Dirac fields in loop quantum gravity and big bang nucleosynthesis

Martin Bojowald

Institute for Gravitation and the Cosmos, The Pennsylvania State University,
104 Davey Lab, University Park, Pennsylvania 16802, USA

Rupam Das

Department of Physics and Astronomy, Vanderbilt University, Nashville, Tennessee 37235
and Institute for Gravitation and the Cosmos, The Pennsylvania State University,
104 Davey Lab, University Park, Pennsylvania 16802, USA

Robert J. Scherrer

Department of Physics and Astronomy, Vanderbilt University, Nashville, Tennessee 37235, USA

Big bang nucleosynthesis requires a fine balance between equations of state for photons and relativistic fermions. Several corrections to equation of state parameters arise from classical and quantum physics, which are derived here from a canonical perspective. In particular, loop quantum gravity allows one to compute quantum gravity corrections for Maxwell and Dirac fields. Although the classical actions are very different, quantum corrections to the equation of state are remarkably similar. To lowest order, these corrections take the form of an overall expansion-dependent multiplicative factor in the total density. We use these results, along with the predictions of big bang nucleosynthesis, to place bounds on these corrections and especially the patch size of discrete quantum gravity states.

I. INTRODUCTION

Much of cosmology is well described by a space-time near a spatially isotropic Friedmann-Robertson-Walker model with line elements

$$ds^2 = -d\tau^2 + a(\tau)^2 \left(\frac{dr^2}{1-kr^2} + r^2(d\vartheta^2 + \sin^2\vartheta d\varphi^2) \right), \tag{1}$$

where $k = 0$ or ± 1, sourced by perfect fluids with equations of state $P = w\rho$. Such an equation of state relates the matter pressure P to its energy density ρ and captures the thermodynamical properties in a form relevant for isotropic space-times in general relativity. Often, one can assume the equation of state parameter w to be constant during successive phases of the evolving universe, with sharp jumps between different phases such as $w = -1$ during inflation, followed by $w = \frac{1}{3}$ during radiation domination and $w = 1$ during matter domination.

Observationally relevant details can depend on the precise values of w at a given stage, in particular, if one uses an effective value describing a mixture of different matter components. For instance, during big bang nucleosynthesis (BBN) one is in a radiation dominated phase mainly described by photons and relativistic fermions. Photons, according to Maxwell theory, have an exact equation of state parameter $w = \frac{1}{3}$ as a consequence of conformal invariance of the equations of motion (such that the stress-energy tensor is trace-free). For fermions the general equation of state is more complicated and nonlinear, but can in relativistic regimes be approximately given by the same value $w = \frac{1}{3}$ as for photons. In contrast to the case of Maxwell theory, however, there is no strict symmetry such as conformal invariance which would prevent w from taking a different value. It is one of the main objectives of the present paper to discuss possible corrections to this value.

For big bang nucleosynthesis, it turns out, the balance between fermions and photons is quite sensitive. In fact, different values for the equation of state parameters might even be preferred phenomenologically [1]. One possible reason for different equations of state could be different coupling constants of bosons and fermions to gravity, for which currently no underlying mechanism is known. In this paper we will explore the possibility of whether quantum gravitational corrections to the equations of state can produce sufficiently different values for the equation of state parameters. In fact, since the fields are governed by different actions, one generally expects different, though small, correction terms which can be of significance in a delicate balance. Note that we are not discussing ordinary quantum corrections of quantum fields on a classical background. Those are expected to be similar for fermions and radiation in relativistic regimes. We rather deal with quantum gravity corrections in the coupling of the fields to the space-time metric, about which much less is known *a priori*. Thus, different proposals of quantum gravity may differ at this stage, providing possible tests.

An approach where quantum gravitational corrections can be computed is loop quantum cosmology [2], which specializes loop quantum gravity [3–5] to cosmological regimes. In such a canonical quantization of gravity, equations of state must be computed from matter Hamiltonians rather than covariant stress-energy tensors. Quantum corrections to the underlying Hamiltonian then imply corrections in the equation of state. This program was carried out for the Maxwell Hamiltonian in [6], and is done here for Dirac fermions. There are several differences between the treatment of fermions and other fields, which from the gravitational point of view are mainly related to the fact that fermions, in a first-order formulation, also couple to torsion and not just the curvature of space-time. After describing the classical derivation of equations of state as well as steps of a loop quantization and its correction terms, we use big bang nucleosynthesis constraints to see how sensitively we can bound quantum gravity parameters.

We have aimed to make the paper nearly self-contained, and included some of the technical details. Section III on the canonical formulation of fermions, Sec. IV on quantum corrections from loop quantum gravity, and Sec. V on the analysis of big bang nucleosynthesis can, however, be read largely independently of each other by readers only interested in some of the aspects covered here. We will start with general remarks on the physics underlying the problem.

II. THE PHYSICAL SETTING

Big bang nucleosynthesis happens at energy scales $E_{\mathrm{BBN}} \sim \mathrm{MeV}$ which are large, but still tiny compared to the Planck energy M_{P}. Also, the universe has already grown large compared to the Planck length ℓ_{P} at this stage, and space-time curvature is small. One may thus question why quantum gravity should play any role. There is certainly a fine balance required for successful big bang nucleosynthesis, but the expected quantum gravity terms of the order E/M_{P}, obtained based on dimensional arguments, would have no effect.

However, dimensional arguments do not always work, in particular, if more than two parameters L_I of the same dimension, or any large dimensionless numbers are involved. Then, precise calculations have to be done to determine which geometric means $\prod_I L_I^{x_I}$ with $\sum_I x_I = 0$ may appear as coefficients, or which powers of dimensionless numbers occur as factors of correction terms. In loop quantum gravity, we are in such a situation: there is the macroscopic length scale L, which in our case we can take as the typical wave length of fields during nucleosynthesis, and also the Planck length $\ell_{\mathrm{P}} = \sqrt{G\hbar}$, which arises due to the presence of Newton's constant G and Planck's constant $\hbar$. In addition, there is a third and, in general, independent scale ℓ given by the microscopic size of elementary spatial patches in a quantum gravity state. This is a new feature of the fundamentally discrete theory, for which the precise state of quantum gravity plays an important role. Although ℓ must be proportional to the Planck length, its specific value for a given state can differ numerically. Then, a detailed calculation must show how L, ℓ, and ℓ_{P} appear in quantum gravity corrections and which numerical values may arise.

Alternatively, one can work with only two length scales, L and ℓ_{P}, but one has to deal with a large dimensionless parameter $\mathcal{N}$ given by the number of discrete patches of the underlying state in the volume considered, for instance, a volume of the size L^3 such that $\mathcal{N} = L^3/\ell^3$. Examples of cosmological phenomena are known where this does play a role for quantum gravity corrections [7,8], and here we analyze which features arise in the presence of fermions and especially for big bang nucleosynthesis.

There are precedents where such considerations have played important roles. Best known is the evidence for the atomic nature of matter derived by Einstein from the phenomenon of Brownian motion. Also here, there are several orders of magnitude between the expected size of molecules and the resolution of microscopes at that time. However, there is also a large number of molecules which by their sheer number can and do leave sizable effects on much larger suspended particles. There is, of course, never a guarantee that something analogous has to happen elsewhere. But this is to be checked by calculations and cannot always be ruled out based only on dimensional arguments.

We can use the result of corrections to the equation of state of radiation derived in [6], but a new analysis is required for fermions and their specific action. Even in relativistic regimes, the coupling of fermions to gravity differs from other fields, e.g. by torsion contributions which arise already from the kinetic term of the Dirac action. One could thus expect that quantum corrections for fermions differ from those to radiation and thus, by throwing off the balance during nucleosynthesis, possibly enhance the effect of quantum gravity corrections. Whether or not this happens cannot be decided without detailed calculations as they are reported and applied here.

III. CANONICAL FORMULATION OF DIRAC FERMIONS

For fermions, one has to use a tetrad e^I_μ rather than a space-time metric $g_{\mu\nu}$, which are related by $e^I_\mu e^I_\nu = g_{\mu\nu}$, in order to formulate an action with the appropriate covariant derivative of fermions. This naturally leads one to a first-order formalism of gravity in which the basic configuration variables are a connection one-form and the tetrad. In vacuum the connection would, as a consequence of field equations, be the torsion-free spin connection compatible with the tetrad. In the presence of matter fields which couple directly to the connection, such as fermions, this is no longer the case and there is torsion [9–11].

A. Nonminimally coupled Einstein-Dirac action

A canonical analysis of gravity minimally coupled to fermions is presented in detail in [11]. In [10], a non-minimally coupled action is proposed in order to eliminate violations of parity. Here, we first present a general canonical analysis providing a consistent canonical formulation for all coupling parameters. At the end of this section, we briefly comment on parity which plays a role in this context.

The basic configuration variables in a Lagrangian formulation of fermionic field theory are the Dirac bispinor $\Psi = (\psi, \eta)^T$ and its adjoint $\bar{\Psi} = (\Psi^*)^T \gamma^0$ with γ^α being the Minkowski Dirac matrices. The 2-component SL(2, $\mathbb{C}$) spinors ψ and η will be used later on in a Hamiltonian decomposition of the action. Being interested in applications to highly relativistic regimes, we only deal with massless fermions. Their nonminimal coupling to gravity can then be expressed, in the notation of [10], by the total action

$$S[e, \omega, \Psi] = S_G[e, \omega] + S_F[e, \omega, \Psi]$$
$$= \frac{1}{16\pi G} \int_M d^4x\, e\, e_I^\mu e_J^\nu P^{IJ}{}_{KL} F_{\mu\nu}{}^{KL}(\omega)$$
$$+ \frac{1}{2} i \int_M d^4x\, e \left[\bar{\Psi} \gamma^I e_I^\mu \left(1 - \frac{i}{\alpha} \gamma_5\right) \nabla_\mu \Psi \right.$$
$$\left. - \overline{\nabla_\mu \Psi} \left(1 - \frac{i}{\alpha} \gamma_5\right) \gamma^I e_I^\mu \Psi \right], \quad (2)$$

where α is the parameter for nonminimal coupling and minimal coupling is reproduced for $\alpha \to \infty$. The action is composed of the matter contribution S_F resulting from the fermion field and the gravitational contribution S_G expressed in terms of

$$P^{IJ}{}_{KL} = \delta_K^{[I} \delta_L^{J]} - \frac{1}{\gamma} \frac{\epsilon^{IJ}{}_{KL}}{2},$$
$$P^{-1}{}_{IJ}{}^{KL} = \frac{\gamma^2}{\gamma^2 + 1}\left(\delta_I^{[K} \delta_J^{L]} + \frac{1}{\gamma} \frac{\epsilon_{IJ}{}^{KL}}{2}\right) \quad (3)$$

where γ is the Barbero-Immirzi parameter [12,13].

The specific form of the gravitational action is the one given by Holst [14], formulated in terms of a tetrad field e_I^μ with inverse e_μ^I, whose determinant is e. (Thus, the space-time metric is $g_{\mu\nu} = e_\mu^I e_\nu^I$. For all space-time fields, $I, J, \ldots = 0, 1, 2, 3$ denote internal Lorentz indices and $\mu, \nu, \ldots = 0, 1, 2, 3$ space-time indices.) The Lorentz connection ω_μ^{IJ} in this formulation is an additional field independent of the triad, and $F_{\mu\nu}^{KL}(\omega) = 2\partial_{[\mu} \omega_{\nu]}^{IJ} + [\omega_\mu, \omega_\nu]^{IJ}$ is its curvature. It also determines the covariant derivative ∇_μ of Dirac spinors by

$$\nabla_\mu \equiv \partial_\mu + \frac{1}{4}\omega_\mu^{IJ}\gamma_{[I}\gamma_{J]}, \qquad [\nabla_\mu, \nabla_\nu] = \frac{1}{4}F_{\mu\nu}^{IJ}\gamma_{[I}\gamma_{J]} \quad (4)$$

in terms of Dirac matrices γ_I (which will always carry an index such that no confusion with the Barbero-Immirzi parameter γ can arise).

B. Dirac Hamiltonian

As usual, a Hamiltonian formalism of gravity requires a space-time foliation Σ_t: $t = $ const such that one can introduce fields and their rates of change, which will provide canonical variables. This is done by referring to a time function t as well as a time evolution vector field t^μ such that $t^\mu \nabla_\mu t = 1$. Rates of change of spatial fields will then be associated with their derivatives along t^a.

For convenience, one decomposes t^μ into normal and tangential parts with respect to Σ_t by defining the lapse function N and the shift vector N^a such that $t^\mu = N n^\mu + N^\mu$ with $N^\mu n_\mu = 0$. Here, n^μ is the unit normal vector field to the hypersurfaces Σ_t. The space-time metric $g_{\mu\nu}$ induces a spatial metric $q_{\mu\nu}(t)$ on Σ_t by the formula $g_{\mu\nu} = q_{\mu\nu} - n_\mu n_\nu$. This is one of the basic fields of a canonical formulation, and its momentum will be related to $\dot{q}_{ab}$ defined as the Lie derivative of q_{ab} along t^a. Since contractions of $q_{\mu\nu}$ and N^μ with the normal n^μ vanish, they give rise to spatial tensors q_{ab} and N^a.

In our case, we are using a tetrad formulation, where e_μ^I provides a map from the tangent space of space-time to an internal Minkowski space. The space-time foliation thus requires an associated space-time splitting of the Minkowski space. This takes the form of a partial gauge fixing on the internal vector fields of the tetrad: the directions (or rather boosts) of tetrad fields can no longer be chosen arbitrarily. Instead, we decompose the tetrad into a fixed internal unit timelike vector field and a triad on the space Σ_t. We choose the internal vector field to be constant, $n_I = -\delta_{I,0}$ with $n^I n_I = -1$. Then, we allow only those tetrads which are compatible with the fixed n^I in the sense that $n^a = n^I e_I^a$ is the unit normal to the given foliation. This implies that $e_I^a = \mathcal{E}_I^a - n^a n_I$ with $\mathcal{E}_I^a n_a = \mathcal{E}_I^a n^I = 0$ so that $\mathcal{E}_I^a$ is a spatial triad.

Now, using $e_I^a = \mathcal{E}_I^a - n^a n_I$ with $n_I = -\delta_{I,0}$ and $n^a = N^{-1}(t^a - N^a)$, we can decompose the nonminimally coupled Dirac action and write it in terms of spatial fields only:

$$S_{\text{Dirac}} = \frac{i}{2} \int_M d^4x\, N\sqrt{q}\left(\bar{\Psi}\gamma^I(\mathcal{E}_I^a + N^{-1}(t^a - N^a))\right.$$
$$\times \left(1 - \frac{i}{\alpha}\gamma_5\right)\nabla_a \Psi - \overline{\nabla_a \Psi}\left(1 - \frac{i}{\alpha}\gamma_5\right)$$
$$\left. \times \gamma^I(\mathcal{E}_I^a + N^{-1}(t^a - N^a))\Psi\right), \quad (5)$$

where the space-time determinant is factorized as $e = N\sqrt{q}$ with the determinant q of the spatial metric. These terms can be decomposed into several terms containing the SL(2, $\mathbb{C}$) spinors ψ and η instead of the Dirac spinor Ψ:

$$\frac{i}{2}N\sqrt{q}n^a(\bar{\Psi}\gamma^0\nabla_a\Psi - \text{c.c.}) + \frac{i}{2}N\mathcal{E}_i^a(\bar{\Psi}\gamma^i\nabla_a\Psi - \text{c.c.})$$

$$= \sqrt{q}\left(-\frac{1}{2}i(\psi^\dagger\dot{\psi} + \eta^\dagger\dot{\eta} - \text{c.c.}) + \frac{1}{4}\epsilon_{mnk}\omega_t{}^{mn}J^k\right) - \sqrt{q}N^a\left(-\frac{1}{2}i(\psi^\dagger\partial_a\psi + \eta^\dagger\partial_a\eta - \text{c.c.}) + \frac{1}{4}\epsilon_{mnk}\omega_a{}^{mn}J^k\right)$$

$$- N\sqrt{q}\mathcal{E}_i^a\left(-\frac{1}{2}i(-\psi^\dagger\sigma^i\partial_a\psi + \eta^\dagger\sigma^i\partial_a\eta - \text{c.c.}) - \left(\frac{1}{4}\epsilon^i{}_{mn}\omega_a{}^{mn}(\psi^\dagger\psi - \eta^\dagger\eta) + \frac{1}{2}\epsilon^i{}_{mn}\omega_a{}^{m0}J^n\right)\right), \tag{6}$$

and similarly an expansion of the terms involving the nonminimal coupling parameter α in the action (5) gives

$$\frac{1}{2\alpha}\sqrt{q}((t^a - N^a)(\bar{\Psi}\gamma^0\gamma_5\nabla_a\Psi - \overline{\nabla_a\Psi}\gamma_5\gamma^0\Psi) + N\mathcal{E}_i^a(\bar{\Psi}\gamma^i\gamma_5\nabla_a\Psi - \overline{\nabla_a\Psi}\gamma_5\gamma^i\Psi))$$

$$= \frac{1}{2\alpha}\sqrt{q}(t^a - N^a)(\psi^\dagger\partial_a\psi - \eta^\dagger\partial_a\eta + (\partial_a\psi^\dagger)\psi - (\partial_a\eta^\dagger)\eta - \omega_a{}^{j0}J^j)$$

$$+ \frac{N}{2\alpha}\sqrt{q}\mathcal{E}_i^a(\psi^\dagger\sigma^i\partial_a\psi + \eta^\dagger\sigma^i\partial_a\eta + (\partial_a\psi^\dagger)\sigma^i\psi + (\partial_a\eta^\dagger)\sigma^i\eta + \omega_a{}^{ik}J_k - \omega_a{}^{i0}J_0), \tag{7}$$

where c.c. denotes complex conjugation and we have introduced the fermion current $J^k := \psi^\dagger\sigma^k\psi + \eta^\dagger\sigma^k\eta$ and the time component $J_0 := \psi^\dagger\psi - \eta^\dagger\eta$. Details of this calculation as well as the corresponding canonical decompositions of the gravitational part can be found in [11].

We also refer to [11] for the definition of canonical gravitational variables as they appear also in the matter action. In terms of the connection components ω_μ^{IJ} in Eq. (4), we define

$$\Gamma_b^i := -\tfrac{1}{2}\epsilon^i{}_{jk}\omega_b^{jk}, \qquad K_b^i := -\omega_b{}^{i0} \tag{8}$$

and the Ashtekar-Barbero connection [12,15]

$$A_a^i := \Gamma_a^i + \gamma K_a^i = -\tfrac{1}{2}\epsilon_{kl}^i\omega_a{}^{kl} - \gamma\omega_a{}^{i0} \tag{9}$$

with the Barbero-Immirzi parameter γ [13]. This connection is important because it provides a convenient canonical structure with Poisson brackets

$$\{A_a^i(x), E_j^b(y)\} = 8\pi\gamma G\delta_a^b\delta_j^i\delta(x - y) \tag{10}$$

with the densitized triad $E_j^b := \sqrt{q}\mathcal{E}_j^b$. Moreover, following the steps presented in [11], it is straightforward to show that the torsion contribution to the spin connection Γ is given by

$$C_a^j := \frac{\gamma\kappa}{4(1 + \gamma^2)}\left(\left(1 - \frac{\gamma}{\alpha}\right)\epsilon^j{}_{kl}e_a^kJ^l - \left(\gamma + \frac{1}{\alpha}\right)e_a^jJ^0\right). \tag{11}$$

One can then show that $\Gamma_a^i = \tilde{\Gamma}_a^i + C_a^i$, with $\tilde{\Gamma}_a^i$ being the torsion-free spin connection compatible with the cotriad e_a^i.

For a fixed value of γ, the above equation for torsion reduces to that for minimal coupling as $\alpha \to \infty$. As noted in [10], based on a Lagrangian analysis, another interesting case is $\alpha = \gamma$, which reduces C_a^j to $-\frac{\kappa}{4}e_a^iJ^0$, making torsion γ independent. The nonminimal coupling is important for the behavior of the action under parity reversal, and it is the value $\alpha = \gamma$ which provides parity invariance of the combined system of gravity and fermions. This may seem surprising because neither vacuum gravity nor the minimal Dirac action on a fixed background violate parity, and thus the extra term of nonminimal coupling seems to introduce parity violation. As [10] in the Lagrangian picture and [11] in the Hamiltonian picture show, however, the torsion introduced by coupling fermions to gravity also introduces parity violation in the Holst action, which has to be canceled by the additional term of nonminimal coupling. (The action with minimal coupling was called "not fully consistent" in [10], while our analysis is consistent for any value of α. To avoid potential confusion, one should first note that [10] starts from the Einstein-Cartan action with minimal coupling, not directly from a Holst action as here. Parity properties are different in both cases, and the observation of [10] refers only to this behavior. The Holst action with minimal coupling would indeed be inconsistent with parity preservation, but it does not present an inconsistency of the overall framework unless one explicitly requires parity preservation. In fact, parity violation is expected at the onset of BBN due to weak interactions. Anyway, the following discussion in our paper can remain unchanged for different values of α and thus applies directly to the parity preserving case.)

Upon inserting (11) in (6) and (7), the action (5) takes the form

$$S_{\text{Dirac}}(P_i^a, \Gamma_a^i, \psi, \psi^\dagger, \eta, \eta^\dagger) = -\int dt \int_{\Sigma_t} d^3x \sqrt{q} \Bigg(\Big(i(\theta_L(\psi^\dagger \dot{\psi} - \dot{\eta}^\dagger \eta) - \theta_R(\dot{\psi}^\dagger \psi - \eta^\dagger \dot{\eta})) - \frac{1}{4}\epsilon_{mnk}\omega_t^{\,mn}J^k + \frac{1}{2\alpha}\omega_t^{\,j0}J_i \Big)$$

$$- N^a \Big(i(\theta_L(\psi^\dagger \mathcal{D}_a \psi - \overline{\mathcal{D}_a \eta}\,\eta) + \theta_R(\eta^\dagger \mathcal{D}_a \eta - \overline{\mathcal{D}_a \psi}\,\psi)) - \frac{1}{2}\Big(\gamma + \frac{1}{\alpha}\Big)K_a^i \sqrt{q} J_i \Big)$$

$$+ N\Big(E_i^a(i\theta_L(\psi^\dagger \sigma^i \mathcal{D}_a \psi + \overline{\mathcal{D}_a \eta}\,\sigma^i \eta) - i\theta_R(\eta^\dagger \sigma^i \mathcal{D}_a \eta + \overline{\mathcal{D}_a \psi}\,\sigma^i \psi))$$

$$- \frac{1}{2}\Big(\gamma + \frac{1}{\alpha}\Big)E_i^a K_a^i J^0 + \frac{1}{2}\epsilon_{lkn}K_a^l E_k^a\Big(1 - \frac{\gamma}{\alpha}\Big)J^n \Big) \Bigg), \tag{12}$$

where $\theta_L := \frac{1}{2}(1 + \frac{i}{\alpha})$, $\theta_R := \frac{1}{2}(1 - \frac{i}{\alpha})$, and we have used the covariant derivatives, $\mathcal{D}_a = \partial_a + A_a^l \tau_l$, related to the Ashtekar-Barbero connection, and $\tau_l = -\frac{1}{2}i\sigma_l$. The first part of the action in this form shows that momenta of the fermion field are $\pi_\xi = -i\xi^\dagger$ of $\xi = \sqrt[4]{q}\psi$ and $\pi_\chi = -i\chi^\dagger$ of $\chi = \sqrt[4]{q}\eta$. Here, we are using half-densitized spinor fields to avoid complex-valued canonical variables [16].

Moreover, the action provides the Hamiltonian $H = \int d^3x N(x)\delta S/\delta N$. Together with fermion-dependent terms resulting from the gravitational action, this provides derivative terms and self-interaction terms in the Hamiltonian,

$$H_{\text{Dirac}} = \int_{\Sigma_t} d^3x N\Bigg(-\frac{\beta E_i^a}{\sqrt{q}}\mathcal{D}_a(\pi_\xi^T \tau^i \xi + \pi_\chi^T \tau^i \chi) - i\frac{2E_a^i}{\sqrt{q}}(\theta_L \pi_\xi^T \tau^i \mathcal{D}_a \xi - \theta_R \pi_\chi^T \tau^i \mathcal{D}_a \chi - \text{c.c.})$$

$$+ \frac{\gamma\kappa\beta}{2\sqrt{q}(1 + \gamma^2)}\Big(3 - \frac{\gamma}{\alpha} + 2\gamma^2\Big)(\pi_\xi^T \tau_l \xi + \pi_\chi^T \tau_l \chi)(\pi_\xi^T \tau^l \xi + \pi_\chi^T \tau^l \chi) + \frac{3\gamma\kappa\theta}{8\alpha\sqrt{q}}(\pi_\xi^T \xi - \pi_\chi^T \chi)(\pi_\xi^T \xi - \pi_\chi^T \chi) \Bigg) \tag{13}$$

with $\beta := \gamma + \frac{1}{\alpha}$ and $\theta := 1 - \frac{\gamma}{\alpha}$. The top line of this expression is the most important one because its derivative terms are dominant in relativistic regimes. In addition to those, we highlight the presence of four-fermion interactions in the second line, which we summarize as

$$B := \frac{\gamma\kappa\beta}{2(1 + \gamma^2)}\Big(3 - \frac{\gamma}{\alpha} + 2\gamma^2\Big)(\pi_\xi^T \tau_l \xi + \pi_\chi^T \tau_l \chi)$$

$$\times (\pi_\xi^T \tau^l \xi + \pi_\chi^T \tau^l \chi) + \frac{3\gamma\kappa\theta}{8\alpha}(\pi_\xi^T \xi - \pi_\chi^T \chi)$$

$$\times (\pi_\xi^T \xi - \pi_\chi^T \chi) \tag{14}$$

multiplying $q^{-1/2}$.

C. Equation of state

From the Hamiltonian we can determine energy and pressure and formulate the equation of state. The matter Hamiltonian is directly related to energy density by

$$\rho = \frac{1}{\sqrt{q}}\frac{\delta H_{\text{Dirac}}}{\delta N} \tag{15}$$

and thus, from (13), the energy density is

$$\rho = \frac{2E_i^a}{q}\Bigg(-\frac{\beta}{2}\mathcal{D}_a(\pi_\xi^T \tau^i \xi + \pi_\chi^T \tau^i \chi)$$

$$+ i(-\theta_L \pi_\xi^T \tau^i \mathcal{D}_a \xi + \theta_R \pi_\chi^T \tau^i \mathcal{D}_a \chi - \text{c.c.}) \Bigg) + \frac{B}{q}. \tag{16}$$

The canonical formula for pressure is

$$P = -\frac{2}{3N\sqrt{q}}E_i^a \frac{\delta H_{\text{Dirac}}}{\delta E_i^a} \tag{17}$$

as shown by a straightforward adaptation of the calculation done in [6] for metric variables. Now using the functional derivative

$$\frac{\delta\sqrt{q(x)}}{\delta E_i^a(y)} = \frac{1}{2}e_a^i\delta(x - y), \tag{18}$$

and thus

$$\frac{\delta}{\delta E_j^b(y)}\Big(\frac{2E_i^a(x)}{\sqrt{q(x)}}\Big) = \frac{1}{\sqrt{q}}(2\delta_b^a\delta_i^j - e_i^a e_b^j)\delta(x - y), \tag{19}$$

and inserting (19) in (17), we obtain the pressure

$$P = \frac{2E_i^a}{3q}\Bigg(-\frac{\beta}{2}\mathcal{D}_a(\pi_\xi^T \tau^i \xi + \pi_\chi^T \tau^i \chi) + i(-\theta_L \pi_\xi^T \tau^i \mathcal{D}_a \xi$$

$$+ \theta_R \pi_\chi^T \tau^i \mathcal{D}_a \chi - \text{c.c.}) \Bigg) + \frac{B}{q}. \tag{20}$$

This results in an equation of state

$$w_{\text{Dirac}} = \frac{P}{\rho} = \frac{1}{3} - \frac{2B}{3\rho}. \tag{21}$$

In relativistic regimes, the kinetic term involving partial derivatives ∂_a contained in $\mathcal{D}_a$ is dominant, which leaves us with an equation of state

$$w = \frac{P}{\rho} = \frac{1}{3} + \epsilon \tag{22}$$

whose leading term agrees with the parameter for a

Maxwell field. But there are clearly correction terms for fermions already in the classical first-order theory. They do not arise for the Maxwell field, implying a difference in the coupling to gravity due to torsion, which is present even in relativistic regimes. The order of magnitude of the additional term depends on the fermion current density and is thus not expected to be large unless regimes are very dense. We will not consider this correction further in this article, but highlight its role as a consequence of torsion.

In addition, the canonical analysis performed here provides the stage for a quantization of the theory, resulting in further correction terms from quantum gravity as they occur even for the Maxwell field [6]. Their magnitudes must also be extracted and then compared for the different fields.

IV. QUANTUM CORRECTIONS

We are now in a position to derive the form of quantum gravity corrections for fermion fields in an explicit scenario of quantum gravity, completing the program of Sec. II. We will not require the quantized dynamics of gravity, which makes our aim feasible. But we do need to know how the triad is quantized because it appears in the Dirac Hamiltonian. It appears as a basic variable in the Poisson relations (10), together with the Ashtekar-Barbero connection A_a^i. Loop quantum gravity [3–5] proceeds by forming a quantum representation of these basic fields which turns holonomies $h_e(A) = \mathcal{P} \exp(\int_e \mathrm{d}t \dot{e}^a A_a^i \tau_i)$ along arbitrary spatial curves e into creation operators of geometrical excitations along these curves. Many such creations then excite a discrete state based on the graph formed by all curves used. On such a state, E_i^a is quantized through the fluxes $F_S(E) = \int_S \mathrm{d}^2 y E_i^a n_a \tau^i$ through surfaces S with conormal n_a. Resulting operators have eigenvalues obtained by counting intersections of the surface and the graph in the state, weighed by multiplicities of overlapping curves.

This structure is important for a background-independent representation which replaces the usual Fock representation of quantum fields on a curved background space-time. The Fock representation requires a metric to define basic concepts such as a vacuum state or creation and annihilation operators, which can thus not be used for the nonperturbative gravitational field itself. This mathematical property has important physical implications because extended objects, namely, the integrated holonomies and fluxes, are used as basic quantities. In this way, there is nonlocality and other implications which affect quantum gravity corrections.

In loop quantum gravity, there are three main effects which imply correction terms in effective matter equations. Any Hamiltonian contains inverse powers of triad components such as $1/\sqrt{q}$ in (13), which for fermions is a consequence of the fact that they are quantized through half-densities [16]. The loop quantization, however, leads to triad operators which have discrete spectra containing zero,

thus lacking inverse operators. A proper quantization, along the lines of [17], does give a well-defined operator with the correct semiclassical limit. But there are deviations from the classical behavior on small length scales, which are the first source of correction terms. As in the case of the Maxwell field [6], this is the main effect we include here.

In addition, there are qualitatively different correction terms. First, loop quantum gravity is spatially discrete, with states supported on spatial graphs. Quantizations of Hamiltonians thus lead to a discrete representation of any spatial derivative term as they also occur for fermions. The classical expression arises in a continuum limit, but for any given state the discrete representation implies corrections to the classical derivatives as the leading terms in an expansion. Second, the connection is quantized through holonomies rather than its single components. Thus, the quantum Hamiltonians are formulated in terms of exponentials of line integrals of the connection, which also give the leading classical term plus corrections in an expansion. Finally, whenever a Hamiltonian is not quadratic, there are genuine quantum effects occurring in typical low energy effective actions. They can be computed in a Hamiltonian formulation as well [18,19], contributing yet another source of corrections.

One certainly needs to know the relative magnitude of all corrections in order to see which ones have to be taken into account. For all of them, the magnitude depends on details of the quantum state describing the regime. Here, properties of states have to be taken into account, and dimensional arguments are no longer sufficient. For instance, discretization and curvature corrections depend on the patch size occurring in the discrete state underlying quantum gravity. This patch size is typically small compared to scales on which the matter field changes, even in relativistic regimes assumed here. Thus, such corrections can be ignored in a first approximation. What remains are corrections from inverse powers. While other corrections shrink in the continuum limit where the patch size becomes small, inverse corrections actually grow when the patch size approaches the Planck length. The regimes where the two classes of corrections are dominant are thus neatly separated, and we can safely focus on inverse triad corrections only. The relevant formulas are collected in the Appendix; see also [20]. A detailed and complete derivation is not yet available since precise properties of a quantum gravity state would be required. Still, many general qualitative insights can be gained in this way.

In the Dirac Hamiltonian (13) the factor to be quantized containing inverse powers of the densitized triad is

$$\frac{2E_i^a}{\ell_0 \sqrt{q}} = \epsilon^{abc} \epsilon_{ijk} \frac{e_b^j e_c^k}{\ell_0 \sqrt{q}} \approx \epsilon^{abc} \epsilon_{ijk} \frac{\ell_0^2 e_b^j e_c^k}{V_\nu}$$

in terms of the volume $V_\nu \approx \ell_0^3 \sqrt{q(\nu)}$ of one discrete patch at a point ν. We can already notice the close resemblance

to the Maxwell Hamiltonian, where the corresponding expression is $q_{ab}/\ell_0\sqrt{q} = e_a^i e_b^i/\ell_0\sqrt{q}$ which differs only by the additional ϵ tensors. This close relation will, in the end, lead to quite similar quantum corrections for photons and fermions.

We proceed using (A6) for $r = 1/2$, and write

$$\frac{2E_i^a}{\ell_0\sqrt{q}} = \left(\frac{\ell_0}{2\pi G\gamma}\right)^2 \epsilon^{abc}\epsilon_{ijk}\{A_b^j, V_v^{1/2}\}\{A_c^k, V_v^{1/2}\}, \quad (23)$$

which can then be quantized by turning Poisson brackets into commutators of operators. This results in

$$\widehat{\frac{2E_i^a}{\ell_0\sqrt{q}}} = \epsilon^{KIJ}\epsilon_{ijk}(\ell_0\widehat{V_v^{-1/2}e_I^j})(\ell_0\widehat{V_v^{-1/2}e_J^k})$$

$$= \epsilon^{KIJ}\epsilon_{ijk}\hat{C}_{v,I}^{(1/2)}\hat{C}_{v,J}^{(1/2)}\delta_{(I)}^j\delta_{(J)}^k \quad (24)$$

with $\hat{C}_{v,I}^{1/2}$ defined in (A7). As explicit calculations in the Appendix show, eigenvalues (A9) of the operator $\hat{C}_{v,I}^{(1/2)}$, or its expectation values in semiclassical states, do not follow precisely the behavior expected for the classical function it quantizes. Deviations become larger for small values of the triad components, which can be captured in a correction function $\alpha[p^I(v)]$ multiplying the classical expression. In particular, such a correction function will appear in a Hamiltonian operator, thus also correcting expressions for energy density and pressure or the equation of state. The functional form follows from Eqs. (A9) and (A10).

Thus the general expression one can expect for a phenomenological Dirac Hamiltonian including corrections from inverse powers of the triad is

$$H_{\text{phen}} = \int_{\Sigma_t} d^3x N\left(\frac{E_i^a}{\sqrt{q}}\alpha(E_j^b)(-\beta\mathcal{D}_a(\pi_\xi^T\tau^i\xi + \pi_\chi^T\tau^i\chi)\right.$$

$$- 2i(\theta_L\pi_\xi^T\tau^i\mathcal{D}_a\xi - \theta_R\pi_\chi^T\tau^i\mathcal{D}_a\chi - \text{c.c.}))$$

$$\left. + \frac{\sigma(E_j^b)}{\sqrt{q}}B\right) \quad (25)$$

with two possibly different correction functions α and θ. This also affects the energy density and pressure terms, derived by the general expressions (15) and (17). We are mainly interested in the correction to the factor of one-third in the equation of state (21), so we focus on the first term in (25) in what follows. Moreover, for a *nearly* [21] isotropic background geometry, α only depends on the determinant q of the spatial metric and thus $q^{ab}\delta\alpha/\delta q^{ab} = -3q\,d\alpha/dq = -\frac{1}{2}a\,d\alpha/da$ with the scale factor a related to q by $q = \det(q_{ab}) = a^6$. In this case the quantum gravitational expectation for $\alpha(q)$, as per Eqs. (A9) and (A10), simplifies. To use these expressions, we have to relate the scale factor to quantum gravitational excitation levels as they occur in calculations of loop quantum gravity. In the notation of the Appendix, an elementary discrete patch in a nearly isotropic space-time has, on the one hand, an area of $\ell_0^2a^2$ if ℓ_0 is the coordinate diameter of the patch. This can be expressed as $\ell_0^2a^2 = (V_\gamma/\mathcal{N}_\gamma)^{2/3}$ where $\mathcal{N}_\gamma$ is the

number of patches in a box $\mathcal{V}$ of volume V_γ. On the other hand, using (A2) the quantum gravity state assigns a value of $4\pi\gamma\ell_P^2\mu_v$ to this patch via the flux operator, where μ_v is the quantum number of the geometrical excitation of this patch. Thus, we obtain

$$\mu_v = \frac{V_\gamma^{2/3}}{4\pi\gamma\ell_P^2\mathcal{N}_\gamma^{2/3}} =: \frac{a^2}{a_{\text{disc}}^2}$$

where

$$a_{\text{disc}} = 2\sqrt{\pi\gamma}\ell_P\left(\frac{\mathcal{N}_\gamma}{V_0}\right)^{1/3} \quad (26)$$

with the coordinate volume V_0 of the box $\mathcal{V}$. The numerical value of a_{disc} depends on coordinates via V_0, or on the normalization of the scale factor. (It does not depend on the choice of the box $\mathcal{V}$ because a change would multiply $\mathcal{N}_\gamma$ and V_0 by the same factor.) But it is important to note that a_{disc} is not just determined by the Planck length ℓ_P, which appears for dimensional reasons, but also depends on the large number $\mathcal{N}_\gamma$ of discrete patches per volume as given by the quantum gravity state. This is a parameter exactly as expected in the discussion of Sec. II. Replacing μ_v in the equations of the Appendix, we obtain

$$\alpha(a) = 8\sqrt{2}(a/a_{\text{disc}})^2((2(a/a_{\text{disc}})^2 + 1)^{1/4}$$

$$- |2(a/a_{\text{disc}})^2 - 1|^{1/4})^2 \quad (27)$$

where a_{disc} appears, influencing the size of quantum gravity corrections.

The function is plotted in Fig. 1. One can easily see that $\alpha(a)$ approaches the classical value $\alpha = 1$ for $a \gg a_{\text{disc}}/\sqrt{2}$, while it differs from 1 for small a. For $a > a_{\text{disc}}/\sqrt{2}$, the corrections are perturbative in a^{-1},
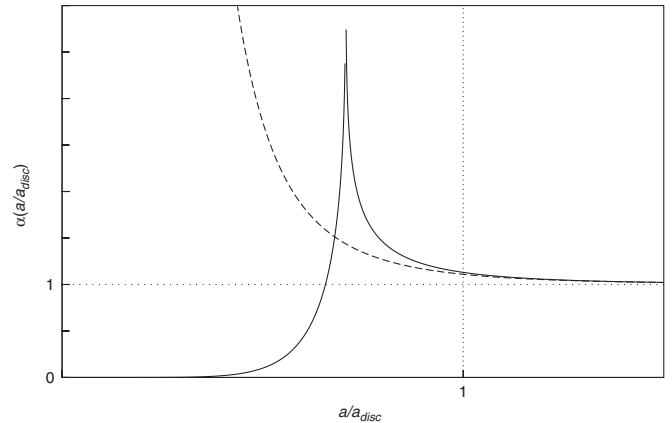

FIG. 1.   The correction function (27) as a function of the scale factor (solid line). The asymptotic form (28) for large a is shown by the dashed line. [The sharp cusp, a consequence of the absolute value appearing in (27), is present only for eigenvalues as plotted here, but would disappear for expectation values of the inverse volume operator in coherent states. This cusp will play no role in the analysis of this paper.]

$$\alpha(a) \sim 1 + \frac{7}{64}\left(\frac{a_{\text{disc}}}{a}\right)^4 + \cdots. \qquad (28)$$

This is the first correction in an asymptotic expansion for eigenvalues. If semiclassical states rather than volume eigenstates are used, powers of a^{-1} in the leading corrections can be smaller. Moreover, via $\mathcal{N}_\gamma$ the discreteness scale a_{disc} is expected to be not precisely constant but a function of a itself because the underlying spatial discreteness of quantum gravity can be refined dynamically during cosmological evolution [22,23]. (Indeed, dynamical refinement is also required for several other phenomenological reasons [24–28].) In our following analysis we will thus assume a functional form

$$\alpha(a) = 1 + c(a/a_0)^{-n} \qquad (29)$$

where we traded the fundamental normalization by a_{disc} for normalization with respect to the present-day value of the scale factor a_0. From the derivation, n is likely to be a small, even integer and c is known to be positive. The constant c depends on a_{disc} and inherits the $\mathcal{N}_\gamma$ factor. It can thus be larger than of order 1. We will treat this parameter as phenomenological, and in the end, formulate bounds on c as bounds for $\mathcal{N}_\gamma$.

Energy density and the pressure then are, ignoring the classical interaction term B,

$$\rho_{\text{eff}} = \frac{2E_i^a}{a^6}\alpha(a)(-\beta\mathcal{D}_a(\pi_\xi^T\tau^i\xi + \pi_\chi^T\tau^i\chi)$$
$$+ i(-\theta_L\pi_\xi^T\tau^i\mathcal{D}_a\xi - \theta_R\pi_\chi^T\tau^i\mathcal{D}_a\chi - \text{c.c.})) \quad (30)$$

and

$$3P_{\text{eff}} = \frac{2E_i^a}{a^6}\alpha(a)\left(1 - \frac{\mathrm{d}\log\alpha}{\mathrm{d}\log a}\right)(-\beta\mathcal{D}_a(\pi_\xi^T\tau^i\xi + \pi_\chi^T\tau^i\chi)$$
$$+ i(-\theta_L\pi_\xi^T\tau^i\mathcal{D}_a\xi - \theta_R\pi_\chi^T\tau^i\mathcal{D}_a\chi - \text{c.c.})). \quad (31)$$

From this, the equation of state w can easily be computed:

$$w_{\text{eff}} = \frac{1}{3}\left(1 - \frac{\mathrm{d}\log\alpha}{\mathrm{d}\log a}\right). \qquad (32)$$

This quantum gravity correction is independent of the specific matter dynamics as in the classical relativistic case. It results in an equation of state which is linear in ρ, but depends on the geometrical scales (and the Planck length) through α. This is the same general formula derived in [6] for radiation. Thus, on an isotropic background, radiation and relativistic fermions are not distinguished by the form of quantum corrections they receive.

With the equation of state and the continuity equation

$$\dot{\rho} + 3\frac{\dot{a}}{a}(\rho + P) = 0, \qquad (33)$$

where a is the scale factor and the dot indicates a proper time derivative, we can determine energy density as a function of the universe size. We first obtain

$$\frac{\mathrm{d}\log\rho(a)}{\mathrm{d}\log a} = -3(1 + w(a)), \qquad (34)$$

explicitly showing the dependence of the equation of state on the scale factor. The solution is

$$\rho(a) = \rho_0 \exp\left[-3\int(1 + w(a))\mathrm{d}\log a\right], \qquad (35)$$

where ρ_0 is the integration constant. Inserting an equation of state of the form (32), we obtain

$$\rho(a) = \rho_0\alpha(a)a^{-4}. \qquad (36)$$

For $\alpha = 1$, we retrieve the classical result $\rho(a) \propto a^{-4}$, but for $\alpha \neq 1$ loop quantum gravity corrections induced by the discreteness of flux operators are reflected in the evolution of energy density in a Friedmann-Robertson-Walker universe.

V. EFFECT ON BIG BANG NUCLEOSYNTHESIS

The production of elements in the early universe is highly sensitive to the expansion rate, given by

$$\frac{\dot{a}}{a} = \left(\frac{8}{3}\pi G\rho\right)^{1/2}, \qquad (37)$$

where ρ is the total density, thus including radiation and fermions. As we have seen here for fermions and in [6] for radiation, the effect of loop quantum gravity corrections is to multiply the effective $\rho(a)$ by a factor $\alpha(a)$. Most importantly, we find that $\alpha(a)$ is the *same* for both bosons and fermions (up to possible quantization ambiguities), so a separate treatment of the two types of particles in the early universe (as in Ref. [1]) is unnecessary here.

In the standard treatment of the thermal history of the universe, the density of relativistic particles (bosons or fermions) is given by

$$\rho = \frac{\pi^2}{30}g_*T^4, \qquad (38)$$

where g_* is the number of spin degrees of freedom for bosons and 7/8 times the number of spin degrees of freedom for fermions, and T is the temperature, which scales as

$$T \propto a^{-1}. \qquad (39)$$

The equation of state parameter is

$$w = 1/3. \qquad (40)$$

Clearly, Eqs. (38)–(40) are inconsistent with Eqs. (32) and (36). There is some ambiguity in determining the correct way to modify the expressions for $\rho(T)$ and w. We have chosen to assume that the modifications are contained in the gravitational sector, so that the density is given by

$$\rho = \alpha(a)\frac{\pi^2}{30}g_*T^4, \qquad (41)$$

with the temperature scaling as in Eq. (39), and the equa-

tion of state w is given by Eq. (32). This guarantees that the standard continuity equation (33) continues to hold. Note that this is not the only way to incorporate Eq. (36) into the calculation, but it seems to us the most reasonable way. This issue requires a consideration of thermodynamics on a quantum space-time, which is a fascinating but not well-studied area. Instead of entering details here, we note that we interpret the α correction as a consequence of a quantum gravity sink to energy and entropy. Thus, quantum gravity implies a nonequilibrium situation which would otherwise imply that ρ must be proportional to T^4 without any additional dependence on $a \propto T^{-1}$.

With these assumptions, we can simply treat $\alpha(a)$ as an effective multiplicative change in the overall value of G. Note that this simplification is only possible because we explicitly derived by our canonical analysis that, unexpectedly, quantum corrections of radiation and fermions appear in similar forms. This makes possible a comprehensive derivation of implications for BBN, bearing on earlier work. In fact, a great deal of work has been done on the use of BBN to constrain changes in G (see, e.g., Refs. [29–33]). The calculation is straightforward, if one has a functional form for the time variation in G. For the loop quantum gravity corrections considered here, the most reasonable functional form is (29). Note that this expression is by construction valid only in the limit where $\alpha(a) - 1 \ll 1$. In terms of the effective gravitational constant, G, one can then write

$$G(a) = G_0[1 + c(a/a_0)^{-n}], \qquad (42)$$

where G_0 is the present-day value of the gravitational constant.

In order to constrain the values of c and n, we calculate the predicted element abundances with the indicated change in G and compare with observational constraints. Big bang nucleosynthesis proceeds first through the weak interactions that interconvert protons and neutrons:

$$n + \nu_e \leftrightarrow p + e^-, \qquad n + e^+ \leftrightarrow p + \bar{\nu}_e,$$
$$n \leftrightarrow p + e^- + \bar{\nu}_e. \qquad (43)$$
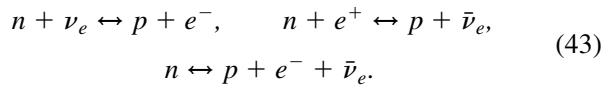
When $T \gtrsim 1$ MeV, the weak-interaction rates are faster than the expansion rate, $\dot{a}/a$, and the neutron-to-proton ratio (n/p) tracks its equilibrium value $\exp[-\Delta m/T]$, where Δm is the neutron-proton mass difference. As the universe expands and cools, the expansion rate becomes too fast for weak interactions to maintain weak equilibrium and n/p freezes out. Nearly all the neutrons which survive this freeze-out are converted into ^{4}He as soon as deuterium becomes stable against photodisintegration, but trace amounts of other elements are produced, particularly deuterium and ^{7}Li (see, e.g., Ref. [34] for a review).

In the standard model, the predicted abundances of all of these light elements are a function of the baryon-photon ratio, η, but any change in G alters these predictions. Prior to the era of precision cosmic microwave background

(CMB) observations (i.e., before WMAP), big bang nucleosynthesis provided the most stringent constraints on η, and modifications to the standard model could be ruled out only if no value of η gave predictions for the light element abundances consistent with the observations. However, the CMB observations now provide an independent estimate for η, which can be used as an input parameter for big bang nucleosynthesis calculations.

Copi *et al.* [32] have recently argued that the most reliable constraints on changes in G can be derived by using the WMAP values for η in conjunction with deuterium observations. The reason is that deuterium can be observed in (presumably unprocessed) high-redshift quasistellar object (QSO) absorption line systems (see Ref. [35] and references therein), while the estimated primordial ^{4}He abundance, derived from observations of low metallicity HII regions, is more uncertain (see, for example, the discussion in Ref. [36]). While we agree with the argument of Copi *et al.* in principle, for the particular model under consideration here it makes more sense to use limits on ^{4}He than on deuterium, in conjunction with the WMAP value for η. The reason is that the ^{4}He abundance is most sensitive to changes in the expansion rate at $T \sim 1$ MeV, when the freeze-out of the weak interactions determines the fraction of neutrons that will eventually be incorporated into ^{4}He. Deuterium, in contrast, is produced in big bang nucleosynthesis only because the expansion of the universe prevents all of the deuterium from being fused into heavier elements. Thus, the deuterium abundance is most sensitive to the expansion rate at the epoch when this fusion process operates ($T \sim 0.1$ MeV). The importance of this distinction with regard to modifications of the standard model was first noted in Ref. [37], and a very nice quantitative analysis was given recently in Ref. [38]. Note that our estimate for the behavior of $G(a)/G_0 - 1$, Eq. (42), is a steeply decreasing function of a. Thus, the change in the primordial ^{4}He abundance will always be much larger than the change in the deuterium abundance. Therefore, we can obtain better constraints on this model by using extremely conservative limits on ^{4}He, rather than by using the more reliable limits on the deuterium abundance. For the same reason, we can ignore any effect on the CMB, since the latter is generated at a much larger value of a, and any change will be minuscule. Hence, we can confidently use the WMAP value for η.

WMAP gives [39]

$$\eta = 6.116^{+0.197}_{-0.249} \times 10^{-10}. \qquad (44)$$

Because the estimated errors on η are so small, we simply use the central value for η; the bounds we derive on c in Eq. (29) change only slightly when η is varied within the range given by Eq. (44). Since c in Eqs. (29) and (42) is thought to be positive, the effect of loop quantum gravity corrections is to increase the primordial expansion rate,

which increases the predicted ^{4}He abundance. We therefore require an observational upper bound on the primordial ^{4}He abundance. As noted earlier, this is a matter of some controversy. We therefore adopt the very conservative upper bound recommended by Olive and Skillman [36]:

$$Y_P \leq 0.258, \tag{45}$$

where Y_P is the primordial mass fraction of ^{4}He. For a fixed value of n in Eq. (42), we determine the largest value of c that yields a primordial ^{4}He abundance consistent with this upper limit on Y_P. Since we are essentially bounding the change in G at $a/a_0 \sim 10^{-10}$, it is convenient to rewrite Eq. (29) as

$$\alpha = 1 + \tilde{c}/a_{10}^n \tag{46}$$

where $a_{10} \equiv 10^{10}(a/a_0)$. This upper bound on $\tilde{c}$ as a function of n is given in Fig. 2. For the special case $n = 4$, we can use these results to place a bound on a_{disc} in Eq. (28). We obtain

$$\frac{a_{\text{disc}}}{a_0} < 2.4 \times 10^{-10}. \tag{47}$$

This is not a strong bound for the parameters of quantum gravity, but clearly demonstrates that quantum corrections are consistent with successful big bang nucleosynthesis.

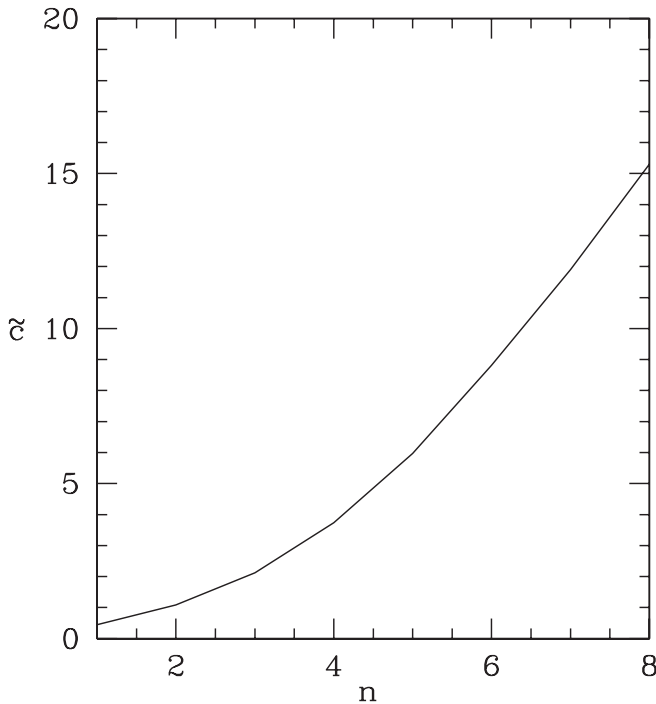In terms of more tangible quantum gravity parameters, we have

FIG. 2. The solid curve gives an upper bound on $\tilde{c}$ as a function of n, for the assumed form for α: $\alpha = 1 + \tilde{c}/a_{10}^n$, where a_{10} is the value of the scale factor in units for which $a_{10} = 10^{10}$ at present.

$$\mathcal{N}_\gamma^{1/3} < \frac{1.2 \times 10^{-10}}{\sqrt{\pi\gamma}} \frac{a_0 V_0^{1/3}}{\ell_P} \tag{48}$$

for the number of patches at the time of big bang nucleosynthesis. In terms of the volume $V_\gamma = (10^{-10}a_0)^3 V_0$ at this time, we have $\mathcal{N}_\gamma < 3V_\gamma/\ell_P^3$ with the value $\gamma \approx 0.24$ of the Barbero-Immirzi parameter as derived from black hole entropy calculations. More meaningfully, if we view $2\sqrt{\pi\gamma}\ell_P$ as the basic length scale as it appears in the spectrum (A2) of loop quantum gravity, the bound becomes more interesting: This gives $\mathcal{N}_\gamma^{1/3} < 2.4V_\gamma^{1/3}/(2\sqrt{\pi\gamma}\ell_P)$. This upper limit is already quite close to what one expects for elementary patch sizes in loop quantum gravity, which would provide $\mathcal{N}_\gamma^{1/3} < V_\gamma^{1/3}/(2\sqrt{\pi\gamma}\ell_P)$ as a fundamental upper limit. Given that these values are close to each other, we see a clear potential of improvements by more precise observational inputs. Moreover, other correction terms from quantum gravity could be used to obtain a lower bound for $\mathcal{N}_\gamma$ such that the allowed window would be reduced to a smaller size.

VI. CONCLUSIONS

Big bang nucleosynthesis is a highly relativistic regime which, to a good approximation, implies identical equations of state for fermions and photons. There are, however, corrections to the simple equation of state $w = \frac{1}{3}$ for fermions, even classically. One observation made here is that the interaction term derived in [40] leads to such a correction and might be more constrained by nucleosynthesis than through standard particle experiments [41]. We have not analyzed this further here because more details of the behavior of the fermion current would be required.

A second source of corrections arises from quantum gravity. Remarkably, while quantum gravity effects on an isotropic background do correct the equations of state, they do so equally for photons and relativistic fermions. Initially, this is not expected for both types of fields due to their very different actions. Thus, quantum gravity effects do not spoil the detailed balance required for the scenario to work, and bounds from big bang nucleosynthesis obtained so far are not strong. But there are interesting limits for the primary parameter, the patch size of a quantum gravity state. It is dimensionally expected to be proportional to the Planck length ℓ_P but could be larger. In fact, current bounds derived here already rule out a patch size of exactly the elementary allowed value in loop quantum gravity. With more precise estimates, these bounds can be improved further.

We have made use of quantum gravity corrections in a form which does not distinguish fermions from radiation. Although the most natural implementation, quite unexpectedly, provides equal corrections as shown here, there are several possibilities for differences which suggest several

further investigations. Small deviations in the equations of state and thus energy densities of fermions and radiation are possible. First, there are always quantization ambiguities, and so far we tacitly assumed that the same basic quantization choice is made for the Maxwell and Dirac Hamiltonians. Such ambiguity parameters can be explicitly included in specific formulas for correction functions; see e.g. [20,42,43]. Independent consistency conditions for the quantization may at some point require one to use different quantizations for both types of fields, resulting in different quantum corrections and different energy densities. Such conditions can be derived from an analysis of anomaly freedom of the Maxwell field and fermions coupled to gravity, which is currently in progress. As shown here, if this is the case it will become testable in scenarios sensitive to the behavior of energy density such as big bang nucleosynthesis. Moreover, assuming the same quantization parameters leads to identical quantum corrections for photons and fermions only on isotropic backgrounds. Small-scale anisotropies have different effects on both types of fields and can thus also be probed through their implications on the equation of state.

For this, it will be important to estimate more precisely the typical size of corrections, which is not easy since it requires details of the quantum state of geometry. The crucial ingredient is again the patch size of underlying lattice states. On the other hand, taking a phenomenological point of view allows one to estimate ranges for patch sizes, which would leave one in agreement with big bang nucleosynthesis constraints. Interestingly, corrections studied here provide upper bounds to the patch size, and other corrections from quantum gravity are expected to result in lower bounds. A finite window thus results, which can be shrunk with future improvements in observations.

APPENDIX: ELEMENTS OF LOOP QUANTUM GRAVITY

We collect the basic formulas required to compute quantum corrections in loop quantum cosmology, referring to [6,20] for further details. We can restrict our attention to perturbative regimes around a spatially flat isotropic solution where one can choose the canonical variables to be given by functions $(\tilde{p}^I(x), \tilde{k}_J(x))$ which determine a densitized triad by $E_i^a = \tilde{p}^{(i)}(x)\delta_i^a$ and extrinsic curvature by

$K_a^i = \tilde{k}_{(i)}(x)\delta_a^i$. This diagonalization implies strong simplifications in the explicit evaluation of formulas [44,45].

Any state is associated with a spatial lattice with U(1) elements $\eta_{v,I}$ attached to its links $e_{v,I}$ starting at a vertex v and pointing in a direction along a translation generator X_I^a of the homogeneous background. In our context of perturbations around Friedmann-Robertson-Walker space-times, this can be seen as a definition of the quantum gravity state on which the quantized E_i^a and K_a^i act. The U(1) elements $\eta_{v,I}$ appear as matrix elements in SU(2) holonomies $h_{v,I} = \mathrm{Re}\,\eta_{v,I} + 2\tau_I \mathrm{Im}\,\eta_{v,I}$ and represent the connection, while fluxes $F_{v,I}$ are used as independent variables for the momenta. An orthonormal basis of the quantum Hilbert space is given by functionals $|\ldots, \mu_{v,I}, \ldots\rangle = \prod_{v,I} \eta_{v,I}^{\mu_{v,I}}$, for all possible choices of $\mu_{v,I} \in \mathbb{Z}$. Together with basic operators, which are represented as holonomies

$$\hat{\eta}_{v,I}|\ldots, \mu_{v',J}, \ldots\rangle = |\ldots, \mu_{v,I} + 1, \ldots\rangle \qquad (A1)$$

for each pair (v, I), and fluxes

$$\hat{\mathcal{F}}_{v,I}|\ldots, \mu_{v',J}, \ldots\rangle = 2\pi\gamma\ell_{\mathrm{P}}^2(\mu_{v,I} + \mu_{v,-I})|\ldots, \mu_{v',J}, \ldots\rangle, \tag{A2}$$

this defines the basic quantum representation. Here, $\ell_{\mathrm{P}} = \sqrt{\hbar G}$ is the Planck length and a subscript $-I$ means that the edge preceding the vertex v in the chosen orientation is taken.

For quantum corrections to classical equations, one important step in addition to computing expectation values of operators is the introduction of a continuum limit. This relates holonomies

$$\eta_{v,I} = \exp\left(i \int_{e_{v,I}} \mathrm{d}t\, \gamma\tilde{k}_I/2\right) \approx \exp(i\ell_0 \gamma\tilde{k}_I(v + I/2)/2) \tag{A3}$$

to continuum fields $\tilde{k}_I$ through midpoint evaluation (denoted by $v + I/2$), and similarly for fluxes

$$F_{v,I} = \int_{S_{v,I}} \tilde{p}^I(y)\mathrm{d}^2y \approx \ell_0^2 \tilde{p}^I(v + I/2). \tag{A4}$$

Here, ℓ_0 is the coordinate length of lattice links, which enters the continuum approximation since we are integrating classical fields in holonomies and fluxes. From flux operators one defines the volume operator $\hat{V} = \sum_v \prod_{I=1}^3 \sqrt{|\hat{\mathcal{F}}_{v,I}|}$, using the classical expression $V = \int \mathrm{d}^3x \sqrt{|\tilde{p}^1\tilde{p}^2\tilde{p}^3|} \approx \sum_v \ell_0^3 \sqrt{|\tilde{p}^1\tilde{p}^2\tilde{p}^3|} = \sum_v \sqrt{|p^1 p^2 p^3|}$. With (A2), its eigenvalues are

$$V(\{\mu_{v,I}\}) = (2\pi\gamma\ell_{\mathrm{P}}^2)^{3/2} \sum_v \prod_{I=1}^3 \sqrt{|\mu_{v,I} + \mu_{v,-I}|}. \tag{A5}$$

The volume operator is central for inverse triad corrections because inverse densitized triads, or a cotriad, can be

quantized using relations such as

$$\{A_a^i, V_v^r\} = 4\pi\gamma GrV_v^{r-1}e_a^i \tag{A6}$$

for $0 < r < 2$. Even though an inverse power of volume, together with a cotriad, occurs on the right-hand side, the left-hand side can be quantized directly in terms of the volume operator, using holonomies for connection components, and turning the Poisson bracket into a commutator [17,46]. Such a quantization leads to

$$\widehat{V_v^{r-1}e_I^i} = \frac{-2}{8\pi i r\gamma\ell_P^2\ell_0}\sum_{\sigma\in\{\pm 1\}}\sigma\,\mathrm{tr}(\tau^i h_{v,\sigma I}[h_{v,\sigma I}^{-1}, \hat{V}_v^r])$$

$$= \frac{1}{2\ell_0}(\hat{B}_{v,I}^{(r)} - \hat{B}_{v,-I}^{(r)})\delta_{(I)}^i =: \frac{1}{\ell_0}\hat{C}_{v,I}^{(r)}\delta_{(I)}^i \tag{A7}$$

where, for symmetry, we use both edges touching the vertex v along direction X_I^a, and $\hat{B}_{v,I}^{(r)}$ is, after taking the trace in (A7),

$$\hat{B}_{v,I}^{(r)} := \frac{1}{4\pi i\gamma G\hbar r}(s_{v,I}\hat{V}_v^r c_{v,I} - c_{v,I}\hat{V}_v^r s_{v,I}) \tag{A8}$$

with

$$c_{v,I} = \frac{1}{2}(\eta_{v,I} + \eta_{v,I}^*) \quad \text{and} \quad s_{v,I} = \frac{1}{2i}(\eta_{v,I} - \eta_{v,I}^*).$$

As in [20] effects of the quantization of triad (metric) coefficients are included by inserting correction functions in the classical Hamiltonian which follow, e.g., from the eigenvalues [20]

$$C_{v,I}^{(1/2)}(\{\mu_{v',I'}\}) = 2(2\pi\gamma\ell_P^2)^{-1/4}|\mu_{v,J} + \mu_{v,-J}|^{1/4}|\mu_{v,K}$$
$$+ \mu_{v,-K}|^{1/4}(|\mu_{v,K} + \mu_{v,-K} + 1|^{1/4}$$
$$- |\mu_{v,K} + \mu_{v,-K} - 1|^{1/4}) \tag{A9}$$

(where indices J and K are defined such that $\epsilon_{IJK} \neq 0$) of operators $\hat{C}_{v,I}^{(1/2)}$.

Classically, we expect $q_{IJ}/\sqrt{q} = \sqrt{|p^1 p^2 p^3|}/p^I p^J$ for this quantity, with a densitized triad $E_i^a = p^{(i)}\delta_i^a$ and using the relation (A2) between labels and flux components. Although for large $\mu_{v,I}$ the eigenvalues indeed approach the function

$$C_{v,I}^{(1/2)}(\{\mu_{v',I'}\})C_{v,J}^{(1/2)}(\{\mu_{v',I'}\}) \sim (2\pi\gamma\ell_P^2)^{-1/2}$$

$$\times \frac{\prod_{K=1}^3\sqrt{|\mu_{v,K} + \mu_{v,-K}|}}{|\mu_{v,I} + \mu_{v,-I}||\mu_{v,J} + \mu_{v,-J}|}$$

from the classical expectation, they differ for values of $\mu_{v,I}$ closer to 1. This deviation can, for an isotropic background, be captured in a single correction function

$$\alpha_{v,K} = \frac{1}{3}\sum_I C_{v,I}^{(1/2)}(\{\mu_{v',I'}\})^2 \cdot \frac{\sqrt{2\pi\gamma\ell_P^2}(\mu_{v,I} + \mu_{v,-I})^2}{\prod_{J=1}^3\sqrt{|\mu_{v,J} + \mu_{v,-J}|}},$$
$$\tag{A10}$$

which would equal 1 in the absence of quantum corrections. This is indeed approached in the limit where all $\mu_{v,I} \gg 1$, but for any finite values there are corrections. If all $\mu_{v,I} > 1$ one can directly check that corrections are positive, i.e. $\alpha_{v,K} > 1$ in this regime. Expressing the labels in terms of the densitized triad through fluxes (A2) results in functionals

$$\alpha[p^I(v)] = \alpha_{v,K}(4\pi\gamma\ell_P^2\mu_{v,I}) \tag{A11}$$

which enter quantum corrections.

[1] J. D. Barrow and R. J. Scherrer, Phys. Rev. D **70**, 103515 (2004).
[2] M. Bojowald, Living Rev. Relativity **8**, 11 (2005), http://relativity.livingreviews.org/Articles/lrr-2005-11/.
[3] C. Rovelli, *Quantum Gravity* (Cambridge University Press, Cambridge, England, 2004).
[4] A. Ashtekar and J. Lewandowski, Classical Quantum Gravity **21**, R53 (2004).
[5] T. Thiemann, arXiv:gr-qc/0110034.
[6] M. Bojowald and R. Das, Phys. Rev. D **75**, 123521 (2007).
[7] M. Bojowald, H. Hernández, M. Kagan, P. Singh, and A. Skirzewski, Phys. Rev. Lett. **98**, 031301 (2007).
[8] M. Bojowald, AIP Conf. Proc. **917**, 130 (2007).
[9] F. W. Hehl, P. von der Heyde, G. D. Kerlick, and J. M. Nester, Rev. Mod. Phys. **48**, 393 (1976).
[10] S. Mercuri, Phys. Rev. D **73**, 084016 (2006).
[11] M. Bojowald and R. Das, arXiv:0710.5722.

[12] J. F. Barbero G., Phys. Rev. D **51**, 5507 (1995).
[13] G. Immirzi, Classical Quantum Gravity **14**, L177 (1997).
[14] S. Holst, Phys. Rev. D **53**, 5966 (1996).
[15] A. Ashtekar, Phys. Rev. D **36**, 1587 (1987).
[16] T. Thiemann, Classical Quantum Gravity **15**, 1487 (1998).
[17] T. Thiemann, Classical Quantum Gravity **15**, 1281 (1998).
[18] M. Bojowald and A. Skirzewski, Rev. Math. Phys. **18**, 713 (2006).
[19] M. Bojowald and A. Skirzewski, Int. J. Geom. Methods Mod. Phys. **4**, 25 (2007).
[20] M. Bojowald, H. Hernández, M. Kagan, and A. Skirzewski, Phys. Rev. D **75**, 064022 (2007).
[21] We are not assuming strict isotropy to compute quantum corrections of inhomogeneous fermion fields. Nevertheless, in leading order corrections one can use the background geometry.
[22] M. Bojowald, Gen. Relativ. Gravit. **38**, 1771 (2006).

[23] M. Bojowald, Gen. Relativ. Gravit. **40**, 639 (2008).

[24] A. Ashtekar, T. Pawlowski, and P. Singh, Phys. Rev. D **74**, 084003 (2006).

[25] M. Bojowald, D. Cartin, and G. Khanna, Phys. Rev. D **76**, 064018 (2007).

[26] W. Nelson and M. Sakellariadou, Phys. Rev. D **76**, 044015 (2007).

[27] W. Nelson and M. Sakellariadou, Phys. Rev. D **76**, 104003 (2007).

[28] M. Bojowald and G. M. Hossain, Phys. Rev. D **77**, 023508 (2008).

[29] J. D. Barrow, Mon. Not. R. Astron. Soc. **184**, 677 (1978).

[30] J. Yang, D. N. Schramm, G. Steigman, and R. T. Rood, Astrophys. J. **227**, 697 (1979).

[31] F. S. Accetta, L. M. Krauss, and P. Romanelli, Phys. Lett. B **248**, 146 (1990).

[32] C. J. Copi, A. N. Davis, and L. M. Krauss, Phys. Rev. Lett. **92**, 171301 (2004).

[33] T. Clifton, J. D. Barrow, and R. J. Scherrer, Phys. Rev. D **71**, 123526 (2005).

[34] K. A. Olive, G. Steigman, and T. P. Walker, Phys. Rep. **333**, 389 (2000).

[35] D. Kirkman, D. Tytler, N. Suzuki, J. M. O'Meara, and D. Lubin, Astrophys. J. Suppl. Ser. **149**, 1 (2003).

[36] K. A. Olive and E. D. Skillman, Astrophys. J. **617**, 29 (2004).

[37] E. W. Kolb and R. J. Scherrer, Phys. Rev. D **25**, 1481 (1982).

[38] C. Bambi, M. Giannotti, and F. L. Villante, Phys. Rev. D **71**, 123524 (2005).

[39] D. N. Spergel *et al.*, Astrophys. J. Suppl. Ser. **170**, 377 (2007).

[40] A. Perez and C. Rovelli, Phys. Rev. D **73**, 044013 (2006).

[41] L. Freidel, D. Minic, and T. Takeuchi, Phys. Rev. D **72**, 104002 (2005).

[42] M. Bojowald, Classical Quantum Gravity **19**, 5113 (2002).

[43] M. Bojowald, Pramana **63**, 765 (2004).

[44] M. Bojowald, Classical Quantum Gravity **17**, 1489 (2000).

[45] M. Bojowald, Classical Quantum Gravity **21**, 3733 (2004).

[46] T. Thiemann, Classical Quantum Gravity **15**, 839 (1998).