Extensive investigation of the generalized dark matter model

Michael Kopp,[*] Constantinos Skordis,[†] and Dan B. Thomas[‡]

Department of Physics, University of Cyprus, 1, Panepistimiou Street, 2109 Aglantzia, Cyprus
(Received 25 May 2016; published 10 August 2016)

The cold dark matter (CDM) model, wherein the dark matter is treated as a pressureless perfect fluid, provides a good fit to galactic and cosmological data. With the advent of precision cosmology, it should be asked whether this simplest model needs to be extended, and whether doing so could improve our understanding of the properties of dark matter. One established parametrization for generalizing the CDM fluid is the generalized dark matter (GDM) model, in which dark matter is an imperfect fluid with pressure and shear viscosity that fulfill certain postulated closure equations. We investigate these closure equations and the three new parametric functions they contain: the background equation of state w, the speed of sound c_s^2 and the viscosity c_{vis}^2. Taking these functions to be constant parameters, we analyze an exact solution of the perturbed Einstein equations in a flat GDM-dominated universe and discuss the main effects of the three parameters on the cosmic microwave background (CMB). Our analysis suggests that the CMB alone is not able to distinguish between the GDM sound speed and viscosity parameters, but that other observables, such as the matter power spectrum, are required to break this degeneracy. In order to elucidate further the meaning of the GDM closure equations, we also consider other descriptions of imperfect fluids that have a nonperturbative definition and relate these to the GDM model. In particular, we consider scalar fields, an effective field theory (EFT) of fluids, an EFT of large-scale structure, nonequilibrium thermodynamics and tightly coupled fluids. These descriptions could be used to extend the GDM model into the nonlinear regime of structure formation, which is necessary if the wealth of data available on those scales is to be employed in constraining the model. We also derive the initial conditions for adiabatic and isocurvature perturbations in the presence of GDM and standard cosmological fluids and provide the result in a form ready for implementation in Einstein-Boltzmann solvers.

I. INTRODUCTION

It is now a century since Einstein proposed his theory of gravity, general relativity (GR). In that time, GR has passed every experimental test [1] and has few, if any, serious competitors. However, this experimental success necessitates the existence of dark matter (DM) and dark energy (DE), collectively called the dark sector, in order for galactic and cosmological observations to be satisfied. Although GR is then consistent with the observations, this implies that the total energy density of the present-day Universe is dominated by the dark sector, for which we do not have any nongravitational evidence.

In order to achieve agreement with the observations [2], it is sufficient to treat DM and DE as two noninteracting perfect fluids with very simple properties. In particular, DM is modeled with zero pressure ($P_c = 0$) and DE is modeled as a cosmological constant Λ with constant energy density $\rho_\Lambda = \frac{\Lambda}{8\pi G}$ and pressure $P_\Lambda = -\rho_\Lambda$. The assumption of vanishing pressure for DM means that the DM is cold, collisionless and single streaming.[1] This simple model of the dark sector, together with GR as the theory of gravity and the Standard Model (SM) describing the known constituents of matter, forms the standard ΛCDM model of cosmology.

While the DE component of the dark sector is a more recent addition to the standard cosmological model, the evidence for DM goes back much further [3,4]. Further evidence comes from a variety of galactic [5–8], galaxy cluster [9–11], gravitational lensing [12,13], CMB [2,14] and large-scale structure observations [15–18]. The low baryonic energy density as inferred from calculations of the big bang nucleosynthesis and observations of the abundance of light elements [19,20] shows that DM cannot be baryonic.

As mentioned above, the evidence for the dark sector is all gravitational in nature. This has led to the consideration

[*]kopp.michael@ucy.ac.cy
[†]skordis@ucy.ac.cy
[‡]thomas.daniel@ucy.ac.cy

[1]Note that once shell crossing occurs on small scales, technically speaking, CDM ceases to be cold in the sense that the phase space distribution that satisfies the collisionless Boltzmann equation develops velocity dispersion. However, initially cold DM that undergoes shell crossing is still commonly referred to as CDM, although a pressureless fluid description is not possible anymore, and one usually resorts to N-body simulations or the so-called effective theory of large-scale structure to solve for the collisionless dynamics of dark matter in this stage; see Sec. IV B.

043512-1

of alternative theories of gravity in lieu of including DM and DE as new components of the universe; see Ref. [21] for a review. For the question of whether phenomena attributed to DM may be due to the gravitational field not correctly described by GR, one particularly interesting observation is the bullet cluster [9,22]. In this system, the baryonic gas appears to be spatially separated from the dominant contribution to the lensing potential. Thus, in a GR framework, the baryonic gas cannot be the source of the gravitational potential, and an additional matter component is required. The lensing potential of the bullet cluster has minima where CDM would be expected to reside, providing further support for the DM hypothesis. If a different theory of gravity from GR is the correct explanation, then it would have to be nonlocal or contain additional degrees of freedom in such a way as to mimic CDM, such as in [23,24].

Although there is no lack of physically motivated particle dark matter candidates [25], it is commonly assumed that all such candidates behave as a pressureless fluid. Therefore, they are indistinguishable in terms of their purely gravitational properties and can all be modeled as a CDM fluid. As mentioned above, this simple modeling of the dark matter as CDM is consistent with the cosmological and galactic observations. However, to date there have been no convincing detections of dark matter in direct and indirect searches, and these searches have already ruled out many theoretically favored regions in parameter space [26–31].

The assumption of a pressureless perfect fluid does not hold for all dark matter candidates. For instance, a massive neutrino can act as warm dark matter [32–34], and it can be modeled as an imperfect fluid with a nonvanishing pressure and viscosity in the regime where linear perturbation theory applies [35]. Another interesting example is an axion Bose-Einstein condensate, which can also be interpreted as a classical scalar field [36]. This behaves similarly to collisionless DM [37,38], but exhibits a scale-dependent quantum pressure. While the background expansion is identical to CDM, small perturbations around the Friedmann background therefore behave like a fluid with nonadiabatic pressure [39–41]. Even a weakly interacting massive particle (WIMP), which is the most widely accepted dark matter candidate, does not behave as a pressureless perfect fluid on all scales and times relevant for structure formation [42,43]. According to the so-called "effective field theory of large-scale structure" (EFTofLSS) [44–47] (see also [48,49]), even ideal CDM, an initially exactly perfect pressureless fluid, is better described as an imperfect fluid at the level of the Friedmann background and linear perturbations, due to unresolved small-scale nonlinearities. In all these cases, the expansion history and evolution of linear dark matter perturbations is modified in a distinctive way. Thus, we could distinguish between and constrain these models using the CMB and other probes of the expansion history and large-scale structure formation.

Interestingly, observed halo properties deviate from expectations of ΛCDM and might hint at dark matter being more complicated than CDM. For instance, many observed halo density profiles have cores in their centers rather than cusps [50], and some have substructures [51] that are at odds with ΛCDM simulations and suggest that DM might not be collisionless. Also, the low observed mass function of small halos seems to be in conflict with expectations from ΛCDM simulations [52–54].

Warm DM [55,56], condensate DM [38,57] or interacting DM [58] can all alleviate some problems of ΛCDM. In light of the lack of a detection of a DM particle, the interest in DM beyond CDM and the improved precision of cosmological data (notably the Planck satellite [2]), it is timely to explore all possible avenues for constraining the nature of dark matter. In general, any deviation away from CDM could introduce new properties for DM and so potentially influence cosmological observables, thus allowing us to investigate the nature of DM.

Searching for signatures beyond ΛCDM in cosmological data requires the specification of an alternative model, which is typically either "fundamental" or phenomenological. The fundamental approach considers a specific model in which, at least in principle, every observable can be worked out. Examples of this include axions [59], collisionless warm dark matter [60,61], collisionless massive neutrinos [35,62], self-interacting massive neutrinos [63,64], DM coupled to dark radiation [65,66], DM coupled to neutrinos or photons [67–70], DM coupled to DE [71–73] or Chaplygin gas [74]. These fundamental (in the sense of specific) models, usually come with a low-dimensional parameter space that can be well constrained by the data. The main downside of the fundamental approach is that each model has to be studied separately. On the other hand, the phenomenological approach introduces, in a more or less *ad hoc* way, some modifications of the ΛCDM model [75–82] that parametrize some basic physical properties shared by a range of fundamental models, but usually without the ability to explicitly map between parameter spaces. Although primarily developed for DE rather than DM, there are also parametrizations that are somewhat in between those two extremes and guarantee a mapping to the parameter space of the fundamental models [83–90]. This usually comes at the price of a very large parameter and free-function space such that only specific subspaces can be studied in practice.

In this paper, we use the generalized dark matter (GDM) model [75], a purely phenomenological approach to constraining DM properties in the linear regime. The model contains one time-dependent free function, the background equation of state parameter $w(a) \equiv \bar{P}_g/\bar{\rho}_g$,[2] and two free functions $c_s^2(k, a)$ (the sound speed) and $c_{\mathrm{vis}}^2(k, a)$ (the viscosity), which are allowed to depend on scale k as well as the scale factor a, but are solution independent. This independence from the solution is why we refer to $w(a)$, $c_s^2(k, a)$ and $c_{\mathrm{vis}}^2(k, a)$ as parameters. The equation of state is not assumed to be of the barotropic form

---

[2]Note that we use w to denote the background equation of state of DM rather than DE.

$P_g \neq P_g(\rho_g)$; i.e., the GDM pressure P_g is not assumed to be a unique function of the GDM energy density ρ_g. Subsequently, the sound speed c_s^2 is not related to w in the standard fashion, where c_s^2 would be equal to the so-called adiabatic sound speed $c_a^2(a) \equiv \dot{P}_g/\dot{\rho}_g$. Considering only scalar perturbations, GDM is determined by these three functions, the "GDM parameters":

$$w(a), \qquad c_s^2(k, a), \qquad c_{\text{vis}}^2(k, a), \qquad (1.1)$$

plus the particular expressions for the linearly perturbed GDM pressure Π_g and shear Σ_g in terms of GDM density and velocity perturbation δ_g and θ_g and parameters; see Sec. II and [75].

GDM has been shown to be a universal tool to constrain the properties of dark matter in a very wide range. For example, it is able to describe ultrarelativistic matter, or a dark fluid that can simultaneously behave as DM and DE [75]. It has also been employed to establish that a large fraction of the ultrarelativistic component is freely streaming, as expected for the cosmic neutrino background [91].

Here, we are interested in GDM as an extension of CDM. Thus, we consider GDM that is close to CDM, in the sense that $w, c_s^2, c_{\text{vis}}^2 \ll 1$. For the case where CDM is replaced by GDM with w as a free parameter and $c_s^2 = c_{\text{vis}}^2 = 0$, w has been constrained using WMAP data to be $|w| < \mathcal{O}(10^{-1})$ [77] at the 95% confidence level (C.L.) and with the Planck 2013 data release [92] to be $|w| < \mathcal{O}(10^{-3})$ [81] at the 99.7% C.L., in both cases combined with various other probes of the expansion history and structure formation. Similar constraints using WMAP have been obtained in [76], although that model slightly differs from GDM; see Sec. II D. In that paper, the case $w = c_s^2$, and $c_{\text{vis}}^2 = 0$ was also constrained, with the result $|w| < \mathcal{O}(10^{-6})$ at the 99.7% C.L.

In a companion paper [93], we presented the first study jointly constraining all three GDM parameters w, c_s^2 and c_{vis}^2. Using only the Planck 2015 data release [94] supplemented by either Hubble space telescope or baryon acoustic oscillations data, we found $|w| < \mathcal{O}(10^{-3})$ and $c_s^2, c_{\text{vis}}^2 < \mathcal{O}(10^{-6})$, both at the 99.7% C.L. In a future work, we intend to extend this analysis to consider degeneracies with other extensions of the base ΛCDM model, such as the curvature Ω_K, the inclusion of isocurvature modes and considering the neutrino mass as a free parameter rather than fixing it to a specific value. We will also allow the GDM parameters to vary with a and k.

Recently our constraints on constant GDM parameters have been confirmed by another group [95]. In that work time-varying GDM parameters proportional to a^{-2}, mimicking warm dark matter, have also been constrained, and their values today are $w, c_s^2, c_{\text{vis}}^2 < \mathcal{O}(10^{-10})$ at the 99% C.L.

If it turns out that nonzero GDM parameters are favored, we would interpret this as evidence that DM is more complicated than CDM.[3] If CDM remains the favored model, it would be worthwhile to extend the analysis to time- and scale-dependent GDM parameters, as well as to also extend the GDM model itself to deal with quasilinear and nonlinear scales. These scales are relevant for galaxy and Lyman-α surveys [15–18,99], which will help to break degeneracies but, on the other hand, are also much harder to employ due to their inherently nonlinear physics.

In this paper, we investigate the GDM parametrization in order to better understand the nature of the GDM parameters. We also explore its relation to several physical models in order to elucidate to which of them the GDM parameters may relate to. This may be used as a guide for possible future improvements and generalizations of it, particularly in the nonlinear regime. Specifically, the models we study are nonequilibrium thermodynamics, effective theories of CDM and fluids, a particular class of scalar field dark matter and tightly coupled fluids.

The structure of the paper is as follows. In Sec. II we define the GDM model along with some notation and some straightforward extensions. We then focus on the cosmological phenomenology of the GDM model in Sec. III. In particular, we derive all possible types of initial conditions and use the adiabatic mode to analyze the perturbations of a simplified GDM model using an exact solution as well as in a more realistic situation containing all known forms of matter and radiation. That analysis is then used to discuss CMB observables calculated with a modified CLASS code [100] in which we implemented GDM and the modified adiabatic and isocurvature initial conditions. The two most important results of this investigation are that the sound speed c_s^2 and viscosity c_{vis}^2 are strongly degenerate in the CMB (for adiabatic initial conditions) and that, unlike CDM, the GDM isocurvature mode is distinguishable from the baryon isocurvature mode. In Sec. IV we consider models that are more fundamental than GDM, in the sense that they are only defined for the background and linear perturbations, but also nonperturbatively. The aim is to better understand in which circumstances those models can be described by GDM in the linear regime. This sheds some light on the interpretation of the GDM closure equations for pressure and shear and serves as a guide for future extensions of GDM into the nonlinear regime of structure formation. Section IV A shows that nonequilibrium thermodynamics allows for shear and pressure perturbations that can be approximated by GDM. We relate the EFTofLSS to GDM in Sec. IV B. In Sec. IV C we review that both monotonically rolling and oscillating scalar fields allow a DM-like behavior that can be mapped to GDM. Section IV

---

[3]In [96,97] it was shown that a GDM model with $c_s^2 = c_a^2$ and $c_{\text{vis}}^2 = 0$ can parametrize completely different physical situations in which DM is CDM, but either interacts with DE energy or gravity behaves differently from GR. This kind of degeneracy can never be eliminated in linear perturbation theory, as has been first exemplified in [98].

D shows that an effective theory of imperfect fluids based on scalar fields contains particular scale-dependent GDM pressure perturbations, although it is, in general, more complex. In Sec. IV E we consider a fluid composed of two tightly coupled adiabatic fluids, which nevertheless gives rise to a nonadiabatic pressure of the GDM type in certain limits. We conclude in Sec. V.

II. A SHORT OVERVIEW OF THE GDM MODEL

The GDM model is a phenomenological description of a fluid where the pressure P_g and shear Σ_g fluid variables are related to the density and velocity variables via two closure equations. As this description is formulated, and is only valid, in a linearly perturbed Friedman-Robertson-Walker (FRW) universe, we first give a short description of cosmological perturbation theory before discussing the defining relations of the model.

Throughout this work we use the conventions of Misner-Thorne-Wheeler [101] where spacetime indices and spatial indices are denoted by lowercase Greek and lowercase Latin letters, respectively.

A. The energy-momentum tensor

The energy-momentum tensor of a general fluid has the form

$$T_{\mu\nu} = (\rho + P)u_\mu u_\nu + P g_{\mu\nu} + \Sigma_{\mu\nu}, \qquad (2.1)$$

where ρ is the energy density, P is the pressure and $\Sigma_{\mu\nu}$ is the symmetric anisotropic stress tensor obeying $u^\mu \Sigma_{\mu\nu} = \Sigma^\mu{}_\mu = 0$. We choose the four-velocity u_μ (normalized to $u^\mu u_\mu = -1$) to be in the Landau-Lifshitz (LL) frame; thus it is defined as the energy eigenvector of the energy-momentum tensor $u_\alpha T^\alpha{}_\nu = -\rho u_\nu$.[4]

Although the GDM fluid may be used in any theory of gravity, we work exclusively within general relativity. The metric $g_{\mu\nu}$ obeys the Einstein equations

$$G_{\mu\nu} = 8\pi G T_{\mu\nu}, \qquad (2.2)$$

which are sourced by the total energy-momentum tensor $T^\mu{}_\nu$ of matter. The latter is a sum of the individual energy-momentum tensors for each matter component indexed by "I" as

$$T^\mu{}_\nu = \sum_I T_I{}^\mu{}_\nu = T_g{}^\mu{}_\nu + T_{\mathrm{DE}}{}^\mu{}_\nu + T_{\mathrm{SM}}{}^\mu{}_\nu + \cdots, \qquad (2.3)$$

where the label "g" stands for GDM, "SM" for Standard Model, and "DE" for dark energy. The Standard Model

fields may be further split into photons, neutrinos and baryons, labeled with "γ", "ν" and "b," respectively. Each individual energy-momentum tensor $T_I{}^\mu{}_\nu$ takes the form (2.1) with density ρ_I, pressure P_I, LL four-velocity $u_I{}^\mu$ and shear $\Sigma_I{}^{\mu\nu}$. Unless otherwise indicated, the energy-momentum tensors are assumed to be separately conserved $\nabla_\mu T_I{}^\mu{}_\nu = 0$, and the conservation of the total energy-momentum tensor is a consequence of (2.2).

The conservation and the Einstein equations do not provide enough information to solve for the pressure P_I and the shear $\Sigma_I{}^{\mu\nu}$. These two fluid quantities have to be specified in terms of the density ρ_I, the four-velocity u_I^μ, the metric $g_{\mu\nu}$ and possibly additional degrees of freedom like the particle number density n_I. The closure equations for P_I and $\Sigma_I{}^{\mu\nu}$ determine the physical properties of the fluid I.

B. The Friedman universe and its perturbations

1. Perturbed metric and matter variables

The perturbed FRW metric to linear order is

$$ds^2 = a^2 \left\{ -(1 + 2\Psi)d\tau^2 - 2\vec{\nabla}_i \zeta d\tau dx^i \right.$$
$$\left. + \left[\left(1 + \frac{1}{3}h\right)\gamma_{ij} + D_{ij}\nu \right] dx^i dx^j \right\}, \qquad (2.4)$$

where $a(\tau)$ is the scale factor of conformal time τ, γ_{ij} is the metric (used to raise and lower three-dimensional indices) of a three-dimensional space of constant curvature κ, $\vec{\nabla}_i$ is the covariant derivative of γ_{ij} and $D_{ij} = \vec{\nabla}_i \vec{\nabla}_j - \frac{1}{3}\gamma_{ij}\vec{\nabla}^2$ is a traceless derivative operator. The perturbed metric contains the four scalar modes Ψ, h, ζ and ν from which we find it useful to define the metric variable

$$\eta = \frac{1}{6}(\vec{\nabla}^2 \nu - h). \qquad (2.5)$$

We omit the four-vector and the two tensor modes as they are not responsible for structure formation. We also find it convenient to work with Fourier-space transfer functions which depend on wave number k. In particular, in flat spacetime we expand a perturbed variable $A(\tau, \vec{x}) = \int \frac{d^3k}{(2\pi)^3} e^{i\vec{k}\cdot\vec{x}} \tilde{A}(\tau, k) \xi_A(\vec{k})$ where $\tilde{A}(\tau, k)$ is the transfer function of variable $A(\tau, \vec{x})$ and $\xi_A(\vec{k})$ the primordial random perturbation. Since there is no confusion arising, we omit the tilde from the Fourier-space variables.

For each fluid component the four-velocity is parametrized as

$$u_0 = -a(1 + \Psi), \qquad u_i = -a\vec{\nabla}_i \theta, \qquad (2.6)$$

---

[4]Note that a heat flux q_ν does not appear in $T_{\mu\nu}$ because of our choice of u_ν to be the LL frame. There is no loss of generality with this choice.

where θ is the scalar velocity perturbation of the fluid and the fluid index was suppressed for brevity.[5] Furthermore, we perturb the density as $\rho = \bar{\rho}(1 + \delta)$ and the pressure as $P = \bar{\rho}(w + \Pi)$ where w is the (background) equation of state and $\Pi = \delta P/\bar{\rho}$ is the normalized pressure perturbation. With these considerations the energy-momentum tensor for each fluid becomes

$$T^0{}_0 = -\bar{\rho}(1 + \delta), \qquad (2.7a)$$

$$T^0{}_i = -(\bar{\rho} + \bar{P})\vec{\nabla}_i\theta, \qquad (2.7b)$$

$$T^i{}_0 = (\bar{\rho} + \bar{P})\vec{\nabla}^i(\theta - \zeta), \qquad (2.7c)$$

$$T^i{}_j = \bar{\rho}(w + \Pi)\delta^i{}_j + (\bar{\rho} + \bar{P})D^i{}_j\Sigma, \qquad (2.7d)$$

where the index "I" on the fluid variables is again suppressed for brevity. Note that on a FRW background $\bar{\Sigma}_{\mu\nu} = 0$; hence, the shear appears only at the perturbed level through the scalar mode Σ[6] (as we have ignored vector and tensor modes). The total energy-momentum tensor is analogously defined using the total variables. For instance, $\bar{\rho}\delta = \sum_I \bar{\rho}_I\delta_I$ and likewise for the other perturbations.

2. The background and perturbed equations

The Einstein equation (2.2) for the unperturbed FRW background becomes the two Friedmann equations

$$3\mathcal{H}^2 + 3\kappa = 8\pi Ga^2\bar{\rho}, \qquad (2.8)$$

$$2\dot{\mathcal{H}} + \mathcal{H}^2 + \kappa = -8\pi Ga^2\bar{P}, \qquad (2.9)$$

where $\mathcal{H} = \frac{\dot{a}}{a}$ and dots denote derivatives with respect to conformal time τ. Once again, $\bar{\rho} = \sum_I \bar{\rho}_I$ and $\bar{P} = \sum_I \bar{P}_I$. For the Ith component energy conservation $\nabla_\mu T_I{}^\mu{}_\nu = 0$ implies that

$$\dot{\bar{\rho}}_I = -3\mathcal{H}(1 + w_I)\bar{\rho}_I, \qquad (2.10)$$

$$w_I \equiv \frac{\bar{P}_I}{\bar{\rho}_I}, \qquad (2.11)$$

and similarly for the total energy-momentum tensor. Related to the equation of state is the adiabatic sound speed defined via

$$c_{aI}^2 \equiv \frac{\dot{\bar{P}}_I}{\dot{\bar{\rho}}_I} = w_I - \frac{\dot{w}_I}{3\mathcal{H}(1 + w_I)}. \qquad (2.12)$$

If w_I is time independent, then $c_{aI}^2 = w_I$.

For notational simplicity we denote the GDM equation of state w_g by w (without the subscript g) and denote the total equation of state parameter $w_{\text{tot}} = \bar{P}/\bar{\rho}$ to distinguish it from w. At the background level, the GDM equation of state is completely determined by a time-dependent function $w(a)$.[7] Likewise, the adiabatic sound speed is also completely determined by $w(a)$.

At the linearized level, in Fourier space, the Einstein equation (2.2) for scalar modes gives the four equations

$$\mathcal{H}(\dot{h} - 2k^2\zeta) - 6\mathcal{H}^2\Psi - 2(k^2 - 3\kappa)\eta = 8\pi Ga^2\bar{\rho}\delta, \qquad (2.13a)$$

$$2\dot{\eta} + 2\mathcal{H}\Psi + \kappa(\dot{\nu} + 2\zeta) = 8\pi Ga^2(\bar{\rho} + \bar{P})\theta, \qquad (2.13b)$$

$$-\ddot{h} - 2\mathcal{H}\dot{h} + 6\mathcal{H}\dot{\Psi} + 6(\mathcal{H}^2 + 2\dot{\mathcal{H}})\Psi - 6\kappa\eta \\ + 2k^2(\eta - \Psi + \dot{\zeta} + 2\mathcal{H}\zeta) = 24\pi Ga^2\bar{\rho}\Pi, \qquad (2.13c)$$

and

$$\frac{1}{2}\ddot{\nu} + \dot{\zeta} + \mathcal{H}(\dot{\nu} + 2\zeta) + \eta - \Psi = 8\pi Ga^2(\bar{\rho} + \bar{P})\Sigma. \qquad (2.13d)$$

For the matter fluids we need to perturb $\nabla_\mu T_I{}^\mu{}_\nu = 0$. This gives two first-order equations: the continuity equation

$$\dot{\delta}_I = 3\mathcal{H}(w_I\delta_I - \Pi_I) - (1 + w_I)\left[k^2(\theta_I - \zeta) + \frac{1}{2}\dot{h}\right] \qquad (2.14)$$

and the Euler equation

$$\dot{\theta}_I = -(1 - 3c_{aI}^2)\mathcal{H}\theta_I + \frac{\Pi_I}{1 + w_I} - \frac{2}{3}(k^2 - 3\kappa)\Sigma_I + \Psi. \qquad (2.15)$$

Up to this point the gauge has not been fixed. Standard gauges are easily obtained: Synchronous gauge requires $\zeta = \Psi = 0$, while conformal Newtonian gauge sets $\nu = \zeta = 0$ and identifies the second Newtonian potential as $\Phi \equiv \eta = -h/6$.

---

[5]Note that our notation for the velocity perturbation θ is related to [75,102] via $\theta = (v - B)_{\text{Hu}}/k = \theta_{\text{MB}}/k^2$.

[6]Note that our notation for the shear Σ is related to [75,102] via $(1 + w)k^2\Sigma = w\pi_{\text{Hu}} = \frac{3}{2}(1 + w)\sigma_{\text{MB}}$.

[7]Specifying $w(a)$ does not determine the functional form $P = P(\rho, \dots)$, such that the (nonperturbative) equation of state is unknown. However, on the background level any equation of state assumes the form $\bar{P} = \bar{P}(\bar{\rho}, \dots) = w(a)\bar{\rho}$ and thus $w(a)$ parametrizes the equation of state relevant for the background.

As is common, and also very useful, we define gauge-invariant variables. Two standard gauge-invariant variables are the Bardeeen potentials $\hat{\Phi}$ and $\hat{\Psi}$ defined as

$$\hat{\Phi} \equiv \eta + \mathcal{H}\left(\frac{1}{2}\dot{\nu} + \zeta\right), \qquad (2.16a)$$

$$\hat{\Psi} \equiv \Psi - \frac{1}{a}\partial_\tau\left[a\left(\frac{1}{2}\dot{\nu} + \zeta\right)\right], \qquad (2.16b)$$

while a third useful gauge-invariant metric variable is

$$\mathcal{R} \equiv \hat{\Phi} + \frac{2}{3}\frac{\dot{\hat{\Phi}} + \mathcal{H}\hat{\Psi}}{(1+w)\mathcal{H}}. \qquad (2.17)$$

Two gauge-invariant variables that we use further below are

$$\hat{\Delta}_g \equiv \delta_g + 3(1+w)\mathcal{H}\theta_g, \qquad (2.18a)$$

$$\hat{\Theta}_g \equiv \theta_g - \zeta - \frac{1}{2}\dot{\nu}, \qquad (2.18b)$$

corresponding to the rest frame or comoving GDM density perturbation and the conformal Newtonian GDM velocity perturbation, respectively.

C. Definition of the GDM model

The variables Π_I and Σ_I are not determined by the fluid equations (2.14) and (2.15). In the case of fluids, the closure equations for Π_I and Σ_I must be specified in terms of metric and other fluid variables. If the fluid is comprised of particles, Π_I and Σ_I can be expressed in terms of the distribution function of the microscopic theory that satisfies a Boltzmann equation. Whether closure equations for Π_I and Σ_I in terms of the other fluid variables can be derived depends on the details of the microscopic theory and the availability of approximations for the evaluation of the phase space integrals. For instance, ultrarelativistic collisionless radiation, such as massless neutrinos, has $\Pi_\nu = \delta_\nu/3$. However, in general, no closed form equation for Σ_ν can be derived without making some approximations. If the microscopic theory is that of a classical field rather than specified in terms of particles, the explicit form of the energy-momentum tensor in terms of the field and its derivatives follows from the field Lagrangian. Alternatively, the equation of state and the closure equation may be postulated to achieve a desired physical behavior, as is the case for the GDM model.

The scalar perturbations $\delta_g, \theta_g, \Pi_g, \Sigma_g$ of GDM satisfy the continuity and Euler equations of (2.13) (with $I = g$) and two postulated closure equations for the pressure perturbation Π_g and the shear Σ_g [75]. These are

$$\Pi_g = c_a^2\delta_g + (c_s^2 - c_a^2)\hat{\Delta}_g \qquad (2.19a)$$

and

$$\dot{\Sigma}_g = -3\mathcal{H}\Sigma_g + \frac{4}{1+w}c_{\mathrm{vis}}^2\hat{\Theta}_g. \qquad (2.19b)$$

Making the gauge invariance explicit is useful as the shear Σ_g and the nonadiabatic pressure

$$\Pi_{\mathrm{nad}} \equiv \Pi_g - c_a^2\delta_g \qquad (2.20)$$

are always gauge invariant independently of their particular definition. The significance of this particular choice of the closure equations (2.19) will be discussed in the next subsection.

We note here that our equation for the shear is slightly different than the form originally postulated in [75]. The difference is in the $-3\mathcal{H}\Sigma_g$ term which in the case of [75] is replaced by $-\frac{3c_a^2}{w}\mathcal{H}\Sigma_g$ in our notation. We chose this modification of the original equation in order to easily allow for crossing the $w = 0$ point if a time-dependent equation of state is used. Clearly if $\dot{w} = 0$ the two formulations agree.

To summarize, the GDM model is defined by designing a conserved energy-momentum tensor $T_g^{\mu\nu}$ of the form (2.7) in the LL frame. The background pressure $\bar{P}_g$ is determined by the time-dependent equation of state parameter w which also gives rise to an adiabatic sound speed (2.12). The normalized pressure perturbation Π_g is algebraically given by (2.19a) and depends on the free function $c_s^2(a, k)$, the sound speed, which determines the equation of state at the level of linear perturbations. The scalar mode of the anisotropic stress, Σ_g, obeys the differential equation (2.19b) which contains the free function $c_{\mathrm{vis}}^2(a, k)$, the viscosity. While the adiabatic sound speed c_a^2 is completely determined once the equation of state $w(a)$ is specified, the sound speed $c_s^2(k, a)$ and the viscosity $c_{\mathrm{vis}}^2(k, a)$ are free functions that can depend on space and time but are independent of the solution, particularly the matter and metric perturbations.

We note that [2,64,103,104] refer to the GDM model [75] but do not include the Hubble friction $-3\mathcal{H}\Sigma_g$ in the shear equation (2.19b). Instead they start with the standard equations for a moment expansion of the Boltzmann equation for all the $F_{n\geq3}$ moments and insert a viscosity parameter in the corresponding shear equation as above while at the same time keeping the F_3 term. However, in [75] the friction term was designed to mimic the missing third moment F_3 of the distribution function in (2.19b), effectively closing the Boltzmann hierarchy through this approximation. This does not mean that the hierarchy $F_{n\geq3}$ is irrelevant, but that the combined effect of the higher moments can be approximated by the friction term. For ultrarelativistic collisionless particles this form can be derived on subhorizon scales from the Boltzmann hierarchy; see Appendix B of [105]. It is also known that the GDM parametrization can model the collisionless Boltzmann equation for nonrelativistic particles [35,62]. In [104] it was noticed that GDM without the friction term does not

provide a good fit to freely streaming massless neutrinos. An independent friction term of the form $R_c \mathcal{H} \Sigma_g$ can arise from the collision term in the Boltzmann equation [102], and if $R_c \gg 1$, then the hierarchy $F_{n \geq 3}$ becomes irrelevant and can be truncated by setting $F_{n \geq 3} = 0$.

D. Simple extensions of GDM

In order to close the continuity (2.14) and Euler (2.15) equations for the generalized dark matter fluid, we postulated two closure equations for the pressure perturbation Π_g and the shear Σ_g (2.19), as proposed in [75]. In this section we discuss simple extensions, or modifications, of these two closure equations.

1. Pressure

Writing (2.19a) explicitly,

$$\Pi_g = c_s^2 \delta_g + 3(1+w)(c_s^2 - c_a^2)\mathcal{H}\theta_g, \qquad (2.21)$$

we see that c_s^2 is proportional to δ_g, so we expect only c_s^2, and not the other variables, to determine the sound speed. The adiabatic sound speed c_a^2 is not *a priori* related to c_s^2 and does not affect the sound speed deep inside the horizon since $\mathcal{H}\theta_g$ is suppressed by a factor $(\mathcal{H}/k)^2$ compared to δ_g. In the case where $c_s^2 = c_a^2$ we recover the standard expression $\Pi_g = c_a^2 \delta_g$. Therefore the nonadiabatic pressure (2.20)[8] of GDM, i.e.,

$$\Pi_{\mathrm{nad}} = (c_s^2 - c_a^2)\hat{\Delta}_g, \qquad (2.24)$$

is a simple ansatz that allows for an effective sound speed c_s^2 if $c_s^2 \neq c_a^2$, but reduces to the standard adiabatic pressure in the case $c_s^2 = c_a^2$.

The above requirements, however, are not sufficient to determine the shape of Π_{nad}. Consider, for instance,

---

[8]Several definitions of the "(intrinsic) entropy perturbation" Γ related to the nonadiabatic pressure Π_{nad} exist in the literature. In particular, [75,106–108] define

$$\Gamma \equiv \frac{\dot{\bar{P}}}{\bar{P}}\left(\frac{\delta P}{\dot{\bar{P}}} - \frac{\delta\rho}{\dot{\bar{\rho}}}\right) = \frac{1}{w}\Pi_{\mathrm{nad}}, \qquad (2.22)$$

while [109] defines

$$\Gamma \equiv \frac{\delta P}{\dot{\bar{P}}} - \frac{\delta\rho}{\dot{\bar{\rho}}} = \frac{\bar{\rho}}{\dot{\bar{P}}}\Pi_{\mathrm{nad}}. \qquad (2.23)$$

As these two different definitions of Γ are not well behaved in situations where $\bar{P}$ and $\dot{\bar{P}}$ can cross zero, we choose to work directly with Π_{nad}.

$$\Pi_{\mathrm{nad}}^{\mathrm{extended}} = (c_s^2 - c_a^2)\Bigg\{(1 - C_1 - C_2)\hat{\Delta}_g$$
$$+ C_1\left[\delta_g + 3\mathcal{H}(1+w)\left(\frac{1}{2}\dot{\nu} + \zeta\right)\right]$$
$$+ C_2[\delta_g - 3(1+w)\eta]\Bigg\}, \qquad (2.25)$$

where C_1 and C_2 are two new parameters which are restricted in the range $0 \leq C_1, C_2 \leq 1$ and the terms multiplying C_1 and C_2 are the gauge-invariant GDM density perturbations in the Newtonian and flat gauges, respectively. One recovers the GDM model by setting $C_1 = C_2 = 0$. All three gauge-invariant density perturbations have the property that c_s^2 becomes the sound speed deep inside the horizon, while the factor $c_s^2 - c_a^2$ ensures that Π_{nad} vanishes for $c_s^2 = c_a^2$. One could add other gauge-invariant variables to $\Pi_{\mathrm{nad}}^{\mathrm{extended}}$; however, if they do not involve δ_g they cannot influence the sound speed. In terms of gauge-invariant variables, Eq. (2.24) may also be written as

$$\Pi_{\mathrm{nad}}^{\mathrm{extended}} = (c_s^2 - c_a^2)[\hat{\Delta}_g - 3\mathcal{H}(1+w)(C_1 + C_2)\hat{\Theta}_g$$
$$- 3(1+w)C_2\hat{\Phi}], \qquad (2.26)$$

where the gauge-invariant potential $\hat{\Phi}$ and gauge-invariant velocity perturbation $\hat{\Theta}$ are defined by (2.16a) and (2.18b), respectively. Interestingly the effective field theory approach of [48] is of this form with $C_1 = 1$ and $C_2 = 0$.

A common justification for the form $\Pi_{\mathrm{nad}} = (c_s^2 - c_a^2)\hat{\Delta}_g$ is described in [75,107,110]. The argument is that the sound speed should be defined in the fluid rest frame[9] as seen by an observer comoving with the fluid. Alternatively,

---

[9]The fluid rest frame is determined by the fluid four-velocity. Usually this is chosen to be the LL frame (used in this work). If however the fundamental degree of freedom is a scalar field, then another natural choice is the scalar frame, or, if there is a particle species with conserved particle number present, a natural choice is the Eckart frame. It should be noted that under a frame change given by a Lorentz boost and to linear order in the boost velocity, Π and δ remain invariant while θ does not. It would then seem that our expressions for $\hat{\Delta}_g$ and Π_{nad}, Eqs. (2.18a) and (2.24), should transform accordingly with the boost velocity (as they contain θ_g). However, $\hat{\Delta}_g$ and Π_{nad} were defined under the assumption of the LL frame, and not in a general frame; in particular, θ_g is the scalar mode contained in the four-velocity of the LL frame of GDM. Once the frame has been fixed, we cannot expect the resulting expressions to be manifestly frame covariant. One also needs to keep in mind that there is a distinction between a frame choice, that is, the physical definition of the four-velocity in the energy-momentum tensor, and a gauge choice, that is, the fixing of the space-time coordinate system. From a practical point of view these two choices have many things in common. Both are necessary to remove redundancy in the description, and also aspects of the choice of gauge can be connected to a four-velocity field [111]. We return to the issue of frame choice in Sec. IV A.

one can simply choose a gauge adapted to the rest frame in which $T_g{}^i{}_0|_{\rm rf} = T_g{}^0{}_i|_{\rm rf} = 0$ (equivalently $\theta_g|_{\rm rf} = \zeta|_{\rm rf} = 0$). In this gauge it is then postulated that $c_s^2 = \delta P_g/\delta\rho_g|_{\rm rf} = \Pi_g/\delta_g|_{\rm rf}$ is a parameter of the theory that does not explicitly depend on the particular solution of Π_g and δ_g. After performing a gauge transformation away from the rest frame, we obtain the GDM form (2.21). A similar argument in which the rest frame is replaced by either the conformal Newtonian or the flat frame leads to the second or third expressions in (2.25), respectively. Since the sound speed is a fluid property, the fluid rest frame is arguably a more natural choice compared to the two geometrical frames. In any case, the assumption that there exists any frame in which $\Pi_g/\delta_g|_{\rm frame}$ is a solution-independent function is quite strong. In Sec. IV we study several models where this happens either exactly or approximately. In those cases where such a frame exists, it turns out to be the fluid rest frame.

In addition to the arbitrariness of which gauge-invariant combination to use in order to define $\Pi_{\rm nad}$, there is no reason to expect that $\Pi_{\rm nad}$ is related to them algebraically. Indeed, as we show in Secs. IV A and IV E, if GDM is thought of as arising from nonequilibrium thermodynamics or from two tightly coupled perfect fluids, $\Pi_{\rm nad}$ satisfies a first-order differential equation similar to that of the GDM shear, Σ_g. This additional degree, however, oscillates with a similar frequency as δ_g albeit with a small phase shift. Therefore we expect that neglecting a possible dynamical contribution to $\Pi_{\rm nad}$ can be compensated for by adjusting c_s^2 and $c_{\rm vis}^2$ in the GDM model.

2. Bulk viscosity

Yet another possible contribution to $\Pi_{\rm nad}$ is bulk viscosity $P_{\rm bulk}$, a contribution to the isotropic stress whose main effect is not to modify the sound speed but to impede the isotropic expansion of the fluid. Note that while the freedom to choose $w(a)$ would easily accommodate bulk viscosity in GDM at the background level, the shape of $\Pi_{\rm nad}$ (2.24) excludes this possibility. The main effect of bulk viscosity could be modeled by adding a term $c_{\rm bulk}^2 \mathcal{H}^{-1}\vec{\nabla}^2\hat{\Theta}_g$ to $\Pi_{\rm nad}$. We expect its main effect to be similar to shear (or anisotropic stress) $\Sigma_g{}^i{}_j$ which impedes shearing flows $D^i{}_j\hat{\Theta}_g$ rather than $\vec{\nabla}^2\hat{\Theta}_g$. In the context of cosmology this has been studied in [43,112–116]. Bulk viscosity is known to be irrelevant for radiation [117]. However there is no *a priori* reason to neglect it in applications to DM [42,43]. We do not study bulk viscosity in detail in the present work.

3. Shear viscosity

The tightly coupled photon-baryon fluid is a well-known example for an imperfect fluid with small shear. The shear is suppressed by the small number $R_c^{-1} = \tau_c\mathcal{H}$, where τ_c is

the mean time between collisions of photons and free electrons. This allows a truncation of the Boltzmann hierarchy of the photon distribution function and justifies the fluid description. This example (see [102,106]) therefore suggests the following generalization of the GDM shear (2.19b):

$$\dot{\Sigma}_g^{\rm extended} = -3\mathcal{H}R_c\Sigma_g + \frac{4}{1+w}\tilde{c}_{\rm vis}^2\hat{\Theta}_g. \qquad (2.27)$$

One could therefore think of $R_c(a)$ as a new parameter, which is set to 1 in [75] in order to match the behavior of freely streaming radiation; see Appendix B of [105]. The limit $R_c = 0$ is realized in elastic dark energy models where $\tilde{c}_{\rm vis}^2$ acts as rigidity rather than viscosity [118] and is therefore of less interest in applications to DM. If $R_c \gg 1$, the shear at leading order in R_c^{-1} becomes algebraically related to the other perturbations [102], which leads to

$$\Sigma_g^{\rm extended} \simeq \frac{4}{(1+w)\mathcal{H}} \frac{\tilde{c}_{\rm vis}^2}{d_{\rm IC} + 3R_c}\hat{\Theta}_g. \qquad (2.28)$$

Here, we introduced by hand a constant parameter $d_{\rm IC} > 0$, the leading-order power $\Sigma_g \propto \tau^{d_{\rm IC}}$ of the solution to Eq. (2.27) for $k\tau \to 0$. This ensures that for $c_{\rm vis}^2 = \tilde{c}_{\rm vis}^2$ the solution of (2.27) will initially agree with (2.29). For adiabatic initial conditions $d_{\rm IC} = 2$ while for isocurvature modes $d_{\rm IC} = 0$ (CDM or baryon isocurvature), $d_{\rm IC} = 2$ (neutrino isocurvature density) and $d_{\rm IC} = 1$ (neutrino isocurvature velocity).

In the case of the photon-baryon plasma we have $R_c \gg 1$, giving rise to an effectively algebraic shear with $c_{\rm vis}^2 \propto R_c^{-1}\tilde{c}_{\rm vis}^2$. In the following we set $R_c = 1$ such that

$$\Sigma_g^{\rm alg} = \frac{4}{(1+w)\mathcal{H}} \frac{c_{\rm vis}^2}{d_{\rm IC} + 3}\hat{\Theta}_g \qquad (2.29)$$

exactly agrees with Hu's (2.19b) at early times, i.e. as $k\tau \to 0$, and approximately at later times. Figure 1 shows a comparison between the GDM shear (2.19b) and the algebraic version (2.29), for adiabatic initial conditions. Both versions qualitatively agree and lead to a similar damping of GDM density perturbations, as is depicted in Fig. 2.

For $R_c = 1$ and $\tilde{c}_{\rm vis}^2 = c_{\rm vis}^2$ Eq. (2.27) becomes the GDM closure equation (2.19b) which was designed to describe the shear in a medium composed of freely streaming particles, where the friction term $-3\mathcal{H}\Sigma_g$ serves as an approximation to the Boltzmann hierarchy [35,75,105].

Following the argument that led us to $\Pi_{\rm nad}^{\rm extended}$ (2.25), we can now extend $\Sigma_g^{\rm alg}$ (2.29) by adding other gauge-invariant combinations of θ_g in addition to $\hat{\Theta}_g$. While in (2.25) we avoided including terms involving $k^2\theta_g$, we now avoid
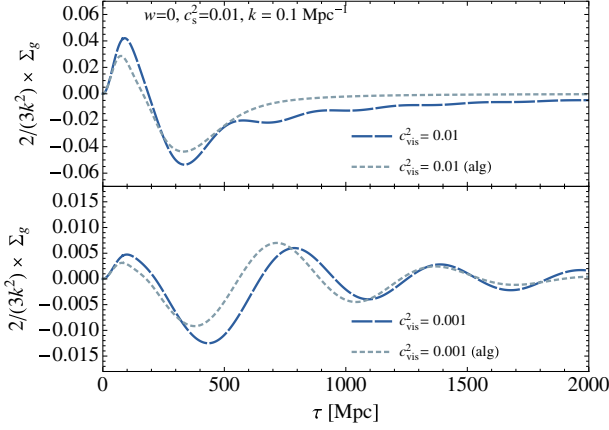
FIG. 1.   Comparison between dynamical (2.19b) and algebraic (alg) (2.29) shear with adiabatic initial conditions for a set of standard cosmological parameters. The upper panel shows the overdamped case $c_{\rm vis}^2 = c_s^2$, the lower panel the case $c_{\rm vis}^2 \ll c_s^2$.

adding terms involving δ_g to Σ_g in order to make the physical effects of Π_g and Σ_g as distinct as possible. The only other gauge-invariant velocity perturbation apart from $\hat{\Theta}_g$ that can be constructed solely from the metric and θ_g is the GDM-comoving curvature perturbation $\mathcal{H}\mathcal{R}_g = \theta_g + \mathcal{H}\eta = \hat{\Theta}_g + \mathcal{H}\hat{\Phi}$, such that

$$\Sigma_g^{\rm extented, alg} = \frac{4}{5(1+w)\mathcal{H}} c_{\rm vis}^2 (\hat{\Theta}_g + C_3 \mathcal{H}\hat{\Phi}) \qquad (2.30)$$
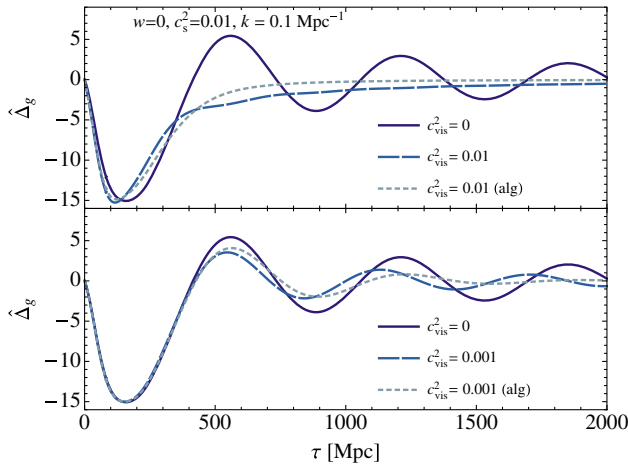
with $0 < C_3 < 1$.

FIG. 2.   Comparison of the time evolution of a single k-mode of the GDM density perturbation $\hat{\Delta}_g$ for $w = 0$ and $c_s^2 = 0.01$. The upper and lower panels compare $c_{\rm vis}^2 = 0$ (solid curve) to $c_{\rm vis}^2 = 0.1 c_s^2$ and $c_{\rm vis}^2 = c_s^2$, respectively. In each panel we show the dynamical and algebraic shear models, (2.19b) and (2.29), with adiabatic initial conditions for a set of standard cosmological parameters.

III. PHENOMENOLOGY OF THE GDM MODEL

In this section we discuss the CMB phenomenology of the GDM model, first analytically and then numerically with CLASS. After determining the growing initial condition modes in Sec. III A, we solve analytically the algebraic GDM model where the shear is given by (2.29) and the universe is purely GDM dominated, in Sec. III B. The main results are that (i) the metric potential $\hat{\Phi}$ necessarily decays below a scale $k_{\rm d}^{-1}$ given by (3.25), that (ii) on an even smaller scale, k_J^{-1}, sound waves may form, and that (iii) on a yet smaller scale $k_{\rm damp}^{-1}$ acoustic oscillations are impossible to form. Section III C outlines the equations for $\hat{\Phi}$ in a universe filled with a realistic mixture of fluids and gives a qualitative discussion for how the CMB observables depend on $\hat{\Phi}$ and the GDM parameters. Finally in Sec. III D we discuss the numerical solution for $\hat{\Phi}$ and various observable CMB power spectra that have been employed in [93] to constrain the full GDM model (2.19).

A. Initial conditions

We start by determining all possible initial condition modes for scalar perturbations. We assume that in the limit $\tau \to 0$, the GDM parameters w, c_s^2 and $c_{\rm vis}^2$ are time independent and much smaller than unity. This assumption is relevant and justified *a posteriori*, given that the constraints obtained on GDM as dark matter strongly constrain $|w| < \mathcal{O}(10^{-3})$ and $c_s^2, c_{\rm vis}^2 < \mathcal{O}(10^{-6})$ [93]. Thus, we construct the initial condition modes as a series expansion in w, c_s^2 and $c_{\rm vis}^2$, keeping only the lowest relevant order. Let us also note that adiabatic initial conditions in the case where $c_{\rm vis}^2 = 0$ have been derived in [119].

In addition to GDM we include all standard fluids which are the baryons, CDM (denoted by a subscript "c"), photons and neutrinos, the latter assumed to be massless in the deep radiation era. These are grouped into radiation (photons and neutrinos; denoted by a subscript "r"), and matter (baryons, CDM and GDM; denoted by a subscript "m"). Keeping CDM in addition to GDM can be useful in studies where DM is a mixture of CDM and GDM, or simply to make a modification of the Boltzmann code tidier. The curvature and the cosmological constant terms can be safely ignored at early times.

When numerically integrating the Einstein-Boltzmann system of equations, one starts the integration on superhorizon scales $\mathcal{H}_k^{-1} \equiv k\mathcal{H}^{-1} \ll 1$. If the initial time is chosen deep enough in the radiation era, such that corrections to $\mathcal{H} = 1/\tau$ are small, then the superhorizon condition simplifies to $x = k\tau \ll 1$. Thus, x may be used as a time coordinate and, in addition, as a series expansion parameter in a way specified below.

1. Background evolution

The background density is the sum of the radiation and matter component $\bar{\rho} = \bar{\rho}_r + \bar{\rho}_m$, which individually evolve as

$$\bar{\rho}_r = \bar{\rho}_{ri}\left(\frac{a_i}{a}\right)^4, \tag{3.1}$$

$$\bar{\rho}_m = \bar{\rho}_{di}\left(\frac{a_i}{a}\right)^3 + \bar{\rho}_{gi}\left(\frac{a_i}{a}\right)^{3(1+w)}, \tag{3.2}$$

where a_i is the scale factor and $\bar{\rho}_{ri}$, $\bar{\rho}_{di}$ and $\bar{\rho}_{gi}$ the radiation, dust (CDM + baryons) and GDM densities, respectively, all evaluated at the initial time. We further define

$$f_{mr} \equiv \frac{\bar{\rho}_{mi}}{\bar{\rho}_{ri}} \quad \text{and} \quad \lambda_k \equiv \frac{\frac{8\pi G}{3}\bar{\rho}_{mi}}{k\sqrt{\frac{8\pi G}{3}\bar{\rho}_{ri}}}a_i \tag{3.3}$$

and the relative species contributions

$$S_X \equiv \frac{\bar{\rho}_{Xi}}{\bar{\rho}_{mi}}, \qquad S_Y \equiv \frac{\bar{\rho}_{Yi}}{\bar{\rho}_{ri}}, \tag{3.4}$$

where $\bar{\rho}_{mi} = \bar{\rho}_{di} + \bar{\rho}_{gi}$ and where X may be c, b or g and Y either γ or ν.

The procedure for obtaining the initial conditions requires an expansion of all variables as a power series in x. While in the standard calculation (without GDM), a series in integer powers of x suffices, the GDM density term which is of the form $a^{3(1+w)} \approx a^3(1 + 3w \ln a)$ requires the addition of terms involving $\ln x$ and powers thereof. We expect that in the limit $w \to 0$ and also as $S_g \to 0$, the standard radiation-matter solution should be reproduced; hence, assuming that all expansion coefficients are w independent, the only plausible expansion is

$$\tilde{a} \equiv \frac{f_{mr}}{a_i\lambda_k}a(x) = \left(1 + \frac{1}{4}\lambda_k x\right)x$$

$$+ wS_g x\left[\sum_{n=1}^{\infty} a_n^{(w)}x^{n-1} + \ln x\sum_{n=1}^{\infty} a_n^{(\ln,w)}x^{n-1} + \cdots\right] + \cdots \tag{3.5}$$

where $a_n^{(w)}$ and $a_n^{(\ln,w)}$ are coefficients to be determined and where we have ignored terms involving higher powers of w and $\ln x$. Note that $(\ln x)^2 \gg |\ln x|$ for small enough x such that it is not clear *a priori* that our ansatz [see also (3.8) below] solves the Einstein and fluid equations, and if so, that the approximate solution is a good solution. However, the full numerical solution of (A1) shows that this is indeed a good approximation.

Inserting (3.5) into the Friedmann equation determines the coefficients as $a_1^{(w)} = a_1^{(\ln,w)} = 0$,

$$a_2^{(\ln,w)} = -\frac{3\lambda_k}{4}, \qquad a_2^{(w)} = \frac{3\lambda_k}{4}\left[\frac{1}{2} - \ln\left(\frac{\lambda_k}{f_{mr}}\right)\right], \tag{3.6a}$$

and for all $n \geq 3$,

$$a_n^{(\ln,w)} = 0, \qquad a_n^{(w)} = -A_n(-\lambda_k)^{n-1}, \tag{3.6b}$$

where

$$A_{n+1} = \frac{n-1}{2(n+1)}A_n - \frac{3}{2^{2n-1}(n+1)(n-1)(n-2)} \tag{3.6c}$$

with $A_3 = \frac{1}{16}$ as the starting value. Ignoring the w and $\ln x$ corrections, which amounts to approximating the GDM component as CDM, incorrectly predicts several leading-order solutions for the matter-type isocurvature perturbations.

2. Perturbations

In order to find the allowed initial conditions for the perturbations and their initial time and scale dependence, we expand all perturbational variables as a series involving the small parameter x following a similar procedure as in [120]. In the standard case without GDM, a power series in x suffices; however, as in the background case, the presence of the background GDM density scaling as $a^{3(1+w)}$ requires the inclusion of powers of $\ln x$. For convenience, we work with the dimensionless variables $\sigma \equiv \frac{2}{3}k^2\Sigma$ and $v \equiv k\theta$.

The problem of finding the initial condition comprises two parts: (i) determine how many regular growing mode solutions exist (corresponding to the adiabatic and various isocurvature modes), and (ii) obtain the solutions to the perturbed field equations as a series in x (and $\ln x$) thereby allowing the numerical integration to start at a convenient time without mixing adiabatic and isocurvature modes.

We adopt the synchronous gauge by setting $\Psi = \zeta = 0$. This gauge has a residual gauge mode which is set to zero by discarding decaying initial conditions.[10]

Following [120] we assume that photons and baryons are tightly coupled through Thomson scattering, such that $v_\gamma = v_b$ and all higher moments of the photon Boltzmann hierarchy vanish. In addition, on superhorizon scales the Boltzmann hierarchy of neutrinos can be truncated at third order (due to free-streaming), keeping only δ_ν, v_ν and σ_ν. The resulting equations are displayed in Appendix A.

In order to construct the initial condition modes, we need to specify an ansatz for the solution of the perturbational variables

---

[10]In the synchronous gauge, the CDM velocity perturbation satisfies $av_c = $ const which is identical for the solution to the residual gauge mode. The residual gauge freedom allows us to set this constant to zero, $v_c \equiv 0$. This is not true for any other type of fluid, including GDM, where v_g has a solution different from the residual gauge mode.

$$\mathcal{P} = \{\eta, h, \delta_b, \delta_c, \delta_\gamma, v_\gamma, \delta_\nu, v_\nu, \sigma_\nu, \delta_g, v_g, \sigma_g\}. \qquad (3.7)$$

By inspection of the x-dependence of the scale factor (3.5) we choose the following ansatz for the solution:

$$\mathcal{P} = \mathcal{P}_0 + \mathcal{P}_1 x + \mathcal{P}_1^{(\varepsilon)} x + \mathcal{P}_1^{(\ln,\varepsilon)} x \ln x$$
$$+ \mathcal{P}_2 x^2 + \mathcal{P}_2^{(\varepsilon)} x^2 + \mathcal{P}_2^{(\ln,\varepsilon)} x^2 \ln x + \cdots \qquad (3.8)$$

where ε is a proxy for the GDM parameters w and c_s^2, assumed to have the same smallness. The coefficients without an ε label are independent of w and c_s^2, and we keep only linear order in ε in the ansatz to avoid higher powers of $\ln x$. In the limit $\varepsilon \to 0$, one recovers the standard ΛCDM initial conditions. We note that the constant term h_0 for the metric variable h can be set to zero by a gauge transformation. An ansatz containing powers like x^{1-3w} as used in [121] does not work if we want to recover all possible modes, adiabatic and isocurvature.

For the GDM density contrast δ_g we also include the term $\delta_{g0}^{(\ln,\varepsilon)} \ln x$, which is necessary to find the GDM isocurvature mode for $w \neq c_s^2$. Thus, $\delta_g = \delta_{g0} + \delta_{g0}^{(\ln,\varepsilon)} \ln x + \cdots$, where the remaining terms follow the expansion in (3.8). This additional term does not introduce a new type of initial condition. When $w = c_s^2$ no pure $\ln(x)$ term is required.

3. Solution method

The ansatz (3.8) is used in the perturbed Einstein and fluid equations (A1), and the coefficients for the same powers of x and $\ln x$ are matched, thus providing a consistent solution. We collect all variables in the set

$$\mathcal{A} = \mathcal{A}^{\varepsilon=0} + \mathcal{A}^\varepsilon, \qquad (3.9)$$

with $\mathcal{A}^{\varepsilon=0} = \{\mathcal{P}_0, \mathcal{P}_1, \mathcal{P}_2, \ldots\}$ containing the zeroth-order coefficients in ε and $\mathcal{A}^\varepsilon = \{\mathcal{P}_1^{(\varepsilon)}, \mathcal{P}_1^{(\ln,\varepsilon)}, \mathcal{P}_2^{(\varepsilon)}, \ldots\}$ containing the correction due to ε. We expand all functions up to order x^n, with the exception of η, σ_ν and σ_g which avoids the introduction of coefficients with label $n+1$.

We chose $n = 4$ and used a brute force method to test for every possible subset $\mathcal{I}_{i,\text{test}}$ of $\mathcal{P}_0$ where $i = 1, \ldots, 2^{|\mathcal{P}_0|}$, whether or not $|\mathcal{A}^{\varepsilon=0} - \mathcal{I}_{i,\text{test}}|$ equals the rank of the system of linear equations with $\varepsilon = 0$. Out of the 2^{11} test sets, there are 72 that fulfill this criterion but only four of them with $\max(|\mathcal{I}_{i,\text{test}}|) = 6$. We choose

$$\mathcal{I}_{\text{modes}} = \{\eta_0, \delta_{\nu,0}, v_{\nu,0}, \delta_{c,0}, \delta_{b,0}, \delta_{g,0}\}. \qquad (3.10)$$

The other three possible sets are obtained by exchanging $\delta_{\nu,0}$ with $\delta_{\gamma,0}$ and $v_{\nu,0}$ with $v_{\gamma,0}$. Finally we solve for the remaining coefficients in $\mathcal{P}$, that is, $\mathcal{A}^{\varepsilon=0} - \mathcal{I}_{\text{modes}}$ and $\mathcal{A}^\varepsilon$, such that they are expressed as functions of $\mathcal{I}_{\text{modes}}$.

In all the modes displayed below and in Appendix A, we only include the leading powers of x unless the leading-order solution is constant or it is suppressed by the product of c_{vis}^2 and ε, in which case we include the next-to-leading order as well. The modes have been checked to agree to reasonable accuracy with the solution which includes all powers up to x^4 as well as with a numerical integration of Eq. (A1). We note that the initial condition modes also hold for the algebraic version of the GDM shear (2.29).

4. Adiabatic (Ad)

Setting $\eta_0 = 1$ (which we can always do via rescaling) and all remaining perturbations in $\mathcal{I}_{\text{modes}}$ (3.10) to zero, the adiabatic mode is

$$\eta = 1 - \frac{5 + 4S_\nu}{12(15 + 4S_\nu)} x^2, \qquad h = \frac{1}{2} x^2,$$

$$\delta_c = \delta_b = -\frac{1}{4} x^2, \qquad \delta_\gamma = \delta_\nu = -\frac{1}{3} x^2,$$

$$\delta_g = \left[-\frac{1}{4} + \frac{3c_s^2 - 5w}{8} \right] x^2,$$

$$v_\gamma = -\frac{1}{36} x^3, \qquad v_\nu = -\frac{23 + 4S_\nu}{36(15 + 4S_\nu)} x^3,$$

$$v_g = -\left[\frac{1}{16} c_s^2 + \frac{2c_{\text{vis}}^2}{3(15 + 4S_\nu)} \right] x^3,$$

$$\sigma_\nu = \frac{2}{3(15 + 4S_\nu)} x^2, \qquad \sigma_g = \frac{8c_{\text{vis}}^2}{3(15 + 4S_\nu)} x^2.$$

The adiabatic initial conditions agree with those presented in [95] upon Taylor expansion in ε and c_{vis}^2. A comparison of terms next-to-leading order in x would reveal differences compared to [95], as our solution contains terms involving $\ln x$ even for the adiabatic mode.

5. Isocurvature modes

There are five growing isocurvature modes in the GDM model: the radiation-type neutrino isocurvature density (NID) and neutrino isocurvature velocity (NIV) and the matter-type CDM isocurvature (CI), baryon isocurvature (BI) and GDM isocurvature (GI). As we do not use these modes in the phenomenology of the rest of this section, we display them in Appendix A.

We remark that in searches for signatures of isocurvature modes within ΛCDM, only one of the BI and the CI is included in the analysis since they are completely degenerate [122,123]. The situation of a GDM isocurvature mode is more interesting than CDM, since the C_ls of BI and GI modes are no longer degenerate if either w or c_s^2 is nonzero.

B. Evolution of GDM perturbations and decay of $\hat{\Phi}$

Let us consider a flat GDM dominated universe with algebraic shear (2.29) such that the 00 equation (2.13a), $0i$ equation (2.13b) and shear may be manipulated into

$$k^2\hat{\Phi} = -4\pi G a^2 \bar{\rho}_g \hat{\Delta}_g, \tag{3.11}$$

$$\dot{\hat{\Phi}} + \mathcal{H}\hat{\Psi} = 4\pi G a^2 \bar{\rho}_g (1+w)\hat{\Theta}_g, \tag{3.12}$$

$$\Sigma_g^{\text{alg}} = \frac{4}{5\mathcal{H}(1+w)} c_{\text{vis}}^2 \hat{\Theta}_g, \tag{3.13}$$

where the gauge-invariant variables $\hat{\Phi}$, $\hat{\Psi}$ and $\mathcal{R}$ are given by (2.16a), (2.16b) and (2.17), respectively.

In this case the ij Einstein equations take the form

$$\mathcal{H}^{-1}\dot{\hat{\Phi}} = \left[\frac{3}{2}(1+w) + \frac{12}{5}c_{\text{vis}}^2\right](\mathcal{R}-\hat{\Phi}) - \hat{\Phi}, \tag{3.14a}$$

$$\mathcal{H}^{-1}\dot{\mathcal{R}} = -\left(\frac{k}{\mathcal{H}}\right)^2 \frac{2}{3(1+w)}\left[c_s^2\hat{\Phi} + \frac{4}{5}c_{\text{vis}}^2(\mathcal{R}-\hat{\Phi})\right]. \tag{3.14b}$$

Using e-folding time N defined by $\partial_N = \mathcal{H}^{-1}\partial_\tau$, denoting ∂_N by a prime and assuming constant w, c_s^2 and c_{vis}^2, Eq. (3.14) assumes the form of a damped harmonic oscillator

$$\hat{\Phi}'' + \left(\frac{k}{\mathcal{H}}\right)^2 \left\{\left[c_s^2 + \frac{8c_{\text{vis}}^2(1+3c_s^2)}{15(1+w)}\right]\hat{\Phi} + \frac{8c_{\text{vis}}^2}{15(1+w)}\hat{\Phi}'\right\}$$
$$+ \left[1 + \frac{3}{2}(1+w) + \frac{12}{5}c_{\text{vis}}^2\right]\hat{\Phi}' = 0. \tag{3.15}$$

The above equation shows that the integrated Sachs-Wolfe (ISW) effect in the GDM dominated universe vanishes for $c_s^2 = c_{\text{vis}}^2 = 0$, irrespective of the value of w. In this case $\dot{\mathcal{R}} = 0$ and the equation admits $\dot{\hat{\Phi}} = 0$ such that $\hat{\Phi}$ freezes during GDM domination.[11] On the other hand, if c_s^2 or c_{vis}^2 is nonzero then $\mathcal{R}$ is sourced, but only on subhorizon scales due to the overall factor $(k/\mathcal{H})^2$. Notice however that once w is time dependent, the analogue of (3.15) contains terms proportional to $\hat{\Phi}$ and therefore will generally admit $\dot{\hat{\Phi}} = 0$. A thorough discussion of the effect of small DM sound speed on the ISW effect can be found in [125]. Let us also emphasise that the coefficients of $\hat{\Phi}$ and $\hat{\Phi}'$ are manifestly non-negative for $w > -1/3$ and positive c_{vis}^2 and c_s^2; hence the potential decays, in general.

Sound waves are possible if $c_{\text{vis}}^2 < c_s^2$, and the effective propagation speed is close to c_s^2 if $c_{\text{vis}}^2 \ll c_s^2$. For $c_{\text{vis}}^2 \gtrsim 0.57c_s^2$ on the other hand, the potential decays without

---

[11]We could allow mildly negative sound speeds, in which case a constant potential may also be achieved if $c_s^2 = -\frac{8c_{\text{vis}}^2(1+3c_s^2)}{15(1+w)}$ and $c_{\text{vis}}^2 \geq 0$. However, we do not allow this possibility as it does not seem natural and requires fine-tuning to ensure stability. This stabilizing property has also been observed for the GDM shear in [124].

oscillations. All of these properties may be extracted from the exact solution to (3.15) as we examine in more detail below.

In order to find the exact solution to (3.15) it is easier to transform back to τ as an independent variable. In a flat GDM-dominated universe the Friedman equation gives $\mathcal{H}^{-1} = \tau(1+3w)/2$ so that (3.15) transforms into

$$\ddot{\hat{\Phi}} + \left\{\frac{6[5(1+w)+4c_{\text{vis}}^2]}{5(1+3w)}\frac{1}{\tau} + \frac{4(1+3w)c_{\text{vis}}^2}{15(1+w)}k^2\tau\right\}\dot{\hat{\Phi}}$$
$$+ k^2\left[c_s^2 + \frac{8c_{\text{vis}}^2(1+3c_s^2)}{15(1+w)}\right]\hat{\Phi} = 0. \tag{3.16}$$

Defining $y = -\gamma c_{\text{vis}}^2 k^2\tau^2$, where

$$\gamma = \frac{2(1+3w)}{15(1+w)}, \tag{3.17}$$

transforms the equation into

$$y\frac{d^2\hat{\Phi}}{dy^2} + (\beta - y)\frac{d\hat{\Phi}}{dy} - \alpha\hat{\Phi} = 0, \tag{3.18}$$

where

$$\alpha = \frac{1+3c_s^2}{1+3w} + \frac{15c_s^2(1+w)}{8c_{\text{vis}}^2(1+3w)}, \tag{3.19}$$

$$\beta = \frac{35+45w+24c_{\text{vis}}^2}{10(1+3w)}. \tag{3.20}$$

Equation (3.18) is Kummer's differential equation whose regular solution is the Kummer confluent hypergeometric function $M(a, b, y)$ such that

$$\hat{\Phi} = A_0 M(\alpha, \beta, -\gamma k^2\tau^2 c_{\text{vis}}^2), \tag{3.21}$$

where A_0 is a constant. The regular solutions (3.21) automatically satisfy $\dot{\hat{\Phi}}(\tau=0) = 0$. The nonregular solution of (3.18) is of the form $B(k\tau)^{-n_d}M$, which in the limit $k\tau \to 0$ behaves as

$$(k\tau)^{-n_d}, \qquad n_d = 1 + \frac{4(5+6c_{\text{vis}}^2)}{5(1+3w)}, \tag{3.22}$$

and is therefore a decaying mode and is of no interest to us.

The general solution (3.21) evolves through four regimes of behavior. For a given Fourier mode k, the solution starts on superhorizon scales from $\tau = 0$ with a constant amplitude which persists even after horizon crossing. It begins to decay around the scale k_d, and then on smaller scales the solution will continue to either decay monotonously or enter an acoustic regime, leading to a period of oscillations. This is determined by the relative magnitude of two further

scales, the Jeans scale k_J and the overdamping scale k_{damp}. Once the Jeans scale is crossed, $\hat{\Phi}$ begins a period of oscillations until the overdamping scale is reached, where oscillations cease and $\hat{\Phi}$ simply decays. As can be seen in Fig. 3, depending on the values of the GDM parameters c_s^2 and c_{vis}^2, the solution may go through only the oscillation regime (thick yellow line) or only through the overdamping regime (thin blue line) or both (red line). The Jeans and the overdamping scales can be estimated by examining the zeros of the confluent hypergeometric function $M(\alpha, \beta, -y)$. In [126] it is proved that they are bounded by $y_- < y < y_+$ where

$$y_\pm = 2\alpha - \beta \pm 2\sqrt{\alpha(\alpha - \beta) - \beta}, \qquad (3.23)$$

which in our case translates to the two scales $k_\pm^{-1} = \frac{\sqrt{7}c_{\mathrm{vis}}\tau}{\sqrt{y_\pm}}$. The scale k_- may be identified with the Jeans scale $k_J = k_-$, while the scale k_+ may be related to the over-damping scale if k_+ is real.

We now discuss several special cases and regimes of (3.21) and use them to estimate the above three scales, namely, k_{d}, k_J, and k_{damp} in terms of the GDM parameters. Without loss of generality we set $\hat{\Phi}(\tau = 0) = A_0 = 1$.

1. Case 1: $c_s^2, c_{\mathrm{vis}}^2 = 0$

The nondecaying solution is $\hat{\Phi} = 1$ as is immediately clear from (3.15). This generalizes the standard CDM solution to the case of nonzero constant w, leading to a zero ISW effect.
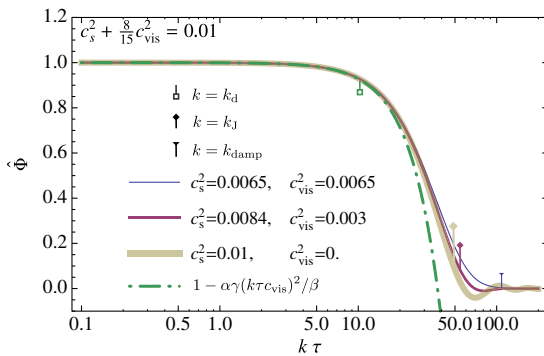
2. Case 2: $c_s k\tau \ll 1$ and $c_{\mathrm{vis}} k\tau \ll 1$

At early times [see Eq. (13.1.2) of [127]] the solution to the potential is

$$\hat{\Phi} \simeq 1 - \frac{\gamma\alpha c_{\mathrm{vis}}^2}{\beta} k^2 \tau^2, \qquad (3.24)$$

which is constant to lowest order in k^2 and decays at next-to-leading order if $w > -1/3$ and $c_s^2, c_{\mathrm{vis}}^2 > 0$. Therefore for reasonable values of GDM parameters, $|w|, c_s^2, c_{\mathrm{vis}}^2 \ll 1$, the potential can only decay. Using (3.24) and taking the limit of small GDM parameters, the comoving scale below which the potential starts to decay is

$$k_{\mathrm{d}}^{-1}(\tau) \equiv \tau\sqrt{c_s^2 + \frac{8}{15}c_{\mathrm{vis}}^2}. \qquad (3.25)$$

The above definition is such that for $k = k_{\mathrm{d}}$, the potential has dropped to $\hat{\Phi} = 13/14 \approx 0.93$. The time evolution of the potential for three different combinations of c_s^2 and c_{vis}^2, keeping the same $k_{\mathrm{d}} = 10/\tau$ and $w = 0$, is shown in Fig. 3. Observables that directly probe the large-scale structure will be sensitive to both the Jeans scale and the over-damping scale, which will be defined further below. However, for the CMB it is mostly the decay scale k_{d}, below which $\hat{\Phi}$ starts to decay, which matters. Therefore one should expect a strong negative degeneracy between c_s^2 and c_{vis}^2 in the CMB spectrum, and this was verified in [93].

3. Case 3: $c_{\mathrm{vis}}^2 = 0$

This is the zero shear viscosity case. The solution may be found by either taking the limit $c_{\mathrm{vis}}^2 \to 0$ of (3.21) with the help of Eq. 13.3.2 of [127] or by setting $c_{\mathrm{vis}}^2 = 0$ in (3.16)
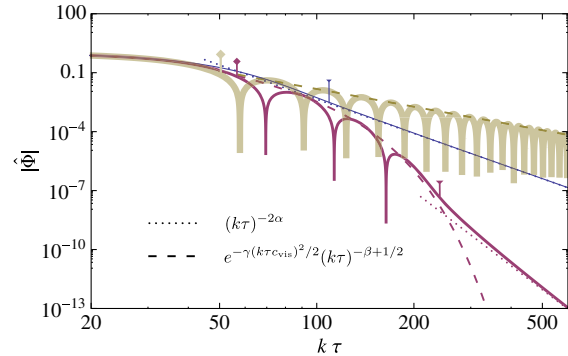

FIG. 3. The solid curves show the exact solution $\hat{\Phi}$ in a flat GDM dominated universe for three GDM parameters as specified in the legend. In all cases we set $w = 0$. The left panel shows that horizon entry at $k\tau \simeq 1$ does not influence $\hat{\Phi}$. However, around $k\tau \simeq k_{\mathrm{d}}\tau = 10$, the potential decays for all three combinations of c_s^2 and c_{vis}^2. Up to this time the solution is well described by (3.24) as is indicated by the dot-dashed curve. The right panel shows the details of the decay. The dashed curves display the envelope of (3.27), which is valid in the acoustic regime starting at k_J, and the dotted curves exhibit the asymptotic behavior in the overdamped regime (3.32) starting well after the last oscillation at k_{damp}. There is no acoustic regime for $c_{\mathrm{vis}}^2 = c_s^2$ (blue, thin line). For $c_{\mathrm{vis}}^2 = 0$ the acoustic oscillations never stop. In the case of nonzero c_{vis}^2 the acoustic regime is accompanied by exponential decay during which the effective sound speed $c_{\mathrm{eff}}^2 < c_s^2$ is reduced.

and transforming it into Bessel's equation. The exact solution in this case simplifies to the well-known result

$$\hat{\Phi} = \frac{A_1 J_n(c_s k\tau)}{(c_s k\tau)^n}, \qquad n = \frac{5 + 3w}{2(1 + 3w)}, \qquad (3.26)$$

where A_1 is a normalization constant and J_n is the Bessel function of order n. The envelope is nearly constant outside the Jeans scale and decays as $\tau^{-n-1/2}$ once $c_s k\tau \geq 1$ as can also be seen through the thick yellow line and its dashed envelope in the right panel of Fig. 3. Deep inside the Jeans scale $c_s k\tau \ll 1$ the potential oscillates with frequency $c_s k$ as is seen by the thick yellow solid curve in Fig. 3.

4. Case 4: $c_{\rm vis} \ll c_s$ and $c_{\rm vis} k\tau \ll 1$

Rather than taking the limit of vanishing viscosity we may expand the exact solution (3.21) in $c_{\rm vis}/c_s \ll 1$ and $c_{\rm vis} k\tau \ll 1$ using Eq. (13.3.7) of [127], leading to

$$\hat{\Phi} \simeq \frac{A_2 e^{-\gamma c_{\rm vis}^2 k^2 \tau^2/2}}{(c_{\rm eff} k\tau)^{\beta-1}} J_{\beta-1}(c_{\rm eff} k\tau), \qquad (3.27)$$

where A_2 is a normalization constant. It may easily be shown that $\beta > 3/2$ for $w > -\frac{1}{3}$, and therefore the solution is always decaying for large $k\tau$, as we discuss further below.

Keeping nonzero $c_{\rm vis}^2$ has a further effect. The solution (3.27) oscillates with frequency $c_{\rm eff} k$ where the *effective*[12] sound speed is

$$c_{\rm eff}^2 = c_s^2 - \frac{2c_{\rm vis}^2}{5(1+w)} \left[1 + 3w - 4\left(c_s^2 - \frac{2}{5} c_{\rm vis}^2 \right) \right] \quad (3.28a)$$

$$\simeq c_s^2 - \frac{2}{5} c_{\rm vis}^2, \qquad (3.28b)$$

where the second line holds for small GDM parameters. We notice that the algebraic shear (on which we have based our calculation) slightly decreases the sound speed, and this is seen in the upper panel of Fig. 2, while the dynamical shear has the opposite effect. Note that the expression (3.28a) determines the effective sound speed beyond the approximation $c_{\rm vis} k\tau \ll 1$, in the sense that the solution for $\hat{\Phi}$ as determined by (3.27) is the lowest term in an expansion in terms of a series of Bessel functions $J_{\beta-1+\beta_n}(c_{\rm eff} k\tau)$ for $\beta_n = 0 \ldots \infty$. It is also worth emphasizing that the effective sound speed is different from c_s^2 even if $c_s^2 = c_a^2$, and hence, even if $P_g = w\rho_g$ for constant w.

Remembering that $-\beta + \frac{1}{2} < -1$, the envelope of (3.27) decays as $e^{-\gamma c_{\rm vis}^2 k^2 \tau^2/2} \tau^{-\beta+1/2}$ once $k \geq k_{\rm d}$, as can be seen by the red dashed curve in the right panel of Fig. 3. This may be derived using Eq. (9.2.1) of [127] which involves the

large-argument expansion of the Bessel function, i.e., $c_{\rm eff} k\tau \gg 1$.[13]

Let us now estimate the Jeans scale k_J. In this regime the relevant parameter that determines the start of the acoustic regime is $c_{\rm eff}^2$, so we will write the sound speed in terms of this quantity. Rearranging (3.28a) and solving for c_s^2 in terms of $c_{\rm vis}^2$, $c_{\rm eff}^2$ and w, we find

$$c_s^2 = \frac{(1+w)c_{\rm eff}^2 + \frac{2(1+3w)c_{\rm vis}^2}{5} + \frac{16c_{\rm vis}^4}{25}}{1 + w + \frac{8c_{\rm vis}^2}{5}} \qquad (3.29)$$

so that c_s^2 is always positive as long as both $c_{\rm vis}^2$ and $c_{\rm eff}^2$ are positive and $w > -1/3$. We now expand (3.23) for small w, $c_{\rm eff}^2$ assuming that they are both of the same order, i.e., $O(w) \sim O(c_{\rm eff}^2)$, and for small $c_{\rm vis}^2$ assuming that it is of order $O(c_{\rm eff}^4)$. This gives the GDM Jeans scale $k_J \simeq k_-$ as

$$k_J^{-1}(\tau) \equiv \frac{2c_{\rm eff}\tau}{\sqrt{105}} \approx 0.2c_{\rm eff}\tau. \qquad (3.30)$$

Note that taking the limit $c_{\rm vis}^2 \to 0$ in (3.27) reproduces (3.26) of case 3, i.e., $c_{\rm vis} = 0$, as expected.

5. Case 5: $c_{\rm vis}/c_s \gtrsim 1$ or $k \gtrsim k_+$

This is the case related to the overdamping regime where the solution decays without any oscillations. In order to determine the scale where this happens, one may start from (3.23), expand in small GDM parameters and associate the overdamping scale with k_+. However, as β always decreases k_+, a better estimate is obtained if we set $\beta = 0$ in (3.23) which leads us to the definition of the over-damping scale as $k_{\rm damp} c_{\rm vis}\tau \equiv 2\sqrt{\alpha/\gamma}$. Once again, we expand this expression for small GDM parameters, now assuming that $O(w) \sim O(c_s^2) \sim O(c_{\rm vis}^2)$; i.e., $c_{\rm vis}^2$ is now assumed to be of the same order as c_s^2. The resulting expression

$$k_{\rm damp}^{-1}(\tau) = \frac{c_{\rm vis}\tau}{\sqrt{30}\sqrt{1 + \frac{15c_s^2}{8c_{\rm vis}^2}}} \approx \frac{0.18c_{\rm vis}\tau}{\sqrt{1 + \frac{15c_s^2}{8c_{\rm vis}^2}}} \qquad (3.31)$$

is now valid not only if $c_{\rm vis}^2 \ll c_s^2$ but also if $c_{\rm vis}^2 \gg c_s^2$ (in which case k_+ is no longer real). Interestingly, for scales below $k_{\rm damp}^{-1}$ the exact solution (3.21) decays with a power law

$$\hat{\Phi} \simeq A_3 (c_{\rm vis} k\tau)^{-2\alpha}, \qquad (3.32)$$

for some constant A_3 [see Eq. (13.1.5) of [127]], rather than exponentially as one might expect from (3.27). This is shown by the red and blue dotted lines in the right panel of

---

[12]We continue to call c_s^2 the sound speed for convenience.

[13]This is valid as long as $c_{\rm eff}$ is sufficiently larger than $c_{\rm vis}$.

Fig. 3. The limit $k\tau \to \infty$ is contained in case 3 through (3.26) and case 5 through (3.32), which shows that $\hat{\Phi} \to 0$ as $k\tau \to \infty$ if either c_s^2 or c_{vis}^2 is nonzero.

The exact solution (3.21) does not admit oscillations and has no exponential decay if $k_+ = k_-$ as may be seen by the blue lines in the right panel of Fig. 3. For small GDM parameters this occurs for

$$c_{\text{vis}}^2 \geq \frac{15}{2(3+\sqrt{105})} c_s^2 \qquad (3.33)$$

which approximates to $c_{\text{vis}}^2 \gtrsim 0.57 c_s^2$.

It is worth noticing that the behavior between k_d^{-1} and $\max(k_J^{-1}, k_{\text{damp}}^{-1})$ is such that for fixed k_d the decay is quickest for $c_{\text{vis}}^2 = 0$ as may be seen in the left panel of Fig. 3.

We remarked already in footnote 11 that a fine-tuned negative sound speed $c_s^2 = -\frac{8}{15} c_{\text{vis}}^2$ can lead to a constant $\hat{\Phi}$ if $c_{\text{vis}}^2 > 0$ because $k_d^{-1} = 0$. This generalizes case 0 to include the possibility $\hat{\Phi} \neq \hat{\Psi}$. If on the other hand the negative sound speed satisfies $|c_s^2| > -\frac{8}{15} c_{\text{vis}}^2$ and therefore $k_d^{-1} > 0$, we are in the regime (3.33) where the potential simply decays as (3.32) below k_d^{-1} without oscillations.

To close this section about the behavior of GDM perturbations, we note that Hu's nonadiabatic pressure Π_{nad} (2.24) is rather special compared to its extended version $\Pi_{\text{nad}}^{\text{extended}}$ (2.26). If we instead use the extended version $\Pi_{\text{nad}}^{\text{extended}}$, Eq. (3.14b) changes to

$$\mathcal{H}^{-1}\dot{\mathcal{R}} = -\frac{2}{3(1+w)}\left(\frac{k}{\mathcal{H}}\right)^2 \left[c_s^2\hat{\Phi} + \frac{4}{5}c_{\text{vis}}^2(\mathcal{R}-\hat{\Phi})\right]$$
$$+ 3(c_a^2 - c_s^2)[(C_1+C_2)(\mathcal{R}-\hat{\Phi}) + C_2\hat{\Phi}], \qquad (3.34)$$

which adds a k-independent source for $\mathcal{R}$, leading to k-independent terms proportional to $\hat{\Phi}$ in the analogue of (3.16). Therefore, the curvature perturbation $\mathcal{R}$ is not conserved on superhorizon scales unless either $C_1 = C_2 = 0$ or $c_s^2 = c_a^2$ (i.e., adiabatic fluid) so that an ISW effect is generated by GDM even for the case $c_s^2 = c_{\text{vis}}^2 = 0$ (baring the trivial case where $w = 0$ in addition). This property of Hu's Π_{nad}, that c_s^2 does not influence superhorizon modes, was observed before in [119].

A similar analysis of the behavior of linear perturbations was performed in [48] for the "Newtonian" model (2.26) with $C_1 = 1$ and $C_2 = 0$ and for a particular time dependence of GDM parameters c_s^2 and c_{vis}^2, which will be discussed in Sec. IV B. There it was also observed that when $w = 0$ the large-scale perturbations are solely sensitive to the combination $c_s^2 + \frac{8}{15}c_{\text{vis}}^2$ as in (3.25).

C. Behavior of $\hat{\Phi}$ for a mix of GDM, baryons and radiation

In this section we qualitatively discuss the evolution of the potential $\hat{\Phi}$ in the presence of a mixture of baryons, photons, neutrinos and GDM, as is relevant for the CMB and large-scale structure formation. This mixture may be treated as a cosmological fluid with equation of state w_{tot}, adiabatic sound speed $c_{a,\text{tot}}^2$ given by $(1+w_{\text{tot}})c_{a,\text{tot}}^2 = \sum_I (1+w_I)\Omega_I c_{aI}^2$ and total nonadiabatic pressure perturbation $\Pi_{\text{nad,tot}} = \Pi - c_{a,\text{tot}}^2\delta$. The GDM does not couple to photons or to baryons; however, it affects the CMB through gravity. Thus, in this section we examine how GDM affects the evolution of the gravitational potential $\hat{\Phi}$ which in turn leaves its imprint on the CMB spectrum, for instance, through the ISW effect, lensing and acoustic driving [128–132].

As in the last subsection, we rewrite the spatial trace Einstein equation (2.13c) for a flat cosmology ($\kappa = 0$) using the traceless Einstein equation (2.13d) to eliminate $\hat{\Psi}$, in terms of the two first-order equations for $\hat{\Phi}$ and $\mathcal{R}$,

$$\hat{\Phi}' = -\hat{\Phi} + 3(1+w_{\text{tot}})\left[\frac{1}{2}(\mathcal{R}-\hat{\Phi}) + \mathcal{H}^2\Sigma\right], \qquad (3.35a)$$

$$\mathcal{R}' = \frac{1}{1+w_{\text{tot}}}\left[\Pi_{\text{nad,tot}} - \frac{2k^2}{3\mathcal{H}^2}c_{a,\text{tot}}^2\hat{\Phi}\right] - \frac{2k^2}{3}\Sigma. \qquad (3.35b)$$

It is then transparent how the evolution of $\hat{\Phi}$ depends on $w_{\text{tot}}, \Pi_{\text{nad,tot}}$ and Σ. We discuss each of these in turn.

1. The equation of state w_{tot}

The total background equation of state w_{tot} depends on the relative abundances and equations of state of the cosmological fluids. It determines the time dependence of a and $\mathcal{H}$ and gives rise to $c_{a,\text{tot}}^2$. It also determines the time of radiation-matter equality when w_{tot} interpolates between $1/3$ and w, and the time of the transition between GDM and Λ domination, with w_{tot} approaching -1 in the latter.

If the right-hand side of (3.35b) vanishes, $\mathcal{R}$ remains constant. However, $\hat{\Phi}$ still retains some temporal evolution if w_{tot} is time dependent. Only in the case where $\Sigma = 0$ and $\dot{w}_{\text{tot}} = 0$ does the potential approach a constant, as was the case in a purely GDM dominated universe with $c_s^2 = c_{\text{vis}}^2 = 0$ studied in the previous subsection. In the realistic universe we consider in this section, w_{tot} is expected to be weakly time dependent even during matter domination since baryons and GDM have a slightly different equation of state, in general. In Fig. 4 we display the evolution of a single k mode of the potential $\hat{\Phi}$, where ΛCDM (black dotted line), $w = 0.01$ (dark-blue dashed line) and $w = -0.01$ (light-blue dot-dashed line) give an approximately constant potential during matter domination which subsequently decays at very late times as Λ eventually comes to
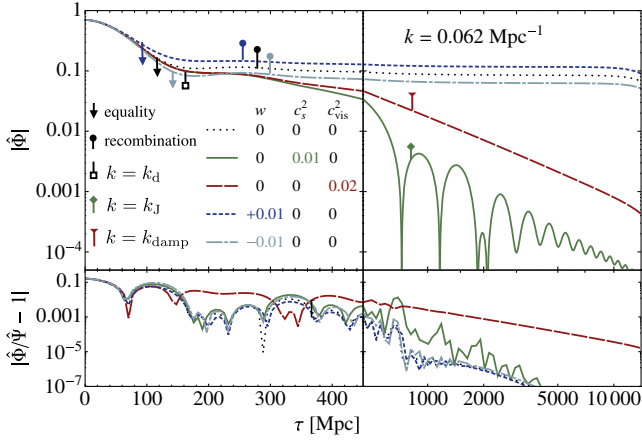
FIG. 4.   Comparison of the effect of w and c_s^2 and c_{vis}^2 on a single k mode of the potential Φ; see the legend in Fig. 5. Compared to Fig. 3, the universe is now filled, in addition, with photons, neutrinos, baryons and a cosmological constant. The lower panel shows the effect of shear Σ causing a $\hat{\Phi}/\hat{\Psi} \neq 1$.

dominate. Observe also that the case $w > 0$ has a larger freeze-out value than $w = 0$, and the opposite happens for $w < 0$. This is easily understood: increasing w shifts the time of radiation-matter equality earlier such that a given k mode spends less time during the era of radiation domination and therefore experiences stronger decay until it freezes out during GDM domination. The opposite is true when w is decreased. Finally, let us note that during GDM domination $c_{a,tot}^2 \simeq c_a^2$ and $\Pi_{nad,tot} \simeq \Pi_{nad}$, such that $c_{a,tot}^2$ has no significant effect on $\hat{\Phi}$.

2. Nonadiabatic pressure $\Pi_{nad,tot}$

Consider a mixture of cosmological fluids that may also be pairwise coupled and therefore exchange energy and momentum. Their energy-momentum tensors would then not be individually conserved, in general, but instead

$$\nabla_\mu T_I{}^\mu{}_\nu = J_{I\nu}, \qquad \sum_I J_{I\nu} = 0, \qquad (3.36)$$

with the background value of the exchange current $J_{I\nu}$ denoted by $Q_I \equiv \bar{J}_{I0}$. Then the total nonadiabatic pressure is given by [106]

$$\Pi_{nad,tot} = \frac{1}{1 + w_{tot}} \sum_{I<J} \Omega_I \Omega_J (1 + w_I)(1 + w_J)$$

$$\times (c_{aI}^2 - c_{aJ}^2) \left(\frac{\delta_I}{1 + w_I} - \frac{\delta_J}{1 + w_J} \right)$$

$$- \frac{1}{3\mathcal{H}\bar{\rho}} \frac{\delta}{1 + w_{tot}} \sum_I c_{aI}^2 Q_I$$

$$+ \frac{1}{\bar{\rho}} \sum_I \bar{\rho}_I \Pi_{Inad}, \qquad (3.37)$$

which is the sum of three terms. The first term (first two lines) is the relative entropy perturbation and vanishes initially for adiabatic initial conditions. It is suppressed when the sound speeds are very similar, when $\Omega_I \ll 1$ or when $\Omega_I \ll \Omega_J$ for all $I < J$. In ΛCDM this is the case during radiation domination when the dominating species, neutrinos and photons, have the same sound speed $c_{a,\nu}^2 = c_{a,\gamma}^2$ and during matter or Λ domination, where the dominating clustering species CDM has $c_{a,c}^2 = 0$ and $1 + w_\Lambda = 0$. The second term (third line), proportional to δ, manifestly modifies the sound speed of the total density perturbation if the fluids exchange energy. This is a subleading effect for standard cosmological fluids, e.g., after recombination when baryons lose a tiny fraction of their energy to photons [133]. The third term (last line), the intrinsic nonadiabatic pressure, is usually assumed to be absent in ΛCDM,[14] but does appear in Λ-GDM. It is given by (2.24) since GDM is the only fluid that admits a sizable intrinsic nonadiabatic pressure.

In a nutshell, we expect Π_{nad}^{tot} to be a subleading effect in ΛCDM and mostly relevant around the radiation-matter equality, when it is dominated by the relative entropy perturbation between matter and radiation. In Λ-GDM, even well within matter domination, Π_{nad} causes $\hat{\Phi}$ to decay below the scale k_d^{-1} given by (3.25) as can be seen by the green solid curve in Fig. 4. We investigate possible physical origins for Π_{nad} in Sec. IV.

3. The shear Σ

In ΛCDM the shear Σ interpolates between a mixture of mainly neutrino and photon shear during radiation domination and vanishes during matter or Λ domination for massless neutrinos. In Λ-GDM, during matter domination, the GDM shear Σ_g provides the dominant contribution to the total shear Σ leading to potential decay, as is displayed by the red dashed curve in Fig. 4. In addition, the total shear causes a difference between $\hat{\Phi}$ and $\hat{\Psi}$ (see the lower panel in Fig. 4), such that $\dot{\Sigma}$ adds a contribution to the ISW effect. The same effect occurs for the lensing potential [132]; any line-of-sight projection of $\hat{\Psi} + \hat{\Phi}$ will be affected by Σ as well as by $\hat{\Phi}$.

D. How GDM affects the CMB

We now discuss the effects that a GDM component may have on the CMB in the case of adiabatic initial conditions. We present the CMB power spectra and compare ΛCDM (black dots) to four cases of Λ-GDM: $c_s^2 = 0.01$ (solid green line), $c_{vis}^2 = 0.02$ (long-dashed red line), $w = 0.01$ (short-dashed dark-blue line) and $w = -0.01$ (dot-dashed

---

[14]Although any fluid with internal degrees of freedom, for instance a baryon fluid, has, in general, some internal nonadiabatic pressure.

light-blue line). We fix the standard cosmological parameters to the best-fit Planck values [2] in all cases. All spectra were produced using a version of the CLASS code [93,100] modified to incorporate GDM.

1. Effect of w

In the case of pure CDM, the most distinctive effect on $\mathcal{D}_l^{TT} = l(l+1)C_l^{TT}/2\pi$ is a modification of the heights of the first few acoustic peaks that depends on ω_c, the dimensionless CDM density [128,134]. This is because the CDM abundance affects the time of radiation-matter equality and therefore which modes enter the horizon during radiation domination. During radiation domination, $\hat{\Phi}$ decays and boosts the observed CMB temperature [128,134] due to acoustic driving. Increasing the CDM density pushes radiation-matter equality earlier, which reduces acoustic driving and lowers the amplitude of the peaks. Indeed, one of the best pieces of evidence for dark matter comes from the CMB spectrum, as the absence of CDM would introduce large acoustic driving, boosting the peak amplitude and leading to a spectrum that completely disagrees with observations.

In the case of GDM, increasing the dimensionless GDM density ω_g gives a rather similar effect to CDM since the equation of state is taken to be small, $|w| \ll 1$ [75–77,81,93,95]. Larger values for w will result in GDM behaving more like radiation, in effect creating large acoustic driving and boosting the CMB peaks to values inconsistent with observations.

Even though w is taken to be small, its actual value is still of importance as the GDM density approximately scales as

$$a^3\bar{\rho}_g \tilde{\propto} \omega_g(1 + 3w\ln(1 + z)). \tag{3.38}$$

In particular, its greatest effect is to shift the time of radiation-matter equality for fixed ω_g. Increasing w raises the amount of GDM in the past (leading to smaller acoustic driving which in turn reduces the peak heights), and this is similar to increasing the dimensionless GDM density ω_g. Accordingly, we expect w and ω_g to be anticorrelated. This effect has been discussed in [75] and observationally shown in [77,93].

In Fig. 5 we compare the temperature and E-mode polarization power spectra, $\mathcal{D}_l^{TT}$ and $\mathcal{D}_l^{EE}$, in a ΛCDM model to two Λ-GDM models with $w = \pm 0.01$. The dotted curve is the reference ΛCDM model with all GDM parameters set to zero and the remaining parameters taken from Planck [94]. The $w = 0.01$ model (dark-blue dashed line) is below that of ΛCDM for the first few peaks, and the opposite is true for $w = -0.01$ (light-blue dot-dashed line). Note that the time difference in horizon entry $\Delta\tau_{k=\mathcal{H}} \simeq$ 0.1 Mpc is much smaller than the shift in the time of radiation-matter equality $\Delta\tau_{\rm eq} \simeq 25$ Mpc. Therefore, the main reason for the modification of the peak heights when
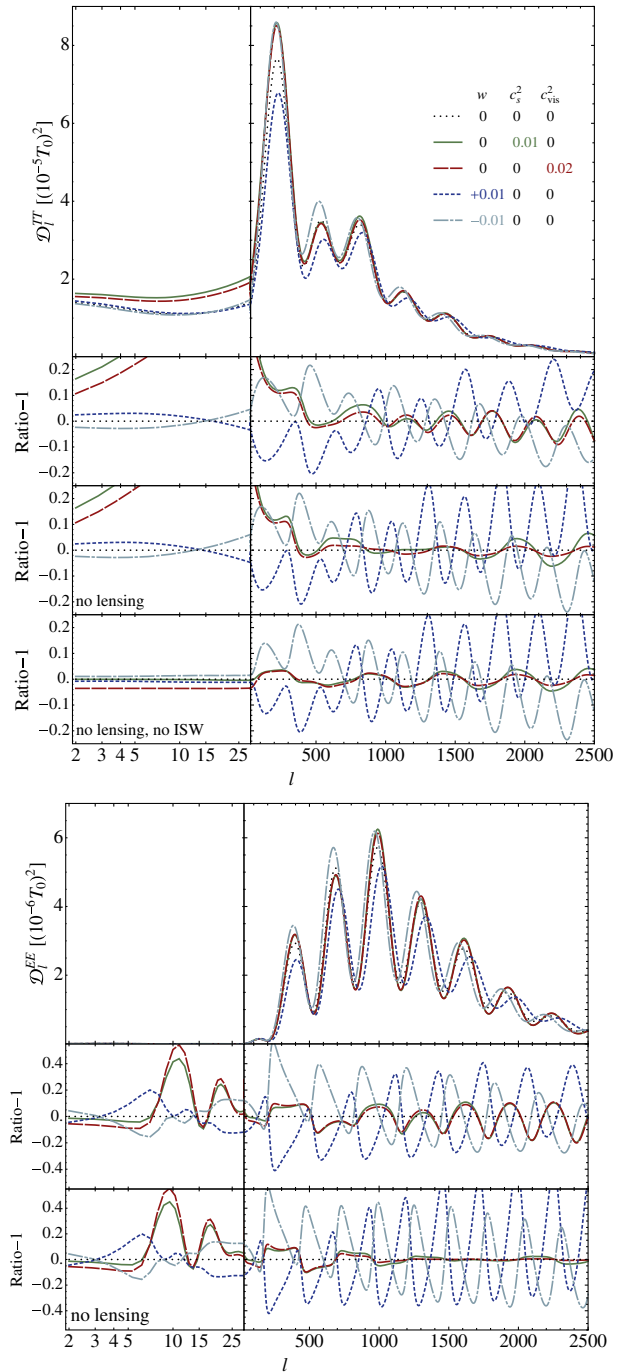

FIG. 5. Comparison of the effect of w and c_s^2 and $c_{\rm vis}^2$ on temperature power spectrum $\mathcal{D}_l^{TT}$ and E-mode polarization power spectrum $\mathcal{D}_l^{EE}$. T_0 is the mean CMB temperature. The lower panels show ration between the cases with a nonzero GDM parameter and a ΛCDM reference model. The panels labeled "no lensing" and "no lensing, no ISW" have been calculated without the effect of lensing and without the ISW effect.

w is varied is a shift of the radiation-matter equality time, denoted by arrows in Fig. 4. In that plot we show the time evolution of a single k-mode of $\hat{\Phi}$ that corresponds to the third C_l peak. In addition, the time of recombination is

shifted by $\Delta\tau_* \simeq \Delta\tau_{\rm eq}$, and therefore the size of the sound horizon at recombination is reduced for positive w. Since the decrease of the sound horizon is accompanied by a decrease in the angular diameter distance to recombination (as varying w directly affects the Hubble parameter H), the change in the peak positions is rather moderate compared to the case where ω_g is varied. Nevertheless, the peaks move slightly to the left (right) for negative (positive) w.

Panels 2–4 of the $\mathcal{D}_l^{TT}$ part of Fig. 5 and panels 2–3 of the $\mathcal{D}_l^{EE}$ part, show the ratio of the C_ls with nonzero GDM parameters to the reference model C_ls, making the change of relative peak heights and also the shift of peak positions more visible. More specifically, the C_l ratio is displayed without the effect of lensing (panel 3 of either part) and without the ISW effect (panel 4 in the $\mathcal{D}_l^{TT}$ part). These C_ls have been calculated by artificially removing the ISW and/or lensing terms in CLASS. It is clear that it is mostly the first few peaks that are affected by the ISW effect as well as all scales larger than the first peak, while the higher peaks are affected by lensing. At low l, the ISW effect for the $w = 0.01$ model is slightly larger than ΛCDM, while for the $w = -0.01$ model it is slightly smaller because the potential freezes to a slightly larger constant value in the former. This fairly small effect was discussed in the previous subsection (see also Fig. 4). The effect of the equation of state w on the lensing amplitude is shown by the dark-blue dashed ($w = 0.01$) and light-blue dot-dashed ($w = -0.01$) curves in Fig. 6. This can be understood from Fig. 4; a positive w allows $\hat{\Phi}$ to freeze out earlier and therefore at a larger value.

2. Effect of c_s^2 and $c_{\rm vis}^2$

Let us now turn to the effects of the perturbative GDM parameters, namely, the sound speed c_s^2 and viscosity $c_{\rm vis}^2$. An important property of CDM is that during CDM domination, $\hat{\Phi}$ freezes to a constant value. For a GDM dominated universe we saw in (3.25) that $\hat{\Phi}$ will be time dependent and decay below

$$k_{\rm d}^{-1}(\tau) = \tau \sqrt{c_s^2 + \frac{8}{15} c_{\rm vis}^2},$$

as long as c_s^2 or $c_{\rm vis}^2$ is nonzero. We therefore expect these two parameters to be degenerate in the CMB, and indeed the cases $c_s^2 = 0.01$ and $c_{\rm vis}^2 = 0.02$ lead to very similar CMB observables. Two further GDM scales that we have uncovered in the previous subsections are the Jeans and overdamping scales $k_J^{-1} = 0.2 c_{\rm eff}\tau$ [see (3.30)] and $k_{\rm damp}^{-1} = \frac{2}{15} c_{\rm vis}^2 k_{\rm d}^{-1}$ [by combining (3.31) with (3.25)], respectively. All three scales are marked in Fig. 4 at $\tau = \tau_*$, the conformal time at recombination. However, these scales are not visible in the CMB. The reason for this is that the CMB spectra are mostly determined by the photon
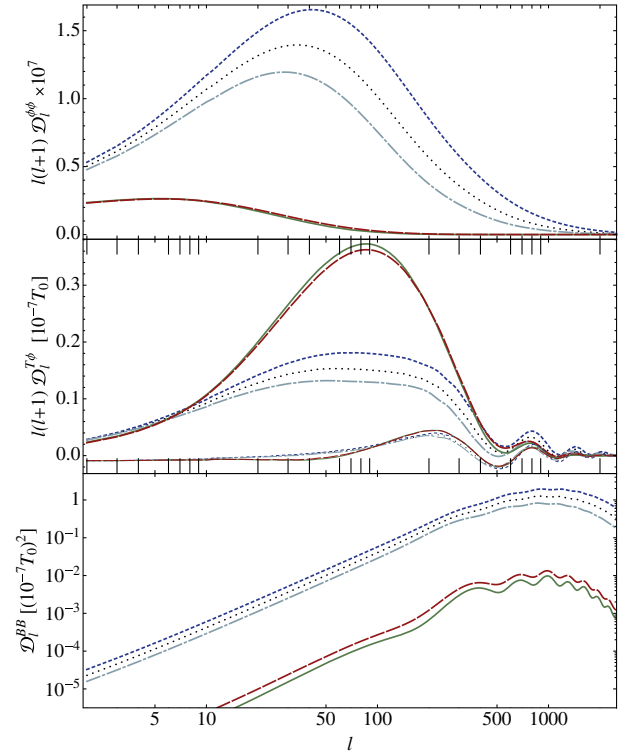

FIG. 6. Lensing potential power spectrum $\mathcal{D}_l^{\phi\phi}$, lensing temperature cross spectrum $\mathcal{D}_l^{T\phi}$ and lensing B-mode power spectrum $\mathcal{D}_l^{BB}$. The faint lines in the second panel show $\mathcal{D}_l^{T\phi}$ without the ISW effect. The different lines correspond to the legend in Fig. 5.

temperature $\delta_g/4$ which is only indirectly sensitive to GDM dynamics, while the potentials play a lesser role; moreover, their effects (such as ISW) are convolved over a wide range of time scales. This makes the GDM scales invisible by eye in the CMB spectra even though the size of the residuals compared to ΛCDM is mainly determined by $k_{\rm d}^{-1}(\tau_*)$. In contrast, at $z = 0$, the potential decay scale $k_{\rm d}(z = 0)$ and the Jeans scale $k_J(z = 0)$ are clearly visible in the matter power spectrum, as we see in Fig. 7.
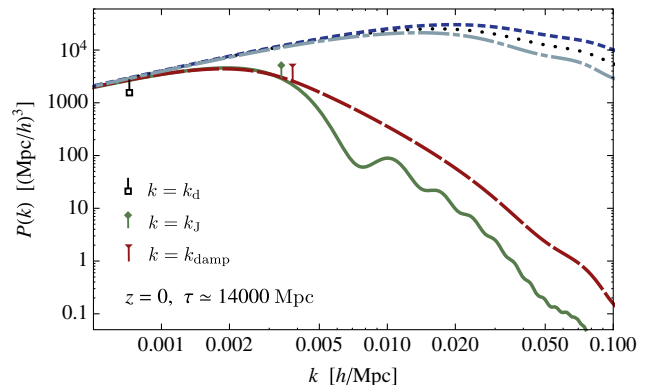

FIG. 7. Total matter power spectrum $P(k)$ at $z = 0$, where the lines correspond to the legend in Fig. 5.

Potential decay for nonzero c_s^2 and c_{vis}^2 leads to smaller CMB lensing compared to ΛCDM and, at the same time, larger (and continuous across time) ISW. This is observed by comparing panels 2 and 4 in the $\mathcal{D}_l^{TT}$ part for the effect on ISW and panels 2 and 3 in both temperature and polarization parts of Fig. 5 for the effect on lensing. Neither the "no lensing" nor the "no ISW" C_ls are directly observable, but the lensing potential power spectrum $\mathcal{D}_l^{\phi\phi}$ and the temperature-lensing cross correlation $\mathcal{D}_l^{T\phi}$, displayed for all models in Fig. 6, are directly observable. We observe that nonzero c_s^2 or c_{vis}^2 leads to a reduction of the lensing potential power spectrum $\mathcal{D}_l^{\phi\phi}$ (upper panel) and lensing B-mode power $\mathcal{D}_l^{BB}$ (lower panel). The lensing-temperature cross correlation $\mathcal{D}_l^{T\phi}$ (middle panel), however, is boosted for nonzero c_s^2, c_{vis}^2 because a larger fraction of the temperature anisotropies are caused by the ISW effect. This is clear from the fainter lines in the $\mathcal{D}_l^{T\phi}$ panel which have been calculated by artificially removing the ISW term. During the radiation matter transition, a nonzero c_s^2 or c_{vis}^2 leads to a quicker decay of the potential and therefore can boost the acoustic driving of the observed temperature of the first couple of CMB peaks. For the chosen parameter values $c_s^2 = 0.01$, $c_{\text{vis}}^2 = 0.02$, this is a subdominant effect compared to lensing and the ISW effect, as may be seen in the "no lensing, no ISW" panel in Fig. 5. Since k_{d}^{-1} is a length scale appearing in the perturbations, we do not expect that varying k_{d}^{-1} will affect the size of the various CMB imprints to the same degree. Indeed for much smaller constant parameters, such as $c_s^2 = 10^{-6}$, the only remaining effect on the CMB spectra is the reduced lensing compared to ΛCDM [93]. On the other hand, if c_s^2 and c_{vis}^2 grow with redshift, i.e., as $c_s^2, c_{\text{vis}}^2 \propto a^{-2}$, then the CMB will be mostly sensitive to k_{d}^{-1} at early times; see the discussion in [95].

The total linear matter power spectrum at $z = 0$ is shown in Fig. 7. The scales k_{d}^{-1} where the potentials start to decay, k_J^{-1} below which GDM oscillates, and k_{damp}^{-1} below which GDM is overdamped, are also shown. We expect the constraints on c_s^2 and c_{vis}^2 to improve considerably and their degeneracy to be broken, if small-scale late-time structure formation data are combined with the CMB. However, to fully utilize these data would require an extension of the GDM model into the nonlinear regime. This is one of the motivations for the comparison of GDM to other models in the next section.

We remark that the case of negative c_s^2 has been studied in [74,76]. We do not think that it makes sense to consider negative c_s^2, c_{vis}^2 since they lead to exponential instabilities unless one fine-tunes the viscosity as discussed at the end of Sec. III B. We checked that, for small enough negative values $|c_s^2|, |c_{\text{vis}}^2| < 10^{-6}$, the numerical integration works and gives rise to reasonably looking results. Within some range of parameter space, the potential $\hat{\Phi}$ grows slightly

without exploding, but only when numerical integration is restricted to times and scales relevant for the CMB. This qualitatively new and phenomenologically interesting feature of growing potentials might be expected in alternative theories of gravity (see [135]) but not from dark matter. We therefore suggest using a parametrization suited for alternative theories of gravity for this purpose [83–85,135].

IV. CONNECTION BETWEEN COVARIANT IMPERFECT NONADIABATIC FLUIDS AND GDM

It is certainly possible that not all dark matter models can be brought into the GDM form. As one would like to use the GDM model to test alternative DM models and determine whether they are allowed or even favored by the CMB, we have to assess which realistic particle and field-based DM models can actually be brought into the GDM form. For instance, in the case of particle-based models, one concern may be that the phase-space distribution function $f_g(x^\mu, p_\nu)$ and its dynamics, as governed by the Boltzmann equation, does not allow for a truncation or closure of the hierarchy at $l_{\max} = 2$. In this case, additional cumulants of $f_g(x^\mu, p_\nu)$ beyond the first three (δ_g, θ_g and Σ_g) may be necessary. The collisionless case includes warm DM which can be described as GDM in the linear regime of structure formation [35]. For the collisional case parametrizations based on the Boltzmann equation were recently presented in [136,137]. We leave it to future work to investigate the connection of GDM to the phase space description of collisionless and collisional DM and therefore the connection of GDM to specific models of particle DM.

As we discuss below, if DM has internal degrees of freedom, then a GDM description may be possible in certain circumstances. Such is the case for nonequilibrium thermodynamics, the effective theory of fluids of Ballesteros [138] and the case of tightly coupled interacting adiabatic fluids. Alternatively, the GDM model may arise as an effective description of pure CDM once small-scale modes are integrated out [44], and lastly, as an effective fluid reformulation of scalar field models.

Most of these models have a nonperturbative definition. This is desirable if the model is to also be used in the mildly nonlinear and fully nonlinear regimes of structure formation. It is known that higher-order perturbation theory based on imperfect fluids improves the modeling of CDM in the mildly nonlinear regime [44,48]. Similarly it is known that even in the fully nonperturbative regime of structure formation, a self-gravitating scalar field is a viable alternative to particle dark matter [38].

A. GDM arising from thermodynamics

In this subsection, we consider nonequilibrium fluids that are close to thermal equilibrium, such that thermodynamic relations still hold. Fluids of this kind are well-known instances of imperfect fluids and therefore offer a

clear physical interpretation of the GDM parameters and serve as candidates for extensions of the GDM model into the nonlinear regime.

Fluids that are not in thermal equilibrium can develop (i) bulk viscosity, a special kind of nonadiabatic pressure proportional to $\nabla_\mu u^\mu$ that hampers the fluid expansion,[15] (ii) shear viscosity, proportional to the trace-free part of $\nabla_{(\mu} u_{\nu)}$ that impedes shearing flows, and (iii) diffusion flux, which is proportional to the gradient of a particular thermodynamic potential and which acts to smooth out those gradients. Bulk viscosity[16] arises when the collision times between particles are long, when the fluid consists of a mixture of relativistic and nonrelativistic particles, or when the particles have internal degrees of freedom [139–143]. Similarly, shear viscosity is related to the free-streaming time τ_c between collisions, with the photon-baryon fluid and freely streaming massless neutrinos being well-known examples of this.[17] Diffusion (or heat) flux exists whenever the energy flux is not exactly aligned with the particle flux, which happens for instance in the photon-baryon fluid at next-to-leading order in the tight-coupling approximation. Bulk and shear viscosity as well as diffusion flux are related to entropy production and give rise to imperfect terms in the energy-momentum tensor [140,146] and, as we will see, nonadiabatic pressure. It is also known that WIMP dark matter, although usually described as a pressureless perfect fluid, is better modeled by an imperfect fluid with shear and bulk viscosity, as well as pressure [43,147].

The shape of the energy-momentum tensor depends on the definition of the fluid four-velocity u^μ, in other words, the frame. A natural choice is the so-called energy or Landau-Lifshitz (LL) frame defined through $u_\alpha T^\alpha{}_\nu = -\rho u_\nu$. We adopt this frame throughout this paper, and we denote the corresponding four-velocity by u^μ. This frame enforces the constraint $u_\alpha \Sigma^\alpha{}_\nu = 0$, and prevents the occurrence of a term $q_{(\mu} u_{\nu)}$, where q_μ is the heat flux, in the energy-momentum tensor (2.1).

If a conserved particle current

$$N^\nu = n u^\nu + j^\nu \qquad (4.1)$$

exists, the equation of motion for n is found from

$$\nabla_\nu N^\nu = 0. \qquad (4.2)$$

In general, N^ν is not aligned with the energy flux ρu^ν, in which case the diffusion flux j^ν is nonzero. Since $n = -N^\mu u_\mu$ is the number density, we have $j^\nu u_\nu = 0$, such that $\bar{j}^\nu$ vanishes on FRW and δj^ν is gauge invariant in linear

perturbation theory. If N^ν exists and is nonzero, it is common to choose the Eckart frame [148] defined via $N^\nu = n n^\nu$, rather than the LL frame. The Eckart frame requires a term proportional to $q^{(\mu} n^{\nu)}$ to be added to the energy-momentum tensor (2.1).

Under a frame transformation given by a Lorentz boost, the four-velocity and the spatial vectors q_ν and j^ν do not remain invariant to linear order in the boost velocity, while all other functions entering N^ν and $T_{\mu\nu}$ remain frame invariant [149,150]. The combination

$$\tilde{q}_\nu = q_\nu - j_\nu \frac{\rho + p}{n} \qquad (4.3)$$

is also frame invariant to linear order in the boost velocity [149] and can therefore be interpreted as a frame-independent definition of the heat flux in linear perturbation theory. The information in the generalized heat flux $\tilde{q}_\nu$ is stored entirely in the diffusion flux j^ν when the LL frame is adopted, while in the Eckart formulation it is stored entirely in the (standard) heat flux q_ν. Note that q_μ directly enters the energy-momentum tensor $T_{\mu\nu}$, while j^ν does not. Therefore, whether a heat-diffusion-type departure from a perfect fluid is included in the energy-momentum tensor depends on the frame chosen. In the Eckart frame, $j^\nu = 0$ and hence N^ν assumes a perfect fluid form, while $T_{\mu\nu}$ develops the additional imperfect term $q_{(\mu} n_{\nu)}$. In the LL frame, $T_{\mu\nu}$ retains its perfect fluid form but N^ν receives imperfect corrections through j^μ.

We stick to the Landau-Lifshitz (or energy) frame unless otherwise stated. One good reason to choose the LL frame is that it always exists, regardless of the existence of a species with conserved particle number. There are some other good reasons for this choice, which will be discussed further below.

In the GDM model, if $c_s^2 \neq c_a^2$ the total GDM pressure cannot be obtained from a barotropic equation of state, i.e., $P \neq P(\rho)$. Therefore, it is natural to assume that the pressure has to depend on some other quantity as well, for instance, the particle number density n, the chemical potential μ, the temperature T, or the entropy S, such that $P = P(\rho, n, \mu, T, S, \ldots)$. The obvious complication with this idea is that the GDM model contains neither of those additional degrees of freedom. However, we will assume that the thermodynamic relations (4.4) are valid, allowing us to assume that the equation of state is given by $P = P(\rho, S)$ in the absence of bulk viscosity.

The main results of this section are that δS, although in general dynamical, is sourced only by $\hat{\Delta}_g$ [see (4.23)] and that a mapping to GDM is possible in two limits: (1) where the heat conduction vanishes and δS is nondynamical (see Sec. IV A 3) or (2) where the heat conduction becomes very large and the δS becomes algebraically related to $\hat{\Delta}_g$ (see Sec. IV A 4). That the pressure is, in general, dynamical is also expected from kinetic theory [151].

---

[15]Note that within this subsection we do not display the subscript g on the nonequilibrium imperfect fluid quantities for notational simplicity.

[16]Often also called second, volume or dilatational viscosity.

[17]See also [113,116,117,144,145] for a discussion of viscosity in a cosmological context.

1. Landau-Lifshitz imperfect fluid

Let us first review the LL imperfect fluid derivation adapted to our notation. We assume in the following that the thermodynamic relations (which are guaranteed to hold in local thermal equilibrium)

$$\rho + p = \mu n + Ts, \tag{4.4a}$$

$$d\rho = \mu dn + Tds \quad \text{(Gibbs relation)}, \tag{4.4b}$$

$$dp = nd\mu + sdT \quad \text{(Gibbs-Duhem relation)}, \tag{4.4c}$$

still hold in situations that are slightly off-equilibrium. Here, s is the entropy density, n is the conserved particle number density and p is the thermodynamic pressure. In the absence of bulk viscosity the thermodynamic pressure would equal P, which suggests the definition

$$P_{\text{bulk}} = P - p, \tag{4.5}$$

for the bulk pressure P_{bulk}.

The derivation of the LL imperfect fluid equations uses the conservation equations in the form $\nabla_\alpha T^\alpha{}_\nu = 0$ and $\nabla_\nu N^\nu = 0$. Making use of (4.4a), the energy-momentum tensor can be written as

$$T^\mu{}_\nu = (\mu n + Ts)u^\mu u_\nu + P_{\text{bulk}}u^\mu u_\nu + P\delta^\mu{}_\nu + \Sigma^\mu{}_\nu. \tag{4.6}$$

With the help of (4.2), (4.4c) and the normalization condition $u^\nu u_\nu = -1$, the expression for $u^\nu \nabla_\alpha T^\alpha{}_\nu$ gives

$$u^\nu \nabla_\alpha T^\alpha{}_\nu = -T\nabla_\nu(su^\nu) + \mu\nabla_\nu j^\nu + u^\nu\nabla_\alpha(\Sigma^\alpha{}_\nu + P_{\text{bulk}}q^\alpha{}_\nu), \tag{4.7}$$

where $q_{\mu\nu} = g_{\mu\nu} + u_\mu u_\nu$ is the projector on u^ν-orthogonal hypersurfaces. Energy-momentum conservation plus the identities $\Sigma^\alpha{}_\nu u^\nu = q^\alpha{}_\nu u^\nu = 0$ give rise to the evolution equation

$$\nabla_\nu S^\nu = -j^\nu\nabla_\nu\frac{\mu}{T} - \frac{1}{T}\Sigma^\alpha{}_\nu\nabla_\alpha u^\nu - \frac{P_{\text{bulk}}}{T}\nabla_\nu u^\nu, \tag{4.8}$$

for the entropy current,

$$S^\nu \equiv su^\nu - \frac{\mu}{T}j^\nu. \tag{4.9}$$

The definition of S^ν is suggested by the fact that it takes this form in local thermal equilibrium within kinetic theory [152].[18] In order to guarantee $\nabla_\nu S^\nu \geq 0$, Landau and Lifshitz postulate the following constitutive relations:

$$\Sigma_{\mu\nu} = -2\eta_{\text{LL}}\left(q^\alpha{}_\mu q^\beta{}_\nu - \frac{1}{3}q_{\mu\nu}q^{\alpha\beta}\right)\nabla_{(\alpha}u_{\beta)}, \tag{4.10a}$$

$$P_{\text{bulk}} = -\zeta_{\text{LL}}\nabla_\beta u^\beta, \tag{4.10b}$$

$$j_\mu = -\kappa_{\text{LL}}\left(\frac{nT}{\rho + p}\right)^2 q^\nu{}_\mu\nabla_\nu\frac{\mu}{T}$$
$$=: -\tilde{\kappa}_{\text{LL}}q^\nu{}_\mu\nabla_\nu\frac{\mu}{T}. \tag{4.10c}$$

The non-negative coefficients η_{LL}, ζ_{LL} and κ_{LL} are known as shear viscosity, bulk viscosity and heat conduction, respectively. In the last line we defined $\tilde{\kappa}_{\text{LL}}$ for later convenience.

We now briefly return to the discussion of the frame choice. It is not well known in the cosmology literature that the Eckart and LL theories are not equivalent [154]. This inequivalence points to a flaw of the theory of nonequilibrium thermodynamics since a physical state should never depend on a frame choice.[19] A remedy to this puzzle was recently put forward by Ván and Biró [155], where it was suggested to modify the thermodynamic relations (4.4) if the frame of the fluid and the frame of the thermometer that measures T are different from the LL frame. It was shown that a particular generalization of (4.4), explicitly containing the fluid and thermometer velocities, leads to a manifestly frame-covariant set of closure relations involving one equation for the frame-independent quantity $\tilde{q}_\mu$, rather than two separate equations for j_μ and q_μ as in [154]. This set of closure equations then reduces to (4.10) once the LL frame is chosen, while it does not reduce to the closure equations of Eckart in the Eckart frame. By modifying the thermodynamic relations (4.4) according to [155], the solution obtained in the Eckart frame can by mapped to a solution obtained in the LL frame through a boost, which immediately follows from the frame covariance of the conservation equations, as was shown in [155].

With the standard Gibbs relations (4.4), the Eckart frame leads to unphysical instabilities. Choosing the Eckart frame with the Gibbs relations of [155], however, leads to a stable solution that is not equivalent to the solution obtained by Eckart [117,148]. Support for the LL frame also comes from kinetic theory [156,157] and its stability properties compared to other frames [154,158,159].[20] The most conservative and reasonable frame choice therefore seems to us to be the LL frame [162].[21]

---

[18]It is exactly this equation that receives quadratic corrections Q^ν in causal nonequilibrium thermodynamics [149,153]. Note that in [154] it was proven that all first-order theories apart from LL are unstable, where first order here means that S^ν depends linearly on the energy-momentum tensor and the particle flux.

[19]For the same reason why a physical state cannot depend on the gauge choice. In both cases, frame and gauge choice, the mathematical result depends on these choices, but the physical state must be invariant.

[20]Although the LL theory contains superluminal effects, they are unimportant [158]. Making the theory causal [149,153] comes at the price of having more differential equations and free functions while giving rise to only unobservable small corrections compared to the LL theory [160,161].

Let us now continue with our task to connect the LL theory with the GDM model. One might wonder what the physical significance of the parameter κ_{LL} in (4.10c) is, since j_ν does not directly affect the energy-momentum tensor: Both n and j_ν can affect ρ and u^ν only via the equation of state $p = p(\rho, n)$, but they have no effect if the equation of state is barotropic $p = p(\rho)$. We discuss this further below where we perturb the LL theory around a FRW background.

2. Linear perturbations of the LL theory

From now on we set the bulk viscosity P_{bulk} to zero. This simplifies the notation in the following paragraphs and also makes it manifest that nonadiabatic pressure does not require bulk viscosity, which is, in any case, not part of the GDM model. Nevertheless bulk viscosity might not be negligible in some situations (see the discussion in [117] and [43,112] in the context of CDM); we plan to add this to GDM in future work.

In linear perturbation theory, taking into account only scalar modes, the LL closure relations (4.10) give

$$\Sigma = \frac{2\eta_{\mathrm{LL}}}{a\bar{\rho}(1+w)}\hat{\Theta}, \qquad (4.11a)$$

$$j = \tilde{\kappa}_{\mathrm{LL}}(\dot{\bar{\xi}}\theta - \delta\xi), \qquad (4.11b)$$

where we have defined the normalized chemical potential

$$\xi \equiv \frac{\mu}{T} \qquad (4.12)$$

and the gauge-invariant scalar perturbation j via $j_i = \vec{\nabla}_i j$. It is also useful to rewrite the thermodynamic relations in terms of ξ and the entropy per particle

$$S = \frac{s}{n} \qquad (4.13)$$

as

$$\rho + p = nT(S + \xi), \qquad (4.14a)$$

$$dS = (S + \xi)\left(\frac{d\rho}{\rho + p} - \frac{dn}{n}\right), \qquad (4.14b)$$

$$d\xi = (S + \xi)\left(\frac{dp}{\rho + p} - \frac{dT}{T}\right). \qquad (4.14c)$$

The entropy evolution equation (4.8) on a linearly perturbed FRW spacetime then reads $\nabla_\nu S^\nu = 0$ since both j_ν

and $\Sigma^\mu{}_\nu$ are spatial tensors and vanish at the background level. Explicitly this gives

$$\dot{\bar{S}} = 0, \qquad (4.15a)$$

$$\dot{\delta S} = -\frac{\tilde{\kappa}_{\mathrm{LL}}k^2}{a\bar{n}}(\bar{S} + \bar{\xi})(\dot{\bar{\xi}}\theta - \delta\xi). \qquad (4.15b)$$

The first result means that there is no entropy production within linear perturbation theory. This is a direct consequence of discarding bulk viscosity. Nonetheless, entropy perturbations are generally nonzero for nonvanishing κ_{LL} and are dynamical.

The perturbed LL equations are known to be relevant in cosmology: Heat conduction and shear viscosity have similar and equally important effects in the photon-baryon plasma. They are proportional to the mean free time of photons τ_c [117,129,130] giving rise to Silk damping of baryon acoustic oscillations [164]. The photon-baryon fluid is also an example where the bulk viscosity can be neglected, since its magnitude compared to the shear viscosity is suppressed by the large number of photons per baryon [117].

Equations (4.15) can only play a role in the evolution of the density and velocity perturbations if the pressure P also depends on S.[22] Assuming a general $P = P(\rho, S)$ we obtain

$$\dot{\bar{P}} = \left.\frac{\partial\bar{P}}{\partial\bar{\rho}}\right|_{\bar{S}}\dot{\bar{\rho}}, \qquad (4.16)$$

$$\delta P = \left.\frac{\partial\bar{P}}{\partial\bar{\rho}}\right|_{\bar{S}}\delta\rho + \left.\frac{\partial\bar{P}}{\partial\bar{S}}\right|_{\bar{\rho}}\delta S. \qquad (4.17)$$

Eliminating $\frac{\partial\bar{P}}{\partial\bar{\rho}}|_{\bar{S}}$ we find

$$\Pi = c_a^2\delta + \frac{1}{\bar{\rho}}\left.\frac{\partial\bar{P}}{\partial\bar{S}}\right|_{\bar{\rho}}\delta S, \qquad (4.18)$$

$$\Pi_{\mathrm{nad}} = \frac{1}{\bar{\rho}}\left.\frac{\partial\bar{P}}{\partial\bar{S}}\right|_{\bar{\rho}}\delta S. \qquad (4.19)$$

At this point we cannot conclude that c_a^2 is the sound speed since δS might have nontrivial dynamics similar to δ. In the absence of bulk viscosity, the entropy perturbation has the straightforward interpretation of a relative entropy between δ and δn,

---

[21]There are however different opinions on this matter: In [163] it was argued that the Eckart frame is more physical than the LL frame.

[22]We could equally assume an equation of state of the form $P = P(\rho, n)$, $P = P(\rho, \xi)$, $P = P(n, S)$, or any other combination of ρ, n, S, ξ, T. They can be shown to lead to identical results (4.25) and (4.27). To show this, Eq. (4.2) and the thermodynamic relations IV A 2 have to be employed, in particular, the Maxwell relations following from $ddS = 0$ and $dd\xi = 0$.

$$\delta S = (\bar{S} + \bar{\xi})\left(\frac{\delta}{1+w} - \frac{\delta n}{n}\right) \qquad (4.20)$$

[see (4.14b)], and therefore we expect that, in general, δS can modify the sound speed. The relation (4.20) shows that the relative entropy perturbation between ρ and n is in fact an entropy perturbation in the thermodynamic sense if the fluid is in a thermal state and also explains why Π_{nad} is known as the "entropy perturbation." A system in local thermal equilibrium defined by two state variables may equally be expressed by any other set of two linearly independent state variables due to the relations (4.4). We assume this property to be true also off equilibrium, such that we may assume $\xi = \xi(\rho, S)$ and therefore $d\xi = \frac{\partial \xi}{\partial \rho}|_S d\rho + \frac{\partial \xi}{\partial S}|_\rho dS$. On a linearly perturbed FRW spacetime without bulk viscosity, this leads to

$$\dot{\bar{\xi}} = -3\mathcal{H}\frac{\partial \bar{\xi}}{\partial \bar{\rho}}\bigg|_{\bar{S}}(1+w)\bar{\rho}, \qquad (4.21)$$

$$\delta \xi = \frac{\partial \bar{\xi}}{\partial \bar{\rho}}\bigg|_{\bar{S}}\bar{\rho}\delta + \frac{\partial \bar{\xi}}{\partial \bar{S}}\bigg|_{\bar{\rho}}\delta S. \qquad (4.22)$$

Inserting this into (4.15b) gives

$$\dot{\delta S} = \frac{\tilde{\kappa}_{\text{LL}}k^2}{\bar{n}a}(\bar{S} + \bar{\xi})\left(\bar{\rho}\frac{\partial \bar{\xi}}{\partial \bar{\rho}}\bigg|_{\bar{S}}\hat{\Delta}_g + \frac{\partial \bar{\xi}}{\partial \bar{S}}\bigg|_{\bar{\rho}}\delta S\right). \qquad (4.23)$$

This result shows that Π_{nad} in (4.19) is, in general, a dynamical degree of freedom and sourced by $\hat{\Delta}_g$. In the remainder of this section we will investigate under which conditions $\delta S = 0$ and $\delta S \propto \hat{\Delta}_g$ and therefore establish a connection to GDM.

3. GDM as a LL perfect fluid with a conserved particle number

For a perfect fluid, $\eta_{\text{LL}} = \zeta_{\text{LL}} = \kappa_{\text{LL}} = j = 0$, and (4.23) simplifies to

$$\dot{\delta S} = 0, \qquad (4.24)$$

showing that δS is constant in time and does not have a large impact on the dynamics of δ and θ. The dynamics of the perfect fluid variables ρ, u^μ with a general $P = P(\rho, S)$ are thus modeled by a particular GDM model where

$$w \equiv \frac{\bar{P}(\bar{\rho}, \bar{S})}{\bar{\rho}}, \qquad (4.25a)$$

$$c_s^2 = c_a^2 = \frac{\partial \bar{P}}{\partial \bar{\rho}}\bigg|_{\bar{S}}, \qquad (4.25b)$$

$$c_{\text{vis}}^2 = 0, \qquad (4.25c)$$

with a corresponding adiabatic sound speed (2.12). It is thus clear why $c_a^2 = \frac{\partial \bar{P}}{\partial \bar{\rho}}|_{\bar{S}}$ is called the adiabatic sound speed: It is calculated from a general equation of state with the entropy held fixed. The relation to GDM is a good approximation since δS is constant in time. Furthermore, the relation to GDM becomes exact for adiabatic initial conditions, i.e., $\delta S = 0$.

Note that in order to arrive at this result, we do not have to use any thermodynamic relations, and we could have equally derived (4.25) by assuming an equation of state $P = P(\rho, n)$ and showing that the particular combination $\delta/(1+w) - \delta n/\bar{n}$ is slowly varying compared to δ using the perturbed particle conservation equation (4.2) and the continuity equation (2.14). Therefore the result (4.25) holds for any perfect fluid with a conserved particle number and does not require the additional assumption of being in a thermal state. Also note that $c_s^2 = c_a^2$ holds even for nonlinear perturbations [165].

A discussion of the equation of state $p(\rho, S)$ of an ideal nonrelativistic gas in the context of cosmological perturbation theory can be found in [166].

4. GDM as a LL imperfect fluid with a conserved particle number

As we discussed above, although the GDM model lacks a particle conservation or alternatively an entropy evolution equation, it may still be used to describe a perfect fluid even for the case $P = P(\rho, S)$, since δS is either time independent or zero. It is clear, however, that the GDM model cannot, in general, describe an *imperfect* fluid completely, as in that case δS will be dynamical. Fortunately, as we show here, there are situations where the GDM model can be used to describe imperfect fluids as an approximation, by effectively removing the additional degree of freedom (usually associated with S) that is present in the LL theory.

Equation (4.23) may be solved using an approximation scheme analogous to the tight-coupling approximation for two interacting fluids (see Sec. IV E). In the limit of large $\tilde{\kappa}_{\text{LL}}$ the last bracket in (4.23) has to be parametrically smaller than $\dot{\delta S}$ in order for linear perturbation theory to apply. Therefore at leading order in an expansion in $\tilde{\kappa}_{\text{LL}}^{-1}$ the rest-frame density perturbations $\hat{\Delta}_g$ and the entropy perturbation δS become proportional to each other,

$$\delta S = -\frac{\frac{\partial \bar{\xi}}{\partial \bar{\rho}}\big|_{\bar{S}}}{\frac{\partial \bar{\xi}}{\partial \bar{S}}\big|_{\bar{\rho}}}\bar{\rho}\hat{\Delta}_g + \mathcal{O}(\tilde{\kappa}_{\text{LL}}^{-1}). \qquad (4.26)$$

Inserting this leading-order solution into (4.18) gives the GDM pressure equation (2.19a) with sound speed

$$c_s^2 = c_a^2 - \frac{\partial \bar{P}}{\partial \bar{S}}\bigg|_{\bar{\rho}}\frac{\frac{\partial \bar{\xi}}{\partial \bar{\rho}}\big|_{\bar{S}}}{\frac{\partial \bar{\xi}}{\partial \bar{S}}\big|_{\bar{\rho}}} = \frac{\partial \bar{P}}{\partial \bar{\rho}}\bigg|_{\bar{\xi}}. \qquad (4.27a)$$

Thus in the large $\tilde{\kappa}_{LL}$ limit the sound speed is given by $c_s^2 = \frac{\partial \bar{P}}{\partial \bar{\rho}}|_{\bar{\xi}}$ which should be contrasted with the perfect fluid case $\kappa_{LL} = 0$ where $c_s^2 = c_a^2 = \frac{\partial \bar{P}}{\partial \bar{\rho}}|_{\bar{S}}$. The other two GDM parameters, equation of state w and viscosity c_{vis}^2 arise as

$$w = \frac{\bar{P}(\bar{\rho}, \bar{\xi})}{\bar{\rho}}, \qquad (4.27b)$$

$$c_{vis}^2 = \frac{d_{IC} + 3}{2} \frac{\mathcal{H}\eta_{LL}}{a\bar{\rho}}, \qquad (4.27c)$$

where we have mapped the LL shear (4.11a) directly into the form of the algebraic GDM shear (2.29), which explicitly depends on the initial conditions (to remind the reader, for adiabatic initial conditions $d_{IC} = 2$).

Let us point out that it is only as a matter of convenience that we use the dimensionless[23] "viscosity speed" squared c_{vis}^2 rather than η_{LL} as it is precisely that combination of variables that appears in the phenomenology: First of all it is known that for freely streaming ultrarelativistic radiation $c_{vis}^2 = c_s^2 = 1/3$ [75]. In addition, as we have shown, for the algebraic shear the combination $c_s^2 + \frac{8}{15}c_{vis}^2$ determines the scale where the potential decays and that the effective sound speed is $c_{eff}^2 \simeq c_s^2 - \frac{2}{5}c_{vis}^2$ (see Sec. III).

Observe how (4.27a) offers an interpretation for Π_{nad} as the *thermodynamic* entropy perturbation, clearly deserving the name "entropy perturbation," which Π_{nad} is often referred to as. Since $\dot{S} = 0$, this is not necessarily related to entropy production, as the linearized entropy fluctuations average to zero when integrated over all space. However, entropy is indeed produced at second order in perturbation theory.

For nonrelativistic particles of mass m the chemical potential satisfies $S = m/T - \xi + 5/2$ such that $\dot{\bar{\xi}} = -m\dot{\bar{T}}/\bar{T}^2$ for $\dot{\bar{S}} = 0$; hence, a nonzero $\dot{\bar{\xi}}$ seems natural. However, it is less clear whether the large κ_{LL} limit can be naturally achieved in a dark matter model.

In closing this subsection, we remark that there are other approaches to nonequilibrium thermodynamics [167,168] or imperfect fluids [169] that might be better-suited candidates for an extension of GDM into the nonlinear regime of structure formation.

B. GDM arising from an effective theory of CDM large-scale structure

As the Einstein and fluid equations are intrinsically nonlinear, the FRW background and the linear perturbations should both be affected by the small-scale nonlinearities (backreaction), generating imperfect contributions to the CDM energy-momentum tensor as well as pressure [44]. We

therefore expect that the CDM background and linear perturbations should be described as GDM with (nonzero) GDM parameters that increase with time as the nonlinear scale grows in the late universe and that are approximately scale independent on the linear scales under consideration [170].

The form of the effective energy-momentum tensor can be derived through a coarse-graining of the microscopic equations (the lowest two moments of the Boltzmann hierarchy) and a subsequent gradient expansion [44–47]. In [44,171] it was argued that this leads to a LL-type imperfect fluid energy-momentum tensor whose time-dependent coefficients (equation of state, sound speed and viscosities) can be extracted by matching to the microscopic theory.

It was later emphasized in [46,172], that the effective energy-momentum tensor is a spatially local function of ρ, u^μ and the Riemann tensor because there exists a hierarchy of spatial scales $kv_p\tau_{fs} \ll 1$ (where $v_p \ll 1$ is the average particle velocity and τ_{fs} is the free-streaming time of a particle) such that $kv_p\tau_{fs} \ll 1$ means that scales of interest are larger than the mean free path. On the other hand, the stress-energy-momentum tensor cannot be a local function in time due to the absence of a temporal hierarchy of scales since the free-streaming time is of the same order of magnitude as the age of the universe $\tau_{fs}\mathcal{H} = \mathcal{O}(1)$.[24] Nevertheless a local-in-time approximation of the energy-momentum tensor turns out to be a good approximation for certain applications in perturbation theory [172,173].

The relevance of the effective field theory of large-scale structure (EFTofLSS) in the context of GDM is that it shows that even "ordinary" CDM has a FRW background and linear perturbations that are more completely described by an imperfect fluid with nonzero w, c_s^2 and c_{vis}^2, and a bulk viscosity term with parameter $c_{bulk}^2 = -\bar{P}_{bulk}/\bar{\rho}$.

As mentioned above, these GDM-type terms arise in the EFTofLSS because both linear perturbations and the background get renormalized by small-scale physics that has been integrated out. The numerical values and their time, scale and cosmology dependence (in particular, the normalization of the matter power spectrum) can be estimated using perturbation theory (see Appendix D of [44] and [171]), or more accurately using N-body simulations [47,172,174]. At $z = 0$

$$w, c_s^2, c_{vis}^2 \simeq \mathcal{O}(10^{-6}) \simeq (10 \times k_{nl}^{-1}\mathcal{H})^2, \qquad (4.28)$$

and they scale approximately with redshift like the variance of the peculiar velocity in linear perturbation theory

---

[23]We set the speed of light to 1.

[24]This is in contrast to a collisional fluid where usually a small mean free path $kv_p\tau_c \ll 1$ is accompanied by a small mean free time $\tau_{fs}\mathcal{H} \ll 1$ leading to an energy-momentum tensor that is temporally and spatially local function of the δ and θ.

$(fD\mathcal{H})^2$, where D is the linear growth function and $f = d\ln D/d\ln a$ the linear growth rate and $k_{\rm nl} \simeq 4.6\,{\rm h\,Mpc}^{-1}$ is the nonlinear scale below which the EFT breaks down [172].[25] The second relation in (4.28) shows that $k_{\rm d}^{-1} \simeq 10 k_{\rm nl}^{-1}$. Therefore, the largest characteristic scale $k_{\rm d}^{-1}$ of the imperfect fluid is within the range of validity of the effective theory.

We note that the shear in GDM is nonlocal in time since (2.19b) can be formally integrated, $\Sigma_g = \int^\tau g(\tau,\tau', \hat{\Theta}_g(\tau'))d\tau'$. Nonetheless, we saw that the qualitative behavior is well captured by the local-in-time algebraic version (2.29); see Figs. 1 and 2. In the EFTofLSS the stress tensor and therefore Π_g, Σ_g and bulk viscosity are nonlocal functions in time of $\hat{\Phi}$, $\hat{\Delta}_g$, $\hat{\Theta}_g$. We find a similar effect in our investigation of tightly coupled fluids further below; see Fig. 10.

Those two examples suggest that, for the search of signatures of pressure and imperfect fluid behavior of dark matter, it is sufficient to focus on one specific choice of parametrization of the stress tensor in terms of $\hat{\Phi}$, $\hat{\Delta}_g$, $\hat{\Theta}_g$ and a set of free functions: w and c_s^2 for the pressure and $c_{\rm vis}^2$ for the viscosity. We find in [93], using Planck and baryon acoustic oscillations data, that the constraints on c_s^2, $c_{\rm vis}^2 < \mathcal{O}(10^{-6})$ have a similar magnitude as the best-fitting parameters of the EFTofLSS. However, we note that the proximity of those numbers is an accident and has no immediate consequence for EFTofLSS. This is because we assumed parameters to be constant in time, while those of EFTofLSS decrease with increasing redshift, making the CMB less sensitive to EFTofLSS parameters at early times. Constraining GDM parameters with particular time dependence and via principal components is left to future work. Then it might be possible to measure the parameters of the EFTofLSS in data.

We also note that a similar approach for an EFT of LSS has been put forward in [48] where a parametric set of equations similar to the algebraic GDM model was used from the outset, albeit with a small difference (see Sec. II D). The shear viscosity was assumed to be of LL form (4.10a), such that the parametrization could be applied to higher-order perturbation theory.

C. GDM arising from scalar fields

Scalar fields have often been linked to effective fluids on a cosmological background. Here we reexamine this relation, connect it to the GDM model and discuss further possibilities beyond GDM. As it turns out, the effective behavior depends on whether the value of the scalar field ϕ crosses zero; hence we consider two possibilities

separately: a case with no oscillations in the background value of ϕ and the opposite.

1. No oscillations in the background value of ϕ

It is well known that quintessence scalar fields with a canonical kinetic term $X = -\frac{1}{2}\nabla_\mu\phi\nabla^\mu\phi$ and potential $V(\phi)$ can be described by an effective fluid. In the appendix of [75] it was already noted that a quintessence scalar field is described by a GDM model with an arbitrary (and, in general, time-dependent) equation of state $w = \frac{\bar{X}-V}{\bar{X}+V}$, sound speed $c_s^2 = 1$ and viscosity $c_{\rm vis}^2 = 0$.

A generalization of the standard quintessence field by introducing a noncanonical kinetic term $K(\phi, X)$, hence dubbed *k-essence*, was proposed in [175]. The action takes the form

$$\mathcal{I} = \int d^4x \sqrt{-g}\left[\frac{1}{16\pi G}R + K(\phi, X) + \mathcal{L}_{\rm m}\right]. \quad (4.29)$$

One may define a fluid velocity

$$\tilde{u}_\mu = -\frac{1}{\sqrt{2X}}\nabla_\mu\phi \quad (4.30)$$

provided $X > 0$ (and $\dot{\phi} > 0$). For instance, although this condition holds on a cosmological background, it does not hold in the static spherically symmetric case. Hence, the fluid description is not generally applicable in all situations. If $X > 0$, then it is clear that $\tilde{u}_\mu\tilde{u}^\mu = -1$, such that $\tilde{u}^\mu$ provides a natural vector field representing the fluid velocity. The frame defined by $\tilde{u}^\mu$ is called the scalar frame.

The association to a fluid is valid both on a FRW background and at the linear perturbation level, and this is sufficient to make a connection to GDM. The relevant variables are [175]

$$w = \frac{K}{2\bar{X}K_X - K}, \quad (4.31a)$$

$$c_s^2 = \frac{K_X}{2\bar{X}K_{XX} + K_X}, \quad (4.31b)$$

$$c_{\rm vis}^2 = 0, \quad (4.31c)$$

where $K_X \equiv \frac{\partial K}{\partial X}$. If $K(\phi, X) = X - V(\phi)$ then one recovers the quintessence case. Let us note that the sound speed in the k-essence case is, in general, time dependent; however, it is always spatially constant.

The k-essence model has traditionally been used in the context of inflation or dark energy. However, by carefully choosing K one can design models which are more suitable for dark matter. It was shown by Scherrer [176] that for shift-symmetric k-essence [$K = K(X)$ only], it is possible to obtain models which approach ΛCDM, albeit with

---

[25]We get the estimate (4.28) and the time dependence $(fD\mathcal{H})^2$ for c_s^2 by inspection of Eqs. (3,51,84) of [45]. That w and $c_{\rm vis}^2$ should be of the same order of magnitude as c_s^2 follows from Appendix D of [44].

$c_s^2 \approx 0$. In particular, for any $K(X)$ which has an extremum at $X = X_0$, we may expand it as $K(X) \approx K_0 + K_2(X - X_0)^2 + \cdots$. The field equations for ϕ may then be integrated once to get $\sqrt{X} K_X = F_0 a^{-3}$ where F_0 is an integration constant.[26] Then one obtains $\rho = -K_0 + 2F_0\sqrt{X_0}a^{-3}$ and $P = K_0 + \frac{F_0^2}{4K_2 X_0}a^{-6}$, which is valid as long as $F_0 K_2^{-1} X_0^{-3/2} a^{-3} \ll 1$.

Identifying $\rho_\Lambda = -K_0$ and separating out the cosmological constant leaves us with a GDM component with $\rho_g = \rho_{g,0} a^{-3}$ where $\rho_{g,0} = 2F_0\sqrt{X_0}$. The sound speed and equation of state obey the strict relation

$$c_s^2 = 2w = \frac{F_0}{4K_2 X_0^{3/2}} a^{-3} \qquad (4.32)$$

and are always time dependent. Thus, given K_2 and X_0, one can match the required GDM energy density today by choosing the integration constant F_0 appropriately. This in turn fixes w and c_s^2 completely.

For the k-essence action above, it may be shown that $\tilde{u}^\mu$ coincides with the LL velocity u^μ; however, this is not the case for more general actions of the Horndeski class. In [86] it was shown that more general scalar field actions necessarily lead to imperfect fluids and, in particular, the appearance of shear and bulk viscosities as well as heat flux. For instance, k-essence that is nonminimally coupled to gravity via a term $\int d^4x \sqrt{-g} \frac{1}{16\pi G} e^{\varkappa(\phi)} R$ in the action necessarily leads to bulk viscosity and is therefore a model beyond GDM. The addition of a cubic term $\int d^4x \sqrt{-g} G^{(1)}(\phi, X) \Box \phi$ in the action leads to a nonadiabatic pressure that is more general than the form considered here in (2.25); however, it still leads to zero shear just like k-essence. Nonzero shear arises when the quartic and quintic terms of the Horndeski action are included. It is unknown at the moment whether there exists a subset of the Horndeski action that is more general than k-essence, but which still conforms to the GDM template (with perhaps shear viscosity).

A different type of scalar field model that is not of the Horndeski class is the imperfect dark matter model [177], which extends the mimetic dark matter model of [178]. It seems plausible that it also has a close correspondence with GDM.

2. Background value of ϕ oscillates

If the background value of the scalar $\bar{\phi}$ is oscillating around a potential minimum, then the results (4.31) do not apply. This is because $\partial_\mu \bar{\phi}$ changes sign and $\bar{X}$ momentarily vanishes such that (4.30) is not a well-defined four-velocity. It was shown in [23,36] that oscillating scalar fields provide a working alternative to particle dark matter.

In the appendix of [75] it was pointed out that a GDM fluid may still provide an effective description if one averages the Einstein equations over several oscillation periods. A very interesting example is an oscillating real classical Klein-Gordon field with $P_g = K = X - m^2\phi^2/2$, which describes certain types of axion dark matter [179]. While the background expansion is identical to CDM on cosmologically relevant time scales, small perturbations around the Friedmann background behave like a fluid with nonadiabatic pressure [39–41]. The sound speed is only solution independent in the fluid comoving frame, the nonadiabatic pressure is of the GDM form [180] and the approximate mapping to GDM is given by

$$w = 0, \qquad (4.33a)$$

$$c_s^2 = \left(1 + \left(\frac{k}{2am}\right)^{-2}\right)^{-1} \simeq \left(\frac{k}{2am}\right)^2, \qquad (4.33b)$$

$$c_{\text{vis}}^2 = 0, \qquad (4.33c)$$

for scales much larger than the Compton wavelength $k \ll k_C \equiv am$. When the Klein-Gordon scalar ϕ is split into a slowly varying complex field ψ and a high frequency part e^{imt} [39,181], it is easy to see that ψ solves the Schrödinger-Poisson equation and that a dustlike behavior emerges above the Jeans scale

$$k_{\psi,\text{J}}^{-1} \simeq a^{-1}(G\bar{\rho}_g)^{-1/4}m^{-1/2}, \qquad (4.34)$$

which is the de Broglie wavelength of a k-mode of ψ.[27] It is guaranteed that there is a range of modes within the Jeans scale for which (4.33) applies if the envelope of ϕ is much smaller than the Planck mass [184]. A new method to numerically solve the Klein-Gordon equation without time averaging and without employing the nonrelativistic limit was developed in [185], where it was also implemented in the CMB code CLASS.

It is remarkable that for both the nonoscillating and oscillating background scalars, the nonadiabatic pressure is of GDM type, i.e., $C_1 = C_2 = 0$ in (2.26); see [86] and [180], respectively.

---

[26]The integration constant F_0 may easily be related to an initial condition for X at a specific initial time.

[27]The authors of [182] disagree with (4.33) and (4.34), and find that for an oscillating scalar field the Jeans scale is the Compton scale. Their approach does not involve averaging over time scales m^{-1}. Moreover, in [106] the authors argue, in Sec. VI-4, that the dynamics of scalar perturbations may be qualitatively different if averaged background quantities are used in the perturbation equations. Therefore, there appears to be no consensus on whether a GDM fluid with (4.33), and thus a Schrödinger field with (4.34), describes perturbations of an oscillating real Klein-Gordon field. However, the majority of the axion literature agrees with the view presented in this article, for instance, [37,39–41,57,59,179–181,183]. In particular, recent numerical studies [184,185] found (4.33) and (4.34) to be accurate for scales larger than the Compton wavelength.

D. GDM arising from effective field theory for fluids

In [138,186], the authors studied the class of actions of three scalars φ^a, $a = 1$, 2, 3, which are invariant under volume-preserving internal diffeomorphisms that send $\varphi^a \to \tilde{\varphi}^a(\varphi^b)$ with $\det(\partial \tilde{\varphi}^a / \partial \varphi^b) = 1$. See [187] for a review of the pullback formalism and [72] for applications to the coupling of dark matter to dark energy. The fields $\varphi^a(\tau, x^i)$ label the Lagrangian fluid volume elements such that $x^i(\tau, \varphi^a)$ are the trajectories of the fluid volume element labeled by φ^a. The assumed symmetry leads to the automatic conservation of the current

$$S^\mu \equiv s\breve{u}^\mu, \qquad (4.35)$$

where

$$s \equiv \sqrt{\det(B^{ab})}, \qquad B^{ab} \equiv g^{\mu\nu} \partial_\mu \varphi^a \partial_\nu \varphi^b, \qquad (4.36)$$

and the four-velocity $\breve{u}^\mu$ is defined as

$$\breve{u}^\mu \equiv -\frac{1}{6n} \epsilon^{\mu\alpha\beta\gamma} \tilde{\epsilon}_{abc} \partial_\alpha \varphi^a \partial_\beta \varphi^b \partial_\gamma \varphi^c, \qquad (4.37)$$

with the totally antisymmetric symbols having the conventions $\epsilon^{0123} = -1/\sqrt{-g}$ and $\tilde{\epsilon}_{123} = 1$. We discuss the physical meaning of the conserved current S^μ further below. Actions where φ^a is accompanied by only one derivative have been studied in [188–191] and give rise to perfect fluids without the need for Lagrange multipliers. They are therefore interesting starting points for general and consistent parametrizations of fluids.

In order to go beyond perfect fluids, more than one derivative per φ^a is necessary [138,192,193]. The most general action compatible with the assumed symmetry can be expanded in the number of gradients $\partial_\mu \ll \Lambda_c$ acting on each field, where Λ_c is the cutoff scale of the effective theory. At leading order (LO) and next-to-leading order (NLO) the most general action is [138]

$$\mathcal{I} = \int d^4x \sqrt{-g} \left[\frac{1}{16\pi G} R + F(s) + \frac{1}{\Lambda_c^2} \sum_{i=0}^{4} h_I(s) f_I + \mathcal{L}_{\mathrm{m}} \right], \qquad (4.38a)$$

where F and h_I are smooth functions of n and

$$f_0 = (g^{\mu\nu} + \breve{u}^\mu \breve{u}^\nu)\nabla_\mu \breve{u}^\alpha \nabla_\nu \breve{u}_\alpha, \qquad f_1 = (\breve{u}^\mu \nabla_\mu s)^2,$$
$$f_2 = \nabla_\mu s \nabla^\mu s, \qquad f_3 = \nabla_\mu \breve{u}^\nu \nabla_\nu \breve{u}^\mu,$$
$$f_4 = \epsilon^{\alpha\beta\mu\nu} \nabla_\alpha \breve{u}_\beta \nabla_\mu \breve{u}_\nu. \qquad (4.38b)$$

If h_4 were a constant, the term $\sqrt{-g} h_4 f_4$ would be a pure boundary term in (4.38). In general, $h_4(s)$ contributes only to vector modes, while the background evolution and scalar modes are unaffected [138], so we drop it in what follows. It was shown in [192] that S^μ, denoted there by $\mathcal{J}^\mu$, indeed fulfills the criteria of a conserved entropy current of a nonequilibrium thermal fluid. However, the somewhat unusual combination of imperfect stress-energy-momentum and a conserved S^μ means that the EFT of fluids describes nondissipative imperfect fluids [138,192]. Comparing to Sec. IV A, the imperfect contributions to the stress-energy-momentum tensor of LL theory are strictly dissipative, $\nabla_\mu S^\mu > 0$, while those of the EFT of fluids are strictly nondissipative, $\nabla_\mu S^\mu = 0$.[28] In order to simplify the subsequent discussion and to emphasize the connection to GDM, we set to zero the combination

$$h_0 + 3s^2(h_1 - h_2) + h_3 = 0, \qquad (4.39)$$

thereby eliminating all NLO corrections to the FRW background. This leaves only two free functions,

$$\alpha_{\mathrm{EFT}} - 1 = \frac{16\pi G}{\Lambda_c^2}(h_0 + h_3), \qquad (4.40)$$

$$\gamma_{\mathrm{EFT}} - 1 = \frac{48\pi G}{\Lambda_c^2} s^2 h_2, \qquad (4.41)$$

relevant for the scalar perturbations. The leading-order action, with all $h_I = 0$, gives rise to an adiabatic perfect fluid and has been used in the context of cosmology before [72,186]. By disregarding vector perturbations (and therefore 2 of the available 3 d.o.f. provided by the φ^a), and using the results for the background $\bar{\varphi}^a = \delta_j^a x^j$ and functional metric derivatives [187],

$$\frac{\delta \breve{u}^\mu}{\delta g_{\alpha\beta}} = \frac{1}{2} \breve{u}^\mu \breve{u}^\alpha \breve{u}^\beta, \qquad \frac{\delta s}{\delta g_{\alpha\beta}} = -\frac{s}{2}(g^{\alpha\beta} + \breve{u}^\alpha \breve{u}^\beta), \qquad (4.42)$$

we obtain $\bar{s} = a^{-3}$, the background energy density

$$\bar{\rho}_g = -\bar{F} = -F(\bar{s}),$$

and equation of state

$$w = -1 - \frac{1}{3} \frac{d \ln(-\bar{F})}{d \ln a}. \qquad (4.43a)$$

Note that $\bar{\rho}_g$ does not contain any time derivatives of the fields φ^a. Therefore, in contrast to conventional scalar field theories, no differential equation has to be solved for the background dynamics of any of the three fields φ^a.

---

[28]We thank G. Ballesteros for pointing this out to us.

The adiabatic sound speed c_a^2 is related to w as usual [(2.12)].

The scalar perturbations can be parametrized by a single scalar φ [138] as $\delta\varphi^a = \delta_j^a \vec{\nabla}_j \varphi$.[29] The number density perturbation $\delta_s = \delta s / \bar{s}$ and the entropy velocity perturbation then assume the following form:

$$\delta \breve{u}_i = -a \vec{\nabla}_i \left(\dot{\hat{\varphi}} + \frac{1}{2}\dot{\nu} + \zeta \right), \tag{4.44}$$

$$\delta_s = 3\eta + \vec{\nabla}^2 \hat{\varphi}, \tag{4.45}$$

where $\hat{\varphi} = \varphi - \nu/2$ is gauge invariant. These expressions agree with [138] if the conformal Newtonian gauge is chosen. The components of the perturbed energy-momentum tensor (2.7) take the form[30]

$$\delta_g = (1+w)\delta_s, \tag{4.46a}$$

$$\theta_g = \dot{\hat{\varphi}} + \frac{1}{2}\dot{\nu} + \zeta + \frac{(\bar{\gamma}_{\text{EFT}} - 1)\mathcal{H}}{8\pi G(1+w)a^2 \bar{\rho}_g}\hat{\Delta}_s, \tag{4.46b}$$

$$\Pi_g = c_a^2 \delta_g + \frac{\bar{\gamma}_{\text{EFT}} - 1}{24\pi G a^2 \bar{\rho}_g}\vec{\nabla}^2 \hat{\Delta}_s, \tag{4.46c}$$

$$\Sigma_g = \frac{[\dot{\bar{\alpha}}_{\text{EFT}} + 2(\bar{\alpha}_{\text{EFT}} - 1)\mathcal{H}]\dot{\hat{\varphi}} + (\bar{\alpha}_{\text{EFT}} - 1)\ddot{\hat{\varphi}}}{8\pi G a^2 \bar{\rho}_g (1+w)}, \tag{4.46d}$$

where we defined the gauge-invariant number density perturbation in the entropy frame

$$\hat{\Delta}_s = \delta_s + 3\mathcal{H}\left(\dot{\hat{\varphi}} + \frac{1}{2}\dot{\nu} + \zeta \right).$$

The LL frame $\delta u_i = -a\vec{\nabla}_i \theta_g$ agrees with the entropy frame $\delta\breve{u}_i$ only to LO, such that the NLO contribution to θ_g in

(4.46b) may be interpreted as heat flux.[31] The nonadiabatic pressure $\Pi_{\text{nad}} = \Pi_g - c_a^2 \delta_g$ is given by

$$\Pi_{\text{nad}} = \frac{\bar{\gamma}_{\text{EFT}} - 1}{24\pi G a^2 \bar{\rho}_g}\vec{\nabla}^2 \hat{\Delta}_s \tag{4.47}$$

and turns out to be proportional to the divergence of the heat flux. Since $\hat{\Delta}_s$ describes fluctuations of the entropy density, the name nonadiabatic pressure is justified. Since both Π_{nad} and Σ_g have no LO contribution, we can eliminate $\dot{\hat{\varphi}}$ and $\hat{\Delta}_s$ with their LO expressions: $\hat{\Theta}_g = \dot{\hat{\varphi}}$ and $\hat{\Delta}_g = (1+w)\hat{\Delta}_s$, to obtain closure equations for Π_{nad} and Σ_g in terms of $\hat{\Delta}_g$ and $\hat{\Theta}_g$. We get

$$\Pi_{\text{nad}} = \frac{\bar{\gamma}_{\text{EFT}} - 1}{24\pi G a^2 (1+w)\bar{\rho}_g}\vec{\nabla}^2 \hat{\Delta}_g, \tag{4.48a}$$

$$\Sigma_g = \frac{\bar{\alpha}_{\text{EFT}} - 1}{8\pi G a^2 (1+w)\bar{\rho}_g}\left[\left(\frac{\dot{\bar{\alpha}}_{\text{EFT}}}{\bar{\alpha}_{\text{EFT}} - 1} + \mathcal{H} \right)\hat{\Theta}_g \right.$$
$$\left. + \frac{c_a^2 \hat{\Delta}_g}{1+w} + \hat{\Psi} \right]. \tag{4.48b}$$

The NLO correction to the pressure takes exactly the GDM form. However, it is a particular subclass of all allowed Π_{nad} of GDM: the time dependence of $c_s^2 - c_a^2$ can be chosen freely via γ_{EFT}, but the scale dependence is fixed to $c_s^2 - c_a^2 \propto k^2$.

One limitation is that $\Pi_{\text{nad}} \ll \hat{\Delta}_g$ in order for the EFT expansion to be valid. This is not a problem if the EFT is applied to dark matter where we expect $\Pi_g \ll \delta_g$. The fact that nonadiabatic corrections to the sound speed are proportional to $(k/\mathcal{H})^2$ is a consequence of the EFT being a gradient expansion that describes a perfect fluid at leading order. Since $\bar{s} = a^{-3}$, choosing $F \propto s$ gives rise to $w = 0$, while any other w can be achieved by specifying an appropriate $F(s)$. Therefore, as in GDM, one is completely free to choose any time dependence of w.

In the expression for Σ_g [(4.48b)] we used the LO Euler equation to eliminate $\ddot{\hat{\varphi}}$. The shear cannot be brought into either dynamical or algebraic GDM form, i.e., proportional solely to $\hat{\Theta}_g$. Nonetheless, the coefficient of $\hat{\Theta}_g$ in (4.48b) can be matched with the algebraic GDM shear (2.29) such that, at least approximately,

---

[29]In [138] the symbol s is used for the scalar mode but here we use φ in order to avoid conflict with the entropy density s.

[30]We found a few typos in the equations of [138] and urge caution when comparing our results to that work. Two typos concern the shear: the right-hand side of Eq. (60) of [138] is missing an overall minus sign [restoring this sign in Eq. (60) makes the equation consistent with Eqs. (72) and (76)]. The right-hand side of Eq. (61) is missing a factor $1/a^2$. Restoring the factor $1/a^2$ and the missing minus sign makes both equations consistent with Eqs. (72) and (76) of [138] and our (4.46d). There is also a typo in the expression for the energy frame velocity perturbation θ, which in [138] is denoted by θ_R/k^2 and can be constructed from Eqs. (66), (67), (68) and (73). The resulting expression deviates from our (4.46b) by a sign difference in the NLO part. To check that our perturbed energy-momentum tensor (4.46) is correct, we confirmed that the continuity equation (2.14) is identically satisfied and that the Euler equation (2.15) agrees with the equation of motion for $\hat{\varphi}$.

[31]This heat flux is not related to $\tilde{q}_\mu$ (4.3), since there are no conserved particles in the EFT of fluids. In the LL frame the entropy current reads $S^\nu = su^\nu + Q^\nu$, where Q^μ is second order in deviations from thermal equilibrium, but with contributions linear in perturbations around FRW [149,192]. This should be contrasted to the corresponding expression in LL theory (4.9) for which $Q^\mu = 0$. If we do not insist on a thermodynamic interpretation of the EFT of fluids, then we can interpret the conserved current (4.35) as particle current N^μ and the four-velocity (4.37) as Eckart frame n^μ, in which case Q^ν simply becomes the diffusion flux j^ν.

$$w = -1 - \frac{1}{3}\frac{d\ln(-\bar{F})}{d\ln a}, \tag{4.49a}$$

$$c_s^2 = c_a^2 - \frac{\mathcal{H}^2}{24\pi G(1+w)a^2\bar{\rho}_g}(\bar{\gamma}_{\mathrm{EFT}} - 1)\left(\frac{k}{\mathcal{H}}\right)^2, \tag{4.49b}$$

$$c_{\mathrm{vis}}^2 = \frac{(d_{\mathrm{IC}}+3)\mathcal{H}}{32\pi G a^2\bar{\rho}_g}[\dot{\bar{\alpha}}_{\mathrm{EFT}} + \mathcal{H}(\bar{\alpha}_{\mathrm{EFT}} - 1)]. \tag{4.49c}$$

We note that the corresponding GDM model is, in general, not a good approximation as, once $\alpha_{\mathrm{EFT}} \neq 1$, the appearance of $\hat{\Delta}_g$ in (4.48b) will give rise to a modification of the effective sound speed. *A priori* this modification is as important as the corrections coming from γ_{EFT} in (4.48a) since both terms are of the form $k^4\delta_g$ in the Euler equation (2.15). However, in applications to dark matter, where c_a^2 and $\bar{\alpha}_{\mathrm{EFT}} - 1$ are both small, the leading departure from GDM is due to the $\hat{\Psi}$ term appearing in (4.48b) which is then very similar to the proposed $\Sigma_g^{\mathrm{extended,alg}}$ Eq. (2.30).

Cosmological perturbation theory with only h_2 nonzero in the NLO action has been studied in [193]. Since this violates our assumption (4.39) this theory corresponds to a different subclass of the full theory. In the most general case where $6n^2h_1 \neq (\Lambda_c^2/8\pi G)(\alpha_{\mathrm{EFT}} - \gamma_{\mathrm{EFT}})$ and therefore (4.39) does not hold, the background receives NLO corrections similar to bulk viscosity which also complicates the structure of Π_{nad}. The behavior of the general theory with independent h_I is beyond the scope of this paper, in which case both the background and the perturbations receive corrections reminiscent of bulk viscosity. In particular, (4.46a) ceases to hold, signaling the presence of intrinsic entropy perturbations.

E. GDM arising from two interacting adiabatic fluids

1. Definition of the model

General description.—Interacting fluids have been investigated in the context of dark matter coupled to one of the known species, for instance, neutrinos and photons [67–70] or to dark energy [71–73]. We do not follow this approach here, but a similar one where the interaction is assumed to be between two dark species, and we investigate whether their combined effect can be effectively described by GDM. This happens, for instance, if dark matter is tightly coupled to dark radiation as in [65,66].

It was shown in [194,195] how two perfect fluids can be combined into a single imperfect fluid with anisotropic stress and heat flow. This framework for creating an imperfect fluid from perfect fluids, however, is not useful in a situation where the background four-velocities of the constituent fluids are the same and the misalignment between them is purely perturbative. The situation where several fluids are coupled in linear perturbation theory is treated in [106,145,196,197] and is our starting point.

In the following we use the labels 1 and 2 for the two coupled adiabatic (but otherwise unspecified) fluids. For simplicity, we also assume that their respective equations of state are specified by constant-w parameters, w_1 and w_2, as this is sufficient to obtain a GDM-like pressure.

Our formulation closely follows [90], where an interaction of a dark matter and a dark energy component was studied in the so-called parametrized post-Friedmann framework. Here, however, we assume that the DE is uncoupled, and instead we use the coupled set of equations for the purpose of obtaining a combined GDM behavior, as we show further below. For all components, including the combined fluid, we assume the LL frame.

The combined stress-energy-momentum tensor is

$$T_g{}^\mu{}_\nu = T_1{}^\mu{}_\nu + T_2{}^\mu{}_\nu, \tag{4.50}$$

with $\nabla_\mu T_g{}^\mu{}_\nu = 0$. The stress-energy-momentum tensors of the two constituents are not individually conserved since the two constituents exchange energy and momentum via the current $J_{I\mu} \equiv -s_I J_\mu$. Here, $s_1 = 1$ and $s_2 = -1$, such that $\nabla_\mu T_2{}^\mu{}_\nu = J_\nu = -\nabla_\mu T_1{}^\mu{}_\nu$ and all other $I \neq 1$, 2 in (3.36) have $J_{I\mu} = 0$. For the two constituents, the coupling current J_ν can be split into a background part $Q \equiv \bar{J}_0$ (as $\bar{J}_i = 0$) and two linear scalar perturbations $q_{\mathrm{int}} \equiv \delta J_0$ and $\bar{\nabla}_i S_{\mathrm{int}} = \delta J_i$. Let us point out that although we do not specify the current J_ν nonperturbatively, the model with pure momentum exchange has a straightforward nonperturbative extension.

Equations of motion for the constituents.—The background current Q describes an energy transfer between the two components

$$\dot{\bar{\rho}}_I + 3\mathcal{H}\bar{\rho}_I(1+w_I) = s_I Q. \tag{4.51}$$

Perturbatively, each component's density contrast δ_I evolves according to

$$\dot{\delta}_I = -(1+w_I)\left[k^2(\theta_I - \zeta) + \frac{1}{2}\dot{h}\right] + \frac{s_I}{\bar{\rho}_I}[q_{\mathrm{int}} - Q\delta_I], \tag{4.52a}$$

while the momentum divergence θ_I evolves as

$$\dot{\theta}_I = -(1-3w_I)\mathcal{H}\theta_I + \frac{w_I}{1+w_I}\delta_I - \frac{2}{3}(k^2 - 3\kappa)\Sigma_I$$
$$+ \Psi + \frac{s_I}{\bar{\rho}_I(1+w_I)}[S_{\mathrm{int}} - Q(1+w_I)\theta_I]. \tag{4.52b}$$

The mixture variables.—For the mixture we define the total (background) density and pressure according to (4.50) as $\bar{\rho}_g = \bar{\rho}_1 + \bar{\rho}_2$ and $\bar{P}_g = \bar{P}_1 + \bar{P}_2$, respectively. The total equation of state w of the mixture is equal to the average equation of state over the two components

$$w = \sum_{I=1,2} r_I w_I \quad \text{with} \quad r_I = \frac{\bar{\rho}_I}{\bar{\rho}_g}. \tag{4.53}$$

Note that $0 < r_I < 1$ and $\sum_{I=1,2} r_I = 1$. Although each individual component has a constant-w equation of state, the mixture's equation of state is evolving so that its adiabatic sound speed is

$$c_a^2 = \frac{w_1 + R_{\text{mix}} w_2}{1 + R_{\text{mix}}} - \frac{w_{12} Q}{3(1+w)\mathcal{H}\bar{\rho}_g}, \tag{4.54}$$

where we defined $w_{12} = w_1 - w_2$ and

$$R_{\text{mix}} \equiv \frac{\bar{\rho}_2(1+w_2)}{\bar{\rho}_1(1+w_1)} = \frac{r_2(1+w_2)}{r_1(1+w_1)}. \tag{4.55}$$

With the above definitions the scalar perturbations of the mixture energy-momentum tensor (4.50) are related to the components through

$$\delta_g = \sum_{I=1,2} r_I \delta_I, \tag{4.56a}$$

$$(1+w)\theta_g = \sum_{I=1,2} r_I (1+w_I)\theta_I, \tag{4.56b}$$

$$\Pi_g = \sum_{I=1,2} r_I \Pi_I \to \sum_{I=1,2} r_I w_I \delta_I, \tag{4.56c}$$

$$(1+w)\Sigma_g = \sum_{I=1,2} r_I (1+w_I)\Sigma_I. \tag{4.56d}$$

From (4.54) and (4.56), or by making use of (3.37) with $Q_2 = -Q_1 = Q$, we find that the nonadiabatic pressure of the mixture, $\Pi_{\text{nad}} = \Pi_g - c_a^2 \delta_g$, is given by

$$\Pi_{\text{nad}} = w_{12}\left[\frac{Q}{3(1+w)\mathcal{H}\bar{\rho}_g}\hat{\Delta}_g + \frac{R_{\text{mix}}(1+w)}{(1+R_{\text{mix}})^2}S_{12}\right], \tag{4.57}$$

where the gauge-invariant variable S_{12} is defined by

$$S_{12} = \frac{\delta_1}{1+w_1} - \frac{\delta_2}{1+w_2} - Q\left[\frac{1}{\rho_1(1+w_1)} + \frac{1}{\rho_2(1+w_2)}\right]\theta_g. \tag{4.58}$$

The variables δ_g, θ_g, S_{12} and $\theta_{12} = \theta_1 - \theta_2$ provide a complete set of alternative dynamical variables describing the mixture.

2. Equations of motion for the combined fluid

The total background energy density of the mixture evolves as usual according to (2.10), while the combined variables δ_g and θ_g obey the usual uncoupled fluid equations (2.14) and (2.15), respectively.

The equations of motion for the new set of variables, S_{12} and θ_{12}, can be found in [106,197]. The latter reference contains the fully general equations where the constituent fluids are themselves allowed to have GDM-type non-adiabatic pressure. We adapt those equations here in the case of constant-w constituents. The equations of motion for the two difference variables S_{12} and θ_{12} follow from IV E 1 and are

$$\dot{S}_{12} = -k^2\theta_{12} - \frac{Q\bar{\rho}_g(1+w_1+w_2)}{\bar{\rho}_1\bar{\rho}_2(1+w_1)(1+w_2)}\hat{\Delta}_g + Q\left[\frac{1}{\bar{\rho}_2} - \frac{1}{\bar{\rho}_1}\right]S_{12}$$

$$+ Q\left[\frac{1}{\bar{\rho}_1(1+w_1)} + \frac{1}{\bar{\rho}_2(1+w_2)}\right] \times$$

$$\times \left[\frac{q_{\text{int}}}{Q} - \Psi + \left(\mathcal{H} - \frac{\dot{Q}}{Q}\right)\theta_g + \frac{2}{3}(k^2 - 3\kappa)\Sigma_g\right], \tag{4.59a}$$

for S_{12} and

$$\dot{\theta}_{12} = \left\{ \mathcal{H}[r_1(1+w_1)(3w_2-1) + r_2(1+w_2)(3w_1-1)] \right.$$

$$\left. - \frac{Qr_2(1+w_2)}{\bar{\rho}_1} + \frac{Qr_1(1+w_1)}{\bar{\rho}_2}\right\}\frac{\theta_{12}}{1+w} + \frac{w_{12}}{1+w}\hat{\Delta}_g$$

$$+ \frac{w_1 r_2(1+w_2) + w_2 r_1(1+w_1)}{1+w}S_{12} - \frac{2}{3}(k^2 - 3\kappa)\Sigma_{12}$$

$$+ \left[\frac{1}{\bar{\rho}_1(1+w_1)} + \frac{1}{\bar{\rho}_2(1+w_2)}\right](S_{\text{int}} - Q\theta_g), \tag{4.59b}$$

for θ_{12}, where $\Sigma_{12} = \Sigma_1 - \Sigma_2$.

Note that forming the pressure perturbation via $\Pi_g = c_a^2 \delta_g + \Pi_{\text{nad}}$ by using (4.57) and (4.54) results in

$$\Pi_g = c_a^2|_{Q=0}\delta_g + \frac{w_{12}Q}{\bar{\rho}_g}\theta_g + \frac{w_{12}R_{\text{mix}}(1+w)}{(1+R_{\text{mix}})^2}S_{12}. \tag{4.60}$$

This means that if $S_{12} = 0$ then the pressure assumes the GDM form (2.21) with $c_s^2 = c_a^2|_{Q=0} \neq c_a^2$, even though $Q \neq 0$. This is reminiscent of the thermodynamics studied in Sec. IV A 3, where the sound speed $c_s^2 = \frac{\partial P}{\partial \rho}|_S = c_a^2|_{S=\text{const}}$ even if the entropy S is not constant.

We now show whether and how GDM behavior emerges from the system of two interacting adiabatic fluids. We see that the Π_{nad} in (4.57), built out of sum and difference variables, already has a very suggestive form: If the second term were absent then we would be left with the exact GDM expression (2.24). The first term, however, disappears if $Q = 0$, and the only way to obtain a GDM-type Π_{nad} is to find $S_{12} \propto \hat{\Delta}_g$. We thus consider these two cases separately: with energy exchange $Q \neq 0$ and with no energy exchange ($Q = q_{\text{int}} = 0$).

3. Energy exchange: $Q \neq 0$

The suggestive form of Π_{nad} in (4.57) when $Q \neq 0$ indicates that when $S_{12} \to 0$ the GDM model is recovered. Since in this case there is net energy flow between the constituent fluids, this means that the two fluids are not in equilibrium and it is not surprising to find that $c_s^2 \neq c_a^2$.[32]

In order to effectively remove the S_{12} degree of freedom we assume that a situation exists where the two fluids are tightly coupled. In particular, assuming a tight-coupling relation of the form

$$q_{\mathrm{int}} = Q\left[\Psi - \left(\mathcal{H} - \frac{\dot{Q}}{Q}\right)\theta_g + \frac{2}{3}(k^2 - 3\kappa)\Sigma_g\right] + \frac{\bar{\rho}_g \mathcal{H} R_c}{(1+w)}S_{12},$$

(4.61)

where R_c is a tight-coupling parameter such that in the limit $R_c^{-1} \to 0$ and using (4.59a), the condition $S_{12}^{(0)} = 0$ is enforced to leading order in R_c^{-1}.

Since we work at lowest order in tight coupling, discarding all $\mathcal{O}(R_c^{-1})$ terms in Π_g, the θ_{12} degree of freedom does not enter. Thus, within this approximation, justifiable for the case $R_c^{-1} \ll Q/(\bar{\rho}_g \mathcal{H})$, we do not have to enforce $\theta_{12}^{(0)} = 0$ in addition to $S_{12}^{(0)} = 0$. Therefore, assuming only (4.61) we get at lowest order in R_c^{-1} a nonadiabatic pressure of the GDM form (2.24) resulting in

$$w = r_1 w_1 + r_2 w_2,$$

(4.62a)

$$c_s^2 = \frac{w_1 + R_{\mathrm{mix}} w_2}{1 + R_{\mathrm{mix}}} = c_a^2|_{Q=0}.$$

(4.62b)

The next order in R_c^{-1} introduces corrections in c_s^2 which depend on the (still) dynamical $\theta_{12}^{(0)}$, spoiling the GDM template. The situation is similar to the large-$\tilde{\kappa}_{\mathrm{LL}}$ limit in Sec. IV A 4, where δS becomes dynamical and the diffusion flux becomes nonzero at next to leading order in the expansion in $\tilde{\kappa}_{\mathrm{LL}}^{-1}$. Here we have the option to ensure that $\theta_{12}^{(0)} = 0$ in addition to $S_{12}^{(0)} = 0$, as discussed in Appendix B.

4. No energy exchange: $Q = q_{\mathrm{int}} = 0$

The $Q = q_{\mathrm{int}} = 0$ assumption is justified for the photon-baryon fluid tightly coupled through Thomson scattering,

---

[32]A situation of energy exchange exists for baryons after recombination and therefore outside the realm of the two-fluid GDM, when the Compton cooling of baryons modifies the baryon sound speed [133]. Another situation might be an interaction of dark matter with dark energy [90]. In those cases, however, there is no tight coupling. In addition, this would necessarily require an extension of the GDM model since the GDM component is not conserved; the baryons lose energy in the first scenario and the DM loses energy in the second.

when thermal equilibrium is assumed and justified [106], and second-order perturbative effects like thermalization of acoustic oscillations can be neglected [198]. Here, we take a more general approach which reduces to the photon-baryon case when $w_1 = 1/3$ and $w_2 = 0$.

Equations (4.59) simplify to

$$\dot{S}_{12} = -k^2 \theta_{12}$$

(4.63a)

and

$$\dot{\theta}_{12} = \mathcal{H}[r_1(1+w_1)(3w_2-1) + r_2(1+w_2)(3w_1-1)]\frac{\theta_{12}}{1+w}$$
$$+ \frac{w_{12}}{1+w}\hat{\Delta}_g + \frac{w_1 r_2(1+w_2) + w_2 r_1(1+w_1)}{1+w}S_{12} +$$
$$- \frac{2}{3}(k^2 - 3\kappa)\Sigma_{12} + \left[\frac{1}{\bar{\rho}_1(1+w_1)} + \frac{1}{\bar{\rho}_2(1+w_2)}\right]S_{\mathrm{int}},$$

(4.63b)

and Π_{nad} is determined by S_{12} as

$$\Pi_{\mathrm{nad}} = \frac{w_{12} R_{\mathrm{mix}}(1+w)}{(1 + R_{\mathrm{mix}})^2}S_{12}.$$

(4.64)

In order to proceed further, we need to specify the variable S_{int} in terms of other perturbations. Naturally we must have a term which imposes the tight-coupling condition $\theta_1 = \theta_2$ in a certain limit. Hence, without loss of generality we set

$$S_{\mathrm{int}} = -\mathcal{H}R_c(1+w_1)\bar{\rho}_1\theta_{12} + \tilde{S}_{\mathrm{int}} + \mathcal{O}(R_c^{-1}),$$

(4.65)

where $\tilde{S}_{\mathrm{int}}$ is still unspecified and R_c is a function of time only. Let us point out that the first term could be obtained from the following nonperturbative definition:

$$J^\nu = \frac{1}{3}R_c \nabla_\nu u^\nu (\rho_1 + P_1)(u_1^\mu - u_2^\mu).$$

The parameter R_c can be interpreted as collision efficiency related to the mean free time $\tau_c = 1/(R_c \mathcal{H})$, or opacity τ_c^{-1}. In the case of the photon-baryon fluid, S_{int} can be calculated from kinetic theory [106] and also leads to a friction term, like in (2.19b), for the shear Σ_g. In the limit $\mathcal{H}R_c \to \infty$ we get $\theta_{12} \to 0$, which is the tight-coupling condition. Hence to zeroth order in tight coupling, we have $\theta_{12}^{(0)} = 0$ such that (4.63a) gives $\dot{S}_{12}^{(0)} = 0$. This means that $S_{12}^{(0)}$ is a time-independent function that is related to a choice of initial conditions. We can choose adiabatic initial conditions such that $S_{12}^{(0)} = 0$. Hence to zeroth order, we find that the mixture is purely adiabatic, i.e., $\Pi_{\mathrm{nad}} = 0$.

In order to find the solution to order R_c^{-1}, we follow a similar approach as in the case of the photon-baryon fluid.

From (4.55) it follows that $1 + R_{\text{mix}} = (1 + w)/[r_1(1 + w_1)]$. We then rearrange the $\dot{\theta}_{12}$ in Eq. (4.63b) to get, to lowest order in R_c^{-1},

$$\theta_{12}^{(1)} = \frac{R_{\text{mix}}}{(1 + R_{\text{mix}})\mathcal{H}R_c} \left[\frac{w_{12}}{1 + w} \hat{\Delta}_g - \frac{2}{3}(k^2 - 3\kappa)\Sigma_{12} \right.$$
$$\left. + \frac{1 + R_{\text{mix}}}{\bar{\rho}_2(1 + w_2)} \tilde{S}_{\text{int}} \right]. \tag{4.66}$$

This is the next-to-leading-order correction to the tight-coupling solution. Inserting the above equation into the one for $\dot{S}_{12}$ gives the first correction for S_{12} as

$$S_{12}^{(1)} = -k^2 \int_0^\tau d\tau' \frac{R_{\text{mix}}}{(1 + R_{\text{mix}})\mathcal{H}R_c} \left[\frac{w_{12}}{1 + w} \hat{\Delta}_g \right.$$
$$\left. - \frac{2}{3}(k^2 - 3\kappa)\Sigma_{12} + \frac{1 + R_{\text{mix}}}{\bar{\rho}_2(1 + w_2)} \tilde{S}_{\text{int}} \right], \tag{4.67}$$

where the integrand is evaluated at time τ'.

It does not seem that (4.67) reproduces the GDM pressure relation even though the appearance of the rest frame density perturbation $\hat{\Delta}_g$ is promising. We have already argued that it is natural that Π_{nad} is dynamical, in other words, a temporally nonlocal function of δ_g and θ_g. The effect of the nonlocality of S_{12} is that Π_{nad} is slightly out of phase with $\hat{\Delta}_g$ in the acoustic regime, leading to damping in addition to viscosity, which can be interpreted as heat-diffusion flux; see Sec. IV A 4 and [117,130,154,199,200].

To get an idea of how well $S_{12}^{(1)}$ approximates the exact S_{12}, we can study the case where the two fluids are given by photons $(I = 1)$ and baryons $(I = 2)$ that are tightly coupled via Thomson scattering before recombination. The variable R_c in this case can be calculated from kinetic theory [106]

$$\tau_c^{-1} = \mathcal{H}R_c = an_e\sigma_T \equiv X_e a^{-2}\tilde{\sigma}_T, \tag{4.68}$$

where σ_T is the Thomson cross section, n_e is the number density of free electrons and $X_e = n_e/(n_H + n_p)$ is the free electron fraction. The last equality defines $\tilde{\sigma}_T = a^3\sigma_T n_e/X_e \sim 2.3048 \times 10^{-5}(1 - Y_{\text{He}})\omega_b$ for helium fraction Y_{He} and dimensionless baryon density ω_b. The resulting equation for $\theta_{\gamma b} = \theta_{12}$ agrees with [102,105].[33]

In this case $\tilde{S}_{\text{int}} = 0$ while [102,105]

$$\Sigma_g^{(1)} = \frac{8}{15\mathcal{H}R_c} \hat{\Theta}_g \sim \mathcal{O}(R_c^{-1}), \tag{4.69}$$

which is of the algebraic GDM (or the LL) shear form. In Fig. 8 we compare the exact numerical solution from CLASS (see [105]) to $\Sigma_g^{(1)}$.

---

[33]Note that the variable R used in [102,105] is the reciprocal of our $R_{\text{mix}} = R^{-1}$.
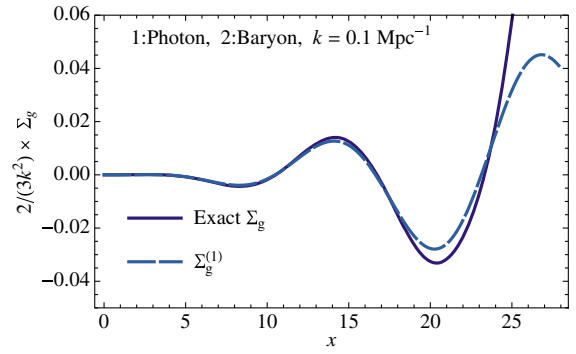

FIG. 8. Comparison of the exact shear $\Sigma_g = \Sigma_\gamma$, when GDM is set to describe the tightly coupled photon-baryon fluid to its algebraic approximation that arises at next-to-leading order in tight coupling.

Since the integrand is suppressed by $(\mathcal{H}R_c)^{-1}$, we expect S_{12} to be small and thus that the sound speed will be nearly adiabatic. On scales larger than the sound horizon, $\hat{\Delta}_g = D(\tau)\hat{\Delta}_g^i$ such that

$$S_{12} = -\frac{1}{12}k^2 \left[\frac{1}{D(\tau)} \int_0^\tau d\tau' \frac{R_{b\gamma}(3 + 4R_{b\gamma})}{(1 + R_{b\gamma})^2} \frac{D(\tau')}{\mathcal{H}R_c} \right] \hat{\Delta}_g, \tag{4.70}$$

where

$$R_{b\gamma} = R_{\text{mix}} = \frac{3\bar{\rho}_b}{4\bar{\rho}_\gamma} = \frac{3S_b f_{mr}}{4S_\gamma} \frac{a}{a^i} \tag{4.71}$$

for the baryon-to-photon ratio. Here, f_{mr} is the ratio of energy density in the form of nonrelativistic matter and relativistic matter, S_b is the fraction of nonrelativistic matter in the form of baryons and similarly S_γ is the fraction of relativistic matter in the form of photons at some initial time with scale factor a^i. Figure 9 compares the exact and approximate solution for S_{12} for a single wave number $k = 0.1$ Mpc^{-1}. Instead of conformal time τ, we use the time-dependent R_c^{-1} as a time variable on the x axis. We see that $S_{12} = 0$ initially, and until $R_c^{-1} = 0.005$ both solutions agree well.

Having determined the form of S_{12}, the GDM functions are found to be

$$w = \frac{1}{3 + 4R_{b\gamma}}, \qquad c_a^2 = \frac{1}{3(1 + R_{b\gamma})}, \tag{4.72a}$$

$$c_s^2 = c_a^2 - \frac{k^2 R_{b\gamma}}{9(1 + R_{b\gamma})(3 + 4R_{b\gamma})}$$
$$\times \frac{1}{D(\tau)} \int_0^\tau d\tau' \frac{R_{b\gamma}(3 + 4R_{b\gamma})D(\tau')}{(1 + R_{b\gamma})^2 \mathcal{H}R_c}. \tag{4.72b}$$
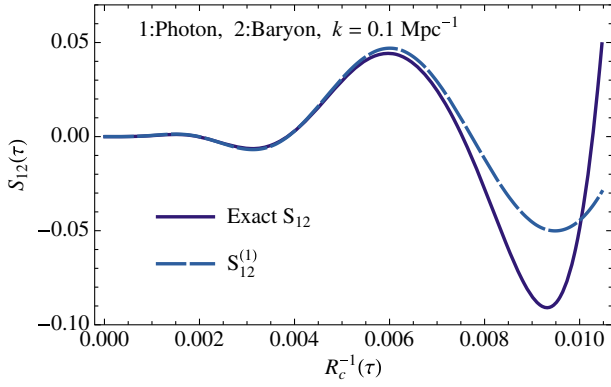
FIG. 9.　Comparison of the exact solution S_{12} of (4.63) and the next-to-leading-order solution $S_{12}^{(1)}$ (4.67) for the photon-baryon case. The x axis uses R_c^{-1} as time rather than τ to give an idea of how well the tight-coupling solution works given some value of R_c. The right end of the R_c^{-1} axis corresponds to recombination.
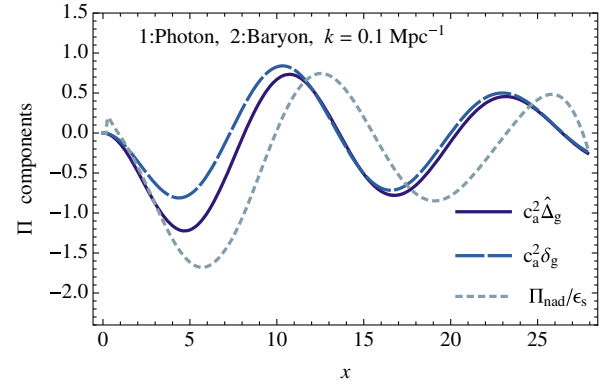

FIG. 10.　Comparison of the adiabatic and nonadiabatic components of Π_g, if GDM is the tightly coupled photon-baryon fluid. We normalized Π_{nad} by ϵ_s. This makes it visible when compared to $c_a^2\delta_g$ and also shows that ϵ_s, which was estimated in the limit $x \ll 1$, works well also for large x. The kink at $x = 0.2$ is caused by CLASS using $S_{\gamma b} = 0$ until $\tau = 2$ Mpc.

The key lesson from the photon-baryon example is that a situation where a dark matter species is tightly coupled to dark radiation [65] can be described as a GDM. It also shows that we only expect mild deviations from the adiabatic sound speed.

It is interesting to note that in the effective theory of fluids [138], the nonadiabatic pressure has exactly the same form $c_s^2 - c_a^2 \propto k^2$. To judge the importance of this term, we can estimate the ratio $(c_s^2 - c_a^2)/c_a^2$. In other words, switching to the dimensionless variable $x = k\tau$, we need to evaluate $\epsilon_s \equiv \frac{c_s^2}{c_a^2} - 1$, which is given analytically by

$$\epsilon_s = -\frac{R_{b\gamma}}{3(3 + 4R_{b\gamma})}\frac{1}{D(\tau)}\int_0^x dx'\, \frac{R_{b\gamma}(3 + 4R_{b\gamma})D(\tau')}{(1 + R_{b\gamma})^2 \mathcal{H}_k R_c},$$
(4.73)

and determine its size.

We now expand (4.73) in small $x = k\tau$ and use the adiabatic initial conditions from Appendix III A 4 in order to get

$$\epsilon_s \to -\frac{k}{32X_e\tilde{\sigma}_T}\left(\frac{\lambda_k^2 a^i S_b}{f_{mr}S_\gamma}\right)^2 x^5,$$
(4.74)

thus ϵ_s scales as x^5. How big or small it is in the early Universe depends on the constants we need to include. Assuming $X_e \sim 1$ and standard cosmological parameters we find

$$\epsilon_s \simeq -0.024\left(\frac{k}{0.1 \text{ Mpc}^{-1}}\right)^6\left(\frac{\tau}{\tau_{\mathrm{rec}}}\right)^5,$$
(4.75)

where $\tau_{\mathrm{rec}} \approx 281$ Mpc is the conformal time of recombination where the tight-coupling approximation breaks down.

In Fig. 10 we compare the two components of Π_g, the adiabatic $c_a^2\delta_g$ and the nonadiabatic Π_{nad} [(4.64)]. We divided Π_{nad} by ϵ_s for two reasons. First, note that ϵ_s was derived in the limit $x \ll 1$ in which Π_{nad} can be written as $\Pi_{\mathrm{nad}} = c_a^2\epsilon_s\hat{\Delta}_g$ such that only in this limit $c_a^2\hat{\Delta}_g$ and $\Pi_{\mathrm{nad}}/\epsilon_s$ are expected to agree. But as is clear from Fig. 10 their magnitudes still agree for $x \gg 1$. Thus, ϵ_s is a good proxy for the relative importance of Π_{nad}. Second, we observe that although Π_{nad} has a slightly shifted phase compared to δ_g, it might still be a good approximation to assume that $\Pi_{\mathrm{nad}} \simeq c_a^2\epsilon_s\hat{\Delta}_g$. The damping caused by the Π_{nad} being slightly out of phase with $\hat{\Delta}_g$ could be taken into account by adjusting c_{vis}^2.

V. CONCLUSION

We have presented an extensive investigation of the generalized dark matter model, first proposed by Hu [75]. The GDM model extends the commonly used pressure-less perfect fluid that describes cold dark matter in a linearly perturbed FRW universe. GDM describes a phenomenological imperfect fluid with two particular closure equations (2.19) and three parametric functions: its equation of state w, sound speed c_s^2 and viscosity c_{vis}^2. Note that CDM is recovered for $w = c_s^2 = c_{\mathrm{vis}}^2 = 0$. We placed strong constraints on these parameters in a companion paper [93], finding them to be consistent with CDM.

We have calculated the adiabatic and isocurvature initial conditions, and these are presented in Sec. III A and in Appendix A. To understand the imprints of the GDM model parameters on the CMB, we analytically analyzed a simplified yet very similar version of GDM (2.29) and found that the evolution of the gravitational

potential in a GDM dominated universe with small w, c_s^2 and c_{vis}^2 is mainly determined by c_s^2 and c_{vis}^2. For physical values of these parameters $(c_s^2, c_{vis}^2 \geq 0)$, they can only cause the gravitational potential to decay and not to grow. This decay occurs on scales below $k_d^{-1} \simeq \tau \sqrt{c_s^2 + 8c_{vis}^2/15}$; see (3.25). The parameters c_s^2 and c_{vis}^2 cause further, less degenerate, effects at the Jeans (3.30) and damping (3.31) scales, which are both on smaller scales. We expect the CMB to be less sensitive to these smaller scales.

We numerically investigated the CMB power spectra in Sec. III D. We found that c_s^2 and c_{vis}^2 appear to be very degenerate in all CMB power spectra with adiabatic initial conditions, consistent with the expectation from above that the CMB is mostly affected by c_s^2 and c_{vis}^2 through the combination k_d^{-1}. The decay of the potential below this scale predominantly affects the CMB through the ISW effect and lensing. The effect of the equation of state w on the CMB spectra can be understood through its effect on the time of radiation-matter equality.

We also investigated several alternatives to the GDM model (see Sec. IV), most of which are defined nonperturbatively. In principle, nonperturbative models such as these are able to describe the nonlinear regime of structure formation.

Thus, these models may be useful to look for signatures beyond CDM in data like in [201,202] that probe the mildly nonlinear and nonlinear regimes. Similarly, these models can be employed in forecasts of GDM parameter constraints that will be possible in the future with LSS surveys like Euclid [99]. We leave GDM constraints and forecasts involving nonlinearities for a future study.

In this paper we focused on the linear regime and showed how these models are related to GDM and, when possible, how these models can be mapped to the GDM parametric functions. In total we examined five models: We considered the theory of nonequilibrium thermodynamics of Landau and Lifshitz and pointed out that the presence of a conserved particle current and its perturbations can be accounted for by GDM in the perfect fluid limit (4.25) or when the heat conduction is very large (4.27). We presented the mapping to GDM parameters if DM is modeled by a monotonously moving (4.31) or oscillating scalar field (4.33). The latter case is important if DM is a low mass axion. We also investigated the imperfect fluid arising at next to leading order in an effective field theory expansion based on the pullback formalism of fluids and found that a certain subclass of this theory can be modeled by GDM (4.49). According to the EFTofLSS [44], even CDM develops imperfect fluid behavior on linear scales. Here, we clarified the connection of EFTofLSS parameters to GDM (4.28). Finally, we considered the case where two fluids are tightly coupled and therefore can be described by a single fluid.

In the tight-coupling limit with energy exchange, this combined fluid has a nonadiabatic pressure of GDM form (2.24) with GDM parameters (4.62). This two-fluid model is the only model considered here that is not defined nonperturbatively. However, other two-fluid models can be defined nonperturbatively, such as the model in Sec. IV E 4 possessing only momentum exchange.

APPENDIX A: INITIAL CONDITIONS FOR SCALAR MODES

1. Einstein and fluid equations

For convenience, we multiply all equations with suitable powers of $\tilde{a}$ [see (3.5)] in order to avoid inverse powers of x once we insert our anatz (3.8). The resulting necessary equations used in the calculation of the initial condition modes are as follows:

00-Einstein equation,

$$
\begin{aligned}
\tilde{a}\tilde{a}'h' &- 2\tilde{a}^2\eta \\
&= 3S_\gamma\delta_\gamma + 3S_\nu\delta_\nu \\
&+ 3\lambda_k\tilde{a}\left\{ S_g\left[1 - 3w\ln\left(\frac{\lambda_k}{f_{mr}}\tilde{a}\right)\right]\delta_g + S_c\delta_c + S_b\delta_b \right\}.
\end{aligned}
$$
(A1a)

0i-Einstein equation,

$$
\begin{aligned}
2\tilde{a}^2\eta' &= 3\lambda_k\tilde{a}\left\{ S_g\left[1 + w - 3w\ln\left(\frac{\lambda_k}{f_{mr}}\tilde{a}\right)\right]v_g + S_b v_\gamma \right\} \\
&+ 4S_\gamma v_\gamma + 4S_\nu v_\nu.
\end{aligned}
$$
(A1b)

CDM continuity,

$$
\delta_c' = -\frac{1}{2}h'.
$$
(A1c)

Baryon continuity,

$$
\delta_b' = -v_\gamma - \frac{1}{2}h'.
$$
(A1d)

Photon continuity,

$$\delta_\gamma' = -\frac{4}{3}v_\gamma - \frac{2}{3}h'. \tag{A1e}$$

Baryon-photon Euler,

$$\left(1 + \frac{3}{4}\frac{S_b}{S_\gamma}\lambda_k\tilde{a}\right)v_\gamma' = \frac{1}{4}\delta_\gamma - \frac{3}{4}\frac{S_b}{S_\gamma}\lambda_k\tilde{a}'v_\gamma. \tag{A1f}$$

Neutrino continuity,

$$\delta_\nu' = -\frac{4}{3}v_\nu - \frac{2}{3}h'. \tag{A1g}$$

Neutrino Euler,

$$v_\nu' = \frac{1}{4}\delta_\nu - \sigma_\nu. \tag{A1h}$$

Neutrino closure,

$$\sigma_\nu' = \frac{4}{15}v_\nu + \frac{2}{15}(h' + 6\eta'). \tag{A1i}$$

GDM continuity,

$$\tilde{a}^4\delta_g' = 3\tilde{a}^3\tilde{a}'(w - c_s^2)\delta_g - (1 + w)[\tilde{a}^4 + 9(\tilde{a}\tilde{a}')^2(c_s^2 - w)]v_g$$
$$-\frac{1}{2}\tilde{a}^4(1 + w)h'. \tag{A1j}$$

The GDM Euler equation,

$$\tilde{a}v_g' = (3c_s^2 - 1)\tilde{a}'v_g + \tilde{a}\left(\frac{c_s^2}{1 + w}\delta_g - \sigma_g\right). \tag{A1k}$$

GDM shear equation,

$$\tilde{a}\sigma_g' + 3\tilde{a}'\sigma_g = \tilde{a}\frac{8c_{\rm vis}^2}{3(1 + w)}\left(v_g + \frac{1}{2}h' + 3\eta'\right). \tag{A1l}$$

where we have used (2.19) to substitute for the GDM pressure.

2. Isocurvature modes

Here we list all nondecaying isocurvature modes. These are the radiation-type neutrino isocurvature density (NID) and neutrino isocurvature velocity (NIV) and the matter-type CDM isocurvature (CI), baryon isocurvature (BI) and GDM isocurvature (GI).

a. Neutrino isocurvature density

Setting $\delta_{\nu,0} = 1 = -\frac{S_\gamma}{S_\nu}\delta_{\gamma,0}$ and all remaining perturbations in $\mathcal{I}_{\rm modes}$ [(3.10)] to zero, the neutrino isocurvature density mode is

$$\eta = -\frac{S_\nu}{6(15 + 4S_\nu)}x^2, \qquad h = \frac{S_bS_\nu}{40S_\gamma}\lambda_kx^3,$$

$$\delta_c = -\frac{1}{2}h, \qquad \delta_b = \frac{S_\nu}{8S_\gamma}x^2,$$

$$\delta_\gamma = -\frac{S_\nu}{S_\gamma}\delta_\nu, \qquad \delta_\nu = 1 - \frac{x^2}{6},$$

$$\delta_g = \frac{S_\nu}{5}x^2\left[\frac{3c_{\rm vis}^2(w - c_s^2)}{15 + 4S_\nu} - \frac{S_b(1 - c_s^2 + 2w)}{16S_\gamma}\lambda_kx\right],$$

$$v_\gamma = -\frac{S_\nu}{S_\gamma}v_\nu, \qquad v_\nu = \frac{1}{4}x, \qquad v_g = \frac{2c_{\rm vis}^2S_\nu}{15(15 + 4S_\nu)}x^3,$$

$$\sigma_\nu = \frac{1}{2(15 + 4S_\nu)}x^2, \qquad \sigma_g = -\frac{8c_{\rm vis}^2S_\nu}{15(15 + 4S_\nu)}x^2.$$

b. Neutrino isocurvature velocity

Setting $v_{\nu,0} = 1 = -\frac{S_\gamma}{S_\nu}v_{\gamma,0}$ and all remaining perturbations in $\mathcal{I}_{\rm modes}$ [(3.10)] to zero, the neutrino isocurvature velocity mode is

$$\eta = -\frac{4S_\nu}{3(5 + 4S_\nu)}x, \qquad h = \frac{3S_b}{8S_\gamma}\lambda_kS_\nu x^2,$$

$$\delta_c = -\frac{1}{2}h, \qquad \delta_b = \frac{S_\nu}{S_\gamma}x,$$

$$\delta_\gamma = -\frac{S_\nu}{S_\gamma}\delta_\nu, \qquad \delta_\nu = -\frac{4}{3}x,$$

$$\delta_g = \left[\frac{8c_{\rm vis}^2(w - c_s^2)}{5 + 4S_\nu} - \frac{3(2 - 3c_s^2 + 5w)S_b}{32S_\gamma}\lambda_kx\right]S_\nu x,$$

$$v_\gamma = -\frac{S_\nu}{S_\gamma}\left(1 - \frac{3\lambda_kS_b}{4S_\gamma}x\right), \qquad v_\nu = 1 - \frac{9 + 4S_\nu}{6(5 + 4S_\nu)}x^2,$$

$$v_g = \frac{8c_{\rm vis}^2S_\nu}{9(5 + 4S_\nu)}x^2,$$

$$\sigma_\nu = \frac{4}{3(5 + 4S_\nu)}x, \qquad \sigma_g = -\frac{8c_{\rm vis}^2S_\nu}{3(5 + 4S_\nu)}x.$$

c. CDM isocurvature

Setting $\delta_{c,0} = 1$ and all remaining perturbations in $\mathcal{I}_{\rm modes}$ [(3.10)] to zero, the CDM isocurvature mode is

$$\eta = -\frac{1}{6}\lambda_k S_c x, \qquad h = \lambda_k S_c x,$$

$$\delta_c = 1 - \frac{1}{2}\lambda_k S_c x, \qquad \delta_b = -\frac{1}{2}\lambda_k S_c x,$$

$$\delta_\gamma = \delta_\nu = -\frac{2}{3}\lambda_k S_c x, \qquad \delta_g = -\frac{S_c(1 - 3c_s^2 + 4w)}{2}\lambda_k x,$$

$$v_\gamma = v_\nu = -\frac{1}{12}\lambda_k S_c x^2, \qquad v_g = -\frac{1}{6}c_s^2 \lambda_k S_c x^2,$$

$$\sigma_\nu = -\frac{S_c}{6(15 + 2S_\nu)}\lambda_k x^3, \qquad \sigma_g = \frac{c_{\rm vis}^2(15 - 4S_\nu)}{9}\sigma_\nu.$$

Note that had we assumed a pure radiation background (without the matter corrections to the scale factor evolution), σ_ν would (incorrectly) seem to grow as x^2 rather than the standard result (x^3), even in the case of a vanishing GDM component. One finds similar deviations when the w corrections to the scale factor are neglected.

d. Baryon isocurvature

The structure of the baryon isocurvature mode is identical to the CDM isocurvature mode from which it is obtained with the mappings $\delta_b \leftrightarrow \delta_c$ and $S_c \leftrightarrow S_b$.

e. GDM isocurvature

Setting $\delta_{g,0} = 1$ and all remaining perturbations in $\mathcal{I}_{\rm modes}$ [(3.10)] to zero, the GDM isocurvature mode is

$$\eta = -\frac{1}{6}\left[x^{-3c_s^2} + c_s^2 - 3w\ln\left(\frac{\lambda_k}{f_{mr}}\right)\right]\lambda_k S_g x, \qquad h = -6\eta,$$

$$\delta_c = \delta_b = -\frac{1}{2}h, \qquad \delta_\gamma = \delta_\nu = -\frac{2}{3}h,$$

$$\delta_g = x^{3(w-c_s^2)} + \frac{1}{4}\left[-2S_g x^{-3c_s^2} + 3(w - c_s^2) + 4S_g(c_s^2 - 2w)\right.$$
$$\left. + 6wS_g \ln\left(\frac{\lambda_k}{f_{mr}}\right)\right]\lambda_k x,$$

$$v_\gamma = v_\nu = -\frac{1}{12}\left[x^{-3c_s^2} + \frac{5}{2}c_s^2 - 3w\ln\left(\frac{\lambda_k}{f_{mr}}\right)\right]\lambda_k S_g x^2,$$

$$v_g = \frac{1}{2}c_s^2 x,$$

$$\sigma_\nu = -\frac{1}{6(15 + 2S_\nu)}\left[x^{-3c_s^2} + \frac{3c_s^2(65 + 4S_\nu)}{4(15 + 2S_\nu)}\right.$$
$$\left. - 3w\ln\left(\frac{\lambda_k}{f_{mr}}\right)\right]\lambda_k S_g x^3,$$

$$\sigma_g = c_{\rm vis}^2\left[\frac{4}{15}c_s^2 x^2 - \frac{15 - 4S_\nu}{54(15 + 2S_\nu)}\lambda_k S_g x^{3(1-c_s^2)}\right].$$

Note that for $w \neq c_s^2$ the value of the parameter $\delta_{g,0}$ does not really specify the value of δ_g in the limit $x \to 0$ due to the pure log term in the expansion ansatz (3.8) for δ_g. To ameliorate this problem, we have rewritten $a\ln(x)$ as $x^a - 1$ which converges for $x \to 0$ and gives better numerical results.

APPENDIX B: GENERALIZED TIGHT COUPLING

In this appendix we consider a more general way to impose the tight-coupling conditions in the case of two interacting fluids. In particular, we allow $q_{\rm int}$ and $S_{\rm int}$ to be linear combinations of θ_{12} and S_{12} parametrized by an angle β_c in the range $0 \leq \beta_c \leq \pi$. The relevant relations are

$$q_{\rm int} + Q\left[\left(\mathcal{H} - \frac{\dot{Q}}{Q}\right)\theta_g - \Psi + \frac{2}{3}(k^2 - 3\kappa)\Sigma_g\right] = \frac{\bar{\rho}_g \mathcal{H} R_c}{(1 + w)}[\cos(\beta_c)S_{12} + \sin(\beta_c)\mathcal{H}\theta_{12}] \tag{B1}$$

and

$$S_{\rm int} - Q\theta_g - \frac{2}{3}(k^2 - 3\kappa)\frac{\bar{\rho}_1\bar{\rho}_2(1 + w_1)(1 + w_2)}{\bar{\rho}_g(1 + w)}\Sigma_{12} = -(1 + w_1)\bar{\rho}_1 R_c[-\sin(\beta_c)S_{12} + \cos(\beta_c)\mathcal{H}\theta_{12}]. \tag{B2}$$

This immediately implies that $S_{12}^{(0)} = \theta_{12}^{(0)} = 0$. Rearranging the equations of motion for S_{12} and θ_{12} and keeping only the lowest-order terms, we find

$$S_{12}^{(1)} = \frac{1}{R_c}\left[\frac{Q(1 + w_1 + w_2)}{\bar{\rho}_g \mathcal{H}}\cos(\beta_c) - \frac{R_{\rm mix}w_{12}}{(1 + R_{\rm mix})(1 + w)}\sin(\beta_c)\right]\hat{\Delta}_g \tag{B3}$$

and

$$\mathcal{H}\theta_{12}^{(1)} = \frac{1}{R_c}\left[\frac{Q(1+w_1+w_2)}{\bar{\rho}_g\mathcal{H}}\sin(\beta_c) + \frac{R_{\mathrm{mix}}w_{12}}{(1+R_{\mathrm{mix}})(1+w)}\cos(\beta_c)\right]\hat{\Delta}_g. \tag{B4}$$

The $S_{12}^{(1)}$ relation then leads to a sound speed

$$c_s^2 = \frac{w_1 + R_{\mathrm{mix}}w_2}{1 + R_{\mathrm{mix}}} + \frac{w_{12}R_{\mathrm{mix}}(1+w)}{R_c(1+R_{\mathrm{mix}})^2}\left[\frac{Q(1+w_1+w_2)}{\bar{\rho}_g\mathcal{H}}\cos(\beta_c) - \frac{R_{\mathrm{mix}}}{(1+R_{\mathrm{mix}})}\frac{w_{12}}{1+w}\sin(\beta_c)\right], \tag{B5}$$

which now depends on the angle β_c. For general β_c the term which includes $\cos(\beta_c)$ is expected to be parametrically smaller than the term including $\sin(\beta_c)$ because it is suppressed by $Q/(\bar{\rho}_g\mathcal{H}) \ll 1$. We note that the tight-coupling condition $\theta_{12}^{(0)} = 0$ is unnecessary when keeping only the lowest-order terms in R_c^{-1}; however, it is necessary when including the next-to-leading order, as otherwise the dynamical $\theta_{12}^{(0)}$ will contribute to Π_g and spoil the GDM template.

---

[1] C. M. Will, Living Rev. Relativ. **9** (2006).
[2] P. A. R. Ade, N. Aghanim, M. Arnaud *et al.* (Planck Collaboration), arXiv:1502.01589.
[3] F. Zwicky, Helv. Phys. Acta **6**, 110 (1933).
[4] S. Smith, Astrophys. J. **83**, 23 (1936).
[5] V. C. Rubin and W. K. Ford, Jr., Astrophys. J. **159**, 379 (1970).
[6] W. J. G. de Blok and S. S. McGaugh, Mon. Not. R. Astron. Soc. **290**, 533 (1997).
[7] M. A. Zwaan, J. M. van der Hulst, W. J. G. de Blok, and S. S. McGaugh, Mon. Not. R. Astron. Soc. **273**, L35 (1995).
[8] F. Iocco, M. Pato, and G. Bertone, Nat. Phys. **11**, 245 (2015).
[9] D. Clowe, M. Bradač, A. H. Gonzalez, M. Markevitch, S. W. Randall, C. Jones, and D. Zaritsky, Astrophys. J. **648**, L109 (2006).
[10] A. Vikhlinin, A. Kravtsov, W. Forman, C. Jones, M. Markevitch, S. S. Murray, and L. Van Speybroeck, Astrophys. J. **640**, 691 (2006).
[11] A. Vikhlinin *et al.*, Astrophys. J. **692**, 1060 (2009).
[12] A. Refregier, Annu. Rev. Astron. Astrophys. **41**, 645 (2003).
[13] R. Massey, J. Rhodes, R. Ellis *et al.*, Nature (London) **445**, 286 (2007).
[14] D. N. Spergel, L. Verde, H. V. Peiris *et al.*, Astrophys. J. Suppl. Ser. **148**, 175 (2003).
[15] J. A. Peacock, S. Cole, P. Norberg, C. M. Baugh, J. Bland-Hawthorn, T. Bridges, R. D. Cannon, M. Colless, C. Collins, W. Couch *et al.*, Nature (London) **410**, 169 (2001).
[16] M. Tegmark *et al.*, Phys. Rev. D **74**, 123507 (2006).
[17] M. Viel, J. S. Bolton, and M. G. Haehnelt, Mon. Not. R. Astron. Soc. **399**, L39 (2009).
[18] J. Kwan, C. Sanchez, J. Clampitt, J. Blazek, M. Crocce, B. Jain, J. Zuntz, A. Amara, M. Becker, G. Bernstein, C. Bonnett *et al.*, arXiv:1604.07871.
[19] R. I. Epstein, J. M. Lattimer, and D. N. Schramm, Nature (London) **263**, 198 (1976).
[20] R. H. Cyburt, B. D. Fields, K. A. Olive, and T.-H. Yeh, Rev. Mod. Phys. **88**, 015004 (2016).
[21] T. Clifton, P. G. Ferreira, A. Padilla, and C. Skordis, Phys. Rep. **513**, 1 (2012).
[22] D. Harvey, R. Massey, T. Kitching, A. Taylor, and E. Tittley, Science **347**, 1462 (2015).
[23] A. A. Starobinskii, Sov. Astron. Lett. **4**, 82 (1978).
[24] A. H. Chamseddine, V. Mukhanov, and A. Vikman, J. Cosmol. Astropart. Phys. 06 (2014) 017.
[25] G. Bertone, D. Hooper, and J. Silk, Phys. Rep. **405**, 279 (2005).
[26] E. Aprile, M. Alfonsi, K. Arisaka *et al.*, Phys. Rev. Lett. **109**, 181301 (2012).
[27] E. Aprile, F. Agostini, M. Alfonsi *et al.*, Phys. Rev. D **90**, 062009 (2014).
[28] J. Buckley, D. F. Cowen, S. Profumo *et al.*, arXiv:1310.7040.
[29] K. A. Olive *et al.* (Particle Data Group), Chin. Phys. C **38**, 090001 (2014).
[30] F. Reindl, G. Angloher, A. Bento *et al.* (CRESST Collaboration), arXiv:1509.09124.
[31] D. S. Akerib, H. M. Araújo, X. Bai *et al.* (LUX Collaboration), Phys. Rev. Lett. **116**, 161301 (2016).
[32] S. Dodelson and L. M. Widrow, Phys. Rev. Lett. **72**, 17 (1994).
[33] S. Colombi, S. Dodelson, and L. M. Widrow, Astrophys. J. **458**, 1 (1996).
[34] X. Shi and G. M. Fuller, Phys. Rev. Lett. **82**, 2832 (1999).
[35] J. Lesgourgues and T. Tram, J. Cosmol. Astropart. Phys. 09 (2011) 032.
[36] P. J. E. Peebles, Astrophys. J. **534**, L127 (2000).
[37] L. M. Widrow and N. Kaiser, Astrophys. J. **416**, L71 (1993).
[38] H.-Y. Schive, T. Chiueh, and T. Broadhurst, Nat. Phys. **10**, 496 (2014).
[39] W. Hu, R. Barkana, and A. Gruzinov, Phys. Rev. Lett. **85**, 1158 (2000).
[40] P. Sikivie and Q. Yang, Phys. Rev. Lett. **103**, 111301 (2009).
[41] C.-G. Park, J.-c. Hwang, and H. Noh, Phys. Rev. D **86**, 083535 (2012).

[42] C. Bœhm, P. Fayet, and R. Schaeffer, Phys. Lett. B **518**, 8 (2001).

[43] S. Hofmann, D. J. Schwarz, and H. Stöcker, Phys. Rev. D **64**, 083507 (2001).

[44] D. Baumann, A. Nicolis, L. Senatore, and M. Zaldarriaga, J. Cosmol. Astropart. Phys. 07 (2012) 051.

[45] J. J. M. Carrasco, M. P. Hertzberg, and L. Senatore, J. High Energy Phys. 09 (2012) 82.

[46] S. M. Carroll, S. Leichenauer, and J. Pollack, Phys. Rev. D **90**, 023518 (2014).

[47] S. Foreman and L. Senatore, J. Cosmol. Astropart. Phys. 04 (2016) 033.

[48] D. Blas, S. Floerchinger, M. Garny, N. Tetradis, and U. A. Wiedemann, J. Cosmol. Astropart. Phys. 11 (2015) 049.

[49] D. Blas, M. Garny, M. M. Ivanov, and S. Sibiryakov, arXiv:1512.05807.

[50] B. Moore, Nature (London) **370**, 629 (1994).

[51] M. J. Jee, H. Hoekstra, A. Mahdavi, and A. Babul, Astrophys. J. **783**, 78 (2014).

[52] M. Boylan-Kolchin, J. S. Bullock, and M. Kaplinghat, Mon. Not. R. Astron. Soc. **415**, L40 (2011).

[53] E. Papastergis, R. Giovanelli, M. P. Haynes, and F. Shankar, Astron. Astrophys. **574**, A113 (2015).

[54] A. Klypin, I. Karachentsev, D. Makarov, and O. Nasonova, Mon. Not. R. Astron. Soc. **454**, 1798 (2015).

[55] P. Bode, J. P. Ostriker, and N. Turok, Astrophys. J. **556**, 93 (2001).

[56] M. R. Lovell, C. S. Frenk, V. R. Eke, A. Jenkins, L. Gao, and T. Theuns, Mon. Not. R. Astron. Soc. **439**, 300 (2014).

[57] D. J. E. Marsh and A.-R. Pop, Mon. Not. R. Astron. Soc. **451**, 2479 (2015).

[58] D. N. Spergel and P. J. Steinhardt, Phys. Rev. Lett. **84**, 3760 (2000).

[59] R. Hlozek, D. Grin, D. J. E. Marsh, and P. G. Ferreira, Phys. Rev. D **91**, 103512 (2015).

[60] C. Armendariz-Picon and J. T. Neelakanta, J. Cosmol. Astropart. Phys. 03 (2014) 049.

[61] O. F. Piattella, L. Casarini, J. C. Fabris, and J. A. de Freitas Pacheco, J. Cosmol. Astropart. Phys. 02 (2016) 024.

[62] M. Shoji and E. Komatsu, Phys. Rev. D **82**, 089901(E) (2010).

[63] F.-Y. Cyr-Racine and K. Sigurdson, Phys. Rev. D **90**, 123533 (2014).

[64] I. M. Oldengott, C. Rampf, and Y. Y. Y. Wong, J. Cosmol. Astropart. Phys. 04 (2015) 016.

[65] F.-Y. Cyr-Racine and K. Sigurdson, Phys. Rev. D **87**, 103515 (2013).

[66] R. Diamanti, E. Giusarma, O. Mena, M. Archidiacono, and A. Melchiorri, Phys. Rev. D **87**, 063509 (2013).

[67] C. Bœhm, A. Riazuelo, S. H. Hansen, and R. Schaeffer, Phys. Rev. D **66**, 083505 (2002).

[68] P. Serra, F. Zalamea, A. Cooray, G. Mangano, and A. Melchiorri, Phys. Rev. D **81**, 043507 (2010).

[69] R. J. Wilkinson, J. Lesgourgues, and C. Bœhm, J. Cosmol. Astropart. Phys. 04 (2014) 026.

[70] R. J. Wilkinson, C. Bœhm, and J. Lesgourgues, J. Cosmol. Astropart. Phys. 05 (2014) 011.

[71] L. Amendola, Mon. Not. R. Astron. Soc. **312**, 521 (2000).

[72] A. Pourtsidou, C. Skordis, and E. J. Copeland, Phys. Rev. D **88**, 083505 (2013).

[73] G. D'Amico, T. Hamill, and N. Kaloper, arXiv:1605.00996.

[74] H. B. Sandvik, M. Tegmark, M. Zaldarriaga, and I. Waga, Phys. Rev. D **69**, 123524 (2004).

[75] W. Hu, Astrophys. J. **506**, 485 (1998).

[76] C. M. Müller, Phys. Rev. D **71**, 047302 (2005).

[77] E. Calabrese, M. Migliaccio, L. Pagano, A. others Melchiorri, and P. Natoli, Phys. Rev. D **80**, 063539 (2009).

[78] S. Kumar and L. Xu, Phys. Lett. B **737**, 244 (2014).

[79] L. Xu, Phys. Rev. D **87**, 043503 (2013).

[80] H. Wei, Z.-C. Chen, and J. Liu, Phys. Lett. B **720**, 271 (2013).

[81] L. Xu and Y. Chang, Phys. Rev. D **88**, 127301 (2013).

[82] W. Li and L. Xu, Eur. Phys. J. C **74**, 2765 (2014).

[83] C. Skordis, Phys. Rev. D **79**, 123527 (2009).

[84] J. K. Bloomfield, É. É. Flanagan, M. Park, and S. Watson, J. Cosmol. Astropart. Phys. 08 (2013) 010.

[85] G. Gubitosi, F. Piazza, and F. Vernizzi, J. Cosmol. Astropart. Phys. 02 (2013) 032.

[86] I. Sawicki, I. D. Saltas, L. Amendola, and M. Kunz, J. Cosmol. Astropart. Phys. 01 (2013) 004.

[87] T. Baker, P. G. Ferreira, and C. Skordis, Phys. Rev. D **87**, 024015 (2013).

[88] R. A. Battye and J. A. Pearson, J. Cosmol. Astropart. Phys. 03 (2014) 051.

[89] B. Soergel, T. Giannantonio, J. Weller, and R. A. Battye, J. Cosmol. Astropart. Phys. 02 (2015) 037.

[90] C. Skordis, A. Pourtsidou, and E. J. Copeland, Phys. Rev. D **91**, 083537 (2015).

[91] R. Trotta and A. Melchiorri, Phys. Rev. Lett. **95**, 011305 (2005).

[92] P. A. R. Ade et al. (Planck Collaboration), Astron. Astrophys. **571**, A1 (2014).

[93] D. B. Thomas, M. Kopp, and C. Skordis, arXiv:1601.05097.

[94] P. A. R. Ade, N. Aghanim, M. Arnaud et al. (Planck Collaboration), arXiv:1502.01589.

[95] M. Kunz, S. Nesseris, and I. Sawicki, Phys. Rev. D **94**, 023510 (2016).

[96] H. Wei, J. Liu, Z.-C. Chen, and X.-P. Yan, Phys. Rev. D **88**, 043510 (2013).

[97] H. Velten, H. Borges, and T. R. P. Caramês, Phys. Rev. D **93**, 063503 (2016).

[98] M. Kunz and D. Sapone, Phys. Rev. Lett. **98**, 121301 (2007).

[99] A. Refregier, A. Amara, T. D. Kitching, A. Rassat, R. Scaramella, J. Weller, and f. t. Euclid (Imaging Consortium), arXiv:1001.0061.

[100] J. Lesgourgues, arXiv:1104.2932.

[101] C. W. Misner, K. S. Thorne, and J. A. Wheeler, Gravitation (W.H. Freeman and Co., San Francisco, 1973).

[102] C.-P. Ma and E. Bertschinger, Astrophys. J. **455**, 7 (1995).

[103] M. Archidiacono, E. Calabrese, and A. Melchiorri, Phys. Rev. D **84**, 123008 (2011).

[104] E. Sellentin and R. Durrer, Phys. Rev. D **92**, 063012 (2015).

[105] D. Blas, J. Lesgourgues, and T. Tram, J. Cosmol. Astropart. Phys. 07 (2011) 034.

[106] H. Kodama and M. Sasaki, Prog. Theor. Phys. Suppl. **78**, 1 (1984).

[107] R. Bean and O. Doré, Phys. Rev. D **69**, 083503 (2004).

[108] R. Durrer, *The Cosmic Microwave Background* (Cambridge University Press, Cambridge, England, 2008).

[109] D. Wands, K. A. Malik, D. H. Lyth, and A. R. Liddle, Phys. Rev. D **62**, 043527 (2000).

[110] J. Väliviita, E. Majerotto, and R. Maartens, J. Cosmol. Astropart. Phys. 07 (2008) 020.

[111] M. Bruni, P. K. S. Dunsby, and G. F. R. Ellis, Astrophys. J. **395**, 34 (1992).

[112] C. Bœhm and R. Schaeffer, Astron. Astrophys. **438**, 419 (2005).

[113] O. F. Piattella, J. C. Fabris, and W. Zimdahl, J. Cosmol. Astropart. Phys. 05 (2011) 029.

[114] H. Velten and D. J. Schwarz, Phys. Rev. D **86**, 083501 (2012).

[115] H. Velten, T. R. P. Caramês, J. C. Fabris, L. Casarini, and R. C. Batista, Phys. Rev. D **90**, 123526 (2014).

[116] S. Floerchinger, N. Tetradis, and U. A. Wiedemann, Phys. Rev. Lett. **114**, 091301 (2015).

[117] S. Weinberg, Astrophys. J. **168**, 175 (1971).

[118] R. A. Battye and A. Moss, Phys. Rev. D **76**, 023005 (2007).

[119] G. Ballesteros and J. Lesgourgues, J. Cosmol. Astropart. Phys. 10 (2010) 014.

[120] M. Bucher, K. Moodley, and N. Turok, Phys. Rev. D **62**, 083508 (2000).

[121] E. Majerotto, J. Väliviita, and R. Maartens, Mon. Not. R. Astron. Soc. **402**, 2344 (2010).

[122] K. Moodley, M. Bucher, J. Dunkley, P. G. Ferreira, and C. Skordis, Phys. Rev. D **70**, 103520 (2004).

[123] P. A. R. Ade, N. Aghanim, M. Arnaud *et al.* (Planck Collaboration), arXiv:1502.02114.

[124] T. Koivisto and D. F. Mota, Phys. Rev. D **73**, 083502 (2006).

[125] D. Bertacca, N. Bartolo, and S. Matarrese, Adv. Astron. **2010**, 904379 (2010).

[126] S. Ahmed, J. Approx. Theory **34**, 335 (1982).

[127] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions* (Dover Publications, Mineola, New York, 1965).

[128] W. Hu, Lect. Notes Phys. **470**, 207 (1996).

[129] W. Hu and N. Sugiyama, Astrophys. J. **471**, 542 (1996).

[130] W. Hu and M. White, Astrophys. J. **471**, 30 (1996).

[131] M. Zaldarriaga and U. Seljak, Phys. Rev. D **55**, 1830 (1997).

[132] A. Lewis and A. Challinor, Phys. Rep. **429**, 1 (2006).

[133] A. Lewis, Phys. Rev. D **76**, 063001 (2007).

[134] W. Hu, arXiv:astro-ph/9508126.

[135] W. Hu and I. Sawicki, Phys. Rev. D **76**, 104043 (2007).

[136] F.-Y. Cyr-Racine, K. Sigurdson, J. Zavala, T. Bringmann, M. Vogelsberger, and C. Pfrommer, Phys. Rev. D **93**, 123527 (2016).

[137] M. Vogelsberger, J. Zavala, F.-Y. Cyr-Racine, C. Pfrommer, T. Bringmann, and K. Sigurdson, Mon. Not. R. Astron. Soc. **460**, 1399 (2016).

[138] G. Ballesteros, J. Cosmol. Astropart. Phys. 03 (2015) 001.

[139] L. Tisza, Phys. Rev. **61**, 531 (1942).

[140] L. D. Landau and E. M. Lifshitz, *Fluid Mechanics (Course of Theoretical Physics)*, 2nd ed., Course of Theoretical Physics (Butterworth-Heinemann, Oxford, 1987), Vol. 6.

[141] X. Chen, H. Rao, and E. A. Spiegel, Phys. Lett. A **271**, 87 (2000).

[142] X. Chen and E. A. Spiegel, Mon. Not. R. Astron. Soc. **323**, 865 (2001).

[143] K. Paech and S. Pratt, Phys. Rev. C **74**, 014901 (2006).

[144] J. Bernstein, *Kinetic Theory in the Expanding Universe* (Cambridge University Press, Cambridge/New York, 1988), p. 157.

[145] M. Giovannini, Classical Quantum Gravity **22**, 5243 (2005).

[146] R. Maartens, arXiv:astro-ph/9609119.

[147] A. Loeb and M. Zaldarriaga, Phys. Rev. D **71**, 103520 (2005).

[148] C. Eckart, Phys. Rev. **58**, 919 (1940).

[149] W. Israel and J. M. Stewart, Ann. Phys. (N.Y.) **118**, 341 (1979).

[150] G. Ballesteros, L. Hollenstein, R. K. Jain, and M. Kunz, J. Cosmol. Astropart. Phys. 05 (2012) 038.

[151] E. Bertschinger, Phys. Rev. D **74**, 063509 (2006).

[152] W. Israel, J. Math. Phys. (N.Y.) **4**, 1163 (1963).

[153] I. Müller, Z. Phys. **198**, 329 (1967).

[154] W. A. Hiscock and L. Lindblom, Phys. Rev. D **31**, 725 (1985).

[155] P. Ván and T. S. Biró, AIP Conf. Proc. **1578**, 114 (2014).

[156] K. Tsumura and T. Kunihiro, Phys. Rev. E **87**, 053008 (2013).

[157] K. Tsumura, Y. Kikuchi, and T. Kunihiro, Phys. Rev. D **92**, 085048 (2015).

[158] P. Kostädt and M. Liu, Phys. Rev. D **62**, 023003 (2000).

[159] W. A. Hiscock and T. S. Olson, Phys. Lett. A **141**, 125 (1989).

[160] R. Geroch, J. Math. Phys. (N.Y.) **36**, 4226 (1995).

[161] L. Lindblom, Ann. Phys. (N.Y.) **247**, 1 (1996).

[162] P. Ván (private communication).

[163] S. A. Hayward, arXiv:gr-qc/9803007.

[164] J. Silk, Astrophys. J. **151**, 459 (1968).

[165] M. Alcubierre, *Introduction to 3+1 Numerical Relativity* (Oxford University Press, Oxford, 2008).

[166] E. Bertschinger, NASA STI/Recon Technical Report No. **96**, 1995.

[167] H. Öttinger, *Beyond Equilibrium Thermodynamics* (Wiley, Hoboken, New Jersey, 2005).

[168] I. Müller and T. Ruggeri, *Rational Extended Thermodynamics* (Springer Science & Business Media, Berlin, 2013), Vol. 37.

[169] M. M. Disconzi, T. W. Kephart, and R. J. Scherrer, arXiv:1510.07187.

[170] T. Baldauf, L. Mercolli, and M. Zaldarriaga, Phys. Rev. D **92**, 123007 (2015).

[171] M. P. Hertzberg, Phys. Rev. D **89**, 043521 (2014).

[172] J. J. M. Carrasco, S. Foreman, D. Green, and L. Senatore, J. Cosmol. Astropart. Phys. 07 (2014) 057.

[173] A. Akbar Abolhasani, M. Mirbabayi, and E. Pajer, J. Cosmol. Astropart. Phys. 05 (2016) 063.

[174] S. Foreman, H. Perrier, and L. Senatore, J. Cosmol. Astropart. Phys. 05 (2016) 027.

[175] J. Garriga and V. F. Mukhanov, Phys. Lett. B **458**, 219 (1999).

[176] R. J. Scherrer, Phys. Rev. Lett. **93**, 011301 (2004).

[177] L. Mirzagholi and A. Vikman, J. Cosmol. Astropart. Phys. 06 (2015) 028.

[178] A. H. Chamseddine and V. Mukhanov, J. High Energy Phys. 11 (2013) 135.

[179] A. H. Guth, M. P. Hertzberg, and C. Prescod-Weinstein, Phys. Rev. D **92**, 103513 (2015).

[180] J.-C. Hwang and H. Noh, Phys. Lett. B **680**, 1 (2009).

[181] E. Seidel and W.-M. Suen, Phys. Rev. D **42**, 384 (1990).

[182] M. Alcubierre, A. de la Macorra, A. Diez-Tejedor, and J. M. Torres, Phys. Rev. D **92**, 063508 (2015).

[183] T.-P. Woo and T. Chiueh, Astrophys. J. **697**, 850 (2009).

[184] J. A. R. Cembranos, A. L. Maroto, and S. J. Núñez Jareño, J. High Energy Phys. 03 (2016) 013.

[185] L. A. Ureña-López and A. X. Gonzalez-Morales, arXiv:1511.08195.

[186] G. Ballesteros and B. Bellazzini, J. Cosmol. Astropart. Phys. 04 (2013) 001.

[187] N. Andersson and G. L. Comer, Living Rev. Relativ. **10** (2007).

[188] J. Kijowski, A. Smólski, and A. Górnicka, Phys. Rev. D **41**, 1875 (1990).

[189] J. D. Brown, Classical Quantum Gravity **10**, 1579 (1993).

[190] G. L. Comer and D. Langlois, Classical Quantum Gravity **10**, 2317 (1993).

[191] S. Dubovsky, L. Hui, A. Nicolis, and D. T. Son, Phys. Rev. D **85**, 085029 (2012).

[192] J. Bhattacharya, S. Bhattacharyya, and M. Rangamani, J. High Energy Phys. 02 (2013) 153.

[193] N. A. Koshelev, arXiv:1512.07097.

[194] P. S. Letelier, Phys. Rev. D **22**, 807 (1980).

[195] J. P. Krisch and E. N. Glass, J. Math. Phys. (N.Y.) **52**, 102503 (2011).

[196] P. K. S. Dunsby, M. Bruni, and G. F. R. Ellis, Astrophys. J. **395**, 54 (1992).

[197] N. A. Koshelev, J. Cosmol. Astropart. Phys. 04 (2011) 021.

[198] J. Chluba, R. Khatri, and R. A. Sunyaev, Mon. Not. R. Astron. Soc. **425**, 1129 (2012).

[199] P. J. E. Peebles and J. T. Yu, Astrophys. J. **162**, 815 (1970).

[200] W. Hu and S. Dodelson, Annu. Rev. Astron. Astrophys. **40**, 171 (2002).

[201] D. Parkinson, S. Riemer-Sørensen, C. Blake, G. B. Poole, T. M. Davis, S. Brough, M. Colless, C. Contreras, W. Couch, S. Croom et al., Phys. Rev. D **86**, 103518 (2012).

[202] P. McDonald, U. Seljak, S. Burles, D. J. Schlegel, D. H. Weinberg, R. Cen, D. Shih, J. Schaye, D. P. Schneider, N. A. Bahcall et al., Astrophys. J. Suppl. Ser. **163**, 80 (2006).