2 = |\Gamma_\varphi\vec{e}_\Gamma + \gamma K_\varphi\vec{e}_K|^2 = \Gamma_\varphi^2 + \gamma^2 K_\varphi^2\,. \tag{2.210}$$

The term in (2.196) containing spatial derivatives of the connection can now be written as

$$\begin{aligned} E^2A_3' - E^3A_2' &= \vec{e}_1\cdot(\vec{E}\times\vec{A}') = E^\varphi\vec{e}_1(\vec{e}_E\times(A_\varphi\vec{e}_A)') \\ &= E^\varphi(-A_\varphi'\sin(\alpha)+A_\varphi\beta'\cos(\alpha))\,. \end{aligned}$$

We then express connection terms through spin connection and extrinsic curvature, using

$$A_\varphi\sin(\alpha) = A_\varphi\vec{e}_A\cdot\vec{e}_\Gamma = \Gamma_\varphi \tag{2.211}$$

and

$$A_\varphi\cos(\alpha) = A_\varphi\vec{e}_A\cdot\vec{e}_K = \gamma K_\varphi\,. \tag{2.212}$$

Therefore,

$$\begin{aligned} E^2A_3' - E^3A_2' &= E^\varphi(-(A_\varphi\sin(\alpha))+A_\varphi(\alpha'+\beta')\cos(\alpha)) \\ &= E^\varphi\left(-\Gamma_\varphi'+\gamma K_\varphi(\alpha'+\beta')\right)\,. \end{aligned} \tag{2.213}$$

The angles in the last term can be combined with a similar contribution from the second term in (2.196), which adds A_1 to $\alpha'+\beta'$. (In (2.196), A_1 is multiplied with $A_2E^2 + A_3E^3 = \vec{A}\cdot\vec{E} = \gamma K_\varphi E^\varphi$, which does not depend on Γ_φ because $\vec{e}_\Gamma\cdot\vec{e}_E = 0$.) Since $\alpha'+\beta' = -\Gamma_1$ [39] and $A_1 - \Gamma_1 = \gamma K_1$, we have

$$E^2A_3' - E^3A_2' + A_1(A_2E^2+A_3E^3) = E^\varphi\left(-\Gamma_\varphi'+\gamma^2K_\varphi K_1\right)\,. \tag{2.214}$$

Thus, by using variables invariant under transformations generated by the Gauss constraint, we have been led to an expression in which all spatial derivatives of the connection have been replaced by spatial derivatives of the triad (through Γ_φ).

Again in [39], the Poisson brackets

$$\{K_\varphi(x),E^\varphi(y)\} = G\delta(x,y), \quad \{K_1(x),E^1(y)\} = 2G\delta(x,y) \tag{2.215}$$

for the new gauge-invariant variables have been derived. (Note that this is only consistent if (2.207) is true.) If we express the diffeomorphism and Hamiltonian constraints in these variables, we restrict the previous theory to the solution space of the Gauss constraint. We obtain

$$D[N^x] = \frac{1}{2G}\int \mathrm{d}x N^x\left(2E^\varphi K_\varphi' - K_1(E^1)'\right) \tag{2.216}$$

and

$$H[N] = \frac{1}{2G}\int \mathrm{d}x \frac{N}{\sqrt{E^1}}\Big(K_\varphi^2 E^\varphi(\epsilon-\gamma^2) + 2\epsilon K_\varphi K_1 E^1 + (\Gamma_\varphi^2+\gamma^2 K_\varphi^2 - 1)E^\varphi - 2E^1\Gamma'_\varphi\Big)\,. \tag{2.217}$$

2.6.2 Modified constraint with classical brackets

In the Hamiltonian constraint, the two terms with $\gamma^2 K_\varphi^2$ cancel out, showing that, for $\epsilon = -1$, we obtain the Hamiltonian constraint as considered in [39]. Our calculation here extends this result to Euclidean signature, $\epsilon = 1$. Since all γ-dependent terms drop out of the final expression, it is no longer clear why $\gamma^2 = \epsilon$ should lead to different options for modified constraints. Nevertheless, the previous distinction between $\gamma^2 = \epsilon$ and $\gamma^2 \neq \epsilon$ can still be realized if we do not cancel the γ-dependent terms in (2.217) *before* we try to modify the constraint. In particular, the previous modification, using an arbitrary function of $f(A_2^2+A_3^2-1)$, can still be implemented in the invariant version if we recognize the combination $\Gamma_\varphi^2+\gamma^2K_\varphi^2-1$ as the correct substitute of $A_2^2+A_3^2-1 = A_\varphi^2-1$. We therefore consider the modified constraint

$$H[N] = \frac{1}{2G}\int \mathrm{d}x \frac{N}{\sqrt{E^1}}\Big(K_\varphi^2 E^\varphi(\epsilon-\gamma^2) + 2\epsilon K_\varphi K_1 E^1 + f(\Gamma_\varphi^2+\gamma^2 K_\varphi^2 - 1)E^\varphi - 2E^1\Gamma'_\varphi\Big)\,. \tag{2.218}$$

Given the form of this new constraint, it is not obvious that it can lead to closed brackets for f not equal to its classical form because, compared with our previous derivation, we now have up to second-order spatial derivatives of the triad (through Γ_φ) instead of first-order derivatives of its momenta.

Thanks to antisymmetry of the Poisson bracket, the only terms that give non-zero contributions to $B_{NM} := \{H[N], H[M]\}$ are combinations of a term from $H[N]$ depending on one of the K_i and a term from $H[M]$ depending on a (first or second order) spatial derivative of one of the E_i, or vice versa. Therefore,

$$\begin{aligned} B_{NM} = \frac{1}{4G^2}\int \mathrm{d}x\mathrm{d}y \frac{N(x)M(y)}{\sqrt{E^1(x)E^1(y)}}\Bigg(&-(\epsilon-\gamma^2)\{K_\varphi^2(x),(E^\varphi)'\}\frac{E^1(y)E^1(y)'E^\varphi(x)}{(E^\varphi(y))^2} \\ &-2\epsilon\{K_\varphi(x),E^\varphi(y)'\}K_1(x)\frac{E^1(x)E^1(y)E^1(y)'}{(E^\varphi(y))^2} \\ &-2\epsilon K_\varphi(x)\{K_1(x),E^1(y)'\}\frac{E^1(x)E^1(y)E^\varphi(y)'}{(E^\varphi(y))^2} \end{aligned}$$

$$
-\{f, E^\varphi(y)'\}\frac{E^\varphi(x)E^1(y)E^1(y)'}{(E^\varphi(y))^2} + 2\epsilon K_\varphi(x)\{K_1(x), f\}E^1(x)E^\varphi(y)
$$
$$
\left. +2\epsilon K_\varphi(x)\frac{E^1(y)}{E^\varphi(y)}\{K_1(x), E^1(y)''\}E^1(x)\right) - (N \leftrightarrow M)\,. \tag{2.219}
$$

Integrating by parts, we obtain

$$
\begin{aligned}
B_{NM} &= \frac{1}{4G}\int \mathrm{d}x NM'\left((2(\epsilon-\gamma^2)K_\varphi\frac{(E^1)'}{E^\varphi} + 2\epsilon\frac{E^1}{(E^\varphi)^2}K_1(E^1)' + 4\epsilon K_\varphi(E^\varphi)'\frac{E^1}{(E^\varphi)^2}\right.\\
&\quad \left. -4\epsilon\frac{E^1}{(E^\varphi)^2}E^\varphi K_\varphi' - 4\epsilon K_\varphi\frac{E^1(E^\varphi)'}{(E^\varphi)^2} + \frac{\partial f}{\partial K_\varphi}\frac{(E^1)'}{E^\varphi} - 4\epsilon K_\varphi E^\varphi\frac{\partial f}{\partial (E^1)'}\right) - (N \leftrightarrow M)\\
&= \frac{-\epsilon}{2G}\int \mathrm{d}x\frac{E^1}{(E^\varphi)^2}(NM'-N'M)(2E^\varphi K_\varphi' - K_1(E^1)')\\
&\quad +\frac{1}{4G}\int \mathrm{d}x(NM'-N'M)\left(2(\epsilon-\gamma^2)K_\varphi\frac{(E^1)'}{E^\varphi} + \frac{\partial f}{\partial K_\varphi}\frac{(E^1)'}{E^\varphi} - 4\epsilon K_\varphi E^\varphi\frac{\partial f}{\partial (E^1)'}\right)\\
&= -\epsilon D\left[\frac{E^1}{(E^\varphi)^2}(NM'-N'M)\right] \qquad\qquad (2.220)\\
&\quad +\frac{1}{4G}\int \mathrm{d}x(NM'-N'M)\left(2(\epsilon-\gamma^2)K_\varphi\frac{(E^1)'}{E^\varphi} + \frac{\partial f}{\partial K_\varphi}\frac{(E^1)'}{E^\varphi} - 4\epsilon K_\varphi E^\varphi\frac{\partial f}{\partial (E^1)'}\right).
\end{aligned}
$$

For a closed bracket, therefore,

$$
2(\epsilon-\gamma^2)K_\varphi\frac{(E^1)'}{E^\varphi} + \frac{\partial f}{\partial K_\varphi}\frac{(E^1)'}{E^\varphi} - 4\epsilon K_\varphi E^\varphi\frac{\partial f}{\partial (E^1)'} = 0\,. \tag{2.221}
$$

Since f depends on K_φ and $(E^1)'$ only through $\frac{1}{4}(E^{1\prime})^2/(E^\varphi)^2 + \gamma^2K_\varphi^2 - 1$, the chain rule implies that

$$
\frac{\partial f}{\partial K_\varphi} = 2\gamma^2K_\varphi\dot{f} \quad \text{and} \quad \frac{\partial f}{\partial (E^1)'} = \frac{1}{2(E^\varphi)^2}(E^1)'\dot{f}\,, \tag{2.222}
$$

and (2.221) is equivalent to

$$
2(\epsilon-\gamma^2)K_\varphi\frac{(E^1)'}{E^\varphi}\left(1-\dot{f}\right) = 0\,. \tag{2.223}
$$

If $\gamma^2 = \epsilon$, the equation holds identically for any f. If $\gamma^2 \neq \epsilon$, however, $\dot{f} = 1$, and only the classical case is allowed. The modification found in [29] can therefore be found also in gauge-invariant variables, in which case the Hamiltonian constraint contains second-order derivatives of the triad, with the same restriction that it is allowed only for a specific value of γ.

2.6.3 Modified brackets

A generic modification that does not require a specific value of γ can be obtained for the theories considered here, as has been known for some time for real variables [9, 13]. Since the Hamiltonian constraint in real variables has the same form as the general spherically symmetric constraint in gauge-invariant variables, the same modification can be transferred also to self-dual type variables ($\gamma^2 = \epsilon$) provided we implement it at the gauge-invariant level. At the level of variables that are not gauge invariant, this new modification (compared with [29]) is possible provided we use the Gauss constraint to reintroduce second-order derivatives of triads in the Hamiltonian constraint.

Starting with (2.217), the new modification is derived in a way very similar to the case of real variables, found in [9]. Nevertheless, we reproduce the calculation of brackets here for the sake of completeness. We modify (2.217) to

$$\begin{aligned} H[N] =& \frac{1}{2G}\int \mathrm{d}x N(x)(E^1)^{-1/2}\Big(\epsilon f_1(K_\varphi)E^\varphi + 2\epsilon f_2(K_\varphi)E^1K_1 \\ &+\left(\frac{(E^{1\prime})^2}{4(E^\varphi)^2}-1\right)E^\varphi + \frac{E^1(E^1)''}{E^\varphi} - \frac{E^1(E^1)'(E^\varphi)'}{(E^\varphi)^2}\Big) \end{aligned} \tag{2.224}$$

with two functions, f_1 and f_2, that will be restricted further by the condition of having closed brackets. We first interpret this modification based on arguments within canonical effective field theory. We are now allowing for a non-quadratic dependence of the Hamiltonian constraint on K_φ. If K_φ is still a first-order time derivative, a non-quadratic dependence would be non-generic unless we also allow for higher-order spatial derivatives of the densitized triad, which we do not do in (2.224).

However, modifying the Hamiltonian constraint in this form also modifies the equations of motion that classically imply the first-order nature of K_φ. An analysis of these modified equations of motion should then be performed in order to reveal the derivative order of the Hamiltonian constraint. Schematically, we obtain the modified derivative dependence of K_φ from the equation of motion

$$\dot{E}^1 = 2N\sqrt{E^1}f_2(K_\varphi) + N^1(E^1)' \tag{2.225}$$

$$\begin{aligned} \dot{E}^\varphi =& N\sqrt{E^1}K_1\frac{\mathrm{d}f_2(K_\varphi)}{\mathrm{d}K_\varphi} + \frac{NE^\varphi}{2\sqrt{E^1}}\frac{\mathrm{d}f_1(K_\varphi)}{\mathrm{d}K_\varphi} \\ &+(N^1E^\varphi)' \end{aligned} \tag{2.226}$$

provided we can invert the function f_2. This can explicitly be done only in examples, which we restrict here to the common case of $f_1(K_\varphi) = \sin^2(K_\varphi)$, which implies $f_2(K_\varphi) = \sin(K_\varphi)\cos(K_\varphi)$ or $f_2(K_\varphi)^2 = f_1(K_\varphi)(1 - f_1(K_\varphi))$. The latter equation can be solved for

$$\begin{aligned} f_1(K_\varphi) &= \frac{1}{2}\left(1 \pm \sqrt{1 - 4f_2(K_\varphi)^2}\right) \\ &= f_2(K_\varphi)^2 + f_2(K_\varphi)^4 + \cdots . \end{aligned} \tag{2.227}$$

According to (2.225), $f_2(K_\varphi)$ is strictly of first order in derivatives, but $f_1(K_\varphi)$ is not polynomial in $f_2(K_\varphi)$, and therefore a derivative expansion of $f_1(K_\varphi)$ does not terminate. Similarly,

$$\frac{\mathrm{d}f_2(K_\varphi)}{\mathrm{d}K_\varphi} = \cos(2K_\varphi) = 1 - 2f_1(K_\varphi) = \sqrt{1 - f_2(K_\varphi)^2} \tag{2.228}$$

has a derivative expansion that does not terminate. Therefore, K_1 has a non-terminating derivative expansion because $K_1\sqrt{1 - f_2(K_\varphi)^2}$ must be of first order according to (2.226).

We conclude that the constraint (2.224) contains a derivative expansion in both space and time derivatives, which can consistently be truncated at any finite derivative order. The resulting effective theory is therefore meaningful, but it may not be the most general one because the derivative expansion results only from the K-dependent terms in (2.224) while we have not included higher-derivative corrections of the E-dependent terms. (Higher spatial derivatives may be expected from an expansion of non-local holonomies used in the Hamiltonian constraints for models of loop quantum gravity; see for instance [40, 41]. However, it is difficult to find consistent constraint brackets with such modifications [13].) The mismatch does not violate (deformed) covariance because the constraint brackets still close. However, unless the symmetries implied by the closed constraints select only this specific derivative structure, the modified theory is not generic. (It resembles Born–Infeld type theories.) Since no other consistent modifications are known as of now, it remains unclear whether the apparently non-generic model is selected by symmetries.

In order to confirm that the constraint brackets can be closed, we compute

$$\begin{aligned} \{H[N], H[M]\} &= \frac{1}{4G^2}\int \mathrm{d}x\mathrm{d}y \frac{N(x)M(y)}{\sqrt{E^1(x)E^1(y)}}\left(-\epsilon\frac{E^\varphi(x)E^1(y)E^1(y)'}{(E^\varphi)^2(y)}\{f_1(K_\varphi(x)), E^\varphi(y)'\}\right. \\ &\quad -2\epsilon\frac{E^1(x)E^1(y)E^1(y)'K_1(x)}{(E^\varphi)^2(y)}\{f_2(K_\varphi(x), E^\varphi(y)'\} \\ &\quad +\epsilon\frac{f_2(K_\varphi(x))E^1(x)}{2E^\varphi(y)}\{K_1(x), (E^1(y)')^2\} \end{aligned}$$

$$+2\epsilon f_2(K_\varphi(x))E^1(x)\frac{E^1(y)}{E^\varphi(y)}\{K_1(x),E^1(y)''\}$$
$$-2\epsilon f_2(K_\varphi(x))E^1(x)\frac{E^1(y)E^\varphi(y)'}{E^\varphi(y)^2}\{K_1(x),E^1(y)'\}\Big)-(N\leftrightarrow M) \quad (2.229)$$

writing only terms that produce non-zero contributions. All terms are multiplied with ϵ, and therefore the possibility of modifications does not depend on the space-time signature.

The first two lines contain Poisson brackets of $f_1(K_\varphi)$ and $f_2(K_\varphi)$ and therefore lead to derivatives of the modification functions:

$$\frac{1}{G}\frac{E^\varphi(x)E^1(y)E^1(y)'}{(E^\varphi)^2(y)}\{f_1(K_\varphi(x)),E^\varphi(y)'\}=\frac{E^\varphi(x)E^1(y)E^1(y)'}{(E^\varphi)^2(y)}\frac{\mathrm{d}f_1(K_\varphi)}{\mathrm{d}K_\varphi}\partial_y\delta(x,y) \quad (2.230)$$

and

$$\frac{2}{G}\frac{E^1(x)E^1(y)E^1(y)'K_1(x)}{(E^\varphi)^2(y)}\{f_2(K_\varphi(x)),E^\varphi(y)'\}=2\frac{E^1(x)E^1(y)E^1(y)'K_1(x)}{(E^\varphi)^2(y)}\frac{\mathrm{d}f_2(K_\varphi)}{\mathrm{d}K_\varphi}\partial_y\delta(x,y)\,. \quad (2.231)$$

Another derivative of $f_2(K_\varphi)$ results from the second-order derivative of the delta function obtained after evaluating $\{K_1,(E^1)''\}$ in the fourth line of (2.229). This contribution follows from

$$2f_2(K_\varphi(x))\frac{E^1(x)E^1(y)}{E^\varphi(y)}\{K_1(x),E^1(y)''\}=4f_2(K_\varphi(x))\frac{E^1(x)E^1(y)}{E^\varphi(y)}\partial_y^2\delta(x,y)\,. \quad (2.232)$$

Upon integrating by parts twice in the resulting expression in (2.229), we initially produce a term with $N(x)M(y)''$ times a delta function without derivatives. Integrating over y, the delta function is eliminated and we can integrate by parts once again to obtain a term with $N'M'$ (which cancels out in the antisymmetric bracket) and a term with NM' times the derivative of the entire coefficient in (2.232):

$$-4\left(f_2(K_\varphi)\frac{(E^1)^2}{E^\varphi}\right)'=-4\left(\frac{\mathrm{d}f_2}{\mathrm{d}K_\varphi}K_\varphi'\frac{(E^1)^2}{E^\varphi}+f_2(K_\varphi)\left(2\frac{E^1(E^1)'}{E^\varphi}-\frac{(E^1)^2(E^\varphi)'}{(E^\varphi)^2}\right)\right)\,. \quad (2.233)$$

The last term (containing $(E^\varphi)'$) cancels out with the fifth line of (2.229), while only half the second term cancels out with the third line of (2.229), for any f_2. In order for the remaining terms to be proportional to the diffeomorphism constraint, only expressions proportional to K_1 or K_φ' can remain. Therefore, the other half of the second term in

(2.233) must cancel out with (2.230), which requires

$$f_2(K_\varphi) = \frac{1}{2}\frac{\mathrm{d}f_1(K_\varphi)}{\mathrm{d}K_\varphi}\,. \tag{2.234}$$

Only two terms are then left, (2.231) and the first contribution in (2.233). They are both proportional to $\mathrm{d}f_2(K_\varphi)/\mathrm{d}K_\varphi$ and combine to form the diffeomorphism constraint:

$$\begin{aligned}\{H[N],H[M]\} &= -\frac{\epsilon}{2G}\int \mathrm{d}x N'M\frac{E^1}{(E^\varphi)^2}\frac{\mathrm{d}f_2}{\mathrm{d}K_\varphi}(2E^\varphi K_\varphi' - K_1(E^1)') - (N\leftrightarrow M)\\ &= -\epsilon D\left[\frac{\mathrm{d}f_2(K_\varphi)}{\mathrm{d}K_\varphi}\frac{E^1}{(E^\varphi)^2}(NM'-N'M)\right]\,. \end{aligned} \tag{2.235}$$

This modification, following [9, 13], differs from the modification of [29] in that it modifies not only the constraints but also their brackets (while the latter remain closed). It, therefore, implies a new non-classical space-time structure [20, 21]. This modification is consistent for all γ and is therefore generic. From this perspective, the modification of [29], which preserves the brackets, requires $\gamma^2 = \epsilon$ and is not generic; it does not provide a way to avoid non-classical space-time structures without fine-tuning. Our derivations have shown that the different outcomes of [29] versus [9, 13] are not a consequence of working with self-dual connections (used in [29]) or real variables (used in [9, 13]). The crucial difference is that modified constraints with unmodified brackets, as in [29], can be obtained only for specific γ, while modifications of constraints as well as brackets exist for all γ.

2.7 Conclusion

We have shown that deformations of the classical space-time structure appear generically in spherically symmetric models of loop quantum gravity. For self-dual variables or Euclidean gravity with $\gamma = \pm 1$, we have derived the most general form of the quadratic Hamiltonian constraint free of triad derivatives, such that a system with unmodified closed brackets is obtained. This rigidity result, just like the setting of [29] which it generalizes, relies on the absence of derivative terms of the triad. However, from the point of view of an effective field theory, this result is not generic because it depends on a restriction of derivative terms even within the classical structure of second-order derivatives. Moreover, this rigidity result can be obtained only for specific values of the Barbero–Immirzi parameter γ.

The results of [29] have sometimes been interpreted as saying that deformations arising in the hypersurface-deformation brackets, obtained originally using holonomy modifications in real-valued variables, might be avoided in the self-dual case. Self-dual variables represent a specific choice for the Immirzi parameter, and therefore do not lead to generic results. These variables (or the values of γ they correspond to) are not distinguished intrinsically by symmetries because constraint brackets, which define the symmetries of a canonical theory, can be closed for any γ.

Moreover, we have shown that the possibility of modifications, even within a self-dual setting, formally depends on the derivative structure which can be changed by adding multiples of the Gauss constraint or its spatial derivatives to the Hamiltonian constraint. This ambiguity can be eliminated by solving the Gauss constraint explicitly, following [39], in which case the same derivative structure is obtained in self-dual type variables and in real variables, which agrees with the form originally used in an analysis of modified brackets [9, 13]. We therefore conclude that modified brackets and non-classical space-time structures are generic in any spherically symmetric model with holonomy modifications, even for self-dual variables. We also pointed out that the currently known modifications may not be generic from the point of view of canonical effective theory introduced here: After translating momenta into time derivatives, different derivative orders appear in the terms of a modified Hamiltonian constraint. This observation suggests that there is room for further explorations of possibly new models. A likely candidate for a generic extension is the inclusion of canonical quantum back-reaction effects [42–44], which in an action formulation provide higher-curvature terms with generic higher derivatives. However, quantum back-reaction on its own does not modify the hypersurface-deformation brackets of constraints [45] and is therefore unlikely to change our conclusions about modified space-time structures.

Euclidean and self-dual type variables are special also in an analysis of cosmological perturbations [46,47], in which case non-generic modifications of constraint brackets have been observed as well. Our results present useful indications for operator calculations [48–54] which have demonstrated the possibility of off-shell closure of commutators of constraint operators, mainly in the Euclidean case. So far, these investigations have not yet given rise to indications that the commutators of constraint operators may be modified, in contrast to effective derivations as well as the operator constructions in [14, 55]. (However, it is not always clear how to read off modifications of structure functions in the operator setting, which should be some function of a spatial metric or densitized triad and therefore requires a suitable notion of states of a semi-classical geometry which does not yet exist in the operator formulation.) Our results show that the Euclidean setting is, in fact, inconclusive as regards modifications of structure functions because it is a non-generic case that allows closed brackets with and without modifications. Current effective and operator treatments are therefore consistent with one another. For a complete picture of space-time structures in loop quantum gravity, it will be important to extend off-shell operator calculations to the full Lorentzian constraint.

While the issues above remain open, they share roots with other common attempts to quantize gravity and are not unique to loop quantum gravity. For example, even in string theory, the construction of stable semi-classical states, such as ones corresponding to the quasi de Sitter space-time, has been a longstanding challenge. Ambiguities of the state propagate down to the resulting effective theory and force us to cast doubt on the traditional predictions in quantum cosmology—stability of de Sitter, equation of state of dark energy, and what inflaton models are allowed, just to name a few. Canonical effective methods provide a systematic way of analyzing consistencies and determine how much the notion of space-time still holds. In the next chapter, we will also see how we can parameterize the state in hopes of alleviating some of the state ambiguities. Canonical effective methods will play a key role in describing the evolution of these parameters and provide us with a time-dependent way to describe effective potentials.

Chapter 3
Cosmology

3.1 Introduction

As we have alluded to in the previous chapter, in many physical situations of interest, it can be cumbersome or even impossible to specify the exact quantum state. As an alternative, we look for an effective description that parameterizes our state with a set of observables. A set of observables are the quantum fluctuations (also known as moments) associated with a state. In principle, a quantum state can be reconstructed if we know the moments of all the operators of the Hilbert space. In practice, we can make predictions with just a few important ones—among them are the second-order moments (2-point functions) and the non-Gaussianities (3-point functions). In a Lagrangian framework, these moments typically have non-linear couplings between their equations of motion. Moreover, within a given truncation of the order of moments, not all of them are true independent degrees of freedom. We will see how these difficulties are naturally resolved in the canonical framework. Consequently, an effective system with an enlarged number of degrees-of-freedom (DoFs) is produced. The new DoFs characterize time-dependent corrections to the low energy effective potential (which is traditionally approximated by assuming strict spatially and temporally homogeneous vacuum in its computation).

There are more practical reasons why treating quantum corrections as effective DoFs in phase space can be helpful. They help support certain single-field inflaton models and offer ways to evade trans-Planckian issues without resorting to fine-tuning or new physics.

It is a natural requirement that self-consistent inflationary models should be largely independent of the high energy quantum gravity theory, viewed in an effective field theory framework. However, an exact decoupling of scales relevant for inflation from high-energy modes can happen only if the low-energy Lagrangian consists entirely of terms that are renormalizable using Wilsonian effective actions). This condition restricts single-field models of inflation to be of the chaotic type with quartic potentials.

If the inflationary action contains terms beyond mass-dimension four, then the theory is liable to be affected by as yet unknown high-energy physics. In fact, one even has to rely on ultraviolet physics in order to derive a suitable higher-order form of the potential. In common single-field inflation, this problem can rarely be avoided as the models preferred by observations [56] depend crucially on non-renormalizable terms in the potential, for instance in Starobinsky inflation [57]. Fundamentally, such terms have to be understood as remnants in an effective description of some underlying theory of gravity and matter, such as quantum gravity or string theory, but specific top-down justifications of suitable forms of the potential are usually hard to come by.

Alternatively, if chaotic-type potentials, which have been ruled out by data as single-field models, can somehow be resurrected, then the burden of explaining these potentials does not have to fall on quantum gravity. Motivated by this observation, we begin with a Higgs-inspired classical potential,

$$V_{\rm cl}(\psi) = M^4\left(1 - 2\frac{\psi^2}{v^2} + \frac{\psi^4}{v^4}\right) \tag{3.1}$$

with two parameters, M and v, assumed to be positive. While the only known scalar to have been discovered to date is the standard-model Higgs particle, it is well-known that this type of an inflaton potential, by itself, is found to be inconsistent with cosmological observations. To make matters worse, even renormalization-group improvements do not suffice to make Higgs-like potentials compatible with data [58–60]. The only observationally consistent formulations proposed up until now have been based on a scalar field non-minimally coupled to the Ricci scalar [61, 62], modifying the kinetic term of the Higgs field. Non-minimal coupling terms, however, mean that one is forced to modify the nature of the standard model at high energies [63], amongst other issues [64].

In the present work, we will preserve the simple nature of a minimally coupled field with a quartic classical potential (3.1). Applying a canonical formalism of effective theory which, crucially, remains valid in non-adiabatic regimes. Heuristically, this formalism includes effects of higher time derivatives in the quantum effective action without requiring a derivative expansion. We emphasize that this notion of non-adiabatic behavior refers to the background state of the scalar field, describing its homogeneous contribution, rather than its inhomogeneous modes which may well remain largely adiabatic if the slow-roll regime is sufficiently long. By these methods, the classical potential will be quantum extended to a two-field model with a specific potential derived from (3.1). The second field, φ, will be shown to be an authentic field degree of freedom representing quantum fluctuations of the background inflaton, ψ. (Our application of quantum fluctuations in the context of inflation is different from the way they appear in stochastic or eternal inflation [65–70]. We are using back-reaction effects implied by quantum fluctuations in the deterministic evolution of wave functions, rather than stochastic properties implied by fluctuations for the measurement process. Our model could certainly be extended by including suitable stochastic terms in the equations of motion [66, 71–74], but we will not attempt to do so in the present paper. As such, it is subject to uncertainty relations that will be used to obtain important lower bounds on its initial value. Initial evolution is then non-adiabatic, but it automatically sets the stage for a long slow-roll phase (in a so-called waterfall regime of the two-field model) that is consistent with observational constraints. A final non-adiabatic phase automatically ends inflation with just the right number of e-folds in a large region of the parameter space.

Coefficients of the two-field potential are determined by the same two parameters, M and v, that appear in the single-field model (3.1). In addition, there are new coefficients derived from moments of the inflaton state, such as parameters for non-Gaussianity of the background state. In inflation models, this is a new kind of non-Gaussianity different from what one usually refers to in primordial fluctuations during inflation. In our case, non-Gaussianity is present already in the wave function of the homogeneous quantum inflaton field (referred to here as the background state), and not only in the perturbation spectrum. It is therefore possible to put constraints on the two-field potential based on known properties of states, or conversely, to determine conditions on suitable inflaton states based on observational constraints. An important finding is that constraints on the spectral index, its running, and the tensor-to-scalar ratio prefer small background non-Gaussianity.

In Section 3.2, we present a review of relevant methods of non-adiabatic quantum dynamics, which have appeared in various forms in fields as diverse as quantum field theory, quantum chaos, quantum chemistry, and quantum cosmology. The same section presents a comparison with Gaussian methods and shows how non-adiabatic dynamics can include non-Gaussian states. These methods are applied to cosmology in Section 3.3, focusing on Higgs-like inflation. The results are, however, more general and can easily be adapted to any potential. This section will demonstrate the importance of going beyond Gaussian dynamics, including higher-order moments, and maintaining non-adiabatic regimes. A detailed cosmological analysis, including numerical simulations and analytical approximations, is performed in Section 3.4, where observational implications are discussed. The derivations in the present paper justify the more concise physical discussion presented in [75].

3.2 Canonical effective potentials

Our construction is based on canonical effective methods for non-adiabatic quantum dynamics, which in a leading-order treatment has appeared several times independently in various fields [2, 76–80], including quantum chaos, quantum chemistry, and quantum cosmology, but has only recently been worked out to higher orders using systematic methods of Poisson manifolds [81, 82]. While higher orders go beyond Gaussian dynamics, the leading-order effects are closely related to Gaussian approximations and can therefore be used for an illustration of the method.

Throughout this paper, the term "adiabatic" will by default refer to the concepts discussed in the present section, unless otherwise stated. That is, adiabatic behavior is by definition realized when it is possible to capture crucial physical phenomena in a derivative expansion, for instance in a quantum effective action or, as used below, in the equations of motion for expectation values and moments of a state. The behavior is non-adiabatic when a derivative expansion does not faithfully capture the dynamics. In this case, new non-classical degrees of freedom play an important role, which may be given by auxiliary fields in a non-local effective action, or independent moments of a quantum state. In general, this notion has no relationship with the concept of adiabatic modes which is often used in cosmology. Later on, we will however briefly use an adiabatic combination of fields when deriving observables in the context of multi-field inflation.

3.2.1 Relation to the time-dependent variational principle

In order to illustrate our claim that quantum fluctuations can provide an independent degree of freedom that can influence the inflationary dynamics, we first consider a canonical formulation of the time-dependent variational principle for Gaussian states.

The most general parametrization of Gaussian fluctuations around the homogeneous field ψ can be represented by the wave function [76]

$$\Psi(\psi'|\psi,\pi_\psi,\varphi,\pi_\varphi) \quad = \quad \frac{1}{\left(2\pi\varphi^2\right)^{1/4}}\exp\left(-\tfrac{1}{4}\varphi^{-2}(1-2i\varphi\pi_\varphi)(\psi'-\psi)^2\right)$$

$$\times \exp(i\pi_\psi(\psi' - \psi)) \exp(-\tfrac{1}{2}i\varphi\pi_\varphi) \,. \qquad (3.2)$$

The notation is such that Ψ is a wave function depending on a free variable ψ' for any choice of the parameters ψ, π_ψ, φ and π_φ, which determine a specific ψ'-dependent wave function. Despite its lengthy form, this variational wave function has some useful properties: It is normalized, $\langle\Psi|\Psi\rangle = 1$, and has basic expectation values

$$\langle\Psi|\hat{\psi}|\Psi\rangle = \psi \quad , \quad \langle\Psi|\hat{\pi}_\psi|\Psi\rangle = \pi_\psi \qquad (3.3)$$

and variances

$$\langle\Psi|(\hat{\psi} - \psi)^2|\Psi\rangle = \varphi^2 \quad , \quad \langle\Psi|(\hat{\pi}_\psi - \pi_\psi)^2|\Psi\rangle = \pi_\varphi^2 + \frac{1}{4\varphi^2} \qquad (3.4)$$

where operators are defined with respect to the dependence of Ψ on ψ'. Moreover, Ψ obeys the conditions

$$i\langle\Psi|\partial/\partial\psi|\Psi\rangle = \pi_\psi \quad , \quad i\langle\Psi|\partial/\partial\varphi|\Psi\rangle = \pi_\varphi \qquad (3.5)$$

$$\langle\Psi|\partial/\partial\pi_\psi|\Psi\rangle = 0 \quad , \quad \langle\Psi|\partial/\partial\pi_\varphi|\Psi\rangle = 0 \,. \qquad (3.6)$$

The equations of motion for the variational parameters, ψ, φ, π_ψ and π_φ, are given by the variation of the action

$$\begin{aligned} S &= \int \mathrm{d}t \left\langle \Psi \left| \left(i\partial_t - \hat{H}\right) \right| \Psi \right\rangle \\ &= \int \mathrm{d}t \left(i\dot{\psi} \langle\Psi|\partial/\partial\psi|\Psi\rangle + i\dot{\varphi} \langle\Psi|\partial/\partial\varphi|\Psi\rangle - \langle\Psi|\hat{H}|\Psi\rangle \right) \qquad (3.7) \end{aligned}$$

using the chain rule. The identities obeyed by Ψ therefore allow us to write the action in canonical form,

$$S = \int \mathrm{d}t \left(\dot{\psi}\pi_\psi + \dot{\varphi}\pi_\varphi - H_\mathrm{G} \right) \qquad (3.8)$$

where we defined the Gaussian Hamiltonian $H_\mathrm{G} = \langle\Psi|\hat{H}|\Psi\rangle$. The variation of this action gives Hamilton's equations

$$\dot{\psi} = \frac{\partial H_\mathrm{G}}{\partial \pi_\psi} \quad , \quad \dot{\pi}_\psi = -\frac{\partial H_\mathrm{G}}{\partial \psi} \quad , \quad \dot{\varphi} = \frac{\partial H_\mathrm{G}}{\partial \pi_\varphi} \quad , \quad \dot{\pi}_\varphi = -\frac{\partial H_\mathrm{G}}{\partial \varphi} \,. \qquad (3.9)$$

For example, if we consider the Hamilton operator

$$\hat{H} = \frac{1}{2}\hat{\pi}_\psi^2 + M^4 \left(1 - 2\frac{\hat{\psi}^2}{v^2} + \frac{\hat{\psi}^4}{v^4}\right) \tag{3.10}$$

with the Higgs-like potential, the Gaussian Hamiltonian is

$$H_{\mathrm{G}} = \frac{1}{2}\pi_\psi^2 + \frac{1}{2}\pi_\varphi^2 + \frac{1}{8\varphi^2} + M^4 \left(1 - 2\frac{\psi^2}{v^2} + \frac{\psi^4}{v^4} + 6\frac{\psi^2\varphi^2}{v^4} - 2\frac{\varphi^2}{v^2} + 3\frac{\varphi^4}{v^4}\right) . \tag{3.11}$$

3.2.2 Canonical effective methods

While the Gaussian approximation is useful in a wide range of applications a more general class of states is relevant for our application to inflation where non-Gaussianities should be included in the analysis. Canonical effective methods [42, 43] provide a good alternative because they allow for generally non-Gaussian states while still retaining the canonical structure that makes Gaussian states attractive. Importantly, it is not required to find a specific representation of non-Gaussian states as wave functions, which would be much more involved than (3.2). Instead, one can formulate states of a quantum system in terms of expectation values and moments assigned by a generic state to the basic operators $\hat{\psi}$ and $\hat{\pi}_\psi$. The evolution of a state is then formulated as a dynamical system for the basic expectation values $\psi = \langle\hat{\psi}\rangle$ and $\pi_\psi = \langle\hat{\pi}_\psi\rangle$ as well as the moments

$$\Delta(\psi^a \pi_\psi^b) = \left\langle (\hat{\psi} - \langle\hat{\psi}\rangle)^a (\hat{\pi}_\psi - \langle\hat{\pi}_\psi\rangle)^b \right\rangle_{\mathrm{Weyl}} , \tag{3.12}$$

using Weyl (or completely symmetric) ordering in order to avoid overcounting degrees of freedom.

The basic expectation values and moments inherit a Poisson structure from the commutator,

$$\left\{\langle\hat{A}\rangle, \langle\hat{B}\rangle\right\} = \frac{1}{i\hbar}\left\langle[\hat{A}, \hat{B}]\right\rangle , \tag{3.13}$$

augmented by the Leibniz rule in an application to moments. The equations of motion for some phase space function, $F(\psi, \pi_\psi, \Delta(\cdot))$, are then given in the form of the usual Hamilton's equations,

$$\dot{F}(\psi, \pi_\psi, \Delta(\cdot)) = \{F, H_{\mathrm{Q}}\} \tag{3.14}$$

with a quantum Hamiltonian $H_{\mathrm{Q}} = \langle\hat{H}\rangle$ defined as the expectation value of the Hamilton operator $\hat{H}$ in a generic (not necessarily Gaussian) state. For a Hamiltonian of the form

$\hat{H} = \frac{1}{2}\hat{\pi}_\psi^2 + \hat{V}(\psi)$, this definition implies the quantum Hamiltonian

$$H_{\rm Q} = \langle \hat{H} \rangle = \frac{1}{2}\pi_\psi^2 + \frac{1}{2}\Delta(\pi_\psi^2) + V(\psi) + \sum_{n=2}^{\infty} \frac{1}{n!}\frac{\partial^n V}{\partial \psi^n}\Delta(\psi^n)\,. \tag{3.15}$$

The formulation of the system in terms of expectation values and moments allows for a systematic canonical analysis at the semiclassical level. Written directly for moments as coordinates on the quantum phase space, the Poisson structure, based on (3.13) together with the Leibniz rule, is rather complicated. For instance, one can see that the Poisson bracket of two moments is not constant and not linear in general [42, 83]. Using moments as coordinates on a phase space, therefore, leads to a more complicated inflationary analysis lacking a clear separation between configuration and momentum variables. It is then unclear how to determine kinetic and potential energies or a unique relationship between specific phenomena and individual degrees of freedom.

In order to make the semiclassical analysis more clear, it is preferable to choose a coordinate system on phase space that puts the Poisson bracket in canonical form as in the variables used in (3.11), but possibly extended to higher orders in moments. The Darboux theorem [84] or its extension to Poisson manifolds [85] guarantees the existence of such coordinates, but explicit constructions are in general difficult. For second-order moments, the moment phase space is 3-dimensional and can be handled more easily than in the general context. In this case, a canonical mapping has been found several times independently [76–79]. It is accomplished by the coordinate transformation

$$\Delta(\pi_\psi^2) = \pi_\phi^2 + \frac{U}{\varphi^2} \quad , \quad \Delta(\psi\pi_\psi) = \varphi\pi_\varphi \quad , \quad \Delta(\psi^2) = \varphi^2 \tag{3.16}$$

where $\{\varphi, \pi_\varphi\} = 1$. The parameter $U = \Delta(\psi^2)\Delta(\pi_\psi^2) - \Delta(\psi\pi_\psi)^2$ is a conserved quantity (or a Casimir variable of the algebra of second-order moments), restricted by Heisenberg's uncertainty relation to obey the inequality $U \geq \hbar^2/4$. Direct calculations show that the transformation (3.16) is a canonical realization of the algebra of second-order moments. At this stage, we already have a departure from the Gaussian states, because the uncertainty for a pure Gaussian equals $\hbar^2/4$, while we retain the uncertainty as a free (but bounded) parameter.

Additional non-Gaussianity parameters, relevant for inflation, are revealed by an extension of the canonical mapping to higher-order moments. Considering higher order semiclassical corrections implies more canonical degrees of freedom. (For a single classical degree of freedom, the moments up to order N form a phase space of dimension

$D = \sum_{j=2}^{N}(j+1) = \frac{1}{2}(N^2+3N-4)$.) A canonical mapping for these higher-order semiclassical degrees of freedom has only recently been derived in [81, 82] up to the fourth order. For the relevant moments, the results are

$$\Delta(\pi_\psi^2) = \sum_{i=1}^{5} \pi_{\varphi_i}^2 + \sum_{i>j} \frac{U}{(\varphi_i - \varphi_j)^2} \tag{3.17}$$

$$\Delta(\psi^2) = \sum_{i=1}^{5} \varphi_i^2 \tag{3.18}$$

$$\Delta(\psi^3) = C \sum_{i=1}^{5} \varphi_i^3 \tag{3.19}$$

$$\Delta(\psi^4) = C^2 \sum_{i=1}^{5} \varphi_i^4 + \sum \varphi_i^2 \varphi_j^2 \tag{3.20}$$

while all other moments up to fourth order can be derived from the relevant ones using suitable Poisson brackets. There are now five canonical pairs, $(\varphi_i, \pi_{\varphi_i})$ and two Casimir variables, U and C, forming a 12-dimensional phase space of moments.

In order to parametrize the entire fourth-order semiclassical phase space we had to introduce a total of five pairs of canonical degrees of freedom and two Casimir variables, U and C. In principle, we could consider all ten non-constant semiclassical degrees of freedom, but in order to keep the analysis simple, we take inspiration from some more terrestrial applications [82, 86, 87] and choose a moment closure, thereby approximating higher-order moments in terms of lower-order ones. In particular, we choose $\Delta(\pi_\psi^2) = \pi_\varphi^2 + U/\varphi^2$, $\Delta(\psi^2) = \varphi^2$, $\Delta(\psi^3) = a_3$ (or, alternatively, $a_3\varphi^3$) and $\Delta(\psi^4) = a_4\varphi^4$. This closure corresponds to (3.17) written in higher dimensional spherical coordinates with the assumption that the angular momenta are small enough to be ignored. The parameter values $U = \hbar^2/4$, $a_3 = 0$ and $a_4 = 3$ correspond to the Gaussian case. We can therefore think of this closure as describing the non-Gaussianities by three parameters, U, a_3, and $\delta = a_4 - 3$, while maintaining the same number of degrees of freedom as in the Gaussian case.

Considering a Higgs-inspired matter field coupled to a classical and isotropic spacetime background with spatial metric $h_{ij} = a(t)^2\delta_{ij}$ in terms of proper time t, the standard Lagrangian

$$L = \int \mathrm{d}^3x \sqrt{\det h} \left(\frac{1}{2}\dot{\psi}^2 - \frac{1}{2}h^{ij}\partial_i\psi\partial_j\psi - V(\psi) \right) \tag{3.21}$$

is first reduced to homogeneous form by assuming spatially constant ψ and integrating:

$$L_{\rm hom} = \frac{1}{2}a(t)^3 V_0 \dot{\psi}^2 - a(t)^3 V_0 V(\psi)\,. \tag{3.22}$$

The new parameter V_0, defined as the coordinate volume of the spatial region in which inflation takes place, does not have physical implications but merely ensures that the combination $a(t)^3 V_0$ represents the spatial volume in a coordinate-independent way. (The value of $a(t)^3 V_0$ would be determined by the maximum length scale on which approximate homogeneity may be assumed in the early universe just before inflation [88, 89].) This Lagrangian implies the scalar momentum

$$\pi_\psi = \frac{\partial L_{\rm hom}}{\partial \dot{\psi}} = a(t)^3 V_0 \dot{\psi} \tag{3.23}$$

such that the Hamiltonian is given by

$$H = \frac{1}{2a(t)^3 V_0}\pi_\psi^2 + a(t)^3 V_0 V(\psi)\,. \tag{3.24}$$

Quantizing the scalar field, using our explicit potential (3.1), the Hamilton operator is

$$\hat{H} = \frac{1}{2a(t)^3 V_0}\hat{\pi}_\psi^2 + a(t)^3 V_0 M^4 \left(1 - \frac{\hat{\psi}^2}{v^2}\right)^2 , \tag{3.25}$$

keeping the background scale factor $a(t)$ classical. The closure we choose here implies the reduced version

$$\begin{aligned} H_{\rm Q}^{\rm closure} \;=\; & \frac{1}{2a(t)^3 V_0}\pi_\psi^2 + \frac{1}{2a(t)^3 V_0}\pi_\varphi^2 + \frac{U}{2a(t)^3 V_0 \varphi^2} \\ & + a(t)^3 V_0 M^4 \left(1 + \left(\frac{6\varphi^2}{v^4} - \frac{2}{v^2}\right)\psi^2 + \frac{\psi^4}{v^4} - 2\frac{\varphi^2}{v^2} + \frac{a_4\varphi^4}{v^4} + 4\frac{a_3\psi}{v^4}\right) \end{aligned} \tag{3.26}$$

of the quantum Hamiltonian. Hamilton's equations generated by $H_{\rm Q}^{\rm closure}$ are, as usual, deterministic, even though here they contain variables representing quantum fluctuations and higher moments. This dynamics presents an approximation of the deterministic evolution of a wave function that is implicitly determined by the momemts. We therefore do not include stochastic effects of fluctuations that would be present if the inflaton were somehow measured while inflation is still going on.

While parameterizing some higher moments through a moment closure is required for a tractable model, keeping at least one quantum degree of freedom, φ, independent is crucial for a description of non-adiabatic phases. In this way, our quantum Hamiltonian goes beyond effective potentials of low-energy type, in particular, the Coleman–Weinberg potential [90]. As shown in [91], it is possible to derive the Coleman–Weinberg potential from a field-theory version of (3.26) if one minimizes the Hamiltonian with respect to φ. This step eliminates all independent quantum degrees of freedom and, in the traditional treatment, is equivalent to ignoring non-adiabatic effects by using a low-order truncation of the derivative expansion, in addition to the semiclassical expansion also applied here. In this sense, by including the new variable φ as an authentic degree of freedom we retain non-adiabatic information of our dynamics.

In our cosmological scenario, this degree of freedom will be relevant at the beginning and end of inflation. Since the long, intermediate phase of slow-roll inflation remains by necessity adiabatic, a traditional low-energy effective action or a derivative expansion of a quantum field theory for the inflaton may be applied. As shown in [42–44], the background contribution of such an effective theory [92] is equivalent to an adiabatic approximation applied to moment corrections in a quantum Hamiltonian. All relevant phases are therefore included in our formalism.

The effective Hamiltonian (3.26) is very similar to the Gaussian Hamiltonian (3.11), which also retains an independent quantum variable, but it is more general because of the presence of the new parameters U, a_3 and a_4. As shall be shown later, the characteristics of our inflationary phase depend crucially on these parameters. In particular for a Gaussian state, inflation never ends, but if we consider small non-Gaussianities parametrized by U, a_3 and a_4, we can obtain a phenomenologically viable inflationary phase. Moreover, these parameters are determined by the quantum state of the early universe, and so constraining them with data would shed light on the character of the quantum state of the early universe.

3.3 Two-field model

After our transformation to canonical moment variables, we can uniquely extract an effective potential from (3.26),

$$\begin{aligned}\frac{1}{M^4}V_{\rm eff}(\psi,\varphi) &= 1+\frac{U}{2M^4a^6V_0^2\varphi^2}+\left(6\frac{\varphi^2}{v^4}-\frac{2}{v^2}\right)\psi^2+\frac{\psi^4}{v^4}-2\frac{\varphi^2}{v^2}+\frac{4a_3\psi}{v^4}+a_4\frac{\varphi^4}{v^4}\\ &\approx 1+2\left(\frac{\varphi^2-\varphi_c^2}{\varphi_c^2}\right)\frac{\psi^2}{v^2}+\frac{4a_3\,\psi}{v^4}+\frac{\psi^4}{v^4}-\frac{2}{3}\frac{\varphi^2}{\varphi_c^2}+a_4\frac{\varphi^4}{v^4}\,, \qquad (3.27)\end{aligned}$$

where $\varphi_c^2 \equiv v^2/3$. By construction, the second field, φ, represents the quantum fluctuation associated with the classical field ψ. As explained earlier, the additional parameters, U, a_3 and a_4 describe a possibly non-Gaussian quantum state of the background inflaton.

3.3.1 Initial conditions and the trans-Planckian problem

In the second line of the equation, we ignored the U-term $U/(2M^4 a^6 V_0^2 \varphi^2)$ in an approximation valid for sufficiently large scale factors (or, rather, averaging volumes $a^3 V_0$). The origin of this term is purely quantum and represents a potential barrier that enforces Heisenberg's uncertainty relation for the fluctuation variable φ. This term can be easily ignored after a few e-folds of inflation, but at early times its presence necessitates φ to start out with non-zero values. The subsequent non-adiabatic phase will be crucial for our model, and therefore this term alleviates our need to fine-tune the initial condition for φ.

The main effect of this repulsive term in the potential is to push out φ to large values to begin with, after which we are always able to neglect it throughout inflation. The initial φ obtained in this way is indeed consistent with requirements on inflation models. In particular, we can easily obtain the initial condition $\varphi > \varphi_c$ of hybrid inflation [93]: We expect the initial φ to be large and can therefore restrict the effective potential (3.27) to the term quartic in φ, together with the U-term relevant at early times. This restricted potential has a local minimum at

$$\varphi = \sqrt[6]{\frac{Uv^4}{4a^6 V_0^2 M^4 a_4}} \,. \tag{3.28}$$

We do not know much about the volume $a^3 V_0$ of the initial spatial region that was meant to expand in an inflationary way. But in order to avoid the trans-Planckian problem [94–96], we should require that $a^3 V_0 > \ell_{\rm P}^3$. This lower bound implies the upper bound

$$\varphi_{\rm ini} < \frac{1}{\ell_{\rm P}} \sqrt[6]{\frac{Uv^4}{4a_4 M^4}} \tag{3.29}$$

for (3.28). For parameters of the order $v \sim \mathcal{O}(M_{\rm P})$ and $M^4 \ll M_{\rm P}^4$, as common in hybrid models and used in our analysis to follow, the upper bound on $\varphi_{\rm ini}$ is much greater than φ_c.

3.3.2 Waterfall: Phase transitions

Our effective potential (3.27), depending on the classical field ψ and its fluctuation, φ is of the hybrid-inflation type. These models typically produce a blue-shifted tilt when one starts with a large φ and small ψ [93]. Inflation in this scenario essentially relies on the near-constant vacuum energy of ψ. However, there is an alternative scenario in the same model, the so-called waterfall regime [97, 98], realized at a later stage in our model in which φ has moved to and stays close to a minimum while ψ gradually inches away from its vacuum value that has by then become an unstable equilibrium position.

As we will show, initial conditions for the waterfall regime to take place are generated in our extension of the model by a non-adiabatic phase in which φ is still large. The subsequent waterfall regime then generates a significant number of e-folds and leads to a red-shifted tilt for a wide range of parameters. For this scenario to take place, it is important that our effective potential differs from the traditional hybrid one in that we have an $a_4\varphi^4$ term as well as a Z_2-breaking term $a_3\psi$, which is assumed to be small but not exactly zero. The symmetry breaking term is parameterized by a_3, which represents background non-Gaussianity. In a perfect vacuum, such a term would vanish. In a less fine-tuned state, this term relieves us of the burden of supplying a non-zero initial value for ψ, which is required to start the dynamics of the waterfall regime, as we shall demonstrate later. Because both new terms depend on state parameters in our semiclassical approximation, the resulting description of inflation is characterized by an intimate link between observational features and properties of quantum states.

Another difference with the traditional hybrid model is that the hierarchy between our set of parameters is more rigid, leaving less room for tuning and ambiguity and making our results more robust. The traditional potential has three parameters that can be adjusted independently, while in our case only two (non-state) parameters are independent. This is so because we do not have a generic two-field model but rather a single-field model accentuated by its quantum fluctuation. As opposed to the traditional hybrid model [97], we have two phase transitions characterized by non-adiabatic behavior, and the majority of e-folds are created in between.

As in the original hybrid model, we start with some $\varphi > \varphi_c$ with φ quickly rolling down to its minima under an effective φ^4 term. This phase is driven by a simplified potential of the form

$$\frac{V_{\text{eff}}^{\varphi}}{M^4} = 1 - \frac{2}{3}\frac{\varphi^2}{\varphi_c^2} + a_4\frac{\varphi^4}{v^4} \tag{3.30}$$

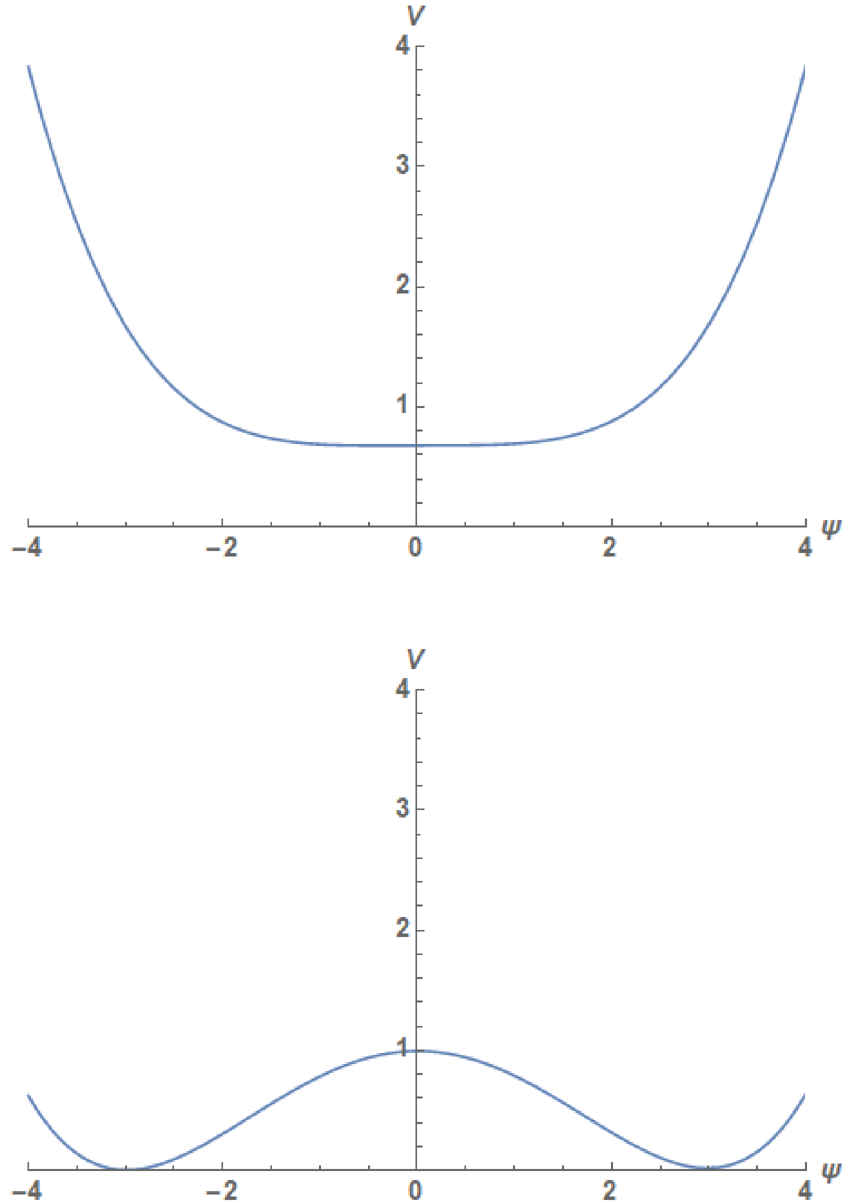

Figure 3.1. Shape of the potential $V(\psi)$ for constant φ at early (top) and late times (bottom), defined relative to the time when φ crosses φ_c.

since ψ sits in its local minimum at the origin during this time and therefore all ψ-terms can be ignored. Once φ crosses φ_c, the new true minima of ψ are displaced from the origin due to a tachyonic term in its effective potential, of the form

$$\frac{V_{\rm eff}^{\psi}}{M^4} = 1 + 2\left(\frac{\varphi^2 - \varphi_c^2}{\varphi_c^2}\right)\frac{\psi^2}{v^2} + \frac{4a_3\psi}{v^4} + \frac{\psi^4}{v^4} - \frac{2}{3}\frac{\varphi^2}{\varphi_c^2} + a_4\frac{\varphi^4}{v^4}\,. \tag{3.31}$$

Due to the a_3 term, the Z_2 symmetry of ψ is broken and the field starts slowly

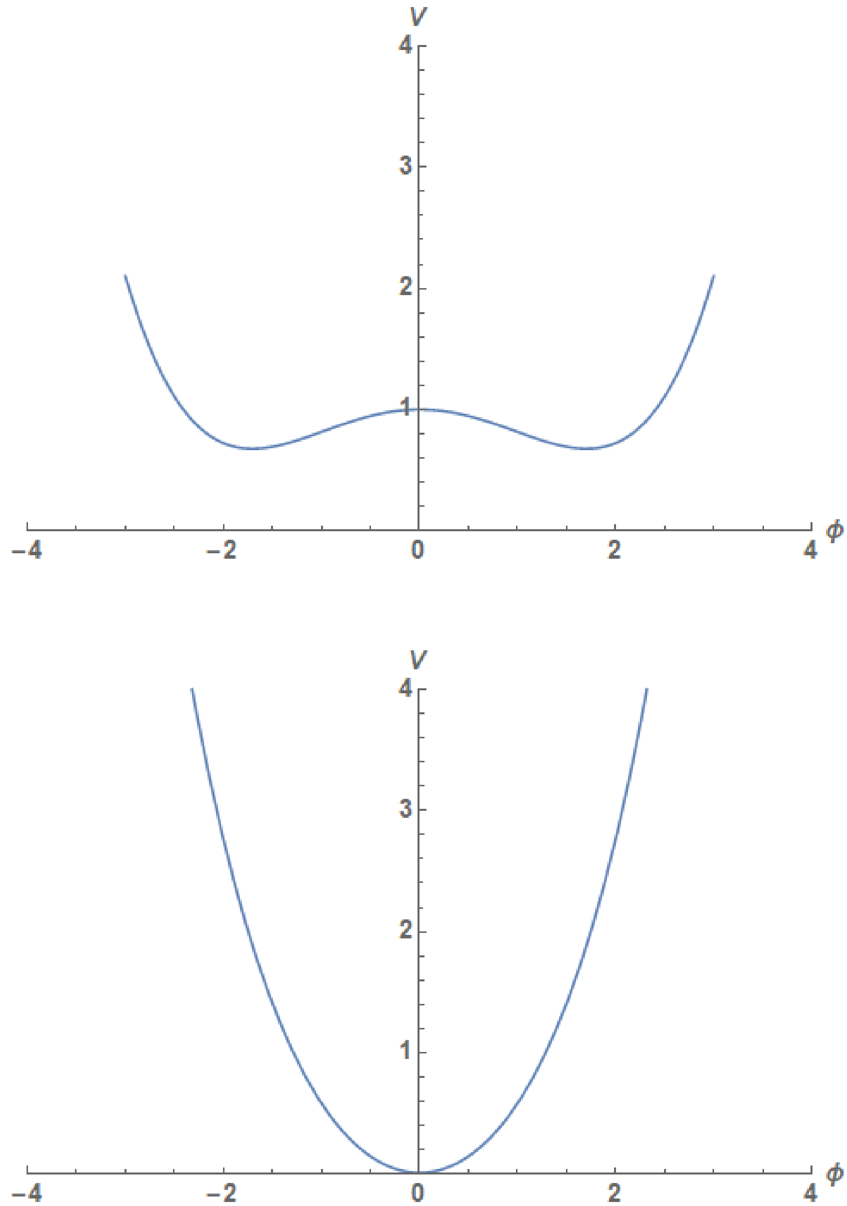

Figure 3.2. Shape of the potential $V(\varphi)$ for constant ψ at early (top) and late times (bottom), defined relative to the time when φ crosses φ_c. Here, we are ignoring the contribution from the uncertainty principle that prevents φ from crossing $\varphi = 0$.

rolling away from the origin. This gradual change enables φ to closely follow its vacuum expectation value, φ_*. (Its gradual nature also means that the back-reaction of homogeneous fluctuations φ on the homogeneous expectation value ψ is small, justifying our semiclassical approximation. The combined system of ψ and φ has a pronounced effect on the background space-time, driving its expansion. However, since energy densities always remain sub-Planckian, our model is semiclassical also in the sense of quantum

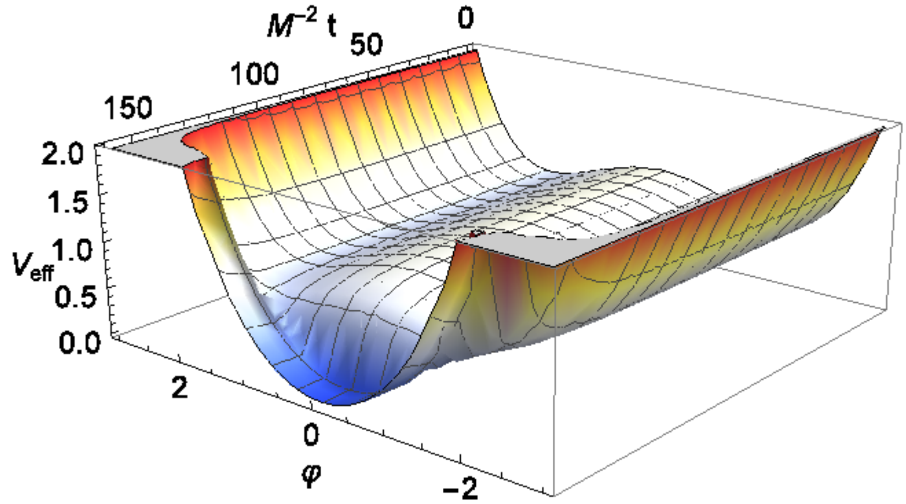

Figure 3.3. After a brief non-adiabatic phase when it rolls down a steep potential wall, φ traces its minimum for the majority of inflation. The growth of ψ^2 moves the φ-minima closer to zero, causing another non-adiabatic phase that ends with an approximate symmetry restoration for φ. The parameters used are $v = 3$, $a_3 = 0.05$ and $\delta = 0.1$.

gravity and we are justified in keeping the scale factor a unquantized.) Eventually, φ_* approaches zero but never reaches it due to the uncertainty principle, thereby almost restoring the symmetry for φ; this is the second phase transition mentioned above. As shown in Figs. 3.1 and 3.2, φ causes the traditional phase transition when it crosses φ_c, and then the slow roll of ψ down its tachyonic hilltop will end in a second phase transition. The whole process is clarified further by examining how the effective potential changes in time, shown in Figs. 3.3 and 3.4.

The hilltop phase generates the dominant number of e-folds, and it ends automatically when ψ reaches its new minimum. This is a new feature compared to the traditional hybrid inflation and relies on the existence of a φ^4 term in our effective potential. Our model is not a variant of the original hybrid model [99], such as the inverted-hybrid model [100] or a modified hilltop model [101], or having corrections to the potential coming from supergravity-embedding of the model [102]; rather, we start with a Higgs-like model and include effects from an initial quantum state that turn it into a hybrid model with some additional terms.

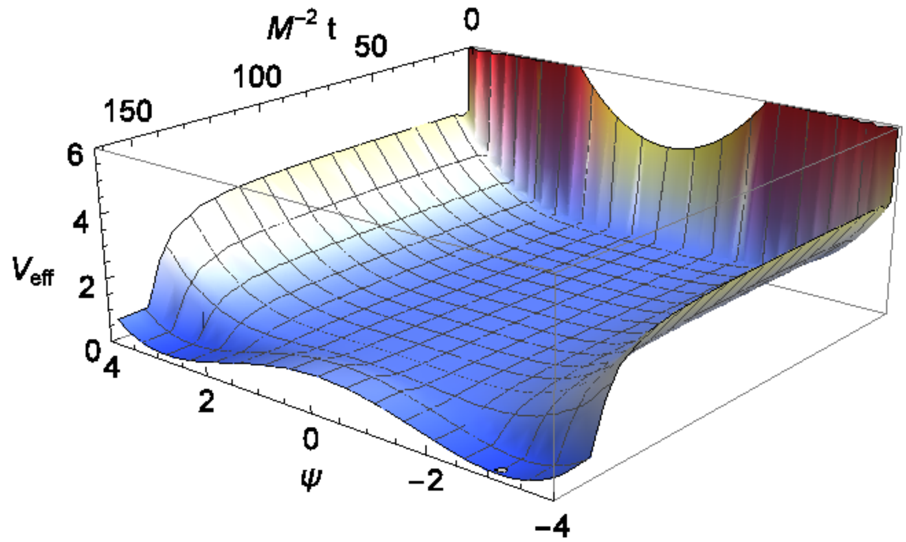

Figure 3.4. During the initial non-adiabatic phase, a phase transition akin to traditional hybrid models occurs. Reflection symmetry in the potential is slightly broken by the a_3-term (which is not apparent in the figure due to its smallness). This non-Gaussianity term drives ψ to its new stable point where ψ^2 approaches v^2. The parameters are the same as Fig. 3.3.

3.3.3 UV-completion and the swampland

One of the conceptual requirements for inflation models is that they should have a well-defined quantum completion. One way to implement this is to derive specific forms of inflationary potentials from string theory constructions as was done, for instance, in the case of natural inflation. Another recent idea has been that of the swampland, a complement of the string landscape, which stems from the fact that not all low-energy effective field theories can be consistently completed in the ultraviolet into a quantum theory of gravity [103,104]. In order for an effective field theory to be consistent, it would have to satisfy the eponymous swampland constraints. This is a much more general way in which quantum gravity may restrict the form of the potential, amongst other things, in the low-energy effective field theory used as the starting point for inflation. More specifically, it has been argued that many models of (at least) single-field inflation are not consistent with the swampland conjectures since the latter require either a large value for the slope of the potential, $|V'|/V > \mathcal{O}(1)$, or large tachyonic directions, $V''/V < -\mathcal{O}(1)$ [105, 106].

Taken together, these conjectures severely restrict the lifetime of metastable (quasi-)de Sitter spacetimes that can be built from string theory. In order to obtain an estimate for the numbers of order one that appear in one of them, the so-called de-Sitter conjecture, one has to resort to fundamental properties of quantum gravity such as the absence of eternal inflation [107,108] or the trans-Planckian censorship conjecture [109,110]. The latter has put a more concrete bound on the duration of inflation which, when combined with the observed power spectrum, imposes severe constraints on the allowed models for inflation. It has been shown that only hilltop type of models, which generically allow for a small slow-roll parameter ϵ but a big η, are the ones that survive amongst all single-field models unless one invokes additional degrees of freedom as in non-Bunch Davies initial states or warm inflation. Even for hilltop potentials, which seem to be the most compatible with the swampland, one has to resort to an arbitrary steepening of the potential to end inflation so as not to have too many e-folds since that would once again make the model incompatible with the constraints. To date, there are no string theory realizations of any such single-field potential that can abruptly stop inflation after a finite amount of time.

The remarkable feature of our new model is that it is able to give a viable inflationary cosmology as well as a graceful exit with a tachyonic (p)reheating, all starting from a Higgs-like single-field potential as the main input. We are using only standard quantum mechanics in a non-adiabatic semiclassical approximation and do not have to rely on unknown features of quantum gravity. In addition, by virtue of the fact that the classical field ψ plays the role of the inflaton relevant for observable scales, this model is essentially of the hilltop type which has recently been shown to be preferred by the swampland and to be able to ameliorate the η-problem [111]. Quantum effects imply that the single-field classical potential is, upon quantization, no longer a single-field model that would have to be tuned in order to avoid having too many e-folds of inflation or require any additional mechanism to achieve stability against radiative corrections [112]. Moreover, our detailed derivations below reveal that the model maintains a large value of the slow-roll parameter η throughout inflation (in addition to a small ϵ, as is usually the case for a prototype hilltop model). Indeed, it is when the value of η becomes too large that inflation ends in this model, once again thanks to effects of quantum fluctuations of the classical field (as opposed to a generic second field). All of this is possible even though we start with a single-field model with a monomial potential, but then take into account the effects of quantum fluctuations in a systematical manner.

3.4 Analysis

The effective Hamiltonian (3.26) describes a two-field model with standard kinetic terms in an expanding universe and an interaction potential similar to hybrid models. A numerical analysis can be applied directly to Hamilton's equations for ψ and ϕ generated by $H_{\mathrm{Q}}^{\mathrm{closure}}$, (3.26), using suitable initial values. We will present such solutions in comparison with a slow-roll approximation to be developed first.

3.4.1 Slow-roll approximation

For inflationary applications of (3.26), we are interested in a long phase of slow roll that can be generated by ψ staying near its initially stable and then metastable equilibrium position at $\psi = 0$. As long as $\psi^2 \ll v^2$ and $\varphi^2 \approx \varphi_*^2$ is near a local minimum, the slow-roll approximation can be used and evaluated analytically. This phase is adiabatic and therefore does not require all terms in (3.26) that are implied by semiclassical methods for non-adiabatic quantum dynamics. However, as we have already seen, the remaining terms are essential in achieving suitable initial values for the slow-roll phase and to end it early enough. Throughout this analysis, we will also assume small background non-Gaussianity. As our results will show, this assumption is justified by observational constraints on the spectral index.

Given these conditions, the slow-roll parameters can be approximated as

$$\epsilon_\varphi \equiv \frac{1}{2}M_{\mathrm{P}}^2\left(\frac{V_\varphi}{V}\right)^2 \approx \frac{1}{2}M_{\mathrm{P}}^2\left(\frac{M^4}{P}\right)^2\left(\frac{4\varphi}{3\varphi_c^2}\left(1-\frac{3\psi^2}{v^2}\right)-\frac{4a_4\varphi^3}{v^4}\right)^2 \quad (3.32)$$

$$\epsilon_\psi \equiv \frac{1}{2}M_{\mathrm{P}}^2\left(\frac{V_\psi}{V}\right)^2 \approx \frac{1}{2}M_{\mathrm{P}}^2\left(\frac{M^4}{P}\right)^2\left(\frac{4\psi}{v^2}\left(\frac{\varphi^2}{\varphi_c^2}+\frac{\psi^2}{v^2}-1\right)+\frac{4a_3}{v^4}\right)^2 \quad (3.33)$$

$$\eta_{\varphi\varphi} \equiv M_{\mathrm{P}}^2\frac{V_{\varphi\varphi}}{V} = -\frac{M^4}{P}\left(\frac{4}{3\varphi_c^2}\left(1-\frac{3\psi^2}{v^2}\right)-\frac{12a_4\varphi^2}{v^2}\right) \quad (3.34)$$

$$\eta_{\psi\psi} \equiv M_{\mathrm{P}}^2\frac{V_{\psi\psi}}{V} = \frac{M^4}{P}\frac{4}{v^2}\left(\frac{\varphi^2-\varphi_c^2}{\varphi_c^2}+\frac{3\psi^2}{v^2}\right) \quad (3.35)$$

$$\eta_{\psi\varphi} \equiv M_{\mathrm{P}}^2\frac{V_{\varphi\psi}}{V} = \frac{M^4}{P}\frac{8\psi\varphi}{v^2\varphi_c^2}\,, \quad (3.36)$$

where $V_\varphi = \partial V/\partial\varphi$ and $V_\psi = \partial V/\partial\psi$, iterated for higher derivatives. The constant P is the initial potential energy, evaluated when $\varphi \approx \varphi_c$ and $\psi \approx 0$. In the following we set $M_{\mathrm{P}} = 1$. We will see later that small background non-Gaussianity ensures that $\varphi^2/\varphi_c^2 - 1 \ll 1$. Along with the adiabatic approximation for φ, this inequality can ensure that ϵ_ψ and $\eta_{\psi\psi}$ are very small. However $\eta_{\varphi\varphi}$ is not necessarily small, even though $\ddot{\varphi} \ll 3H\dot{\varphi}$ and $\dot{\varphi}^2 \ll V$.

Our equations of motion, under slow roll, then read

$$\frac{3H\dot{\varphi}}{M^4} = \frac{4\varphi}{3\varphi_c^2}\left(1 - \frac{3\psi^2}{v^2}\right) - \frac{4a_4\varphi^3}{v^4} \tag{3.37}$$

$$\frac{3H\dot{\psi}}{M^4} = -\frac{4\psi}{v^2}\left(\frac{\varphi^2 - \varphi_c^2}{\varphi_c^2} + \frac{\psi^2}{v^2}\right) - \frac{4a_3}{v^4}\,. \tag{3.38}$$

where we can make M implicit by rescaling $t \to t/M^2$. The regime covered by our approximations can be split into two phases followed by an end phase.

3.4.1.1 Phase 1

In early stages, we have $\psi^2 \ll v^2$ and can thus ignore the term $3\psi^2/v^2$ in (3.37). Therefore, the constant $\varphi^2 \approx \varphi_*^2 \approx 3\varphi_c^2/a_4$ is an approximate solution. Adiabaticity ensures that we can expand the equation of motion around the critical point φ_* where $V_\varphi(\varphi_*) = 0$:

$$\dot{\varphi} \approx -\frac{1}{3H}V_{\varphi\varphi}(\varphi_*)(\varphi - \varphi_*)\,. \tag{3.39}$$

Defining $\varphi' := \mathrm{d}\varphi/\mathrm{d}N$ where N is the number of e-folds, we obtain

$$\varphi' \approx -\eta_{\varphi\varphi}(\varphi = \varphi_*, \psi \approx 0)(\varphi - \varphi_*)\,. \tag{3.40}$$

For small background non-Gaussianity, we have $a_4 = 3 + \delta$ with $\delta \ll 1$. Choosing the initial value $\varphi(0) = \varphi_c$ for Phase 1 therefore implies

$$\varphi_1(N) \approx \frac{\varphi_c\delta}{2a_4}\exp(-\eta_{\varphi\varphi}(\varphi_*, 0)N) + \varphi_*\,. \tag{3.41}$$

Note that small non-Gaussianity also implies $\varphi_*^2 = \varphi_c^2 + O(\delta) + O(\psi^2)$.

We can expect $\varphi^2/\varphi_c^2 - 1 \approx -\delta/a_4$ to be much bigger than ψ^2/v^2 at early times. This reduces the second equation of motion, (3.38), to

$$\psi' \approx \frac{1}{P}\frac{4}{v^2}\left(\frac{\delta}{a_4}\psi - \frac{a_3}{v^2}\right) \tag{3.42}$$

which is solved by

$$\psi_1(N) \approx -\frac{a_3a_4}{\delta v^2}\left(\exp\left(\frac{4\delta}{v^2a_4P}N\right) - 1\right) \tag{3.43}$$

for an initial ψ_1 at the origin. To summarize, Phase 1 is characterized mathematically by the possibility to ignore the ψ^2/v^2 terms in (3.37) and (3.38).

3.4.1.2 Phase 2

As ψ moves away from its metastable position at $\psi = 0$, the terms ψ^2/v^2 in the equations of motion will eventually have noticeable effects even while they may still be small. In particular, the local minimum of φ at

$$\varphi_*(\psi(t))^2 = \frac{v^4}{3\varphi_c^2 a_4}\left(1 - \frac{3\psi(t)^2}{v^2}\right) \tag{3.44}$$

is then time-dependent. The solution for φ in Phase 2 can therefore be obtained directly from (3.41) by inserting the time-dependent ψ and φ_*,

$$\varphi_2(N) = \varphi_1(N)|_{\psi\to\psi(N)}\,, \tag{3.45}$$

using the solution for $\psi(N) \equiv \psi_2(N)$ to be derived now. As implied by adiabaticity, we still have $\varphi^2 \approx \varphi_*^2$, tracking the local minimum.

Our phase now is described by the first two terms of (3.38) dominating over the a_3-term. Therefore,

$$\begin{aligned} \psi' &\approx -\frac{1}{P}\frac{4\psi}{v^2}\left(\frac{\varphi_*(\psi(t))^2 - \varphi_c^2}{\varphi_c^2} + \frac{\psi^2}{v^2}\right) \\ &= \frac{1}{P}\frac{4\psi}{v^2}\left(\frac{\delta}{a_4} + \frac{2\psi^2}{v^2} + O(\delta\psi^2/v^2)\right)\,. \end{aligned} \tag{3.46}$$

which is solved by

$$\psi_2(N) \approx -\mathrm{sgn}(a_3)\sqrt{\frac{\delta}{(2a_4/v^2 + \delta/\psi_{\mathrm{g}}^2)\exp(-8\delta(N-N_{\mathrm{g}})/(v^2Pa_4)) - 2a_4/v^2}}\,. \tag{3.47}$$

(Although a_3 does not appear in our approximate equation (3.46), its sign determines the direction in which ψ starts moving as a consequence of reflection symmetry breaking.) Here, the subscript "g" denotes the value of solutions at the "gluing" point of the two phases, defined as the point where the cubic term in (3.38) is on the order of the a_3-term; see Fig. 3.6 below for an illustration.

3.4.1.3 End phase

Even though Phase 1 and Phase 2 are sufficient to describe the majority of inflation, finding the point at which inflation ends requires a qualitatively different approximation compared with the above two phases. The physics is also quite different. To see this, note that if we extend the approximations of Phase 2 too far, we arrive at two wrong conclusions. First, ψ will eventually cross the point $\psi^2 = v^2/3$, such that the two minima of $V_{\rm eff}(\varphi)$ meet at $\varphi_* = 0$. Second, this behavior causes φ to approach zero, such that the field ψ ends up at its new $V_{\rm eff}(\psi)$-minimum, $\psi_{\rm min} = -v$ (assuming a_3 is positive). The former ($\varphi \to 0$) is forbidden by the uncertainty principle, embodied in our U-term in $V_{\rm eff}$ neglected so far in the slow-roll analysis, and the latter is erroneous since it implies that once everything has settled, H^2, which is proportional to $V_{\rm eff}$ during slow roll, would seem to approach a negative value $4a_3\psi/v^4 < 0$.

However, this last conclusion certainly cannot be correct because our classical potential (3.1), a complete square $V_{\rm cl}(\psi) = M^4(1-\psi^2/v^2)^2$, is positive semidefinite. Therefore, it is quantized to a positive, self-adjoint operator $\hat{V}$ which cannot possibly have a negative expectation value $V_{\rm eff} = \langle \hat{V} \rangle$ in any admissible state. In terms of moments used in our canonical effective description, after ψ crosses the value $v^2/3$, the fluctuation variable φ shrinks. Therefore, according to our moment closure introduced after equation (3.17), the variance $\Delta(\psi^2) = \varphi^2$ as well as the fourth-order moment $\Delta(\psi^4) = a_4\varphi^4$ approach zero, while $\Delta(\psi^3) = a^3$ has so far been assumed constant. This latter assumption violates higher-order uncertainty relations for small φ.

We will not require a precise form of such higher-order uncertainty relations, or a specific decreasing behavior of $\Delta(\psi^3)$ because, referring to positivity, we know that the magnitude of the a_3-term in the potential is not allowed to be larger than the sum of the rest of the terms in $V_{\rm eff}$. (But see the next subsection for numerical examples with decreasing $\Delta(\psi^3)$.) This observation places an implicit bound on non-Gaussianity parameters when our potential energy decreases at the end of inflation. Taking this outcome into account, our effective potential eventually becomes

$$\frac{V_{\rm eff}}{M^4} \approx \left(1 - \frac{\psi^2}{v^2}\right)^2 + \frac{2}{3}\frac{\varphi^2}{\varphi_c^2}\left(\frac{3\psi^2}{v^2} - 1\right) + \frac{U}{2M^4a^6V_0^2\varphi^2}\,, \tag{3.48}$$

where we have neglected the φ^4 and a_3 terms for small fluctuations. The corrected values

φ_* of the two φ-minima are now

$$\varphi_* \approx \pm \left(\frac{u}{K(\psi^2)} \varphi_c^2 \right)^{1/4} \tag{3.49}$$

where

$$u = \frac{U}{M^4 a^6 V_0^2} \quad \text{and} \quad K(\psi^2) = \frac{4}{3} \left(\frac{3\psi^2}{v^2} - 1 \right) . \tag{3.50}$$

Since u is extremely small after 60 e-folds, we have $|\varphi_*| \ll 1$. The symmetry restoration for φ is therefore only an approximate one. In addition, we neglected the $O(\delta\psi^2/v^2)$-term in (3.46), but kept δ/a_4. These two terms become comparable around $\psi^2 = v^2/3$ for our chosen parameters. However, as we will see later in a comparison with numerical solutions, setting $\varphi = 0$ and using the $\psi(N)$ expression of Phase 2 during the end phase gives a sufficiently accurate number of e-folds. This also means that only a negligible amount of e-folds forms in the non-adiabatic phases of our evolution, and while non-adiabatic effects are crucial for the beginning and the end of inflation, they do not affect observations directly.

3.4.2 Comparison of analytical and numerical solutions

Our analytical solutions were obtained with certain approximations, but they generally agree well with numerical solutions of the full equations,

$$\ddot{\varphi} + 3H\dot{\varphi} = \frac{4\varphi}{3\varphi_c^2} \left(1 - \frac{3\psi^2}{v^2} \right) - \frac{4a_4\varphi^3}{v^4} \tag{3.51}$$

$$\ddot{\psi} + 3H\dot{\psi} = -\frac{4\psi}{v^2} \left(\frac{\varphi^2 - \varphi_c^2}{\varphi_c^2} + \frac{\psi^2}{v^2} \right) - \frac{4a_3}{v^4} , \tag{3.52}$$

in situations relevant for inflation. To be specific, we choose parameters $v = 3$, $\delta = 0.1$ and $a_3 = 0.05$ in our numerical solutions. Figure 3.5 shows a representative example of full numerical evolution. To test our analytical assumptions, Fig. 3.6 shows the magnitudes of individual terms that contribute to the equation of motion (3.52) for ψ, while Figs. 3.7 and 3.8 compare analytical and numerical solutions of both equations.

Cosmological parameters relevant for inflation are shown in the next figures, Fig. 3.9 for the slow-roll parameter $\eta_{\psi\psi}$ which eventually ends inflation, Fig. 3.10 for the spectral index according to both analytical and numerical solutions, as well as its running in Fig. 3.11. As shown by these figures, the parameters easily imply solutions compatible with observational constraints. It is also shown how $\eta_{\psi\psi}$ increases at an opportune time

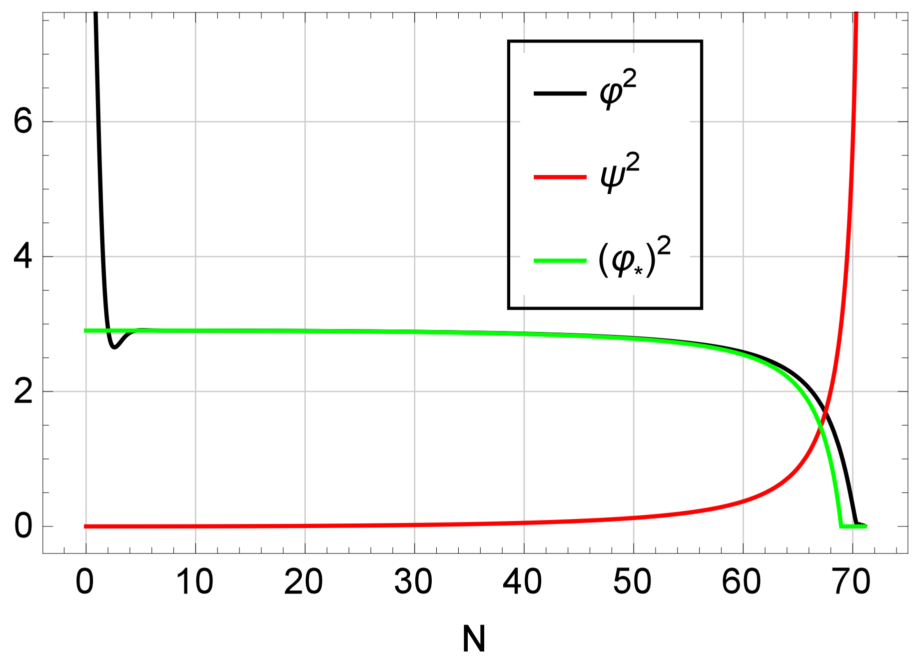

Figure 3.5. Overview of full numerical evolution. The field ψ remains small during inflation while φ follows its vacuum expectation value φ_* very closely throughout the whole evolution. After inflation ends, ψ^2 approaches v^2, a value cut off in this presentation. While the fields may take Planckian values, of the order one in natural units, except for very early times they hover near their potential minima where they imply sub-Planckian energy densities. Quantum-gravity effects are therefore negligible during inflation. The field ψ^2 increases at the end of inflation, but it merely approaches its new minimum seen in Fig. 3.4 and is not a run-away solution.

to end inflation with just the right number of e-folds in order to avoid the trans-Planckian problem.

The role of non-Gaussianity parameters can also be studied. For instance, parameterizing $a_3 = 0.01\varphi^3$ instead of a constant $a_3 = 0.05$ leads to comparable results, as shown for the number of e-folds in Fig. 3.12. The effects of different choices of $\delta = a_4 - 3$ on the spectral index and the tensor-to-scalar ratio (computed as $r \approx 16\epsilon_\sigma$, σ being the effective adiabatic field [97]) are shown in Figs. 3.13 and 3.14. An important new result is that the non-Gaussianity parameters effectively control the onset and duration of inflation, such that observationally preferred numbers of e-folds can be obtained for reasonable choices of background non-Gaussianity. In particular, only small deviations from a nearly Gaussian ground state are required.

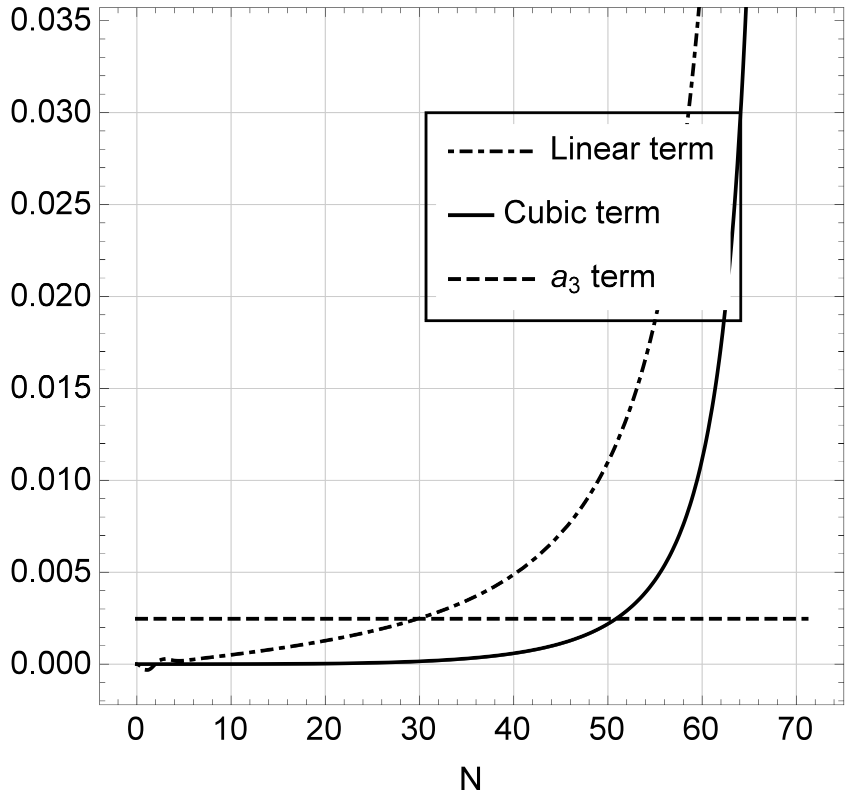

Figure 3.6. The magnitudes of individual terms in (3.38) as functions of N. The term ψ^3/v^4 in (3.38) approaches the order of a_3/v^4 around $N = 50$, marking the transition point to Phase 2.

3.4.3 Analytical results for cosmological observables

As we have argued previously, our system is essentially a two-field model, such that we may directly apply tools from multi-field inflation to predict the number of e-folds N (starting from the crossing of $\phi = \phi_c$) and the spectral index. Using our analytical solutions for the background variables, we may obtain approximate analytical expressions for the observables, which are based on perturbative inhomogeneity. The standard treatment of inflation quantizes the inflaton fields, subject to a given potential, on an expanding space-time and computes power spectra from correlation functions of inhomogeneous modes. Here, we have already used quantum properties to generate our extended 2-field potential. As a consequence, the variable φ, derived from quantum fluctuations of ψ, cannot give rise to a quantum field that could imply correlation functions to be used in a multi-field calculation of power spectra. In a field quantization of our model, there would be only one field operator, $\hat{\psi}$, rather than two quantum fields.

Nevertheless, we are able to formulate our model in a multi-field manner even for

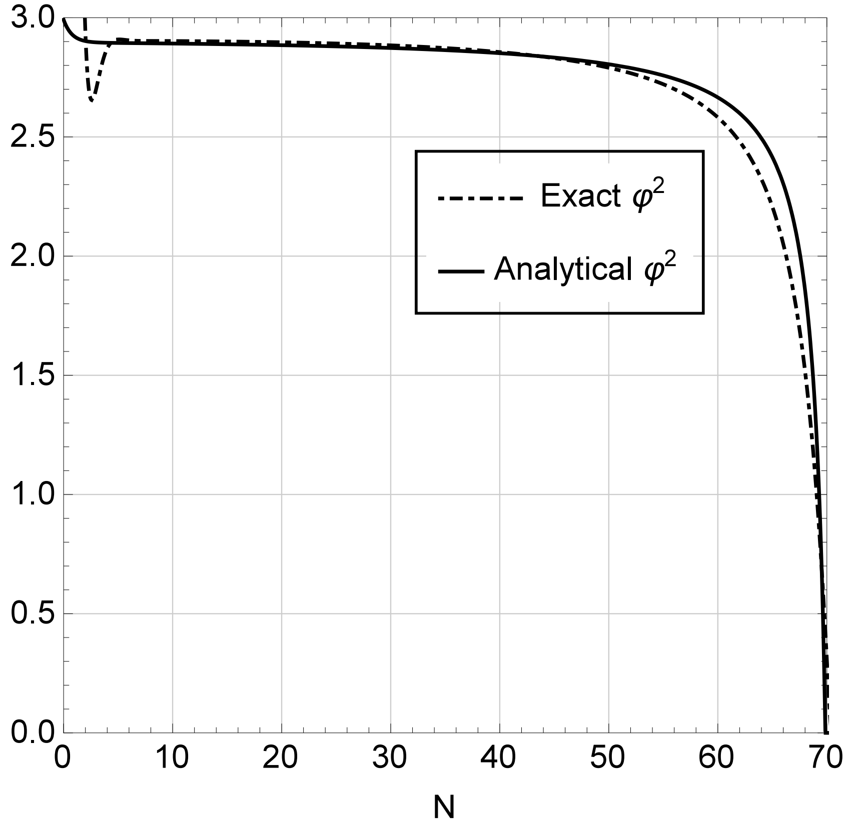

Figure 3.7. Comparison of analytical and numerical solutions for $\varphi(N)$. Our analytical solution for $\varphi(N)$ agrees well with the full numerical one, justifying the adiabatic approximation during inflation.

perturbative inhomogeneity. In our formalism, we would describe the full system of background variables and perturbative modes within the same setting of canonical effective theory. As before, such a framework would be based on moments which, now, also include the sought-after correlation functions of modes. While a complete treatment is well beyond the scope of the present paper, it is not difficult to see that the correct field degrees of freedom would be present. In particular, instead of deriving correlation functions for a quantized fluctuation field φ, which does not exist in our model, we can describe relevant correlation functions through higher moments: Standard correlation functions are quadratic expressions in modes of φ, which as a fluctuation is itself quadratic in the original field ψ. Suitable fourth-order moments of modes of the field ψ, which is associated with a quantum field, can therefore be used as correlation functions for the derived field φ. Since higher-order moments are subdominant for near-Gaussian states, as encountered here, the mode dynamics do not include terms beyond those relevant for the required correlation functions. We are therefore able to apply standard

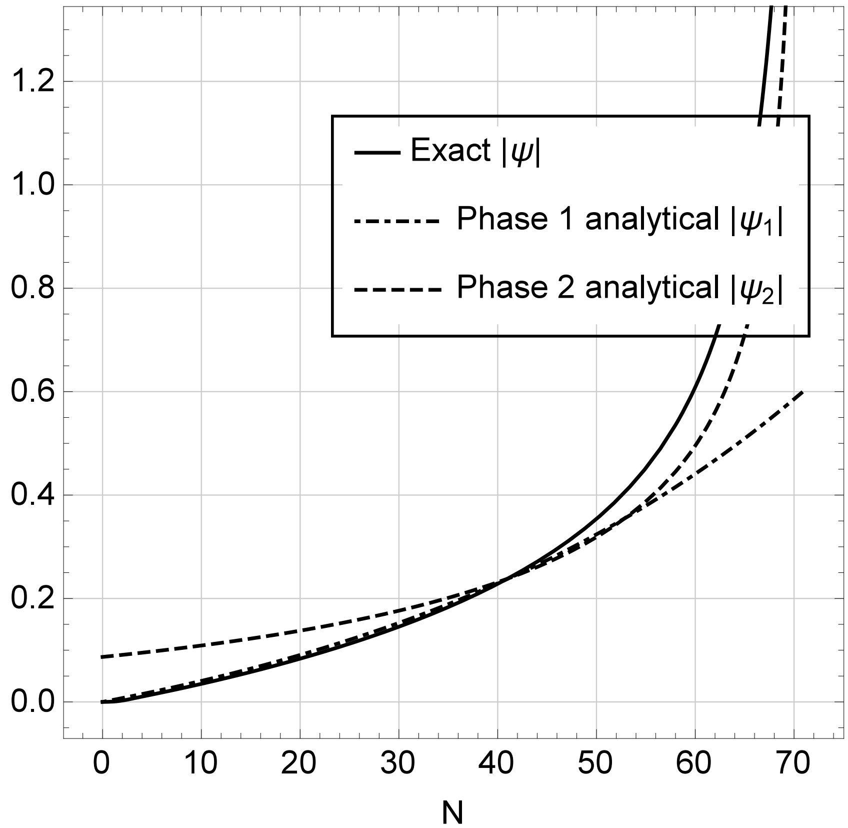

Figure 3.8. Comparison of analytical and numerical solutions for $\psi(N)$. The analytical solution agrees extremely well with the exact one in Phase 1 (before $N = 50$), while small deviations occur in ψ_2 occur Phase 2 (after about $N = 50$).

methods from multi-field inflation.

3.4.3.1 Perturbation modes

In our model, both the classical background variable ψ and its quantum fluctuation φ undergo slow-roll evolution in different phases of the dynamics. Therefore, they should both contribute to the curvature perturbation once inhomogeneous modes are included and one can write down the effective adiabatic field σ as a combination of both these fields, ψ and ϕ. (Here, we use the term "adiabatic" in its standard meaning applied to modes of perturbative inhomogeneity.)

In terms of the adiabatic field σ, consider the spectral index at around horizon exit,

$$n_s = 1 - 6\epsilon_\sigma + 2\eta_{\sigma\sigma} \,. \tag{3.53}$$

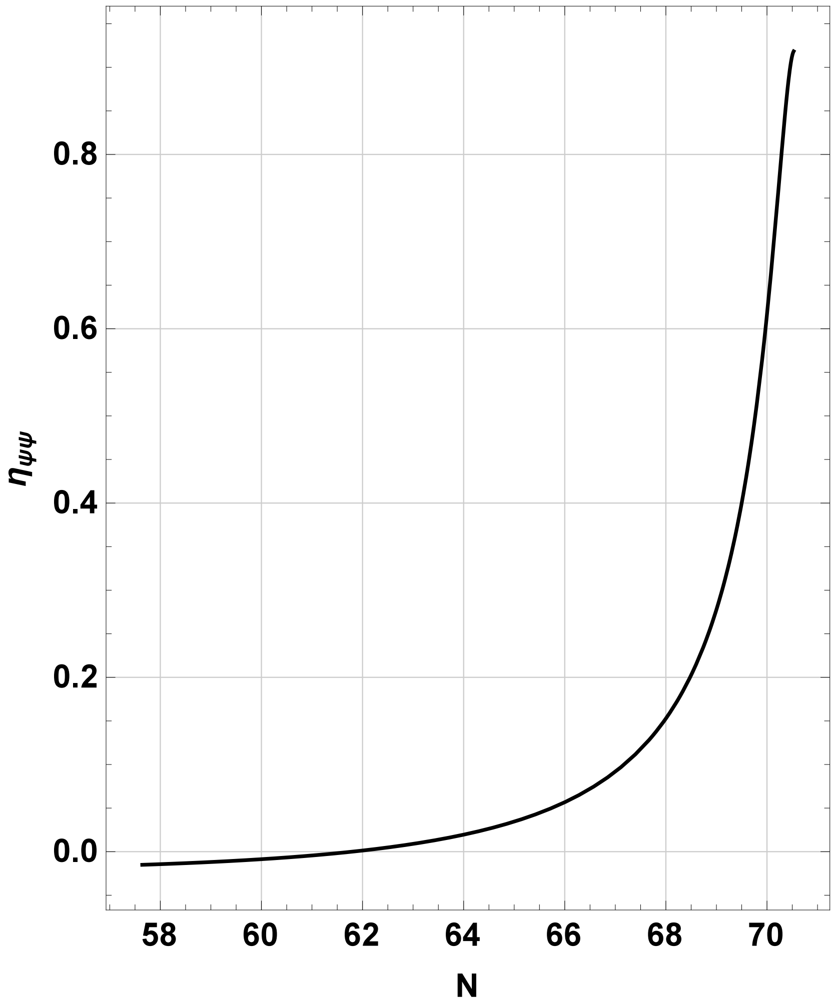

Figure 3.9. Late time behavior (Phase 2) of $\eta_{\psi\psi}(N)$ obtained from analytical solutions for $\psi(N)$ and $\varphi(N)$. The slow-roll assumption starts being violated around $N \sim 70$, effectively ending inflation.

At early times, using the slow-roll approximation for ϕ and small ψ, we have

$$\epsilon_\sigma = \epsilon_\psi + \epsilon_\varphi \approx 0 + O(\psi^2, \delta^2, a_3^2)\,. \tag{3.54}$$

For $\eta_{\sigma\sigma}$ we have [97]

$$\eta_{\sigma\sigma} \;=\; \eta_{\varphi\varphi}\cos^2\theta + \eta_{\psi\psi}\sin^2\theta + 2\eta_{\varphi\psi}\sin\theta\cos\theta \tag{3.55}$$

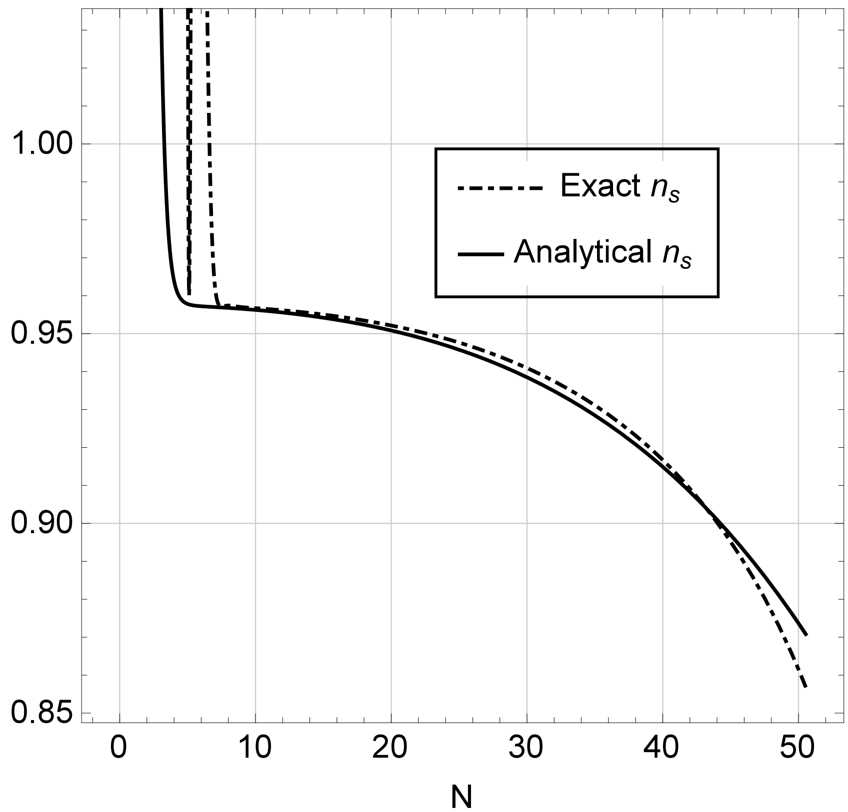

Figure 3.10. Analytical and numerical solutions for the spectral index $n_\mathrm{s}(N)$ in Phase 1. Since Hubble exit takes place at least a $\Delta N \sim 60$ prior to the end of inflation, it can only occur in Phase 1. Importantly, $n_\mathrm{s} \approx 0.96$ at $\Delta N \sim 60$.

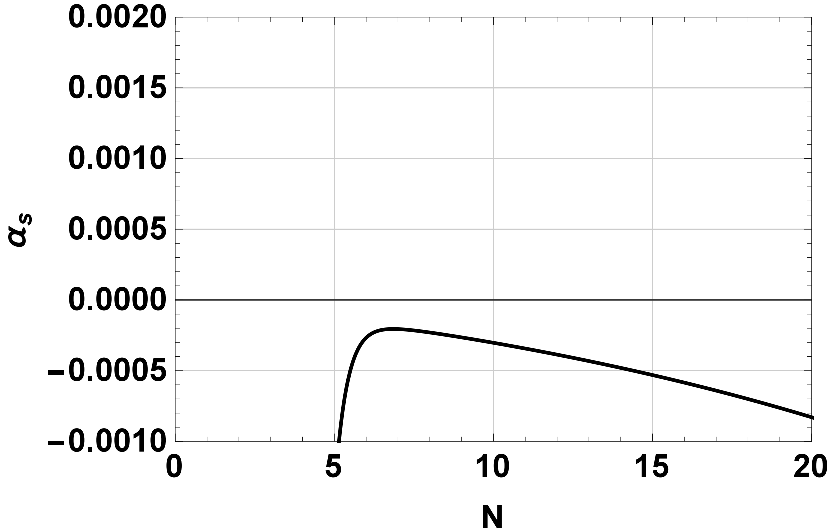

Figure 3.11. Analytical solution for the running $\alpha_\mathrm{s} \approx \mathrm{d}n_\mathrm{s}/\mathrm{d}N$ [1] at early times, using a non-Gaussianity parameter $a_3 = 0.05$. Estimating Hubble exit at $N \sim 10$, α_s is well within Planck's upper bound on the magnitude ($\sim 10^{-3}$).

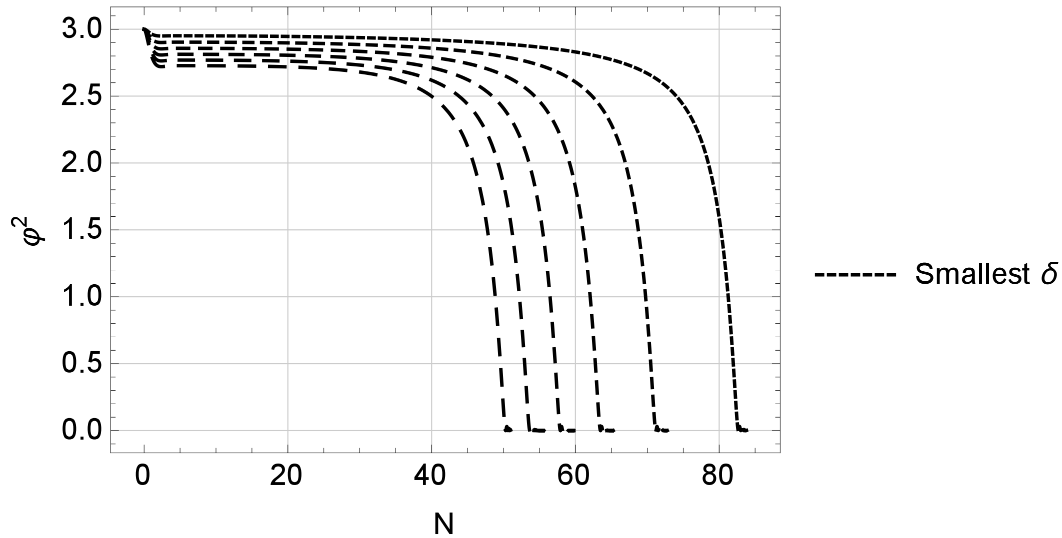

Figure 3.12. Evolution of $\varphi(N)^2$, from numerical solutions using $a_3 = 0.01\varphi^3$. Inflation ends at N_e where $\varphi(N_e) \approx 0$. Different curves correspond to different values of a_4, or $\delta = a_4 - 3$, where $\delta = 0.05, 0.1, 0.15, 0.2, 0.25, 3$. Smaller δ increase the duration of inflation.

where θ is defined such that

$$\cos\theta = \frac{\dot{\varphi}}{\sqrt{\dot{\varphi}^2 + \dot{\psi}^2}} \quad , \quad \sin\theta = \frac{\dot{\psi}}{\sqrt{\dot{\varphi}^2 + \dot{\psi}^2}} \,. \tag{3.56}$$

Based on the slow roll equations of motion for ψ and $\dot{\varphi} \approx \dot{\varphi}_* = -3\psi(a_4\varphi_*)^{-1}\dot{\psi}$ we obtain

$$\cos\theta \approx -\frac{3\psi}{a_4\varphi_*}\sin\theta \quad , \quad \sin\theta \approx 1 \,, \tag{3.57}$$

where we used $V_\psi \gg V_\varphi \approx 0$. To leading order of ψ, we therefore have

$$\eta_{\varphi\varphi}\cos^2\theta \approx 0 + O(\delta^2, \alpha_3^2, \psi^2) \tag{3.58}$$

$$\eta_{\varphi\psi}\sin\theta\cos\theta \approx 0 + O(\psi^2) \tag{3.59}$$

$$\eta_{\psi\psi}\sin^2\theta \approx -\frac{4\delta}{a_4 P v^2} + O(\psi^2) \,, \tag{3.60}$$

such that

$$n_{\mathrm{s}} \approx 1 - \frac{8\delta}{a_4 P v^2} \,. \tag{3.61}$$

Evaluating

$$P \equiv V(\varphi_*(\psi = 0), \psi = 0) = 1 - \frac{1}{a_4} \approx \frac{2}{3} + \frac{\delta}{9} + O(\delta^2) \tag{3.62}$$

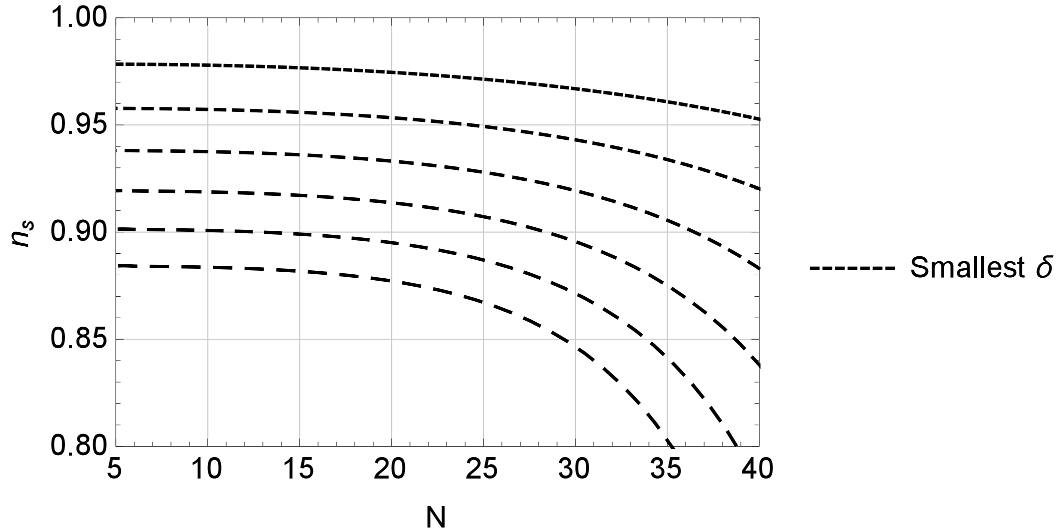

Figure 3.13. Spectral index $n_\mathrm{s}(N)$ as a function of e-folds N at Hubble exit from numerical solutions, using $a_3 = 0.01\varphi^3$. Different curves correspond to different values of a_4, or $\delta = a_4 - 3$, where $\delta = 0.05, 0.1, 0.15, 0.2, 0.25, 3$. Smaller δ brings the spectral index closer to one.

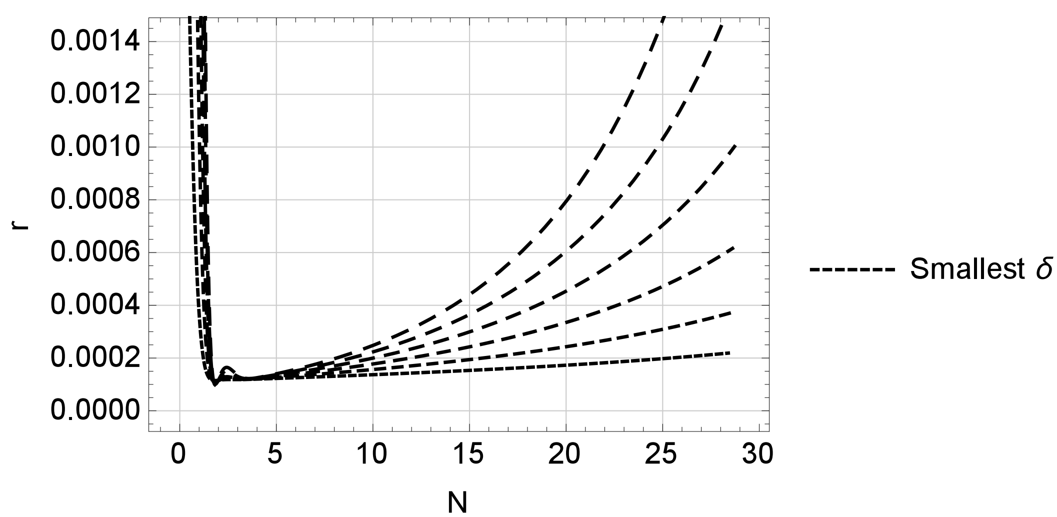

Figure 3.14. Tensor-to-scalar ratio $r(N)$ as a function of e-folds at Hubble exit from numerical solutions, using $a_3 = 0.01\varphi^3$. Different curves correspond to different values of a_4, or $\delta = a_4 - 3$, where $\delta = 0.05, 0.1, 0.15, 0.2, 0.25, 3$. A smaller δ decreases r.

leads to the final expression

$$n_\mathrm{s} \approx 1 - 12\frac{\delta}{a_4 v^2} + O(\delta^2)\,. \tag{3.63}$$

Imposing a slow-roll condition such as $\eta_{\psi\psi} \sim 10^{-2}$ requires $v^2/\delta \sim O(10^2)$, which implies typical values of n_s in the range $0.9 < n_\mathrm{s} < 1$.

3.4.3.2 Number of e-folds

Now, for total number of e-folds N_e we first need to find the value ψ_e of ψ at which inflation ends. Approximately, this stage occurs when

$$\eta_{\psi\psi}(\varphi_*, \psi_e) = \frac{V_{\psi\psi}}{V}|_{\varphi=\varphi_*,\psi=\psi_e} \approx 1 \tag{3.64}$$

during the end phase. Under the approximation $\varphi \approx 0$, we have

$$\frac{V_{\psi\psi}}{V} \approx \frac{4}{v^2}\frac{3\psi^2/v^2 - 1}{(1-\psi^2/v^2)^2}\,. \tag{3.65}$$

Then $V_{\psi\psi} \approx V(\varphi_*, \psi)$ gives

$$\frac{\psi_e^2}{v^2} \approx 1 + \frac{6}{v^2} \pm 2\sqrt{\frac{9}{v^4} + \frac{2}{v^2}} \tag{3.66}$$

$$= 1 + \frac{6}{v^2}\left(1 - \sqrt{1 + \frac{2v^2}{9}}\right)\,, \tag{3.67}$$

where we chose the minus sign in the second line. From the above expression we see that typically $\psi_e^2/v^2 - 1/3 \sim O(10^{-1})$. Then using

$$\Delta\psi \sim -\frac{V_\psi}{V}\Delta N \sim O(1)\Delta N\,, \tag{3.68}$$

we see that beyond $\psi^2/v^2 = 1/3$, we do not get many e-folds before reaching the point $\eta_{\psi\psi} \approx 1$, effectively ending inflation. In terms of the total number of e-folds, it is therefore justified to approximate

$$\psi_2(N)^2 \approx v^2/3 \quad \text{such that} \quad \varphi_*^2 = 0 \tag{3.69}$$

as the end point of inflation.

Since our analytical solution consists of ψ_1 and ψ_2, to find the total number of e-folds N_e at $\psi_2^2 = v^2/3$, we must first find the number of e-folds $N_{\rm g}$ at the gluing point. By definition of the latter,

$$\psi_{\rm g}^3 \equiv \psi_1(N_{\rm g})^3 = -a_3\,. \tag{3.70}$$

Denoting

$$\eta \equiv |\eta_{\psi\psi}(\varphi_1, \psi_1)| \approx \frac{4\delta}{a_4 P v^2} \approx \frac{6\delta}{a_4 v^2} + O(\delta^2)\,, \tag{3.71}$$

we have

$$2\eta = 1 - n_{\rm s}\,. \tag{3.72}$$

Using (3.43),

$$\exp(\eta N_{\rm g}) = \frac{v^2}{\psi_{\rm g}^2}\frac{\delta}{a_4} + 1 \tag{3.73}$$

which, inserted in (3.47), using (3.72) and setting $\psi_2(N_e)^2 = v^2/3$, implies

$$N_e = \frac{1}{1-n_{\rm s}}\left(\log\left(\frac{2}{v^2} + \frac{1-n_{\rm s}}{12}\chi\right) + 2\log\left(\frac{1-n_{\rm s}}{12}\chi v^2 + 1\right) - \log\left(\frac{2}{v^2} + \frac{1-n_s}{4}\right)\right)\,, \tag{3.74}$$

where $\chi \equiv v^2/\psi_g^2$. The relationship (3.74) is illustrated in Fig. 3.15.

Aside from the parameter v that appears in common Higgs-like or hybrid models, our observables depend on two new parameters a_3 and δ which describe the non-Gaussianity of the background state. Background non-Gaussianity effectively controls the amount of non-adiabatic evolution due to its modulation on the shifting of local φ-minima at φ_*. The dependence of the number of e-folds on the non-Gaussianity parameter a_3 is shown in Fig. 3.16, using the analytical solutions.

The dependence (3.74) of N_e on $n_{\rm s}$ is more complicated than in non-minimal Higgs models, but it is nevertheless related. To facilitate a comparison, we rewrite the expression as

$$N_e \approx \frac{f(1-n_{\rm s}, v, a_3)}{1-n_{\rm s}} \tag{3.75}$$

where the function f describes a weak, logarithmic dependence on $1-n_{\rm s}$. In non-minimal Higgs inflation, the analog of the function $f(1-n_{\rm s}, v, \alpha_3)$ is constant ($f = 2$) [61]. Here, the function increases logarithmically with growing $1 - n_{\rm s}$, taking values in the range $1 \lesssim f(1-n_{\rm s}, v, a_3) \lesssim 5$ for typical parameter values considered in our analysis. (An abbreviated derivation of (3.75) can be found in [75].)

3.5 Conclusions

Typically, potentials for the inflaton field are postulated so as to match existing observations. On the other hand, one of the most remarkable successes of inflation is that it explains the large-scale structure of the universe as originating from quantum fluctuations. It is inconceivable to quantize the fluctuations of the inflaton field *alone* without taking into account the quantum corrections to the background field potential. In other words, one cannot simply express the inflationary potential in terms of expectation val-

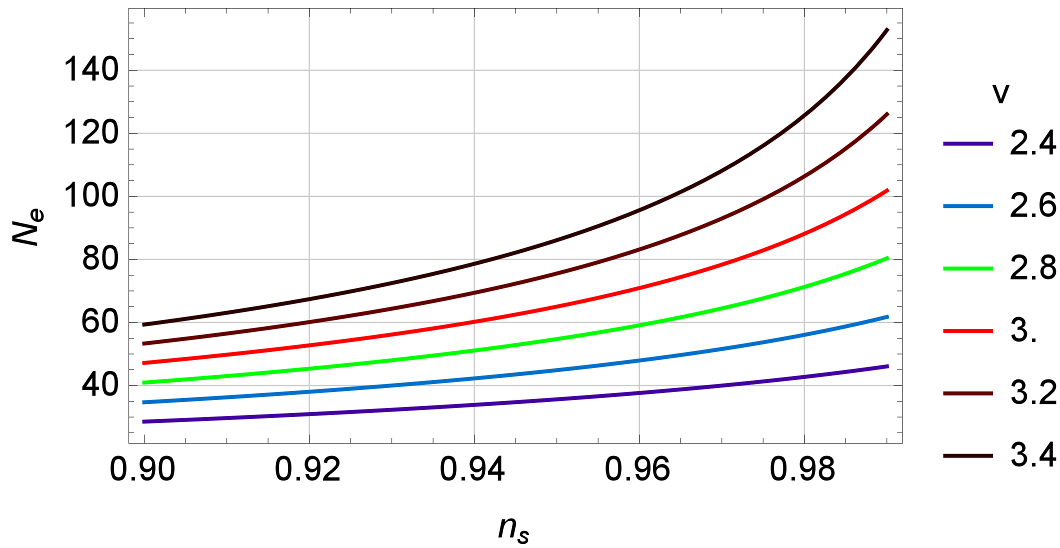

Figure 3.15. The number of e-folds, N_e, increases as a function of the spectral index $n_{\rm s}$, using the approximate relation (3.74). The function is shown for varying parameters v in the potential, while $a_3 = 0.05$. As a function of the non-Gaussianity parameters, the number of e-folds decreases; see Fig. 3.16. (Note that in the analytical relation (3.63), the variation of $n_{\rm s}$ mirrors the non-Gaussianity ratio $\delta/(a_4 v^2)$.)

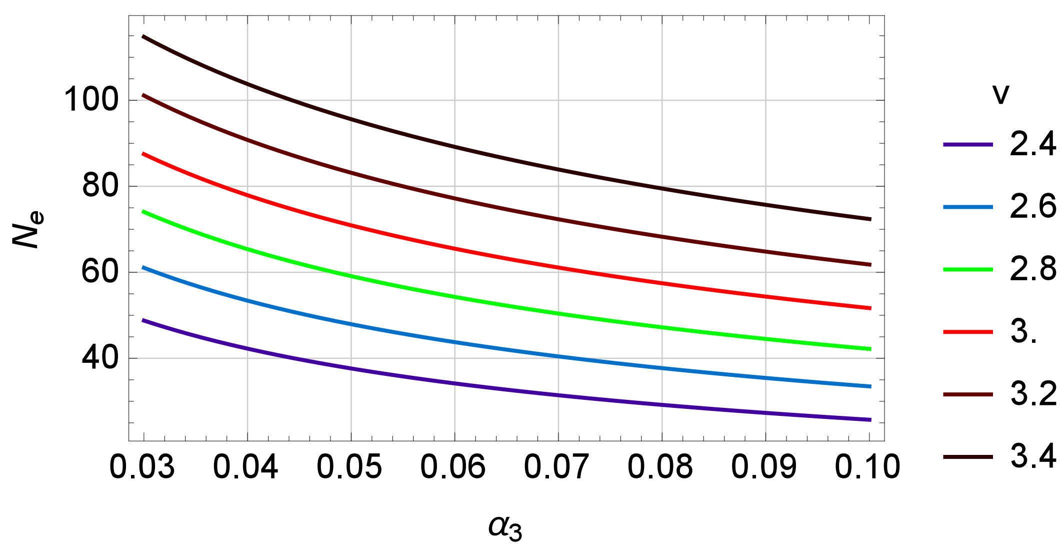

Figure 3.16. The number of e-folds, N_e, decreases with the amount of non-Gaussianity, parameterized by a_3, shown here for fixed $n_{\rm s} \approx 0.96$, $\delta = 0.1$ and using (3.74). Background non-Gaussianities increases the departure from adiabatic evolution, effectively ending inflation earlier than desired.

ues of the homogeneous background field, but should also take fluctuations and higher moments of the quantum state into account. It is customary to express the resulting effective potential in a derivative expansion (of the Coleman-Weinberg type); however, this method is not sufficient if one has to consider non-adiabatic evolution of quantum fluctuations. An adiabatic approximation is certainly valid during a slow-roll regime, but, as shown here, it can miss important features at the beginning and the end of slow-roll. Non-adiabaticity can play a crucial role in setting up the initial conditions for a slow roll phase as well as helping to ending it at the right time. We have presented a more general procedure for calculating the effects of such non-adiabatic evolution in the context of early-universe cosmology.

Using non-adiabatic effective methods, we have constructed an observationally consistent extension of Higgs-like inflation by introducing non-adiabatic quantum effects in a semiclassical approximation, although our formalism is applicable more generally for any inflationary potential. As shown, these effects imply that the classical potential is not only corrected in its coefficients but is also amended by new terms for independent quantum degrees of freedom, in particular the quantum fluctuation of the Higgs field. The original single-field model is therefore turned into a multi-field model. The multi-field terms incorporate quantum corrections of the background field, corresponding to backreaction of radiative corrections. Since the single-field potential is renormalizable, our quantum scenario is robust from the perspective of quantum field theory.

New interaction terms in the multi-field potential have coupling constants that depend on the background state, parameterizing its non-Gaussianity. They imply two new non-adiabatic phases that cannot be seen in low-energy potentials or in cosmological studies based completely on slow-roll approximations. In particular, an initial non-adiabatic phase, combined with the uncertainty relation for the fluctuation degree of freedom, sets successful initial conditions for inflation to take place, and a second non-adiabatic phase ends inflation after the right number of e-folds. In an indirect way, observational constraints show that background non-Gaussianity should be small, but it must be non-zero for the non-adiabatic phases to be realized. (The observational input we use here is not a limit on statistical non-Gaussianity in the inhomogeneity spectrum. Rather, the new link between the number of e-folds and background non-Gaussianity, shown in Fig. 3.16, makes it possible to use readily available limits on the number of e-folds and, in conjunction with Fig. 3.15, the spectral index in order to limit quantum non-Gaussianity of the background state of the inflaton.) Our model is highly constrained because this non-Gaussianity is bounded from below, but we are nevertheless able to derive successful inflation in the range of parameters available to us.

Our model presents a new picture of the role of the quantum state in inflationary cosmology. Quantum fluctuations not only provide the seeds of structure as initial conditions for perturbative inhomogeneity, they also play a crucial role in guiding the inflationary dynamics of the background state. With further analysis and observations, it may be possible to further constrain the quantum state of the inflaton based on cosmological investigations.

Chapter 4
Backreactions

4.1 Introduction

As we have argued in previous chapters, given a generic quantum system, it is often impossible to solve the theory exactly. In extreme cases—like those involving gravity—we are currently not even able to write down a self-consistent quantum equation of motion. Fortunately, in situations where curvature and energies are not too high, we may still make reliable predictions by treating the theory as a part-quantum and part-classical system. In such a treatment we often assume that the quantum contributions of the system evolve on a classical background. Discrimination between classical evolution and quantum evolution is often implicitly assumed—the quantum degrees-of-freedom (DoFs) are directly affected by the classical background but not the other way around. How quantum DoFs influence, or *backreact* on, the classical background still remains an open question.

In the context of cosmological backreactions, we are concerned with the question of how an evolving or fluctuating quantum field in an expanding universe might affect the expansion rate of the background geometry on which it is defined. In the absence of a complete and consistent quantum theory of gravity or cosmology, the background is mainly treated as a classical system. Such is the route typically taken in a standard analysis of cosmological perturbations in the inflationary epoch of our early universe. During this stage, we divide the system into a homogeneous background and inhomogeneous perturbations. The classical potential of the background is often an effective one obtained by integrating out quantum contributions or heavy DoFs. In principle, it will contain contributions from non-local effects. However, these non-local contributions are often assumed to be small if we are only interested in a low energy effective potential. As for the inhomogeneous perturbations, they are quantized and ultimately used to compute the n-point functions, which are then compared to observations. The significance of the backreaction problem is now obvious. It is relevant not only for observational and conceptual questions in cosmology (see for instance [74, 113–115] for recent contributions), it also offers general lessons for the classical-quantum correspondence [116–128] and the validity of hybrid schemes in which only one part of an interacting system is quantum [129–132]. In this chapter, we apply recent advances in canonical descriptions of semiclassical expansions to a system that has been used recently as a cosmological model [2], to what turns out to be the leading semiclassical order of our systematic scheme.

We will utilize a systematic inclusion of quantum fluctuations to construct effective models for early-universe cosmology which may be interpreted as being of hybrid type. While we will not directly address the quantum-to-classical transition, we will be able to shed new light on the related homogeneity problem of inflationary structure formation [124], which states that it should be impossible for translation-invariant quantum evolution to generate inhomogeneity out of an initial homogeneous vacuum state. We will examine the assumption of homogeneity more closely and enter a crucial new ingredient into this discussion, given by quantum non-locality. We will argue that quantum non-locality is able to imply a gradual process of transfer of super-horizon inhomogeneity—suggested to occur generically according to the Belinskii–Khalatnikov–Lifshitz (BKL) scenario [133])—down to smaller distance scales and eventually into the cosmological horizon. As an implication of quantum non-locality, this process may happen even in the quantization of a classical covariant theory. As we will observe for the first time in an application to quantum field theory, canonical moment methods give efficient access to spatial non-locality. (As has been known for some time, moment methods allow one to describe quantum non-locality in time by foregoing an adiabatic approximation of usual effective potentials.)

Let us first consider the canonical moments method in a quantum-classical correspondence context. For comparison sake, we will use a model considered in the treatment of [2]

$$H = \frac{1}{2}p_x^2 + \frac{1}{2}p_z^2 - ax + \frac{1}{2}(\omega^2 + \lambda x^2)z^2 \tag{4.1}$$

with two degrees-of-freedom x and z, where a, ω and λ are positive constants. The degree-of-freedom x remains completely classical in the original treatment and serves as a background which, while rolling down its linear potential $-ax$, excites the oscillator given by z via the interaction term $\frac{1}{2}\lambda x^2 z^2$. In [2], z is quantized and its backreaction effects on the roll-down rate of x are studied. This dynamics could be considered a toy model representing the effects of particle creation on a classical background.

Here, we extend the model in a way that goes beyond a strict background treatment of x. In the cosmological situation, the usual separation into background and perturbations exists only as an approximation of some inhomogeneous dynamics. Because both background and perturbation variables depend on the same fundamental degrees of freedom, given by the space-time metric and suitable matter fields, a separate quantization has a certain limited range of validity. To make matters worse, there may potentially be a problem of re-quantization as the potential of the matter system is often a quantum-corrected effective one—it either describes the vacuum, which is not applicable for inhomogeneous systems, or it includes inhomogeneous effects meaning that inhomogeneity has already been quantized before. Finally, even classically, the separation is not compatible with full general covariance but is maintained only by coordinate transformations that are small in the same sense in which inhomogeneity is considered a small perturbation on the background.

A complete treatment would have to be done within an elusive theory of quantum gravity, but while such a theory is still being constructed, some generic implications can be tested. In particular, if background and perturbations depend on the same fundamental fields, upon quantization there may well be correlations between them. Such correlations might even be required for general covariance to hold at the quantum level: A small coordinate change from t to $t+\xi$ with an inhomogeneous $\xi \ll t$, allowed at the perturbative level, maps a pure background variable such as the scale factor $a(t)$ into an inhomogeneously perturbed quantity, $a(t+\xi) \approx a(t) + \dot{a}\xi$. Assuming vanishing quantum correlations between background and perturbations is therefore too restrictive for a covariant quantum theory of the system.

The task then is to introduce correlation parameters between background and perturbations while maintaining a suitable level of tractability. We solve this problem here by utilizing a quasiclassical formulation of quantum dynamics [76–79] which introduces (or derives) canonical variables for moments of a quantum state. (As we will also demonstrate, the constructions of [2] are equivalent to these quasiclassical methods up to a certain order.) Recent advances in [81, 82] have extended these methods from a single degree of freedom to a pair of coupled and possibly correlated degrees of freedom. This extension makes it possible to derive a systematic formulation of cosmological perturbations with background correlations.

In [134], a quantum-field version of the model of [2] was introduced and studied, concluding that it was not possible to solve the homogeneity problem. We describe and extend related models by applying systematic quasiclassical methods to a quantum field on a semiclassical background. Also here, we are able to introduce correlations between the field and the background. Importantly, we will be able to reveal a new effect implied by the evolution of cross-correlations of different modes of a quantum field which can be used to address the homogeneity problem. In particular, we will observe that the usual assumption of an exactly homogeneous initial state is not always justified because super-Hubble inhomogeneity (as implied at early times by the BKL scenario) may trickle down to within the Hubble radius when quantum correlations are considered. We interpret this new effect as a consequence of quantum non-locality.

4.2 Canonical description of quantum dynamics

In order to facilitate the mathematical description of a quantum degree of freedom interacting with a classical variable, the construction presented in [2] uses a semiclassical approximation. Motivated by the well-known form of energy eigenstates of the harmonic oscillator, in particular of the Gaussian ground state, the authors observe that the first semiclassical correction of an oscillator can be described by doubling its degrees of freedom. In an energy eigenstate (in which the expectation value of $\hat{z}$ vanishes), instead of a single classical variable z, one may use a pair of fluctuation degrees of freedom, χ and ξ with momenta p_χ and p_ξ, such that the quadratic potential z^2 is replaced by $\chi^2+\xi^2$, and similarly for p_z^2. One can think of these replacements as an approximation inspired by a harmonic oscillator system (with potentially time-dependent frequency). The equations of motion in a system with Hamiltonian (4.1) are then given by

$$\ddot{x} = a - x\lambda(\xi^2+\chi^2), \quad \ddot{\xi} = -(\omega^2+\lambda x^2)\xi \quad \ddot{\chi} = -(\omega^2+\lambda x^2)\chi\,. \tag{4.2}$$

In addition, the angular momentum in the new (χ,ξ)-plane is constrained to be non-zero and equal to $\xi p_\chi - \chi p_\xi = 1/2$, restricting the allowed initial values. It is therefore impossible for only one of the new variables to be non-zero, in which case we would have equivalence with the classical formulation.

In [2], this doubling of degrees of freedom of a semiclassical oscillator in an energy eigenstate is obtained by including the variance as an independent degree of freedom, in addition to the expectation value of a basic variable such as z (which vanishes in an energy eigenstate). The expectation value of $\hat{p}_z^2$ in an uncorrelated Gaussian state with variance σ^2 is given by

$$\langle \hat{p}_z^2 \rangle = \langle \hat{p}_z \rangle^2 + \frac{\hbar^2}{4\sigma^2} \tag{4.3}$$

while the expectation value of $\hat{z}^2$ (in any state) equals

$$\langle \hat{z}^2 \rangle = \langle \hat{z} \rangle^2 + \sigma^2 \,. \tag{4.4}$$

The Gaussian expectation value of (4.1) in which only (z, p_z) has been quantized is therefore

$$\langle \hat{H} \rangle = \frac{1}{2} p_x^2 + \frac{1}{2} \langle \hat{p}_z \rangle^2 + \frac{\hbar^2}{8\sigma^2} - ax + \frac{1}{2}(\omega^2 + \lambda x^2)(\langle \hat{z} \rangle^2 + \sigma^2) \,. \tag{4.5}$$

The contribution $\hbar^2/(8\sigma^2)$ may be interpreted as a centrifugal potential of planar motion expressed in polar coordinates with radius σ, together with a spurious angle that does not appear in the potential of (4.5). Transforming to Cartesian coordinates implies the two degrees of freedom, χ and ξ such that $\chi^2 + \xi^2 = \sigma^2$, with the condition that their angular momentum has to equal $\hbar/2$ for $\hbar^2/(8\sigma^2)$ in (4.5) to be the correct centrifugal potential. The harmonic potential is then turned into $\frac{1}{2}(\omega^2 + \lambda x^2)(\langle \hat{z} \rangle^2 + \xi^2 + \chi^2)$, which implies the equations of motion (4.2) for vanishing expectation value $\langle \hat{z} \rangle$, as assumed in (4.2). To complete this construction, we also need a kinetic energy of the new variable σ or its Cartesian analogs ξ and χ, which will be provided naturally by our more general treatment below.

For comparison, we briefly recall an alternative derivation of the equations of motion closer to the approach of [2]: The two degrees of freedom can be derived directly by introducing Heisenberg operators

$$\hat{z}(t) = z(t)^* \hat{a}_0 + z(t) \hat{a}_0^\dagger \tag{4.6}$$

$$\hat{p}_z(t) = \dot{z}(t)^* \hat{a}_0 + \dot{z}(t) \hat{a}_0^\dagger \tag{4.7}$$

with the time-independent annihilation operator $\hat{a}_0$ and a time-dependent complex function $z = \xi + i\chi$. Commutation relations then imply that

$$\hat{z}^2 = z^* z(1 + 2\hat{n}_0) + (z^*)^2 \hat{a}_0^2 + z^2 (a_0^\dagger)^2 \tag{4.8}$$

$$\hat{p}_z^2 = \dot{z}^* \dot{z}(1 + 2\hat{n}_0) + (\dot{z}^*)^2 \hat{a}_0^2 + \dot{z}^2 (\hat{a}_0^\dagger)^2 \tag{4.9}$$

where $\hat{n}_0$ is the number operator associated with a_0. In the ground state, which we assume initially and where $\langle\hat{z}\rangle = 0 = \langle\hat{p}_z\rangle$, we have

$$\Delta(z^2) \equiv \langle(\hat{z}-\langle\hat{z}\rangle)^2\rangle = z^* z = \xi^2 + \chi^2 \tag{4.10}$$
$$\Delta(p_z^2) \equiv \langle(\hat{p}_z-\langle\hat{p}_z\rangle)^2\rangle = \dot{z}^* \dot{z} = \dot{\xi}^2 + \dot{\chi}^2\,, \tag{4.11}$$

valid at all times. Fluctuations therefore imply two dynamical real degrees of freedom. Moreover, the covariance (in the standard sense) equals

$$\Delta(zp_z) = \frac{1}{2}\langle\hat{z}\hat{p}_z + \hat{p}_z\hat{z}\rangle - \langle\hat{z}\rangle\langle\hat{p}_z\rangle = \xi\dot{\xi} + \chi\dot{\chi} \tag{4.12}$$

such that the uncertainty product

$$\Delta(z^2)\Delta(p_z^2) - \Delta(zp_z)^2 = (\xi\dot{\chi} - \dot{\xi}\chi)^2 \tag{4.13}$$

implies a constant $\xi\dot{\chi} - \dot{\xi}\chi = \hbar/2$ in the Gaussian ground state. The specific angular momentum required by (4.5) is therefore closely related to the uncertainty relation.

We point out that the canonical formulation underlying this procedure has been known for some time, and has in fact been (re)discovered independently in various fields, including quantum field theory [76], quantum chaos [78] and quantum chemistry [77,79]. Leading semiclassical corrections of a particle moving in one dimension with coordinate z and momentum p_z can be described by coupling the basic expectation values $\langle\hat{z}\rangle$ and $\langle\hat{p}_z\rangle$ to three additional variables, the quantum fuctuations $\Delta(z^2)$ and $\Delta(p_z^2)$, as well as the covariance $\Delta(zp_z) = \frac{1}{2}\langle\hat{z}\hat{p}_z + \hat{p}_z\hat{z}\rangle - \langle\hat{z}\rangle\langle\hat{p}_z\rangle$. It is useful to introduce a uniform notation that can easily be extended to higher moments, which we write, following [42, 135, 136], as

$$\Delta(z^a p_z^b) = \langle(\hat{z}-\langle\hat{z}\rangle)^a(\hat{p}_z-\langle\hat{p}_z\rangle)^b\rangle_{\mathrm{Weyl}} \tag{4.14}$$

in completely symmetric, or Weyl ordering.

According to [42, 43], the expectation values and moments form a phase space equipped with a Poisson bracket defined by

$$\{\langle\hat{A}\rangle, \langle\hat{B}\rangle\} = \frac{\langle[\hat{A},\hat{B}]\rangle}{i\hbar} \tag{4.15}$$

and extended to moments by using the Leibniz rule. As a simple consequence, the Poisson bracket of basic expectation values equals the classical Poisson bracket, $\{\langle\hat{z}\rangle, \langle\hat{p}_z\rangle\} = 1$, and moments have zero Poisson brackets with basic expectation values. A closed-form expression exists for the bracket of two moments [42, 83], but it is rather complicated.

In particular, it is not canonical.

For instance, for second-order moments, we have the brackets

$$\{\Delta(z^2),\Delta(zp_z)\}=2\Delta(z^2)\quad,\quad\{\Delta(zp_z),\Delta(p_z^2)\}=2\Delta(p_z^2)\quad,\quad\{\Delta(z^2),\Delta(p_z^2)\}=4\Delta(zp_z)\,. \tag{4.16}$$

With hindsight, the semiclassical formulation of [76–79] can be interpreted as a mapping from the 3-dimensional Poisson manifold with brackets (4.16) to canonical, or Casimir–Darboux coordinates. Explicitly, defining the mapping from

$$(\Delta(z^2),\Delta(zp_z),\Delta(p_z^2))\longrightarrow(s,p_s,U)$$

by

$$s=\sqrt{\Delta(z^2)}\quad,\quad p_s=\frac{\Delta(zp_z)}{\sqrt{\Delta(z^2)}}\quad,\quad U=\Delta(z^2)\Delta(p_z^2)-\Delta(zp_z)^2 \tag{4.17}$$

or its inverse,

$$\Delta(z^2)=s^2\quad,\quad\Delta(zp_z)=sp_s\quad,\quad\Delta(p_z^2)=p_s^2+\frac{U}{s^2}\,, \tag{4.18}$$

one can see that we have the canonical Poisson bracket $\{s,p_s\}=1$, while $\{s,U\}=\{p_s,U\}=0$.

These equations hold for all states as they only rely on kinematical properties coming from (4.15). If we make the additional assumption that second-order moments provide a good approximation of quantum dynamics at least for some time, we may insert (4.18) in the expectation value of the harmonic Hamiltonian, taken in an arbitrary semiclassical state. We then obtain the effective Hamiltonian

$$H_{\rm eff}=\langle\hat{H}\rangle=\frac{1}{2}p_x^2-ax+\frac{1}{2}p_z^2+\frac{1}{2}(\omega^2+\lambda x^2)z^2+\frac{1}{2}\left(p_s^2+\frac{U}{s^2}\right)+\frac{1}{2}(\omega^2+\lambda x^2)s^2\,, \tag{4.19}$$

still quantizing only (z,p_z). This Hamiltonian is equivalent to (4.5) if $U=\hbar^2/4$, the minimum value allowed by Heisenberg's uncertainty relation. It is more general if U is allowed to be greater than this value, in which case we are no longer restricted to Gaussian states. The derivation shows how the conserved quantity U is related to the uncertainty relation as well as angular momentum in an effective description after transforming to Cartesian coordinates. Transforming s as the radial coordinate in an auxiliary plane (together with a spurious angle) to Cartesian coordinates (ξ,χ) on this plane, the centrifugal potential $U/(2s^2)$ can be eliminated by doubling the fluctuation degree of freedom s:

$$H_{\text{Cartesian}} = \frac{1}{2}p_x^2 - ax + \frac{1}{2}\left(p_z^2 + p_\xi^2 + p_\chi^2\right) + \frac{1}{2}(\omega^2 + \lambda x^2)(z^2 + \xi^2 + \chi^2)\,. \tag{4.20}$$

The kinetic energy of s, or ξ and χ, is automatically provided by (4.19). In general, angular momentum for motion on the plane is bounded from below but not required to equal $\hbar/2$ for generic states.

The effective Hamiltonian generates equations of motion for x, z, and s, as well as their momenta. Semiclassical aspects of quantum evolution can therefore be described by an enlarged phase space of classical type. Compared with the classical equations, solutions require additional initial values which partially encode properties of quantum states. The specification of an arbitrary state would require infinitely many parameters, for instance, all moments required for the Hamburger problem that asks how a probability density can be reconstructed from all its moments. A semiclassical approximation replaces this infinite number with finitely many values, given by a minimum of three non-classical parameters s, p_s, and U at leading semiclassical order.

A simple initial state, which may be Gaussian but would not be required to stay so in an interacting system, can be specified by the choice

$$p_x(0) = x(0) = 0 \quad , \quad p_z(0) = z(0) = 0 \quad , \quad s(0) = \frac{1}{\sqrt{2\omega}} \quad , \quad p_s(0) = 0\,. \tag{4.21}$$

The specific value chosen for s mimics the ground state of a harmonic oscillator with frequency ω. We may leave the Casimir variable U as a free parameter, which is restricted by the inequality $U \geq \hbar^2/4$ but need not saturate it if the state is not required to be Gaussian. It would then be more difficult to find a specific wave function or density matrix that belongs to these parameters, but semiclassical evolution based on the equations given here can be performed without problems. As shown in Fig. 4.1, the dynamics of the model applied in [2] to Gaussian states are indeed sensitive to the value of U.

4.3 Correlations with the background

Having established the close relationship between [2] and canonical effective methods, we now use extensions of the latter to generalize the dynamics by including correlation degrees of freedom.

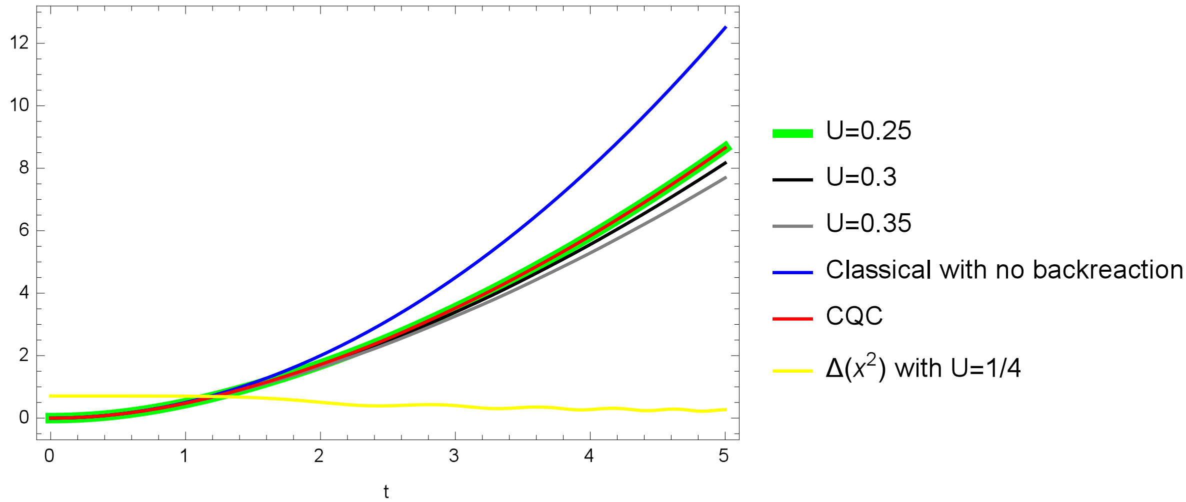

Figure 4.1. Classical evolution (blue) of the background variable $x(t)$, coupled to a semiclassical oscillator according to (4.1), compared with semiclassical evolution generated by (4.19) for the values $U = 0.25$, 0.3 and 0.35. Minimal uncertainty ($U = 0.25$, thick green) agrees with the CQC formulation of [2] (red), while other values of U lead to different dynamics. The yellow curve shows the evolving quantum fluctuation s of the coupled oscillator for the case of $U = 0.25$. To fix units, the same choices $\omega = \lambda = \hbar = 1$ as in [2] have been made.

4.3.1 Motivation

In a traditional background treatment, the variable x is treated completely classically, as in [2] and reviewed in the preceding section. However, if our ultimate goal is to analyze gravitational systems, the separation between background and quantum perturbation becomes a non-trivial issue; this is mainly because gravitational systems are subject to diffeomorphism symmetries that influence both the space-time geometry and matter. Consequently, the notion of perturbations on top of a homogeneous background is formal and ambiguous—in practice it can mean different things prior to fixing a gauge. At this stage mixing between background and perturbations seem almost unavoidable. We expect the mixing to carry over to the quantum theory in the form of correlations. When quantizing cosmological perturbations, we either quantize the system after fixing a gauge and solving the constraints, or we look for gauge-invariant objects, where a clear-cut separation of background and perturbations is often not guaranteed. For gauge-invariant properties, it is therefore of interest to go beyond a pure background treatment.

In particular, one commonly uses curvature perturbations or Mukhanov–Sasaki variables in order to put perturbative inhomogeneity into "gauge-invariant" form [137]. (In the present discussion, we may ignore the fact that these expressions are not fully gauge-invariant, but are so only with respect to a subset of all gauge transformations even in the linearized setting [138, 139].) These variables combine scalar modes of the metric with scalar fields. The former degrees of freedom form a single tensorial object together with the background metric, and therefore pure matter perturbations, distinct from the background, can be obtained only if a gauge is chosen in which the scalar modes of the metric vanish. In a gauge-invariant treatment, by contrast, it is not clear in which sense background and perturbations may be considered sufficiently independent to justify the assumption of vanishing correlations in a generic state upon quantization. (Thanks to general covariance, the quantum description of cosmological perturbations on an expanding background is not the same as quantum field theory on a curved background space-time, in which form it is often presented.)

In fact, the expressions for curvature perturbations in terms of metric and matter fields depend on the background scale factor and the Hubble parameter. Background and perturbations are therefore not independent in a framework that may derive quantized perturbations from some fundamental, unperturbed quantum theory of gravity. Such a derivation would, of course, be challenging, but it suggests that background correlations should be relevant. Unless it can be shown that quantum gravity could not possibly lead to quantum correlations between background and perturbations, it is not justified to assume that such correlations are absent or can be ignored. In this section, we derive a new theory of canonical effective equations with correlations, still applied to the toy Hamiltonian (4.1).

4.3.2 Canonical variables for second-order moments of two degrees of freedom

The technical task is then to generalize the simple mapping (4.18) to the moments of two classical degrees of freedom, x with momentum p_x and z with momentum p_z. To second order, each canonical pair has three individual moments, given by two fluctuations and a position-momentum covariance. These six individual moments are accompanied by four cross-covariances that contain one variable from each pair, such as $\Delta(xz)$. We therefore have a ten-dimensional Poisson manifold, which requires some work to put into canonical form with Darboux coordinates and Casimir variables. This derivation has been completed only recently, in [81,82], where explicit expressions for the ten second-order moments in terms of four pairs of Darboux coordinates — (s_1, p_{s_1}), (s_2, p_{s_2}), (α, p_α) and (β, p_β) — and two Casimir variables, C_1 and C_2, have been provided.

The resulting expressions are rather complicated and, in contrast to (4.18), have momentum variances that cannot be put into a form suitable for canonical kinetic energies with constant coefficients. (A proof of this claim has been provided in [81].) Fortunately, as we will show here, they can be reduced in an approximate way that provides canonical kinetic energies and retains a single independent correlation parameter, β.

We will quote here only the relevant moments. The new fluctuation parameters s_1 and s_2 as well as the correlation parameter β are introduced by the equations

$$\Delta(x^2) = s_1^2 \quad , \quad \Delta(z^2) = s_2^2 \quad , \quad \Delta(xz) = s_1 s_2 \cos\beta \,, \tag{4.22}$$

straightforwardly generalizing (4.18). Momentum variances also have a form similar to (4.18), given by

$$\Delta(p_z^2) = p_{s_2}^2 + \frac{U_2}{s_2^2} \tag{4.23}$$

for the variance of p_z. In contrast to (4.18), however, U_2 is not constant but rather depends on the canonical (Darboux) coordinates in the complicated form

$$U_2 = (p_\alpha - p_\beta)^2 + \frac{1}{2\sin^2\beta}\left((C_1 - 4p_\alpha^2) - \sqrt{C_2 - C_1^2 + (C_1 - 4p_\alpha^2)^2}\,\sin(\alpha+\beta)\right) \tag{4.24}$$

where C_1 and C_2 are the two Casimir variables, and therefore conserved. A similar expression exists for $\Delta(p_x^2)$. This rather long expression has been derived in [82] from the conditions that the moments, expressed in canonical variables, (i) obey the required Poisson brackets and (ii) are represented in a one-to-one manner without losing degrees of freedom in the canonical parameterization. This process, in particular the second condition, requires the inclusion of a canonical pair (α, p_α) in (4.24) whose physical interpretation is not as clear as that of the fluctuation and correlation parameters s_1, s_2 and β together with their momenta.

A possible interpretation can be obtained from the fact that moments can be defined for any pure or mixed state. While the known meaning of s_1, s_2 and β as well as their momenta shows that they are free parameters even if one restricts oneself to pure states, there must be additional parameters in a parameterization of all states that determine how much they deviate from a pure state. The new canonical coordinates α and p_α, as well as C_2, are candidates for such impurity parameters, a conjecture which has been tested with some success in a new canonical derivation of low-energy effective potentials given in [82].

4.3.3 Reduction of degrees of freedom

A minimal model that goes beyond the mapping for a single degree of freedom and retains a correlation parameter can be constructed by assuming small correlations. It turns out that many terms in (4.24) are then suppressed. Consequently, a minimal correlation model can be built as follows: We first assume that p_α and $\sqrt[4]{C_2}$ are much smaller than p_β and $\sqrt{C_1}$. The square root in (4.24) is then small, suppressing the dependence on α via $\sin(\alpha+\beta)$. The variable α therefore need not be assumed small, and it may in fact grow because, according to the cross-term of the first square in U_2, any effective Hamiltonian to which $\Delta(p_z^2)$ contributes in the kinetic energy, generates an equation of motion of the form $\dot{\alpha} \propto p_\beta + \cdots$ where dots indicate terms independent of p_β. It is therefore impossible for α to be exactly zero if p_β (a correlation parameter like β) is non-zero, while we need a non-zero p_β in order to consistently ignore p_α unless this variable is exactly zero. However, α appears in a bounded function in (4.24) that is suppressed by a small square root if our assumption about p_α and C_2 is satisfied, such that the dependence on α can be ignored in this case. The interpretation of α, p_α and C_2 as impurity parameters suggests that our approximation should be valid whenever a state is close to being pure.

We then have the simplified expression

$$\Delta(p_2^2) = p_{s_2}^2 + \frac{p_\beta^2}{s_2^2} + \frac{C_1}{2s_2^2 \sin^2\beta} \tag{4.25}$$

which, in a generalization of the momentum terms in (4.19), can be interpreted as (twice) the kinetic energy of a particle moving in three dimensions, expressed in spherical coordinates (s_2, β, φ) with a spurious degree of freedom φ. In contrast to the Cartesian version of (4.19), one of the angles, β, now is physically meaningful. It is, in fact, the correlation parameter relevant for our present aims.

As shown by the last term in (4.25), the momentum p_φ of the spurious angle is constrained to equal the constant $\sqrt{C_1/2}$. If we include the second degree of freedom in the mapping, given by the background variable x, it would have a similar kinetic energy with the same angle β and its momentum p_β. The two 3-dimensional systems would therefore be subject to constraints.

Here, we apply the 3-dimensional model only to the oscillating degree of freedom, z, while the background degree of freedom x is extended only by a fluctuation variable, s, as in a mapping for a single degree of freedom. In this way, we are able to construct a minimal extension of the model by parameterizing the correlation variable, β. For x, we therefore have the kinetic contribution

$$\frac{1}{2}\left(p_x^2+p_s^2+\frac{U}{s^2}\right)=\frac{1}{2}(p_x^2+p_X^2+p_Y^2) \tag{4.26}$$

as before, see (4.20), implying the contribution

$$-ax+\frac{1}{2}\lambda(x^2+s^2)z^2=-ax+\frac{1}{2}\lambda(x^2+X^2+Y^2)z^2 \tag{4.27}$$

to the effective potential in Cartesian coordinates for quantum fluctuations, in the style of [2] but now applied also to the background.

For z, we transform the kinetic energy implied by (4.25) to 3-dimensional Cartesian coordinates (ξ,χ,ζ), such that

$$\frac{1}{2}\left(p_z^2+p_{s_z}^2+\frac{U_2}{s_z^2}\right)=\frac{1}{2}(p_z^2+p_\xi^2+p_\chi^2+p_\zeta^2)\,. \tag{4.28}$$

This variable contributes several terms to the effective potential:

$$\frac{1}{2}(\omega^2+\lambda x^2)(z^2+\xi^2+\chi^2+\zeta^2)+2\lambda xzs\zeta \tag{4.29}$$

where the last term comes from the correlation (4.22). In our Cartesian-inspired coordinates $\Delta(xz)=s_1s_2\cos(\beta)\equiv s\zeta$, so the coordinate ζ is a direct indication of the amount of correlation between "background" x and the "quantum" DoF z. The Hamiltonian is therefore

$$\begin{aligned} H &= \frac{1}{2}(p_x^2+p_z^2+p_X^2+p_Y^2+p_\xi^2+p_\chi^2+p_\zeta^2)-ax+\frac{1}{2}(\omega^2+\lambda x^2)(z^2+\xi^2+\chi^2+\zeta^2) \\ &\quad +\frac{1}{2}\lambda\left(z^2(X^2+Y^2)+4xz\sqrt{X^2+Y^2}\zeta\right)\,. \end{aligned} \tag{4.30}$$

For zero cross-correlations, we have $\zeta=0$ and the Hamiltonian is reduced to a strict background model if we also set $X^2+Y^2=0=p_X^2+p_Y^2$ (to obtain vanishing x-fluctuation).

4.3.4 Diagonalization

Like many prototypical models, the consequences of a dynamical background include a time-dependent frequency for the perturbations. However, unlike the typical models, our system also has additional degrees-of-freedom representing quantum fluctuations. Indeed, our new effective Hamiltonian (4.30) can be interpreted as a system of four harmonic oscillators—z, ξ, χ, and ζ—with frequencies that depend on time through the background variable x and its fluctuation parameter, $\sqrt{X^2+Y^2}$. The term $xz\sqrt{X^2+Y^2}\zeta$ in (4.30) implies that z and ζ are not normal coordinates of the oscillator system.

We may attempt to (time-dependently) diagonalize this coupling between z and the component ζ of its fluctuation/correlation. In order for the diagonalization to be feasible, we assume that the typical time scale of evolution for z and ζ is much smaller than the time scale of x and its fluctuations. We will then be able to treat the coefficients as approximately time-independent, allowing a straightforward diagonalization of the quadratic form.

Considering the entire Hamiltonian H, the time scale for x is of the order $1/a \sim O(1)$ (noting that we set the mass to one and use natural units), while the time scale for $\Delta(x^2) = X^2+Y^2$ is of the order $1/(\lambda z^2)^{1/2} \gg 1$, as will be confirmed in Fig. 4.5. Turning to z and ζ, we see that their time scales are roughly $1/(\lambda x^2)^{1/2}$ and $1/(x\sqrt{X^2+Y^2})^{1/2}$ which are typically very small because x^2 and $\Delta(x^2)$ grow large at late times. The hierarchy in time scales, along with the assumption $\omega^2 \sim O(1)$, such that $\omega^2 \ll \lambda x^2$, justifies the following approximation.

We may rewrite the $z-\zeta$ part of (4.30) as

$$H_{z-\zeta} = K + \frac{1}{2}\lambda\left((x^2+\delta^2)z^2 + x^2\zeta^2\right) + 2\lambda x\delta z\zeta \tag{4.31}$$

where

$$\delta = \sqrt{X^2+Y^2} \quad \text{and} \quad K = \frac{1}{2}(p_z^2 + p_\zeta^2)\,. \tag{4.32}$$

Upon diagonalization we obtain the normal (angular) frequencies

$$\omega_1^2 = \lambda\left(x^2 + \frac{1}{2}\delta^2 - \delta\sqrt{16x^2+\delta^2}\right) \tag{4.33}$$

$$\omega_2^2 = \lambda\left(x^2 + \frac{1}{2}\delta^2 + \delta\sqrt{16x^2+\delta^2}\right)\,. \tag{4.34}$$

Since δ is the fluctuation of the classical degree of freedom, x, we expect it to be much smaller than x in magnitude. This result implies that the typical behavior of evolution of z and ζ can be described as a fast, nearly harmonic oscillation with frequency $\omega_1 + \omega_2 \approx \sqrt{\lambda}x$, modulated by a slow oscillation with frequency $\omega_2 - \omega_1 \approx 4\sqrt{\lambda}\delta$. The introduction of the correlation parameter β (or equivalently ζ) therefore gives rise to beat-like behavior, which is new and only present if we capture the effect of the background degree of freedom and its fluctuation using the 2-particle mapping.

The "normal coordinates" are given by

$$e_1 = N_1\left(\left(\frac{\delta}{4x} - \sqrt{1+\frac{\delta^2}{16x^2}}\right)z + \left(\frac{\delta}{4x} + \sqrt{1+\frac{\delta^2}{16x^2}}\right)\zeta\right) \tag{4.35}$$

$$e_2 = N_2(z+\zeta)\,, \tag{4.36}$$

where N_1 and N_2 are normalization constants. We see in Fig. 4.2 that the coefficients of z and ζ in the first line grow to be of similar magnitude but opposite signs, as a consequence of a decreasing δ/x such that the background degree of freedom is becoming more and more classical.

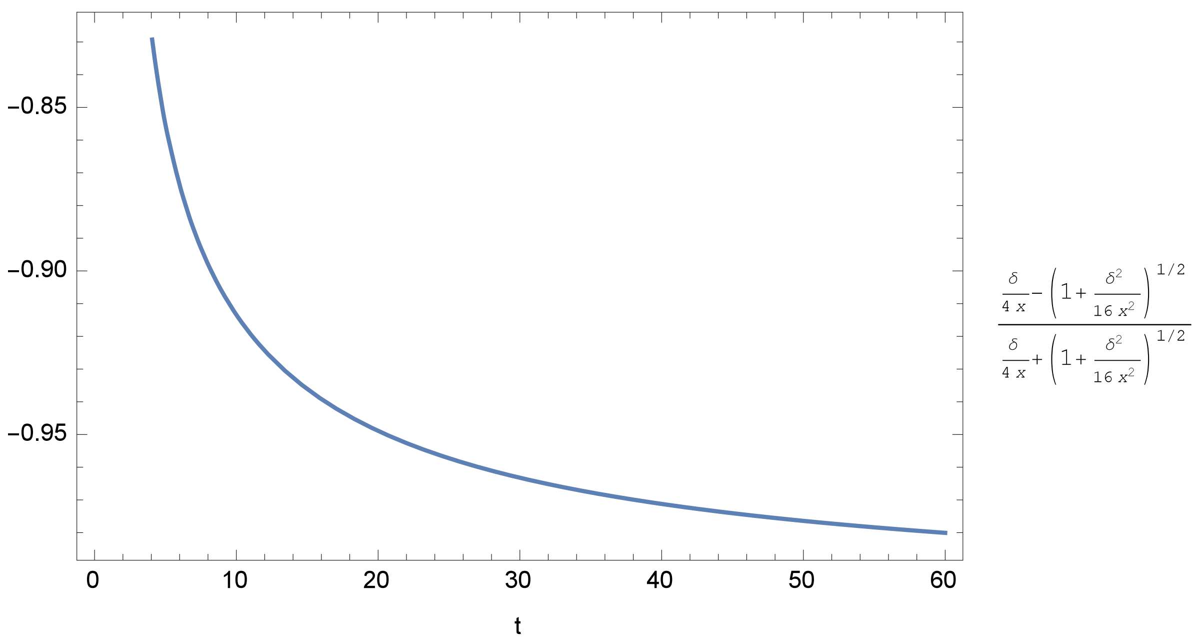

Figure 4.2. Behavior of the normal coordinate e_1 in (4.35). The general trend indicates that $e_1 \to -z+\zeta$ asymptotically. The parameters and initial values used here are specified in Section 4.3.5.

4.3.5 Dynamical implications

At the beginning of this section, we showed that the semiclassical equations of motion agree with those of [2], but only if $z(0) = p_z(0) = 0$ as appropriate for a system in an initial vacuum state and only if a single degree of freedom, z, is quantized. We can mimic the same initial conditions for the case where both degrees of freedom are quantized if

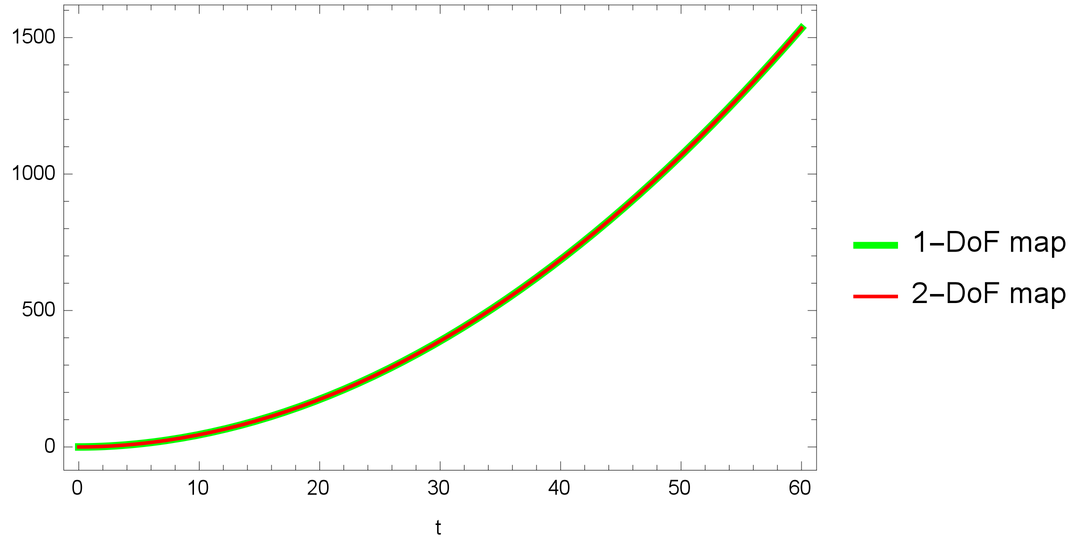

Figure 4.3. Numerical evolution for $\langle x(t)\rangle$ with (red) and without (green) background correlations, respectively, and equivalence conditions imposed. Starting with the same initial values, the evolutions do not deviate from each other.

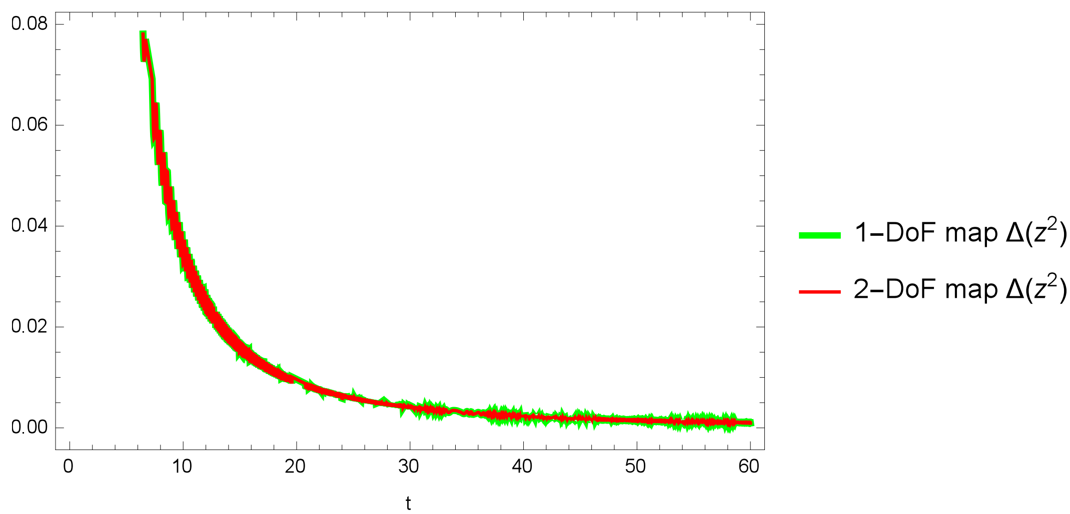

Figure 4.4. Fluctuations $\Delta(z^2)$ with (red) and without (green) background correlations, respectively, imposing equivalence conditions.

we also impose $\zeta(0) = p_\zeta(0) = 0$, which we call the *equivalence conditions*. We also choose $Y(0) = \chi(0) = \frac{1}{\sqrt{2}} = p_X(0) = p_\xi(0)$ and other variables initially 0. The last two equalities for momenta are required by the interpretation of the Casimir variables identified with angular momentum.

Numerical simulations with these conditions do not reveal any additional features as seen in Figs. 4.3 and 4.4. Therefore, the equivalence conditions turn the mapping for two degrees of freedom into a system equivalent with [2]. Even though we do introduce

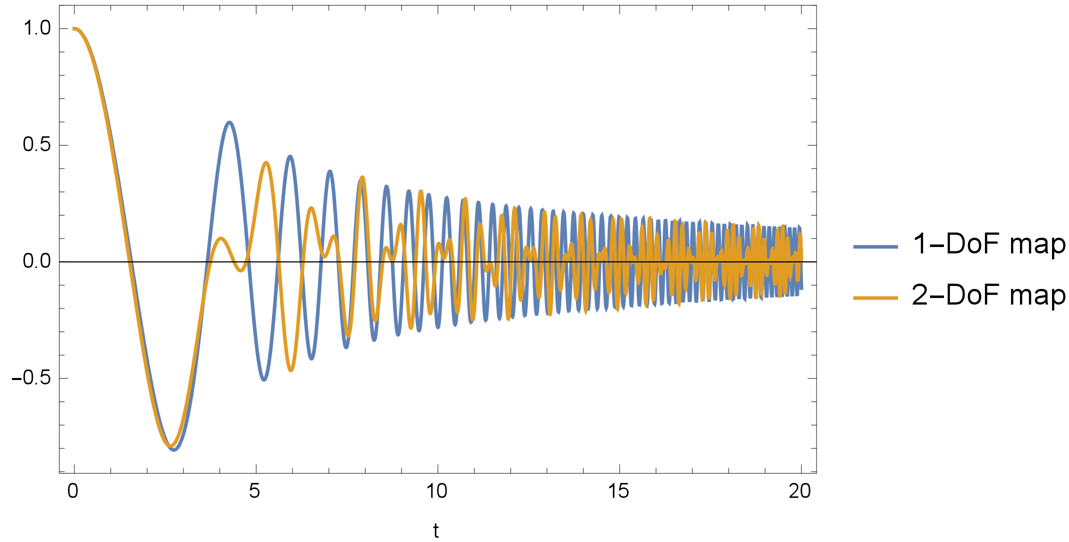

Figure 4.5. Expectation value $\langle z(t)\rangle$with (blue) and without (orange) background correlations, respectively. Using the more general mapping for two degrees of freedom, we see a slow frequency modulation of the original the fast oscillations. This is the beat-like behavior mentioned in the text.

couplings of X, Y and ζ to x, the initial conditions $z(0) = 0$ and hence $\dot{p}_\zeta(0) = 0$ imply $\zeta(t) = 0$ throughout evolution, meaning that x effectively couples only to $\xi^2 + \chi^2$, exactly as in the mapping for a single degree of freedom.

Interesting new effects are, however, obtained if the initial value of z is not zero, such that we do not start in a vacuum state. Physically, it could be justified to use such initial values after a phase transition, where the inflaton field acquires a non-zero vacuum expectation value. For instance, using $z(0) = 1$, Figs. 4.5 and 4.6 show that both $\langle z\rangle$ and $\Delta(z^2)$ obtain new oscillatory features from the mapping for two degrees of freedom. In addition to fast oscillations, $z(t)$ is also modulated by a low-frequency oscillation. The x-fluctuation, $\Delta(x^2)$, also increases with time. Using (4.30), we can think of $\Delta(x^2)$ as a particle in a central-force problem, in which $\Delta(x^2)$ is subject to a central force that decreases with time due to the decrease of z^2 and $xz\zeta$. Since $\Delta(x^2)$ started off with a non-zero initial momentum $p_X^2 + p_Y^2$ due to the uncertainty relation, the particle (fluctuation) will eventually escape to a larger distance from the center.

The correlation between the two degrees of freedom is shown in Fig. 4.7. Two interesting features are the boundedness of $\Delta(x^2)\Delta(z^2)$ in Fig. 4.8, and the upper bound of the background correlation between z and x for a given state in Fig. 4.7. The latter is explained by the interpretation of the correlation parameter β as a spherical angle in an auxiliary space, implied by the appearance of (4.25) in the form of kinetic energy

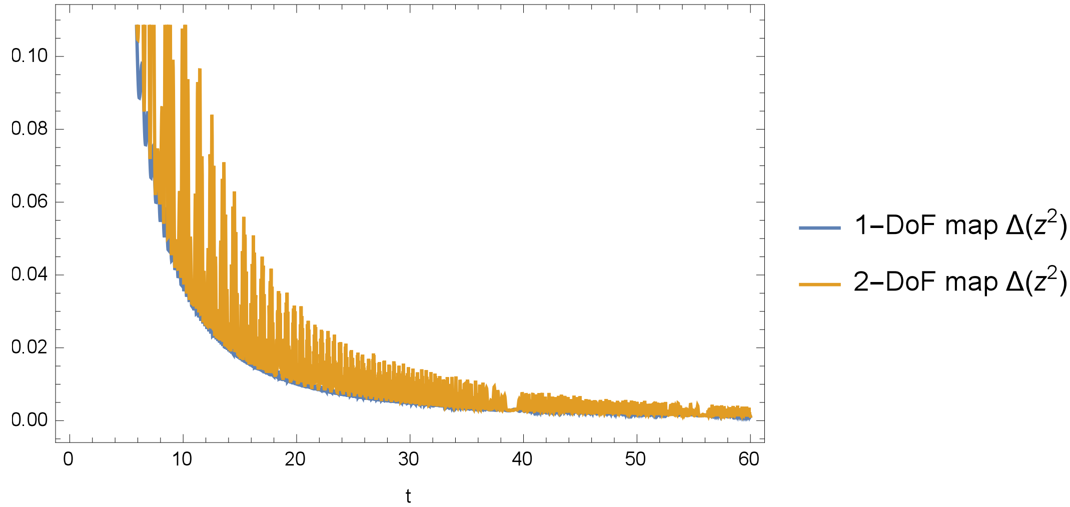

Figure 4.6. Fluctuation $\Delta(z^2)$ with (blue) and without (orange) background correlations, respectively. There is again an enhanced oscillation behavior due an additional dimension in the fluctuation space given by the ζ direction.

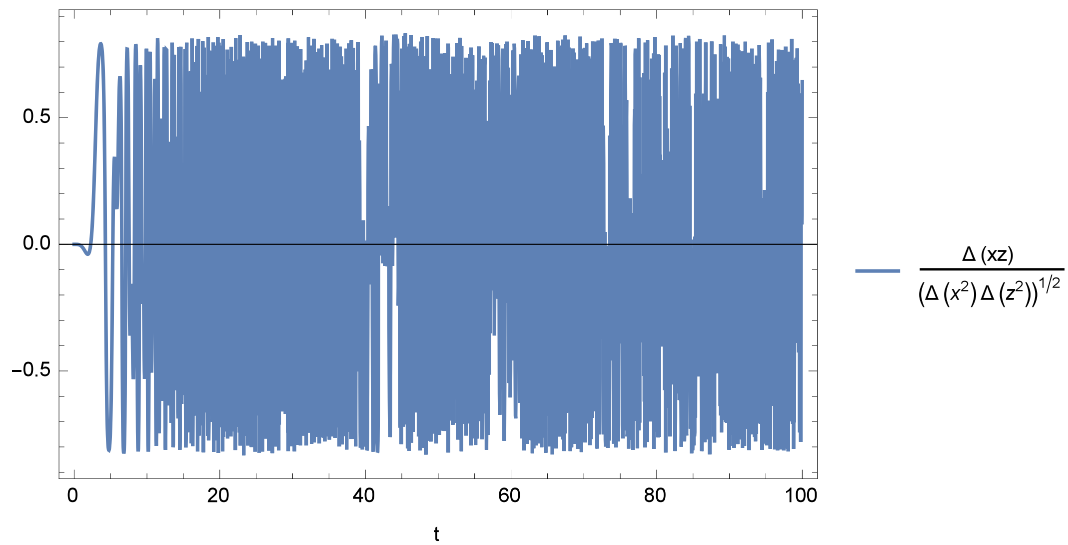

Figure 4.7. Time-dependent background correlation $\rho_{x,z} = \Delta(xz)/\sqrt{\Delta(x^2)\Delta(z^2)}$. Its local maxima are near but not equal to one at late times. There is therefore a maximum correlation for the state implied by our initial conditions.

in spherical coordinates. The angular momentum $\sqrt{C_1/2}$ in this auxiliary space is conserved and generically non-zero for a given initial state. For a non-zero value, it is then impossible for the correlation angle to get arbitrarily close to the poles of the spherical system, such that $\cos\beta$ keeps a certain distance from its general limiting values ± 1. This function equals the combination of moments plotted in Fig. 4.7, confirming the reduced upper bound for the given state implied by our initial values.

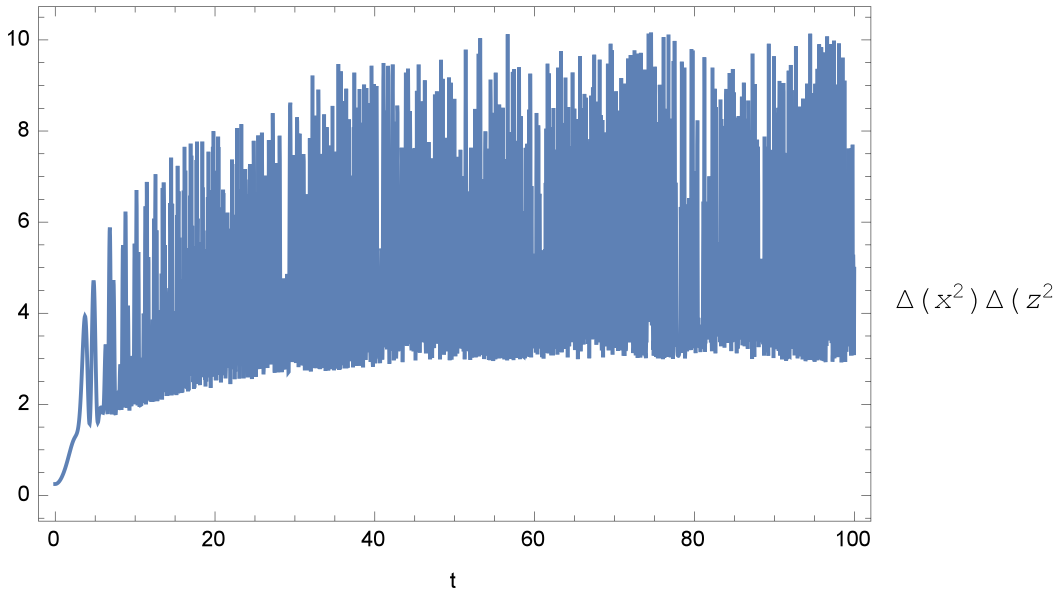

Figure 4.8. Boundedness of $\Delta(x^2)\Delta(z^2)$. Since $\Delta(x^2)$ increases while $\Delta(z^2)$ decreases, the balance between the two in their product is unexpected.

Let us briefly summarize what we have learned so far. Firstly, we have shown that canonical methods can be extended to multi-DoF systems to describe correlations between the background and the quantum DoFs. The multi-DoF model reduces to the one in [2] if we set correlations to zero. Moreover, we find that the magnitude of correlations is bounded from above. The boundedness is related to the Casimir variable C_1, which is generically not zero unless we fine-tune our initial state.

4.4 Field theory model

The formalism of [2] has also been applied to a quantum field back-reacting on a classical homogeneous background [80]. In this section, we will show how these methods are related to moments of a quantum field; see [91] for a general formulation of quantum fields by moments with a derivation of the Coleman–Weinberg potential [90]. The Coleman–Weinberg potential is conceptually related to the setting considered in [80] as it results from an expansion of a quantum field around a homogeneous background expectation value. The backreaction equations of [80] can therefore be embedded in a canonical effective theory by a suitable extension of [91]. In particular, the equations correspond to a leading-order formulation of moment equations by canonical Darboux coordinates. In contrast to the previous section, however, a complete Darboux formulation of a quantum field is challenging because a single quantum field implies a multitude of independent degrees of freedom, which are hard to describe by canonical variables for moments if all possible cross-correlations are included. Nevertheless, an embedding of [2] is feasible.

4.4.1 Modes on a compact homogeneous background

We consider a two-field model with a single spatial dimension, which will be reduced to a quantum field ψ and a classical field ϕ. Both fields are scalar and real. Extending the interactions of (4.1), we introduce the classical action

$$S = \int \mathrm{d}t\mathrm{d}x \left(\frac{1}{2}\left(\dot{\phi}^2 - (\partial_x\phi)^2\right) - V(\phi) + \frac{1}{2}\left(\dot{\psi}^2 - (\partial_x\psi)^2\right) - \frac{1}{2}(m^2 + \lambda\phi^2)\psi^2 \right) \quad (4.37)$$

with the mass m of ψ and a coupling constant λ that may be interpreted as providing a ϕ-dependent correction to the mass of ψ. The field ϕ moves in a potential $V(\phi)$ which, generically, may also include a mass term. As an extension of the preceding section, however, we will continue to assume that $V(\phi)$ is linear in our numerical examples.

In a typical effective potential calculation, one seeks to integrate out the "quantum" degrees-of-freedom ψ, resulting in a ϕ-parameterized correction to the potential energy $V(\phi)$. In this process, a crucial assumption of space-time homogeneity of the state is often made to make the explicit integration tractable. This assumption effectively neglects higher-order derivative terms resulting from the integration—the real-time corrections as well as non-local effects are thus lost. This is not desirable when we wish to discuss dynamic backreactions and inhomogeneity growth. Canonical methods applied to quantum fluctuations offer a way to circumvent the need to assume a complete homogeneous state. The price we pay is that the backreactions usually have to be calculated numerically. But this is already common practice for calculations done in curved space-time, especially in cosmology. In the following, we will first analyze the backreaction of model (4.37) assuming background spatial homogeneity but allowing ψ to depend on position. Then we will discuss how the backreactions in the canonical description may generate the growth of homogeneity.

In the quantum-mechanical model of the previous section, backreaction of z on x causes an exchange of energy: the quantum degree of freedom z backreacts on the classical one and saps its energy, causing x to roll down more slowly on its linear potential. We expect similar transfer in energy in the field version, where the energy lost by the classical field excites particle production for the quantum field.

Using Legendre transformation we derive the Hamiltonian density

$$\mathcal{H} = \frac{1}{2}\left(\Pi_\phi^2 + (\partial_x\phi)^2\right) + V(\phi) + \frac{1}{2}\left(\Pi_\psi^2 + (\partial_x\psi)^2 + \Omega_\phi(t,x)^2\psi^2\right) , \tag{4.38}$$

introducing

$$\Omega_\phi(t,x) = m^2 + \lambda\phi^2(t,x) . \tag{4.39}$$

In order to facilitate an analysis of particle production, we should expand ψ in Fourier modes with respect to a 1-dimensional, spatial wave number k. The resulting system can then be interpreted as a background field, ϕ, coupled to a large number of oscillators with ϕ-dependent mass and frequency.

Fourier transforms of the basic canonical field are given by

$$\psi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathrm{d}k e^{ikx}\tilde{\psi}(k) \tag{4.40}$$

$$\Pi_\psi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathrm{d}k e^{-ikx}\tilde{\Pi}_\psi(k) , \tag{4.41}$$

making time dependence implicit in this notation. Note that we have chosen to write the $\Pi_\psi(x)$ expansion (4.41) in terms of canonical conjugates to $\tilde{\psi}(k)$, which accounts for the negative sign in the exponential. The modes obey the reality conditions

$$\tilde{\psi}(-k) = \tilde{\psi}(k)^* \quad \text{and} \quad \tilde{\Pi}_\psi(-k) = \tilde{\Pi}_\psi(k)^* . \tag{4.42}$$

Choosing opposite signs in the exponentials used to transform ψ and Π_ψ, respectively, simplifies the canonical structure of modes. In particular, the calculation

$$\int \mathrm{d}x\dot{\psi}\Pi_\psi = \frac{1}{2\pi}\int \mathrm{d}x \int_{-\infty}^{\infty}\mathrm{d}k\int_{-\infty}^{\infty}\mathrm{d}l e^{i(k-l)x}\dot{\tilde{\psi}}(k)\tilde{\Pi}_\psi(l) = \int_{-\infty}^{\infty}\mathrm{d}k\dot{\tilde{\psi}}(k)\tilde{\Pi}_\psi(k) \tag{4.43}$$

implies that $\tilde{\Pi}_\psi(k)$ is canonically conjugate to $\tilde{\psi}(k)$.

Keeping the ϕ-Hamiltonian density $\mathcal{H}_\phi$ unchanged, the Hamiltonian for Fourier modes is then given by

$$\begin{aligned} H &= \int \mathrm{d}x\mathcal{H}_\phi + \frac{1}{4\pi}\int \mathrm{d}x\mathrm{d}k\mathrm{d}l\left(e^{-i(k+l)x}\tilde{\Pi}_\psi(k)\tilde{\Pi}_\psi(l) + (-kl + \Omega_\phi(t,x)^2)e^{i(k+l)x}\tilde{\psi}(k)\tilde{\psi}(l)\right) \\ &= \int \mathrm{d}x\mathcal{H}_\phi + \frac{1}{2}\int_{-\infty}^{\infty}\mathrm{d}k\left(|\tilde{\Pi}_\psi(k)|^2 + k^2|\tilde{\psi}(k)|^2 + N_k[\psi]\right) \end{aligned} \tag{4.44}$$

with the non-local (in k-space) contribution

$$N_k[\psi] = \tilde{\psi}(k)\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{d}l\int \mathrm{d}x\Omega_\phi(t,x)^2 e^{i(l+k)x}\tilde{\psi}(l)\,. \tag{4.45}$$

In order to decouple different k and obtain a local k-Hamiltonian, we now assume that $\Omega_\phi(t,x)$ is homogeneous in x, such that the x-integration in $N_k[\psi]$ results in a delta function that removes non-locality. Since $\Omega_\phi(t,x)$ depends on the background field $\phi(x)$ according to (4.39), the background field is assumed to be spatially homogeneous from now on. It may, however, be time-dependent. With this assumption, the classical Hamiltonian

$$H = \int \mathrm{d}x\mathcal{H}_\phi + \frac{1}{2}\int_{-\infty}^{\infty}\mathrm{d}k\left(|\tilde{\Pi}_\psi(k)|^2 + (k^2+\Omega_\phi(t)^2)|\tilde{\psi}(k)|^2\right) \tag{4.46}$$

is local.

A final transformation introduces real fields in k-space by splitting $\tilde{\psi}(k)$ and $\tilde{\Pi}_\psi(k)$ into real and imaginary parts,

$$\tilde{\psi}(k) = \frac{1}{\sqrt{2}}(A(k)+iB(k)) \tag{4.47}$$

$$\tilde{\Pi}_\psi(k) = \frac{1}{\sqrt{2}}(C(k)-iD(k))\,. \tag{4.48}$$

Reality conditions imply that $A(k)$ and $C(k)$ are even functions while $B(k)$ and $D(k)$ are odd. Continuing the calculation in (4.43), we have

$$\begin{aligned}\int_{-\infty}^{\infty}\mathrm{d}k\dot{\tilde{\psi}}(k)\tilde{\Pi}_\psi(k) &= \frac{1}{2}\int_{-\infty}^{\infty}\mathrm{d}k(\dot{A}(k)+i\dot{B}(k))(C(k)-iD(k)) \qquad (4.49)\\ &= \frac{1}{2}\int_{-\infty}^{\infty}\mathrm{d}k\left(\dot{A}(k)C(k)+\dot{B}(k)D(k)-i(\dot{A}(k)D(k)-\dot{B}(k)C(k))\right)\,.\end{aligned}$$

The imaginary contribution vanishes because it integrates an odd function $\dot{A}(k)D(k)-\dot{B}(k)C(k)$ over the full real range of k. The real contribution is even and can therefore be restricted to only positive k, such that

$$\int_{-\infty}^{\infty}\mathrm{d}k\dot{\tilde{\psi}}(k)\tilde{\Pi}_\psi(k) = \int_0^{\infty}\mathrm{d}k\left(\dot{A}(k)C(k)+\dot{B}(k)D(k)\right)\,. \tag{4.50}$$

If we restrict to positive k, we therefore have the momenta $\Pi_A(k)=C(k)$ and $\Pi_B(k)=D(k)$ of $A(k)$ and $B(k)$, respectively.

Our final expression for the classial Hamiltonian in terms of real modes is therefore

$$H = \int \mathrm{d}x \mathcal{H}_\phi + \frac{1}{2}\int_0^\infty \mathrm{d}k \left(\Pi_A(k)^2 + \Pi_B(k)^2 + \omega_\phi^2(k)(A(k)^2 + B(k)^2)\right) , \tag{4.51}$$

providing two harmonic oscillators per mode k, each with a time-dependent frequency

$$\omega_\phi(k) = \sqrt{k^2 + \Omega_\phi(t)^2} \tag{4.52}$$

that depends parameterically on the homogeneous background field ϕ. The ψ-contribution to the Hamiltonian can directly be quantized to

$$\hat{H} = \int \mathrm{d}x \mathcal{H}_\phi + \frac{1}{2}\int_0^\infty \mathrm{d}k \left(\hat{\Pi}_A(k)^2 + \hat{\Pi}_B(k)^2 + \omega_\phi(k)^2(\hat{A}(k)^2 + \hat{B}(k)^2)\right) . \tag{4.53}$$

Numerical simulations of our dynamics will be simpler if we replace the continuum of modes obtained so far with a discrete set by introducing periodic boundary conditions in space. We therefore assume that space is compactified to a circle with circumference L. This finite size also makes the ϕ-Hamiltonian well-defined for a homogeneous ϕ. (The ϕ-Hamiltonian is described by a minisuperspace treatment, in which the averaging size L would play an important role in quantum corrections if ϕ were quantized; see [88, 89]. Here, it is relevant for a consistent technical implementation of the background.) Our values of k are then restricted to the discrete set

$$k = \frac{2\pi n}{L} \tag{4.54}$$

with a positive integer n. The Hamiltonian for discrete modes can be derived by following the previous steps but replacing $\int \mathrm{d}k$ with $(2\pi/L)\sum_n$, $\delta(k-l)$ with $(L/2\pi)\delta_{kl}$, $\tilde{\psi}(k(n))$ with $\sqrt{L/(2\pi)}\psi_n$. (These choices are not unique. We can always rescale them with field redefinition so long as the integrated quantities stay the same.)

We should also adjust the ϕ-Hamiltonian to its minisuperspace form. Starting with the original action (4.37) and introducing homogeneity such that $\frac{1}{2}\int \mathrm{d}t\mathrm{d}x\dot{\phi}^2 = \frac{1}{2}L\int \mathrm{d}t\dot{\phi}^2$, we see that the minisuperspace momentum is given by $\Pi_\phi = L\dot{\phi}$ and depends on L. The ϕ-Hamiltonian of the minisuperspace contribution therefore differs from (4.38) in that Π_ϕ^2 is replaced by Π_ϕ^2/L^2. The Hamiltonian operator that combines a minisuperspace

ϕ-contribution with a discrete set of ψ-oscillators is then

$$\hat{H} = \frac{\Pi_\phi^2}{2L} + LV(\phi) + \frac{1}{2}\sum_{n=1}^{\infty}\left(\hat{\Pi}_{A,n}^2 + \hat{\Pi}_{B,n}^2 + \omega_\phi(k(n))^2(\hat{A}_n^2 + \hat{B}_n^2)\right) + \frac{1}{2}(\hat{\Pi}_{A,0}^2 + \omega_\phi(0)^2\hat{A}_0^2)\,. \tag{4.55}$$

Compared with a continuum of modes, we have to be careful with $n = 0$ because the zero mode $\psi(0)$ is real and therefore implies only one oscillator, A_0.

4.4.2 Effective mode equations

As before, the Hamilton operator $\hat{H}$ implies a quantum Hamiltonian $H_{\rm Q} = \langle\hat{H}\rangle$ evaluated in a generic state. We evaluate this Hamiltonian to second semiclassical order and, in a first step, ignore all cross-correlations. While this assumption constitutes a restriction on the class of states that can be studied with the model, we will show that it is self-consistent. The assumption relies on the condition that the k-modes of ψ are initially decoupled in terms of moments or cross-correlations. The consequence is that cross-correlations will remain 0 throughout evolution. The assumption is quite natural in low-energy effective models and will allow us to utilize canonical coordinates for fields. (Inclusion of correlations will be discussed in later sections.)

The modes are decoupled in our classical Hamiltonian. A sufficient condition for self-consistency of our assumption at the semiclassical level is then that all quadratic moments that involve different ks remain zero if they vanish in an initial state. The relevant equations of motion are obtained from Poisson brackets, derived from (4.15), of the form $\{\Delta(s_{n_1}s_{n_2}), \Delta_H\}$, where s_N denotes any degree-of-freedom and Δ_H is a moments that appears in the quantum Hamiltonian $\langle\hat{H}\rangle$.

For a second-order expansion of $\langle\hat{H}\rangle$ with an H free of classical interactions between the modes, any Δ_H is of the form of either $\Delta(\Pi_N^2)$ or $\Delta(s_N^2)$ where each Π_N or s_N refers to a single mode. In this case, we have

$$\{\Delta(s_{n_1}s_{n_2}), \Delta(\Pi_N^2)\} = 2\Delta(s_{n_1}\Pi_N)\delta_{n_2N} + 2\Delta(s_{n_2}\Pi_N)\delta_{n_1N}\,. \tag{4.56}$$

For a cross-covariance $\Delta(s_{n_1}s_{n_2})$, we have $n_1 \neq n_2$. Therefore, any moment that may appear on the right-hand side of (4.56) with a non-zero coefficient is a cross-covariance. Analogous arguments hold for cross-covariances $\Delta(s_{n_1}\Pi_{n_2})$ and $\Delta(\Pi_{n_1}\Pi_{n_2})$ of different modes. In general, therefore, calling the set of these mixed moments of k-modes $\mathcal{M}$, Hamilton's equations generated by $\langle\hat{H}\rangle$ using the Poisson bracket for moments are necessarily of the form $\{m, \langle\hat{H}\rangle\} \propto \sum_{m'\in\mathcal{M}} a_{m'}m'$ for any $m \in \mathcal{M}$, with moment-independent coefficients $a_{m'}$. Therefore, if all $m \in \mathcal{M}$ vanish initially, they remain zero at all times in this model.

Moments of the state, on which $H_{\rm Q}$ depends, can therefore self-consistently be expressed in canonical Darboux variables by using the same mapping (4.18) known for a single degree of freedom, but applied independently to each mode. This procedure leads to

$$\begin{aligned} H_{\rm Q} &= \frac{\Pi_\phi^2}{2L} + LV(\phi) + \frac{1}{2}\sum_{n=1}^{\infty}\left(\Pi_{A,n}^2 + \Pi_{B,n}^2 + \omega_\phi(k(n))^2(A_n^2 + B_n^2)\right) + \frac{1}{2}(\Pi_{A,0}^2 + \omega_\phi(0)^2 A_0^2) \\ &+ \frac{1}{2}\sum_{n=1}^{\infty}\left(p_{A,n}^2 + p_{B,n}^2 + \frac{U_{A,n}}{s_{A,n}^2} + \frac{U_{B,n}}{s_{B,n}^2} + \omega_\phi(k(n))^2(s_{A,n}^2 + s_{B,n}^2)\right) \qquad (4.57) \\ &+ \frac{1}{2}\left(p_{A,0}^2 + \frac{U_{A,0}}{s_{A,0}^2} + \omega_\phi(0)^2 s_{A,0}^2\right) + O(\hbar^{3/2}) \end{aligned}$$

with canonical quantum degrees of freedom $(s_{A,n}, p_{A,n})$, $(s_{B,n}, p_{B,n})$, $U_{A,n}$ and $U_{B,n}$ such that

$$\Delta(A_n^2) = s_{A,n}^2 \quad , \quad \Delta(\Pi_{A,n}^2) = p_{A,n}^2 + \frac{U_{A,n}}{s_{A,n}^2} \qquad (4.58)$$

$$\Delta(B_n^2) = s_{B,n}^2 \quad , \quad \Delta(\Pi_{B,n}^2) = p_{B,n}^2 + \frac{U_{B,n}}{s_{B,n}^2}\,. \qquad (4.59)$$

All other variables in (4.57) are understood as expectation values of the basic mode, taken in the same state in which moments are computed.

Using canonical Poisson brackets for all variables in (4.57) except for the constant $U_{A,n}$ and $U_{B,n}$, we derive second-order equations of motion

$$\ddot\phi + V'(\phi) + \frac{\lambda\phi}{L}\left(\sum_{n=1}^{\infty}(A_n^2 + B_n^2 + s_{A,n}^2 + s_{B,n}^2) + A_0^2 + s_{A,0}^2\right) = 0 \qquad (4.60)$$

$$\ddot A_n + \omega_\phi^2(k(n))A_n = 0 \quad \text{and} \quad \ddot B_n + \omega_\phi^2(k(n))B_n = 0 \ (n>0) \qquad (4.61)$$

$$\ddot s_{A,n} - \frac{U_{A,n}}{s_{A,n}^3} + \omega_\phi^2(k(n))s_{A,n} = 0 \quad \text{and} \quad \ddot s_{B,n} - \frac{U_{B,n}}{s_{B,n}^3} + \omega_\phi^2(k(n))s_{B,n} = 0 \ (n>0) \qquad (4.62)$$

$$\ddot A_0 + \omega_\phi^2(k(0))A_0 = 0\,. \qquad (4.63)$$

In the first line, we have used the specific frequency (4.39).

Equations (4.60)–(4.63) are coupled and hard to solve analytically, but numerical solutions can be obtained for specific initial values. We assume a background potential $V(\phi) = -\frac{1}{2}\phi$ in what follows and choose the moments of each k-mode to correspond to the Gaussian ground state inititally, with frequency $\omega_\phi(k)|_{t=0} = \sqrt{k^2 + m^2 + \lambda\phi(0)^2}$. In

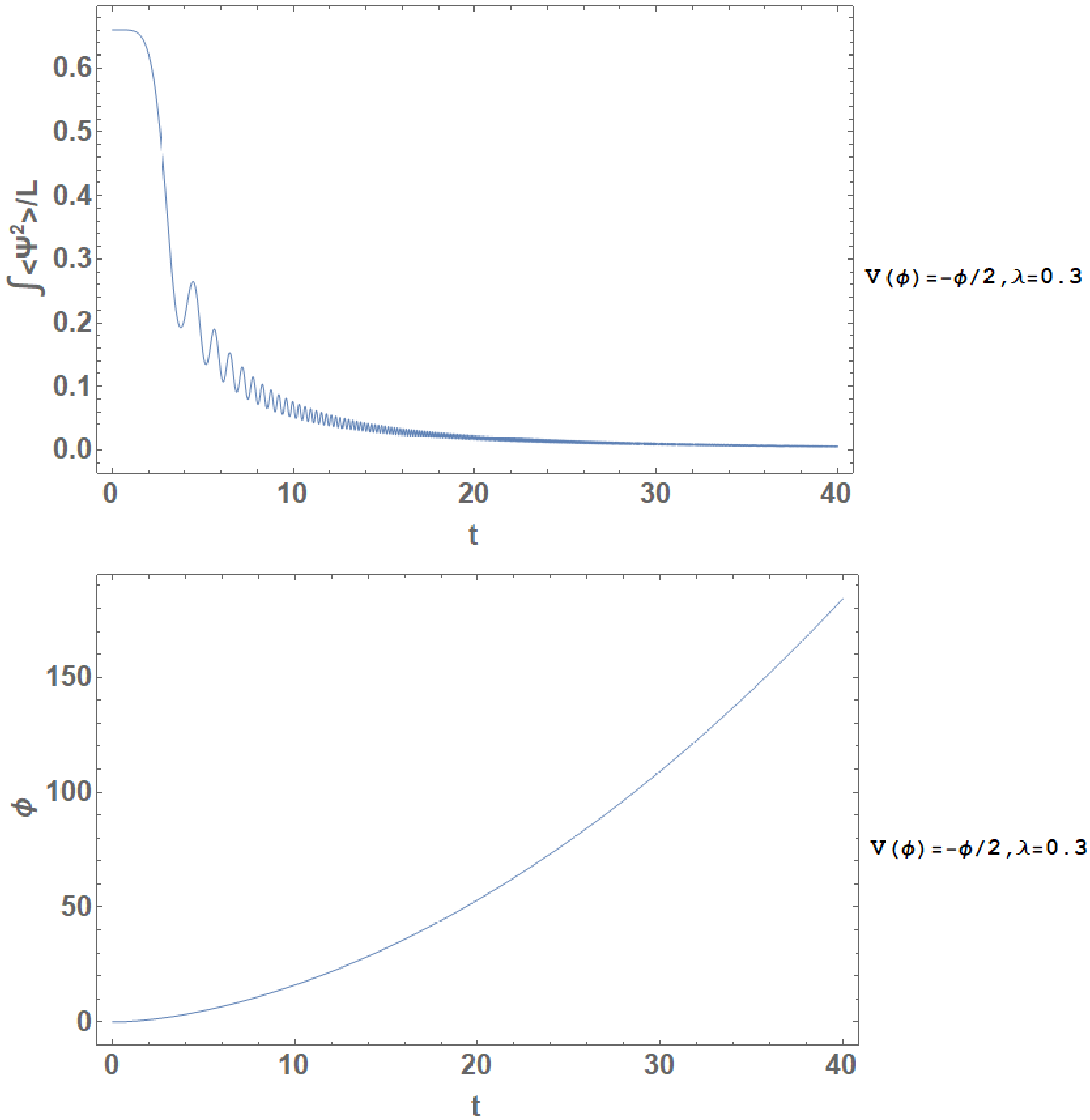

Figure 4.9. The backreaction term $\langle\hat{\psi}^2\rangle$ and background evolution $\phi(t)$ as functions of t, using $\lambda = 0.3$.

particular, $U_{A,n} = U_{B,n} = \hbar^2/4$ for all n. The initial value for ϕ that appears in the frequencies is assumed to vanish, as are all other dynamical variables.

Figures 4.9 and 4.10 show the background evolution $\phi(t)$ and the magnitude $\int \mathrm{d}x\langle\hat{\psi}^2\rangle \approx \sum_{k>0}(A_k^2 + B_k^2 + s_{A,k}^2 + s_{B,k}) + A_0^2 + s_{A,0}^2$ of back-reaction, using a momentum cutoff of $k_\Lambda = 50 \times 2\pi/L$ and the parameters $m = 0.1$, $L = 100$, $\lambda = 0.3$. Except for the momentum cutoff, these parameters match the ones used in Figures 3–5 of [80]. Their

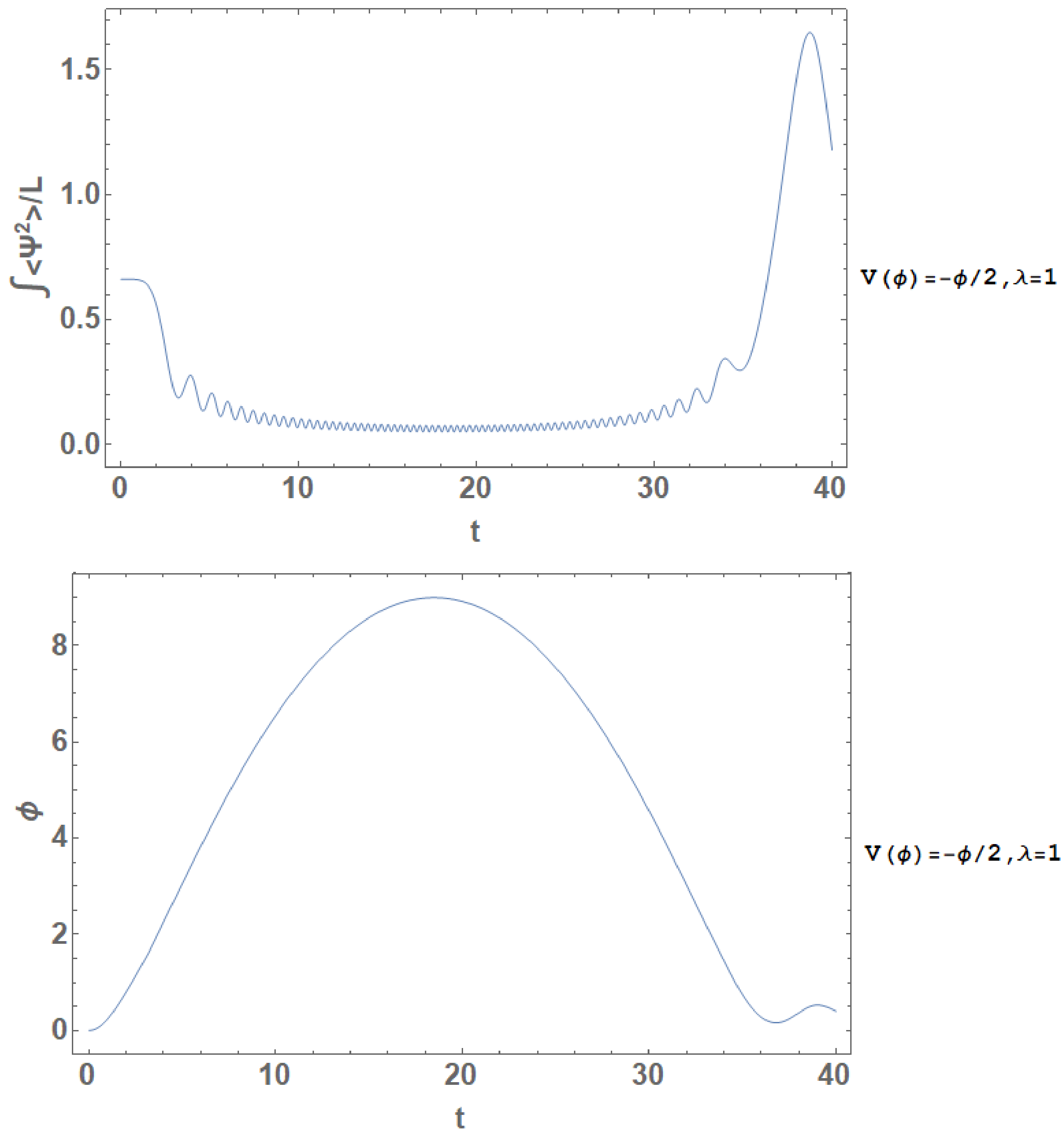

Figure 4.10. The backreaction term $\langle\hat{\psi}^2\rangle$ and background evolution $\phi(t)$ as functions of t, using $\lambda = 1.0$. With this value, compared with Fig. 4.9, back-reaction is strong enough to turn around ϕ before it grows large.

momentum cutoff is $k'_\Lambda = 4$ while for us it is $k_\Lambda = 3.96$. There are also differences in the treatment of quantum fields, which explains why numerical evolutions in these two approaches do not align precisely in quantitative terms. In addition, [80] also considers non-linear potentials $V(\phi)$ in detail, from which we refrain here in a first analysis. For smaller λ, it takes longer and longer for ϕ to turn around as a consequence of back-

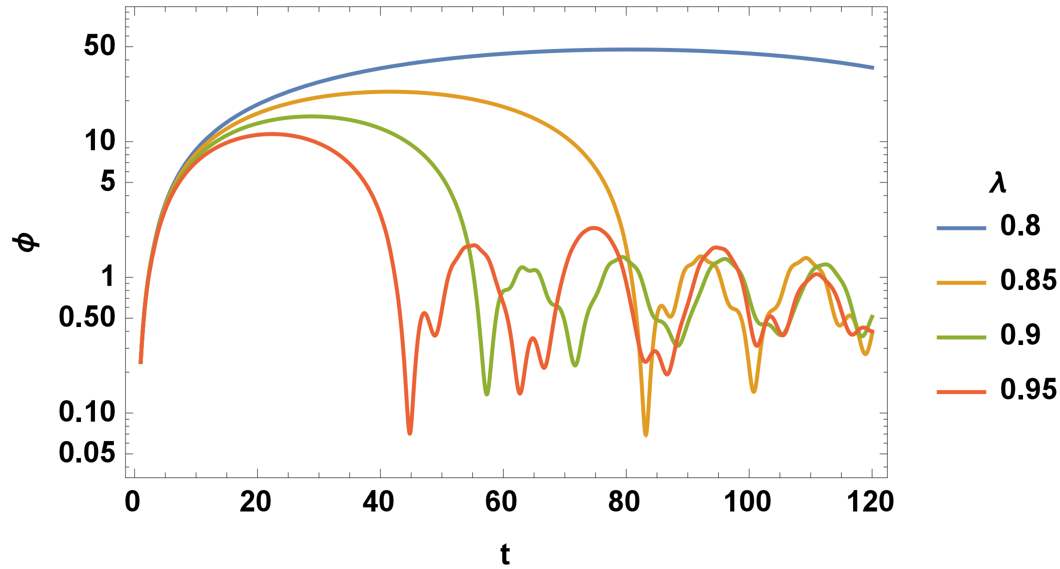

Figure 4.11. Background evolution $\phi(t)$ as functions of t for various values of λ. The turn-around of ϕ is delayed for smaller λ, implying weaker back-reaction.

reaction. The longer phase of increasing ϕ makes it difficult to resolve the turn-around numerically for very small λ, but the general trend of a delayed turn-around is illustrated in Figure 4.11.

Nevertheless, our numerical results are qualitatively comparable with those of [80]. The models are not identical because we evolve mode equations on a compact space, while [80] considers a lattice approximation to evolve spatial fields. The homogeneity assumption on the background is shared by both approaches, except for a small excursion into inhomogeneous backgrounds at the end of [80]. The underlying equations are also identical because equation (2) of [80], given by

$$\Box\phi + V'(\phi) + \lambda\langle\hat{\psi}^2\rangle\phi = 0\,, \tag{4.64}$$

is equivalent to our mode equations for a homogeneous background. In particular, our mode equations correspond to the Klein–Gordon equation

$$\Box\phi + V'(\phi) + \frac{\lambda}{L}\int \mathrm{d}x\langle\hat{\psi}^2\rangle\phi = 0 \tag{4.65}$$

with a source term for back-reaction. This equation is the same as equation (9) of [80] for a homogeneous ϕ.

4.4.3 Interactions and correlations with the background

The homogeneity problem in inflationary cosmology is concerned with the question of how it can be possible for inhomogeneity to develop out of an initial homogeneous vacuum state, subject to translation-invariant dynamics that should not break this symmetry. (See for instance [116, 117, 124, 126–128].) Effective equations for moments of a quantum field or their canonical variables can be a powerful tool to determine conditions under which inhomogeneity may build up. Similar considerations show that quantum correlations with the background may be relevant in certain situations. Specific details of such back-reaction effects may also be relevant for observational questions in cosmological eras at later times [74, 113–115].

4.4.3.1 Inhomogeneity from a back-reacting vacuum

In this section we will show how minimal ingredients from the quantum sector can give rise to inhomogeneity growth of the background field ϕ. We will assume homogeneous initial conditions for the background as part of the following initial conditions

$$\langle\psi_{k\neq 0}\rangle = 0 \quad \text{and therefore} \quad \frac{\partial}{\partial x}\langle\psi\rangle = 0 \quad \text{when} \quad t = 0\,, \tag{4.66}$$

$$\phi = \phi_0(t) + \sum_{K\neq 0}\phi_K(t)e^{iKx} \quad \text{with} \quad \phi_K(0) = 0\,, \tag{4.67}$$

where we have absorbed factors of π into the mode functions as they are not important for our present analysis.

At this point, it is important to note that choosing a specific initial state poses a condition on initial values for evolution, but it should not restrict the general dynamics of a background interacting with a quantum field. Therefore, in order to consider the possibility of background inhomogeneity being generated, the Hamiltonian should contain kinetic, potential, and interaction terms for a generic inhomogeneous background field $\phi(x,t)$, including its momentum $\Pi(x,t)$ in a canonical formulation. If we were to represent the background as a homogeneous minisuperspace model as in (4.55), background inhomogeneity would be excluded by fiat. In order to see how inhomogeneity may be generated, we should start with the more general Hamiltonian (4.53) with a field Hamiltonian density $\mathcal{H}_\phi$ and study the evolution it generates starting with symmetric initial conditions (4.66) for a homogeneous (background) initial state. The expectation value $H_{\rm eff} = \langle\hat{H}\rangle$ in a generic field state is a suitable effective Hamiltonian for this analysis. Here we encounter the important distinction between quantum field theory on a homogeneous background, which would be described by our minisuperspace background, and a more complete treatment of an interacting field and background.

In order to have growth of inhomogeneity we need to non vanishing values for Π_K, the canonical conjugate to ϕ_K. Clearly, with our initial conditions, terms that are diagonal in K-modes are not going to contribute to initial growth of Π_K. The only source of inhomogeneity growth is in the coupling term

$$\int \mathrm{d}x \lambda \phi(x)^2 \psi(x)^2 \sim \int \mathrm{d}x \sum_{k,l} \psi_k \psi_l \times \left(\phi_0^2 e^{i(k+l)x} + \sum_{K\neq 0} \phi_0 \phi_K e^{i(K+k+l)x} + \sum_{K,L\neq 0} \phi_K \phi_L e^{i(K+L+k+l)x} \right) . \tag{4.68}$$

Upon taking the expectation value of the Hamiltonian and computing $\dot{\Pi}_K$, the classical part (terms containing only $\langle\phi\rangle$ and $\langle\psi\rangle$) of (4.68) will not contribute due to our initial conditions (4.66). The terms of interest, after integrating $\int \mathrm{d}x$ and using the resulting delta functions, are

$$\begin{aligned} \dot{\Pi}_K \equiv& \{\Pi_K, \langle H \rangle\} \\ \sim& -\sum_k \left(\phi_0 \Delta(\psi_k \psi_{-k-K}) + 2 \sum_L \phi_L \Delta(\psi_k \psi_{-k-K-L}) \right) , \end{aligned} \tag{4.69}$$

where the notation $\mathcal{O} = \langle \hat{\mathcal{O}} \rangle$ is again used for linear field operators. Imposing our homogeneous initial conditions we have

$$\dot{\Pi}_K(0) \sim -\sum_k \phi_0(0) \Delta(\psi_k \psi_{-k-K}) .$$

Now we see how quantum correlations of field ψ may induce the growth of mode functions for our background field ϕ_K. Even for an initially homogeneous ϕ, the coupling $\phi^2\langle\psi^2\rangle$ implies that the background field ϕ, seen as a harmonic oscillator, may have a position-dependent frequency $\langle\psi^2\rangle$. If the initial state is completely homogeneous, the frequency $\langle\psi^2\rangle$ is homogeneous and no background inhomogeneity is generated, in accordance with the expectation based on translation-invariant dynamics. However, in cosmological space-times, this condition on the initial state is rather strong because it imposes homogeneity in all of space at an initial time, potentially far beyond the Hubble radius that would be accessible by causality. This assumption is not consistent with the generic behavior of space-time expected near a spacelike singularity such as the big bang according to the Belinskii–Khalatnikov–Lifshitz (BKL) scenario [133].

There should therefore be a certain mode number at which strict homogeneity breaks down. On these scales, the background field oscillates differently at different places and cannot stay homogeneous. One might expect that such large-scale inhomogeneity far outside the Hubble radius may not be relevant, but we will now show that moment dynamics of the field ψ implies that large-scale inhomogeneity eventually trickles down to smaller scales, at which the background inhomogeneity it generates will be relevant.

It is perhaps surprising that moment terms from a quantum field allow inhomogeneity to travel from large scales into a cosmological horizon. Our results show that this is indeed possible even in a quantization of a classically causal theory because we start with the standard covariant action of a scalar field. In fact, as our derivation shows, inhomogeneity builds up on smaller scales through a gradual process of growing cross-correlations, rather than direct propagation that would be impossible in a causal theory.

A useful interpretation of the gradual process we derived is as an effect of quantum non-locality. Moment methods make it possible to formulate effective field theories without performing a derivative expansion. This feature has been known for some time in an application to quantum mechanics, such as in quantum chemistry [79] where they give access to non-adiabatic reaction dynamics. In [42–44], it was shown explicitly that a combination of moment methods with an adiabatic approximation yields results equivalent to the low-energy effective potential [92] or higher-derivative corrections. Similarly, in an application to scalar field theories, a combination of moment methods and an adiabatic approximation [91] allows one to rederive the Coleman–Weinberg potential [90]. In the present paper, however, we have applied moment methods to quantum field theories without using an adiabatic approximation.

Instead of higher-derivative corrections, our effective Hamiltonians implement quantum corrections through coupling terms to non-classical fields as new independent degrees of freedom, such as the modes $s_{A,k}$ and $s_{B,k}$ in canonical terms or $\Delta(\psi_k\psi_l)$ directly in terms of moments. It is possible to interpret such an extended field theory as a local formulation of a non-local theory without extra degrees of freedom: If one were able to solve equations of motion for the new fields as functions of the classical fields and to insert them in the extended Hamiltonian, local terms would be replaced by integrals representing solutions of partial differential equations. It would be cumbersome to do so in practice, and it is in fact easier to analyze the local extended system. Nevertheless, the argument demonstrates that the build-up of inhomogeneity derived here is a consequence of quantum non-locality, a crucial new ingredient in early-universe cosmological models.

4.4.3.2 Field correlations

In the previous section, we have seen how quantum correlations $\Delta(\psi_k\psi_{-k-K})$ can back-react on the background and seed the growth of mode ϕ_K. In this section, we will look at the reverse—how background inhomogeneity open up interaction channels for oscillator moments that are not contained in models with strictly homogeneous background. We will show how correlations $\Delta(\psi_a\psi_b)$ (here a, b denote modes) will grow by adding a single mode to ϕ. For our purpose, the initial conditions will be

$$\begin{aligned}\phi(0,x) =&\phi_0(0) + \phi_K(0)e^{iKx} + \phi_{-K}(0)e^{-iKx}\\ \langle\psi_k\rangle(0) =&0\\ \Delta(\psi_k\psi_l) \neq&0 \quad \text{initially only for} \quad k=-l\,.\end{aligned} \tag{4.70}$$

The term that will source correlations is still the coupling term $\lambda\phi^2\psi^2$. Direct computation shows the growth of $\Delta(\psi_a\psi_b)$ is

$$\dot{\Delta}(\psi_a\psi_b) \sim \Delta(\psi_a\Pi_{-b}) + (a \leftrightarrow b)\,.$$

Given our initial conditions, the growth of the right hand side is directly dependent on mode $\phi_{\pm K}$

$$\dot{\Delta}(\psi_a\Pi_{-b}) \sim -\,\phi_K^2\Delta(\psi_a\psi_{b-2K}) - \phi_0\phi_K\Delta(\psi_a\psi_{b-K}) + (K\to -K)\;. \tag{4.71}$$

Since initial conditions (4.70) only allows $\Delta(\psi_k\psi_l)$ to be non-zero if $k+l=0$, equation (4.71) tells us that we will have

$$\begin{aligned}\ddot{\Delta}(\psi_a\psi_{b=-a\pm 2K}) \neq 0\\ \ddot{\Delta}(\psi_a\psi_{b=-a\pm K}) \neq 0\,.\end{aligned} \tag{4.72}$$

Thus, we have shown how correlations $\Delta(\psi_a\psi_{-a\pm 2K})$ and $\Delta(\psi_a\psi_{-a\pm K})$ will grow given a seed of $\phi_{\pm K}$.

Combined with the analysis in the previous section and result (4.69), we now find the following surprising interplay between quantum correlations and background inhomogeneity

$$\phi_{\pm K} \xrightarrow{\text{seed}} \Delta(\psi_a\psi_{-a\pm K}) \text{ and } \Delta(\psi_a\psi_{-a\pm 2K}) \xrightarrow{\text{backreaction}} \phi_{\pm K}, \phi_{\pm 2K} \text{ growth}\,. \tag{4.73}$$

Namely, the inhomogeneous mode ϕ_K have trickled down to ϕ_{2K} due to backreactions. Continuing the same analysis, we will find growth in more infrared modes. Even if the initial ψ_K was zero, therefore, background inhomogeneity develops in a large number of modes (all integer multiples of K) provided there was some initial inhomogeneity in just one mode, K. If the wavelength corresponding to the initial K is super-Hubble, as suggested by the BKL scenario, correlations in ψ and then inhomogeneity in ϕ trickles down to smaller wavelengths (larger k) that eventually enter the Hubble radius and become relevant for the history of our visible universe. These results provide a consistent picture of the generation of background inhomogeneity in cosmology. (We note that the situation in cosmology is different from quantum phase transitions in which structure may also form, as studied for instance in [140]. In the latter case, an infinite-volume limit is required for a mathematical description of the phase transition, and homogeneity of the pre-transition state may be assumed even in the limit. In cosmology, the BKL scenario prevents one from making the same assumption.)

4.4.3.3 Background correlations in quantum field theories

In addition to the indirect effects of field correlations on the background dynamics as just discussed, a general quantum description of background and field modes may contain direct quantum correlations between the background and the field. Such correlations are also excluded by fiat in a strict background treatment, but they can be included in an extended model using our new methods. We briefly present here such models, but for now refrain from analyzing them.

The restricted second-order Hamiltonian (4.30) suggests an extension of the background-field model to multiple canonical fields for background and oscillators. Some of the new fields represent quantum degrees of freedom for fluctuations, as before, and others cross-correlations between background and oscillator fields. Promoting all variables in (4.30) to fields, we obtain the Hamiltonian density

$$\begin{aligned}\mathcal{H} &= \frac{1}{2}\left(\Pi_\phi^2+\Pi_{\phi_1}^2+\Pi_{\phi_2}^2+(\partial_x\phi)^2+(\partial_x\phi_1)^2+(\partial_x\phi_2)^2\right)+V(\phi)\\ &\quad+\frac{1}{2}\left(\Pi_\psi^2+\Pi_{\psi_1}^2+\Pi_{\psi_2}^2+\Pi_{\psi_3}^2+(\partial_x\psi)^2+(\partial_x\psi_1)^2+(\partial_x\psi_2)^2+(\partial_x\psi_3)^2\right)\end{aligned} \tag{4.74}$$

$$+\Omega_\phi(t,x)^2\left(\psi^2+\psi_1^2+\psi_2^2+\psi_3^2\right)+\frac{1}{2}\lambda(\phi_1^2+\phi_2^2)\psi^2+2\lambda\phi\sqrt{\phi_1^2+\phi_2^2}\psi\psi_3$$

with five new fields, ϕ_1, ϕ_2, ψ_1, ψ_2 and ψ_3, and the same $\Omega_\phi(t,x)$ as before. The field ψ_3 describes background correlations.

For a homogeneous background (ϕ,ϕ_1,ϕ_2), a mode expansion is still possible, resulting in local, decoupled Hamiltonians for the modes. The potential

$$\begin{aligned}&\Omega_\phi(t)^2\left(\psi^2+\psi_1^2+\psi_2^2+\psi_3^2\right)+\frac{1}{2}\lambda(\phi_1^2+\phi_2^2)\psi^2+2\lambda\phi\sqrt{\phi_1^2+\phi_2^2}\psi\psi_3\\ =\;&\Omega_1(t)^2\psi^2+\Omega_2(t)^2(\psi_1^2+\psi_2^2)+\Omega_2(t)^2\psi_3^2+\omega(t)\psi\psi_3\end{aligned}\qquad(4.75)$$

then implies three different frequencies,

$$\Omega_1(t)^2=\Omega_\phi(t)^2+\frac{1}{2}\lambda(\phi_1(t)^2+\phi_2(t)^2)\qquad(4.76)$$

for ψ and

$$\Omega_2(t)^2=\Omega_\phi(t)^2\qquad(4.77)$$

for $\sqrt{\phi_1^2+\phi_2^2}$ as well as ψ_3, and a rotation coefficient

$$\omega(t)=2\lambda\phi(t)\sqrt{\phi_1(t)^2+\phi_2(t)^2}\qquad(4.78)$$

between ϕ and ϕ_3. As before, beat-like effects are expected by mode mixing.

If the background is homogeneous, a single correlation field ψ_3 is sufficient to extend the model to background correlations. In addition, there may be correlations between different modes of the field ψ, which at present is hard to describe in canonical quasi-classical form because a complete set of Casimir–Darboux for a quantum field remains unknown. Instead, one may construct a completely correlated background-field model by keeping the background correlation field ψ_3 in canonical quasiclassical form, while fully quantizing the field ψ or its modes as before. In (4.74), ψ, ψ_1, and ψ_2 as well as their momenta would then be replaced by a single field operator ψ and some momenta Π, while ψ_3 remains a single field of classical type. (We hope to stress that we are not re-quantizing (4.74). Instead, we are starting from a quantized system of interest, then analyze directly the quantum Hamiltonian but with the coupling term between different field mapped to a classical quantity like ψ_3 or its function.) In this formulation, hybrid methods of classical-quantum dynamics, such as [129–132], would find a useful place in cosmology.

4.5 Conclusions

We have significantly extended a back-reaction model considered in [2] in order to study the implications of backreactions and the build-up of inhomogeneity in cosmological evolution. To this end, we first showed that the formalism of [2] is equivalent to a special case of non-adiabatic semiclassical approximations obtained from the dynamics of moments parameterizing an evolving state. The restricted nature of the formalism developed in [2] implies that there are key distinctions between the two approaches that provide different advantages, depending on what physical system is being studied.

The formalism of [2] is efficient in tackling quadratic Hamiltonians and Gaussian approximations. For anharmonic systems, its approximate equations of motion remain quite simple, but there is no self-consistent way to determine whether the approximation is reliable. This shortcoming is related to a lack of physical interpretation of the degrees of freedom, ξ and χ, used in [2] to describe quantum variables. The derivation of how these variables appear in a Hamiltonian or in equations of motion requires a wave function ψ, usually assumed Gaussian, but the reduction of infinitely many quantum degrees of freedom implicitly described by the functional dependence of ψ to just two relevant ones is not systematic. In cases in which no suitable wave function is known, for instance in situations of particle creation that lead one away from a vacuum state, the predictability of the formalism remains unclear. No recipe for going beyond leading order (in some kind of loop expansion) has been developed.

Our embedding of the formalism of [2] in the systematic framework of canonical effective theory helps to make the approximations much more systematic. The quantum degrees of freedom are now physically interpreted as moments of a state, not just in Gaussian or near-Gaussian situations. There is a clear extension to higher orders of moments, mimicking the loop expansion of interacting theories. To low orders, as we have shown, explicit canonical realizations are available and can be used to extend the model of [2] to correlation degrees of freedom. As we have argued, these degrees of freedom are crucial for cosmological back-reaction in which no sharp physical separation between background and perturbations exists, owing to general covariance. Correlation degrees of freedom are then seen to give rise to new beat-like effects that may be relevant in particle production.

Extended models also show how background inhomogeneity may be generated out of a symmetric sub-horizon initial state, and what role may be played by various forms of quantum correlations. We recognized that the extension to moments as independent degrees of freedom can be interpreted as a local formulation of quantum non-locality, which in this form has not been considered before in early-universe models. By an explicit calculation of equations of motion for moments we showed that quantum non-locality implies a gradual build-up of inhomogeneity within a cosmological horizon, even if the initial sub-horizon state is completely homogeneous but inhomogeneity exists on much larger scales. This build-up of inhomogeneity is a consequence of quantum non-locality and does not violate causal propagation. The new effect is therefore able to solve the homogeneity problem of inflationary structure formation.

Chapter 5 | Conclusions, discussions, and outlooks

In this dissertation, we have shown how quantum corrections can be examined in a canonical effective setting. We have mainly focused on quantum corrected gauge symmetries in gravity, effective potentials in inflation, and backreactions between fields.

In chapter 2, we focused on the canonical structure of a gauge theory to discuss first-class constraints and their algebras; these algebras are intimately linked to the gauge symmetries of a given theory. In gravity, the quantum corrected constraint algebras are interpreted as modifications to (and possibly the destruction of) the notion of space-time covariance. As the quantization procedure of gravity remain elusive (and canonical transformations in a quantum field theory are non-trivial), the use of different basic variables can lead to seemingly different conclusions in a low energy limit, each with a different derivatives structure. (Derivative structures also affect gravity uniquely as the derivatives in the Hamiltonian are essentially what affect the $H-H$ bracket in the hypersurface deformation algebras.) By demanding the closure of constraint algebras, we constructed a strategy to find the most general form of symmetry preserving Hamiltonians up to a given derivative order. A surprising role is played by the Gauss constraint, which can be used to relate different types of derivatives to each other. (An intuition we gain is that we can expect any first-class constraint that generates effects of rotations or shifts in space-time to be able to mix derivatives.) Consequently, we can change the derivative structure and possibly induce superficially different space-time modifications. In the context of loop quantum gravity, space-time modifications can disappear when using special types of canonical pairs such as the self-dual variables. Our analysis shows that one cannot completely trust these conclusions because there are no symmetries that select out self-dual variables as the special canonical pair. This highlights the need for gauge-invariant variables and restriction to the solution space of the Gauss constraint. Only then are we able to remove ambiguities of space-time deformations arising from the Gauss constraint. Finally, using again the strategy of algebra closure, we find that generic modifications of covariance still exist.

In chapter 3 we showed how one can obtain an effective description of a quantum system by parameterizing quantum corrections with quantum fluctuations. These quantum fluctuations are naturally interpreted as moments of a quantum state. They also make up authentic degrees of freedom in phase space. So while there is only one quantum DoF, reduction to an effective semi-classical model naturally produces one classical DoF along with many parameters containing information about the state. The resulting effective theory not only produces a quantum-corrected classical equation of motion but also contains equations that describe the evolution of state-dependent parameters. However, not all quantum fluctuations are independent DoFs—a naive Lagrangian analysis would produce equations of motion that contain redundancies. Hence, a canonical analysis is in order so as to reveal the symplectic structures of the effective theory and its canonical coordinates. The search for canonical coordinates for higher-order moments or systems with multiple background DoFs is a difficult task of solving non-linearly coupled partial differential equations. But, as we have shown, it is often sufficient to find approximate solutions for cosmological applications.

In cosmology, we first applied our canonical effective methods to cosmic inflation. At the background level, instead of relying on traditional effective potentials that often neglect non-adiabatic corrections—the higher-order derivative term contributions in the quantum action—we use a real-time corrected effective potential. The real-time corrections originate from the evolution of moments, which we describe with canonical coordinates. We then analyze a Higgs-inspired inflation model. The single field model transmutes into a multi-field model, with the additional DoFs coming from quantum fluctuations. We find that the uncertainty principle naturally induces a non-zero field value for the fluctuation field. This initial value, along with the potential deformation induced by non-adiabatic corrections, ignites the slow-roll and inflationary expansion. Observables are computed both analytically and numerically. The two methods are shown to match up well. Observables such as the number of e-folds, the spectral index, and the tensor-to-scalar index are shown to depend on the non-Gaussianity of the background field. Even without excess fine-tuning, the observables conform with current observations.

Our analysis serves as a proof of concept that the often neglected higher derivative contributions in an effective action can qualitatively change an inflationary model. (One should not confuse higher-order derivatives of the effective action with ones in the original action; one can have negative-mass dimension interactions in the effective action. While we would like the original action to be renormalizable, renormalizability on its own is not something we require of the effective action since its tree-level results already describe the all-orders scattering amplitude of the original theory.) With the inclusion of (background) non-adiabatic contributions, many ruled-out inflationary models might have a second chance. However, one shortcoming of our analysis is that we have not examined in detail how the inhomogeneous perturbations of our effective model, which by construction already contain quantum corrections, are related to the quantization of inhomogeneities in traditional literature. In principle, we expect the latter to contain some overlap with the quantum corrections considered in our work. If the overlap is too large, then one should not naively apply the formulas of multi-field inflation to the calculation of observables. Instead, one ought to carry out the quantization of inhomogeneous perturbations from scratch.

As a second application of our canonical effective method, in chapter 4 we looked at quantum corrections to models that contain multiple degrees of freedom prior to quantization. These models are especially important for cosmic backreactions, where splitting of DoFs is assumed. We compare our methods to backreaction calculations that rely on Gaussian approximations. We also show one can overcome the difficulty of generalizing Gaussian approximations to include anharmonic interactions. By using the canonical mapping for multiple degrees of freedom, distinct beat-like oscillation features arise due to correlations. In a cosmological setting, these correlation-induced features can appear in the background field when backreactions from perturbations are considered. As an extension of our analysis, one can look for distinct backreaction oscillation features induced by quantum gravity. We can expect these features to manifest themselves in the form of noise. The analysis can be embedded in the framework of stochastic inflation, where noise affects observables. Traditionally, only the noise coming from short-wavelength modes of the inflaton field is considered. They are often assumed to be white noise, which differs from the model-dependent backreaction noise. Therefore, a distillation of quantum gravitational backreaction signals in this context is both interesting and possible.

Another interesting effect of backreactions emerges when we extend our analysis to include inhomogeneous fields. Specifically, we considered a two-field model with anharmonic coupling—one field plays the role of the background while the other is the to-be-quantized field. The set-up mirrors the splitting of fields in cosmology, where only the perturbations are quantized. We have shown that a positive feedback loop exists between the internal cross-correlations of the perturbation field and the inhomogeneity growth of the background field. A large-scale inhomogeneity will source internal cross-correlations, which then source inhomogeneity of a smaller scale. The existence of cosmological horizons, combined with its evolution, ensures that there are initial infrared inhomogeneities that will trickle down to the ultraviolet. This result exemplifies the genericness of structure growth—unless the entire universe, equip with causal horizons and interactions between fields, is in an exact homogeneous vacuum, inhomogeneities will grow.

Appendix
Covariance in spherically symmetric euclidean gravity

1 Restrictions on the coefficients of semi-symmetric Gaussian terms

We list the solutions to partial differential equations resulting from the $\mathcal{H}$-$\mathcal{G}$ and $\mathcal{H}$-$\mathcal{D}$ brackets. These will give us the so called *semi-symmetric Gaussian* terms. Denoting $(E^\varphi)^2 = E^{22} + E^{33}$, for β^{ij} we have

$$\begin{cases} \beta^{11} = \beta^{11}(E^1) \\ \beta^{12} = E^3 \tilde{C}_\beta(E^1) + E^2 \bar{C}_\beta(E^1) \\ \beta^{13} = E^3 \bar{C}_\beta(E^1) - E^2 \tilde{C}_\beta(E^1) \end{cases}$$

$$\begin{cases} \beta^{22} = 1/2[-8\bar{C}_{\beta^{23}}(E^1)E^{23} + (C_\Sigma(E^1) + \tilde{C}_{\beta^{23}}(E^1))E^{22} + (C_\Sigma(E^1) - \tilde{C}_{\beta^{23}}(E^1))E^{33}] \\ \beta^{33} = 1/2[8\bar{C}_{\beta^{23}}(E^1)E^{23} + (C_\Sigma(E^1) + \tilde{C}_{\beta^{23}}(E^1))E^{33} + (C_\Sigma(E^1) - \tilde{C}_{\beta^{23}}(E^1))E^{22}] \\ \beta^{23} = \tilde{C}_{\beta^{23}}(E^1)E^{23} + 2(E^{22} - E^{33})\bar{C}_{\beta^{23}}(E^1) \end{cases}$$

For γ^i we have

$$\begin{cases} \gamma^1 = \gamma^1(E^1) \\ \gamma^2 = E^3 \tilde{C}_\gamma(E^1) + E^2 \bar{C}_\gamma(E^1) \\ \gamma^3 = E^3 \bar{C}_\gamma(E^1) - E^2 \tilde{C}_\gamma(E^1) \end{cases}$$

For α^i we have

$$\begin{cases}\bar{\alpha}^1 = C_{\alpha^1}(E^1)E^\varphi \\ \bar{\alpha}^2 = (\tilde{C}_{\bar{\alpha}}(E^1)E^3 + \bar{C}_{\bar{\alpha}}(E^1)E^2)E^\varphi \\ \bar{\alpha}^3 = (-\tilde{C}_{\bar{\alpha}}(E^1)E^2 + \bar{C}_{\bar{\alpha}}(E^1)E^3)E^\varphi\end{cases}$$

$$\begin{cases}\alpha_1^1 = \alpha_1^1(E^1) \\ \alpha_1^2 = E^3\tilde{C}_{\alpha_1^2}(E^1) + E^2\bar{C}_{\alpha_1^2}(E^1) \\ \alpha_1^3 = E^3\bar{C}_{\alpha_1^2}(E^1) - E^2\tilde{C}_{\alpha_1^2}(E^1) \\ \alpha_2^1 = (E^2\tilde{C}_{\alpha_2^1}(E^1) + E^3\bar{C}_{\alpha_2^1}(E^1))\dfrac{1}{(E^\varphi)^2} \\ \alpha_3^1 = (-E^2\bar{C}_{\alpha_2^1}(E^1) + E^3\tilde{C}_{\alpha_2^1}(E^1))\dfrac{1}{(E^\varphi)^2}\end{cases} \quad \begin{cases}\alpha_2^2 = (-\tilde{C}_{\alpha_2^2}(E^1)E^{23} + \bar{C}_{\alpha_2^2}(E^1)E^{33})\dfrac{1}{(E^\varphi)^2} \\ \alpha_3^3 = (\tilde{C}_{\alpha_2^2}(E^1)E^{23} + \bar{C}_{\alpha_2^2}(E^1)E^{22})\dfrac{1}{(E^\varphi)^2} \\ \alpha_3^2 = (-\bar{C}_{\alpha_2^2}(E^1)E^{23} + \tilde{C}_{\alpha_2^2}(E^1)E^{22})\dfrac{1}{(E^\varphi)^2} \\ \alpha_2^3 = (-\bar{C}_{\alpha_2^2}(E^1)E^{23} - \tilde{C}_{\alpha_2^2}(E^1)E^{33})\dfrac{1}{(E^\varphi)^2}\end{cases}$$

For Q we have

$$\begin{cases}\bar{Q} = (E^\varphi)^2 C_{\bar{Q}}(E^1) \\ a_1 = E^\varphi C_{a_1}(E^1) \\ a_2 = \dfrac{E^3}{E^\varphi}C_{a_2}(E^1) \\ a_3 = -\dfrac{E^2}{E^\varphi}C_{a_2}(E^1)\end{cases} \begin{cases}c_1 = c_1(E^1) \\ c_2 = \dfrac{E^3}{(E^\varphi)^2}C_k(E^1) \\ c_3 = -\dfrac{E^2}{(E^\varphi)^2}C_k(E^1)\end{cases} \begin{cases}b_{11} = b_{11}(E^1) \\ b_{12} = (-c_1(E^1)E^2/2 + E^3C_b(E^1))\dfrac{1}{(E^\varphi)^2} \\ b_{13} = (-c_1(E^1)E^3/2 - E^2C_b(E^1))\dfrac{1}{(E^\varphi)^2}\end{cases}$$

$$\begin{cases}b_{22} = (E^{33}C_{b_{22}}(E^1) - 3E^{23}C_k(E^1))\dfrac{1}{(E^\varphi)^4} \\ b_{33} = (E^{22}C_{b_{22}}(E^1) + 3E^{23}C_k(E^1))\dfrac{1}{(E^\varphi)^4} \\ b_{23} = [\dfrac{3}{2}C_k(E^1)(E^{22} - E^{33}) - E^{23}C_{b_{22}}(E^1)]\dfrac{1}{(E^\varphi)^4}\end{cases}$$

We also have mixing conditions

$$\begin{cases}C_k(E^1) = -\gamma^1(E^1) = \tilde{C}_{\alpha_2^1}(E^1) \\ C_{a_2}(E^1) = -C_{\alpha^1}(E^1) \\ C_b(E^1) = -\dfrac{1}{2}\alpha_1^1(E^1) \\ C_{b_{22}}(E^1) = -\dfrac{1}{2}\bar{C}_{\alpha_2^1}(E^1)\end{cases} \quad \begin{cases}\bar{C}_{\alpha_2^1}(E^1) = -2\beta^{11}(E^1) \\ -\bar{C}_{\alpha_2^2}(E^1) = 2\tilde{C}_\beta(E^1) - \bar{C}_\gamma(E^1) \\ \tilde{C}_{\alpha_2^2}(E^1) = 2\bar{C}_\beta(E^1) + \tilde{C}_\gamma(E^1)\end{cases}$$

2 Some useful identities

In calculating the $\{H[N(x)], H[M(x)]\}$ bracket, we can often make use of anti-symmetry and integration by parts to simplify our calculations. Suppose we only have one canonical pair (K, E), then typically we have

$$H[N(x)] \sim \int \mathrm{d}x N(x)[\cdots + f(E(x), K(x))n(x) + \dots]$$

where $n(x)$ is a function of phase-space variables depending on x. Plugging this form of Hamiltonian into the Poisson bracket, we obtain the non-trivial term

$$\begin{aligned}\{H[N(x)], H[M(x)]\} \ni \int \mathrm{d}x\mathrm{d}y\{&N(x)M(y)[n(x)\{f(E(x), K(x)), \partial_y^n E(y)\}m(y)] \\ &- (N \leftrightarrow M)\}\end{aligned}$$

Here $m(y)$ is again some phase-space function, which came along with $\partial_y^n E(y)$. Denote $\dot{f}(x) \equiv \partial f(E(x), K(x))/\partial K(x)$ and $K_{NM}^{(n)}$ for the above integral term (including the $(N \leftrightarrow M)$), then for $n = 1$ we have

$$K_{NM}^{(1)} = -\int \mathrm{d}x[M'(x)N(x) - N'(x)M(x)]n(x)m(x)\dot{f}(x)\,.$$

For n=2 we have

$$K_{NM}^{(2)} = \int \mathrm{d}x[M'(x)N(x) - N'(x)M(x)][n(x)\dot{f}(x)m'(x) - m(x)(n(x)\dot{f}(x))']\,.$$

Bibliography

[1] J. L. Lehners and E. Wilson-Ewing. Running of the scalar spectral index in bouncing cosmologies. *JCAP*, 10:038, 2015.

[2] T. Vachaspati and G. Zahariade. A classical-quantum correspondence and backreaction. *Phys. Rev. D*, 98:065002, 2018.

[3] M. Bojowald, S. Brahma, D. Ding, and M. Ronco. Deformed covariance in spherically symmetric vacuum models of loop quantum gravity: Consistency in Euclidean and self-dual gravity. *Phys. Rev. D*, 101:026001, 2020.

[4] M. Bojowald, S. Brahma, S. Crowe, D. Ding, and J. McCracken. Multi-field inflation from single-field models. *JCAP*, 08:047, 2021.

[5] M. Bojowald, S. Brahma, S. Crowe, D. Ding, and J. McCracken. Quantum Higgs Inflation. *Phys. Lett. B*, 816:136193, 2021.

[6] M. Bojowald and D.Ding. Canonical description of cosmological backreaction. *JCAP*, 03:083, 2021.

[7] M. Bojowald. *Canonical Gravity and Applications: Cosmology, Black Holes, and Quantum Gravity.* Cambridge University Press, Cambridge, 2010.

[8] J. Maldacena. Non-gaussian features of primordial fluctuations in single field inflationary models. *JHEP*, 0305:013, 2003.

[9] J. D. Reyes. *Spherically Symmetric Loop Quantum Gravity: Connections to 2-Dimensional Models and Applications to Gravitational Collapse.* PhD thesis, The Pennsylvania State University, 2009.

[10] M. Bojowald, J. D. Reyes, and R. Tibrewala. Non-marginal ltb-like models with inverse triad corrections from loop quantum gravity. *Phys. Rev. D*, 80:084002, 2009.

[11] A. Kreienbuehl, V. Husain, and S. S. Seahra. Modified general relativity as a model for quantum gravitational collapse. *Class. Quantum Grav.*, 29:095008, 2012.

[12] A. Kreienbuehl, V. Husain, and S. S. Seahra. Model for gravitational collapse in effective quantum gravity. 2011.

[13] M. Bojowald, G. M. Paily, and J. D. Reyes. Discreteness corrections and higher spatial derivatives in effective canonical quantum gravity. *Phys. Rev. D*, 90:025025, 2014.

[14] S. Brahma. Spherically symmetric canonical quantum gravity. *Phys. Rev. D*, 91:124003, 2015.

[15] M. Bojowald, S. Brahma, and J. D. Reyes. Covariance in models of loop quantum gravity: Spherical symmetry. *Phys. Rev. D*, 92:045043, 2015.

[16] M. Bojowald and S. Brahma. Covariance in models of loop quantum gravity: Gowdy systems. *Phys. Rev. D*, 92:065002, 2015.

[17] M. Bojowald and S. Brahma. Signature change in loop quantum gravity: Two-dimensional midisuperspace models and dilaton gravity. *Phys. Rev. D*, 95:124014, 2017.

[18] M. Bojowald and S. Brahma. Signature change in 2-dimensional black-hole models of loop quantum gravity. *Phys. Rev. D*, 98:026012, 2018.

[19] M. Bojowald and G. M. Paily. Deformed general relativity and effective actions from loop quantum gravity. *Phys. Rev. D*, 86:104018, 2012.

[20] M. Bojowald, S. Brahma, U. Büyükçam, and F. D'Ambrosio. Hypersurface-deformation algebroids and effective space-time models. *Phys. Rev. D*, 94:104032, 2016.

[21] M. Bojowald, S. Brahma, and D.-H. Yeom. Effective line elements and black-hole models in canonical (loop) quantum gravity. *Phys. Rev. D*, 98:046015, 2018.

[22] J. Ben Achour, F. Lamy, H. Liu, and K. Noui. Polymer schwarzschild black hole: an effective metric. *Eur. Phys. L.*, 123:20006, 2018.

[23] A. Barrau, M. Bojowald, G. Calcagni, J. Grain, and M. Kagan. Anomaly-free cosmological perturbations in effective canonical quantum gravity. *JCAP*, 05:051, 2015.

[24] M. Bojowald, G. Hossain, M. Kagan, and S. Shankaranarayanan. Anomaly freedom in perturbative loop quantum gravity. *Phys. Rev. D*, 78:063547, 2008.

[25] M. Bojowald, G. Hossain, M. Kagan, and S. Shankaranarayanan. Gauge invariant cosmological perturbation equations with corrections from loop quantum gravity. *Phys. Rev. D*, 79:043505, 2009.

[26] T. Cailleteau, J. Mielczarek, A. Barrau, and J. Grain. Anomaly-free scalar perturbations with holonomy corrections in loop quantum cosmology. *Class. Quant. Grav.*, 29:095010, 2012.

[27] T. Cailleteau, A. Barrau, J. Grain, and F. Vidotto. Consistency of holonomy-corrected scalar, vector and tensor perturbations in loop quantum cosmology. *Phys. Rev. D*, 86:087301, 2012.

[28] T. Cailleteau, L. Linsefors, and A. Barrau. Anomaly-free perturbations with inverse-volume and holonomy corrections in loop quantum cosmology. *Class. Quantum Grav.*, 31:125011, 2014.

[29] J. Ben Achour, S. Brahma, and A. Marciano. Spherically symmetric sector of self dual ashtekar gravity coupled to matter: Anomaly-free algebra of constraints with holonomy corrections. *Phys. Rev. D*, 96:026002, 2017.

[30] J. Ben Achour and S. Brahma. Covariance in self dual inhomogeneous models of effective quantum geometry: Spherical symmetry and gowdy systems. *Phys. Rev. D*, 97:126003, 2018.

[31] A. Ashtekar. New hamiltonian formulation of general relativity. *Phys. Rev. D*, 36(6):1587–1602, 1987.

[32] T. Thiemann and H. A. Kastrup. Canonical quantization of spherically symmetric gravity in ashtekar's self-dual representation. *Nucl. Phys. B*, 399:211–258, 1993.

[33] H. A. Kastrup and T. Thiemann. Spherically symmetric gravity as a completely integrable system. *Nucl. Phys. B*, 425:665–686, 1994.

[34] J. F. Barbero G. Real ashtekar variables for lorentzian signature space-times. *Phys. Rev. D*, 51(10):5507–5510, 1995.

[35] J. Samuel. Is barbero's hamiltonian formulation a gauge theory of lorentzian gravity? *Class. Quant. Grav.*, 17:L141–L148, 2000.

[36] G. Immirzi. Real and complex connections for canonical gravity. *Class. Quantum Grav.*, 14:L177–L181, 1997.

[37] S. A. Hojman, K. Kuchař, and C. Teitelboim. Geometrodynamics regained. *Ann. Phys. (New York)*, 96:88–135, 1976.

[38] P. A. M. Dirac. The theory of gravitation in hamiltonian form. *Proc. Roy. Soc. A*, 246:333–343, 1958.

[39] M. Bojowald and R. Swiderski. Spherically symmetric quantum geometry: Hamiltonian constraint. *Class. Quantum Grav.*, 23:2129–2154, 2006.

[40] M. Bojowald, H. Hernández, M. Kagan, P. Singh, and A. Skirzewski. Hamiltonian cosmological perturbation theory with loop quantum gravity corrections. *Phys. Rev. D*, 74:123512, 2006.

[41] E. Alesci, S. Bahrami, and D. Pranzetti. Quantum evolution of black hole initial data sets: Foundations. *Phys. Rev. D*, 98:046014, 2018.

[42] M. Bojowald and A. Skirzewski. Effective equations of motion for quantum systems. *Rev. Math. Phys.*, 18:713–745, 2006.

[43] M. Bojowald and A. Skirzewski. Quantum gravity and higher curvature actions. *Int. J. Geom. Meth. Mod. Phys.*, 4:25–52, 2007. Proceedings of "Current Mathematical Topics in Gravitation and Cosmology" (42nd Karpacz Winter School of Theoretical Physics), Ed. Borowiec, A. and Francaviglia, M.

[44] M. Bojowald, S. Brahma, and E. Nelson. Higher time derivatives in effective equations of canonical quantum systems. *Phys. Rev. D*, 86:105004, 2012.

[45] M. Bojowald and S. Brahma. Effective constraint algebras with structure functions. *J. Phys. A: Math. Theor.*, 49:125301, 2016.

[46] J. Ben Achour, S. Brahma, J. Grain, and A. Marciano. A new look at scalar perturbations in loop quantum cosmology: (un)deformed algebra approach using self dual variables. 2016.

[47] J.-P. Wu, M. Bojowald, and Y. Ma. Anomaly freedom in perturbative models of euclidean loop quantum gravity. *Phys. Rev. D*, 98:106009, 2018.

[48] A. Henderson, A. Laddha, and C. Tomlin. Constraint algebra in lqg reloaded : Toy model of a $u(1)^3$ gauge theory i. *Phys. Rev. D*, 88:044028, 2013.

[49] A. Henderson, A. Laddha, and C. Tomlin. Constraint algebra in lqg reloaded : Toy model of an abelian gauge theory – ii spatial diffeomorphisms. *Phys. Rev. D*, 88:044029, 2013.

[50] M. Varadarajan. Towards an anomaly-free quantum dynamics for a weak coupling limit of euclidean gravity: Diffeomorphism covariance. *Phys. Rev. D*, 87:044040, 2013.

[51] C. Tomlin and M. Varadarajan. Towards an anomaly-free quantum dynamics for a weak coupling limit of euclidean gravity. *Phys. Rev. D*, 87:044039, 2013.

[52] M. Varadarajan. The constraint algebra in smolins' $g \to 0$ limit of 4d euclidean gravity. *Phys. Rev. D*, 97:106007, 2018.

[53] A. Laddha and M. Varadarajan. The diffeomorphism constraint operator in loop quantum gravity. *Class. Quant. Grav.*, 28:195010, 2011.

[54] A. Laddha. Hamiltonian constraint in euclidean lqg revisited: First hints of off-shell closure. 2014.

[55] A. Perez and D. Pranzetti. On the regularization of the constraints algebra of quantum gravity in $2 + 1$ dimensions with non-vanishing cosmological constant. *Class. Quantum Grav.*, 27:145009, 2010.

[56] P. A. R. Ade and others [Planck Collaboration]. Planck 2013 results. xxii. constraints on inflation. *Astron. Astrophys.*, 571:A22, 2014.

[57] A. A. Starobinsky. A new type of isotropic cosmological models without singularity. *Phys. Lett. B*, 91:99–102, 1980.

[58] G. Isidori, V. S. Rychkov, A. Strumia, and N. Tetradis. Gravitational corrections to standard model vacuum decay. *Phys. Rev. D*, 77:025034, 2008.

[59] M. Fairbairn, P. Grothaus, and R. Hogan. The problem with false vacuum higgs inflation. *JCAP*, 06:039, 2014.

[60] Y. Hamada, H. Kawai, and K. Y. Oda. Minimal higgs inflation. *PTEP*, 2014:023B02, 2014.

[61] F. L. Bezrukov and M. E. Shaposhnikov. The standard model higgs boson as the inflaton. *Phys. Lett. B*, 659:703–706, 2008.

[62] C. Steinwachs. Higgs field in cosmology. In S. De Bianchi and C. Kiefer, editors, *100 Years of Gauge Theory. Past, present and future perspectives.* Springer International Publishing, 2020.

[63] F. Bezrukov and M. Shaposhnikov. Standard model higgs boson mass from inflation: Two loop analysis. *JHEP*, 07:089, 2009.

[64] C. P. Burgess, H. M. Lee, and M. Trott. Comment on higgs inflation and naturalness. *JHEP*, 07:007, 2010.

[65] A. A. Starobinsky. Dynamics of phase transition in the new inflationary universe scenario and generation of perturbations. *Phys. Lett. B*, 117:175–178, 1982.

[66] A. A. Starobinsky. Stochastic de sitter (inflationary) stage in the early universe. *Lect. Notes Phys.*, 246:107–126, 1986.

[67] Y. Nambu and M. Sasaki. Stochastic stage of an inflationary universe model. *Phys. Lett. B*, 205:441, 1988.

[68] P. J. Steinhardt. Natural inflation. In *The Very Early Universe*, pages 251–266, Cambridge, UK, 1983. Cambridge University Press.

[69] A. Vilenkin. The birth of inflationary universes. *Phys. Rev. D*, 27:2848, 1983.

[70] A. D. Linde. Eternally existing self-reproducing chaotic inflationary universe. *Phys. Lett. B*, 175:395–400, 1986.

[71] A. Vilenkin. Quantum fluctuations in the new inflationary universe. *Nucl. Phys. B*, 226:527–546, 1983.

[72] E. Calzetta and B. L. Hu. Noise and fluctuations in semiclassical gravity. *Phys. Rev. D*, 49:6636–6655, 1994.

[73] M. Morikawa. Dissipation and fluctuation of quantum fields in expanding universes. *Phys. Rev. D*, 42:1027, 1990.

[74] L. Perreault-Levasseur and E. McDonough. Backreaction and stochastic effects in single field inflation. *Phys. Rev. D*, 91:063513, 2015.

[75] M. Bojowald, S. Brahma, S. Crowe, D. Ding, and J. McCracken. Quantum higgs inflation. *Phys. Lett. B*, 816:136193, 2021.

[76] R. Jackiw and A. Kerman. Time dependent variational principle and the effective action. *Phys. Lett. A*, 71:158–162, 1979.

[77] F. Arickx, J. Broeckhove, W. Coene, and P. van Leuven. Gaussian wave-packet dynamics. *Int. J. Quant. Chem.: Quant. Chem. Symp.*, 20:471–481, 1986.

[78] R. A. Jalabert and H. M. Pastawski. Environment-independent decoherence rate in classically chaotic systems. *Phys. Rev. Lett.*, 86:2490–2493, 2001.

[79] O. Prezhdo. Quantized hamiltonian dynamics. *Theor. Chem. Acc.*, 116:206, 2006.

[80] M. Mukhopadhyay and T. Vachaspati. Rolling with quantum fields. 2019.

[81] B. Baytaş, M. Bojowald, and S. Crowe. Faithful realizations of semiclassical truncations. *Ann. Phys.*, 420:168247, 2020.

[82] B. Baytaş, M. Bojowald, and S. Crowe. Effective potentials from canonical realizations of semiclassical truncations. *Phys. Rev. A*, 99:042114, 2019.

[83] M. Bojowald, D. Brizuela, H. H. Hernandez, M. J. Koop, and H. A. Morales-Técotl. High-order quantum back-reaction and quantum cosmology with a positive cosmological constant. *Phys. Rev. D*, 84:043514, 2011.

[84] V. I. Arnold. *Mathematical Methods of Classical Mechanics*. Springer, 1997.

[85] A. Cannas da Silva and A. Weinstein. *Geometric models for noncommutative algebras*, volume 10 of *Berkeley Mathematics Lectures*. Am. Math. Soc., Providence, 1999.

[86] O. Prezhdo and Yu.V. Pereverzev. Quantized hamilton dynamics. *J. Chem. Phys.*, 113:6557, 2000.

[87] C. Kühn. *Moment Closure—A Brief Review*, pages 253–271. Springer International Publishing, 2016.

[88] M. Bojowald and S. Brahma. Minisuperspace models as infrared contributions. *Phys. Rev. D*, 92:065002, 2015.

[89] M. Bojowald. The bkl scenario, infrared renormalization, and quantum cosmology. *JCAP*, 01:026, 2019.

[90] S. Coleman and E. Weinberg. Radiative corrections as the origin of spontaneous symmetry breaking. *Phys. Rev. D*, 7:1888–1910, 1973.

[91] M. Bojowald and S. Brahma. Canonical derivation of effective potentials. 2014.

[92] F. Cametti, G. Jona-Lasinio, C. Presilla, and F. Toninelli. Comparison between quantum and classical dynamics in the effective action formalism. In *Proceedings of the International School of Physics "Enrico Fermi", Course CXLIII*, pages 431–448, Amsterdam, 2000. IOS Press.

[93] A. D. Linde. Hybrid inflation. *Phys. Rev. D*, 49:748–754, 1994.

[94] J. Martin and R. H. Brandenberger. The trans-planckian problem of inflationary cosmology. *Phys. Rev. D*, 63:123501, 2001.

[95] R. H. Brandenberger and J. Martin. The robustness of inflation to changes in super-planck-scale physics. *Mod. Phys. Lett. A*, 16:999–1006, 2001.

[96] J. C. Niemeyer. Inflation with a planck-scale frequency cutoff. *Phys. Rev. D*, 63:123502, 2001.

[97] H. Kodama, K. Kohri, and K. Nakayama. On the waterfall behavior in hybrid inflation. *Prog. Theor. Phys.*, 126:331–350, 2011.

[98] S. Clesse. Hybrid inflation along waterfall trajectories. *Phys. Rev. D*, 83:063518, 2011.

[99] E. D. Stewart. Mutated hybrid inflation. *Phys. Lett. B*, 345:414–415, 1995.

[100] D. H. Lyth and E. D. Stewart. More varieties of hybrid inflation. *Phys. Rev. D*, 54:7186–7190, 1996.

[101] K. Kohri, C. M. Lin, and D. H. Lyth. More hilltop inflation models. *JCAP*, 12:004, 2007.

[102] R. Jeannerot and M. Postma. Confronting hybrid inflation in supergravity with cmb data. *JHEP*, 05:071, 2005.

[103] Cumrun Vafa. The String landscape and the swampland. 2005.

[104] E. Palti. The swampland: Introduction and review. *Fortsch. Phys.*, 67:1900037, 2019.

[105] H. Ooguri, E. Palti, G. Shiu, and C. Vafa. Distance and de sitter conjectures on the swampland. *Phys. Lett. B*, 788:180–184, 2019.

[106] S. K. Garg and C. Krishnan. Bounds on slow roll and the de sitter swampland. *JHEP*, 11:075, 2019.

[107] S. Brahma and S. Shandera. Stochastic eternal inflation is in the swampland. *JHEP*, 11:016, 2019.

[108] T. Rudelius. Conditions for (no) eternal inflation. *JCAP*, 08:009, 2019.

[109] A. Bedroya and C. Vafa. Trans-planckian censorship and the swampland. 2019.

[110] A. Bedroya, R. H. Brandenberger, M. Loverde, and C. Vafa. Trans-planckian censorship and inflationary cosmology. *Phys. Rev. D*, 101:103502, 2020.

[111] S. Brahma, R. H. Brandenberger, and D. H. Yeom. Swampland, trans-planckian censorship and fine-tuning problem for inflation: Tunnelling wavefunction to the rescue. *JCAP*, 10:037, 2020.

[112] N. Kaloper, M. König, A. Lawrence, and J. H. C. Scargill. On hybrid monodromy inflation (hic sunt dracones). 2020.

[113] R. H. Brandenberger. Back reaction of cosmological perturbations and the cosmological constant problem.

[114] R. H. Brandenberger, L. L. Graef, G. Marozzi, and G. P. Vacca. Back-reaction of super-hubble cosmological perturbations beyond perturbation theory. *Phys. Rev. D*, 98:103523, 2018.

[115] C. Armendariz-Picón. On the expected backreaction during preheating. *JCAP*, 05:035, 2020.

[116] J. J. Halliwell. Correlations in the wave function of the universe. *Phys. Rev. D*, 36:3626–3640, 1987.

[117] J. J. Halliwell. Decoherence in quantum cosmology. *Phys. Rev. D*, 39:2912, 1989.

[118] D. Polarski and A. A. Starobinsky. Semiclassicality and decoherence of cosmological perturbations. *Class. Quant. Grav.*, 13:377–392, 1996.

[119] J. Lesgourgues, D. Polarski, and A. A. Starobinsky. Quantum-to-classical transition of cosmological perturbations for non-vacuum initial states. *Nucl. Phys. B*, 497:479–510, 1997.

[120] C. Kiefer, D. Polarski, and A. A. Starobinsky. Quantum-to-classical transition for fluctuations in the early universe. *Int. J. Mod. Phys. D*, 7:455–462, 1998.

[121] C. Kiefer and D. Polarski. Emergence of classicality for primordial fluctuations: Concepts and analogies. *Annalen Phys.*, 7:137–158, 1998.

[122] A. O. Barvinsky, A. Yu. Kamenshchik, and C. Kiefer. Effective action and decoherence by fermions in quantum cosmology. *Nucl. Phys. B*, 552:420–444, 1999.

[123] A. O. Barvinsky, A. Yu. Kamenshchik, C. Kiefer, and I. V. Mishakov. Decoherence in quantum cosmology at the onset of inflation. *Nucl. Phys. B*, 551:374–396, 1999.

[124] A. Perez, H. Sahlmann, and D. Sudarsky. On the quantum origin of the seeds of cosmic structure. *Class. Quantum Grav.*, 23:2317–2354, 2006.

[125] C. Kiefer, I. Lohmar, D. Polarski, and A. A. Starobinsky. Pointer states for primordial fluctuations in inflationary cosmology. *Class. Quantum Grav.*, 24:1699–1718, 2007.

[126] A. De Unánue and D. Sudarsky. Phenomenological analysis of quantum collapse as source of the seeds of cosmic structure. *Phys. Rev. D*, 78:043510, 2008.

[127] G. León and D. Sudarsky. The slow roll condition and the amplitude of the primordial spectrum of cosmic fluctuations: Contrasts and similarities of standard account and the "collapse scheme". *Class. Quant. Grav.*, 27:225017, 2010.

[128] J. Berjon, E. Okon, and D. Sudarsky. Debunking prevailing explanations for the emergence of classicality in cosmology. 2020.

[129] E. C. G. Sudarshan. Interaction between classical and quantum systems and the measurement of quantum observables. *Pramana*, 6:117, 1976.

[130] R. Kapral. Progress in the theory of mixed quantum–classical dynamics. *Ann. Rev. Phys. Chem.*, 57:129–57, 2006.

[131] D. I. Bondar, F. Gay-Balmaz, and C. Tronci. Koopman wavefunctions and quantum–classical correlation dynamics. *Proc. R. Soc. A*, 475:20180879, 2019.

[132] F. Gay-Balmaz and C. Tronci. Madelung transform and probability densities in hybrid classical-quantum dynamics. *Nonlinearity*, 33:5383, 2020.

[133] V. A. Belinskii, I. M. Khalatnikov, and E. M. Lifschitz. A general solution of the einstein equations with a time singularity. *Adv. Phys.*, 31:639–667, 1982.

[134] T. Vachaspati and G. Zahariade. Classical-quantum correspondence for fields. *JCAP*, 09:015, 2019.

[135] M. Bojowald and A. Tsobanjan. Effective constraints for relativistic quantum systems. *Phys. Rev. D*, 80:125008, 2009.

[136] A. Tsobanjan. Semiclassical states on lie algebras. *J. Math. Phys.*, 56:033501, 2015.

[137] V. F. Mukhanov, H. A. Feldman, and R. H. Brandenberger. Theory of cosmological perturbations. *Phys. Rept.*, 215:203–333, 1992.

[138] J. M. Stewart. Perturbations of friedmann–robertson–walker cosmological models. *Class. Quantum Grav.*, 7:1169–1180, 1990.

[139] M. Bojowald. Non-covariance of the dressed-metric approach in loop quantum cosmology. *Phys. Rev. D*, 102:023532, 2020.

[140] M. Mukhopadhyay, T. Vachaspati, and G. Zahariade. Quantum formation of topological defects. *Phys. Rev. D*, 102:116002, 2020.

Vita

Ding Ding

Ding Ding was born in Zhangjiagang, China on December 12th, 1992. He obtained his bachelor's degree in physics from Lanzhou University in 2016. He joined the Ph.D. program at Pennsylvania State University in August 2016. He has been conducting research under the supervision of Prof. Martin Bojowald since 2017.