Quantum cosmological backreactions. I. Cosmological space adiabatic perturbation theory

S. Schander[1,2,*] and T. Thiemann[1,†]

[1]*Institute for Quantum Gravity, Friedrich-Alexander Universität Erlangen–Nürnberg, Staudtstrasse 7, 91058 Erlangen, Germany*

[2]*Laboratoire de Physique Subatomique et de Cosmologie, Université Grenoble–Alpes, CNRS/IN2P3, 53 avenue des Martyrs, 38026 Grenoble cedex, France*

(Received 2 July 2021; accepted 28 March 2022; published 16 May 2022)

Quantum cosmology, including quantum cosmological inhomogeneities, is a promising framework for describing the very early universe in which all degrees of freedom are being considered as dynamical and quantum. However, many previous discussions on this subject consider the quantum cosmological perturbations as test fields on a curved spacetime with effective quantum modifications and thus neglect certain interactions of the subsystems, namely the backreaction of the inhomogeneous quantum fields on the cosmological background. In a series of four papers, of which this is the first, we aim at improving on the treatment of quantum effects that arise due to backreactions between matter and geometry as well as between the cosmological perturbations and the homogeneous degrees of freedom. We employ the technique of space adiabatic perturbation theory in the form developed by Panati, Spohn, and Teufel which relies on the familiar Born-Oppenheimer approximation. We extend the scheme to quantum field theory in the cosmological perturbative setting and show that this leads to presently neglected correction terms in the quantum Friedmann equations. In this first article, we provide a detailed introduction to the iterative scheme and examine the generic challenges encountered in its application to perturbative quantum cosmology. Our results will allow for a direct comparison to cosmological observations.

DOI: 10.1103/PhysRevD.105.106009

I. INTRODUCTION

Both empirical and theoretical cosmology have experienced a substantial upswing during the past decades. On the theoretical side, the interaction of Einstein's general relativity [1,2] with the Standard Model of particle physics [3] has led to a prolific modeling of the Universe. They form the basis of the cosmological standard model which characterizes the Universe according to cosmological data as being almost spatially homogeneous and isotropic, as well as constantly expanding on its largest scales and throughout its entire history. On the empirical side, today's measurement technology permits one to register electromagnetic and gravitational radiation from astronomical objects that are up to 32 billion light-years away from us [4], and hence provide information about the Universe from about 13.4 billion years ago. The Cosmic Microwave Background (CMB) radiation offers an even more ancient relict of cosmic history from the time of decoupling [5,6]. These measurement data give an astonishingly simple account of the Universe and its history: The ΛCDM model can be described by only six parameters [7,8].

The corresponding data and numerous numerical simulations [9–13] provide evidence that the structures in our present Universe result from small inhomogeneities in the formerly very dense and hot universe [14,15] (potentially due to an inhomogeneous quantum "inflaton" scalar field that has initially filled the universe and eventually decayed into other matter species [16–20] according to the concordance ΛCDM model). For the theoretical description of these fluctuations, one employs quantum field theory on curved spacetimes (QFT on CST) [21] which is based here on the following idea: The classical degrees of freedom of general relativity plus matter split into a homogeneous background and linear perturbations thereof. After having solved the homogeneous background dynamics *independently* of the perturbations, one employs their classical "fixed" background solution in the equations of motion for the *quantum* perturbations. Most importantly, the dynamical evolution of the background has already been fixed in the first stage such that the effects of the perturbations on the background are not taken into account.

This paradigm serves as the basis for the current standard model of cosmology and has proven to be an extremely well-adapted and promising model of the Universe in many respects. Nevertheless, many unsolved questions remain, both on the theoretical level and with respect to cosmological

*susanne.schander@gravity.fau.de
†thomas.thiemann@gravity.fau.de

data. Regarding the latter aspect, we point to the open questions concerning dark matter [22], and dark energy [23,24], and the lack of a satisfying theoretical framework for these phenomena. Moreover, a tension of different measurement methods for the Hubble constant has been detected recently, which might indicate the failure of the assumptions of the ΛCDM model for the very early universe [7,25] (see for example Refs. [26,27] for a different assessment of the measured data). With a view to the inherent problems of the existing theory, we stress that *classical* relativity and the Standard *quantum* Model are mathematically incompatible [2].

It seems hence very timely to scrutinize the cosmological standard model and the assumptions on which it relies. One of the most foundational approaches for this endeavor, which would directly address the issue of incompatibility, is to develop a theory of quantum gravity, uniting the underlying ideas of classical general relativity and QFT. As it is well known, we already have a number of different approaches to quantum gravity at our disposal among which are asymptotic safety [28–31], string theory (in particular in its AdS/CFT incarnation) [32,33], causal dynamical triangulations [34,35], as well as spin foam and loop quantum gravity (LQG) [36–39] (see also Ref. [40] for an extensive overview). We emphasize that remarkable progress has been made in the theoretical and phenomenological elaboration of these theories during the past years. However, none of them has so far been able to provide a uniform model for describing the cosmological data situation, which is partly due to quantization ambiguities of the dynamics, and the tremendous nonlinearity of Einstein's field equations. Another problem which hinders progress in identifying a viable candidate for quantum gravity is the lack of relevant measurement data owing to the fact that the large Planck mass suppresses any quantum gravity effects in Earth based investigations. However, there is hope that the increasing abundance of cosmological data can improve the situation.

In order to make contact between empirical data and a fully fledged theory of quantum gravity, one would, in the first place, seek the cosmological sector of such a theory, which goes hand in hand with examining the semiclassical QFT on CST limit of the theory. Therefore, one has to identify suitable semiclassical, maybe coherent, states of the theory which have only small fluctuations off a classical (cosmological) solution, preserved under the evolution generated by the Hamiltonian, which follow the classical trajectory for reasonably long time intervals. This enterprise is of course very difficult. For example, in LQG promising coherent states have been defined [41–44], but they fail to obey all conditions of a semiclassical state, in particular for gravity coupled to matter [45,46].

Different strategies could help in this situation. First, one could implement a perturbation theory with respect to an inherent perturbation parameter. For gravity coupled to matter, a natural candidate for such a parameter would be the ratio of the gravitational Planck mass $M_{\mathrm{Pl}} = \sqrt{\hbar c/(8\pi G)} \approx 2.43 \times 10^{18}\ \mathrm{GeV}/c^2$ where G is Newton's constant and a typical Standard Model particle mass, the heaviest being the Higgs with a mass $M_{\mathrm{H}} \approx 125.09\ \mathrm{GeV}/c^2 \ll M_{\mathrm{Pl}}$. Indeed, several proposals for such a perturbative formalism have been pursued, for example within the paradigm of quantum geometrodynamics; see [47] and references therein as well as in LQG [48]. The latter, and likewise many of the other discussions of this subject, relies on a quantum gravitational Born-Oppenheimer approximation. This makes perfect sense as the latter approach also relies on the occurrence of two different mass scales in a quantum system.

Indeed, the original Born-Oppenheimer scheme [49,50] was introduced to simplify the computation of molecular spectra. At its core, it relies on exploiting different timescales within a quantum system, which are due to two very different mass scales: The nuclear degrees of freedom serve as the slow subsystem while the electrons act as the fast degrees of freedom. The very small ratio of electron and nuclei masses of about $1/2000$ together with the equipartition theorem justify the separation of the system. Intuitively, the electrons "adiabatically" adapt their fast motion to the slow motion of the nuclear degrees of freedom without much disturbance. Following the reasoning of the Born-Oppenheimer procedure, this means that a quantum state of the system dynamically remains in its initial electronic subspace up to small errors. In the leading order of that mass ratio approximation, one neglects the transition of electronic eigenstates generated by the nuclear part of the Hamiltonian.

To be very precise, this allows for an approximate solution of the full spectral problem in two steps: First, by considering the nuclear phase space variables as mere parameters, it is possible to define an "effective" *electronic* quantum problem. Solving this problem enables one to select one of the electronic quantum numbers and to project the Hamiltonian onto the corresponding quantum state. In a second step, this yields an effective Hamilton operator for the nuclei alone which usually allows for simpler solutions. In fact, one considers the electronic and the nucleonic quantum problem at two different steps of the procedure, and thereby simplifies the quest for approximate solutions to the whole problem. Such a formalism, which formally treats the problems of the gravity and the matter sector separately, could be very profitable for the problem of quantum gravity and the examination of its semiclassical limit. This is precisely the idea of the above-mentioned considerations in quantum gravity that employ the Born-Oppenheimer scheme, and for which the gravity sector appears as the adiabatic ("slow") sector while the matter part corresponds to the "fast" electronic subsystem.

Unfortunately, the Born-Oppenheimer scheme also comes with severe limitations. First, the original Born-Oppenheimer

approximation is quite restrictive in the sense that it offers only a *first* order scheme and cannot be extended to provide better higher order estimates with respect to the perturbation parameter. As mentioned above, in the case of quantum gravity this perturbative parameter would arise as the ratio of some Standard Model matter mass value M_{SM} of the system and the Planck mass M_{Pl}, which should be reasonably small ($\varepsilon^2 \coloneqq M_{\mathrm{SM}}/M_{\mathrm{Pl}} \lesssim 10^{-32}$). Besides, the Born-Oppenheimer formalism applies only to a very narrow class of quantum mechanical problems, in particular to those that admit commuting coupling operators with respect to the slow subsystem. This restriction prevents, however, the application of the scheme to certain models of quantum gravity, e.g., to LQG and as we will see to quantum cosmological perturbation theory.

Fortunately, there exists a solution provided by the scheme of space adiabatic perturbation theory (SAPT) as introduced by Panati, Spohn, and Teufel [51,52] and inspired by former work in [53–55]. The physical reasoning of this scheme is the same as of the Born-Oppenheimer approach, and, indeed, one recovers the latter at the zeroth order of the perturbation theory. Speaking in terms of the molecular example, the basic concept of SAPT consists in iteratively constructing an electronic subspace which is better adapted to the dynamics of the full Hamilton operator than the initial assumption in the Born-Oppenheimer approximation. Namely, these adapted subspaces remain invariant under the full dynamics up to perturbatively decreasing errors. In a second step, the scheme proposes to construct a unitary operator which "rotates" this subspace to a simpler reference subspace such that an effective resulting Hamiltonian acts exclusively on the nuclei. The final step of the scheme consists in the construction of this effective Hamilton operator for the nuclei which includes the backreaction from one or several of the electron energy bands. The dynamics generated by this effective Hamiltonian agrees with the dynamics of the full Hamiltonian up to the desired level of precision with respect to the perturbation parameter.

In order to avoid confusion, let us briefly interpose that the adiabatic perturbative expansion of SAPT has nothing to do with the so-called adiabatic vacua of given order encountered in QFT on CST [56,57]. The latter defines (approximate) Hadamard two-point functions in the sense of QFT on cosmological spacetimes with a fixed time dependence, while in our case the time dependence is *a priori* unspecified as we allow for backreaction between the homogeneous quantum scale factor of the cosmological background and the first order inhomogeneities in both the metric and the inflaton fields. To extract the time dependence of the *interacting* quantum system is precisely the objective of the SAPT scheme.

We also note that the "space" adiabatic scheme is more general than the "time" adiabatic one for explicitly time-dependent Hamilton operators in that it includes the latter (see, e.g., Refs. [52,58]). This becomes obvious by simply regarding the time variable as another degree of freedom in the system's phase space and treating it as a dynamical "spatial" degree of freedom. One imposes then the constraint that the sum of the Hamilton operator and the momentum conjugate to time vanishes, which basically corresponds to the time-dependent Schrödinger equation. Consequently, the theory applies to fully constrained (gravitational) systems without further ado.

Furthermore, the SAPT scheme applies more generally to a much larger class of quantum systems, in particular to those that have noncommuting coupling operators within the slow subsystem. This makes SAPT an ideal candidate for implementing the very idea of the Born-Oppenheimer formalism to quantum gravity, and hence to make progress in a formulation of a QFT on CST limit of quantum gravity.

Indeed, Stottmeister and Thiemann [59–61] could exploit all the above-mentioned advantages of the SAPT formalism in their analysis within the framework of LQG. Since in LQG, the coupling operators of gravity are mutually noncommuting as was already realized by [48], the Born-Oppenheimer scheme could not be rigorously implemented before. The idea of Stottmeister and Thiemann is to employ SAPT for deriving QFT on CST from full LQG, motivated by the idea that the SAPT scheme naturally implements a split between the gravitational and the matter degrees of freedom. Thereby, they identify the matter degrees of freedom with the fast subsystem and develop methods to implement the geometric degrees of freedom of LQG as the slow variables. The concrete implementation, however, turns out to be difficult due to the particular structure of LQG (more precisely of its underlying phase space), and ambiguities related to its quantum representation.

Despite these difficulties arising for the implementation of the SAPT method to fully fledged quantum gravity, we can hope to make progress by first considering simpler models that might eventually lighten our way to the full theory. As in many other situations, it could be helpful to analyze symmetry-reduced models as a first step. In gravity, and in particular in its application to the very early universe, the cosmological principle represents a guiding assumption in such situations. Indeed, the spatially homogeneous and isotropic Friedmann-Lemaître-Robertson-Walker (FLRW) models have been studied extensively, and its solutions are well known, at least at the classical level. To be more realistic, spatial inhomogeneities should be included right from the beginning of such an endeavor. Since, however, the inhomogeneities can be modeled as linear perturbations for a large period of the cosmic history [7], it seems reasonable to employ linear cosmological perturbation theory.

Such theories have been studied extensively at the classical level (see for example Ref. [62], going back to earlier works such as [63]) and led to a sound formulation of linear cosmological perturbation theory. The conventional approach of these cosmological perturbation theories consists in developing the field action of the model up to and

including the second order in the linear cosmological perturbations. Such a second order cosmological perturbation theory is meaningful for models with compact spatial sections (as used in this work) since the zero modes can be isolated in this case. A Legendre transformation then generates the constraints of the system. The zeroth order constraints of this formalism provide the well-known Friedmann equations. As mentioned earlier, the standard method in perturbative cosmology then suggests to first solve these homogeneous constraints independently, and then to insert these solutions into the linear equations of motion for the perturbations. If in addition quantum effects are to be taken into account, it is common to use a QFT on CST formalism.

Unfortunately, these approaches neglect backreaction effects from the perturbations on the homogeneous background. In situations in which the effects of the perturbations on the background are negligible and if the background behaves almost classically, these theories might provide realistic models for the early and medium stages of the Universe, but it is not clear whether, and to which extent, they adequately model the earliest stages of our Universe. Our application of SAPT to such models aims exactly at improving this status quo by treating both the homogeneous and the inhomogeneous degrees of freedom as dynamical and quantum, and by incorporating the resulting interactions. For this enterprise, however, the constraint analysis must be properly adapted.

To the best of our knowledge, this classical program has been carried out for the first time (to second order in the perturbations) in [64–67], and we will adopt it for our purposes. The problem and the proposed solution look as follows: In the first place, the scheme of (canonical) cosmological perturbation theory introduces canonical variables in the homogeneous and in the inhomogeneous sector which are not gauge invariant in general. In order to restore the covariance for the perturbations, the standard scheme applies transformations for the perturbative sector to gauge-invariant variables, for example the Mukhanov-Sasaki variables [68]. These transformations, however, depend also on the homogeneous variables, and hence, the new system of variables breaks the canonical structure. Therefore, the idea is to impose further transformations on the full phase space, including the homogeneous degrees of freedom, with the main intent to restore canonicity and such that the constraints still form a first class system. This provides us with a new set of phase space transformations which is canonical up to second order in the cosmological perturbations [64,67,69].

With these prerequisites, we have a suitable (almost) canonical formulation of cosmological perturbation theory in which all homogeneous and inhomogeneous degrees of freedom are equally treated as canonical and dynamical variables, and with gauge-invariant perturbation variables. This is in fact necessary for our purpose of implementing the SAPT scheme in quantum cosmology since the latter envisions a canonical quantization of the entire phase space. We are thus left with choosing a suitable quantization scheme for our models.

In this respect, as in standard quantum cosmological perturbation theory, it is advisable to choose a Fock quantization with respect to the inhomogeneities since these are treated at a linear level. For the homogeneous degrees of freedom, however, the field equations are nonlinear, and a different quantization scheme is required. This idea of using different quantization methods within quantum cosmology goes back to the hybrid quantization approach in loop quantum cosmology (LQC) [64–66]. There, the homogeneous degrees of freedom are subject to a quantization in the lines of LQC [70–72]. This scheme is a minisuperspace quantization of just the homogeneous degrees of freedom using techniques from LQG. In this series of papers, we adopt the basic idea of the hybrid quantization scheme but will not employ the LQC quantization procedure.

More precisely, we implement a standard Weyl quantization which turns out to be the most natural choice when considering SAPT methods. In fact, the formalism relies on a phase space or deformation quantization [73–76] with respect to the slow, or in our case the homogeneous, degrees of freedom. This phase space quantization scheme allows one to establish a rigorous perturbation theory with respect to the so-called adiabatic perturbation parameter due to its use of the Moyal or star product [74]. Note that this is merely a different formulation of quantum theory but leads to the very same results as the standard operator formalism. We will provide further details on phase space quantization, the star product, and its implementation into the SAPT formalism in the next section.

With such a rigorous (almost) canonical formulation of gauge-invariant cosmological perturbation theory, and a specific choice of (in fact, the most conservative) quantization scheme for such a model, we are now ready to implement the methods of SAPT into quantum cosmology. Before we come to the actual realization of the SAPT methods, let us quickly review its advantages and get to the heart of our goals and achievements. Why is SAPT such an ideal tool in order to gain insights into quantum gravity–matter systems and its inherent interactions? The idea is very simple.

In fact, SAPT and our extensions thereof provide an unambiguous iterative scheme with respect to the inverse Planck mass that provides us with a series of effective quantum Hamilton operators, respectively constraints, for full quantum cosmological perturbation theory but whose structure is considerably simpler than the original problem. While these effective Hamiltonians manifestly take the interactions between the two subsystems into account (in our case, between the homogeneous and the inhomogeneous sectors), the solutions of this operator should be

much easier to determine, and most importantly approximate the solutions of the original problem up to an iteratively small error. Thereby, it proceeds similarly to the Born-Oppenheimer scheme as it provides us in a first step with an effective Hamiltonian for the homogeneous degrees of freedom that includes the corresponding backreaction of the cosmological quantum perturbations. Such backreaction has, however, been neglected in many of the studies pursued so far in quantum cosmological perturbation theory.

In fact, we point to several seminal works within quantum cosmological perturbation theory that aim at describing the interactions in the very early universe. The idea of those approaches is to capture the quantum fluctuations of the homogeneous subsystem in effective background metric and matter fields, and to pursue then the standard analysis of (quantum) field theory on curved spacetimes for the cosmological perturbations. We thereby refer to the dressed metric approach [77], the rainbow metric approach [78], the deformed algebra approach [79], and the hybrid approach [66] in LQC. While these approaches represent an important step toward understanding the primordial universe, they do not take the backreactions as defined above into account. Besides, they rely on introducing certain (semiclassical) ansatz states in order to reduce the homogeneous quantum sector to an effectively classical theory (e.g., by computing expectation values of the Hamilton constraint). Most important, the perturbations then propagate as test fields on the previously determined homogeneous background without affecting the latter. Let us also point out that it is possible to choose a background different from the solution of the homogeneous cosmology, which then contains backreaction in a certain sense.

To the best of our understanding, the aforementioned methods, however, do not incorporate backreaction effects from the perturbations onto the homogeneous degrees of freedom in the sense of a Born-Oppenheimer approximation. Also, to the best of our understanding, the various assumptions that went into these approximations are not easy to control; see Refs. [64,80] for a detailed discussion of these assumptions in the hybrid and the dressed metric approach. In this series of papers, we would like to convince the reader that SAPT provides powerful tools for achieving just that. The basic idea is that the homogeneous modes can be associated with the slow degrees of freedom while the inhomogeneous ones play the role of the fast degrees of freedom. Roughly, this happens because the homogeneous mode, being the integral (or sum) over the inhomogeneous modes, corresponds to a center of mass mode. It therefore comes with the large total mass of the system rather than the small individual ones as will be explained in Sec. III B of this paper.

As outlined before, this paper starts in Sec. II by providing the reader with a detailed introduction into the underlying mathematical structure of SAPT, the necessary tools for its application, and a presentation of the scheme itself. We will also explicitly perform the iterative construction and thereby show that the computations are consistent order by order. In the companion papers, we apply SAPT to a toy model consisting of two oscillators and to a first cosmological purely homogeneous model [81]. In [82], we consider an inhomogeneous cosmological model in which the geometry is purely homogeneous, supplemented by deparametrizing dust and scalar field perturbations. Finally in [69], we are ready to apply SAPT to the standard paradigm of gauge-invariant quantum cosmological perturbation theory. The series of papers hence shows that the methods of SAPT, initially established for *unconstrained* quantum systems, can be extended to *constrained* Hamiltonian systems, and most importantly to problems in quantum cosmology. Our results provide the possibility to unambiguously include the effects of backreaction into quantum cosmology and, eventually, to take all interactions between the different subsystems into account. As the theory naturally includes a split between the homogeneous and the inhomogeneous variables, it is ideally suited to investigate the semiclassical limit of quantum cosmology in a second step. Let us also point to a concise summary of our findings complemented by short overviews of the status of backreaction in classical and semiclassical cosmology in [83].

On our way to a rigorous implementation of SAPT into (inhomogeneous) quantum cosmology, we have encountered several challenges. These will be addressed in great detail in Sec. III of this paper. In summary, these are as follows:

(1) Cosmological perturbation theory is a field theory and hence encounters an infinite number of degrees of freedom while SAPT has *a priori* been developed for quantum systems with a finite number of degrees of freedom. Rather innocent looking assumptions such as that the Hilbert space has the form of a tensor product are no longer granted to make any sense for the quantum field case. This was already remarked in [59–61]. A solution can be provided by employing the almost canonical transformations introduced in [64,67] or modifications thereof.

(2) Many of the theorems proven in [51,52] rely on the assumption that the Hamiltonian satisfies certain conditions regarding its boundedness and its form with respect to the variables of the slow sector. The Hamiltonians in cosmology do not meet these conditions without further ado. We introduce a specific cutoff technique inspired by [51] that alleviates the problem, however, by introducing a physically different model.

(3) After applying the transformations of variables in order to obtain gauge-invariant variables at the perturbative level as well as a canonical system up to second order in the perturbations, we obtain new effective mass terms for the perturbation fields.

These mass squared terms can generically depend on negative powers of the configuration degrees of freedom and of their canonical momenta, and may not be positive definite. Interestingly, one already encounters such issues in standard perturbative cosmology when using gauge-invariant inhomogeneities. This raises domain issues both in the inhomogeneous as well as in the homogeneous sectors. We will propose several strategies for how to circumvent these problems. In this respect, we emphasize once more that *the occurrence of indefinite mass squared functions is not due to the SAPT scheme itself, but already appears in standard gauge-invariant perturbation theory* [62], in particular when one employs Mukhanov-Sasaki variables.

(4) While in (unconstrained) quantum mechanics, the adiabatic perturbation parameter gives directly a rise to two very different timescales for the two subsystems, respectively, and such a conclusion cannot be drawn in a constrained situation. There is no *a priori* notion of time. In contrast to the familiar former case, we cannot rely on (i) the equipartition theorem [84], (ii) the ergodicity assumption, and finally (iii) on the assumption that the Hamiltonian is of second order in the momenta. If this were the case, the time average would equal the phase space average due to ergodicity in say the canonical ensemble. Hence, the time averages of all the kinetic terms in the Hamiltonian have the same magnitude such that the light degrees of freedom are much faster than the heavy ones in average. However, as we will argue in Sec. III, the constraint itself dictates a corresponding behavior in certain regions of phase space.

With these introductory remarks, let us summarize the subsequent architecture of this paper: In Sec. II, we present the space adiabatic perturbation program in a self-contained fashion for the case of a finite dimensional phase space. This also serves to introduce a simple notation aiming at the highest possible transparency. In Sec. III, we address the conceptual and mathematical complications when applying SAPT to quantum cosmological perturbation theory. We comment on the possible solutions, thus preparing the ground for the remaining papers of this series. In Sec. IV, we summarize and give an outlook to the applications discussed in the other papers of this series.

II. BASICS OF SPACE ADIABATIC PERTURBATION THEORY

In Sec. II A, we introduce our notation and motivate the use of operator-valued symbol functions for the SAPT scheme. We closely follow the original work by Panati, Spohn, and Teufel [51] and the more detailed book by Teufel [52]. In Sec. II B, we reproduce the conditions for its application. In Sec. II C, we derive the essential inductive formulas underlying the adiabatic expansion.

A. Notation and symbol functions

In a nutshell, SAPT provides perturbative equations of motion for coupled quantum systems in order to compute the approximate energy spectra and the eigensolutions of these systems. The physical basis of the scheme is inspired by the conventional Born-Oppenheimer approximation, and likewise, the scheme splits the system into a "fast" and a "slow" subsystem. In this work, we denote the set of phase space variables associated with the fast subsystem by (x, y) while the slow phase space variables are labeled by (q, P). It will be sufficient to consider a simple four-dimensional phase space, $\Gamma = \Gamma_s \times \Gamma_f = \mathbb{R}^4$ where Γ_s and Γ_f denote the slow and the fast subspaces, respectively. The generalization to higher dimensional phase spaces proves to be straightforward. We briefly mention that to a certain extent, it is also possible to generalize the scheme to finite dimensional phase spaces which are not vector spaces [59–61]. For the presented model, the phase space structure is simply provided by the only nonvanishing Poisson brackets $\{q, P\} = \{x, y\} = 1$.

As a first step, the perturbative scheme requires us to define a dimensionless, perturbative parameter which we refer to as the "adiabatic" parameter ε. In the standard Born-Oppenheimer theory for molecules, this parameter is given by the mass ratio of the light electron mass m_e, and the heavy nuclei mass M_n, namely by $\varepsilon^2 := \frac{m_e}{M_n}$. This parameter can, however, be defined as any dimensionless composition of coupling constants or mass values of the respective subsystems. The classical Hamilton function providing the dynamics of the theory will be denoted by $H(q, x, P, y)$ and is supposed to be a smooth function of the phase space variables.

We then proceed by formulating a quantum theory of the system, labeling the operators associated with the slow subsystem by hats, i.e., by $(\hat{q}, \hat{P})$ while bold letters stand for the operators of the fast subsystem $(\boldsymbol{x}, \boldsymbol{y})$. The slow quantum subsystem supports a Schrödinger representation on the Hilbert space $\mathcal{H}_s := L^2(\mathbb{R}, \mathrm{d}q)$. Similarly, we describe the fast quantum system with a Schrödinger representation on the Hilbert space $\mathcal{H}_f := L^2(\mathbb{R}, \mathrm{d}x)$. For both subsystems, position and momentum operators shall act on the space of Schwartz functions in the well-known way and satisfy the formal commutation relations $[\boldsymbol{x}, \boldsymbol{y}] = i\mathbf{1}_f$, $[\hat{q}, \hat{P}] = i\hat{1}_s$ where we set $\hbar \equiv 1$. The tensor product of the Hilbert spaces $\mathcal{H} := \mathcal{H}_s \otimes \mathcal{H}_f$ models the total Hilbert space where "$\otimes$" denotes the tensor product of Hilbert spaces. The simplest form of a bounded operator on this Hilbert space is given as $\hat{A} \otimes \boldsymbol{B}$, for $\hat{A} \in \mathcal{B}(\mathcal{H}_s)$ and $\boldsymbol{B} \in \mathcal{B}(\mathcal{H}_f)$. The Weyl elements of the canonical position and momentum operators become lifted to operators on the

full Hilbert space by tensor multiplying with the respective other unity operator.

In order to make SAPT work at the technical level, the scheme resorts to a phase space or deformation quantization procedure with respect to the slow subsystem [76,85–87]. We emphasize that this representation is physically equivalent to the well-known Hilbert space representation of quantum mechanics. It assigns a phase space function, denoted as a "symbol" function, to a classical observable instead of associating with an operator on some dense domain of the Hilbert space. Indeed, the whole calculus for the quantum theory remains on the classical phase space, for which one replaces the usual operator product by a noncommuting Moyal or star product [73,74] on the space of phase space functions. This replaces then the commutative pointwise multiplication of classical phase space functions. As in standard quantum theory, it is necessary to choose an ordering prescription for the quantum operators, and we will stick here to the symmetric Weyl ordering procedure [75,88].

To make this more precise, let us quickly recapitulate the relation between phase space and Hilbert space quantizations using the Weyl elements of the slow quantum theory and Fourier transforms. Before we move on to this, we wish to clarify that in the present case, we are dealing with a coupled system and that we retain the Hilbert space representation for the fast system as introduced above. As a consequence, the theory deals with phase space functions on Γ_s *with values in the bounded operators on* $\mathcal{H}_f$, for example $\boldsymbol{B}(q,P) \in C^\infty(\Gamma_s, \mathcal{B}(\mathcal{H}_f))$. Intuitively, a symbol $(q,P) \mapsto \boldsymbol{A}(q,P)$ can be considered as arising from a function $A(q,P,x,y)$ on the total phase space by just quantizing the fast sector and choosing an operator ordering to obtain $\boldsymbol{A}(q,P) \coloneqq \boldsymbol{A}(q,P,\boldsymbol{x},\boldsymbol{y})$.

If we now want to clarify the relation between these semiclassical symbol functions that are subject to a phase space quantum formalism with respect to the slow subsystem and a pure operator representation of the theory, we consider a complete Hilbert space representation of a quantum operator associated with a classical phase space function $f \in C^\infty(\Gamma)$. It can be obtained by additionally Weyl quantizing the corresponding symbol function $\boldsymbol{f}(q,p)$ with respect to the slow system; i.e., we write $\hat{\mathcal{W}}(\boldsymbol{f}(q,P)) =: \hat{\boldsymbol{f}}$. These are then operators on the total product Hilbert space. Now, it is essential to note that there is a peculiarity regarding the Weyl quantization for SAPT. As we know, the momentum operator translates into a derivation with an additional factor of $\hbar$ (which we set directly to unity) in the conventional Schrödinger representation. However, the structure of the system, in particular the concrete form of the Hamilton operator, suggests to define a rescaled momentum operator, namely $p \coloneqq \varepsilon P$. Thereby, the canonical commutation relation picks up an additional factor of ε such that $[\hat{q},\hat{p}] = i\varepsilon \hat{1}_s$. Also the Weyl elements and in particular the star product are affected by this, as we will see in the following [52]. For this, let us define the Weyl elements of the slow quantum theory which are the representations of the Weyl group in Schrödinger representation labeled by two reals $(k,l) \in \mathbb{R}^2$ and given with the correct ε rescaling here by [75],

$$\hat{\mathcal{W}}(k,l) \coloneqq \exp\left(i\frac{k\hat{q}+l\hat{p}}{\varepsilon}\right), \tag{1}$$

where (k,l) have the dual dimension of (q,p) so that the respective products are dimensionless. Using these Weyl elements, it is possible to write the totally symmetric quantization of the symbol function $\boldsymbol{A}(q,p) \in C^\infty(\mathbb{R}^2, \mathcal{B}(\mathcal{H}_f))$ in the form [52]

$$\hat{\mathcal{W}}(\boldsymbol{A}) = \hat{\boldsymbol{A}} = \frac{1}{(2\pi\varepsilon)^2}\int_{\mathbb{R}^4} \mathrm{d}k\mathrm{d}l\mathrm{d}q\mathrm{d}p\boldsymbol{A}(q,p) \times \exp\left(i\frac{k(\hat{q}-q)+l(\hat{p}-p)}{\varepsilon}\right). \tag{2}$$

The relation to the standard Hilbert space representation of quantum operators arises by using the Baker-Campbell-Hausdorff formula and by employing that the translation of a wave function can be written as $\psi(q-l) = \exp(-i\frac{l\hat{p}}{\varepsilon})\psi(q)$. Indeed, it holds true that [52]

$$(\hat{\boldsymbol{A}}\psi)(q) = \frac{1}{(2\pi\varepsilon)}\int_{\mathbb{R}^2} \mathrm{d}k\mathrm{d}\tilde{q}\boldsymbol{A}\left(\frac{1}{2}(q+\tilde{q}),k\right) \times \exp\left(i\frac{k(q-\tilde{q})}{\varepsilon}\right)\psi(\tilde{q}), \tag{3}$$

for a Schwartz function $\psi \in \mathcal{S}(\mathbb{R})$. Obviously, this equation provides a direct link between quantum observables considered as operators on the total Hilbert space of the theory (the left-hand side) and the symbol function subject to the slow phase space quantization scheme (on the right-hand side). Since this association is unique, the quantization schemes are indeed equivalent.

As it can be shown [89], the representation of operators as phase space symbols gives rise to an operator product analog on phase space. As indicated above, this pullback of the operator product to phase space is denoted as a star or Moyal product, and we will label it by "$\star_\varepsilon$." The product is a bilinear operator and has the form of an asymptotic expansion in the adiabatic parameter ε [52]. As should be clear from definition (3), the identification of phase space symbols and operators, and hence the star product, only applies to a certain class of symbol functions [89]. In particular, one has to guarantee the convergence of the above integral.

Besides, the (star) multiplication, transposition, and adjoints of the symbol functions should evaluate to well-defined and closed operations. Indeed, the theory of

pseudodifferential calculus [90–92] (and references therein) addresses the definition of such classes of symbol functions [89]. The most commonly used symbol classes go back to Hörmander [91,92]. These originally scalar-valued function classes extend directly to the operator-valued case [52]. Here, we use the class $S^m_\rho(\Gamma_s, \mathcal{B}(\mathcal{H}_f)) =: S^m_\rho$ with $m \in \mathbb{R}$ and $0 \leq \rho \leq 1$. Then, a phase space symbol function $\boldsymbol{A}(q,P)$ belongs to S^m_ρ if for every $\alpha, \beta \in \mathbb{N}$, there exists a positive constant $C_{\alpha,\beta}$ such that for every $P \in \mathbb{R}$ [52],

$$\sup_{q\in\mathbb{R}} \|(\partial_q^\alpha \partial_p^\beta \boldsymbol{A})(q,P)\|_{\mathcal{B}(\mathcal{H}_f)} \leq C_{\alpha,\beta} \langle P\rangle^{m-\rho|\beta|}, \tag{4}$$

where $\langle P\rangle := (1+P^2)^{\frac{1}{2}}$. Besides, we assume that the symbols have a common, invariant, and dense domain for all $(q,P) \in \Gamma_s$ and are smooth in (q,P).

This being said, we can give a precise definition for the symmetric Moyal product applied to the symbol functions $\boldsymbol{B} \in S^{m_1}_\rho$ and $\boldsymbol{C} \in S^{m_2}_\rho$ by [52]

$$(\boldsymbol{B} \star_\varepsilon \boldsymbol{C})(q,p) = \sum_{k=0}^{\infty} \frac{(\frac{i\varepsilon}{2})^k}{k!} (\partial_q\partial_\xi - \partial_\rho\partial_p)^k \boldsymbol{B}(q,p)\cdot \boldsymbol{C}(\rho,\xi)\Big|_{\rho=q,\xi=p} =: \boldsymbol{D}(q,p). \tag{5}$$

Then, the symbol function $\boldsymbol{D}$ is in $S^{m_1+m_2}_\rho$, and it corresponds to a well-defined operator $\hat{\boldsymbol{D}}$ in Hilbert space representation. Note that the star product is associative but not commutative. Besides, the corresponding commutator

$$[\boldsymbol{B},\boldsymbol{C}]_{\star_\varepsilon} = \boldsymbol{B}\star_\varepsilon \boldsymbol{C} - \boldsymbol{C}\star_\varepsilon \boldsymbol{B} = [\boldsymbol{B},\boldsymbol{C}] + \mathcal{O}_0(\varepsilon) \tag{6}$$

is just the usual commutator of the symbol functions plus corrections in ε, and the former is of zeroth order in ε if the symbols are so. We emphasize that this commutator is not vanishing in general, in contrast to the classical (standard) case of *scalar* deformation quantization.

Eventually, Weyl quantization thus serves several purposes at the same time: First, it allows us to write any operator on the full Hilbert space $\mathcal{H}_s \otimes \mathcal{H}_f$ in the form of a slow phase space integral over operators of the form $A_s(q,p)\boldsymbol{B}_f(q,p) \in S^m_\rho(\Gamma_s, \mathcal{B}(\mathcal{H}_f))$ where m is in the reals and $0 \leq \rho \leq 1$. This allows one to simplify the full spectral problem as soon as the spectral problem of the operator $\boldsymbol{B}_f(q,p)$ is known in closed form. Second, the Moyal product admits a systematic power expansion in terms of the adiabatic parameter ε, and consequently enables us to set up a perturbative diagonalization scheme.

In this respect, let us stress once more that the adiabatic perturbative expansion of SAPT has nothing to do with the so-called adiabatic vacua of a given order encountered in QFT on CST [56,57]. The latter defines (approximate) Hadamard two-point functions in the sense of QFT on cosmological spacetimes with a fixed time dependence.

B. System requirements

In what follows, we describe the system requirements proposed in [51] in order to ensure convergence of the adiabatic perturbation series with respect to an appropriate topology. One of the requirements, in particular the boundedness of the operators relevant for the fast sector, is not met in the physical situation that we envisage. Yet we discuss these system requirements here for the following reasons: First, for completeness, and second, because by using an energy cutoff, one can easily meet the system requirements which, thus, together with the adiabatic theorem serves to motivate the SAPT program. In the actual unbounded case, we can still use the SAPT scheme to provide a formal power expansion. However, the convergence of the adiabatic series (say in the weak operator topology) then must be analyzed in an additional step. As we consider only first and second order adiabatic corrections in this series of papers, we will not be concerned with convergence issues and rather leave them for future investigations. With these cautionary remarks out of the way, we reproduce the conditions accounted for in [51], which split into four categories:

(C1) *The state space* of the system can be decomposed as

$$\mathcal{H} = L^2(\mathbb{R}) \otimes \mathcal{H}_f = L^2(\mathbb{R}, \mathcal{H}_f), \tag{7}$$

where one identifies the slow Hilbert space $\mathcal{H}_s$ with $L^2(\mathbb{R})$ and $\mathcal{H}_f$ must be separable; i.e., it contains a dense countable subset.

(C2) *The Hamilton operator* $\hat{\boldsymbol{H}}$ of the system is the Weyl quantization of a symbol function $\boldsymbol{H}(q,p) \in S^m_\rho(\Gamma_s, \mathcal{B}(\mathcal{H}_f))$ where the Weyl quantization refers to the slow subsystem. The function $\boldsymbol{H}(q,p)$ has values in the space of bounded *self-adjoint* operators on $\mathcal{H}_f$ and constitutes an asymptotic expansion in ε,

$$\boldsymbol{H}(q,p) \asymp \sum_{i=0}^{\infty} \varepsilon^i \boldsymbol{H}_i(q,p). \tag{8}$$

(C3)γ *Gap condition*. For any fixed $(q,p) \in \mathbb{R}^2$, the spectrum $\sigma(q,p)$ of the zeroth order symbol $\boldsymbol{H}_0(q,p)$ of the total Hamiltonian $\boldsymbol{H}(q,p)$ contains at least one isolated subset $\sigma_\nu(q,p)$ labeled by one fixed quantum number $\nu \in \mathbb{N}$, which is uniformly separated from the remainder $\sigma_{\text{rem}}(q,p) := \sigma(q,p)\backslash\sigma_\nu(q,p)$. In particular, the minimal distance between the elements of σ_ν and the remainder is

nonzero for every single $(q,p)\in\mathbb{R}^2$. More precisely, we can define two continuous "enclosing" functions $f_\pm:\mathbb{R}^2\to\mathbb{R}$ with $f_-\le f_+$ such that

(G1) For every $z\in\mathbb{R}^2$, the spectral component $\sigma_\nu(q,p)$ is entirely contained in the interval $I(q,p):=[f_-(q,p),f_+(q,p)]$.

(G2) The distance between the remainder $\sigma_{\rm rem}(q,p)$ and the interval $I(q,p)$ is uniformly bounded away from zero, i.e., $\mathrm{dist}[\sigma_{\rm rem}(q,p),I(q,p)]\ge C_{\rm g}\langle p\rangle^{\gamma}$.

(G3) The width of $I(q,p)$ is uniformly bounded: $\sup_{(q,p)\in\mathbb{R}^2}|f_+(q,p)-f_-(q,p)|\le C_{\rm w}<\infty$.

(C4) *Convergence condition*. If the system satisfies the gap condition $(\mathrm{C3})_\gamma$ for some $\gamma\in\mathbb{R}$, the Hamilton symbol $\boldsymbol{H}$ must be in S^γ_ρ. If $\rho=0$, also γ must vanish. If $\rho>0$, γ can be any real number but $\hat{\boldsymbol{H}}$ must be essentially self-adjoint on $\mathcal{S}(\mathbb{R},\mathcal{H}_{\rm f})$.

For the systems considered here, condition (C1) is trivially satisfied by definition. Regarding condition (C2), we note that the requirement of boundedness is violated for many interesting systems. However, it is often possible to adapt the Hamiltonian by cutoff strategies [51,81]. Condition (C3) requires the principal symbol $\boldsymbol{H}_0(q,p)$ to be diagonalizable on the Hilbert space $\mathcal{H}_{\rm f}$ and to have (q,p)-pointwise discrete eigenvalues.

This assumption is indeed essential for the scheme as it yields an orthonormal basis of eigensolutions in $\mathcal{H}_{\rm f}$ for the Hamiltonians $\boldsymbol{H}_0(q,p)$ for every point $(q,p)\in\Gamma_{\rm s}$. We henceforth denote this eigenbasis by $\{\xi_{n,a}(q,p)\}_{n\in\mathbb{N},a=1,\ldots,d_n}$, where a is a degeneracy label and d_n the associated degree of degeneracy. We assume that d_n is a constant with respect to (q,p). The eigenvalue equation for $\boldsymbol{H}_0$ is accordingly given by

$$\boldsymbol{H}_0(q,p)\xi_{n,a}(x,q,p)=E_n(q,p)\xi_{n,a}(x,q,p),\qquad(9)$$

where $\{E_n(q,p)\}_{n\in\mathbb{N}}$ constitutes the spectrum of $\boldsymbol{H}_0(q,p)$ for any of the points $(q,p)\in\mathbb{R}^2$. We emphasize that in Schrödinger representation, the functions $\xi_{n,a}$ are vectors in $L^2(\mathbb{R},\mathrm{d}x)$ and have hence a representation as x-dependent functions. The (q,p)-dependence here should be seen as a mere parametric dependence. For every fixed $(q,p)\in\mathbb{R}^2$, there is an independent eigenvalue problem, and we assume that the correspondent eigenvalues $\{E_n(q,p)\}_{n\in\mathbb{N}}$ constitute a pure point spectrum.

In order to approach a solution of the full quantum mechanical problem, it is crucial to know how these fast eigenvalue problems are related for slow *dynamical* variables, i.e., for variable q, p because up to now these were considered as being *fixed*. The space adiabatic scheme will give a perturbative answer to this question, namely, its first step consists in constructing an operator $\boldsymbol{\pi}_\varepsilon\in\mathcal{B}(\mathcal{H}_{\rm f})$ on $\mathcal{H}_{\rm f}$ which takes the dynamics of the (q,p) variables into account. In particular, this "Moyal projection" symbol gives rise to a projection operator $\hat{\boldsymbol{\Pi}}\in\mathcal{B}(\mathcal{H})$ on the total Hilbert space $\mathcal{H}$ which is almost invariant under the dynamics of the *total* Hamiltonian. Indeed, the existence of such a projection operator is guaranteed by the *space adiabatic theorem* [51] which builds the basis for the SAPT scheme. It states that if conditions (C1)–(C4) are fulfilled, there exists an orthogonal projection operator $\hat{\boldsymbol{\Pi}}\in\mathcal{B}(\mathcal{H})$ on the total Hilbert space such that

$$[\hat{\boldsymbol{H}},\hat{\boldsymbol{\Pi}}]=\mathcal{O}_0(\varepsilon^\infty).\qquad(10)$$

Thereby, the estimate $\mathcal{O}_0(\varepsilon^\infty)$ means that for all $m\in\mathbb{N}$, there exists a constant $C_m\ge 0$ such that $\|[\hat{\boldsymbol{H}},\hat{\boldsymbol{\Pi}}]\|_{\mathcal{B}(\mathcal{H}_{\rm f})}\le C_m\varepsilon^m$. Furthermore, and very importantly for the scheme, the theorem assures that there exists a symbol function $\boldsymbol{\pi}_\varepsilon\in S^0_\rho(\varepsilon)$ whose Weyl quantization gives rise to the construction of the above projection operator such that $\hat{\boldsymbol{\Pi}}=\hat{\mathcal{W}}(\boldsymbol{\pi}_\varepsilon)+\mathcal{O}_0(\varepsilon^\infty)$. Thereby, $S^0_\rho(\varepsilon)$ refers to the class of the so-called semiclassical symbols which asymptotically approach a certain series expansion, in particular $\boldsymbol{\pi}_\varepsilon\asymp\sum_i\varepsilon^i\boldsymbol{\pi}_i$ with $\boldsymbol{\pi}_i\in S^{-i\rho}_\rho$ in the given case. The relation between $\hat{\boldsymbol{\Pi}}$ and $\hat{\boldsymbol{\pi}}_\varepsilon$ can be established by means of resolvent methods [52,55], and it assures one that these operators are indeed very close in norm.

The first goal of SAPT is consequently to establish an approximation to the projection operator $\hat{\boldsymbol{\Pi}}$ by constructing the underlying symbol function. In particular, the first stage provides the means to inductively compute a sequence of symbol functions $\{\boldsymbol{\pi}_i(q,p)\}_{i\in\mathbb{N}}$ such that $\boldsymbol{\pi}_i\in S^{-i\rho}_\rho(\mathcal{B}(\mathcal{H}_{\rm f}))$ for any $i\in\mathbb{N}$. This sequence eventually determines a well-defined adiabatic perturbation series $\boldsymbol{\pi}:=\sum_{i\ge0}\varepsilon^i\boldsymbol{\pi}_i\in S^0_\rho(\mathcal{B}(\mathcal{H}_{\rm f}))$. Then, the semiclassical symbol $\boldsymbol{\pi}_\varepsilon$ can be constructed as the [up to $\mathcal{O}_0(\varepsilon^\infty)$] unique resummation of $\boldsymbol{\pi}$. Then, we can relate the Weyl quantization of the symbol function $\boldsymbol{\pi}_\varepsilon$ to the projection operator $\hat{\boldsymbol{\Pi}}$. In order to determine the sequence $\{\boldsymbol{\pi}_i\}_{i\in\mathbb{N}}$ of operator-valued symbol functions, SAPT starts by considering the projections associated with the (q,p)-dependent eigenvalue problem of $\boldsymbol{H}_0(q,p)$ given above. Any of the eigenfunctions $\xi_{n,a}(q,p)\in\mathcal{H}_{\rm f}$ gives rise to an orthogonal projection operator,

$$\boldsymbol{\pi}_{0,n}(q,p):=\sum_{a=1}^{d_n}\xi_{n,a}(q,p)\langle\xi_{n,a}(q,p),\cdot\rangle_{\mathcal{H}_{\rm f}}.\qquad(11)$$

The $\boldsymbol{\pi}_{0,n}(q,p)$ are bounded operators on the fast Hilbert space $\mathcal{H}_{\rm f}$ for every point $(q,p)\in\mathbb{R}^2$. SAPT then necessitates to choose one of the fast subspaces whose quantum number will be denoted by $\nu\in\mathbb{N}$. The corresponding projection operator carries the symbol $\boldsymbol{\pi}_0:=\boldsymbol{\pi}_{0,\nu}(q,p)$. This selection is arbitrary and one could also choose a set of several quantum numbers whose spectrum is distinct

from the remaining spectrum. Indeed, the crucial ingredient is that we can separate a part of the spectrum according to the gap condition (C3). Nevertheless, the restriction to one single isolated eigenvalue simplifies the computations significantly while it does not narrow the results. As the results are valid for any ν, the totality of the backreactions results from performing the computations for all the ν's separately.

We first notice that the symbol $\boldsymbol{\pi}_0$ does not lead to a true projection operator after Weyl quantization, i.e., $\hat{\boldsymbol{\pi}}_0^2 \neq \hat{\boldsymbol{\pi}}_0$. To approach a true projection operator, we correct $\boldsymbol{\pi}_0$ order by order in ε. To make computations accessible, we switch to the phase space quantization scheme by using the Moyal product. This is indeed reasonable, as the above inequality can be estimated with respect to the Moyal product by $\boldsymbol{\pi}_0 \star_\varepsilon \boldsymbol{\pi}_0 - \boldsymbol{\pi}_0 \sim \varepsilon$. SAPT guarantees the existence of a unique series expansion $\boldsymbol{\pi} = \sum_{i\geq 0} \varepsilon^i \boldsymbol{\pi}_i$ such that the zeroth order contribution can be identified with $\boldsymbol{\pi}_0$ and such that the following natural properties hold [52]:

$$(\mathrm{S1}-1)\quad \boldsymbol{\pi}\star_\varepsilon\boldsymbol{\pi} = \boldsymbol{\pi}, \qquad (\mathrm{S1}-2)\quad \boldsymbol{\pi}^* = \boldsymbol{\pi},$$
$$(S1-3)\quad \boldsymbol{H}\star_\varepsilon\boldsymbol{\pi} - \boldsymbol{\pi}\star_\varepsilon\boldsymbol{H} = 0.$$

As we will explicitly work out in the next subsection, these equations together with the power series expansion of the star product in ε give rise to an inductive sequence of equations to determine the $\boldsymbol{\pi}_i$ for $i \geq 1$. The next subsection will deal with solving these algebraic equations up to some order $I \in \mathbb{N}$ in ε. For the time being, we assume that we are able to establish $\sum_{i\geq 0} \varepsilon^i \boldsymbol{\pi}_i$. The Weyl quantization of its resummation $\hat{\boldsymbol{\pi}}_\varepsilon$ is then a bounded operator on $\mathcal{H}$ due to the theorem by Calderon-Vaillancourt [52]. The latter proves to be an almost projector and to be almost commuting with $\hat{\boldsymbol{H}}$, i.e., [52]

$$(\mathrm{S1}-1)'\quad \hat{\boldsymbol{\pi}}_\varepsilon^2 = \hat{\boldsymbol{\pi}}_\varepsilon + \mathcal{O}_{-\infty}(\varepsilon^\infty),$$
$$(\mathrm{S1}-2)'\quad \hat{\boldsymbol{\pi}}_\varepsilon^* = \hat{\boldsymbol{\pi}}_\varepsilon,$$
$$(S1-3)'\quad \hat{\boldsymbol{H}}\hat{\boldsymbol{\pi}}_\varepsilon - \hat{\boldsymbol{\pi}}_\varepsilon\hat{\boldsymbol{H}} = \mathcal{O}_{-\infty}(\varepsilon^\infty),$$

where the $\mathcal{O}_{-\infty}(\varepsilon^\infty)$ means that the equations hold up to corrections that are asymptotically close to zero in the respective symbol class $S^m_\rho(\varepsilon)$. Note that the last estimate only holds for an unbounded $\hat{\boldsymbol{H}}$ with $\boldsymbol{H} \in S^m_\rho$, $m > 0$ if ρ is strictly bigger than zero. Otherwise, the corresponding commutator will not be small in the norm of bounded operators. For a convergent series $\boldsymbol{\pi}$, of course, the resummation $\boldsymbol{\pi}_\varepsilon$ simply equals the series itself, and (S1)' holds without the additional error terms in (S1–1)' and (S1–3)'. In the divergent case, the true projection operator $\hat{\boldsymbol{\Pi}}$ emerges from $\hat{\boldsymbol{\pi}}_\varepsilon$ using the resolvent methods already mentioned [52].

Equipped with the Moyal projector, the scheme aims at restricting the Hamilton operator to the corresponding subspace and thus to investigate the simplified dynamics on this subspace $\hat{\boldsymbol{\Pi}}\mathcal{H}$. Unfortunately, this will not be possible without further ado. On the one hand, $\hat{\boldsymbol{\Pi}}\mathcal{H}$ might not even be close to an ε-independent subspace. Besides, for the applications that we have in mind, we push the perturbation theory to the second order but not further. Hence, we compute

$$\boldsymbol{\pi}_{(I)} := \sum_{i=0}^{I} \varepsilon^i \boldsymbol{\pi}_i, \tag{12}$$

for some finite $I \in \mathbb{N}$. The Weyl quantization of $\boldsymbol{\pi}_{(I)}$ is by construction *not* an exact projector on the total Hilbert space $\mathcal{H}$. In particular, the restricted operator $\hat{\boldsymbol{H}}|_{\hat{\boldsymbol{\pi}}_{(I)}} := \hat{\boldsymbol{\pi}}_{(I)}\hat{\boldsymbol{H}}\hat{\boldsymbol{\pi}}_{(I)}$ does *not* preserve the subspace $\hat{\boldsymbol{\pi}}_{(I)}\mathcal{H}$. Hence, it would remain unclear how to perform a spectral analysis for this operator on $\hat{\boldsymbol{\pi}}_{(I)}\mathcal{H}$. Even worse, the operator $\hat{\boldsymbol{H}}|_{\hat{\boldsymbol{\pi}}_{(I)}}$ maps its domain within $\hat{\boldsymbol{\pi}}_{(I)}\mathcal{H}$ outside of $\hat{\boldsymbol{\pi}}_{(I)}\mathcal{H}$.

As a solution to this dilemma, the space adiabatic scheme suggests to resort to a suitable "reference" space by means of a unitary symbol $\boldsymbol{u}$. The theory will construct the symbol as a perturbation series similar to the procedure for the Moyal projector.

To begin with, it considers a symbol $\boldsymbol{u}_0(q,p)$ which maps the subspace $\boldsymbol{\pi}_0(q,p)\mathcal{H}_\mathrm{f}$ to a subspace of $\mathcal{H}_\mathrm{f}$ that does not depend on (q,p). A simple proposal for this would be the following: Fix a specific pair of slow phase space variables $(q_0,p_0) \in \mathbb{R}^2$ [52]. The choice can be physically motivated and depends on the problem under consideration. Then, consider the eigenbasis $\{\xi_{n,a}(q,p)\}_{n\in\mathbb{N}}$ of the (q,p)-dependent eigenvalue problem associated with the zeroth order Hamilton symbol $\boldsymbol{H}_0(q,p)$, and denote the basis at the point (q_0,p_0) by $\{\zeta_{n,a}\}_{n\in\mathbb{N}} := \{\xi_{n,a}(q_0,p_0)\}_{n\in\mathbb{N}}$. Eventually, this gives rise to the unitary symbol,

$$\boldsymbol{u}_0(q,p) = \sum_{n=0}^{\infty}\sum_{a=1}^{d_n} \zeta_{n,a}\langle\xi_{n,a}(q,p),\cdot\rangle_{\mathcal{H}_\mathrm{f}}. \tag{13}$$

Of course, the $\{\zeta_{n,a}\}_{n\in\mathbb{N}}$ could be any orthonormal basis of $\mathcal{H}_\mathrm{f}$. The important point is that the reference vectors $\zeta_{n,a}$ do *not* depend on (q,p). With their help, the scheme defines the reference projection associated with the quantum number ν as above [52],

$$\boldsymbol{\pi}_\mathrm{p} := \sum_{a=1}^{d_\nu} \zeta_{\nu,a}\langle\zeta_{\nu,a},\cdot\rangle_{\mathcal{H}_\mathrm{f}}. \tag{14}$$

The technical relevance of this reference structure is that $\boldsymbol{\pi}_\mathrm{p}$ in contrast to $\boldsymbol{\pi}_0$ does not receive adiabatic corrections throughout the application of the space adiabatic perturbation scheme, and thus always defines an *exact* projector on the total Hilbert space,

$$\hat{\boldsymbol{\pi}}_{\mathrm{p}} = \hat{\mathcal{W}}(\boldsymbol{\pi}_{\mathrm{p}}) = \boldsymbol{\pi}_{\mathrm{p}} \otimes \hat{\mathrm{I}}_{\mathrm{s}}, \tag{15}$$

which will be crucial for the adiabatic expansion and its spectral analysis. Note also that $\boldsymbol{u}_0$ is in fact independent of the choice of the fast quantum number ν while its higher order adiabatic corrections will depend on ν. Analogous to the perturbative construction of $\boldsymbol{\pi}$ as a power series in ε, we build $\boldsymbol{u}$ according to

$$\boldsymbol{u}(q,p) = \sum_{n=0}^{\infty} \varepsilon^i \boldsymbol{u}_i(q,p). \tag{16}$$

The scheme identifies the zeroth order of this expansion with the symbol in (13), and it follows directly from (11) with the choice $n \equiv \nu$ and with Eqs. (13) and (14) that $\boldsymbol{u}_0$ indeed intertwines the symbols $\boldsymbol{\pi}_0$ and $\boldsymbol{\pi}_{\mathrm{p}}$ according to $\boldsymbol{u}_0 \boldsymbol{\pi}_0 \boldsymbol{u}_0^* = \boldsymbol{\pi}_{\mathrm{p}}$. To extend this equation to the Moyal projector $\boldsymbol{\pi}$ and to make the symbol $\boldsymbol{u}$ a Moyal unitary, the scheme requires the following properties [52]:

$$(\mathrm{S2}-1)\quad \boldsymbol{u}^* \star_\varepsilon \boldsymbol{u} = \mathbf{1}_{\mathrm{f}}, \qquad (\mathrm{S2}-2)\quad \boldsymbol{u} \star_\varepsilon \boldsymbol{u}^* = \mathbf{1}_{\mathrm{f}},$$
$$(\mathrm{S2}-3)\quad \boldsymbol{u} \star_\varepsilon \boldsymbol{\pi} \star_\varepsilon \boldsymbol{u}^* = \boldsymbol{\pi}_{\mathrm{p}}.$$

Indeed, SAPT assures, given the conditions (C1)–(C4) and the symbols $\boldsymbol{\pi}_0, \boldsymbol{u}_0$ as initial values, that there exists a formal series expansion $\boldsymbol{u}$ as in Eq. (16) which satisfies the properties (S2). Furthermore, one can deduce that the coefficients of the series are $\boldsymbol{u}_i \in S_\rho^{-i\rho}(\mathcal{B}(\mathcal{H}_{\mathrm{f}}))$ for any $i \in \mathbb{N}$. Equations (S2) give rise to a series of algebraic equations that determine $\boldsymbol{u}_i$ order by order in ε by using the series expansion of the star product.

Let us assume that it is possible to establish the whole series $\sum_{i\geq 0} \varepsilon^i \boldsymbol{u}_i$ with this method. Then, there exists a [n up to $\mathcal{O}_0(\varepsilon^\infty)$] unique resummation $\boldsymbol{u}_\varepsilon$ of $\boldsymbol{u}$ whose Weyl quantization $\hat{\boldsymbol{u}}_\varepsilon$ is a bounded operator on $\mathcal{H}$ which satisfies [52]

$$(\mathrm{S2}-1)'\quad \hat{\boldsymbol{u}}_\varepsilon^\dagger \hat{\boldsymbol{u}}_\varepsilon = \hat{\mathbf{1}} + \mathcal{O}_{-\infty}(\varepsilon^\infty),$$
$$(\mathrm{S2}-2)'\quad \hat{\boldsymbol{u}}_\varepsilon \hat{\boldsymbol{u}}_\varepsilon^\dagger = \hat{\mathbf{1}} + \mathcal{O}_{-\infty}(\varepsilon^\infty),$$
$$(\mathrm{S2}-3)'\quad \hat{\boldsymbol{u}}_\varepsilon \hat{\boldsymbol{\pi}}_\varepsilon \hat{\boldsymbol{u}}_\varepsilon^\dagger = \hat{\boldsymbol{\pi}}_{\mathrm{p}} + \mathcal{O}_{-\infty}(\varepsilon^\infty).$$

Obviously, the resummation operator $\hat{\boldsymbol{u}}_\varepsilon$ is only almost unitary and intertwines $\hat{\boldsymbol{\pi}}_\varepsilon$ and $\hat{\boldsymbol{\pi}}_{\mathrm{p}}$ only up to some ε-dependent error. One can proceed to modify $\hat{\boldsymbol{u}}_\varepsilon$ by an $\mathcal{O}_0(\varepsilon^\infty)$ term which makes it a true unitary operator $\hat{\boldsymbol{U}}$. This operator is then also a true intertwining operator in the sense that $\hat{\boldsymbol{U}}\hat{\boldsymbol{\Pi}}\hat{\boldsymbol{U}}^\dagger = \hat{\boldsymbol{\pi}}_{\mathrm{p}}$ holds [52]. As for the Moyal projector, we will be concerned with determining the Moyal unitary up to some finite order $I \in \mathbb{N}$, i.e.,

$$\boldsymbol{u}_{(I)}(q,p) = \sum_{i=0}^{I} \varepsilon^i \boldsymbol{u}_i(q,p). \tag{17}$$

Finally, with these ingredients, it is possible to set up the "effective" Hamilton operator, which captures the backreactions from the νth quantum state of the fast subsystem onto the slow subsystem. Again, its explicit construction works using the symbolic calculus as we will show in the following subsection. Therefore, we assume the effective Hamiltonian to have the form of a perturbative power series in ε,

$$\boldsymbol{h}_{\mathrm{eff}}(q,p) = \sum_{i=0}^{\infty} \varepsilon^i \boldsymbol{h}_{\mathrm{eff},i}(q,p). \tag{18}$$

This series expansion is then defined as the unitary mapping of $\boldsymbol{H}$ on the reference space by means of $\boldsymbol{u}$, namely by [52]

$$(\mathrm{S3})\quad h_{\mathrm{eff}} \coloneqq \boldsymbol{u} \star_\varepsilon \boldsymbol{H} \star_\varepsilon \boldsymbol{u}^*.$$

Finally, we denote the Weyl quantization of its resummation $\hat{\boldsymbol{h}}_{\mathrm{eff},\varepsilon}$ as the effective Hamiltonian on the reference space. Besides, $\boldsymbol{h}_{\mathrm{eff},\varepsilon}$ is a semiclassical symbol in $S_\rho^m(\varepsilon)$ which is guaranteed by using the product rule for semiclassical symbols with $\boldsymbol{u} \in S_\rho^0(\varepsilon)$ and $\boldsymbol{H} \in S_\rho^m(\varepsilon)$. We also note that $\hat{\boldsymbol{h}}_{\mathrm{eff},\varepsilon}$ is in fact essentially self-adjoint on the Schwartz space $\mathcal{S}(\mathbb{R}, \mathcal{H}_{\mathrm{f}})$ [52]. Furthermore, the dynamics generated by $\hat{\boldsymbol{h}}_{\mathrm{eff},\varepsilon}$ almost corresponds to the dynamics of the original Hamilton operator $\hat{\boldsymbol{H}}$ in the sense that

$$e^{-i\hat{\boldsymbol{H}}s} - \hat{\boldsymbol{u}}_\varepsilon^\dagger e^{-i\hat{\boldsymbol{h}}_{\mathrm{eff},\varepsilon} s} \hat{\boldsymbol{u}}_\varepsilon = \mathcal{O}_0(\varepsilon^\infty |s|). \tag{19}$$

This theorem ensures that it is reasonable to consider the significantly simpler effective Hamiltonian to extract the dynamics and the solutions of the theory. In our applications of SAPT in the subsequent papers of this series, we will primarily consider effective Hamiltonian restricted to the reference space, which means that we will work with the symbol function $\boldsymbol{\pi}_{\mathrm{p}} \boldsymbol{h}_{\mathrm{eff}} \boldsymbol{\pi}_{\mathrm{p}}$. This is reasonable since the construction scheme for $\hat{\boldsymbol{u}}$ guarantees that $\hat{\boldsymbol{h}}_{\mathrm{eff}}$ preserves the subspace $\hat{\boldsymbol{\pi}}_{\mathrm{p}}\mathcal{H}$. In the applications where we push the perturbation theory up to some order I, this statement can be translated into the assertion that the reduced operator $\hat{\boldsymbol{h}}_{\mathrm{eff},(I)} \coloneqq \hat{\boldsymbol{u}}_{(I)} \hat{\boldsymbol{H}} \hat{\boldsymbol{u}}_{(I)}^\dagger$ preserves the subspace $\hat{\boldsymbol{\pi}}_{\mathrm{p}}\mathcal{H}$ up to corrections of order ε^{I+1}. It thus coincides there up to corrections of order ε^{I+1} with the operator $\hat{\boldsymbol{h}}_{\mathrm{eff},(I),\mathrm{p}} \coloneqq \hat{\boldsymbol{\pi}}_{\mathrm{p}} \hat{\boldsymbol{u}}_{(I)} \hat{\boldsymbol{H}} \hat{\boldsymbol{u}}_{(I)}^\dagger \hat{\boldsymbol{\pi}}_{\mathrm{p}}$ on the Hilbert subspace $\hat{\boldsymbol{\pi}}_{\mathrm{p}}\mathcal{H} \cong \mathcal{H}_{\mathrm{s}} \otimes \mathbb{C}^{d_n}$. On the other hand, we emphasize that the seemingly more natural operator $\hat{\boldsymbol{\pi}}_{(I)} \hat{\boldsymbol{H}} \hat{\boldsymbol{\pi}}_{(I)}$ does not preserve the subspace $\hat{\boldsymbol{\pi}}_{(I)}\mathcal{H}$ because $\hat{\boldsymbol{\pi}}_{(I)}$ is not an exact projector.

As a result, $\hat{\boldsymbol{h}}_{\mathrm{eff},(I),\mathrm{p}}$ provides the perturbative adiabatic decoupling that we wanted to achieve and will consequently be the object of interest in the sequel. The spectrum of $\hat{\boldsymbol{h}}_{\mathrm{eff},(I),\mathrm{p}}$, denoted by $E_{\mathrm{eff},(I)}$, is referred to as the νth energy band (recall that we restricted the backreactions to the fast quantum number ν right from the beginning). If $\Psi_{(I),\nu,\lambda}$ is a generalized eigenvector of $\hat{\boldsymbol{h}}_{\mathrm{eff},(I),\mathrm{p}}$ with eigenvalue λ then up to corrections of order ε^{I+1}, the vector $\tilde{\Psi}_{(I),\nu,\lambda} = \hat{\boldsymbol{u}}^{\dagger}_{(I)}\Psi_{(I),\nu,\lambda}$ is a generalized eigenvector of the original Hamiltonian $\hat{\boldsymbol{H}}$ with the same eigenvalue λ. This can be easily reproduced with the following computation in which we dropped any term of order $\mathcal{O}(\varepsilon^{I+1})$:

$$\begin{aligned}\hat{\boldsymbol{H}}\tilde{\Psi}_{(I),\nu,\lambda} &= \hat{\boldsymbol{H}}\hat{\boldsymbol{u}}^{\dagger}_{(I)}\hat{\boldsymbol{\pi}}_{\mathrm{p}}\hat{\boldsymbol{u}}_{(I)}\hat{\boldsymbol{u}}^{\dagger}_{(I)}\Psi_{(I),\nu,\lambda} = \hat{\boldsymbol{H}}\hat{\boldsymbol{\pi}}_{(I)}\hat{\boldsymbol{u}}^{\dagger}_{(I)}\Psi_{(I),\nu,\lambda}\\ &= \hat{\boldsymbol{\pi}}_{(I)}\hat{\boldsymbol{H}}\hat{\boldsymbol{u}}^{\dagger}_{(I)}\Psi_{(I),\nu,\lambda} = \hat{\boldsymbol{u}}^{\dagger}_{(I)}(\hat{\boldsymbol{\pi}}_{\mathrm{p}}\hat{\boldsymbol{u}}_{(I)}\hat{\boldsymbol{H}}\hat{\boldsymbol{u}}^{\dagger}_{(I)}\hat{\boldsymbol{\pi}}_{(I)})\Psi_{(I),\nu,\lambda}\\ &= \hat{\boldsymbol{u}}^{\dagger}_{(I)}\hat{\boldsymbol{h}}_{\mathrm{eff},(I),\mathrm{p}}\Psi_{(I),\nu,\lambda} = \lambda\tilde{\Psi}_{(I),\nu,\lambda}. \end{aligned} \tag{20}$$

The approximate eigenvector $\tilde{\Psi}_{(I),\nu,\lambda}$ is an element of the approximately invariant subspace $\hat{\boldsymbol{\pi}}_{(I)}\mathcal{H}$ up to corrections of order $\mathcal{O}(\varepsilon^{I+1})$ because it holds that by dropping again the corresponding corrections

$$\tilde{\Psi}_{(I),\nu,\lambda} = \hat{\boldsymbol{u}}^{\dagger}_{(I)}\hat{\boldsymbol{\pi}}_{\mathrm{p}}\hat{\boldsymbol{u}}_{(I)}\hat{\boldsymbol{u}}^{\dagger}_{(I)}\Psi_{(I),\nu,\lambda} = \hat{\boldsymbol{\pi}}_{(I)}\tilde{\Psi}_{(I),\nu,\lambda}. \tag{21}$$

In this way, the $\hat{\boldsymbol{u}}_{(I)}$ are displayed as an auxiliary structure introduced in order to solve the spectral problem including backreactions. But they have no further fundamental relevance as is also clear from the fact that they are not uniquely determined by the perturbative scheme.

Besides, we note that the $\hat{\boldsymbol{u}}_{(I)}$ is not to be confused with the unitary map $\hat{\boldsymbol{V}}$ that maps $\mathcal{H}$ to $L^2(\sigma(\hat{\boldsymbol{H}}), \mathrm{d}\mu)$, granted to exist by the spectral theorem, where $\sigma(\hat{\boldsymbol{H}})$ is the spectrum of $\hat{\boldsymbol{H}}$ and μ its spectral measure. This is already clear from the fact that $\hat{\boldsymbol{u}}_{(I)}$ generically depends on ν while $\hat{\boldsymbol{V}}$ does not. The fact that the $\hat{\boldsymbol{\pi}}_{(I)}$ approximately commute with $\hat{\boldsymbol{H}}$ and are approximate projections displays them as approximants of spectral projections of $\hat{\boldsymbol{H}}$ on the part $E_{\mathrm{eff},(I)}$ of the spectrum. The spectral projections are of course not necessarily mutually orthogonal even if the gap condition holds unless the energy bands are mutually disjoint. For instance, $\boldsymbol{H}(q,p)$ could have a pure point spectrum but $\hat{\boldsymbol{H}}$ could have an absolutely continuous spectrum. With these ideas in mind about the general concept of SAPT, we move on to the explicit construction scheme.

C. Perturbative construction

The objective of the space adiabatic construction is to iteratively compute the Moyal projector $\boldsymbol{\pi}(q,p)$, the Moyal unitary $\boldsymbol{u}(q,p)$, and finally the effective Hamilton symbol $\boldsymbol{h}_{\mathrm{eff}}(q,p)$ as alluded to in the previous subsection. The construction is based on the assumption that these three symbols appear as power series with respect to the adiabatic perturbation parameter ε, and we are going to concentrate on the determination of the power series coefficients up to some finite order $I \in \mathbb{N}$. That is, we establish the symbols $\boldsymbol{\pi}_{(I)} = \sum_{i\leq I}\varepsilon^i\boldsymbol{\pi}_i$, $\boldsymbol{u}_{(I)} = \sum_{i\leq I}\varepsilon^i\boldsymbol{u}_i$, and $\boldsymbol{h}_{\mathrm{eff,p},(I)} = \sum_{i\leq I}\varepsilon^i\boldsymbol{h}_{\mathrm{eff},i}$. Based on the construction rules (S1), (S2), and (S3), we can spell out the algebraic equations for determining any of the coefficients up to order $I \in \mathbb{N}$, in particular for the Moyal projector $\boldsymbol{\pi}$,

$$\begin{aligned}&(\mathrm{S1}-1)'' \quad \boldsymbol{\pi}_{(I)}\star_\varepsilon\boldsymbol{\pi}_{(I)} - \boldsymbol{\pi}_{(I)} = \mathcal{O}_0(\varepsilon^{I+1}),\\ &(\mathrm{S1}-2)'' \quad \boldsymbol{\pi}^*_{(I)} - \boldsymbol{\pi}_{(I)} = \mathcal{O}_0(\varepsilon^{I+1}),\\ &(\mathrm{S1}-3)'' \quad \boldsymbol{H}\star_\varepsilon\boldsymbol{\pi}_{(I)} - \boldsymbol{\pi}_{(I)}\star_\varepsilon\boldsymbol{H} = \mathcal{O}_0(\varepsilon^{I+1});\end{aligned}$$

for the Moyal unitary $\boldsymbol{u}$,

$$\begin{aligned}&(\mathrm{S2}-1)'' \quad \boldsymbol{u}^*_{(I)}\star_\varepsilon\boldsymbol{u}_{(I)} - \mathbf{1}_{\mathrm{f}} = \mathcal{O}_0(\varepsilon^{I+1}),\\ &(\mathrm{S2}-2)'' \quad \boldsymbol{u}_{(I)}\star_\varepsilon\boldsymbol{u}^*_{(I)} - \mathbf{1}_{\mathrm{f}} = \mathcal{O}_0(\varepsilon^{I+1}),\\ &(\mathrm{S2}-3)'' \quad \boldsymbol{u}_{(I)}\star_\varepsilon\boldsymbol{\pi}_{(I)}\star_\varepsilon\boldsymbol{u}^*_{(I)} - \boldsymbol{\pi}_{\mathrm{p}} = \mathcal{O}_0(\varepsilon^{I+1});\end{aligned}$$

and for the effective Hamiltonian $\boldsymbol{h}_{\mathrm{eff}}$,

$$(\mathrm{S3})'' \quad \boldsymbol{h}_{\mathrm{eff},(I)} - \boldsymbol{u}_{(I)}\star_\varepsilon\boldsymbol{H}\star_\varepsilon\boldsymbol{u}^*_{(I)} = \mathcal{O}_0(\varepsilon^{I+1}).$$

Regarding the $\mathcal{O}_0(\varepsilon^{I+1})$ estimate, we note that SAPT assures that the symbol coefficients $\boldsymbol{\pi}_I$, $\boldsymbol{u}_I$, and $\boldsymbol{h}_{\mathrm{eff},I}$ can be built in such a way that the coefficients of order ε^I and smaller vanish on the right-hand sides of these equations. Besides, the product rule for semiclassical symbols assures that by plugging in the correspondent series expansions up to order I on the left-hand side, the right-hand sides will be symbols in the class $S^0_\rho(\mathcal{B}(\mathcal{H}_{\mathrm{f}}))$ at most and so evaluate to bounded operators on the total Hilbert space. In Eq. (S3), this necessarily requires $\boldsymbol{H}$ to be in S^m_ρ with ρ strictly bigger than zero. In the following subsections, we provide the inductive construction rules for $\boldsymbol{\pi}_{(I)}$, $\boldsymbol{u}_{(I)}$, and $\boldsymbol{h}_{\mathrm{eff},(I)}$, based on the original scheme in [51,52], and therefore recall that the Moyal product for two symbols $\boldsymbol{B} \in S^{m_1}_\rho$ and $\boldsymbol{C} \in S^{m_2}_\rho$ is given at leading order by

$$(\boldsymbol{B}\star_\varepsilon\boldsymbol{C})(q,p) = \boldsymbol{B}(q,p)\boldsymbol{C}(q,p) + \mathcal{O}_0(\varepsilon). \tag{22}$$

1. Construction of the Moyal projector

Based on the construction rules (S1)", the Moyal projection symbol is to be determined iteratively up to order I. Therefore, the symbol function $\boldsymbol{\pi}_0(q,p)$ serves as the starting point for the induction scheme. The expansion of the Moyal product in (S1)" simply gives the standard operator product on $\mathcal{B}(\mathcal{H}_{\mathrm{f}})$ at zeroth order such that its restriction to zeroth order yields

$$(\mathrm{S1}-1)\quad \boldsymbol{\pi}_0^2-\boldsymbol{\pi}_0=0,\qquad (\mathrm{S1}-2)\quad \boldsymbol{\pi}_0^*-\boldsymbol{\pi}_0=0,$$
$$(S1-3)\quad [\boldsymbol{H}_0,\boldsymbol{\pi}_0]=0. \tag{23}$$

These equations are fulfilled by construction of $\boldsymbol{\pi}_0$ such that the basis case is granted. Let us then assume that it is possible to construct $\boldsymbol{\pi}_{(I-1)}$ and turn to the solution of the coefficient $\boldsymbol{\pi}_I$. The construction rule (S1–1)" provides the diagonal parts of $\boldsymbol{\pi}_I$. The induction scheme suggests to split $\boldsymbol{\pi}_{(I)}$ into its highest order contribution $\varepsilon^I\boldsymbol{\pi}_I$, and the remainder $\boldsymbol{\pi}_{(I-1)}$. The induction scheme allows one to assume that we already found $\boldsymbol{\pi}_{(I-1)}$ such that $\boldsymbol{\pi}_{(I-1)}\star_\varepsilon\boldsymbol{\pi}_{(I-1)}-\boldsymbol{\pi}_{(I-1)}=\mathcal{O}_0(\varepsilon^I)$ is satisfied. We then denote the terms of (S1–1)" at order ε^I which only include the already known symbols $\boldsymbol{\pi}_{(I-1)}$ by $\boldsymbol{a}_{I-1}$. We then have $\boldsymbol{\pi}_{(I-1)}\star_\varepsilon\boldsymbol{\pi}_{(I-1)}-\boldsymbol{\pi}_{(I-1)}=:\varepsilon^I\boldsymbol{a}_{I-1}+\mathcal{O}_0(\varepsilon^{I+1})$. Importantly, $\boldsymbol{a}_{I-1}$ is already explicitly determined because of the induction assumption. For the induction step, we then consider the rule (S1–1)" *including* the symbol $\boldsymbol{\pi}_I$. By carrying all the terms with ε's of order $I+1$ and higher to $\mathcal{O}_0(\varepsilon^{I+1})$, (S1–1)" takes the form

$$\begin{aligned}\mathcal{O}_0(\varepsilon^{I+1})&=\boldsymbol{\pi}_{(I-1)}\star_\varepsilon\boldsymbol{\pi}_{(I-1)}-\boldsymbol{\pi}_{(I-1)}+\varepsilon^I\boldsymbol{\pi}_I\boldsymbol{\pi}_0+\varepsilon^I\boldsymbol{\pi}_0\boldsymbol{\pi}_I-\varepsilon^I\boldsymbol{\pi}_I\\&=\varepsilon^I(\boldsymbol{a}_{I-1}+\boldsymbol{\pi}_I\boldsymbol{\pi}_0+\boldsymbol{\pi}_0\boldsymbol{\pi}_I-\boldsymbol{\pi}_I).\end{aligned} \tag{24}$$

To extract $\boldsymbol{\pi}_I$, we first define the orthogonal complement of $\boldsymbol{\pi}_0$ in $\mathcal{H}_\mathrm{f}$ as $\boldsymbol{\pi}_0^\perp:=\mathbf{1}_\mathrm{f}-\boldsymbol{\pi}_0$. Then, by projecting (24) to the block diagonal pieces, this yields indeed the determining equations for the diagonal parts of $\boldsymbol{\pi}_I$,

$$\begin{aligned}\boldsymbol{\pi}_I^{\mathrm{D},0}&:=\boldsymbol{\pi}_0\boldsymbol{\pi}_I\boldsymbol{\pi}_0=-\boldsymbol{\pi}_0\boldsymbol{a}_{I-1}\boldsymbol{\pi}_0,\\\boldsymbol{\pi}_I^{\mathrm{D},\perp}&:=\boldsymbol{\pi}_0^\perp\boldsymbol{\pi}_I\boldsymbol{\pi}_0^\perp=\boldsymbol{\pi}_0^\perp\boldsymbol{a}_{I-1}\boldsymbol{\pi}_0^\perp.\end{aligned} \tag{25}$$

For the projection on the off-diagonal parts of $\boldsymbol{\pi}_I$, the following consistency conditions arise:

$$\boldsymbol{\pi}_0\boldsymbol{a}_{I-1}\boldsymbol{\pi}_0^\perp=0=\boldsymbol{\pi}_0^\perp\boldsymbol{a}_{I-1}\boldsymbol{\pi}_0. \tag{26}$$

This identity follows from the defining equation (S1–1)" by projecting to the off-diagonal pieces and pushing all terms of order ε^{I+1} and higher into the error term. Furthermore, the associativity of the star product is exploited to obtain

$$\begin{aligned}\varepsilon^I\boldsymbol{\pi}_0\boldsymbol{a}_{I-1}\boldsymbol{\pi}_0^\perp&=\boldsymbol{\pi}_0(\boldsymbol{\pi}_{(I-1)}\star_\varepsilon\boldsymbol{\pi}_{(I-1)}-\boldsymbol{\pi}_{(I-1)})\boldsymbol{\pi}_0^\perp+\mathcal{O}_0(\varepsilon^{I+1})\\&=\boldsymbol{\pi}_{(I-1)}(\boldsymbol{\pi}_{(I-1)}\star_\varepsilon\boldsymbol{\pi}_{(I-1)}-\boldsymbol{\pi}_{(I-1)})\boldsymbol{\pi}_{(I-1)}^\perp+\mathcal{O}_0(\varepsilon^{I+1})\\&=\boldsymbol{\pi}_{(I-1)}\star_\varepsilon(\boldsymbol{\pi}_{(I-1)}\star_\varepsilon\boldsymbol{\pi}_{(I-1)}-\boldsymbol{\pi}_{(I-1)})\star_\varepsilon\boldsymbol{\pi}_{(I-1)}^\perp+\mathcal{O}_0(\varepsilon^{I+1})\\&=(\boldsymbol{\pi}_{(I-1)}\star_\varepsilon\boldsymbol{\pi}_{(I-1)}-\boldsymbol{\pi}_{(I-1)})\star_\varepsilon\boldsymbol{\pi}_{(I-1)}\star_\varepsilon\boldsymbol{\pi}_{(I-1)}^\perp+\mathcal{O}_0(\varepsilon^{I+1})\\&=-(\boldsymbol{\pi}_{(I-1)}\star_\varepsilon\boldsymbol{\pi}_{(I-1)}-\boldsymbol{\pi}_{(I-1)})\star_\varepsilon(\boldsymbol{\pi}_{(I-1)}\star_\varepsilon\boldsymbol{\pi}_{(I-1)}-\boldsymbol{\pi}_{(I-1)})+\mathcal{O}_0(\varepsilon^{I+1})\\&=\mathcal{O}_0(\varepsilon^{I+1}),\end{aligned} \tag{27}$$

where we used in the last step that $\varepsilon^{2I}\boldsymbol{a}_{I-1}\star_\varepsilon\boldsymbol{a}_{I-1}$ is evidently of higher order than ε^{I+1} and so fits into the error term. As a result, the diagonal terms are determined by (25) while the off-diagonal contributions to $\boldsymbol{\pi}_I$ still need to be fixed. (S1–3)" provides the algebraic equations to establish them. First, the third equation in (23) ensures that the base clause of the induction is settled. We then assume that (S1–3)" also holds for $\boldsymbol{H}$ and $\boldsymbol{\pi}_{(I-1)}$ with the corresponding error term $\mathcal{O}_0(\varepsilon^I)$. For the iteration step, we perform the split $\boldsymbol{\pi}_{(I)}=\boldsymbol{\pi}_{(I-1)}+\varepsilon^I\boldsymbol{\pi}_I$ and insert this into (S1–3)", and we define a new symbol $\boldsymbol{b}_{I-1}$ such that $[\boldsymbol{\pi}_{(I-1)},\boldsymbol{H}]_{\star_\varepsilon}=:\varepsilon^I\boldsymbol{b}_{I-1}+\mathcal{O}_0(\varepsilon^{I+1})$, where we shifted any contribution of order ε^{I+1} and higher in the Moyal commutator into the error term. This results in

$$\begin{aligned}\mathcal{O}_0(\varepsilon^{I+1})&=\boldsymbol{H}\star_\varepsilon\boldsymbol{\pi}_{(I)}-\boldsymbol{\pi}_{(I)}\star_\varepsilon\boldsymbol{H}\\&=\boldsymbol{H}\star_\varepsilon\boldsymbol{\pi}_{(I-1)}-\boldsymbol{\pi}_{(I-1)}\star_\varepsilon\boldsymbol{H}+\varepsilon^I(\boldsymbol{H}\star_\varepsilon\boldsymbol{\pi}_I-\boldsymbol{\pi}_I\star_\varepsilon\boldsymbol{H})\\&=\varepsilon^I(-\boldsymbol{b}_{I-1}+\boldsymbol{H}\star_\varepsilon\boldsymbol{\pi}_I-\boldsymbol{\pi}_I\star_\varepsilon\boldsymbol{H})\\&=\varepsilon^I(-\boldsymbol{b}_{I-1}+[\boldsymbol{H}_0,\boldsymbol{\pi}_I]).\end{aligned} \tag{28}$$

Hence, the term in the brackets should vanish. To extract the off-diagonal contributions of $\boldsymbol{\pi}_I$ with this relation, it is advisable to multiply it by $\boldsymbol{\pi}_0$ from the left as well as by its orthogonal complement $\pi_0^\perp$ from the right, and to repeat the procedure with these operators exchanged. For this, we define the two off-diagonal contributions of $\boldsymbol{\pi}_I$ as $\boldsymbol{\pi}_I^{\mathrm{OD},1}:=\boldsymbol{\pi}_0\boldsymbol{\pi}_I\boldsymbol{\pi}_0^\perp$ and $\boldsymbol{\pi}_I^{\mathrm{OD},2}:=\boldsymbol{\pi}_0^\perp\boldsymbol{\pi}_I\boldsymbol{\pi}_0$. Besides, let us denote the restriction of the zeroth order Hamilton symbol $\boldsymbol{H}_0$ which excludes the preselected energy band E_ν by $\boldsymbol{H}_0^\perp:=\boldsymbol{H}_0\boldsymbol{\pi}_0^\perp$. Using that $\boldsymbol{\pi}_0$ and $\boldsymbol{\pi}_0^\perp$, respectively, commute with $\boldsymbol{H}_0$ as operators on $\mathcal{H}_\mathrm{f}$, this yields for the first off-diagonal part

$$\begin{aligned}\varepsilon^I\boldsymbol{\pi}_0\boldsymbol{b}_{I-1}\boldsymbol{\pi}_0^\perp+\mathcal{O}_0(\varepsilon^{I+1})&=\varepsilon^I(\boldsymbol{\pi}_0\boldsymbol{H}_0\boldsymbol{\pi}_I\boldsymbol{\pi}_0^\perp-\boldsymbol{\pi}_0\boldsymbol{\pi}_I\boldsymbol{H}_0\boldsymbol{\pi}_0^\perp)\\&=\varepsilon^I\boldsymbol{\pi}_I^{\mathrm{OD},1}(E_\nu\mathbf{1}_\mathrm{f}-\boldsymbol{H}_0^\perp).\end{aligned} \tag{29}$$

By restricting to the terms of order ε^I in (29), this gives finally for $\boldsymbol{\pi}_I^{\mathrm{OD},1}$ and similarly for $\boldsymbol{\pi}_I^{\mathrm{OD},2}$

$$\begin{aligned}\boldsymbol{\pi}_I^{\mathrm{OD},1}&=\boldsymbol{\pi}_0\boldsymbol{b}_{I-1}(E_\nu\mathbf{1}_\mathrm{f}-\boldsymbol{H}_0^\perp)^{-1}\boldsymbol{\pi}_0^\perp,\\\boldsymbol{\pi}_I^{\mathrm{OD},2}&=-(E_\nu\mathbf{1}_\mathrm{f}-\boldsymbol{H}_0^\perp)^{-1}\boldsymbol{\pi}_0^\perp\boldsymbol{b}_{I-1}\boldsymbol{\pi}_0.\end{aligned} \tag{30}$$

Again, consistency with the former derivation of $\boldsymbol{a}_{I-1}$ for the diagonal part of the Moyal projector requires one to show that the diagonal part of Eq. (28) is indeed vanishing. We split this task into two steps and first derive an expression for $\boldsymbol{\pi}_0\boldsymbol{b}_{I-1}\boldsymbol{\pi}_0$ and in a second step for $\boldsymbol{\pi}_0[\boldsymbol{H}_0,\boldsymbol{\pi}_I]\boldsymbol{\pi}_0$, namely

$$\begin{aligned}\varepsilon^I\boldsymbol{\pi}_0\boldsymbol{b}_{I-1}\boldsymbol{\pi}_0 &= \boldsymbol{\pi}_0(\boldsymbol{\pi}_{(I-1)}\star_\varepsilon\boldsymbol{H}-\boldsymbol{H}\star_\varepsilon\boldsymbol{\pi}_{(I-1)})\boldsymbol{\pi}_0+\mathcal{O}_0(\varepsilon^{I+1})\\ &= \boldsymbol{\pi}_{(I-1)}\star_\varepsilon(\boldsymbol{\pi}_{(I-1)}\star_\varepsilon\boldsymbol{H}-\boldsymbol{H}\star_\varepsilon\boldsymbol{\pi}_{(I-1)})\star_\varepsilon\boldsymbol{\pi}_{(I-1)}+\mathcal{O}_0(\varepsilon^{I+1})\\ &= (\varepsilon^I\boldsymbol{a}_{I-1}+\boldsymbol{\pi}_{(I-1)})\star_\varepsilon\boldsymbol{H}\star_\varepsilon\boldsymbol{\pi}_{(I-1)}-\boldsymbol{\pi}_{(I-1)}\star_\varepsilon\boldsymbol{H}\star_\varepsilon(\varepsilon^I\boldsymbol{a}_{I-1}+\boldsymbol{\pi}_{(I-1)})+\mathcal{O}_0(\varepsilon^{I+1})\\ &= \varepsilon^I(\boldsymbol{a}_{I-1}\star_\varepsilon\boldsymbol{H}\star_\varepsilon\boldsymbol{\pi}_{(I-1)}-\boldsymbol{\pi}_{(I-1)}\star_\varepsilon\boldsymbol{H}\star_\varepsilon\boldsymbol{a}_{I-1})+\mathcal{O}_0(\varepsilon^{I+1})\\ &= \varepsilon^I(\boldsymbol{a}_{I-1}\boldsymbol{H}_0\boldsymbol{\pi}_0-\boldsymbol{\pi}_0\boldsymbol{H}_0\boldsymbol{a}_{I-1})+\mathcal{O}_0(\varepsilon^{I+1})\\ &= \mathcal{O}_0(\varepsilon^{I+1}).\end{aligned}\tag{31}$$

Indeed, the bracket term in the last line vanishes since $\boldsymbol{H}_0\boldsymbol{\pi}_0 = E_\nu\mathbf{1}_\mathrm{f}$ is proportional to the one in $\mathcal{H}_\mathrm{f}$ and hence commutes with $\boldsymbol{a}_{I-1}$. On the other hand, we have for $\boldsymbol{\pi}_0[\boldsymbol{H}_0,\boldsymbol{\pi}_I]\boldsymbol{\pi}_0$

$$\begin{aligned}\varepsilon^I\boldsymbol{\pi}_0[\boldsymbol{H}_0,\boldsymbol{\pi}_I]\boldsymbol{\pi}_0 &= \varepsilon^I(\boldsymbol{\pi}_0\boldsymbol{H}_0\boldsymbol{\pi}_I\boldsymbol{\pi}_0-\boldsymbol{\pi}_0\boldsymbol{\pi}_I\boldsymbol{H}_0\boldsymbol{\pi}_0)\\ &= \varepsilon^I E_\nu(\boldsymbol{\pi}_I^{\mathrm{D},1}-\boldsymbol{\pi}_I^{\mathrm{D},1})=0,\end{aligned}\tag{32}$$

with the same reasoning. Employing the same relations, it is straightforward to derive that also

$$\varepsilon^I\boldsymbol{\pi}_0^\perp\boldsymbol{b}_{I-1}\boldsymbol{\pi}_0^\perp=\mathcal{O}_0(\varepsilon^{I+1})\tag{33}$$

holds true, and we will consequently not write it down explicitly. This proves the consistency of the relations for the off-diagonal contributions for $\boldsymbol{\pi}_I$. Collecting all terms, the final result for the coefficient of the Moyal projector at order I reads

$$\begin{aligned}\boldsymbol{\pi}_I &= \boldsymbol{\pi}_I^{\mathrm{D},0}+\boldsymbol{\pi}_I^{\mathrm{D},\perp}+\boldsymbol{\pi}_I^{\mathrm{OD},1}+\boldsymbol{\pi}_I^{\mathrm{OD},2}\\ &= -\boldsymbol{\pi}_0\boldsymbol{a}_{I-1}\boldsymbol{\pi}_0+\boldsymbol{\pi}_0^\perp\boldsymbol{a}_{I-1}\boldsymbol{\pi}_0^\perp+\boldsymbol{\pi}_0\boldsymbol{b}_{I-1}\\ &\quad\times(E_\nu\mathbf{1}_\mathrm{f}-\boldsymbol{H}_0^\perp)^{-1}\boldsymbol{\pi}_0^\perp-(E_\nu\mathbf{1}_\mathrm{f}-\boldsymbol{H}_0^\perp)^{-1}\boldsymbol{\pi}_0^\perp\boldsymbol{b}_{I-1}\boldsymbol{\pi}_0,\end{aligned}\tag{34}$$

where $\varepsilon^I\boldsymbol{a}_{I-1}=(\boldsymbol{\pi}_{I-1}\star_\varepsilon\boldsymbol{\pi}_{I-1}-\boldsymbol{\pi}_{(I-1)})|_I$ and $\varepsilon^I\boldsymbol{b}_{I-1}=[\boldsymbol{\pi}_{(I-1)},\boldsymbol{H}]_{\star_\varepsilon}|_I$. This closes the construction of the Moyal projector, and we move on to the inductive construction of the Moyal unitary.

2. Construction of the Moyal unitary

The construction of the Moyal unitary relies on the rules (S2)" in Sec. II C, and as before, we proceed iteratively to build $\boldsymbol{u}_{(I)}$. We assume a formal power series for the Moyal unitary up to order I, namely $\boldsymbol{u}_{(I)}=\sum_{i\le I}\varepsilon^I\boldsymbol{u}_i$. The symbol function $\boldsymbol{u}_0$ from (13) serves as the starting point for the construction scheme. Then, at zeroth order the rules (S2)" evaluate to

$$\begin{aligned}&(\mathrm{S2}-1)\quad \boldsymbol{u}_0^*\cdot\boldsymbol{u}_0-\mathbf{1}_\mathrm{f}=0,\\ &(\mathrm{S2}-2)\quad \boldsymbol{u}_0\cdot\boldsymbol{u}_0^*-\mathbf{1}_\mathrm{f}=0,\\ &(\mathrm{S2}-3)\quad \boldsymbol{u}_0\cdot\boldsymbol{\pi}_0\cdot\boldsymbol{u}_0^*-\boldsymbol{\pi}_\mathrm{p}=0.\end{aligned}\tag{35}$$

These equations are satisfied by construction of $\boldsymbol{u}_0$ and $\boldsymbol{\pi}_\mathrm{p}$. By induction, we assume that the rules (S2)" are satisfied for the symbol $\boldsymbol{u}_{(I-1)}$. Let us start with the first and second construction rules. The induction assumption yields that $\boldsymbol{u}^*_{(I-1)}\star_\varepsilon\boldsymbol{u}_{(I-1)}-\mathbf{1}_\mathrm{f}=\mathcal{O}_0(\varepsilon^I)$, and likewise $\boldsymbol{u}_{(I-1)}\star_\varepsilon\boldsymbol{u}^*_{(I-1)}-\mathbf{1}_\mathrm{f}=\mathcal{O}_0(\varepsilon^I)$. To determine $\boldsymbol{u}_I$, we define the operators $\boldsymbol{c}_{I-1}$ and $\boldsymbol{d}_{I-1}$ as the contributions of order I to these equations, in particular $\boldsymbol{u}^*_{(I-1)}\star_\varepsilon\boldsymbol{u}_{(I-1)}-\mathbf{1}_\mathrm{f}=:\varepsilon^I\boldsymbol{c}_{I-1}+\mathcal{O}_0(\varepsilon^{I+1})$ and $\boldsymbol{u}_{(I-1)}\star_\varepsilon\boldsymbol{u}^*_{(I-1)}-\mathbf{1}_\mathrm{f}=:\varepsilon^I\boldsymbol{d}_{I-1}+\mathcal{O}_0(\varepsilon^{I+1})$. Then consider the corresponding equations for $\boldsymbol{u}_{(I)}$ and its adjoint $\boldsymbol{u}^*_{(I)}$. They give

$$\boldsymbol{u}^*_{(I)}\star_\varepsilon\boldsymbol{u}_{(I)}-\mathbf{1}_\mathrm{f}=\varepsilon^I(\boldsymbol{c}_{I-1}+\boldsymbol{u}_I^*\boldsymbol{u}_0+\boldsymbol{u}_0^*\boldsymbol{u}_I)+\mathcal{O}_0(\varepsilon^{I+1}),\tag{36}$$

$$\boldsymbol{u}_{(I)}\star_\varepsilon\boldsymbol{u}^*_{(I)}-\mathbf{1}_\mathrm{f}=\varepsilon^I(\boldsymbol{d}_{I-1}+\boldsymbol{u}_I\boldsymbol{u}_0^*+\boldsymbol{u}_0\boldsymbol{u}_I^*)+\mathcal{O}_0(\varepsilon^{I+1}),\tag{37}$$

and we require that the terms in the brackets vanish separately. Assuming that this holds true, it is possible to extract $\boldsymbol{u}_I^*$ in both equations according to

$$\boldsymbol{u}_I^*=-(\boldsymbol{c}_{I-1}\boldsymbol{u}_0^*+\boldsymbol{u}_0^*\boldsymbol{u}_I\boldsymbol{u}_0^*)=-(\boldsymbol{u}_0^*\boldsymbol{d}_{I-1}+\boldsymbol{u}_0^*\boldsymbol{u}_I\boldsymbol{u}_0^*).\tag{38}$$

By comparing the two defining terms, one can relate the operators $\boldsymbol{c}_{I-1}$ and $\boldsymbol{d}_{I-1}$ by

$$\boldsymbol{c}_{I-1}=\boldsymbol{u}_0^*\boldsymbol{d}_{I-1}\boldsymbol{u}_0,\tag{39}$$

which is identically satisfied by the induction assumption, namely, we can show that

$$\begin{aligned}\varepsilon^I(\boldsymbol{c}_{I-1}\boldsymbol{u}_0^* - \boldsymbol{u}_0^*\boldsymbol{d}_{I-1}) &= (\boldsymbol{u}^*_{(I-1)}\star_\varepsilon\boldsymbol{u}_{(I-1)} - \mathbf{1}_\mathrm{f})\boldsymbol{u}_0^* - \boldsymbol{u}_0^*(\boldsymbol{u}_{(I-1)}\star_\varepsilon\boldsymbol{u}^*_{(I-1)} - \mathbf{1}_\mathrm{f}) + \mathcal{O}_0(\varepsilon^{I+1})\\ &= (\boldsymbol{u}^*_{(I-1)}\star_\varepsilon\boldsymbol{u}_{(I-1)} - \mathbf{1}_\mathrm{f})\star_\varepsilon\boldsymbol{u}^*_{(I-1)} - \boldsymbol{u}^*_{(I-1)}\star_\varepsilon(\boldsymbol{u}_{(I-1)}\star_\varepsilon\boldsymbol{u}^*_{(I-1)} - \mathbf{1}_\mathrm{f}) + \mathcal{O}_0(\varepsilon^{I+1})\\ &= \mathcal{O}_0(\varepsilon^{I+1}),\end{aligned}$$

where we used the associativity of the star product for the last manipulation.

In a next step, we consider the rule (S2–3)". Since the zeroth order part of this equation is satisfied [cf. Eq. (35)], we assume by induction that we found $\boldsymbol{u}_{(I-1)}$ such that $\boldsymbol{u}_{(I-1)}\star_\varepsilon\boldsymbol{\pi}_{(I-1)}\star_\varepsilon\boldsymbol{u}^*_{(I-1)} - \boldsymbol{\pi}_\mathrm{p} = \mathcal{O}_0(\varepsilon^I)$ is satisfied. We define a new operator $\boldsymbol{e}_{I-1}$ to extract the contributions of order ε^I of this equation, namely $\boldsymbol{u}_{(I-1)}\star_\varepsilon\boldsymbol{\pi}_{(I-1)}\star_\varepsilon\boldsymbol{u}^*_{(I-1)} - \boldsymbol{\pi}_\mathrm{p} =: \varepsilon^I\boldsymbol{e}_{I-1} + \mathcal{O}_0(\varepsilon^{I+1})$. Then the induction step consists in considering equation (iii) for $\boldsymbol{u}_{(I)}$, which yields

$$\begin{aligned}\boldsymbol{u}_{(I)}\star_\varepsilon\boldsymbol{\pi}_{(I)}\star_\varepsilon\boldsymbol{u}^*_{(I)} - \boldsymbol{\pi}_\mathrm{p} = \varepsilon^I(&\boldsymbol{e}_{I-1} + \boldsymbol{u}_I\boldsymbol{\pi}_0\boldsymbol{u}_0^* + \boldsymbol{u}_0\boldsymbol{\pi}_I\boldsymbol{u}_0^*\\ &+ \boldsymbol{u}_0\boldsymbol{\pi}_0\boldsymbol{u}_I^*) + \mathcal{O}_0(\varepsilon^{I+1}).\end{aligned} \tag{40}$$

Consequently, we require the terms in the brackets to vanish. In this case and by means of (38), it holds true that

$$\boldsymbol{e}_{I-1} = -\boldsymbol{u}_0\boldsymbol{\pi}_I\boldsymbol{u}_0^* - \boldsymbol{u}_I\boldsymbol{\pi}_0\boldsymbol{u}_0^* + \boldsymbol{u}_0\boldsymbol{\pi}_0(\boldsymbol{c}_{I-1}\boldsymbol{u}_0^* + \boldsymbol{u}_0^*\boldsymbol{u}_I\boldsymbol{u}_0^*) \tag{41}$$

$$= -\boldsymbol{u}_0\boldsymbol{\pi}_I\boldsymbol{u}_0^* - \boldsymbol{u}_I\boldsymbol{u}_0^*\boldsymbol{\pi}_\mathrm{p} + \boldsymbol{u}_0\boldsymbol{\pi}_0\boldsymbol{c}_{I-1}\boldsymbol{u}_0^* + \boldsymbol{\pi}_\mathrm{p}\boldsymbol{u}_I\boldsymbol{u}_0^*, \tag{42}$$

which transforms into an equation determining $\boldsymbol{u}_I$, namely

$$[\boldsymbol{\pi}_\mathrm{p}, \boldsymbol{u}_I\boldsymbol{u}_0^*] = \boldsymbol{e}_{I-1} + \boldsymbol{u}_0\boldsymbol{\pi}_I\boldsymbol{u}_0^* - \boldsymbol{u}_0\boldsymbol{\pi}_0\boldsymbol{c}_{I-1}\boldsymbol{u}_0^*. \tag{43}$$

We then consider projecting this equation onto the diagonal or off-diagonal parts with respect to $\boldsymbol{\pi}_\mathrm{p}$ and its orthogonal complement $\boldsymbol{\pi}_\mathrm{p}^\perp := \mathbf{1}_\mathrm{f} - \boldsymbol{\pi}_\mathrm{p}$. As the commutator on the left-hand side is antisymmetric, this gives

$$0 = \boldsymbol{\pi}_\mathrm{p}(\boldsymbol{e}_{I-1} + \boldsymbol{u}_0\boldsymbol{\pi}_I\boldsymbol{u}_0^* - \boldsymbol{d}_{I-1})\boldsymbol{\pi}_\mathrm{p} =: \boldsymbol{R}_{ed}, \tag{44}$$

$$0 = \boldsymbol{\pi}_\mathrm{p}^\perp(\boldsymbol{e}_{I-1} + \boldsymbol{u}_0\boldsymbol{\pi}_I\boldsymbol{u}_0^* - \boldsymbol{d}_{I-1})\boldsymbol{\pi}_\mathrm{p}^\perp =: \boldsymbol{R}_{ed}^\perp, \tag{45}$$

where we introduced the symbols $\boldsymbol{R}_{ed}$ and $\boldsymbol{R}_{ed}^\perp$ for later convenience. Both equations are identically satisfied by the induction assumption. To show this, recall the result for the Moyal projector in Eq. (34), which presents the projector already as a decomposition into diagonal and off-diagonal parts. Multiplication by $\boldsymbol{u}_0$ from the left and by $\boldsymbol{u}_0^*$ from the right provides the splitting in diagonal and off-diagonal parts with respect to $\boldsymbol{\pi}_\mathrm{p}$ and $\boldsymbol{\pi}_\mathrm{p}^\perp$,

$$\begin{aligned}\boldsymbol{u}_0\boldsymbol{\pi}_I\boldsymbol{u}_0^* = &-\boldsymbol{\pi}_\mathrm{p}\boldsymbol{u}_0\boldsymbol{a}_{I-1}\boldsymbol{u}_0^*\boldsymbol{\pi}_\mathrm{p} + \boldsymbol{\pi}_\mathrm{p}^\perp\boldsymbol{u}_0\boldsymbol{a}_{I-1}\boldsymbol{u}_0^*\boldsymbol{\pi}_\mathrm{p}^\perp\\ &+ \boldsymbol{\pi}_\mathrm{p}\boldsymbol{u}_0\boldsymbol{b}_{I-1}(E_\nu\mathbf{1}_\mathrm{f} - \boldsymbol{H}_0^\perp)^{-1}\boldsymbol{u}_0^*\boldsymbol{\pi}_\mathrm{p}^\perp\\ &- \boldsymbol{\pi}_\mathrm{p}^\perp\boldsymbol{u}_0(E_\nu\mathbf{1}_\mathrm{f} - \boldsymbol{H}_0^\perp)^{-1}\boldsymbol{b}_{I-1}\boldsymbol{u}_0^*\boldsymbol{\pi}_\mathrm{p}.\end{aligned} \tag{46}$$

It is now obvious to restrict in Eqs. (44) and (45) to the relevant contributions for $\boldsymbol{\pi}_I$, namely

$$\begin{aligned}\varepsilon^I\boldsymbol{R}_{ed} &= \varepsilon^I\boldsymbol{\pi}_\mathrm{p}(\boldsymbol{e}_{I-1} - \boldsymbol{u}_0\boldsymbol{a}_{I-1}\boldsymbol{u}_0^* - \boldsymbol{d}_{I-1})\boldsymbol{\pi}_\mathrm{p}\\ &= \boldsymbol{\pi}_\mathrm{p}((\boldsymbol{u}_{(I-1)}\star_\varepsilon\boldsymbol{\pi}_{(I-1)}\star_\varepsilon\boldsymbol{u}^*_{(I-1)} - \boldsymbol{\pi}_\mathrm{p}) - \boldsymbol{u}_0(\boldsymbol{\pi}_{(I-1)}\star_\varepsilon\boldsymbol{\pi}_{(I-1)} - \boldsymbol{\pi}_{(I-1)})\boldsymbol{u}_0^*)\boldsymbol{\pi}_\mathrm{p} - \boldsymbol{\pi}_\mathrm{p}(\boldsymbol{u}_{(I-1)}\star_\varepsilon\boldsymbol{u}^*_{(I-1)} - \mathbf{1}_\mathrm{f})\boldsymbol{\pi}_\mathrm{p} + \mathcal{O}_0(\varepsilon^{I+1})\\ &= \boldsymbol{\pi}_\mathrm{p}((\boldsymbol{u}_{(I-1)}\star_\varepsilon\boldsymbol{\pi}_{(I-1)}\star_\varepsilon\boldsymbol{u}^*_{(I-1)} - \boldsymbol{u}_{(I-1)}\star_\varepsilon(\boldsymbol{\pi}_{(I-1)}\star_\varepsilon\boldsymbol{\pi}_{(I-1)} - \boldsymbol{\pi}_{(I-1)})\star_\varepsilon\boldsymbol{u}^*_{(I-1)})\boldsymbol{\pi}_\mathrm{p} - \boldsymbol{\pi}_\mathrm{p}(\boldsymbol{u}_{(I-1)}\star_\varepsilon\boldsymbol{u}^*_{(I-1)})\boldsymbol{\pi}_\mathrm{p} + \mathcal{O}_0(\varepsilon^{I+1})\\ &= \boldsymbol{\pi}_\mathrm{p}\boldsymbol{u}_{(I-1)}\star_\varepsilon(-\boldsymbol{\pi}_{(I-1)}\star_\varepsilon\boldsymbol{\pi}_{(I-1)} + 2\boldsymbol{\pi}_{(I-1)} - \mathbf{1}_\mathrm{f})\star_\varepsilon\boldsymbol{u}^*_{(I-1)}\boldsymbol{\pi}_\mathrm{p} + \mathcal{O}_0(\varepsilon^{I+1})\\ &= -\boldsymbol{\pi}_\mathrm{p}\boldsymbol{u}_{(I-1)}\star_\varepsilon(\boldsymbol{\pi}_{(I-1)} - \mathbf{1}_\mathrm{f})\star_\varepsilon(\boldsymbol{\pi}_{(I-1)} - \mathbf{1}_\mathrm{f})\star_\varepsilon\boldsymbol{u}^*_{(I-1)}\boldsymbol{\pi}_\mathrm{p} + \mathcal{O}_0(\varepsilon^{I+1}).\end{aligned} \tag{47}$$

To show that the contributions of $\boldsymbol{R}_{ed}$ at order ε^I are indeed vanishing, we abbreviate the following $\mathcal{O}_0(\varepsilon^I)$ objects:

$$\boldsymbol{A}_{I-1} := \boldsymbol{\pi}_{(I-1)}\star_\varepsilon\boldsymbol{\pi}_{(I-1)} - \boldsymbol{\pi}_{(I-1)}, \tag{48}$$

$$\boldsymbol{C}_{I-1} := \boldsymbol{u}_{(I-1)}\star_\varepsilon\boldsymbol{u}^*_{(I-1)} - \mathbf{1}_\mathrm{f}, \tag{49}$$

$$\boldsymbol{D}_{I-1} := \boldsymbol{u}^*_{(I-1)}\star_\varepsilon\boldsymbol{u}_{(I-1)} - \mathbf{1}_\mathrm{f}, \tag{50}$$

$$\boldsymbol{E}_{I-1} := \boldsymbol{u}_{(I-1)}\star_\varepsilon\boldsymbol{\pi}_{(I-1)}\boldsymbol{u}^*_{(I-1)} - \boldsymbol{\pi}_\mathrm{p}. \tag{51}$$

With this, we continue to elaborate on $\varepsilon^I\boldsymbol{R}_{ed}$ starting from Eq. (47). In the first step, we use that $\boldsymbol{\pi}_\mathrm{p}$ does not depend on the slow phase space variables, and hence the operator product of any other quantity with $\boldsymbol{\pi}_\mathrm{p}$ equals their Moyal product. Consequently,

$$\begin{aligned}\varepsilon^I \boldsymbol{R}_{ed} &= -\boldsymbol{\pi}_\mathrm{p} \star_\varepsilon \boldsymbol{u}_{(I-1)} \star_\varepsilon (\boldsymbol{\pi}_{(I-1)} - \mathbf{1}_\mathrm{f}) \star_\varepsilon (\boldsymbol{\pi}_{(I-1)} - \mathbf{1}_\mathrm{f}) \star_\varepsilon \boldsymbol{u}^*_{(I-1)} \star_\varepsilon \boldsymbol{\pi}_\mathrm{p} + \mathcal{O}_0(\varepsilon^{I+1}) \\ &= -(\boldsymbol{u}_{(I-1)} \star_\varepsilon \boldsymbol{\pi}_{(I-1)} \star_\varepsilon \boldsymbol{u}^*_{(I-1)} - \boldsymbol{E}_{I-1}) \star_\varepsilon \boldsymbol{u}_{(I-1)} \star_\varepsilon (\boldsymbol{\pi}_{(I-1)} - \mathbf{1}_\mathrm{f}) \star_\varepsilon (\boldsymbol{\pi}_{(I-1)} - \mathbf{1}_\mathrm{f}) \\ &\quad \star_\varepsilon \boldsymbol{u}^*_{(I-1)} \star_\varepsilon (\boldsymbol{u}_{(I-1)} \star_\varepsilon \boldsymbol{\pi}_{(I-1)} \star_\varepsilon \boldsymbol{u}^*_{(I-1)} - \boldsymbol{E}_{I-1}) + \mathcal{O}_0(\varepsilon^{I+1}) \\ &= -\boldsymbol{u}_{(I-1)} \star_\varepsilon \boldsymbol{\pi}_{(I-1)} \star_\varepsilon \boldsymbol{u}^*_{(I-1)} \star_\varepsilon \boldsymbol{u}_{(I-1)} \star_\varepsilon (\boldsymbol{\pi}_{(I-1)} - \mathbf{1}_\mathrm{f})^2_{\star_\varepsilon} \star_\varepsilon \boldsymbol{u}^*_{(I-1)} \star_\varepsilon \boldsymbol{u}_{(I-1)} \star_\varepsilon \boldsymbol{\pi}_{(I-1)} \star_\varepsilon \boldsymbol{u}^*_{(I-1)} \\ &\quad + \boldsymbol{E}_{I-1} \star_\varepsilon \boldsymbol{u}_{(I-1)} \star_\varepsilon (\boldsymbol{\pi}_{(I-1)} - \mathbf{1}_\mathrm{f})^2_{\star_\varepsilon} \star_\varepsilon \boldsymbol{u}^*_{(I-1)} \star_\varepsilon \boldsymbol{u}_{(I-1)} \star_\varepsilon \boldsymbol{\pi}_{(I-1)} \star_\varepsilon \boldsymbol{u}^*_{(I-1)} \\ &\quad + \boldsymbol{u}_{(I-1)} \star_\varepsilon \boldsymbol{\pi}_{(I-1)} \star_\varepsilon \boldsymbol{u}^*_{(I-1)} \star_\varepsilon \boldsymbol{u}_{(I-1)} \star_\varepsilon (\boldsymbol{\pi}_{(I-1)} - \mathbf{1}_\mathrm{f})^2_{\star_\varepsilon} \star_\varepsilon \boldsymbol{u}^*_{(I-1)} \star_\varepsilon \boldsymbol{E}_{I-1} + \mathcal{O}_0(\varepsilon^{I+1}), \end{aligned} \tag{52}$$

where we pushed the terms that are quadratic in $\boldsymbol{E}_{I-1}$ in the remainder $\mathcal{O}_0(\varepsilon^{I+1})$. This strategy applies for any quadratic occurrence of the symbols in (48)–(51), such that the continuation of (52) becomes

$$\begin{aligned}\varepsilon^I \boldsymbol{R}_{ed} &= -\boldsymbol{u}_{(I-1)} \star_\varepsilon \boldsymbol{\pi}_{(I-1)} \star_\varepsilon (\boldsymbol{D}_{I-1} + \mathbf{1}_\mathrm{f}) \star_\varepsilon (\boldsymbol{\pi}_{(I-1)} - \mathbf{1}_\mathrm{f})^2_{\star_\varepsilon} \star_\varepsilon (\boldsymbol{D}_{I-1} + \mathbf{1}_\mathrm{f}) \star_\varepsilon \boldsymbol{\pi}_{(I-1)} \star_\varepsilon \boldsymbol{u}^*_{(I-1)} \\ &\quad + \boldsymbol{E}_{I-1} \star_\varepsilon \boldsymbol{u}_{(I-1)} \star_\varepsilon (\boldsymbol{\pi}_{(I-1)} - \mathbf{1}_\mathrm{f})^2_{\star_\varepsilon} \star_\varepsilon (\boldsymbol{D}_{I-1} + \mathbf{1}_\mathrm{f}) \star_\varepsilon \boldsymbol{\pi}_{(I-1)} \star_\varepsilon \boldsymbol{u}^*_{(I-1)} \\ &\quad + \boldsymbol{u}_{(I-1)} \star_\varepsilon \boldsymbol{\pi}_{(I-1)} \star_\varepsilon (\boldsymbol{D}_{I-1} + \mathbf{1}_\mathrm{f}) \star_\varepsilon (\boldsymbol{\pi}_{(I-1)} - \mathbf{1}_\mathrm{f})^2_{\star_\varepsilon} \star_\varepsilon \boldsymbol{u}^*_{(I-1)} \star_\varepsilon \boldsymbol{E}_{I-1} + \mathcal{O}_0(\varepsilon^{I+1}). \end{aligned}$$

In any of the lines, we can directly eliminate those terms that are quadratic in the operators $\boldsymbol{D}_{I-1}$ and $\boldsymbol{E}_{I-1}$. All the remaining terms admit at least one factor of the form $\boldsymbol{\pi}_{(I-1)} \star_\varepsilon (\boldsymbol{\pi}_{(I-1)} - \mathbf{1}_\mathrm{f})$ or the same with factors interchanged. These factors simply evaluate to $\boldsymbol{A}_{I-1}$, and so all the contributions are at least quadratic in $\boldsymbol{A}_{I-1}$, $\boldsymbol{D}_{I-1}$, and $\boldsymbol{E}_{I-1}$, such that we obtain

$$\begin{aligned}\varepsilon^I \boldsymbol{R}_{ed} &= -\boldsymbol{u}_{(I-1)} \star_\varepsilon \boldsymbol{\pi}_{(I-1)} \star_\varepsilon \boldsymbol{D}_{I-1} \star_\varepsilon (\boldsymbol{\pi}_{(I-1)} - \mathbf{1}_\mathrm{f}) \star_\varepsilon \boldsymbol{A}_{I-1} \star_\varepsilon \boldsymbol{u}^*_{(I-1)} \\ &\quad - \boldsymbol{u}_{(I-1)} \star_\varepsilon \boldsymbol{A}_{I-1} \star_\varepsilon (\boldsymbol{\pi}_{(I-1)} - \mathbf{1}_\mathrm{f}) \star_\varepsilon \boldsymbol{D}_{I-1} \star_\varepsilon \boldsymbol{\pi}_{(I-1)} \star_\varepsilon \boldsymbol{u}^*_{(I-1)} \\ &\quad - \boldsymbol{u}_{(I-1)} \star_\varepsilon \boldsymbol{A}_{I-1} \star_\varepsilon \boldsymbol{A}_{I-1} \star_\varepsilon \boldsymbol{u}^*_{(I-1)} \\ &\quad + \boldsymbol{E}_{I-1} \star_\varepsilon \boldsymbol{u}_{(I-1)} \star_\varepsilon (\boldsymbol{\pi}_{(I-1)} - \mathbf{1}_\mathrm{f}) \star_\varepsilon \boldsymbol{A}_{I-1} \star_\varepsilon \boldsymbol{u}^*_{(I-1)} \\ &\quad + \boldsymbol{u}_{(I-1)} \star_\varepsilon \boldsymbol{A}_{I-1} \star_\varepsilon (\boldsymbol{\pi}_{(I-1)} - \mathbf{1}_\mathrm{f}) \star_\varepsilon \boldsymbol{u}^*_{(I-1)} \star_\varepsilon \boldsymbol{E}_{I-1} + \mathcal{O}_0(\varepsilon^{I+1}) \\ &= \mathcal{O}_0(\varepsilon^{I+1}). \end{aligned} \tag{53}$$

The very same argumentation leads to the vanishing of the orthogonal part $\boldsymbol{R}^\perp_{ed}$, namely

$$\varepsilon^I \boldsymbol{R}^\perp_{ed} = \varepsilon^I \boldsymbol{\pi}^\perp_\mathrm{p} (\boldsymbol{e}_{I-1} + \boldsymbol{u}_0 \boldsymbol{\pi}_I \boldsymbol{u}^*_0 - \boldsymbol{d}_{I-1}) \boldsymbol{\pi}^\perp_\mathrm{p} = \mathcal{O}_0(\varepsilon^{I+1}). \tag{54}$$

Equations (53) and (54) consequently show that the block diagonal parts of $\boldsymbol{u}_I$ with respect to $\boldsymbol{\pi}_\mathrm{p}$ remain undetermined. Without loss of generality, we can choose them to vanish, also because $\boldsymbol{u}$ is only an auxiliary structure here. Finally, we project on the off-diagonal contributions to $\boldsymbol{u}_I$ and obtain with (43) and (46)

$$\begin{aligned}\boldsymbol{\pi}_\mathrm{p} \boldsymbol{u}_I \boldsymbol{u}^*_0 \boldsymbol{\pi}^\perp_\mathrm{p} &= \boldsymbol{\pi}_\mathrm{p} (\boldsymbol{e}_{I-1} + \boldsymbol{u}_0 \boldsymbol{\pi}_I \boldsymbol{u}^*_0 - \boldsymbol{d}_{I-1}) \boldsymbol{\pi}^\perp_\mathrm{p} \\ &= \boldsymbol{\pi}_\mathrm{p} (\boldsymbol{e}_{I-1} + \boldsymbol{u}_0 \boldsymbol{b}_{I-1} (E_\nu \mathbf{1}_\mathrm{f} - \boldsymbol{H}^\perp_0)^{-1} \boldsymbol{u}^*_0 - \boldsymbol{d}_{I-1}) \boldsymbol{\pi}^\perp_\mathrm{p}, \end{aligned} \tag{55}$$

$$\begin{aligned}\boldsymbol{\pi}^\perp_\mathrm{p} \boldsymbol{u}_I \boldsymbol{u}^*_0 \boldsymbol{\pi}_\mathrm{p} &= -\boldsymbol{\pi}^\perp_\mathrm{p} (\boldsymbol{e}_{I-1} + \boldsymbol{u}_0 \boldsymbol{\pi}_I \boldsymbol{u}^*_0 - \boldsymbol{d}_{I-1}) \boldsymbol{\pi}_\mathrm{p} \\ &= -\boldsymbol{\pi}^\perp_\mathrm{p} (\boldsymbol{e}_{I-1} - \boldsymbol{u}_0 (E_\nu \mathbf{1}_\mathrm{f} - \boldsymbol{H}^\perp_0)^{-1} \boldsymbol{b}_{I-1} \boldsymbol{u}^*_0 - \boldsymbol{d}_{I-1}) \boldsymbol{\pi}_\mathrm{p}. \end{aligned} \tag{56}$$

Eventually, the inductive equation for $\boldsymbol{u}_I$ evaluates to

$$\begin{aligned}\boldsymbol{u}_I &= \boldsymbol{\pi}_\mathrm{p} (\boldsymbol{e}_{I-1} + \boldsymbol{u}_0 \boldsymbol{b}_{I-1} (E_\nu \mathbf{1}_\mathrm{f} - \boldsymbol{H}^\perp_0)^{-1} \boldsymbol{u}^*_0 - \boldsymbol{d}_{I-1}) \boldsymbol{\pi}^\perp_\mathrm{p} \boldsymbol{u}_0 \\ &\quad - \boldsymbol{\pi}^\perp_\mathrm{p} (\boldsymbol{e}_{I-1} - \boldsymbol{u}_0 (E_\nu \mathbf{1}_\mathrm{f} - \boldsymbol{H}^\perp_0)^{-1} \boldsymbol{b}_{I-1} \boldsymbol{u}^*_0 - \boldsymbol{d}_{I-1}) \boldsymbol{\pi}_\mathrm{p} \boldsymbol{u}_0. \end{aligned} \tag{57}$$

3. Construction of the effective Hamiltonian

The last step of the scheme consists in computing the effective Hamilton symbol $\boldsymbol{h}_\mathrm{eff}$ and, in particular, the effective Hamiltonian restricted to the fast subspace associated with the projector $\boldsymbol{\pi}_\mathrm{p}$ which will be denoted by $\boldsymbol{h}_{\mathrm{eff,p}}$. As before, the scheme proceeds iteratively, and we have [just repeating the construction rule (S3)"]

$$\boldsymbol{h}_{\mathrm{eff},(I)} = \boldsymbol{u}_{(I)} \star_\varepsilon \boldsymbol{H} \star_\varepsilon \boldsymbol{u}^*_{(I)} + \mathcal{O}_0(\varepsilon^{I+1}), \tag{58}$$

where we can insert $\boldsymbol{u}_{(I)}$ from the previous section. As we are mainly interested in the dynamics within the fast subspace associated with the quantum number $\nu \in \mathbb{N}$, we consider the restriction

$$\boldsymbol{h}_{\mathrm{eff,p},(I)} = \boldsymbol{\pi}_{\mathrm{p}}\boldsymbol{u}_{(I)} \star_{\varepsilon} \boldsymbol{H} \star_{\varepsilon} \boldsymbol{u}^{*}_{(I)}\boldsymbol{\pi}_{\mathrm{p}} + \mathcal{O}_0(\varepsilon^{I+1}). \quad (59)$$

We emphasize that the Weyl quantization $\hat{\boldsymbol{h}}_{\mathrm{eff,p},(I)}$ preserves the subspace $\hat{\boldsymbol{\pi}}_{\mathrm{p}}\mathcal{H}$. This subspace carries the orthonormal basis $\{\zeta_{n,a} \otimes \psi_\alpha\}_{n,a,\alpha}$ where $n \in \mathbb{N}$ is the discrete quantum number of the fast eigenstates $\zeta_{n,a}(x) \in \mathcal{H}_{\mathrm{f}}$ with a possible degeneracy label $a = 1, \ldots, d_n$, and $\psi_\alpha(q)$ denotes a (possibly generalized) orthogonal basis of $\mathcal{H}_{\mathrm{s}}$. As already shown before, the spectrum of $\hat{\boldsymbol{h}}_{\mathrm{eff},(I)}$ gives an approximation of order ϵ^{I+1} of the corresponding energy band of the original Hamiltonian $\hat{\boldsymbol{H}}$. The advantage of $\hat{\boldsymbol{h}}_{\mathrm{eff},(I)}$ is that it is effectively an operator on the rather small Hilbert space $\mathbb{C}^{d_n} \otimes \mathcal{H}_{\mathrm{s}}$ while backreaction effects between the slow and fast sectors are taken care of to the given order of approximation.

III. CHALLENGES OF COSMOLOGICAL ADIABATIC PERTURBATION THEORY

A. Introduction

As already pointed out, SAPT in its original form applies to quantum systems with a finite number of degrees of freedom. Consequently, the generalization of the scheme to second order quantum cosmological perturbation theory meets several challenges as the primary assumptions do not automatically transfer to the quantum field theory case. One important reason for this is the infinite number of inhomogeneous modes present in quantum field theory on cosmological spacetimes. The resulting problems need to be fixed before applying SAPT.

Actually, the problems we encounter here are not new. We use the Hilbert-Schmidt condition to illustrate this: In particular, the Fock spaces for a QFT on a possibly curved, cosmological spacetime fail to be unitary equivalent for different "background" configurations of the homogeneous sector. The unitary equivalence of these Fock spaces is, however, a necessary prerequisite if we wish to apply SAPT. There, the fast subspaces are assumed to represent the fiber spaces over the common slow phase space, and the fast subspaces are all unitarily equivalent, which is *a priori* not satisfied for QFT on CST models. As we will see, these kinds of problems originating from the infinite number of degrees of freedom can be circumvented using a transformation of the total system involving both the homogeneous and inhomogeneous modes. This transformation is an exact canonical transformation up to second order in the cosmological perturbations. The idea for these transformations originally comes from the hybrid approach to LQC [64–66].

These transformations lead us to the next obstacle encountered here, which concerns the occurrence of indefinite mass squared functions of the quantum fields. In fact, the same problems already occur for standard gauge-invariant cosmological perturbation theory, and are hence *not due to the SAPT scheme itself*. In fact, already the transformations giving rise to the gauge-invariant Mukhanov-Sasaki scalar variables lead to effective mass terms that depend very nontrivially on the homogeneous phase space variables. The above-mentioned transformations to meet the Hilbert-Schmidt condition have the same effect: The mass squared functions of both the Mukhanov-Sasaki field and the tensor perturbations are generically neither positive definite nor are they polynomials in the homogeneous degrees of freedom.

In the following, we will present a number of possible solutions. First, there is some amount of freedom in the choice of the almost canonical transformation alluded to before that might help to avoid the indefiniteness of the mass squared term. As another proposal, one can consider a canonical transformation just of the slow sector after which the masses are manifestly non-negative. Then, one declares the phase space as defined in terms of the old variables to be restricted by the positivity of the masses. Another idea is to switch off modes by hand for which the frequency becomes negative, as suggested first in [93]. We also point to Ref. [94] in which the authors were able to in principle solve the tachyonic problem by using transformations that lead to positive definite mass terms. Nonetheless, these transformations and hence also the mass functions are only known up to solving a system of semilinear partial differential equations. For SAPT, we need, however, the explicit expression of the mass functions.

Regarding the nonpolynomiality of the mass terms, we point at the solution strategy for a certain class of operators adopted in LQC. There, it is possible to deal with operators that are nonpolynomial in the momenta of the theory but polynomial with respect to the configuration variables. The basic idea consists in using a representation of the homogeneous sector which is not unitarily equivalent to the Schrödinger representation because one substitutes the unbounded configuration variable by a bounded polynomial of Weyl operators [70–72]. These techniques are inspired by the full theory of LQG [95,96]. Here, however, a simple inspection of the Mukhanov-Sasaki mass squared term reveals that the techniques of LQC for the purely homogeneous contribution to the Hamiltonian constraint and for the hybrid approach to it [64–66] will no longer be sufficient when backreactions are switched on.

This is due to the fact that the adiabatic corrections introduce nonpolynomial functions of *both* momentum and configuration variables. We handle this second problem by taking an unbiased point of view toward quantization of the homogeneous sector and try to stay within the standard Schrödinger representation as suggested by the Weyl quantization method that enters the space adiabatic formalism. In the best case, one should find a dense and invariant domain for the various nonpolynomial operators that appear. This is indeed possible for the model that also involves Gaussian dust [82] by exploiting the existence of a remarkable basis of functions in $L^2(\mathbb{R}, \mathrm{d}q)$ which is smooth and of rapid

decrease both at infinity and at the origin [97]. For the case of the Mukhanov-Sasaki field, we have to content ourselves by providing a dense but not invariant domain.

This section covers all the above-mentioned questions in detail. In Sec. III B, we show that the gravity-matter systems considered in the subsequent papers of this series allow for the definition of a suitable perturbation parameter ε which makes the SAPT scheme work at the technical level. We also show that the inhomogeneous modes can be associated with a small mass while the homogeneous degrees of freedom have a much larger effective mass. In Sec. III C, we display the aforementioned obstacle to apply SAPT in the quantum field context. Therefore, we first prove that the Hilbert-Schmidt condition fails for a simple QFT model. Then, we present a solution up to second order in the perturbation theory of the inhomogeneous, cosmological degrees of freedom. In Sec. III D, we discuss the induced occurrence of nonpositive and nonpolynomial mass squared functions. We apply two solution techniques, based on modifying either the classical slow phase space or the number of physical modes in the Fock space. Finally, in Sec. III E, we sketch a proposal for how to deal with the resulting highly nonpolynomial operators that occur as a result of the canonical transformations and the adiabatic corrections due to the Moyal product.

B. Applicability of space adiabatic perturbation theory

SAPT relies on the identification of two distinct subsystems within the model. This distinction becomes manifest by means of an adiabatic perturbation parameter which we denote by ε. For example, in the standard Born-Oppenheimer approach to molecular physics, the adiabatic perturbation parameter arises as the mass ratio of the light electron mass m_e and the heavy nuclei mass M_n such that $\varepsilon^2 \coloneqq \frac{m_\mathrm{e}}{M_\mathrm{n}} \ll 1$. For simplicity, let us consider a simple hydrogen model with only one electron and one nucleus with respective phase space variables (x, y) and (q, P). With the equipartition theorem and assuming that the system is ergodic, it follows that the kinetic energy contributions of the electron and the nucleus must have the same time average, i.e.,

$$\left\langle \frac{P^2}{2M_\mathrm{n}} \right\rangle \approx \left\langle \frac{y^2}{2m_\mathrm{e}} \right\rangle, \text{ and consequently } \langle \dot{q}^2 \rangle \approx \varepsilon^2 \langle \dot{x}^2 \rangle. \quad (60)$$

Evidently, the nucleus then moves much slower on average than the electron. This also implies that the momentum of the nucleus is much larger than the momentum of the electron, i.e., $\varepsilon^2 \langle P^2 \rangle \approx \langle y^2 \rangle$. This suggests to define a rescaled momentum for the nucleus $p \coloneqq \varepsilon P$ such that $\langle p^2 \rangle \approx \langle y^2 \rangle$. The Poisson bracket of q and p consequently receives an additional factor ε which also carries over to the quantum mechanical commutator. As the space adiabatic scheme employs a phase space quantization for the slow subsystem, this rescaling indeed gives rise to a perturbative treatment of the theory. In particular, the Moyal product on phase space appears then as a formal power series expansion in the perturbation parameter ε. The theory additionally requires that the Hamiltonian with the rescaled momenta p does *not* carry any inverse powers of ε.

This works out for the molecular systems but the situation in cosmology is different. In general, the arguments for deriving the smallness of the nucleus' velocity do not apply anymore because the systems are constrained. The equipartition theorem fails. Nevertheless, it remains possible to rescale one or several of the theory's momenta to make the perturbation scheme work at the technical level.

1. Implementation of a perturbation parameter

In order to clarify this, we recall the different systems of interest in [69,81,82]. In total, we consider four different models which we will label with latin numbers. First, a set of two coupled oscillators (I) serves as a toy model for the subsequent cosmological applications. We then examine a model of homogeneous and isotropic general relativity coupled to the homogeneous mode of a scalar matter field (II). We extend this to an inhomogeneous theory by considering a homogeneous, isotropic geometry coupled to all modes of the scalar matter field and in the presence of a deparametrizing dust field (III). Finally, we study a model of homogeneous and inhomogeneous general relativity with scalar matter in gauge-invariant variables but without the deparametrizing matter fields (IV).

Model (I) consists of a set of two coupled harmonic and anharmonic oscillators [81]. Their respective masses m and M give rise to the perturbation parameter $\varepsilon^2 \coloneqq \frac{m}{M} \ll 1$. The system is unconstrained, and the physical intuition that the heavy anharmonic oscillator moves slowly in comparison to the light harmonic oscillator follows from the equipartition theorem. That is, one can show that the average velocity of the heavy oscillator $\dot{q}$ is much smaller than the average velocity $\dot{x}$ of the light oscillator. Note that this is independent of the frequencies of the two oscillators. Even if the frequency Ω of the heavy anharmonic oscillator was much larger than the frequency ω of the light oscillator, it would still hold true that the velocity of the heavy subsystem is much *smaller* than the velocity of the light oscillator (at least their time averages). This is possible because the amplitude A_x of the light oscillator can be much larger than the amplitude A_q of the heavy subsystem. Therefore, consider the relation between the velocity and the frequency of the two subsystems, namely $\dot{q} = A_q \Omega$ and $\dot{x} = A_x \omega$. Then, with $\dot{q} \approx \varepsilon \dot{x}$ in time average coming from the equipartition theorem, it can still hold true that for example $\omega = \varepsilon \Omega$ (i.e., the heavy oscillator frequency is much larger than the one of the light subsystem) by claiming that $A_q \approx \varepsilon^2 A_x$.

The purely homogeneous and isotropic cosmological model with scalar matter content (II) does not feature two mass parameters in the strict sense. However, the gravitational coupling constant $\kappa := 8\pi G$ and the matter coupling constant λ of the scalar field provide a dimensionless fraction which gives rise to a perturbation parameter $\varepsilon^2 := \frac{\kappa}{\lambda}$. Note that λ generically appears as a global factor for the matter (inflaton) action and is set to unity in most applications to inflationary cosmology. Here, we do not restrict its value to one and thereby simply introduce a reparametrization invariance for the scalar field itself. This is of course compliant with the conventional considerations in inflation and does not alter any of the physical results by itself. We also emphasize that the parameter λ is to be distinguished from the mass parameter m of the Klein-Gordon or inflaton scalar field in this model which is an independent parameter. In the space adiabatic treatment, we assume that $\varepsilon^2 \ll 1$. This is admissible since λ is a free parameter. Inserting this parameter into the Hamiltonian, one must, however, guarantee that ε does not occur in inverse powers as this would impede the perturbation theory. Interestingly, when the Hamiltonian is a constraint, we can simply multiply the constraint by sufficiently large powers of ε. This applies partly to model (II) as it can be seen as the homogeneous truncation of either model (III) or model (IV). In the latter case, the system is constrained and the strategy applies. Note that we can always multiply the constraint by ε before taking the limit $\varepsilon \to 0$ in the result.

In the former case, however, we have to inspect the Hamiltonian more closely. As it turns out (see Ref. [81]), there is one contribution proportional to ε^{-2}, and this is exactly the term due to the cosmological constant Λ. According to cosmological observations, we can hence argue that $\Lambda\varepsilon^{-2}$ is still so small that the inverse order in ε does not affect the theory severely. In any case, it seems admissible to neglect the cosmological constant at earliest times as it scales as a^0 with a being the scale factor.

Nonetheless, it would be interesting to see whether the theory allows for the usual distinction of a slow and a fast sector. Therefore, we first note that the equipartition theorem does not work here. To see this, we introduce the phase space $\Gamma = \mathbb{R}^+ \times \mathbb{R}^3$ which is parametrized by the scale factor a, its conjugate momentum P_a, the homogeneous scalar field ϕ, and its conjugate momentum μ. Then, consider the statistical average of some phase space function $f \in C^\infty(\Gamma)$ on the full phase space Γ,

$$\langle f\rangle := \frac{1}{Z}\int_\Gamma \mathrm{d}a\mathrm{d}P_a\mathrm{d}\phi\mathrm{d}\mu e^{-\beta H} f(a,P_a,\phi,\mu),$$
$$Z := \int_\Gamma \mathrm{d}a\mathrm{d}P_a\mathrm{d}\phi\mathrm{d}\mu e^{-\beta H},$$

where Z is the partition function and $\beta := (k_\mathrm{B}T)^{-1}$ is the reciprocal of the thermodynamic temperature of the system. In order that the statistical average of f makes sense, we must assume that H is bounded from below such that the integrals converge. This also assures that in the integration by parts that one performs to show that $\langle p_a \frac{\partial H}{\partial p_a}\rangle = \beta^{-1}$ no boundary terms appear. Both conditions are violated in model (II) because the gravitational kinetic energy is negative. Nevertheless, we can make the following argument for which we assume that the underlying space manifold is a compact three-torus $\mathbb{T}^3$ with volume L^3. By shifting the kinetic energy of the geometric subsystem and multiplying by the global factor $L^3\lambda$, we obtain

$$\frac{1}{12}\frac{\varepsilon^2 P_a^2}{a} = \frac{\Lambda}{\lambda^2\varepsilon^2}a^3 + \frac{\mu^2}{2a^3} + \frac{1}{2\lambda}m^2a^3\phi^2. \tag{61}$$

Note that in our models, we choose ϕ to be dimension-free. Our motivation for this choice is that we do not want to introduce additional mass scales into more general than quadratic inflaton potentials. As a typical measure of the expansion velocity of the Universe, we introduce the Hubble parameter $H := \frac{\dot a}{a}$ which we accordingly denote as v_H. It relates to the variables (a, P_a) according to $v_H = -\varepsilon^2\frac{\lambda}{6L^3}\frac{P_a}{a^2}$. Likewise, the velocity of the scalar field is given by $v_\phi := \dot\phi = \lambda\frac{\mu}{a^3}$. Inserting this in the constraint (61) gives

$$v_H^2 = \frac{1}{6}\varepsilon^2 v_\phi^2 + \frac{1}{3}\Lambda + \frac{1}{6}\varepsilon^2 m^2\phi^2. \tag{62}$$

A comparison of the respective terms allows for the conclusion that for small cosmological constant and small potentials of the scalar field, we have indeed that the geometric velocity is much smaller than the velocity of the scalar field, namely $v_H \sim \varepsilon v_\phi$. Accordingly, the homogeneous mode of the scalar field can be identified as the fast sector while geometry appears as the slow subsystem owing to our assumption on κ and λ. It transpires, however, that this only holds for small cosmological constant and scalar field potentials. As a consequence, this intuition does not apply for an inflationary phase of the scalar field because there the scalar field potential is assumed to be large.

In model (III), we consider general relativity with a cosmological constant $\Lambda \in \mathbb{R}$ and a real-valued scalar field Φ of Klein-Gordon type. The theory is built on a globally hyperbolic spacetime manifold $\mathcal{M} \cong \mathbb{R}\times\mathcal{B}$ with compact spatial hypersurface $\mathcal{B}$. We reduce the gravitational field to its purely homogeneous and isotropic part with respect to the spatial hypersurfaces such that the scale factor a and its conjugate momentum P_a represent the relevant canonical pair of the gravitational field in the Hamiltonian framework. Indeed, for model (III) it is not useful to divide the scalar field into a homogeneous and an inhomogeneous part with respect to the spatial hypersurfaces. This is because we include additional Gaussian dust to deparametrize the theory and such that the scalar field becomes a gauge-invariant degree of freedom. The constraints have already

been solved using reduced phase space methods. As a consequence, there is no motivation from the gauge perspective to split the scalar field into a homogeneous part and the rest as it is the case for the hybrid treatment of model (IV).

However, this model also exhibits new challenges due to the inhomogeneous modes of the scalar field, and we already pointed out that the Fock space representations of the inhomogeneous scalar field are not unitarily equivalent for different values of the homogeneous sector. In order to overcome this problem, we recall that we perform a canonical transformation up to second order in the perturbations [82]. Thereby, the mass value m of the Klein-Gordon field that appears in the frequency of the field transforms into a new "effective" mass value or rather mass function $M(a, P_a)$. It displays a phase space dependence with respect to the homogeneous gravitational degrees of freedom. As it turns out, this dependence on a and P_a is such that the momentum P_a only appears in the combination $\kappa^2 P_a^2 = \varepsilon^4\lambda^2 P_a^2 = \varepsilon^2\lambda^2 p_a^2$ if we define $p_a \coloneqq \varepsilon P_a$ as before. Hence, only positive powers of ε appear in the effective mass squared M^2 when expressed in terms of p_a. Since the Hamiltonian is still not bounded from below, the equipartition theorem can again not be used to argue that the homogeneous geometry represents the slow sector of the theory compared to all scalar field modes. However, the physical Hamiltonian H (not to be confused with the Hubble parameter) results from a constraint of the form $C = \rho + H = 0$ where ρ is the energy density of the dust. As the dust behaves closely to a field of test observers with zero energy density, ρ is assumed to be very small, and thus H is close to zero, at least classically. Thus, we qualitatively argue that $H \approx 0$ such that we can effectively apply the same strategy as above and identify the corresponding regions in phase space. More precisely, the system including dust has the property that the energy-momentum density of the nondust degrees of freedom are constants of motion with regard to the physical time evolution (see the in-depth discussion in [98,99]). As the sum of dust and nondust energy momentum density is constrained to vanish, small dust energy momentum density is equivalent to small nondust energy momentum density. Thus in any given classical solution the nonvanishing of the nondust contribution to the constraints is given by a constant of motion which acts as an additional (not necessarily spatially homogeneous) cosmological "constant." If it is small, the classical argument of the previous model applies. In the quantum theory, since energy density commutes with the physical Hamiltonian, we can use the spectral theorem to split the Hilbert space into sectors of finite energy density range, and these sectors are preserved by the quantum evolution. We also point out that there is an argument leading to the general statement that any kind of inhomogeneous mode behaves as a fast system compared to the homogeneous sector.

Before we come to this, let us discuss model (IV) which is gravity plus a real-valued scalar field Φ, just as for model (III). Here, however, we allow for perturbations of the gravitational field, and we do not include dust fields to deparametrize the theory. Therefore, it is advisable to split the Klein-Gordon scalar field into its homogeneous and inhomogeneous components with respect to the spatial hypersurface $\mathcal{B}$. We denote these new variables by (ϕ, μ) and (φ, π), respectively. Following the idea of Mukhanov and Sasaki [68,100,101], we perform a canonical transformation with respect to the scalar sector of the linear perturbations in order to obtain gauge-invariant variables. These transformations depend also on the homogeneous, zeroth order variables (a, P_a, ϕ, μ). These new scalar variables are the well-known Mukhanov-Sasaki variables $(\vartheta, \pi_\vartheta)$. The corresponding transformations give rise to a new effective mass squared term $M_{\mathrm{MS}}(a, p_a, \phi, \mu)^2$ for the Mukhanov-Sasaki field. Besides, we include the tensorial part of the linear perturbations (t^{ab}, π^t_{ab}) which is already gauge invariant. Then, we perform further transformations with respect to all degrees of freedom following the hybrid approach mentioned above [64,67,69], which also gives rise to a new mass term for the tensor sector $M_{\mathrm{T}}(a, p_a, \phi, \mu)^2$.

In order to check whether the application of SAPT is admissible, we take a look at the occurrence of a suitable perturbative parameter [69], and consider particularly the inhomogeneous contributions to the Hamiltonian constraint. We define the ratio of κ and the matter coupling constant λ as $\varepsilon^2 \coloneqq \frac{\kappa}{\lambda}$. Then, the second order contributions of the scalar constraint are given in terms of the scalar Mukhanov-Sasaki canonical pair $(\vartheta, \pi_\vartheta)$ and the tensorial canonical variables (t^{ab}, π^t_{ab}) by

$$\mathcal{H}^{\mathrm{s}}_2 = \frac{1}{2a}\left(\lambda\pi_\vartheta^2 + \vartheta\left(-\frac{\Delta}{\lambda} + M^2_{\mathrm{MS}}\right)\vartheta\right),$$
$$\mathcal{H}^{\mathrm{t}}_2 = \frac{1}{2a}\left(\kappa(\pi^{ab}_t\pi^t_{ab}) + t^{ab}(x)\left(-\frac{\Delta}{\kappa} + M^2_{\mathrm{T}}\right)t_{ab}\right), \tag{63}$$

where π_t and π^t denote the same tensorial field. We perform the canonical transformation, $\tilde{\pi}^{ab}_t \coloneqq \varepsilon\pi^{ab}_t$ and $\tilde{t}_{ab} = \varepsilon^{-1}t_{ab}$, such that $\mathcal{H}^{\mathrm{t}}_2$ becomes

$$\tilde{\mathcal{H}}^{\mathrm{t}}_2 = \frac{1}{2a}\left(\lambda(\tilde{\pi}^{ab}_t\tilde{\pi}^t_{ab})(x) + \tilde{t}^{ab}(x)\left(-\frac{\Delta}{\lambda} + (\varepsilon M_{\mathrm{T}})^2\right)\tilde{t}_{ab}(x)\right). \tag{64}$$

Fortunately, an explicit check reveals that both the Mukhanov-Sasaki mass squared M^2_{MS} as well as the tensor mass squared $(\varepsilon M_{\mathrm{T}})^2$ receive only non-negative powers of ε when expressed in terms of a new set of rescaled homogeneous momenta,

$$\tilde{p}_a \coloneqq \varepsilon^2 P_a = \varepsilon p_a, \qquad \tilde{\mu} \coloneqq \varepsilon\mu. \tag{65}$$

In particular, we rescale the geometric momentum by an additional factor of ε and also consider a rescaling of the homogeneous matter momentum by a factor ε. This is in contrast to the choices in model (II) where we simply used p_a and μ. Nonetheless, their respective ratios are still the same, namely $\frac{\tilde{p}_a}{\tilde{\mu}} = \frac{p_a}{\mu}$. Indeed, the theories are equivalent as we could have simply multiplied the whole constraint (II) by a factor ε^2 and then worked with $\tilde{p}_a$ and $\tilde{\mu}$. Now, the only remaining problem is that the homogeneous piece of the Hamiltonian constraint of model (IV) cannot be written in terms of $\tilde{p}_a$ and $\tilde{\mu}$ without picking up negative powers of ε. But this can be repaired by multiplying the *entire* constraint by a factor of ε^2. As a consequence, we perform another canonical transformation for the inhomogeneous fields

$$\breve{\pi}_\vartheta := \varepsilon\pi_\vartheta, \qquad \breve{\vartheta} := \frac{\vartheta}{\varepsilon}, \qquad \breve{\pi}_t^{ab} := \varepsilon\tilde{\pi}_t^{ab} = \varepsilon^2\pi_t^{ab},$$
$$\breve{t}_{ab} := \frac{\tilde{t}_{ab}}{\varepsilon} = \frac{t_{ab}}{\varepsilon^2}. \tag{66}$$

After multiplying the whole constraint by a factor of ε^2, the inhomogeneous contributions to the Hamilton constraint become

$$\breve{\mathcal{H}}_2^{\mathrm{s}} = \frac{1}{2a}\left(\lambda\breve{\pi}_\vartheta^2(x) + \breve{\vartheta}(x)\varepsilon^4\left(-\frac{\Delta}{\lambda} + M_{\mathrm{MS}}^2\right)\breve{\vartheta}(x)\right), \tag{67}$$

$$\breve{\mathcal{H}}_2^{\mathrm{t}} = \frac{1}{2a}\left(\lambda(\breve{\pi}_t^{ab}\breve{\pi}_{ab}^t)(x) + \breve{t}^{ab}(x)\varepsilon^4\left(-\frac{\Delta}{\lambda} + (\varepsilon M_{\mathrm{T}})^2\right)\breve{t}_{ab}(x)\right). \tag{68}$$

Note that this rescaling does not affect the physical problem at hand. In particular, the kinetic and the potential contributions to the given Hamilton functions still have the same relative size of magnitude. As far as the quantization of the theory is concerned, the inhomogeneous modes encounter a standard Fock quantization for both the Mukhanov-Sasaki and the tensor fields. The one-particle Hilbert space associated with these Fock spaces are simply the respective L^2-spaces over the compact three-torus. On this space of square-integrable functions, we define the frequency operator for the Mukhanov-Sasaki field according to $\omega_{\mathrm{MS}}^2 := (-\frac{\Delta}{\lambda} + M_{\mathrm{MS}}^2)$, and likewise for the graviton part. By means of these, the annihilation and creation operators (acting on the Fock space and therefore labeled by bold letters) formally read,

$$\breve{\boldsymbol{a}} := \frac{1}{\sqrt{2}}(\varepsilon\sqrt{\omega_{\mathrm{MS}}}\breve{\boldsymbol{\vartheta}} - i(\varepsilon\sqrt{\omega_{\mathrm{MS}}})^{-1}\breve{\boldsymbol{\pi}}_\vartheta),$$
$$\breve{\boldsymbol{a}}^* := \frac{1}{\sqrt{2}}(\varepsilon\sqrt{\omega_{\mathrm{MS}}}\breve{\boldsymbol{\vartheta}} + i(\varepsilon\sqrt{\omega_{\mathrm{MS}}})^{-1}\breve{\boldsymbol{\pi}}_\vartheta),$$

and we emphasize that ω_{MS} as well as $\breve{\boldsymbol{a}}$ and $\breve{\boldsymbol{a}}^*$ depend on the homogeneous degrees of freedom. Remarkably, $\breve{\boldsymbol{a}}$ is exactly the same annihilation operator as one would have defined before that last canonical transformation (i.e., introducing again ϑ, π_ϑ). Thus, the Fock representations are indeed *identical*. When normal ordered, one finds that the Mukhanov-Sasaki Hamiltonian in terms of the transformed variables becomes

$$\breve{\boldsymbol{C}}_2^{\mathrm{s}} = \varepsilon^2\int_{\mathbb{T}^3} \mathrm{d}x\breve{\boldsymbol{a}}^*\omega_{\mathrm{MS}}\breve{\boldsymbol{a}}. \tag{69}$$

Consequently, the spectrum of the inhomogeneous part of the constraint becomes simply rescaled by ε^2 and one can use all the results of the original Fock representation. This is in fact neat as one would expect that the homogeneous modes alone as corresponding to model (II) and multiplied by ε^2 remains almost undisturbed by the inhomogeneous modes. In summary, all that remains to be done is a simple rescaling as proposed in Eq. (66) in order to use SAPT for model (IV), and this is in fact consistent with the treatment of model (II). Actually, we can also treat model (III) consistently this way by simply multiplying (and, to get the correct spectrum, afterwards dividing) by ε^2. Ergo, it is possible to work consistently in all models (II), (III), and (IV) by using the breved variables from the definition (66).

Finally, let us emphasize that the Hamiltonian in Eq. (69) does *not* admit an instantaneous diagonalization as the creation and annihilation operators depend manifestly on the homogeneous degrees of freedom. In contrast, in simplified terms, what the adiabatic machinery achieves is to write the total Hilbert space as a Hilbert bundle over the homogeneous phase space where each fiber corresponds to a Fock space over the corresponding point in this phase space, with all dependence of that Fock fiber space (vacuum, annihilators, etc.) on that point fully taken into account. The aforementioned Hilbert-Schmidt condition ensures that the fibered Fock representations define the same equivalence class.

2. Identification of effective masses

As already indicated, it would be helpful to have a rationale why it is physically reasonable to interpret the inhomogeneous field modes in model (IV) as fast variables while the homogeneous degrees of freedom can be considered as slow variables. The same reasoning could be applied to model (III) although there it would be sufficient that all Klein-Gordon field modes are equally fast. It turns out that the answer to the question lies in the definition of the modes. Therefore, let us consider the scalar field Φ defined on $\mathcal{M}$. By the homogeneous mode of the field, we loosely speaking mean a component of the field which does not depend on position. This becomes unambiguous only when relating it to the full field. We choose again a compact spatial manifold $\mathcal{B}$ and assume that its topology is flat. We can consider the flat three-torus $\mathbb{T}^3$ without any loss of generality. Indeed, this case comprises any flat topology of $\mathcal{B}$ since all compact flat

manifolds are covered by tori according to a theorem by Bieberbach [102,103]. The torus has respective side lengths l such that its volume is l^3, and we denote its coordinates by x. To describe a field on the torus, we consider the discrete mode system $\{f_k(x) \coloneqq l^{-\frac{3}{2}} \exp(ik \cdot x)\}$ where the modes are labeled by vectors $k \in \Sigma \coloneqq 2\pi\mathbb{Z}^3/l$. They represent an orthonormal basis of $L^2([0,l]^3, \mathrm{d}x)$. With the standard scalar product $\langle f, g\rangle \coloneqq \int \mathrm{d}x f(x)\overline{g(x)}$ for L^2-functions f and g on the torus, we hence see that

$$\langle f_k, f_{k'}\rangle = \delta_{k,k'}, \qquad \sum_{k\in\Sigma} f_k\langle f_k, \cdot\rangle = 1, \tag{70}$$

where periodic boundary conditions are understood. Besides, the mode functions are eigenfunctions of the negative Laplace operator $(-\Delta)$ with eigenvalues $\mu_k^2 \coloneqq (2\pi^2)k^2/l^2$. We then define the homogeneous modes of the scalar field $\Phi(t,x)$ and its conjugate momentum $\Pi(t,x)$ by

$$\phi \coloneqq f_0\langle f_0, \Phi\rangle_{L^2(\mathbb{T}^3)} = \frac{1}{l^3}\int_{[0,l]^3} \mathrm{d}x\Phi(x),$$
$$\mu \coloneqq f_0\langle f_0, \Pi\rangle = \frac{1}{l^3}\int_{[0,l]^3} \mathrm{d}x\Pi(x).$$

The canonical brackets of the fields (Φ, Π) are given for two test functions f and g, on the torus by $\{\Phi(f), \Pi(g)\} = \langle f, g\rangle$. In an informal distributional notation, this is equivalent to $\{\Phi(x), \Pi(y)\} = \delta(x-y)$. As a result, the homogeneous modes are easily checked to have canonical brackets, $\{\phi, \mu\} = \frac{1}{l^3}$. Intuitively, we can say that the local point modes $\Phi(x)$ and $\Pi(x)$ are "infinitely faster" than the homogenous modes because formally speaking "$\delta(x-x) = \infty$." To understand this in more detail, we note that SAPT was developed for quantum mechanical systems with a finite number of degrees of freedom. Therefore, let us use a finite resolution cutoff $\delta_N \coloneqq \frac{l}{N}$ by dividing the torus into N^3 disjoint cubes of volume δ_N^3 each. Any cube is labeled by a vector $j \in \mathbb{N}_N^3 \coloneqq [0, N-1]^3$. We can define characteristic functions $\chi_{N,j}(x)$ for the cubes which evaluate to one in and on the cube with label j but vanish everywhere else. Accordingly, let us define the finite number of position localized degrees of freedom,

$$\varphi_{N,j} \coloneqq \frac{1}{\delta_N^3}\int_{[0,l]^3} \mathrm{d}x\chi_{N,j}(x)\Phi(x),$$
$$\pi_{N,j} \coloneqq \frac{1}{\delta_N^3}\int_{[0,l]^3} \mathrm{d}x\chi_{N,j}(x)\Pi(x). \tag{71}$$

We check that $\{\varphi_{N,j}, \pi_{N,j'}\} = \frac{1}{\delta_N^3}\delta_{j,j'}$ and recall that $\delta_N \le l$. Hence, the norm of the Poisson bracket is larger than for the homogeneous modes and tends to infinity for an infinitely fine resolution of the cubes. Thus, it meets the δ-distribution of the continuous field. Moreover, we can represent the homogeneous scalar field variables ϕ and μ using the regulated field degrees of freedom as

$$\phi = \frac{1}{N^3}\sum_{j\in\mathbb{N}_N^3}\varphi_{N,j}, \qquad \mu = \frac{1}{N^3}\sum_{j\in\mathbb{N}_N^3}\pi_{N,j}. \tag{72}$$

These relations suggest to interpret ϕ and μ as the center of mass coordinates of a system of N^3 coordinates $\varphi_{N,j}$ and $\pi_{N,j}$ of equal mass. Thus, at finite resolution, we have an abstract gas of interacting particles with "position" coordinates $\varphi_{N,j}$ and $\pi_{N,j}$, and it is well known from classical mechanics that the center of mass coordinate acquires the total mass of all particles as its effective mass, thus making it much heavier than the individual particles. To see this in more detail, we note that not all of the $\varphi_{N,j}$ are independent which is due to the above identity. For illustrative purposes, we assume that the Klein-Gordon field mass is small, and we consequently approximate the potential term in the Hamiltonian with the discretized Laplacian. Therefore, we use the unit standard basis vectors of $\mathbb{R}^3$ and denote them as $\{e_i\}_{i\in\{1,2,3\}}$. The discretized Laplacian then acts by shifting the cube fields by e_i according to

$$\int_{[0,l]^3}\mathrm{d}x\Phi(x)(-\Delta)\Phi(x) \to -\delta_N^3\sum_{j\in\mathbb{N}_N^3}\sum_{i=1}^{3}\varphi_{N,j}\frac{\varphi_{N,j+e_i}+\varphi_{N,j-e_i}-2\varphi_{N,j}}{\delta_N^2}$$
$$= \delta_N\sum_{j\in\mathbb{N}_N^3}\sum_{i=1}^{3}(\varphi_{N,j+e_i}-\varphi_{N,j})^2, \tag{73}$$

where the arrow stands for the discretization step. In the second line, the discretized Laplacian only depends on the relative coordinates. This motivates us to choose new variables, $\tilde{\varphi}_{N,j} \coloneqq \varphi_{N,j} - \varphi_{N,0}$, with $j \in \mathbb{N}_N^3\backslash\{0\}$, i.e., the discretized fields which are reset by the field $\varphi_{N,0}$ of the cube at the origin. To complete the set of new variables, we also have that the homogeneous mean field of these new variables is given by the homogeneous field in the former

variables, namely $\tilde{\phi} \coloneqq \phi$. To make the relation with the center of mass variables clearer, we define $M \coloneqq N^3$, and we note that $\varphi_{N,0} = \tilde{\phi} - \sum_{j\neq 0} \frac{\tilde{\varphi}_{N,j}}{M}$.

Let us then define the symplectic potential of the discretized theory. The phase space of the theory consists of the finite number of points $\Gamma \coloneqq \{(\varphi_{N,j}, \pi_{N,j})\}_{j\in\mathbb{N}_N^3}$. The symplectic potential is then a map from the tangent space $T_{(\varphi,\mu)}\Gamma$ at some point $(\varphi, \mu) \in \Gamma$ into $\mathbb{R}$. More precisely, let $\mathrm{d}\varphi_{N,j}$ denote the standard one-form on $T_{(\varphi,\pi)}\Gamma$ for the variable $\varphi_{N,j}$. The symplectic potential is then given by

$$\begin{aligned}\frac{1}{\delta_N^3}\Theta &= \sum_{j\in\mathbb{N}_N^3} \pi_{N,j}\mathrm{d}\varphi_{N,j} = \sum_{j\in\mathbb{N}_N^3} (\pi_{N,j}-\mu)\mathrm{d}\varphi_{N,j} + \sum_{j\in\mathbb{N}_N^3} \mu\mathrm{d}\varphi_{N,j}\\ &= \sum_{j\in\mathbb{N}_N^3\setminus\{0\}} (\pi_{N,j}-\mu)\mathrm{d}\varphi_{N,j} + (\pi_{N,0}-\mu)\mathrm{d}\varphi_{N,0} + \sum_{j\in\mathbb{N}_N^3} \mu\mathrm{d}\varphi_{N,j}\\ &= \sum_{j\in\mathbb{N}_N^3\setminus\{0\}} (\pi_{N,j}-\mu)\mathrm{d}(\varphi_{N,j}-\varphi_{N,0}) + \sum_{j\in\mathbb{N}_N^3\setminus\{0\}} (\pi_{N,j}-\mu)\mathrm{d}\varphi_{N,0} + (\pi_{N,0}-\mu)\mathrm{d}\varphi_{N,0} + \sum_{j\in\mathbb{N}_N^3} \mu\mathrm{d}\varphi_{N,j}\\ &= \sum_{j\in\mathbb{N}_N^3\setminus\{0\}} (\pi_{N,j}-\mu)\mathrm{d}\tilde{\varphi}_{N,j} + N^3\mu\mathrm{d}\tilde{\phi} =: \sum_{j\in\mathbb{N}_N^3\setminus\{0\}} \tilde{\pi}_{N,j}\mathrm{d}\tilde{\varphi}_{N,j} + (M\mu)\mathrm{d}\tilde{\phi},\end{aligned} \tag{74}$$

where we used in an intermediate step the following relations:

$$\sum_{j\in\mathbb{N}_N^3} \mu\mathrm{d}\varphi_{N,j} = N^3\mu\mathrm{d}\phi, \qquad \sum_{j\in\mathbb{N}_N^3\setminus\{0\}} \mu\mathrm{d}\varphi_{N,j} = (N^3-1)\mu\mathrm{d}\phi. \tag{75}$$

Thus, Eq. (74) suggests to define the conjugate momenta of the variables $\tilde{\varphi}_{N,j}$ according to $\tilde{\pi}_{N,j} \coloneqq \pi_{N,j} - \mu$ for every $j \in \mathbb{N}_N^3\setminus\{0\}$ and correspondingly the momentum conjugate to $\tilde{\phi}$ as $\tilde{\mu} \coloneqq M\mu$. Let us then consider the kinetic energy contribution to the Hamilton function of the scalar field. We start with twice the original field kinetic energy and perform a discretization to get

$$\begin{aligned}\int_{[0,l]^3} \mathrm{d}x\Pi(x)^2 \to \delta_N^3 \sum_{j\in\mathbb{N}_N^3} \pi_{N,j}^2 &= \delta_N^3 \left(\sum_{j\in\mathbb{N}_N^3} (\pi_{N,j}-\mu^2) + 2\mu\sum_{j\in\mathbb{N}_N^3} \pi_{N,j} - N^3\mu^2\right)\\ &= \delta_N^3 \left(\sum_{j\in\mathbb{N}_N^3\setminus 0} (\pi_{N,j}-\mu)^2 + (\pi_{N,0}-\mu)^2 + N^3\mu^2\right)\\ &= \delta_N^3 \left(\sum_{j\in\mathbb{N}_N^3\setminus 0} \tilde{\pi}_{N,j}^2 + \left(\sum_{j\in\mathbb{N}_N^3\setminus 0} \tilde{\pi}_{N,j}\right)^2 + \frac{\tilde{\mu}}{M}\right).\end{aligned} \tag{76}$$

A further analysis reveals that the quadratic form of the $\tilde{\pi}_{N,j}$ can be diagonalized by an orthogonal transformation, which can also be extended to a canonical one. With this and the reduction in (76), we directly see that the system consists of $N^3 - 2$ modes of unit mass and one mode of the reduced mass M. It is thus reasonable to think of the homogeneous mode as a heavy center of mass degree of freedom, especially in the limit $M = N^3 \to \infty$, which is what we wanted to show.

However, one may object that the cubic field degrees of freedom $\varphi_{N,j}$ have nothing to do with the Fourier momentum modes $\hat{\Phi}_k \coloneqq \{f_k, \Phi\}$, except for the zero mode $M\tilde{\phi}$. But, in fact, the discretized fields $\varphi_{N,j}$ are approximants of $\Phi(x = \frac{l}{N}j)$, which represents the field discretized on a lattice of $\mathbb{T}^3$ at the lattice point $\frac{l}{N}j$. Hence, we pass to a discretized lattice for the three-torus such that points on the lattice are given by $x \in \Omega \coloneqq [0, \frac{l}{N}, \ldots, \frac{(N-1)l}{N}]^3$. Indeed, we then write for the $\varphi_{N,j}$

$$\begin{aligned}\varphi_{N,j} &= \frac{1}{\delta_N^3}\int_{[0,l]^3} \mathrm{d}x\chi_{N,j}(x)\Phi(x) \to \frac{1}{\delta_N^3}\sum_{x\in\Omega} \delta_N^3\delta_{\frac{l}{N}j,x}\Phi(x)\\ &= \Phi\left(\frac{l}{N}j\right).\end{aligned} \tag{77}$$

In this situation, the Fourier transformation can also be restricted to the modes $\frac{2\pi}{l}j$ with $j \in \mathbb{N}_N^3$. To see that this is

true, we first write the Fourier coefficient $\hat{\Phi}_k \in \mathbb{C}$ for $\Phi(x) \in \mathbb{R}^{3N}$ with the Fourier transformation formula and without specifying $k \in \Sigma$,

$$\hat{\Phi}_k = \frac{1}{N^3} \sum_{x \in \Omega} \exp(-ik \cdot x) \Phi(x). \tag{78}$$

Then, we employ as an ansatz for the inverse discrete Fourier transformation $\Phi(x) = \sum_k \exp(ikx) \cdot \hat{\Phi}_k$ where we set the range of the sum to $k \in \Sigma' \coloneqq \frac{2\pi}{l}[0, \ldots, N-1]^3$. Then, we require that the application of the transform and its inverse give the unity operator, namely

$$\begin{aligned} \Phi(x) &= \frac{1}{N^3} \sum_{x \in \Omega} \Phi(x') \sum_{k \in \Sigma'} \exp\left(ik \cdot (x - x')\right) \\ &= \frac{1}{N^3} \sum_{x \in \Omega} \Phi(x') N^3 \delta_{x,x'}. \end{aligned} \tag{79}$$

Indeed, in order to obtain the correct normalizing factor, the range of k must be restricted to Σ'. One can also argue qualitatively: We need to relate the finite-dimensional vector $\Phi(x) \in \mathbb{R}^{3N}$ to the vector $\hat{\Phi}_k$ by some unitary matrix. As the matrix needs to be invertible, $\hat{\Phi}_k$ must have the same dimension as $\Phi(x) \in \mathbb{R}^{3N}$. Consequently, we have shown that the zero mode can be considered as a center of mass mode with respect to certain linear combinations of discretized position modes which in turn are linear combinations of discretized momentum modes. Accordingly, treating the homogeneous mode as the by far most massive one is physically justifiable from this point of view. That instead of an arbitrarily large relative scale $1/N^3 \to 0$ we just used the finite one $\varepsilon^2 = \frac{\kappa}{\lambda}$ is motivated by the specific combinations of the homogeneous momenta that appear in the Hamiltonian.

C. Transformations for well-defined quantum fields

In this section, we discuss the anomalies that occur in QFT on CST due to the infinite number of degrees of freedom. In particular, it is important to understand how the quantum fields for different configurations of the underlying curved spacetime relate one to another. As it turns out, the corresponding natural Fock representations fail to be unitarily equivalent for different configurations of the homogeneous variables. Before we explain the details of our solution, we illustrate the dilemma with the simplest possible QFT.

Therefore, consider a classical real scalar field Φ of Klein-Gordon type defined on a globally hyperbolic spacetime $\mathcal{M} \cong \mathbb{R} \times \mathcal{B}$ for which the spatial hypersurface $\mathcal{B}$ is a flat and compact manifold. Spacetime then foliates into spatial hypersurfaces, each of which has the topology $\mathcal{B}$. Then, there exists a diffeomorphism which maps any point X to its coordinate representative (t, x). The variable x denotes the local coordinates on $\mathcal{B}$. We split the field into its purely homogeneous and isotropic part with respect to this hypersurface foliation $\phi(t, x) = \phi(t)$, and define the difference $\varphi(t, x) \coloneqq \Phi(t, x) - \phi(t)$ as a linear perturbation of ϕ. Furthermore, the model comprises the homogeneous and isotropic parts of the metric field $g_{\mu\nu}$ of general relativity, as well as a timelike, homogeneous, and isotropic, real scalar dust field u with energy density ρ. The homogeneous and isotropic metric reduces to the time-dependent scale factor a and the lapse function N. Physical dynamics is encoded in the velocity of the scale factor $\dot{a}$. The action splits into a homogeneous and an inhomogeneous, perturbative part $S = S_{\text{hom}} + S_{\text{pert}}$ with,

$$\begin{aligned} S_{\text{hom}}[a, N, u, \phi] &= \int_{\mathbb{R}} \mathrm{d}t N \left(-\frac{a\dot{a}^2}{\kappa N^2} + \frac{a^3 \rho}{2}(\dot{u}^2 - 1) \right. \\ &\quad \left. + \frac{a^3}{2\lambda} \left(\left(\frac{\dot{\phi}}{N} \right)^2 - m^2 \phi^2 \right) \right), \\ S_{\text{pert}}[a, N, \varphi] &= \frac{1}{2\lambda} \int_{\mathbb{R} \times \mathcal{B}} \mathrm{d}t \mathrm{d}x N a^3 \left(\dot{\varphi}^2 - \frac{(\nabla \varphi)^2}{a^2} - m^2 \varphi^2 \right), \end{aligned} \tag{80}$$

where $\kappa = 8\pi G$ is the gravitational coupling constant, λ is the coupling constant of the matter field, m is its mass, and we have introduced the measure $\mathrm{d}x$ of the spatial hypersurfaces. The dust field only serves to deparametrize the model, i.e., to obtain a true Hamilton operator instead of a Hamilton constraint. Namely, after fixing the gauge freedom associated with the spacetime diffeomorphisms, the gravitational and scalar contributions to the Hamilton constraint combine to build a physical Hamiltonian when integrated over the spatial hypersurface [104,105]. The gravitational and the scalar field degrees of freedom become observable fields.

Note that this model almost coincides with the model in the third paper [82]. There, we employ the same dust field and reduce the geometry to its homogeneous and isotropic part. However, we do not account separately for the homogeneous and isotropic scalar field mode ϕ but we include it in the inhomogeneous part of the field Φ. This is possible as all the gravitational and matter degrees of freedom are already Dirac observables thanks to the dust field. It is therefore not necessary to treat the ϕ-variable separately which would serve to define gauge transformations in the purely constrained case as for the fourth model [69]. Here, however, we keep this separation of the homogeneous scalar field mode in order to show that the following transformations are valid for systems with more than one pair of homogeneous variables. With this in mind, a standard Legendre transformation from which the conjugate momenta (p_a, μ, π) arise yields the Hamilton function $H = H_{\text{hom}} + H_{\text{pert}}$ with

$$H_{\mathrm{pert}}(a,\varphi,\pi)=\frac{1}{2\lambda}\int_{\mathcal{B}}\mathrm{d}x\left(\frac{\lambda^2\pi^2}{a^3}-a\varphi(\Delta\varphi)+a^3m^2\varphi^2\right). \quad (81)$$

Note that we set $N\equiv 1$ as the constraint analysis of general relativity establishes that the lapse function is a mere Lagrange multiplier without any relevance for the dynamics of the theory. The homogeneous contribution of the Hamiltonian $H_{\mathrm{hom}}(a,p_a,\phi,\mu)$ has the form of the well-known homogeneous and isotropic cosmological Hamilton constraints but its explicit form is not of importance for this section.

Given these prerequisites, SAPT requires one to work on the Hilbert space $\mathcal{H}=\mathcal{H}_{\mathrm{hom}}\otimes\mathcal{H}_{\mathrm{pert}}$, where the first factor refers to the homogeneous sector and the second one to the perturbative sector in our case. As far as the homogeneous sector is concerned, we adopt a usual Schrödinger representation in accordance with the space adiabatic formalism. Regarding the inhomogeneous part, a standard Fock representation suggests itself since $H_{\mathrm{pert}}(a,\varphi,\pi)$ is quadratic in the fields φ and π.

But which one? After all, the "background" variable "a" is not a fixed function of time but a dynamical quantum variable, and hence displays quantum fluctuations. SAPT allows one to technically consider the scale factor a as a real parameter at the first stage, namely when quantizing $H_{\mathrm{pert}}(a,\varphi,\pi)$ with respect to the inhomogeneous variables (φ,π). In this case, the system will correspond to a standard quantum Klein-Gordon field with an a-dependent frequency, and we establish the necessary aspects of the theory in the following. Therefore, consider the one-particle Hilbert space $L^2(\mathcal{B},\mathrm{d}x)$ on the compact manifold $\mathcal{B}$. It appears natural to consider the corresponding symmetric Fock space $\mathcal{H}_{\mathrm{pert}}=\mathcal{F}_{\mathrm{s}}(L^2(\mathcal{B}))$ as the state space of the inhomogeneous quantum field theory for some fixed $a\in\mathbb{R}^+$. We denote the corresponding vacuum state as $\Omega(a)$. Since $\mathcal{B}$ is a compact manifold, the spectrum of the negative Laplace operator $-\Delta$ has a discrete spectrum on the space of two-times continuously differentiable functions $C^2(\mathcal{B},\mathbb{C})$. Consequently, there exists an orthonormal eigenbasis $\{f_k(x)\}$ for $-\Delta$ with eigenvalues $\{\mu_k^2\}$. We denote the discrete set for the labels k by Σ. The discrete eigenbasis provides an orthonormal basis of $L^2(\mathcal{B},\mathrm{d}x)$. For illustrative purposes, consider the compact, flat three-torus $\mathbb{T}^3$ with respective side lengths l and coordinates x. The set of l-periodic functions $\{f_k(x):=l^{-\frac{3}{2}}\exp(ik\cdot x)\}$ with $k\in\Sigma:=2\pi\mathbb{Z}^3/l$ represents an orthonormal basis of the Hilbert space $L^2([0,l]^3,\mathrm{d}x)$. In the generic notation, the Fock space on $\mathcal{B}$ consists of sequences $\{\psi_{(n)}\}_{n\geq 0}$ of totally symmetric functions with n variables $x_i\in\mathcal{B}$ [106] [p. 10].

The algebra of operators is defined by means of the particle annihilation and creation operators $\boldsymbol{b}(a)$ and $\boldsymbol{b}^*(a)$, and we choose to denote the operators acting on states in $\mathcal{F}_{\mathrm{s}}$ by bold letters as the inhomogeneous subsector identifies with the fast system. The operators $\boldsymbol{b}(a,f)$ and $\boldsymbol{b}^*(a,f)$ "annihilate" and "create" any smooth one-particle state $f(x)\in C^\infty(\mathcal{B})$ in the standard way. The maps $f\mapsto\boldsymbol{b}(a,f)$ and $f\mapsto\boldsymbol{b}^*(a,f)$ are antilinear and linear, respectively, with respect to f. Therefore, it is possible to define the operators more explicitly by means of operator-valued distributions $\boldsymbol{b}(a,x)$ and $\boldsymbol{b}^*(a,x)$ such that $\boldsymbol{b}(a,f)=\int\mathrm{d}x\overline{f(x)}\boldsymbol{b}(a,x)$ and $\boldsymbol{b}^*(a,f)=\int\mathrm{d}x f(x)\boldsymbol{b}^*(a,x)$. The annihilation and creation operators stand in direct relation to the canonical field operators $\boldsymbol{\varphi}$ and $\boldsymbol{\pi}$. In this respect, Eq. (81) suggests to define the one-particle operator $\sqrt{\omega(a)}:C^2(\mathcal{B})\to C(\mathcal{B})$ according to $\omega(a)^2:=-a^2\Delta+m^2a^6$. By duality, this one-particle operator acts on the operator-valued $C^\infty(\mathcal{B})$ distributions. Namely, for some operator-valued distribution $\boldsymbol{\rho}(x)\in C^\infty(\mathcal{B},\mathcal{L}(\mathcal{F}_{\mathrm{s}}))^*$ where $\mathcal{L}(\mathcal{F}_{\mathrm{s}})$ denotes the space of linear operators on the bosonic Fock space and the "$*$" stands for the dual space $\sqrt{\omega(a)}$ operates according to $(\sqrt{\omega(a)}(\boldsymbol{\rho}))(f):=\boldsymbol{\rho}(\sqrt{\omega(a)}(f))$. For the given case, the operator-valued distribution of the annihilation operator then has the explicit form

$$\boldsymbol{b}(a,f):=\frac{1}{\sqrt{2}}[(\sqrt{\omega(a)}(\boldsymbol{\varphi}))-i(\sqrt{\omega(a)}^{-1}(\boldsymbol{\pi}))](f), \quad (82)$$

and accordingly for the creation operator with the minus replaced by a plus. In view of the already mentioned orthonormal basis $\{f_k\}$ of $L^2(\mathcal{B})$ which diagonalizes the Laplace operator, it is convenient to introduce the annihilation and creation operators for the modes f_k, namely $\boldsymbol{b}(a,f_k)$ and $\boldsymbol{b}^*(a,f_k)$. They satisfy the standard commutation relations of annihilation and creation operators,

$$[\boldsymbol{b}(a,f_k),\boldsymbol{b}^*(a,f_{k'})]=\delta_{k,k'}\mathbf{1}_{\mathcal{F}_{\mathrm{s}}}, \quad (83)$$

where δ is the Kronecker delta with respect to the discrete modes k and k'. In order to represent the Hamilton operator $\boldsymbol{H}_{\mathrm{pert}}$ in terms of the mode annihilation and creation operators, we first return to the classical fields $\varphi(x),\pi(x)$ and expand them with respect to the mode basis $\{f_k\}$. With the corresponding Fourier coefficients as position and momentum variables φ_k and π_k, we obtain the Hamiltonian of a discrete but infinite set of independent harmonic oscillators with respective frequency $\omega(a,\mu_k^2):=\sqrt{a^4\mu_k^2+ma^6}$. The quantization of the mode coefficients yields the quantum version of the classical Hamiltonian in (81). Eventually, we perform a transformation to annihilation and creation operators according to Eq. (82) with $f=f_k$ and the analogous relation for $\boldsymbol{b}^*(a,f_k)$. Then, the a-dependent normal ordering of the Fock–quantized perturbation Hamiltonian $H_{\mathrm{pert}}(a,\phi,\pi)$ reads

$$\boldsymbol{H}_{\mathrm{pert}}(a) = \frac{1}{a^3}\sum_{k\in\Sigma}\boldsymbol{b}^*(a,f_k)(\omega(a)\boldsymbol{b}(a))(f_k). \quad (84)$$

Two immediate questions arise:

(1) Are the corresponding Fock representations $(\pi_a, \mathcal{F}_{\mathrm{s}})$, where π_a denotes the group homomorphism between the field Weyl algebra of $(\boldsymbol{\varphi}, \boldsymbol{\pi})$ and the space of bounded linear operators on $\mathcal{F}_{\mathrm{s}}$, all unitarily equivalent to a single representation $(\pi_{\mathrm{f}}, \mathcal{F}_{\mathrm{s}})$? This is one of the innocent looking assumptions of SAPT which is automatically satisfied in the quantum mechanical context.

(2) Assuming that this unitary equivalence between the different representations is granted, let f be a smooth function in the one-particle Hilbert space associated with the Fock space $\mathcal{F}_{\mathrm{s}}$. Then, let $\hat{\boldsymbol{b}}(a,f) = \hat{\mathcal{W}}(\boldsymbol{b}(a,f))$ and $\hat{\boldsymbol{b}}^*(a,f) = \hat{\mathcal{W}}(\boldsymbol{b}^*(a,f))$ be the Weyl quantizations with respect to the homogeneous variables of the annihilation and creation operator symbols $\boldsymbol{b}(a,f)$ and $\boldsymbol{b}^*(a,f)$. Then the second question is whether the complete algebra of the operators $\hat{\boldsymbol{a}}, \hat{\boldsymbol{p}}_a, \hat{\boldsymbol{\phi}}, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{b}}(a,f), \hat{\boldsymbol{b}}^*(a,f)$ is well defined on the full Hilbert space $\mathcal{H}_{\mathrm{hom}} \otimes \mathcal{F}_{\mathrm{s}}$.

It turns out that both questions are tightly related and the answer to both is negative. The underlying effect has been first observed in [64–66] in a related context.

To see the origin of the problem, we note that a necessary condition for an affirmative answer to the first question is that the Fock vacuum $\Omega(a_2)$ can be written as an excited state in $\mathcal{F}_{\mathrm{s}}$ with respect to the representation π_{a_1} for all a_1, a_2 in $\mathbb{R}^+$. In fact, this condition is also sufficient because polynomials of the creation operators $\boldsymbol{b}^*(a_1,f)$ can be written as polynomials of the operators $\boldsymbol{b}(a_2,f)$ and $\boldsymbol{b}^*(a_2,f)$ for some smooth $f \in C^\infty(\mathcal{B})$. To elaborate on this, let (n) denote the collection of occupation numbers $\{n_k\}_{k\in\Sigma}$ of some excited state $\psi_{(n)}$. For the point a_1, we define this excited state as,

$$\psi_{(n)}(a_1) = \prod_{k\in\Sigma}\frac{[\boldsymbol{b}^*(a_1,f_k)]^{n_k}}{\sqrt{n_k!}}\Omega(a_1). \quad (85)$$

We then assume that it is possible to write the vacuum state at a_2 as a linear combination of such excited states, namely

$$\Omega(a_2) = \sum_{(n)} z_{(n)}\cdot\psi_{(n)}(a_1), \quad (86)$$

where the sum is over all (n) with only finitely many n_k different from zero. Then, for the two vacuums to stand in a well-defined relation, it must hold true that $\boldsymbol{b}(a_2,f_k)\Omega(a_2) = 0$ for any f_k. To analyze this equation with respect to the π_{a_1} representation, consider the differential operators $u_\pm(a_1,a_2): C^\infty(\mathcal{B}) \to C^\infty(\mathcal{B})$ defined by

$$u_\pm(a_1,a_2) \coloneqq \frac{1}{2}\left(\sqrt{\frac{\omega(a_1)}{\omega(a_2)}} \pm \sqrt{\frac{\omega(a_2)}{\omega(a_1)}}\right), \quad (87)$$

and we recall that $\omega(a)^2 = -a^4\Delta + m^2a^6$. By extending $u_\pm(a_1,a_2)$ to the space of operator-valued distributions as before, we define the Bogoliubov transformation by

$$\begin{aligned}\boldsymbol{b}(a_2,f_k) \coloneqq{}& (u_+(a_1,a_2)\boldsymbol{b}(a_1))(f_k)\\ &+ (u_-(a_1,a_2)\boldsymbol{b}^*(a_1))(f_k),\end{aligned} \quad (88)$$

for any f_k and linear combinations thereof. The one-particle operator $u_\pm(a_1,a_2)$ is diagonal with respect to the basis states $\{f_k\}_{k\in\Sigma}$. Therefore, let $u_\pm(a_1,a_2,k) \in \mathbb{R}$ be the eigenvalues of $u_\pm(a_1,a_2)$ defined by $u_\pm(a_1,a_2)f_k = u_\pm(a_1,a_2,\mu_k^2)f_k$. It follows by extending this relation to the operator-valued distributions that

$$(u_+(a_1,a_2)\boldsymbol{b}(a_1))(f_k) = u_+(a_1,a_2,\mu_k^2)\boldsymbol{b}(a_1,f_k), \quad (89)$$

and correspondingly for u_- and $\boldsymbol{b}^*(f_k)$. Eventually, we denote the collection of occupation numbers for which only the excitation number associated with f_k equals one and the remaining ones vanish by 1_k. Then, the expression $(n)\pm 1_k$ stands for a new set of occupation numbers in which the excitation number of f_k in the set (n) is raised or lowered, respectively, by one. Then, we require that the defining equation of the vacuum state holds, namely that the application of the annihilation operator $\boldsymbol{b}(a_2,f_k)$ on the vacuum state $\Omega(a_2)$ vanishes,

$$\begin{aligned}\boldsymbol{b}(a_2,f_k)\Omega(a_2) &= \sum_{(n)} z_{(n)}[\sqrt{n_k}u_+(a_1,a_2,\mu_k^2)\psi_{(n)-1_k}(a_1)\\ &\quad + \sqrt{n_k+1}u_-(a_1,a_2,\mu_k^2)\psi_{(n)+1_k}(a_1)]\\ &= \sum_{(n)}[z_{(n)+1_k}u_+(a_1,a_2,\mu_k^2)\sqrt{n_k+1}\\ &\quad + z_{(n)-1_k}u_-(a_1,a_2,\mu_k^2)\sqrt{n_k}]\psi_{(n)}(a_1) = 0.\end{aligned} \quad (90)$$

Since the relation (90) should hold for all $k \in \Sigma$ independently, the coefficients must be of infinite product type

$$z_{(n)} = \prod_k \zeta^k_{n_k}. \quad (91)$$

Then, we define the quotient $\sigma(a_1,a_2,\mu_k^2) \coloneqq u_-(a_1,a_2,\mu_k^2)/u_+(a_1,a_2,\mu_k^2)$. The product ansatz together with this definition transforms equation (90) into the recursion relation

$$\zeta^k_{n_k+1} = -\sqrt{\frac{n_k}{n_k+1}}\sigma(a_1,a_2,\mu_k^2)\zeta^k_{n_k-1}. \quad (92)$$

The right-hand side of the recursion relation vanishes for $n_k = 0$. It follows that $\zeta^{n_k}_{(n)} = 0$ for any odd n_k. For even n_k, the solution of Eq. (92) is given by

$$\begin{aligned}\zeta^k_{2n_k} &= -\sqrt{\frac{2n_k-1}{2n_k}}\sigma(a_1,a_2,\mu_k^2)\zeta^k_{2(n_{k-1})}\\ &= (-\sigma(a_1,a_2,\mu_k^2))^{n_k}\sqrt{\frac{(2n_k)!}{4^{n_k}(n_k!)^2}}\zeta^k_0,\end{aligned} \tag{93}$$

where the ζ^k_0 remain undetermined for the moment and their values are of no consequence for the subsequent argumentation. In order to prove that the vacuum $\Omega(a_1)$ transforms into the vacuum $\Omega(a_2)$ in a well-defined fashion, it is necessary that the $\mathcal{F}_s(a_1)$ norm of $\Omega(a_2)$ has a finite, nonvanishing value. As a working hypothesis, we assume that this is true. Then, it stems from the previous results that

$$\begin{aligned}\|\Omega(a_2)\|^2_{\mathcal{F}_s} &= \sum_{(n)}|z_{(n)}|^2 = \sum_{(n)}\left|\prod_{k\in\Sigma}\zeta^k_{n_k}\right|^2 = \sum_{(n)}\prod_{k\in\Sigma}|\zeta^k_{n_k}|^2\\ &= \prod_{k\in\Sigma}\left(\sum_{n=0}^{\infty}|\zeta^k_0|^2\sigma(a_1,a_2,\mu_k^2)^{2n}\frac{(2n)!}{4^n(n!)^2}\right)\\ &= \left(\prod_{k\in\Sigma}|\zeta^k_0|^2\right)\cdot\left(\prod_{k\in\Sigma}\sum_{n=0}^{\infty}\sigma(a_1,a_2,\mu_k^2)^{2n}\frac{(2n)!}{4^n(n!)^2}\right).\end{aligned} \tag{94}$$

Regarding the n-dependent factors in the last expression, it will prove to hold that

$$\frac{1}{2^n} \le \frac{(2n)!}{4^n(n!)^2} \le 1, \tag{95}$$

for example by using mathematical induction. Inserting the estimate (95) into (94) allows one to rewrite the sum as a geometric series. We obtain an infinite series over n with coefficients $(\sigma(a_1,a_2,\mu_k^2)^{2n})/2^n$ and $\sigma(a_1,a_2,\mu_k^2)^{2n}$, respectively. Since $\sigma(a_1,a_2,\mu_k^2)^2 < 1$ independently of μ_k, this gives rise to the following estimate:

$$\begin{aligned}\prod_{k\in\Sigma}|\zeta^k_0|^2\cdot\prod_{k\in\Sigma}\left(1-\frac{\sigma(\mu_k^2)^2}{2}\right)^{-1} &\le \|\Omega(a_2)\|^2_{\mathcal{F}_s}\\ \le \prod_{k\in\Sigma}|\zeta^k_0|^2\cdot\prod_{k\in\Sigma}(1-\sigma(\mu_k^2)^2)^{-1}.\end{aligned} \tag{96}$$

Thus, a necessary condition for the convergence of (94) is that the infinite product on the left converges to a finite value. Note thereby that $\prod_{k\in\Sigma}|\zeta^k_0|^2$ is a common prefactor in all $|z_{(n)}|^2$, and thus it must be convergent to some finite value by itself as otherwise all $z_{(n)}$ would be meaningless. By taking the logarithm, the convergence of the lower estimate is equivalent to the convergence of the series

$$\sum_{k\in\Sigma}\ln\left(1-\frac{\sigma(a_1,a_2,\mu_k^2)^2}{2}\right), \tag{97}$$

which is also known as the *Hilbert-Schmidt* condition. In order to resolve the estimates, we recall that $\sigma(a_1,a_2,\mu_k^2)$ is determined by the frequency function $\omega(a,\mu_k^2) = \sqrt{a^4\mu_k^2 + a^6m^2}$, namely

$$\begin{aligned}\sigma(a_1,a_2,\mu_k^2)^2 &= \frac{u_-(a_1,a_2,\mu_k^2)^2}{u_+(a_1,a_2,\mu_k^2)^2}\\ &= \frac{((a_1^4-a_2^4)\mu_k^2+(a_1^6-a_2^6)m^2)^2}{(\omega(a_1,\mu_k^2)+\omega(a_2,\mu_k^2))^4}.\end{aligned} \tag{98}$$

In order that the series (97) converges, a necessary condition is that the coefficients of the latter tend to zero for $\mu_k \to \infty$. Keeping track of the logarithm, this is true if $\sigma(a_1,a_2,\mu_k^2)^2$ tends to zero. To check this, note that $\omega(a,\mu_k^2) \to a^2\mu_k$ for large μ_k. Consequently, the expression (97) evaluates to

$$\lim_{\mu_k\to\infty}\sigma(a_1,a_2,\mu_k^2)^2 = \frac{(a_1^2-a_2^2)^2}{(a_1^2+a_2^2)^2}, \tag{99}$$

which is a nonvanishing constant for $a_1 \neq a_2$. The Hilbert-Schmidt condition fails for any two distinct a_1, a_2. Note, however, that according to (98), the problem would disappear if the wave mode term was relieved from its scale factor dependence. In particular, then $\sigma(a_1,a_2,\mu_k^2)^2$ would decay as $1/\mu_k^4$, and thus the series $\sum_k \ln(1-\sigma(a_1,a_2,\mu_k^2)^2/2)$ would converge to a nonzero value.

This answers the first question posed above, namely that the Fock representations $(\pi_a,\mathcal{F}_s)$ are *not* unitarily equivalent for different values of a. The second question, i.e., whether the complete operator algebra of the canonical pairs is well defined on the full Hilbert space, is indeed equivalent to the first one. To see this, we assume that the scale factor a is represented as a self-adjoint operator $\hat{a}$ on the full Hilbert space $\mathcal{H}$. Then, by the spectral theorem, it is approved to display the Hilbert space as a direct integral or a Hilbert bundle subordinate to a, namely

$$\mathcal{H} \cong \int_{\sigma(\hat{a})}^{\oplus} \mathrm{d}\mu(a)\mathcal{H}_{\mathrm{pert}}(a), \tag{100}$$

where μ is the spectral probability measure on the spectrum $\sigma(\hat{a})$ of $\hat{a}$ which is just $\mathbb{R}_0^+$. The previous discussion suggests to identify the fiber spaces $\mathcal{H}_{\mathrm{pert}}(a)$ with the Fock spaces $\mathcal{F}_s(a)$. As a consequence of the spectral theorem, the Hilbert spaces $\mathcal{F}_s(a)$ (of equal dimension) should be chosen identical [107,108]. As we already know, this is, however, *not* the case. Nevertheless, it is instructive to pretend the opposite for the moment. Vectors in the Hilbert bundle are given by measurable fiber Hilbert space

valued functions, $\psi : \sigma(\hat{a}) \mapsto \mathcal{H}$, $a \mapsto \psi(a)$ over the base manifold $\sigma(\hat{a})$. They are equipped with the inner product

$$\langle \psi, \psi' \rangle_{\mathcal{H}} = \int_{\sigma(\hat{a})} \mathrm{d}\mu(a) \langle \psi(a), \psi'(a) \rangle_{\mathcal{F}_s}. \tag{101}$$

By the spectral theorem, $\hat{\boldsymbol{a}}$ acts by multiplication by a in the fiber $\mathcal{F}_s(a)$. Accordingly, the operator $\hat{\boldsymbol{H}}_{\mathrm{pert}}$ acts fiberwise as well by the symbol $\boldsymbol{H}_{\mathrm{pert}}(a)$ in (84). The question is now how the operator $\hat{\boldsymbol{p}}_a$ representing the momentum of the scale factor acts on the direct integral Hilbert space. As the spectrum of $\hat{\boldsymbol{a}}$ is of absolutely continuous type, the momentum acts as $(\hat{\boldsymbol{p}}_a \psi)(a) = (i\partial_a + f(a))\psi(a)$ where we introduce the function $f(a)$ related to the divergence of the measure $\mu(a)$. The function has the purpose to turn $\hat{\boldsymbol{p}}_a$ into a symmetric operator. In fact, in order to obtain a self-adjoint operator, it is advisable to pass to the real-valued triad variable e and work with its conjugate momentum. Nevertheless, the conclusion derived below is not affected by these subtleties. In contrast to the momentum operator, the operators $\hat{\boldsymbol{b}}(a, f)$ and $\hat{\boldsymbol{b}}^*(a, f)$ act fiberwise by the corresponding symbols as they only depend on the scale factor a.

In order to check whether the geometric momentum operator $\hat{\boldsymbol{p}}_a$ is well defined on (some dense subset of) $\mathcal{H}$, it is beneficial to compute the norm of the state $(\hat{\boldsymbol{p}}_a \Omega)(a)$. Therefore, we compute the geometric commutator of $\hat{\boldsymbol{p}}_a$ and $\hat{\boldsymbol{b}}(a, f)$. The integral representation of these Weyl operators with respect to the geometric system provides the means to easily deduce the result. As the operator $\boldsymbol{b}(a, f)$ only depends on the configuration variable a, a partial integration allows one to shift the derivative due to $\hat{\boldsymbol{p}}_a$ on $\boldsymbol{b}(a, f)$. Using Eq. (82), we see that the derivative directly acts on powers of the one-particle operator $\omega(a)$ within $\boldsymbol{b}(a, f)$. Finally, this yields an additional a-dependent prefactor $K(a)$ which is a one-particle operator itself. This operator is then also subject to the geometric quantization procedure

$$\begin{aligned} &[\hat{\boldsymbol{p}}_a, \hat{\boldsymbol{b}}(a, f)] = -i(\hat{K}(a)\hat{\boldsymbol{b}}^*(a))(f) \\ &\text{where } K(a) = \frac{1}{2}\frac{\partial_a \omega(a)}{\omega(a)}. \end{aligned} \tag{102}$$

As a result and with the identity $\hat{\boldsymbol{b}}(a, f)\Omega = 0$ for any f, it holds true that

$$\hat{\boldsymbol{b}}(a,f)(\hat{\boldsymbol{p}}_a\Omega) = [\hat{\boldsymbol{b}}(a,f), \hat{\boldsymbol{p}}_a]\Omega = i(\hat{K}(a)\hat{\boldsymbol{b}}^*(a))(f)\Omega. \tag{103}$$

Using that any one-particle state f has a unique decomposition with respect to some basis $\{f_k\}_{k\in\Sigma}$ and that the above equality must hold for *any* f, the vector $\hat{\boldsymbol{p}}_a\Omega$ has respective to this basis the solution

$$\hat{\boldsymbol{p}}_a\Omega = -\frac{i}{2}\sum_{k\in\Sigma} \hat{\boldsymbol{b}}^*(a, f_k)(\hat{K}(a)\hat{\boldsymbol{b}}^*(a))(f_k)\Omega. \tag{104}$$

By that, it is possible to compute the norm of $\hat{\boldsymbol{p}}_a\Omega$. Therefore, we denote the twofold excitation of the vacuum state $\Omega(a)$ with respect to the mode f_k by $\sqrt{2}\psi_{2_k}(a)$. Then, using the inner product with respect to the geometric quantization, we obtain for the norm square

$$\begin{aligned} \|\hat{\boldsymbol{p}}_a\Omega\|^2 &= \int_{\sigma(\hat{a})} \mathrm{d}\mu(a) \sum_{k\in\Sigma} \langle \psi_{2_k}(a) K(a,\mu_k), K(a,\mu_k)\psi_{2_k}(a)\rangle \\ &= \frac{1}{16}\int_{\sigma(\hat{a})} \mathrm{d}\mu(a) \sum_{k\in\Sigma} \left(\frac{\partial_a \omega(a,\mu_k)^2}{\omega(a,\mu_k)^2}\right)^2. \end{aligned} \tag{105}$$

The norm only admits a finite value if the series over k is almost everywhere finite as a function of a. However, for large k and consequently for large μ_k, any coefficient of the series evaluates to $(4/a)^2$, which is a constant for some fixed value of the scale factor a. Hence, the sum does not converge, and we are left with an infinite norm of $\hat{\boldsymbol{p}}_a\Omega$. Note that (105) is the infinitesimal version of (97), if we divide it by $(a_1 - a_2)^2$ and take the limit $a_2 \to a_1$.

Consequently, SAPT does not immediately apply to the case of QFT on CST. One hope might be that by a different choice of Fock representations, the Hilbert-Schmidt condition can be met. However, if the correspondingly normal-ordered Hamiltonian should remain at least densely defined on the Fock states, this again leads to an obstruction. To understand the origin of this impediment, note that we can satisfy the Hilbert-Schmidt condition by rescaling the classical fields $\varphi(x)$ and $\pi(x)$ according to

$$\tilde{\varphi}(x) \coloneqq a\varphi(x), \qquad \tilde{\pi}(x) = \frac{\pi(x)}{a}; \tag{106}$$

see also Ref. [109] where this strategy has been used previously. These new fields still satisfy the canonical Poisson brackets with respect to the inhomogeneous degrees of freedom and with a being fixed. The perturbation Hamilton function reads with these rescaled fields

$$\begin{aligned} H_{\mathrm{pert}} &= \frac{1}{2a}\int_{\mathcal{B}} \mathrm{d}x(\tilde{\pi}^2 + \tilde{\varphi}\,\tilde{\omega}(a)\tilde{\varphi}), \\ \text{where } \tilde{\omega}(a) &= -\Delta + m^2a^2, \end{aligned} \tag{107}$$

such that the coefficient of the Laplace operator in $\tilde{\omega}(a)^2$ is independent of a. However, the transformation (106) is not a canonical transformation on the *full* phase space. In fact, it is no longer the case that $p_a, \tilde{\varphi}$, and $\tilde{\pi}$ have vanishing Poisson brackets as the above transformation depends on a. Consequently, the fundamental canonical structure for the canonical quantization procedure of SAPT has disappeared. Nevertheless, as first pointed out in the context of hybrid quantum cosmology (see, e.g., Refs. [64–66] and

[110] and references therein), the transformation (106) allows for an *exact* completion by adding a corresponding contact term in the symplectic potential. For completeness, we review this procedure below using our terminology.

To define the symplectic potential for the system, we recall that our system is modeled on an infinite dimensional Banach manifold Q which is given by the Cartesian product of the finite dimensional, the homogeneous configuration space $Q_{\rm hom} = \mathbb{R}^+ \times \mathbb{R}$, and the infinite dimensional, perturbative configuration space of fields $Q_{\rm pert}$. The latter space could for example be the Sobolev space $H^1(\mathcal{B})$ of fields whose first derivative has a finite L^2 norm such that the Hamilton function is well defined. With the cotangent bundle T^*Q (i.e., the phase space) and the projection map $\mathrm{pr}: T^*Q \to Q : \alpha := (p, q) \to q$, which maps any $\alpha \in T^*Q$ to its configuration q, we define the symplectic potential Θ on the manifold T^*Q as a map from the tangent space $T_\alpha(T^*Q)$ into $\mathbb{R}$, where $\alpha \in T^*_qQ$ with $q = \mathrm{pr}(\alpha)$. In particular, for some $v \in T_\alpha(T^*Q)$, we have that $\Theta(\alpha)(v) := \alpha(\mathrm{pr}^*(v))$ where $\mathrm{pr}^*: T(T^*Q) \to TQ$ is the pushforward of the projection pr [111].

In a coordinate representation, this has the standard intuitive form: We denote the coordinates of a point $(q, p) \in T^*_\Phi Q$ in phase space by (Φ, Π), where Φ stands for the generalized fields (a, ϕ, φ). Then, Θ has the coordinate representation, $\int_{\mathcal{B}} \mathrm{d}x \Pi(x) \mathrm{d}\Phi(x)$, where d is the exterior derivative such that $\mathrm{d}\Phi(x)$ denotes a standard one-form on $T_{(\Phi,\Pi)}(T^*Q)$ [112]. To shorten the notation, we simply denote the integral by $\langle \Pi, \mathrm{d}\Phi \rangle$, where $\langle \cdot, \cdot \rangle$ stands for the L^2 norm on $\mathcal{B}$. By splitting the fields into the homogeneous and inhomogeneous components and by executing the integrals over the homogeneous degrees of freedom, the symplectic one-form supports the form

$$\Theta = p\mathrm{d}a + \mu \mathrm{d}\phi + \langle \pi, \mathrm{d}\varphi \rangle, \tag{108}$$

where we used the product rule for the exterior derivative. Then, the transformation to the dashed fields of this one-form produces an additional term proportional to da, namely

$$\Theta = \left(p - \frac{1}{a} \langle \pi, \varphi \rangle \right) \mathrm{d}a + \mu \mathrm{d}\phi + \langle \tilde{\pi}, \mathrm{d}\tilde{\varphi} \rangle. \tag{109}$$

This suggests to define new dashed pairs of variables in the homogeneous sector,

$$\begin{aligned} \tilde{p} &:= p - \frac{1}{a} \langle \pi, \varphi \rangle = p - \frac{1}{a} \langle \tilde{\pi}, \tilde{\varphi} \rangle, \\ \tilde{a} &:= a, \qquad \tilde{\phi} := \phi, \qquad \tilde{\mu} := \mu, \end{aligned} \tag{110}$$

as the completion of that transformation. Unfortunately, now we have to write $H_{\rm hom}$ in terms of $\tilde{p}$ and this causes problems when passing to the quantum realm. In particular, when quantizing with respect to the perturbative Fock fields, as suggested by SAPT, the supplementary term due to the completion introduces first and second powers of an ill-defined (normal ordered) operator. By taking an arbitrary basis $\{f_k\}_{k\in\Sigma}$ of $L^2(\mathcal{B})$, this operator is given by

$$\begin{aligned} \sum_{k\in\Sigma} \tilde{\boldsymbol{\pi}}(f_k) \tilde{\boldsymbol{\varphi}}(f_k) = \frac{i}{2} \sum_{k\in\Sigma} &(\tilde{\boldsymbol{b}}(a, f_k)^2 - \tilde{\boldsymbol{b}}^*(a, f_k)^2 \\ &- 2\tilde{\boldsymbol{b}}^*(a, f_k) \tilde{\boldsymbol{b}}(a, f_k)), \end{aligned} \tag{111}$$

where $\tilde{\boldsymbol{b}}$ is the annihilation operator obtained from (82) by substituting all ingredients by those with a tilde, and likewise for the creation operator $\tilde{\boldsymbol{b}}^*$. The operator in Eq. (111) is obviously ill-defined on the corresponding Fock space. However, the discussion suggests to consider more general transformations in order to avoid the disastrous terms such as (111). We thereby follow the logic of [64–66].

We confine ourselves to perturbation theory up to second order with respect to φ and π which themselves are considered to be of first order. This suggests to restrict to transformations to be linear in φ and π such as (106). This keeps the second order nature of $H_{\rm pert}$ while higher order transformations would be hidden anyway. The corresponding contact terms for the homogeneous degrees of freedom will then be already of second order at leading order as in Eq. (110), and we can restrain the precision of the canonical transformation to second order. Let us emphasize that these linear transformations in the perturbative fields will be supplemented by transformations of the homogeneous variables which are of second perturbative order. Such a transformation of the full phase space allows one to obtain a diagonal perturbative Hamiltonian (meaning that there are no terms quadratic in only annihilation or only creation operators), and hence, a normalizable vacuum state (see Ref. [64]).

To shorten the notation, we use the letters (q^j, p_j) with $j = 1, 2$ for denoting the homogeneous variables (a, p) and (ϕ, μ), and we suppress indices whenever possible. We consider the classical fields $(\varphi, \pi) \in H^3(\mathcal{B}) \times H^2(\mathcal{B})$ and apply a set of transformations, (r, s, t, u), which relate the original fields (φ, π) and the transformed fields $(\tilde{\varphi}, \tilde{\pi})$. Note that these transformations are operators on the space of (a certain class of) functions, or rather fields, on $\mathcal{B}$. We write them as

$$\varphi(x) := (r(q, p)\tilde{\varphi}(q, p))(x) + (s(q, p)\tilde{\pi}(q, p))(x), \tag{112}$$

$$\pi(x) := (t(q, p)\tilde{\varphi}(q, p))(x) + (u(q, p)\tilde{\pi}(q, p))(x). \tag{113}$$

We keep the transformations (r, s, t, u) as generic as possible and let them depend on all (q, p). Furthermore, they can involve the Laplace operator which leads to translation invariant operators on the field space. Regarding the Hilbert-Schmidt condition, it suffices to restrict the transformations to depend on the Laplacian so that they mutually

commute and are symmetric on $L^2(\mathcal{B})$. r, s, t, and u are restricted to be real valued since all the variables are.

The transformations should meet certain conditions. First, the new system of fields should satisfy the standard Poisson bracket relations with respect to the inhomogeneous fields such that the transformations (112) and (113) be canonical. Therefore, we smear the fields with functions $f, g \in L^2(\mathcal{B})$ and require that the new and the original fields satisfy the Poisson relations, $\{\varphi(f), \pi(g)\} = \langle f, g\rangle = \{\tilde{\varphi}(f), \tilde{\pi}(g)\}$. This leads to the condition $ur - st = 1$.

Further conditions for the operators (r, s, t, u) arise from plugging the transformations (112) and (113) into the symplectic potential (108). This generates terms in the symplectic potential which give rise to transformations of the homogeneous variables at second order in the perturbations. Let us denote these second order transformations by $q^j \to q^j + \chi_q^j =: \tilde{q}^j$ and $p_j \to p_j + \chi_{p,j} =: \tilde{p}_j$ such that the symplectic potential reads

$$\begin{aligned}(p_j + \chi_{p,j})\mathrm{d}(q^j + \chi_q^j) &= p_j \mathrm{d}q^j + \chi_{p,j}\mathrm{d}q^j + p_j \mathrm{d}\chi_q^j + \mathcal{O}(\chi^2)\\ &= p_j \mathrm{d}q^j + \chi_{p,j}\mathrm{d}q^j - \mathrm{d}p_j \chi_q^j + \mathcal{O}(\chi^2).\end{aligned} \tag{114}$$

To confine the possible transformations, we plug these results into the Hamilton function and develop the latter up to second order in the perturbations. As the perturbative Hamilton function, H_{pert}, is already of second order, it is allowed to simply replace the original homogeneous variables by the dashed ones. Regarding the homogeneous Hamiltonian H_{hom}, the cropping of higher order terms suggests to Taylor expand with respect to the homogeneous degrees of freedom, namely

$$\begin{aligned}H_{\mathrm{hom}}(q, p) = H_{\mathrm{hom}}(\tilde{q}, \tilde{p}) &- \frac{\partial H_{\mathrm{hom}}}{\partial q^j}(\tilde{q}, \tilde{p})\chi_q^j(\tilde{q}, \tilde{p})\\ &- \frac{\partial H_{\mathrm{hom}}}{\partial p_j}(\tilde{q}, \tilde{p})\chi_{p,j}(\tilde{q}, \tilde{p}).\end{aligned} \tag{115}$$

Using the notation of the above transformations, this gives the second order contributions,

$$\begin{aligned}H_{\mathrm{hom}} = \tilde{H}_{\mathrm{hom}} - \frac{1}{2}(&\langle\tilde{\varphi}, (t\dot{r} - r\dot{t})\tilde{\varphi}\rangle + \langle\tilde{\pi}, (u\dot{s} - s\dot{u})\tilde{\pi}\rangle\\ &+ \langle\tilde{\pi}, (u\dot{r} - r\dot{u} + t\dot{s} - s\dot{t})\tilde{\varphi}\rangle).\end{aligned}$$

To shorten notation, we introduce the functions and operators $b := a^{-3}$, $c := a$, $\omega^2 = c(-\Delta) + k$, and $k := m^2 a^3$. In all these expressions, it is allowed to replace a by $\tilde{a}$ in agreement with the truncation after second order in the perturbations. In total, the second order contributions of the transformed Hamilton function are then given by

$$\begin{aligned}&(H_{\mathrm{hom}} - \tilde{H}_{\mathrm{hom}}) + H_{\mathrm{pert}}\\ &= \frac{1}{2}\langle\tilde{\varphi}, (t^2 b + r^2\omega^2 - (t\dot{r} - r\dot{t}))\tilde{\varphi}\rangle\\ &\quad + \frac{1}{2}\langle\tilde{\pi}, (u^2 b + s^2\omega^2 - (u\dot{s} - s\dot{u}))\tilde{\pi}\rangle\\ &\quad + \langle\tilde{\pi}, (ubt + s\omega^2 r - \frac{1}{2}(u\dot{r} - r\dot{u} + t\dot{s} - s\dot{t}))\tilde{\varphi}\rangle.\end{aligned} \tag{116}$$

The last term is ill-defined on any Fock space; hence its round bracket must vanish. The round bracket in the second line is supposed to be a positive operator in order to have a positive kinetic energy contribution. We therefore denote it as "d^2" and it is allowed to be a function of all the homogeneous variables (q, p) as well as of the Laplace operator. Then, we require the round bracket operator in the first line to be of the form, $d^2(-\Delta + M^2)$, where M^2 is a function of the homogeneous variables to be determined. This guarantees that we can factor out d^2 from the expression of the Hamiltonian, leaving us with a Hamiltonian density of standard form with constant coefficients for the Laplacian such that the Hilbert-Schmidt condition is satisfied.

The simplest choice for the operators (r, s, t, u) is that, first, none of them depends on $-\Delta$ and, second, also d does not depend on $-\Delta$. This selection is in fact unique and provides the following solutions. First, the fact that d^2 should not depend on the Laplace operator implies that $s \equiv 0$ so that the term $s^2\omega^2$ vanishes. Consequently, it holds true that $ru = 1$, and hence the whole round bracket of the kinetic term reduces to $d^2 = u^2 b$. Since b and evidently u^2 are manifestly positive, it follows that the entire term is positive. The corresponding algebraic solution for the operator t is then

$$t = -b^{-1}u^{-2}\dot{u}. \tag{117}$$

Note that we freely interchange the order of the operators as they are commuting. Eventually, we consider the first line and recall that the operator in the round brackets must equal $d^2(-\Delta + M^2)$, where M denotes an effective mass term which depends on the dashed variables but not on the Laplace operator. With the previous results, this condition evaluates to

$$\begin{aligned}d^2(-\Delta + M^2) = u^{-2}(&b^{-1}u^{-2}\dot{u}^2 + c(-\Delta + \tilde{k})\\ &+ \{H_{\mathrm{hom}}, tu\}_{\mathrm{hom}}).\end{aligned} \tag{118}$$

In order to fulfill the Hilbert-Schmidt condition, it is essential that the coefficients of the Laplace operators on both sides of the equations match such that finally they only hold a constant coefficient up to a global prefactor. Using that $d^2 = u^2 b$, this yields

$$u^4 = b^{-1}c = \tilde{a}^4, \tag{119}$$

where we employed the expressions for b and c and directly replaced the original variables by the dashed ones. Ergo, we have that $d^2 = \tilde{a}^{-1}$. The transformed mass term becomes in general form and for the concrete case here

$$M^2 = d^{-2}u^{-2}(b^{-1}u^{-2}\dot{u}^2 + c\tilde{k} + \{H_{\mathrm{hom}}, tu\}_{\mathrm{hom}}) \quad (120)$$

$$= \dot{\tilde{a}}^2 + m^2\tilde{a}^2 - \frac{1}{\tilde{a}}\{H_{\mathrm{hom}}, \tilde{a}^2\dot{\tilde{a}}\}_{\mathrm{hom}}. \quad (121)$$

Thus, it is finally achievable to make the Hamilton symbol $H_{\mathrm{pert}}(\tilde{q}, \tilde{p})$ well-defined on all the Fock spaces $\mathcal{F}_s(\tilde{a})$ for any $\tilde{a} \in \mathbb{R}^+$. Note that the case $s = 0$, $d^2 = \tilde{a}^{-1}$ reflects the unitarity of the Fock quantization (see, e.g., Ref. [109]).

Due to the dependence of the mass term $\tilde{m}^2$ on all the homogeneous degrees of freedom $(\tilde{q}^j, \tilde{p}_j)$, the Born-Oppenheimer method is no longer available. We are *forced* into its generalization, the space adiabatic scheme. However, we notice that the square of the mass term M is *not* manifestly positive and the question arises how to deal with that fact. With the specific choices made here, there is no freedom left to change this without making the coefficients (r, s, t, u) also depend on $-\Delta$. Whether this can be improved by exploiting the complete freedom for those operators will be left for future research.

In this respect, we draw attention to Ref. [94]. There, the starting point is a Hamiltonian of second order in the inhomogeneous degrees of freedom with standard form up to a prefactor depending on the homogeneous degrees of freedom. Furthermore, the mass squared is a generic function of the homogeneous degrees of freedom. Hence, we are precisely in the situation arrived at above after the (almost) canonical transformations. The analysis of [94] investigates the most general Fock representation labeled by the homogeneous variables that supports such a Hamiltonian and at the same time provides a canonical transformation of the homogeneous sector to variables which directly commute with the associated *annihilation and creation variables*.

This procedure has the advantage that the Hilbert-Schmidt condition is trivially solved because the annihilation and creation operators do not depend on the transformed homogeneous degrees of freedom in the sense that together they build the basic canonical building blocks. As such, the strategy is similar in spirit to the present one although the details are different. Their strategy does not allow for an algebraic solution, rather it is necessary to solve a system consisting of two nonlinear (but semilinear) first order partial differential equations for *complex coefficient operators* coming from the Hamiltonian vector field of H_{hom}. These equations then guarantee that all conditions are met including the positivity of the mass term. One of the conditions is equivalent to the fixed point equation of the adiabatic vacua construction [21], and the other determines an otherwise free phase. While those partial differential equations are well-posed and can be solved in principle by the method of characteristics, it is generically very hard to solve the system explicitly given the detailed form of H_{hom}. This, however, is a prerequisite to quantize the homogeneous sector as well. Thus for the purpose of the papers in this series, we stick to the method sketched above, although the possibility to ensure the positivity of the mass squared is very attractive. We deal with the complications that arise for negative mass squared terms more explicitly in the next subsection.

There is also another, independent reason why the approach of [94] is especially attractive: Since annihilation and creation operators commute with the operators of the homogeneous sector, the latter operators preserve the domain of the inhomogeneous part of the Hamiltonian. This is not necessarily the case when we simply assure the Hilbert-Schmidt condition. To see this, suppose that the symbol $K(a)$ that we derived in Eq. (102) is of the Hilbert-Schmidt type and only depends on a. Then, the vector $\hat{\boldsymbol{H}}_{\mathrm{pert}}(\hat{\boldsymbol{p}}_a\Omega)$ can be computed recalling that $\hat{\boldsymbol{H}}_{\mathrm{pert}} = \hat{\mathcal{W}}(\boldsymbol{H}_{\mathrm{pert}})$ is the Weyl quantization of the operator-valued symbol $\boldsymbol{H}_{\mathrm{pert}}$ which was explicitly given in Eq. (84) and by means of the explicit representation of $\hat{\boldsymbol{p}}_a\Omega$ in (104). After shifting the annihilation operator due to the Hamilton operator to the right side of the resulting operator, we obtain

$$\hat{\boldsymbol{H}}_{\mathrm{pert}}(\hat{\boldsymbol{p}}_a\Omega) = -i\hat{\mathcal{W}}(a^{-3}) \times \sum_{k\in\Sigma} \hat{\boldsymbol{b}}^*(a, f_k)(\hat{\omega}(a)\hat{K}(a)\hat{\boldsymbol{b}}^*(a))(f_k)\Omega. \quad (122)$$

The operator symbol $\omega(a)K(a)$ is given with (102) by $\partial_a\omega(a)$ which grows as $|k|$ for large $|k|$ if the coefficient in front of the Laplace operator depends on the scale factor. Even if the Laplace operator does not carry an a-dependent prefactor, the resulting expression decays at most as $1/|k|$. Hence, the infinite sum over all the k's and this falloff property of the above operator prevents $(\hat{\boldsymbol{H}}_{\mathrm{pert}}\hat{\boldsymbol{p}}_a)$ from being a well-defined operator on Fock space. By itself this is not a problem because we want to consider the spectrum of $\hat{\boldsymbol{H}} = \hat{\mathcal{W}}(\boldsymbol{H})$ rather than $\hat{\mathcal{W}}(\boldsymbol{H}_{\mathrm{pert}})$ which does not require one o have the commutator $[\hat{\boldsymbol{p}}_a, \hat{\boldsymbol{H}}_{\mathrm{pert}}]$ defined on the Fock space. Nevertheless, it would be a convenient property to have. Thereby, we recall that once $\hat{\boldsymbol{H}}$ can be constructed as a self-adjoint operator, the existence of a dense and invariant domain is granted [107,108]. Again, to have a domain left invariant by the operators of the homogeneous sector could possibly be achieved within the context of this paper if we considered more generic transformations (r, s, t, u) and in particular made the derivative of the function $d(-\Delta + M^2)d$, with respect to the homogeneous variables to decay faster than $1/|k|$.

Let us finally note that the above-proposed transformations apply to rather generic second order Hamiltonians. More generally, one may have several matter or geometry species, e.g., scalar, vector, tensor, and spinor modes [110]. A straightforward proposition for this consists in applying

an individual canonical transformation for any of the species labeled by "s" which is then parametrized by (r_s, s_s, t_s, u_s). The transformation of the symplectic potential for each of the species simultaneously has the simple effect that the $(\chi_s^j, \chi_{s,j})$ corrections for all the species add up. Since we perturb H_{hom} only linearly in $(\chi_s^j, \chi_{s,j})$ and because it is allowed to drop the $(\chi_{s,j}, \chi_{s,j})$ corrections for the homogeneous variables in the perturbed Hamiltonian H_{pert} in second order perturbation theory, the species contributions never mix. Accordingly, we can consider the Hamiltonian symbols to be well-defined on the corresponding Fock spaces, and the Hilbert-Schmidt conditions are all solved.

D. Indefinite mass squared operators

As alluded to above, the effective mass squared terms M^2 that result from the transformations discussed in the previous section fail to be positive in general on the entire phase space. In what follows, we present several strategies for how to deal with this problem, none of which is entirely satisfactory as they either lead to instabilities or contain *ad hoc* elements. Since this is a new situation, the discussion will be mostly exploratory. We pursue five different lines of thought.

The first suggestion is to exploit the full freedom in the transformations (r, s, t, u) beyond the restricted ansatz of the previous subsection and to make positivity manifest thereby. We emphasize that there should be substantial freedom in the choice of these transformations, and the domain of the phase space where the mass squared functions become negative depends on this freedom. Thereupon, this region should not be of any physical significance and the generalization of the transformations is hence a physically motivated criterion.

A second approach consists in restricting the classical phase space of the homogeneous degrees of freedom to the set of points (q, p) where $M_s^2(q, p) \geq 0$ for all perturbation species s. This restriction can be achieved by defining new variables v_s for the associated homogeneous variables and to set $v_s^2 = M_s^2$. An especially nice situation occurs when the mass terms M_s^2 have mutually vanishing Poisson brackets between them. Namely, in this case, it is reasonable to consider them as action variables and determine the corresponding angle variables as their conjugate variables. This is in particular possible for a single field species as for the Klein-Gordon Hamiltonian in (81). Unfortunately, the procedure already fails for a set of tensorial and scalar field modes present at the same time. More generally, we may be able to write M_s^2 in the form

$$M_s^2(q, p) = (F^s)^2(q, p) v_s^2(q, p), \tag{123}$$

where $(F^s)^2(q, p)$ is a positive function and $v_s^2(q, p)$ may still be indefinite for certain species but the v_s^2 are mutually commuting for all s for which v_s^2 is indefinite. Then, apply the action angle prescription to the new variables v_s^2 and assume that the number of homogeneous variable pairs is at least as large as the number of indefinite mass squared terms. In the most general case, we solve the equations $v_s^2 = v_s^2(q, p)$ for some homogeneous momenta $p_s = F^s(v_s, q^s, (q, p))$ where the (q, p) stands for the remaining phase space variables which are not associated with a perturbative species. The variables $(v_s, q^s, (q, p))$ coordinatize a new phase space with induced symplectic structure. While these variables fail to be canonical coordinates for the system, they are supposed to have full range in some $\mathbb{R}^{2m}$ in contrast to the p_s. Finally, we must pass to suitable Darboux coordinates and hope that they are global in order that we may apply Weyl quantization. To illustrate this, let us compute the mass squared operator for the model (81) with a set of homogeneous variables (a, p_a, ϕ, μ) and for which the homogeneous part of the Hamilton function reads

$$H_{\text{hom}}(a, p_a, \phi, \mu) = -\sigma \frac{p_a^2}{a} + \Lambda a^3 + \frac{\mu^2}{a^3} + m^2 a^3 \phi^2, \tag{124}$$

where we set σ as some constant factor. It follows for the purely homogeneous Hamilton equations of the geometric subsystem

$$\dot{a} = \{H_{\text{hom}}, a\}_{\text{hom}} = -2\sigma \frac{p_a}{a}, \tag{125}$$

$$\dot{p}_a = \{H_{\text{hom}}, p_a\}_{\text{hom}} = -\sigma \frac{p_a^2}{a^2} - 3\Lambda a^2 + \frac{3}{2}\left(\frac{\mu^2}{a^4} - m^2 a^2 \phi^2\right). \tag{126}$$

Thus, from (120) we obtain the new effective mass value after having performed the transformations of the homogeneous and the inhomogeneous fields up to second order as in the last section,

$$M^2 = a^2(m^2 - 6\sigma\Lambda - 3\sigma m^2 \phi^2) + 3\sigma \frac{\mu^2}{a^4} - 2\sigma^2 \frac{p_a^2}{a^2}. \tag{127}$$

Evidently, this mass function is indefinite. To illustrate the new procedure more clearly, suppose for the moment that we would have treated the homogeneous mode of the scalar field on equal footing with the inhomogeneous ones so that the (ϕ, μ)-dependent terms are missing from the Hamilton function (124) and the mass squared (127). Nonetheless, the mass squared function remains indefinite. For this setting, we define a new set of canonical variables (v, b), and we therefore introduce the constants $\delta^2 := 2\sigma^2/F_\Lambda^2$ with $F_\Lambda^2 := m^2 - 6\sigma\Lambda$. Let us assume that the cosmological constant Λ is so small that F_Λ^2 is indeed positive definite. We construe the new variables (v, b) according to

$$a^2 := v^2 + \delta^2 \frac{b^2}{v^2}, \qquad p_a := \frac{a}{v} b. \tag{128}$$

Note that the scale factor a is positive definite, and hence the square root of a^2 in (128) has only one branch. Thus, the new variables in (128) are uniquely defined away from $v = 0$. We compute the Poisson bracket of b and v to illustrate that the transformation is indeed canonical, namely

$$\{b, v\}_{\text{hom}} = \frac{1}{2}\left\{\frac{b}{v}, v^2\right\}_{\text{hom}} = \frac{1}{2}\left\{\frac{b}{v}, v^2 + \delta^2\left(\frac{b}{v}\right)^2\right\}_{\text{hom}} = \frac{1}{2}\left\{\frac{p_a}{a}, a^2\right\}_{\text{hom}} = \{p_a, a\}_{\text{hom}} = 1.$$

With this new set of canonical variables, it is possible to represent the mass term by

$$M^2 = F_\Lambda^2 v^2, \tag{129}$$

which is manifestly positive. However, the transformation (128) restricts the range of the original variables, i.e., $(a, p_a) \in \mathbb{R}_+ \times \mathbb{R}$, to the set of pairs (a, p_a) with $a^4 \geq \delta^2 p_a^2$, $p \in \mathbb{R}$. Hence, as a result of this procedure, the admissible domain of the phase space variables becomes restricted. To conclude, this procedure has the advantage that it stays within the standard quantum field framework with a positive mass squared, however, at the price of making the Hamiltonian even less polynomial and by a possibly nonphysical restriction of the phase space.

The third possibility is to take the indefinite mass terms M^2 seriously as they stand. Accordingly, for certain ranges of the homogeneous variables (q, p) the inhomogeneous symbol operator $\boldsymbol{H}_{\text{pert}}$ defines a *quantum field theory of tachyons*. A possibility to deal with the tachyonic instability was suggested in Ref. [93]: The idea is to construct a Fock space $\mathcal{H}(q, p)$ as before and to allow only those modes corresponding to eigenfunctions f_k of $-\Delta$ such that their eigenvalue of the frequency squared operator $\tilde{\omega}(\tilde{a})^2 = -\Delta^2 + M^2(q, p)$ is bigger or equal to zero, i.e., $k^2 + M^2(q, p) \geq 0$.

Accordingly, the more negative $M^2(q, p)$ becomes, the larger the required infrared cutoff on the admissible modes. Specifically, for the example mass squared from above, Eq. (127) with vanishing (ϕ, μ), we find that $\tilde{m}^2$ gets very negative for $(p_a^2/a^2) \to \infty$. Note that this term is proportional to $\dot{a}^2$. Hence, for a baryon or radiation dominated universe, this term certainly diverges at the classical big bang. For SAPT, this has the following consequence. Recall that for a torus of respective side lengths l in all directions, the mode numbers evaluate to $k^2 = (2\pi/l)^2 z^2$ with $z \in \mathbb{Z}^3 \backslash 0$. Let $S_\pm$ be the subsets of the slow phase space defined by $M^2(q, p) \geq 0$ and $M^2(q, p) < 0$, respectively. Let us consider the spectrum of the Hamilton symbol $\boldsymbol{H}_{\text{pert}}(q, p)$ and thereby use its representation with respect to annihilation and creation operators as in Eq. (84), but subject to the transformations above. The energy eigenvalue of some Fock state with finitely many excitations and at a fixed (q, p) is determined by a handful of numbers. In particular, we need the wave vectors $k = (2\pi/l) \cdot z$, and the number of excitations for one particular mode, denoted by n_z. Let us designate the torus factor $(2\pi/l)$ by τ. We collect the excitation numbers for all nonvanishing modes in a set called N. The spectral value is then given by

$$E_N(q, p) = \sum_{z \in \mathbb{Z}^3} n_z \sqrt{\tau^2 z^2 + M^2(q, p)}. \tag{130}$$

The mode number configurations which give rise to the same $E_N(q, p)$ determine the degeneracy of $E_N(q, p)$. First, we see that varying the wave vectors z and the excitation numbers n_z does not leave the energy (130) invariant [almost everywhere with respect to (q, p)]. Otherwise, the numbers $\sqrt{\tau^2 z^2 + M^2(q, p)}$ would be linearly dependent over the positive rationals which is not the case almost everywhere. It follows that the only degeneracy lies in choosing the entries of z with the same z^2. Since there are at least eight possibilities, the degeneracy of $E_N(q, p)$ is at least $8^{|N|}$ independent of (q, p) when $(q, p) \in S_+$. However, for $(q, p) \in S_-$ we compute $r(q, p)^2 := -M(q, p)^2$ and can only allow the energy bands characterized by a wave vector $\vec{z}$ with $\tau^2 z^2 \geq r(q, p)^2$. We then impose to disregard the eigenenergies $E_z(q, p)$ associated with the wave vectors z for which it holds that $\tau^2 z^2 < r(q, p)^2$.

Consequently, also the mode eigenstates $\psi_z(q, p)$ and their related Moyal projectors and unitaries $\boldsymbol{\pi}_{z,0}(q, p)$ and $\boldsymbol{u}_z(q, p)$ will be neglected. The Fock space $\mathcal{F}_s(q, p)_{r(q,p)}$ restricts to the subspace of $\mathcal{F}_s(q, p)$ spanned by the $\psi_z(q, p)$ with $\tau^2 z^2 \geq r(q, p)^2$. This implies that for some given wave vector z, the eigenenergy function $(q, p) \mapsto E_z(q, p)$ has a discontinuity at the surface $\tau^2 z^2 = r(q, p)^2$ in the phase space. This can be problematic when computing the Moyal products which ask to take derivatives with respect to the homogeneous variables (q, p). Nevertheless, this procedure has the advantage of not worsening the amount of nonpolynomiality and of not modifying the phase space of the slow sector; however, the physical interpretation of the mode off-switching remains obscure: Typically this happens at the classical big bang singularity, and in the limit of the vanishing scale factor, all modes would need to be removed. One could speculate whether this presents a self-regulating effect of a new kind in the sense that the matter density and thus curvature automatically vanish as we approach the singularity, and thus eventually avoid the big bang.

The fourth proposal presented here suggests to modify the Weyl quantization procedure for the homogeneous variables which is part of the space adiabatic perturbation scheme. In particular, the idea is to restrict the phase space

integral over the homogeneous degrees of freedom that enters the Weyl quantization to S_+ in an *ad hoc* manner. This can be achieved by multiplying all symbol functions such as the Hamiltonian $\boldsymbol{H}(q,p)$ with the characteristic function $\chi_{S_+}(q,p)$ of S_+. This is again not differentiable, and it would be more appropriate to substitute χ_{S_+} by a smoothed version of it, i.e., a smooth function that is zero in S_- and smoothly reaches unity within S_+ in an arbitrarily small neighborhood of the boundary ∂S_+. Of course, the quantum theory then will depend on that smoothing procedure which introduces ambiguities and technical challenges because the smoothed version of χ_{S_+} is a highly nonpolynomial function of (q,p).

Finally, the fifth suggestion for how to deal with the indefinite mass squared function is to consider a mode decomposition of the inhomogeneous Hamiltonian contribution $\boldsymbol{H}_{\rm pert}(q,p)$ and to write for $(q,p)\in S_-$, $\boldsymbol{H}_{\rm pert}(q,p)=\boldsymbol{H}^+_{\rm pert}(q,p)+\boldsymbol{H}^-_{\rm pert}(q,p)$ where $\boldsymbol{H}^+_{\rm pert}(q,p)$ is the contribution from all modes k with $k^2\geq r(q,p)^2$. Then, $\boldsymbol{H}^+_{\rm pert}(q,p)$ may be quantized as before, while the quantization of $\boldsymbol{H}^-_{\rm pert}(q,p)$ represents a finite sum of flipped quantum harmonic oscillators. The difference of this strategy compared to the third method, where we performed a mode cutting, is exactly that we do not discard $\boldsymbol{H}^-_{\rm pert}$. We notice that the spectrum of a flipped harmonic oscillator is purely of the absolutely continuous type [113]. Thus, the spectrum of $\boldsymbol{H}^-_{\rm pert}$ changes drastically when we transit from S_+ to S_- with corresponding consequences for the space adiabatic perturbation scheme. Besides, such a theory would be unstable.

For the model (81) with only a Klein-Gordon field as the matter content and a purely homogeneous geometry, the second strategy of performing a canonical transformation in the homogeneous sector seems to be most promising as we will see in the third paper of this series [82].

E. Nonpolynomial operators

The purely homogeneous piece $H_{\rm hom}(a,p_a,\phi,\mu)$ of the Hamiltonian in the previous section, Eq. (124), is nonpolynomial in the scale factor a and contains inverse powers of it. The mass squared corrections coming from the canonical transformation in Sec. III C contain derivatives of $H_{\rm hom}(a,p_a,\phi,\mu)$ which increase that negative power. Furthermore, the adiabatic corrections contain additional derivatives of $H_{\rm hom}(a,p_a,\phi,\mu)$ of arbitrary order coming from the Moyal product which introduces further arbitrarily negative powers of a. Even worse, after the mass squared corrections, we potentially also find inverse powers of arbitrarily high order in the momentum p conjugate to a, the Mukahnov-Sasaki mass term being a prominent example. It transpires that it would be desirable to dispose of a dense set of vectors which is invariant under *any* of the operators corresponding to a^i and p^j, $i,j\in\mathbb{Z}$. In LQC, one deals with negative powers of a by using a nonstandard representation inspired by the representation used in the full LQG theory such that the spectrum of a is pure point rather than absolutely continuous. Hence, the commutator between fractional powers of a and Weyl elements for p both is densely defined and introduces the desired negative powers of a. This comes at the price that the operator corresponding to p_a does not exist and one consequently needs to approximate it by polynomials in Weyl elements. However, negative powers of p_a would then also need to be approximated by inverse polynomials of Weyl elements, and these are not in the domain of a so that for our purpose the representation chosen in LQC is of no direct advantage.

We thus advocate to take an unbiased point of view and ask whether it is possible to choose the above desired domain directly in the Schrödinger representation. The advantage would be that the operators corresponding to a, p exist. We found the following answers:

Theorem 1. Consider the canonical pair of position and momentum operators $(\hat{q},\hat{p})$ and use a Schrödinger representation on the Hilbert space $\mathcal{H}=L^2(\mathbb{R},{\rm d}x)$, i.e., $(\hat{q}\psi)(x)=x\psi(x)$, $(\hat{p}\psi)(x)=-i{\rm d}\psi(x)/{\rm d}x$ for $\psi\in\mathcal{H}$. Then, the following holds true:

(1) There exists a dense and invariant domain $\mathcal{D}\subset\mathcal{H}$ for the operators $\hat{q}^i$ and $\hat{p}^j$ where $i\in\mathbb{Z}$ and $j\in\mathbb{N}_0$. This domain $\mathcal{D}$ consists of smooth functions of rapid decrease, both at $x=0$ and at $x=\pm\infty$.

(2) The domain $\mathcal{D}$ is spanned by functions $\{\xi_n(x)\}$ with $n\in\mathbb{Z}$ whose inner products can be computed analytically in closed form. Correspondingly, an orthonormal basis can be constructed by means of the Gram-Schmidt procedure.

(3) Let $f(x)$ be a function such that both $f(x)$ and $f^{-1}(x)$ are polynomially bounded and smooth except possibly at $x=0$ or $x\pm\infty$. Furthermore, let $f_1(x),\ldots,f_N(x)$ be polynomials in x. Then there exists a common domain $\mathcal{D}_L(f)\subset\mathcal{D}$ for the operators of item 1. and of the operators corresponding to the symbols $f_k(q,p):=|f(q)|^2f_k(q)p^{-k}$ with $k=1,\ldots,N$, in suitable symmetric orderings, where L depends on both N and the degree of the polynomials f_k.

The proof of this theorem can be found in Ref. [97]. Thereby, note that $\hat{p}^{-1}$ is a symmetric operator with distribution kernel

$$(\hat{p}^{-1}\psi)(x)=-\frac{i}{2}\int_{\mathbb{R}}{\rm d}y\,{\rm sgn}(x-y)\psi(y),\qquad(131)$$

which can easily be seen by applying $\hat{p}=-i{\rm d}/{\rm d}x$ from the left and using that ${\rm dsgn}(x-y)/{\rm d}x=-2\delta(x-y)$ in a distributional sense. The domain of $\hat{p}^{-1}$ must be chosen carefully. Even if ψ is a Schwartz function, while $\hat{p}^{-1}\psi$ is smooth, it may not be of rapid decrease any more at infinity.

Likewise, it is a simple corollary that a dense and invariant domain for $\hat{p}^i\hat{q}^j$ with $i \in \mathbb{Z}$, $j \in \mathbb{N}_0$ is given by the Fourier transform of the functions of item (1) but that Fourier transform is not necessarily of rapid decrease in x any more. This is why the statement of item (3) is significantly weaker, in particular $\mathcal{D}_L(f)$ is not an invariant domain for the list of operators stated, and it is presently not clear whether it is dense. It is, however, certain that there exists no function in $\mathcal{D}$ orthogonal to $\mathcal{D}_L(f)$.

The idea for defining the rather singular symbols that we encounter in the homogeneous sector of quantum cosmology is thus as follows (provided that we can factor out a suitable $|f|^2$ as described above): At any order of the adiabatic expansion the terms that involve negative powers of p are of the form described in item (3) and are finite in number. Thus we use the ordering alluded to in item (3) and the domain described there. The other terms not involving negative powers of p are also defined on that domain since $\mathcal{D}_L(f) \subset \mathcal{D}$.

IV. CONCLUSION AND OUTLOOK

In the present first paper of this series, we provided the tools with which we intend to improve on the treatment of backreactions in quantum cosmology, in particular with respect to backreactions from the inhomogeneous cosmological quantum fields on the homogeneous quantum degrees of freedom. In the Introduction, Sec. I, we have provided the reader with an overview of the current state of research regarding the questions of a rigorous account for interactions in quantum gravity–matter and quantum cosmological systems, its relation to the semiclassical limit and possible approximation schemes that attempt to make progress in this direction. We have explained how and why the implementation of SAPT methods in quantum gravity and quantum cosmology can help to find, at least partial, answers to these questions, and how it naturally leads to an implementation of backreaction in quantum cosmology as a first step.

In Sec. II, we gave a detailed introduction to the original SAPT scheme as proposed by Panati, Spohn, and Teufel [52], and executed the iterative construction step by step in order to show that the scheme is consistent. Then, in Sec. III, we pointed to the challenges that occur when implementing SAPT methods within (inhomogeneous) quantum cosmology and suggested several solutions.

These considerations show that an application of SAPT methods within quantum cosmological perturbation theory may indeed be possible, and offer hence the possibility to unambiguously account for backreaction effects in such situations. Thus, we have prepared the ground to approach the various models that are being treated in the subsequent papers of the series. The plan of these subsequent papers is as follows:

In the second paper [81], we treat the two quantum mechanical models labeled as models (I) and (II) in Sec. III C. Model (I) is a standard quantum mechanical problem consisting of a polynomially coupled slow anharmonic and fast harmonic oscillator. They permit one to mimic the situation of model (II) and serve to illustrate the formalism. Model (II) considers the purely homogeneous cosmological sector, i.e., a homogeneous geometry coupled to the homogeneous part of a real-valued scalar matter field. In suitable variables, this model can be displayed as an inverted slow harmonic oscillator which is nonpolynomially coupled to a fast standard harmonic oscillator. The adiabatic perturbation parameter is related to the ratio of the coupling constants for gravity and the scalar field, and this ratio can be assumed to be very small. The adiabatic parameter can also be written as the ratio of the corresponding scalar field and Planck mass scales and is consequently very tiny if we consider the latter to be of the order of the mass scales that appear in the current Standard Model of elementary particle physics. This parameter will also organize the adiabatic perturbation expansion of the third and fourth papers.

In the third paper of the series [82], we consider as matter content a scalar Klein-Gordon field as well as Gaussian dust. The usual Hamiltonian constraint is now a physical Hamiltonian as shown in [105]. The full constraints, not only their perturbations, are already solved—all metric and inflaton degrees of freedom are physical observables. We expand the physical Hamiltonian to second order in the inhomogeneous modes leading to three scalar, one vector, and two tensor modes. For simplicity we consider only the quantization of the inflaton field; i.e., we drop all metric perturbations and keep only the homogeneous metric degrees of freedom.

Finally in the fourth paper [69], we consider as matter content just the real matter scalar field and follow closely the approach taken in [64–66] in order to extract the gauge-invariant observables. These are the well-known Mukhanov-Sasaki field in the scalar sector and the tensor mode perturbations (i.e., gravitations). This model is generically used in order to describe the physics of the early universe and provides hence the possibility to compare our future results with observations, for example, by computing primordial power spectra.

In all papers, we compute the backreaction effects to second order in the adiabatic perturbation parameter, thus displaying their existence and potential phenomenological importance. Note that the model of the second paper can be considered as the purely homogeneous truncation of both, the model of the third and the fourth paper, respectively, just that in the first case it is to be considered as a dynamical system with a true Hamiltonian, and in the second case that Hamiltonian is constrained to vanish. Accordingly, for the second paper, we are interested in the full spectrum of the

Hamiltonian which in appropriate variables can be understood as a harmonic oscillator nonpolynomially coupled to an inverted harmonic oscillator [113].

ACKNOWLEDGMENTS

We thank Beatriz Elizaga Navascues for drawing our attention to the hybrid LQC strategy to use canonical transformations in second order cosmological perturbation theory in order to satisfy the Hilbert-Schmidt condition in the corresponding quantum field theory. We are also very grateful to the anonymous referee for his/her detailed and very helpful comments to the first version of this manuscript. S. S. thanks the Heinrich Böll Foundation e.V. for financial and intellectual support and the German Academic Scholarship Foundation for intellectual support.

[1] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (W. H. Freeman, San Francisco, 1973).
[2] R. Wald, *General Relativity* (The University of Chicago Press, Chicago, 2010).
[3] M. D. Schwartz, *Quantum Field Theory and the Standard Model* (Cambridge University Press, Cambridge, England, 2014).
[4] P. A. Oesch *et al.*, A remarkably luminous galaxy at $z = 11.1$ measured with hubble space telescope grism spectroscopy, Astrophys. J. **819**, 129 (2016).
[5] R. A. Alpher and R. C. Herman, On the relative abundance of the elements, Phys. Rev. **74**, 12 (1948).
[6] P. J. E. Peebles, D. N. Schramm, E. L. Turner, and R. G. Kron, The case for the relativistic hot Big Bang cosmology, Nature (London) **352**, 769 (1991).
[7] N. Aghanim *et al.*, Planck 2018 results. I. Overview and the cosmological legacy of Planck, Astron. Astrophys. **641**, A1 (2020).
[8] D. N. Spergel, Cosmology today, Daedalus (Boston) **143**, 125 (2014).
[9] D. Nelson *et al.*, First results from the IllustrisTNG simulations: The galaxy color bimodality, Mon. Not. R. Astron. Soc. **475**, 624 (2018).
[10] F. Marinacci *et al.*, First results from the IllustrisTNG simulations: Radio haloes and magnetic fields, Mon. Not. R. Astron. Soc. **480**, 5113 (2018).
[11] V. Springel, R. Pakmor, A. Pillepich, R. Weinberger, D. Nelson, L. Hernquist, M. Vogelsberger, S. Genel, P. Torrey, F. Marinacci, and J. Naiman, First results from the IllustrisTNG simulations: Matter and galaxy clustering, Mon. Not. R. Astron. Soc. **475**, 676 (2018).
[12] J. P. Naiman *et al.*, First results from the IllustrisTNG simulations: A tale of two elements—chemical evolution of magnesium and europium, Mon. Not. R. Astron. Soc. **477**, 1206 (2018).
[13] A. Pillepich, D. Nelson, L. Hernquist, V. Springel, R. Pakmor, P. Torrey, R. Weinberger, S. Genel, J. P. Naiman, F. Marinacci, and M. Vogelsberger, First results from the IllustrisTNG simulations: the stellar mass content of groups and clusters of galaxies, Mon. Not. R. Astron. Soc. **475**, 648 (2018).
[14] J. M. Bardeen, P. J. Steinhardt, and M. S. Turner, Spontaneous creation of almost scale—free density perturbations in an inflationary universe, Phys. Rev. D **28**, 679 (1983).
[15] G. R. Blumenthal, S. M. Faber, J. R. Primack, and M. J. Rees, Formation of galaxies and large scale structure with cold dark matter, Nature (London) **311**, 517 (1984).
[16] A. Albrecht and P. Steinhardt, Cosmology for Grand Unified Theories with Radiatively Induced Symmetry Breaking, Phys. Rev. Lett. **48**, 1220 (1982).
[17] A. H. Guth and S. Y. Pi, Fluctuations in the New Inflationary Universe, Phys. Rev. Lett. **49**, 1110 (1982).
[18] S. W. Hawking, The development of irregularities in a single bubble inflationary universe, Phys. Lett. **115B**, 295 (1982).
[19] A. D. Linde, A new inflationary universe scenario: A possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems, Phys. Lett. **108B**, 389 (1982).
[20] A. A. Starobinsky, Dynamics of phase transition in the new inflationary universe scenario and generation of perturbations, Phys. Lett. **117B**, 175 (1982).
[21] S. Fulling, *Aspects of Quantum Field Theory in Curved Space-Time*, London Mathematical Society Student Texts, Vol. 7 (Cambridge University Press, Cambridge, England, 1989).
[22] G. Bertone and D. Hooper, History of dark matter, Rev. Mod. Phys. **90**, 045002 (2018).
[23] S. Perlmutter *et al.*, Measurements of Ω and Λ from 42 high redshift supernovae, Astrophys. J. **517**, 565 (1999).
[24] A. G. Riess *et al.*, Observational evidence from supernovae for an accelerating universe and a cosmological constant, Astron. J. **116**, 1009 (1998).
[25] E. Di Valentino *et al.*, Snowmass2021—Letter of interest cosmology intertwined II: The hubble constant tension, Astropart. Phys. **131**, 102605 (2021).
[26] G. Efstathiou, A Lockdown Perspective on the Hubble Tension (with comments from the SH0ES team), arXiv:2007.10716.
[27] W. L. Freedman, Measurements of the Hubble constant: Tensions in perspective, Astrophys. J. **919**, 16 (2021).
[28] S. Weinberg, Critical phenomena for field theorists, *Understanding the Fundamental Constitutents of Matter*, edited by A. Zichichi (Springer US, Boston, MA, 1978), pp. 1–52.
[29] S. Weinberg, Ultraviolet divergences in quantum theories of gravitation, *General Relativity: An Einstein Centenary Survey*, edited by S. W. Hawking and W. Israel (Cambridge University Press, Cambridge, England, 1980), pp. 790–831.

[30] R. Percacci, Asymptotic safety, *Approaches to Quantum Gravity: Toward a New Understanding of Space, Time and Matter*, edited by D. Oriti (Cambridge University Press, Cambridge, England, 2007), pp. 111–128.
[31] A. Eichhorn, An asymptotically safe guide to quantum gravity and matter, Front. Astron. Space Sci. **5**, 47 (2019).
[32] M. Ammon and J. Erdmenger, *Gauge/Gravity Duality—Foundations and Applications* (Cambridge University Press, Cambridge, England, 2015).
[33] J. Polchinski, *String Theory*, Vol. I, II (Cambridge University Press, Cambridge, England, 2005).
[34] J. Ambjorn, J. Jurkiewicz, and R. Loll, Causal dynamical triangulations and the quest for quantum gravity, *Foundations of Space and Time: Reflections on Quantum Gravity*, edited by J. Murugan, A. Weltman, and G. F. R. Ellis (Cambridge University Press, Cambridge, England, 2012), pp. 321–337.
[35] R. Loll, Quantum gravity from causal dynamical triangulations: A review, Classical Quantum Gravity **37**, 013002 (2020).
[36] A. Ashtekar and J. Pullin, *Loop Quantum Gravity* (World Scientific, Singapore, 2017).
[37] R. Gambini and J. Pullin, *A First Course in Loop Quantum Gravity* (Oxford University Press, New York, 2011).
[38] C. Rovelli, *Quantum Gravity*, Cambridge Monographs on Mathematical Physics (Cambridge University Press, Cambridge, England, 2010).
[39] T. Thiemann, *Modern Canonical Quantum General Relativity* (Cambridge University Press, Cambridge, England, 2008).
[40] D. Oriti, *Approaches to Quantum Gravity: Toward a New Understanding of Space, Time and Matter* (Cambridge University Press, Cambridge, England, 2009).
[41] T. Thiemann, Gauge field theory coherent states (GCS): 1. General properties, Classical Quantum Gravity **18**, 2025 (2001).
[42] T. Thiemann and O. Winkler, Gauge field theory coherent states (GCS). 2. Peakedness properties, Classical Quantum Gravity **18**, 2561 (2001).
[43] T. Thiemann and O. Winkler, Gauge field theory coherent states (GCS): 3. Ehrenfest theorems, Classical Quantum Gravity **18**, 4629 (2001).
[44] T. Thiemann and O. Winkler, Gauge field theory coherent states (GCS) 4: Infinite tensor product and thermodynamical limit, Classical Quantum Gravity **18**, 4997 (2001).
[45] H. Sahlmann and T. Thiemann, Towards the QFT on curved space-time limit of QGR. 1. A General scheme, Classical Quantum Gravity **23**, 867 (2006).
[46] H. Sahlmann and T. Thiemann, Towards the QFT on curved space-time limit of QGR. 2. A Concrete implementation, Classical Quantum Gravity **23**, 909 (2006).
[47] C. Kiefer, *Quantum Gravity, Second Edition* (Oxford University Press, New York, 2007).
[48] K. Giesel, J. Tambornino, and T. Thiemann, Born-Oppenheimer decomposition for quantum fields on quantum spacetimes, arXiv:0911.5331.
[49] M. Born and R. Oppenheimer, Zur Quantentheorie der Molekeln, Ann. Phys. (Berlin) **389**, 457 (1927).
[50] A. Messiah, *Quantum Mechanics*, Vol. I. (North Holland Publishing Company, Amsterdam, 1961).
[51] G. Panati, H. Spohn, and S. Teufel, Space-adiabatic perturbation theory, Adv. Theor. Math. Phys. **7**, 145 (2003).
[52] S. Teufel, *Adiabatic Perturbation Theory in Quantum Dynamics*, Lecture Notes in Mathematics Vol. 1821 (Springer-Verlag, Berlin, 2003).
[53] R. Brummelhuis and J. Nourrigat, Scattering amplitude for Dirac operators, Commun. Part. Differ. Equations **24**, 377 (1999).
[54] A. Martinez and V. Sordoni, On the time-dependent born-oppenheimer approximation with smooth potential, Comp. Rend. Math. **334**, 185 (2002).
[55] G. Nenciu and V. Sordoni, Semiclassical limit for multistate Klein-Gordon systems: Almost invariant subspaces and scattering theory, J. Math. Phys. (N.Y.) **45**, 3676 (2004).
[56] L. E. Parker, Quantized fields and particle creation in expanding universes. I., Phys. Rev. **183**, 1057 (1969).
[57] L. E. Parker and S. A. Fulling, Adiabatic regularization of the energy momentum tensor of a quantized field in homogeneous spaces, Phys. Rev. D **9**, 341 (1974).
[58] A. Messiah, *Quantum Mechanics*, Vol. II (North Holland Publishing Company, Amsterdam, 1962).
[59] A. Stottmeister and T. Thiemann, Coherent states, quantum gravity and the Born- Oppenheimer approximation, I: General considerations, J. Math. Phys. (N.Y.) **57**, 063509 (2016).
[60] A. Stottmeister and T. Thiemann, Coherent states, quantum gravity and the Born- Oppenheimer approximation, II: Compact Lie Groups, J. Math. Phys. (N.Y.) **57**, 073501 (2016).
[61] A. Stottmeister and T. Thiemann, Coherent states, quantum gravity and the Born- Oppenheimer approximation, III: Applications to loop quantum gravity, J. Math. Phys. (N.Y.) **57**, 083509 (2016).
[62] V. Mukhanov, *Physical Foundations of Cosmology* (Cambridge University Press, Cambridge, England, 2005).
[63] J. J. Halliwell and S. W. Hawking, Origin of structure in the Universe, Phys. Rev. D **31**, 1777 (1985).
[64] L. Castelló Gomar, M. Martín-Benito, and G. A. Mena Marugán, Gauge-invariant Perturbations in hybrid quantum cosmology, J. Cosmol. Astropart. Phys. 06 (2015) 045.
[65] L. Castelló Gomar, M. Martín-Benito, and G. A. Mena Marugán, Quantum corrections to the Mukhanov-Sasaki equations, Phys. Rev. D **93**, 104025 (2016).
[66] B. Elizaga Navascués and G. A. Mena Marugán, Hybrid loop quantum cosmology: An overview, Front. Astron. Space Sci. **8**, 81 (2021).
[67] F. Benítez Martínez and J. Olmedo, Primordial tensor modes of the early Universe, Phys. Rev. D **93**, 124008 (2016).
[68] V. F. Mukhanov, Quantum theory of gauge invariant cosmological perturbations, Sov. Phys. JETP **67**, 1297 (1988).
[69] S. Schander and T. Thiemann, this issue, Quantum cosmological backreactions IV: Constrained quantum cosmological perturbation theory, Phys. Rev. D **105**, 106012 (2022).
[70] A. Ashtekar, M. Bojowald, and J. Lewandowski, Mathematical structure of loop quantum cosmology, Adv. Theor. Math. Phys. **7**, 233 (2003).

[71] A. Ashtekar, T. Pawlowski, and P. Singh, Quantum nature of the big bang: Improved dynamics, Phys. Rev. D **74**, 084003 (2006).
[72] M. Bojowald, Loop quantum cosmology, Living Rev. Relativity **11**, 4 (2008).
[73] H. J. Groenewold, On the principles of elementary quantum mechanics, Physica (Amsterdam) **12**, 405 (1946).
[74] J. E. Moyal, Quantum mechanics as a statistical theory, Math. Proc. Cambridge Philos. Soc. **45**, 99 (1949).
[75] H. Weyl, Quantenmechanik und Gruppentheorie, Z. Phys. **46**, 1 (1927).
[76] M. Blaszak and Z. Domanski, Phase space quantum mechanics, Ann. Phys. (Amsterdam) **327**, 167 (2012).
[77] I. Agullo, A. Ashtekar, and W. Nelson, Extension of the quantum theory of cosmological perturbations to the Planck era, Phys. Rev. D **87**, 043507 (2013).
[78] M. Assanioussi, A. Dapor, and J. Lewandowski, Rainbow metric from quantum gravity, Phys. Lett. B **751**, 302 (2015).
[79] A. Barrau, M. Bojowald, G. Calcagni, J. Grain, and M. Kagan, Anomaly-free cosmological perturbations in effective canonical quantum gravity, J. Cosmol. Astropart. Phys. 05 (2015) 051.
[80] I. Agullo, A. Ashtekar, and W. Nelson, The pre-inflationary dynamics of loop quantum cosmology: Confronting quantum gravity with observations, Classical Quantum Gravity **30**, 085014 (2013).
[81] J. Neuser, S. Schander, and T. Thiemann, following paper, Quantum cosmological backreactions II: Purely homogeneous quantum cosmology, Phys. Rev. D **105**, 106010 (2022).
[82] S. Schander and T. Thiemann, this issue, Quantum cosmological backreactions III: Deparametrized quantum cosmological perturbation theory, Phys. Rev. D **105**, 106011 (2022).
[83] S. Schander and T. Thiemann, Backreaction in Cosmology, Front. Astron. Space Sci. **8**, 692198 (2021).
[84] W. Thirring, *A Course in Mathematical Physics 4—Quantum Mechanics of Large Systems* (Springer-Verlag, Wien, 1983).
[85] F. Bayen, M. Flato, C. Fronsdal, A. Lichnerowicz, and D. Sternheimer, Deformation theory and quantization. 1. Deformations of symplectic structures, Ann. Phys. (N.Y.) **111**, 61 (1978).
[86] F. Bayen, M. Flato, C. Fronsdal, A. Lichnerowicz, and D. Sternheimer, Deformation theory and quantization. 2. Physical applications, Ann. Phys. (N.Y.) **111**, 111 (1978).
[87] S. Waldmann, Deformation quantization: Observable algebras, states and representation theory, *Proceedings of the 2nd Summer School in Modern Mathematical Physics* (2003), arXiv:hep-th/0303080.
[88] D. A. Dubin, M. A. Hennings, and T. B. Smith, *Mathematical Aspects of Weyl Quantization and Phase* (World Scientific, Singapore, 2000).
[89] G. B. Folland, *Harmonic Analysis in Phase Space* (Princeton University Press, Princeton, NJ, 1989).
[90] J. J. Kohn and L. Nirenberg, An algebra of pseudo-differential operators, Commun. Pure Appl. Math. **18**, 269 (1965).
[91] L. Hörmander, *The Analysis of Linear Partial Differential Operators II—Differential Operators with Constant Coefficients* (Springer Verlag, Berlin-Heidelberg, 1983).
[92] L. Hörmander, *The Analysis of Linear Partial Differential Operators III—Pseudo- Differential Operators* (Springer Verlag, Berlin–Heidelberg, 1994).
[93] M. J. Radzikowski, Stable, renormalizable, scalar tachyonic quantum field theory with chronology protection, arXiv:0804.4534.
[94] B. Elizaga Navascués, G. A. Mena Marugán, and T. Thiemann, Hamiltonian diagonalization in hybrid quantum cosmology, Classical Quantum Gravity **36**, 185010 (2019).
[95] T. Thiemann, Anomaly-free formulation of nonperturbative, four-dimensional Lorentzian quantum gravity, Phys. Lett. B **380**, 257 (1996).
[96] T. Thiemann, Quantum spin dynamics (QSD), Classical Quantum Gravity **15**, 839 (1998).
[97] T. Thiemann, Properties of a Smooth, Dense, Invariant Domain for Singular Potential Schrödinger Operators (unpublished).
[98] K. Giesel, S. Hofmann, T. Thiemann, and O. Winkler, Manifestly gauge-invariant general relativistic perturbation theory. I. Foundations, Classical Quantum Gravity **27**, 055005 (2010).
[99] K. Giesel, S. Hofmann, T. Thiemann, and O. Winkler, Manifestly Gauge-invariant general relativistic perturbation theory. II. FRW background and first order, Classical Quantum Gravity **27**, 055006 (2010).
[100] H. Kodama and M. Sasaki, Cosmological perturbation theory, Prog. Theor. Phys. Suppl. **78**, 1 (1984).
[101] M. Sasaki, Gauge invariant scalar perturbations in the new inflationary universe, Prog. Theor. Phys. **70**, 394 (1983).
[102] L. Bieberbach, Über die Bewegungsgruppen der Euklidischen Räume, Erste Abhandlung, Math. Ann. **70**, 297 (1911).
[103] L. Bieberbach, Über die Bewegungsgruppen der Euklidischen Räume, Zweite Abhandlung. Die Gruppen mit einem endlichen Fundamentalbereich, Math. Ann. **72**, 400 (1912).
[104] J. D. Brown and K. V. Kuchar, Dust as a standard of space and time in canonical quantum gravity, Phys. Rev. D **51**, 5600 (1995).
[105] K. Giesel and T. Thiemann, Scalar material reference systems and loop quantum gravity, Classical Quantum Gravity **32**, 135015 (2015).
[106] O. Bratteli and D. W. Robinson, *Operator Algebras and Quantum Statistical Mechanics. Vol. 2: Equilibrium States. Models in Quantum Statistical Mechanics* (Springer-Verlag, Berlin, 1996).
[107] M. Reed and B. Simon, *Methods of Modern Mathematical Physics I: Functional Analysis* (Academic Press, Inc., New York, 1980).
[108] M. Reed and B. Simon, *Methods of Modern Mathematical Physics II: Fourier Analysis, Self-Adjointness* (Academic Press, Inc., New York, 1980).
[109] J. Cortez, G. A. Mena Marugán, and J. M. Velhinho, Quantum unitary dynamics in cosmological spacetimes, Ann. Phys. (Amsterdam) **363**, 36 (2015).

[110] B. Elizaga Navascués, G. A. Mena Marugán, and M. Martín-Benito, Fermions in hybrid loop quantum cosmology, Phys. Rev. D **96**, 044023 (2017).

[111] P. R. Chernoff and J. E. Marsden, *Properties of Infinite Dimensional Hamiltonian Systems* (Springer-Verlag, Berlin, 1974).

[112] H. Cartan, *Differential Forms* (Hermann, Houghton Mifflin Company, Boston, 1970).

[113] F. Finster and J. M. Isidro, L^p-Spectrum of the Schrödinger operator with inverted harmonic oscillator potential, J. Math. Phys. (N.Y.) **58**, 092104 (2017).