PHYSICAL REVIEW D **80,** 123536 (2009)

Bimetric MOND gravity

Mordehai Milgrom

The Weizmann Institute Center for Astrophysics, Rehovot 76100, Israel

(Received 19 November 2009; published 30 December 2009)

A new relativistic formulation of MOND is advanced, involving two metrics as independent degrees of freedom: the MOND metric $g_{\mu\nu}$, to which alone matter couples, and an auxiliary metric $\hat{g}_{\mu\nu}$. The main idea hinges on the fact that we can form tensors from the difference of the Levi-Civita connections of the two metrics, $C^{\alpha}_{\beta\gamma} = \Gamma^{\alpha}_{\beta\gamma} - \hat{\Gamma}^{\alpha}_{\beta\gamma}$, and these act like gravitational accelerations. In the context of MOND, we can form dimensionless "acceleration" scalars and functions thereof (containing only first derivatives) from contractions of $a_0^{-1}C^{\alpha}_{\beta\gamma}$. I look at a subclass of bimetric MOND theories governed by the action $I = -(16\pi G)^{-1} \times \int[\beta g^{1/2}R + \alpha\hat{g}^{1/2}\hat{R} - 2(g\hat{g})^{1/4}f(\kappa)a_0^2\mathcal{M}(\tilde{\Upsilon}/a_0^2)] \times d^4x + I_M(g_{\mu\nu}, \psi_i) + \hat{I}_M(\hat{g}_{\mu\nu}, \chi_i)$, with $\tilde{\Upsilon}$ as a scalar quadratic in the $C^{\alpha}_{\beta\gamma}$, $\kappa = (g/\hat{g})^{1/4}$, I_M as the matter action, and allow for the existence of twin matter that couples to $\hat{g}_{\mu\nu}$ alone. Thus, gravity is modified not by modifying the elasticity of the space-time in which matter lives, but by the interaction between that space-time and the auxiliary one. In particular, I concentrate on the interesting and simple choice $\tilde{\Upsilon} \propto g^{\mu\nu}(C^{\gamma}_{\mu\lambda}C^{\lambda}_{\nu\gamma} - C^{\gamma}_{\mu\nu}C^{\lambda}_{\lambda\gamma})$. This theory introduces only one new constant, a_0; it tends simply to general relativity (GR) in the limit $a_0 \to 0$ and to a phenomenologically valid MOND theory in the nonrelativistic limit. The theory naturally gives MOND and "dark energy" effects from the same term in the action, both controlled by the MOND constant a_0. In regards to gravitational lensing by nonrelativistic systems–a holy grail for relativistic MOND theories–the theory predicts that the same potential that controls massive-particle motion also dictates lensing in the same way as in GR: Lensing and massive-particle probing of galactic fields will require the same "halo" of dark matter to explain the departure of the present theory from GR. This last result can be modified with other choices of $\tilde{\Upsilon}$, but lensing is still enhanced and MOND-like, with an effective logarithmic potential.

DOI: 10.1103/PhysRevD.80.123536 PACS numbers: 95.35.+d, 04.80.Cc

I. INTRODUCTION

From the inception of MOND [1], it has been clear that the paradigm needs buttressing by a relativistic formulation. Indeed, efforts to construct such a formulation started shortly thereafter, with the tensor-scalar version sketched in [2]. This was the first in a chain of theories of increasing force, culminating in the advent of the tensor-vector-scalar theory (TeVeS) of Bekenstein [3]. Some landmarks along this track are described in [3–8]; see, in particular, the reviews in [5,8]. All of these theories involve, as independent degrees of freedom, an Einstein metric, whose free action is the standard Einstein-Hilbert action, with additional scalar and/or vector degrees of freedom, with their own actions. These scalar/vector degrees of freedom are used to dress up the Einstein metric into the "physical" metric to which matter couples. TeVeS has a version of the nonrelativistic (NR) theory proposed in [2] as an NR limit.

Another line of relativistic theories that aim to reproduce MOND phenomenology has been propounded in [9,10], based on the omnipresence of a gravitationally polarizable medium proposed in [11].

Here, I propound a new class of relativistic formulations for the MOND paradigm in the form of bimetric MOND (BIMOND) theories. These came to light as follows: I have recently described [12] a new class of nonrelativistic, bi-potential MOND theories, a subclass of which is governed by a Lagrangian density of the form

$$\mathcal{L} = -\frac{1}{8\pi G}\{\beta(\vec{\nabla}\phi)^2 + \alpha(\vec{\nabla}\hat{\phi})^2 - a_0^2\mathcal{M}[(\vec{\nabla}\phi - \vec{\nabla}\hat{\phi})^2/a_0^2]\} + \rho\left(\frac{1}{2}\mathbf{v}^2 - \phi\right), \quad (1)$$

leading to the field equations

$$\vec{\nabla}\cdot[\mu^*(|\vec{\nabla}\phi^*|/a_0)\vec{\nabla}\phi^*] = 4\pi G\rho,$$
$$\mu^*(y) \equiv \beta - \frac{\alpha+\beta}{\alpha}\mathcal{M}'(y^2)$$
$$\Delta\phi = \vec{\nabla}\cdot[(1-\alpha^{-1}\mathcal{M}')\vec{\nabla}\phi^*]$$
$$= 4\pi G\beta^{-1}\rho + \beta^{-1}\vec{\nabla}\cdot(\mathcal{M}'\vec{\nabla}\phi^*), \quad (2)$$

with $\phi^* = \phi - \hat{\phi}$. I also described in detail the requirements from α, β, and $\mathcal{M}(z)$ that lead to the required MOND and Newtonian limits of these theories. In particular, I discussed at length the interesting case $\beta + \alpha = 0$ ($\beta = 1$ then normalizes G to be the Newton constant), which leads to the field equations

$$\Delta\phi^* = 4\pi G\rho,$$
$$\Delta\phi = \vec{\nabla}\cdot[(1+\mathcal{M}')\vec{\nabla}\phi^*] = 4\pi G\rho + \vec{\nabla}\cdot(\mathcal{M}'\vec{\nabla}\phi^*), \quad (3)$$

with $\mathcal{M}'$ as a function of $(\vec{\nabla}\phi^*/a_0)^2$, such that $\mathcal{M}'(z) \to 0$

for $z \to \infty$ ensures the Newtonian limit, and $\mathcal{M}'(z) \approx z^{-1/4}$ in the MOND regime $z \ll 1$. This is a particularly tractable MOND theory, as it requires solving only linear differential equations, with the inevitable MOND nonlinearity entering only algebraically. In all of the NR theories above, matter couples only to one of the potentials: the MOND potential ϕ, while $\hat{\phi}$ is an auxiliary potential, and in the special case of Eq. (3) their difference ϕ^* is exactly the Newtonian potential of the problem.

These NR MOND theories have inspired the construction of closely analogous relativistic MOND theories with two metrics as independent, gravitational degrees of freedom, which I begin to investigate here.[1] This new class of BIMOND theories involves only a_0 as a new constant. They tend to general relativity (GR) in the limit $a_0 \to 0$, which is a desirable trait. And, they tend to a MOND theory compatible with MOND phenomenology in their NR limit.

These theories, like all other relativistic versions of MOND proposed to date, must, I believe, be only approximate, effective theories to be derived from some more fundamental picture that underlies them. This is pointed to by the appearance of an *a priori* unspecified function in all these theories.

The use of two (or more) metrics to describe gravity has a long history. For example, Rosen [13] considered bimetric theories, where the auxiliary metric is forced to be flat. More recently, it was found [14] that ghosts appear in a large class of bimetric theories (apparently not including the present BIMOND). More matter-of-principle questions regarding bimetric gravities are discussed in [15–18], but these authors confined themselves to metric couplings that involve only the metrics, not their derivatives, as in the case of BIMOND.

In Sec. II, I present the formalism underlying the BIMOND theories; in Sec. III, I consider the NR limit of these theories, showing how they lead to NR MOND theories. Section IV demonstrates how the theories go to GR in the limit $a_0 \to 0$. Section V discusses lensing, Sec. VI discusses cosmology briefly, and Sec. VII is a discussion.

II. FORMALISM

The NR theories mentioned above involve two potentials, the MOND potential ϕ felt by matter, and an auxiliary one $\hat{\phi}$. They point to relativistic BIMOND theories involving the MOND metric $g_{\mu\nu}$, to which matter couples, and which reduces to ϕ in the NR limit, and involving, in addition, an auxiliary metric $\hat{g}_{\mu\nu}$.

Working with two metrics enables us to form nontrivial tensors and scalars from the difference in their Levi-Civita connections $\Gamma^\alpha_{\beta\gamma}$ and $\hat{\Gamma}^\alpha_{\beta\gamma}$,

$$C^\alpha_{\beta\gamma} = \Gamma^\alpha_{\beta\gamma} - \hat{\Gamma}^\alpha_{\beta\gamma}, \tag{4}$$

involving only first derivatives of the metrics, which is not possible with a single metric. This is particularly pertinent in the context of MOND, since connections act like gravitational accelerations. So, without introducing new constants in the relativistic formulation, we can write Lagrangian functions of dimensionless scalars constructed from $a_0^{-1} C^\alpha_{\beta\gamma}$ that enable us to interpolate between the GR limit, $a_0 \to 0$, and the MOND limit, $a_0 \to \infty$.

The tensor $C^\alpha_{\beta\gamma}$ is related to covariant derivatives of one metric with the connection of the other (more generally, they relate covariant derivatives of tensors with respect to the two connections):

$$\begin{aligned} g_{\mu\nu:\lambda} &= g_{\alpha\nu} C^\alpha_{\mu\lambda} + g_{\alpha\mu} C^\alpha_{\nu\lambda}, \\ \hat{g}_{\mu\nu;\lambda} &= -\hat{g}_{\alpha\nu} C^\alpha_{\mu\lambda} - \hat{g}_{\alpha\mu} C^\alpha_{\nu\lambda}, \end{aligned} \tag{5}$$

$$\begin{aligned} C^\lambda_{\alpha\beta} &= \tfrac{1}{2} g^{\lambda\rho}(g_{\alpha\rho:\beta} + g_{\beta\rho:\alpha} - g_{\alpha\beta:\rho}) \\ &= -\tfrac{1}{2}\hat{g}^{\lambda\rho}(\hat{g}_{\alpha\rho;\beta} + \hat{g}_{\beta\rho;\alpha} - \hat{g}_{\alpha\beta;\rho}), \end{aligned} \tag{6}$$

where the covariant derivative (;) is taken with the connection $\Gamma^\alpha_{\beta\gamma}$ and (:) with $\hat{\Gamma}^\alpha_{\beta\gamma}$. We can form various scalars out of $C^\alpha_{\beta\gamma}$ and the metrics. One scalar that will be of particular use to us is based on the tensor

$$\Upsilon_{\mu\nu} = C^\gamma_{\mu\lambda} C^\lambda_{\nu\gamma} - C^\gamma_{\mu\nu} C^\lambda_{\lambda\gamma}, \tag{7}$$

with the same index combination that appears in the expression for the Ricci tensor

$$R_{\mu\nu} = \Gamma^\alpha_{\mu\alpha,\nu} - \Gamma^\alpha_{\mu\nu,\alpha} + \Gamma^\gamma_{\mu\lambda}\Gamma^\lambda_{\nu\gamma} - \Gamma^\gamma_{\mu\nu}\Gamma^\lambda_{\lambda\gamma}, \tag{8}$$

and in $\hat{R}_{\mu\nu}$ constructed similarly from $\hat{g}_{\mu\nu}$. One finds

$$R_{\mu\nu} - \hat{R}_{\mu\nu} = C^\lambda_{\mu\lambda;\nu} - C^\lambda_{\mu\nu;\lambda} - \Upsilon_{\mu\nu}. \tag{9}$$

Thus, using well-known manipulations, the scalar $\Upsilon \equiv g^{\mu\nu}\Upsilon_{\mu\nu}$ connects the two Ricci scalars $R = g^{\mu\nu}R_{\mu\nu}$ and the mixed $\hat{R}_m = g^{\mu\nu}\hat{R}_{\mu\nu}$ by

$$\begin{aligned} R - \hat{R}_m = &-\Upsilon + g^{-1/2}(g^{1/2} g^{\mu\nu} C^\lambda_{\mu\lambda})_{,\nu} \\ &- g^{-1/2}(g^{1/2} g^{\mu\nu} C^\lambda_{\mu\nu})_{,\lambda}. \end{aligned} \tag{10}$$

Similarly, interchanging the roles of $g_{\mu\nu}$ and $\hat{g}_{\mu\nu}$,

$$\begin{aligned} \hat{R} - R_m = &-\hat{\Upsilon} - \hat{g}^{-1/2}(\hat{g}^{1/2}\hat{g}^{\mu\nu} C^\lambda_{\mu\lambda})_{,\nu} \\ &+ \hat{g}^{-1/2}(\hat{g}^{1/2}\hat{g}^{\mu\nu} C^\lambda_{\mu\nu})_{,\lambda}, \end{aligned} \tag{11}$$

where $\hat{\Upsilon} = \hat{g}^{\mu\nu}\Upsilon_{\mu\nu}$, $R_m = \hat{g}^{\mu\nu}R_{\mu\nu}$, $\hat{R}$ is the Ricci scalar

[1] In these theories, the two metrics are independent degrees of freedom. Theories like Brans-Dicke, TeVeS, etc., are also sometimes described as being bimetric, because they involve two metrics, but those two metrics are *a priori* related conformally or disformally via other degrees of freedom such as scalars or vectors.

of $\hat{g}_{\mu\nu}$, and g and $\hat{g}$ are minus the determinants of $g_{\mu\nu}$ and $\hat{g}_{\mu\nu}$, respectively.

We can construct gravitational Lagrangian densities using the scalars R, $\hat{R}$, R_m, $\hat{R}_m$, and scalars constructed by contracting powers of $C^{\lambda}_{\mu\nu}$ with the two metrics and their inverses (there are also $\hat{g}/g$, $\bar{\omega} \equiv g^{\mu\nu}\hat{g}_{\mu\nu}$, etc. that can be used). If we only contract with $g_{\mu\nu}$ and $g^{\mu\nu}$, a quadratic scalar is a linear combination (possibly with coefficients depending on scalars such as $g/\hat{g}$ or $\bar{\omega}$) of the following scalars:

$$g^{\mu\nu}C^{\gamma}_{\mu\lambda}C^{\lambda}_{\nu\gamma}, \qquad \bar{C}^{\gamma}C_{\gamma}, \qquad g_{\mu\nu}\bar{C}^{\mu}\bar{C}^{\nu},$$
$$g^{\mu\nu}C_{\mu}C_{\nu}, \qquad g_{\alpha\lambda}g^{\beta\mu}g^{\gamma\nu}C^{\alpha}_{\beta\gamma}C^{\lambda}_{\mu\nu}, \tag{12}$$

where $\bar{C}^{\gamma} \equiv g^{\mu\nu}C^{\gamma}_{\mu\nu}$, $C_{\gamma} \equiv C^{\alpha}_{\gamma\alpha}$. The choice of scalars to be used may be forced on us by various theoretical and phenomenological desiderata (see below). The main point is that $C^{\gamma}_{\mu\nu}$ have only first derivatives of the metrics, that they reduce to derivatives of the potential difference in the Newtonian limit (in the sense to be discussed below), and that we can form dimensionless quantities from them with the MOND acceleration a_0 (or a MOND length $\ell = c^2/a_0$). In regards to the four curvature scalars, is will be advantageous to include them in the action only linearly and to eschew terms such as in the voguish $f(R)$ theories. Such nonlinear terms render the theory a higher derivative one, which I would like to avoid.[2] Another reason to avoid such terms in the MOND context is that they do not naturally lead in their NR limit to a single constant a_0 controlling the dynamics.[3] Neither obstacle appears if we allow functions of scalars made of $C^{\alpha}_{\beta\gamma}$. These contain only first derivatives of the metrics and give NR limits in which only a_0 appears (see below). We see from Eqs. (10) that $g^{1/2}R$ and $g^{1/2}\hat{R}_m$ differ by Υ plus a total derivative, so it is enough to include one of these in the action, as we permit functions of Υ anyway. The same is true of the pair $\hat{g}^{1/2}R_m$ and $\hat{g}^{1/2}\hat{R}$. Because the number of possible combinations is too large to explore here, I limit myself to the subclass of actions of the form

$$I = -\frac{c^4}{16\pi G}\int[\beta g^{1/2}R + \alpha\hat{g}^{1/2}\hat{R} - 2(g\hat{g})^{1/4}f(\kappa) \times \ell^{-2}\mathcal{M}(\ell^m\Upsilon^{(m)}_i)]d^4x + I_M(g_{\mu\nu}, \psi_i) + \hat{I}_M(\hat{g}_{\mu\nu}, \chi_i), \tag{13}$$

where $\ell \equiv c^2/a_0$, $\kappa \equiv (g/\hat{g})^{1/4}$, $f(1) = 1$, and $\Upsilon^{(m)}_i$ are scalars formed by contracting a product of (even) m $C^{\alpha}_{\beta\gamma}$, which can be used in principle. In what follows, I shall confine myself to quadratic scalars.[4] I have included two matter actions: The first, I_M, involves the matter degrees of freedom with which we interact directly, designated symbolically as ψ_i. It contains only the MOND metric $g_{\mu\nu}$ to which matter is coupled in the standard way. The other, $\hat{I}_M$, involves other matter degrees of freedom, χ_i, and only $\hat{g}_{\mu\nu}$, to account for the possibility that $\hat{g}_{\mu\nu}$ controls a matter world of its own. There are no direct (electromagnetic, etc.) interactions between the ψ matter and the twin χ matter.[5]

I make two requirements of the action: a. Require that it gives an NR MOND theory in its NR limit. This means the following: given a nonrelativistic system of slow masses, one can express the metrics solution of the relativistic theory in terms of potentials so that the equations of motion for slow particles in the resulting (multi) potential theory are those required by NR MOND, with the appropriate MOND and Newtonian limits; this is a phenomenological requirement (by itself, it does not dictate the effects on massless particles–e.g., gravitational lensing–even in NR systems). b. Require that the action gives GR in the limit $a_0 \to 0$. This is not a phenomenological necessary (for example, TeVeS does not satisfy it), but I feel that it is highly desirable for various reasons. This automatically causes the theory to agree with the stringent constraints from the solar system and binary pulsars–which are known to agree with GR–because the accelerations in these systems are many orders of magnitude larger than a_0. I also require this limit lest we have to introduce additional constant(s) to the theory, which has to give GR in some limit of its parameters.

When the two metrics are conformally related, which might be the case in certain circumstances, $g_{\mu\nu} = e^{\vartheta(x)}\hat{g}_{\mu\nu}$, we have $C_{\lambda} = 2\vartheta_{,\lambda}$, $\bar{C}^{\lambda} = -g^{\lambda\rho}\vartheta_{,\rho}$, $\Upsilon_{\mu\nu} = (1/2)(g_{\mu\nu}g^{\alpha\beta}\vartheta_{,\alpha}\vartheta_{,\beta} - \vartheta_{,\mu}\vartheta_{,\nu})$, $\Upsilon = (3/2)g^{\mu\nu}\vartheta_{,\mu}\vartheta_{,\nu}$. If we *a priori* constrain our metrics to be conformally related (i.e. vary the action only over such pairs), we get the Brans-Dicke theory with the choice $\mathcal{M}(z) \propto z$ (and appropriate

[2]For the same reason, I avoid scalars that are higher order in the curvature tensors, such as the different possible contractions of $R_{\mu\nu}$ with itself or with $\hat{R}_{\mu\nu}$. These are even less appealing as explained in [19].

[3]For example, to account for dimensions correctly, a function of R has to be introduced as $f(\ell^2R)$, with ℓ as some length scale. The NR limit of R includes $c^{-2}\Delta\phi$ as the dominant term in ϕ/c^2, and second order ones such as $c^{-4}(\vec{\nabla}\phi)^2$ and $c^{-4}\phi\Delta\phi$. Thus, in the argument of f, the second term will give $(\vec{\nabla}\phi/a_0)^2$ with $a_0 = c^2/\ell$, which fits well into the MOND frame. But the first, dominant term would involve a time scale ℓ/c, not an acceleration. When R appears linearly, the first term becomes immaterial in the action, as a complete derivative, and we are left with terms that are welcome in MOND [the $\phi\Delta\phi$ term is also $(\vec{\nabla}\phi)^2$ up to a derivative].

[4]The MOND constant a_0 is normalized so that the mass-asymptotic-velocity relation is $MGa_0 = V^4$. It defines the scale length ℓ that is used in the coefficient and the argument of $\mathcal{M}$. Any dimensionless factors can be absorbed in the definition of $\mathcal{M}$ so that its coefficient is $\ell^{-2} = c^{-4}a_0^2$ and its argument is as prescribed here.

[5]To obviate possible confusion, note that the twin matter is not to play the role of the putative dark matter in galactic systems. This is still fully replaced by MOND effects; see below.

choice of the constants and $f(\kappa)$ and possibly using $\hat{\Upsilon}$ in the argument of $\mathcal{M}$). With a more general form of $\mathcal{M}(z)$, we then get the relativistic MOND theory sketched in [2].

Without the interaction $\mathcal{M}$ term, the theory separates into two disjoint copies of GR. It is important to note that as a combined structure, the theory then enjoys a larger symmetry involving separate coordinate transformations in the two separate actions. This double symmetry has to be brought to bear when solving the field equations of the theory, which now satisfy two sets of Bianchi identities. So, eight gauge conditions can, and have to, be employed. It is the interaction that breaks this larger symmetry, as, generically, it is only invariant to application of the same coordinate transformation to the two metrics. However, under certain circumstances, the interaction is symmetric under a more extended set of coordinate transformation, and we must be careful then to employ the larger gauge freedom. The above mentioned complete decoupling is an example that, as we shall see in Sec. IV, applies in the formal limit $a_0 \rightarrow 0$ of the theory (leading, as we want, to GR). It may also happen, in principle, that the interaction term vanishes only in some limited regions of space-time; for example, if the extreme GR limit applies in some regions. In this case, we must allow for gauge freedom involving coordinate transformations that coincide only outside these regions, but not inside them. We shall see another example in Sec. III, where the NR limit of the theory has such a partial double gauge freedom.

A. Concrete simple example

I shall hereafter concentrate on a simple special case of the class. Some generalizations will be mentioned briefly below, in this section, and in Sec. VI.

In the first place, I take $\mathcal{M}$ to be a function of only one scalar, quadratic in the $C^{\alpha}_{\beta\gamma}$. In particular, I find the scalar Υ defined above a natural choice for this argument, as it has the same structure as the first-derivative part of the Ricci curvature scalar (not itself a scalar)

$$\Gamma^{(2)} \equiv g^{\mu\nu}(\Gamma^{\gamma}_{\mu\lambda}\Gamma^{\lambda}_{\nu\gamma} - \Gamma^{\gamma}_{\mu\nu}\Gamma^{\lambda}_{\lambda\gamma}). \tag{14}$$

It is well known that one can replace R in the Einstein-Hilbert action by $\Gamma^{(2)}$ and still get GR. Here, we can also do this, replacing also $\hat{R}$ by the corresponding $\hat{\Gamma}^{(2)}$, and making $\mathcal{M}$ a function of Υ, which is constructed in the same way from $C^{\alpha}_{\beta\gamma}$. We shall also see that with this choice of scalar argument, the NR limit of the theory is especially simple.

As a further simplification, I take $\alpha + \beta = 0$. This will yield a particularly interesting and simple subclass of theories, which turn out to have the theory (3) as their NR limit for slowly moving masses in a double Minkowski background. I then take $\beta = 1$ for G to be Newton's constant.

Work in units in which $c = 1$, and use $a_0 = \ell^{-1}$ to highlight the connection with MOND. Also, anticipating the expression for NR limit of Υ, I take the argument of $\mathcal{M}$ to be $-\Upsilon/2a_0^2$. The relativistic action I then consider is

$$I = -\frac{1}{16\pi G}\int[g^{1/2}R - \hat{g}^{1/2}\hat{R} - 2(g\hat{g})^{1/4}f(\kappa) \times a_0^2\mathcal{M}(-\Upsilon/2a_0^2)]d^4x + I_M(g_{\mu\nu}, \psi_i) - \hat{I}_M(\hat{g}_{\mu\nu}, \chi_i). \tag{15}$$

[Using Eq. (11), we can replace the first two terms by $(g^{1/2}g^{\mu\nu} - \hat{g}^{1/2}\hat{g}^{\mu\nu})R_{\mu\nu} + \hat{g}^{1/2}\hat{\Upsilon}$.] I take a mixed volume element for the interaction term, with f normalized such that $f(1) = 1$. Note the change of sign in the definition of the twin matter action to match the negative sign for the Hilbert-Einstein action of $\hat{g}_{\mu\nu}$.

Varying over $g^{\mu\nu}$ and over $\hat{g}^{\mu\nu}$, we get, respectively,

$$G_{\mu\nu} + S_{\mu\nu} = -8\pi G\mathcal{T}_{\mu\nu}, \tag{16}$$

$$\hat{G}_{\mu\nu} + \hat{S}_{\mu\nu} = -8\pi G\hat{\mathcal{T}}_{\mu\nu}, \tag{17}$$

where $G_{\mu\nu}$ and $\hat{G}_{\mu\nu}$ are the Einstein tensors of the two metrics,

$$G_{\mu\nu} = R_{\mu\nu} - \tfrac{1}{2}Rg_{\mu\nu}, \qquad \hat{G}_{\mu\nu} = \hat{R}_{\mu\nu} - \tfrac{1}{2}\hat{R}\hat{g}_{\mu\nu}. \tag{18}$$

$\mathcal{T}_{\mu\nu}$ and $\hat{\mathcal{T}}_{\mu\nu}$ are the matter energy-momentum tensors (EMT); e.g., $\delta I_M \equiv -(1/2)\int g^{1/2}\mathcal{T}_{\mu\nu}\delta g^{\mu\nu}$, and $S_{\mu\nu}$, $\hat{S}_{\mu\nu}$ are the functional derivatives (one with an opposite sign) of the interaction term with respect to the two metrics:

$$\delta\int -2(g\hat{g})^{1/4}f(\kappa)a_0^2\mathcal{M}(-\Upsilon/2a_0^2)d^4x \equiv \int(g^{1/2}\delta g^{\mu\nu}S_{\mu\nu} - \hat{g}^{1/2}\delta\hat{g}^{\mu\nu}\hat{S}_{\mu\nu})d^4x. \tag{19}$$

For the present choice of the scalar argument of $\mathcal{M}$, we have

$$S_{\mu\nu} = \kappa_-\mathcal{M}'\Upsilon_{\mu\nu} + (\kappa_-\mathcal{M}'\tilde{\mathcal{S}}^{\lambda}_{\mu\nu})_{;\lambda} - \Lambda_m g_{\mu\nu}, \tag{20}$$

$$\hat{S}_{\mu\nu} = (\kappa_+\mathcal{M}'\hat{\mathcal{S}}^{\lambda}_{\mu\nu})_{:\lambda} - \hat{\Lambda}_m\hat{g}_{\mu\nu}, \tag{21}$$

$$\tilde{\mathcal{S}}^{\lambda}_{\mu\nu} = C^{\lambda}_{\mu\nu} - \delta^{\lambda}_{(\mu}C_{\nu)} + \tfrac{1}{2}(C^{\lambda} - \bar{C}^{\lambda})g_{\mu\nu}, \tag{22}$$

$$\hat{\mathcal{S}}^{\lambda}_{\mu\nu} = q^{\alpha}_{(\mu}C^{\lambda}_{\nu)\alpha} + g^{\lambda\rho}C^{\alpha}_{\rho(\mu}\hat{g}_{\nu)\alpha} - \hat{g}^{\lambda\rho}C^{\alpha}_{\rho\beta}q^{\beta}_{(\mu}\hat{g}_{\nu)\alpha} - q^{\lambda}_{(\mu}C_{\nu)} + \tfrac{1}{2}g^{*}_{\mu\nu}\hat{g}^{\lambda\alpha}C_{\alpha} - \tfrac{1}{2}\bar{C}^{\lambda}\hat{g}_{\mu\nu}, \tag{23}$$

$$\Lambda_m = -\frac{1}{2\kappa}[\kappa f(\kappa)]'a_0^2\mathcal{M}, \qquad \hat{\Lambda}_m = -\frac{\kappa^3}{2}[\kappa^{-1}f(\kappa)]'a_0^2\mathcal{M}. \tag{24}$$

Here,

$$C^\lambda \equiv g^{\alpha\lambda}C_\alpha, \qquad \kappa_\pm \equiv \kappa^{\pm 1}f(\kappa),$$
$$q^\mu_\alpha = g^{\mu\nu}\hat{g}_{\nu\alpha}, \qquad g^*_{\mu\nu} = \hat{g}_{\alpha\mu}g^{\alpha\beta}\hat{g}_{\beta\nu}, \tag{25}$$

and $(\mu\ldots\nu) = \{\mu\ldots\nu + \nu\ldots\mu\}/2$ signifies symmetrization over the two indices.

The tensor

$$\tilde{T}_{\mu\nu} = \frac{1}{8\pi G}[\kappa_- \mathcal{M}'\Upsilon_{\mu\nu} + (\kappa_-\mathcal{M}'\tilde{\mathcal{S}}^\lambda_{\mu\nu})_{;\lambda}] \tag{26}$$

may be viewed as the EMT of the phantom dark matter (DM); whereas, the Λ_m term may roughly be viewed as dark energy. Note that the last term in $\tilde{\mathcal{S}}^\lambda_{\mu\nu}$, which contributes $\frac{1}{2}[\kappa_-\mathcal{M}'(C^\lambda - \bar{C}^\lambda)]_{;\lambda}g_{\mu\nu}$ may also contribute to the dark energy due to its form. Define in analogy with $\tilde{T}_{\mu\nu}$,

$$\hat{T}_{\mu\nu} = \frac{1}{8\pi G}(\kappa_+\mathcal{M}'\hat{\mathcal{S}}^\lambda_{\mu\nu})_{:\lambda}. \tag{27}$$

The Einstein tensors satisfy the usual Bianchi identities $G^\nu_{\mu;\nu} = \hat{G}^\nu_{\mu:\nu} = 0$,[6] derivable from the invariance of the Einstein-Hilbert actions to coordinate transformations. In addition, we have here, for the general action (13), a set of four identities following from the fact that the mixed term is a scalar; these read

$$S^\nu_{\mu;\nu} - \kappa^{-2}\hat{S}^\nu_{\mu:\nu} = 0. \tag{28}$$

Given that the matter EMTs are divergence free (for matter degrees of freedom satisfying their own equations of motion): $\mathcal{T}^\nu_{\mu;\nu} = \hat{\mathcal{T}}^\nu_{\mu:\nu} = 0$, the above identities imply four differential identities satisfied by our 20 field equations. If we write these equations as $Q_{\mu\nu} = 0$ and $\hat{Q}_{\mu\nu} = 0$, respectively, then the four relations

$$Q^\nu_{\mu;\nu} - \kappa^{-2}\hat{Q}^\nu_{\mu:\nu} = 0 \tag{29}$$

hold identically, and, as usual, deprive us of four equations to account for the fact that the solution can be determined only up to a coordinate transformation. This seems to leave us with a tractable Cauchy problem, although this requires more careful checking.[7]

Of course, for solutions of the field equations, we do have separately

$$S^\nu_{\mu;\nu} = \hat{S}^\nu_{\mu:\nu} = 0, \tag{30}$$

which can be used as useful constraints of the solutions (only one set is independent).

Note the useful identities

$$C_\nu = \tfrac{1}{2}g^{\alpha\beta}g_{\alpha\beta:\nu} = -\tfrac{1}{2}\hat{g}^{\alpha\beta}\hat{g}_{\mu\nu;\nu} = 2\kappa_{,\nu}/\kappa,$$
$$\bar{C}^\lambda = -\kappa^{-2}(\kappa^2 g^{\lambda\rho})_{:\rho}C^\lambda = -\bar{C}^\lambda - g^{\lambda\nu}{}_{:\nu}, \tag{31}$$

$$\tilde{\mathcal{S}}^\lambda_{\mu\nu;\lambda} = \hat{R}_{\mu\nu} - \tfrac{1}{2}\hat{R}_m g_{\mu\nu} - (R_{\mu\nu} - \tfrac{1}{2}Rg_{\mu\nu}) - (\Upsilon_{\mu\nu} - \tfrac{1}{2}\Upsilon g_{\mu\nu}),$$
$$(C^\lambda - \bar{C}^\lambda)_{;\lambda} = R - \hat{R}_m + \Upsilon. \tag{32}$$

Identities (32) follow from Eqs. (9) and (10). Similar manipulations are possible for $\hat{\mathcal{S}}^\lambda_{\mu\nu}$, and the field Eq. (17).

Contracting Eq. (16) with $g^{\mu\nu}$ gives

$$R - [\kappa_-\mathcal{M}'(C^\lambda - \bar{C}^\lambda)]_{;\lambda} - \kappa_-\mathcal{M}'\Upsilon + 4\Lambda_m = 8\pi G\mathcal{T}. \tag{33}$$

Contracting Eq. (17) with $\hat{g}^{\mu\nu}$ gives

$$\hat{R} - [\kappa_+\mathcal{M}'(\tfrac{1}{2}\bar{\omega}\hat{g}^{\lambda\rho}C_\rho - \bar{C}^\lambda - \hat{g}^{\lambda\rho}q^\alpha_\mu C^\mu_{\rho\alpha})]_{:\lambda} + 4\hat{\Lambda}_m = 8\pi G\hat{\mathcal{T}}, \tag{34}$$

where $\bar{\omega} = g^{\mu\nu}\hat{g}_{\mu\nu}$. We can thus replace Eq. (16) by

$$R_{\mu\nu} + \kappa_-\mathcal{M}'(\Upsilon_{\mu\nu} - \tfrac{1}{2}\Upsilon g_{\mu\nu}) - [\kappa_-\mathcal{M}'(\delta^\lambda_{(\mu}C_{\nu)} - C^\lambda_{\mu\nu})]_{;\lambda} + \Lambda_m g_{\mu\nu} = -8\pi G(\mathcal{T}_{\mu\nu} - \tfrac{1}{2}\mathcal{T}g_{\mu\nu}), \tag{35}$$

and similarly for Eq. (17). We can also use identities (9)–(11) to write these equations in different forms. Equations (33) and (34) can be used to write possibly useful integral (virial) relations by integrating them over space-time, each with its own volume element.

It was deduced in [14] that under certain assumptions about the theory, bimetric theories generically posses ghosts. One of their assumptions was that to lowest order in departure from double Minkowski the theory is a sum of Pauli-Fierz actions for the different metrics, which are quadratic in the metric departures. This, however, leads to a linear theory in this limit, which is at odds with MOND: MOND phenomenology dictates that at $g_{\mu\nu} = \hat{g}_{\mu\nu} = \eta_{\mu\nu}$, any BIMOND theory (or any relativistic MOND theory for that matter) is not even analytic in the squares of the departures $g_{\mu\nu} - \eta_{\mu\nu}$, $\hat{g}_{\mu\nu} - \eta_{\mu\nu}$ (where the argument of $\mathcal{M}'$ in the above version of the theory vanishes, and $\mathcal{M}'$ diverges). It thus remains to be seen if obstacles similar to these are at all relevant to BIMOND, and if they are to what extent they are deleterious.

For conformally related metrics $g_{\mu\nu} = e^{\vartheta(x)}\hat{g}_{\mu\nu}$, we have $\tilde{\mathcal{S}}^\lambda_{\mu\nu} = e^{\vartheta(x)}\hat{\mathcal{S}}^\lambda_{\mu\nu}$.

[6] For each tensor, indices are raised with the corresponding metric; so, e.g., $G^\nu_\mu = g^{\nu\alpha}G_{\mu\alpha}$, $\hat{G}^\nu_\mu = \hat{g}^{\nu\alpha}\hat{G}_{\mu\alpha}$.

[7] As a result of identities (28) and the Bianchi identities, the four expressions $G^0_\mu + S^0_\mu - \kappa^{-2}(\hat{G}^0_\mu + \hat{S}^0_\mu)$ contain only up to first time derivatives of the metric and cannot be used to propagate the problem in time. Instead, the initial conditions have to satisfy the four equations $Q^0_\mu - \kappa^{-2}\hat{Q}^0_\mu = 0$, and the remaining 16 field equations, with the aid of four gauge conditions, propagate us in time, and insure that these four are always satisfied.

B. Generalizations

Some generalizations of the above simple theory include the following.

1. Instead of using Υ as the argument of $\mathcal{M}$, we can use other scalars, or several scalar variables. A quadratic scalar variable can be written, most generally, as

$$\Xi = Q^{\beta\gamma\mu\nu}_{\alpha\lambda} C^{\alpha}_{\beta\gamma} C^{\lambda}_{\mu\nu}, \tag{36}$$

where $Q^{\beta\gamma\mu\nu}_{\alpha\lambda}$ is built from $g_{\mu\nu}$, $\hat{g}_{\mu\nu}$, their inverses, δ^{α}_{β}, and scalars such as κ and $\bar{\omega}$. In this case, the Λ terms take a more general form, and so do terms that are second order in the $C^{\alpha}_{\beta\gamma}$. The only terms in $S_{\mu\nu}$ and $\hat{S}_{\mu\nu}$ that survive in the NR limit, which we treat below, are those involving $\tilde{\mathcal{S}}^{\lambda}_{\mu\nu}$ and $\hat{\mathcal{S}}^{\lambda}_{\mu\nu}$. For these, we now have, for example,

$$\tilde{\mathcal{S}}^{\lambda}_{\mu\nu} = 2U^{\gamma\lambda}_{(\mu} g_{\nu)\gamma} - U^{\gamma\sigma}_{\rho} g_{\mu\gamma} g_{\nu\sigma} g^{\lambda\rho}, \qquad U^{\mu\nu}_{\lambda} = -2Q^{\beta\gamma\mu\nu}_{\alpha\lambda} C^{\alpha}_{\beta\gamma}, \tag{37}$$

which I shall need in what follows.

For example, taking as the argument of $\mathcal{M}$, $-C^{\gamma}_{\mu\lambda} C^{\lambda}_{\nu\gamma}/2a_0^2$ instead of $-\Upsilon/2a_0^2$, would leave us with only the first term in expression (22) for $\tilde{\mathcal{S}}^{\lambda}_{\mu\nu}$, and with the first three terms in expression (23) for $\hat{\mathcal{S}}^{\lambda}_{\mu\nu}$.

2. One can consider more general α, β values.

3. We can increase the symmetry with respect to the two metrics by taking interaction terms of the form $\mathcal{M}(\Upsilon\hat{\Upsilon})$, $\mathcal{M}(\Upsilon)\mathcal{M}(\hat{\Upsilon})$, etc.

4. One can make $\mathcal{M}$ a function of scalars such as κ and $\bar{\omega}$.

Additional generalizations will be mentioned in Sec. VI.

III. NONRELATIVISTIC LIMIT

Consider now the NR limit of the theory derived from the action (15). This limit applies to systems where all quantities with the dimensions of velocities, such as v, $\sqrt{\phi}$, etc., are much smaller than the speed of light. In the context of GR, this limit is attained by formally taking $c \to \infty$ everywhere in the relativistic theory. In the context of MOND, one has to be more specific, since system attributes with the dimensions of acceleration, such as v^2/R, $\vec{\nabla}\phi$, etc., cannot be assumed very small in the limiting process, even though they have velocities in the numerator. We want to consider systems, such as galaxies, in which these are finite compared with the MOND acceleration, which is also a relevant parameter. The NR limit in MOND is thus formally attained by taking everywhere $c \to \infty$, but at the same time $\ell \to \infty$, so that $a_0 = c^2/\ell$ remains finite.

Take a system of quasistatic (nonrelativistically moving) masses, so that to a satisfactory approximation we can, as usual, neglect all components of the matter EMT except $\mathcal{T}_{00} = \rho$. I also neglect here the possible effects of the presence of twin matter.[8] First, I consider the system in a double Minkowski background. This is aesthetically the most appealing option, which I shall assume. It relies on the possibility that on cosmological scales the two metrics are, somehow, maintained the same from some symmetry. There are indeed versions of BIMOND [made more symmetric in the two metrics than our simple action (15) is] that have cosmological solutions with $\hat{g}_{\mu\nu} = g_{\mu\nu}$, either at all times, or as vacuum solutions, which might be appropriate for today (see Sec. VI). In this case, we have $C^{\alpha}_{\beta\gamma} = 0$ for the cosmological background, and finite $C^{\alpha}_{\beta\gamma}$ values occur only due to local inhomogeneities. We can then take locally, on scales much smaller then cosmological ones, a double Minkowski background. Departures from this assumption will be discussed below.

Write, then, the metrics as slightly perturbed from Minkowski. Because the source system is time-reversal symmetric in the approximation, we treat it (neglecting motions in the source), we are looking for a solution for which the mixed space-time elements of the two metrics vanish.[9] We can then write most generally

$$g_{\mu\nu} = \eta_{\mu\nu} - 2\phi\delta_{\mu\nu} + h_{\mu\nu}, \qquad \hat{g}_{\mu\nu} = \eta_{\mu\nu} - 2\hat{\phi}\delta_{\mu\nu} + \hat{h}_{\mu\nu}, \tag{38}$$

where $h_{0\mu} = h_{\mu 0} = \hat{h}_{0\mu} = \hat{h}_{\mu 0} = 0$. We denote the differences

$$g^{*}_{\mu\nu} = g_{\mu\nu} - \hat{g}_{\mu\nu} = -2\phi^{*}\delta_{\mu\nu} + h^{*}_{\mu\nu}, \tag{39}$$

with $\phi^{*} = \phi - \hat{\phi}$, $h^{*}_{\mu\nu} = h_{\mu\nu} - \hat{h}_{\mu\nu}$. We wish to solve the field equations to first order in the potentials ϕ, $\hat{\phi}$, h_{ij}, $\hat{h}_{ij}$ (Roman letters are used for space indices).

Note that there is a subtlety here (as in all metric MOND theories) due to the fact that the NR MOND potential for an isolated mass diverges logarithmically at infinity; so, strictly speaking we cannot formulate a first-order theory for such an isolated mass assuming $\phi \ll 1$ at all radii. However, we are, in any event, dealing with an effective theory to be understood in the context of the Universe at large, and in this context there are no isolated masses; our approach is meant to work only well within the distance from the central mass to the next comparable mass, where we can assume the first-order theory to be a good approximation.

[8]This is justified if this matter is nonexistent, or of it is smoothly distributed so its local contribution is negligible, or if there does not happen to exist a twin body in the near vicinity of the ψ body under study.

[9]We do not have to assume this *a priori*; if we do not, the equations themselves will tell us that there is a choice of gauge in which the solution satisfies this ansatz; see the end of this subsection. The ansatz simplifies the presentation, and is justified *a posteriori* by our showing below that such a solution exists.

To the required order, the only nonvanishing components of $C^{\alpha}_{\beta\gamma}$ are[10]

$$C^i_{00} = C^0_{0i} = C^0_{i0} = -\tfrac{1}{2}g^*_{00,i} = \phi^*_{,i},$$
$$C^i_{jk} = \tfrac{1}{2}(g^*_{ij,k} + g^*_{ik,j} - g^*_{jk,i}) = \tfrac{1}{2}(h^*_{ij,k} + h^*_{ik,j} - h^*_{jk,i}) + \phi^*_{,i}\delta_{jk} - \phi^*_{,j}\delta_{ik} - \phi^*_{,k}\delta_{ij}. \quad (40)$$

These reflect the same relations between the separate connections with their respective potentials.

The only nonvanishing components of the Ricci tensors (shown here for $R_{\mu\nu}$) are

$$R_{00} = -\Delta\phi, \qquad R_{ij} = \tfrac{1}{2}H_{ij} - \Delta\phi\delta_{ij}, \quad (41)$$

with

$$H_{ij} \equiv \Delta h_{ij} + h_{,i,j} - 2h_{k(i,j),k}, \quad (42)$$

where h is the trace of h_{ij}. The nonvanising components of the Einstein tensor are

$$G_{00} = -2\Delta\phi + \tfrac{1}{4}H, \qquad G_{ij} = \tfrac{1}{2}\Big(H_{ij} - \tfrac{1}{2}H\delta_{ij}\Big) \quad (43)$$

(H is the trace of H_{ij}). The same expressions exist for the hatted and for the starred quantities. We are now ready to use these expressions in the field Eqs. (16) and (17). We neglect the small cosmological-constant terms (in line with our assuming background Minkowski metrics), and note that terms such as $\Upsilon_{\mu\nu}$ are of second order in the potentials, so they can be neglected. Also, $\tilde{\mathcal{S}}^{\lambda}_{\mu\nu}$ and $\hat{\mathcal{S}}^{\lambda}_{\mu\nu}$ are linear in components of the tensor $C^{\alpha}_{\beta\gamma}$, which are first order in the potentials; so everywhere else in these expressions, we can take the metrics as Minkowski, so that $f(\kappa) \approx 1$, $\kappa_{\pm} \approx 1$, $q^{\lambda}_{\mu} \approx \delta^{\lambda}_{\mu}$, etc. Also, for the same reason, the covariant derivatives can be replaced by normal derivatives. All in all, we get that the two terms involving $\tilde{\mathcal{S}}^{\lambda}_{\mu\nu}$ and $\hat{\mathcal{S}}^{\lambda}_{\mu\nu}$ are equal. Thus, taking the difference of the two field equations, we get

$$G^*_{00} = -8\pi G\mathcal{T}_{00}, \qquad G^*_{ij} = 0. \quad (44)$$

Substituting from Eq. (43), we get

$$\Delta\phi^* - \tfrac{1}{8}H^* = 4\pi G\rho, \qquad H^*_{ij} - \tfrac{1}{2}H^*\delta_{ij} = 0. \quad (45)$$

Taking the trace of the second part, we get $H^* = 0$, and substituting in the first, we get

$$\Delta\phi^* = 4\pi G\rho. \quad (46)$$

We impose for ϕ^* the boundary condition at infinity $\phi^* \to 0$, which establishes it as the Newtonian potential of the problem. But the second Eq. (45) does not, in itself, determine h^*_{ij}, because G^*_{ij}, like G_{ij}, satisfy three Bianchi identities $G^*_{ij,j} = 0$, which are the reductions of identities (29) to our case (the fourth identity is automatically satisfied for our choice of vanishing mixed elements of the metrics).

We have only used one of the field equations (or rather their difference). Now consider the first field equation alone in the form (35). Again, neglect the second order Υ terms, etc. to get

$$R_{\mu\nu} + \Big[\mathcal{M}'\Big(\bar{\mathcal{S}}^i_{\mu\nu} - \tfrac{1}{2}\bar{\mathcal{S}}^i\eta_{\mu\nu}\Big)\Big]_{,i} = -4\pi G\rho\delta_{\mu\nu}, \quad (47)$$

where $\bar{\mathcal{S}}^i_{\mu\nu}$ is the NR limit of $\tilde{\mathcal{S}}^{\lambda}_{\mu\nu}$, and $\bar{\mathcal{S}}^i$ its (four) trace. The $(0i)$ components of the equations hold identically (to first order), since $\bar{\mathcal{S}}^i_{0j} = 0$.[11] The (00) and (ij) components give, respectively,

$$-\Delta\phi + [\mathcal{M}'(\bar{\mathcal{S}}^k_{00} + \tfrac{1}{2}\bar{\mathcal{S}}^k)]_{,k} = -4\pi G\rho, \quad (48)$$

$$\tfrac{1}{2}H_{ij} - \Delta\phi\delta_{ij} + [\mathcal{M}'(\bar{\mathcal{S}}^k_{ij} - \tfrac{1}{2}\bar{\mathcal{S}}^k\delta_{ij})]_{,k} = -4\pi G\rho\delta_{ij}. \quad (49)$$

Multiply Eq. (48) by δ_{ij} and subtract from Eq. (49) to get

$$\tfrac{1}{2}H_{ij} + [\mathcal{M}'(\bar{\mathcal{S}}^k_{ij} - \bar{\mathcal{S}}^k_{mm}\delta_{ij})]_{,k} = 0, \quad (50)$$

which I use instead of Eq. (49). Equation (50) does not satisfy any more identities and thus gives six independent equations, which together with the above four unused independent equations, and the remaining freedom to choose three gauge conditions, should determine the remaining 13 potentials ϕ, h_{ij}, h^*_{ij}.

It is beneficial to employ three of these six equations encapsuled in Eq. (50) by taking its divergence, taking the Bianchi identities for H_{ij} into account, to get

$$(\mathcal{M}'\bar{\mathcal{S}}^k_{ij})_{,k,j} = 0. \quad (51)$$

These three equations, together with the three independent equations in the second of (45) now involve only the six h^*_{ij} (so we managed to decouple these from ϕ and h_{ij}; ϕ^*, which also appears in these equations, is already known), and can be solved for these.[12] Once this is done (imposing boundary conditions at infinity), ϕ is determined from Eq. (48) by solving a Poisson equation, and h_{ij} are likewise determined from Eq. (50) with the aid of gauge conditions. The remaining gauge freedom is associated with coordinate transformations that preserve our assumed form of the

[10]Because the metric derivatives, connections, and curvature components are already first order, all the metrics that are used to contract them can be taken as $\eta_{\mu\nu}$.

[11]To see this, note that from Eq. (37), we have $\bar{\mathcal{S}}^i_{0j} \propto (2Q^{\beta\gamma ji}_{\alpha 0} - Q^{\beta\gamma 0i}_{\alpha j})C^{\alpha}_{\beta\gamma}$. Now, the NR limit of the Q tensor is constructed only from $\eta_{\mu\nu}$ and $\eta^{\mu\nu}$. This means that its only nonvanishing components must have three pairs of equal indices. This means, in turn, that the only contributions to $\bar{\mathcal{S}}^i_{0j}$ come from $C^{\alpha}_{\beta\gamma}$ with one or three time indices, but these all vanish.

[12]These are coupled nonlinear equations, since h^*_{ij} appear also in the argument as $\mathcal{M}'$.

metrics, i.e., near-Minkowski and stationary (time independent and lacking mixed elements). These are of the general form

$$t = t', \qquad x^i = x'^i + \epsilon^i(\mathbf{x}'), \tag{52}$$

with $\epsilon^i(\mathbf{x}')$ first order in the potentials. They do not affect $g^*_{\mu\nu}$ and so leave h^*_{ij} and ϕ^* intact,[13] changing only h_{ij} by $\epsilon_{i,j} + \epsilon_{j,i}$.

Everything so far is valid for an arbitrary choice of quadratic scalar argument. I now specialize to my preferred choice of scalar argument $-\Upsilon/2$. In this case, we have

$$\begin{aligned}\bar{\mathcal{S}}^i_{00} &= 2\phi^*_{,i} + \tfrac{1}{2}(h^*_{ij,j} - h^*_{,i}),\\ \bar{\mathcal{S}}^i_{jk} &= \tfrac{1}{2}(h^*_{ij,k} + h^*_{ik,j} - h^*_{jk,i}) + \tfrac{1}{4}[2(h^*_{,i} - h^*_{im,m})\delta_{jk}\\ &\quad - h^*_{,k}\delta_{ij} - h^*_{,j}\delta_{ik}].\end{aligned} \tag{53}$$

What is special about this case is that the space components $\bar{\mathcal{S}}^k_{ij}$ depend only on h^*_{ij}, not on ϕ^*. This greatly simplifies the solution of Eqs. (45) and (51), which has to be $h^*_{ij} = 0$ (with boundary conditions $h^*_{ij} \to 0$ at infinity).[14]

The fact that $h^*_{ij} = 0$ causes the $C^\alpha_{\beta\gamma}$, as given in Eq. (40), to be linear combinations of derivatives of ϕ^*, and the argument of $\mathcal{M}'$ becomes a function of $(\vec{\nabla}\phi^*)^2$. Now Eq. (48) reads

$$\Delta\phi = 4\pi G\rho + \vec{\nabla}\cdot\{\mathcal{M}'[(\vec{\nabla}\phi^*/a_0)^2]\vec{\nabla}\phi^*\}. \tag{54}$$

Equation (50) becomes

$$H_{ij} = 0 \tag{55}$$

and can be used with three gauge conditions, to determine h_{ij}–as would be done in GR, where this equation is always satisfied. This implies, with the appropriate boundary conditions, that there is a gauge in which $h_{ij} = 0$, and we work in this gauge.

The matter action for a system of slowly moving masses is, to our present approximation,

$$I_M \approx \frac{1}{2}\int \rho(\mathbf{v}^2 - \phi)d^3x dt. \tag{56}$$

So the motion of such particles is governed by the potential ϕ, which is thus identified as the MOND potential. It is determined from the field equation in Eq. (54) (with ϕ^* being the Newtonian potential of the system), which is the quasilinear MOND formulation described by Eq. (3), and discussed at length in [12]. Note that it requires solving only linear differential equations. To have the required Newtonian limit, we have to have $\mathcal{M}'(z) \to 0$ for $z \to \infty$ (i.e., $a_0 \to 0$). In the MOND regime $z \ll 1$, we have to have $\mathcal{M}'(z) \approx z^{-1/4}$ to get space-time scale invariance, which is the defining tenet of the NR MOND limit [20] (the normalization is absorbed in the definition of a_0).

To recapitulate, we end up with the simple result in the chosen gauge:

$$g_{\mu\nu} = \eta_{\mu\nu} - 2\phi\delta_{\mu\nu}, \qquad \hat{g}_{\mu\nu} = \eta_{\mu\nu} - 2\hat{\phi}\delta_{\mu\nu}, \tag{57}$$

with $\phi^* = \phi - \hat{\phi}$ being the Newtonian potential, and the MOND potential ϕ determined from the quasilinear MOND Eq. (3). The relation between the first-order MOND metric and the MOND potential is thus exactly the same as that between the first-order GR metric and the Newtonian potential.

Suppose we have not assumed *a priori* that the mixed elements of the first-order metrics vanish. This does not change expressions (40) for $C^\alpha_{\beta\gamma}$, but we now have additional nonvanishing elements

$$C^i_{0j} = (h^*_{0i,j} - h^*_{0j,i})/2, \qquad C^0_{ij} = -(h^*_{0i,j} + h^*_{0j,i})/2. \tag{58}$$

The $(0i)$ component of the difference Ricci tensor is $R^*_{0i} = (\Delta h^*_{0i} - h^*_{0j,i,j})/2$. So Eq. (45) is now complemented by

$$(\Delta h^*_{0i} - h^*_{0j,i,j}) = 0. \tag{59}$$

With our boundary conditions $h^*_{0i} \to 0$ at infinity, the solution of this equation is $h^*_{0i} = v^*_{,i}$ for some $v^*(\mathbf{x})$ (the left-hand side is identically divergence free, which leaves us with only two independent equations). Note now that the first-order limit of the theory, discussed here, enjoys a less obvious symmetry beside the general invariance to simultaneous coordinate transformations: it is invariant under a "small" transformation of the form $t = t' - u(\mathbf{x}')$ applied *separately* to the $g_{\mu\nu}$ and the $\hat{g}_{\mu\nu}$ sectors (with u first order in the potentials). In other words, there is symmetry to transforming $g_{0i} \to g_{0i} + u_{,i}$, $\hat{g}_{0i} \to \hat{g}_{0i} + \hat{u}_{,i}$ (hence $g^*_{0i} \to g^*_{0i} + u^*_{,i}$; $u^* = u - \hat{u}$), with u and $\hat{u}$ free for us to choose (the EMTs are unchanged to lowest order). We thus have the freedom to choose a gauge for $\hat{g}_{\mu\nu}$ alone, in which $v^* = 0$, and hence $h^*_{0i} = 0$. This means that $C^i_{0j} = C^0_{ij} = 0$, and so the $(0i)$ components of Eq. (47) read

$$(\Delta h_{0i} - h_{0j,i,j}) = 0, \tag{60}$$

whose solution is $h_{0i} = v_{,i}$ for some $v(\mathbf{x})$. We still have the gauge freedom to chose $v(\mathbf{x}) = 0$, and so we do. We thus end up with a gauge in which the field equations themselves dictate $h_{0i} = \hat{h}_{0i} = 0$, as we assumed *a priori*.

The double gauge symmetry we use can be seen to apply directly to the first-order equations. It can be traced back to the fact that the NR limit of Υ is invariant to it: If we do not assume *a priori* that the mixed elements h^*_{0i} vanish, then the addition to the lowest order expression for Υ is $C^i_{0k}C^k_{0i}$,

[13] Since $g^*_{\mu\nu}$ is already first order, in our approximation the transformation affects it only to zeroth order; i.e., not at all.

[14] All of the above holds if we use both Υ and Υ^* as variables, because they degenerate into one in the NR limit. Also, note that $\bar{C}^0 = 0$, and $\bar{C}^i = h^*_{ij,j} - (1/2)h^*_{,i}$; so, adding the scalar $g_{\mu\nu}\bar{C}^\mu\bar{C}^n$ should not change this conclusion.

because C^i_{0j} is antisymmetric in i, j, while C^0_{ij} is symmetric. But, under the double gauge transformation, h^*_{0i} changes by $u^*_{,i}$, so C^i_{0k} is invariant, and so is Υ.

It remains to be checked if this symmetry is a remnant of some symmetry enjoyed by the relativistic theory itself.

Anticipating the discussion of the next subsection, note that not all scalar arguments are invariant to this double gauge in their NR limit. For example, the change induced in the scalar $g_{\mu\nu}\bar{C}^\mu\bar{C}^\nu$ is $-h^*_{0i,i}\Delta u^*$. For the more general case, Eq. (59) is still valid, and again gives $h^*_{0i} = v^*_{,i}$, while instead of Eq. (60), we have more generally

$$(\Delta h_{0i} - h_{0j,j,i})/2 + [\mathcal{M}'\bar{\mathcal{S}}^k_{0i}]_{,k} = 0. \tag{61}$$

The first term is identically divergence free, so we can write one of these three equations as

$$[\mathcal{M}'\bar{\mathcal{S}}^k_{0i}]_{,k,i} = 0. \tag{62}$$

Now, $\bar{\mathcal{S}}^k_{0i}$ is linear in v^* (which may also appear in the argument of $\mathcal{M}'$), so this equation generically dictates $v^* = 0$. For the scalar Υ, this does not work, because at this stage we already have $\bar{\mathcal{S}}^k_{0i} = 0$, but in return we have the double gauge freedom to help us remove v^*. For the scalar $g_{\mu\nu}\bar{C}^\mu\bar{C}^\nu$, we do not have the double gauge freedom, but we can write $\bar{\mathcal{S}}^k_{0i} \propto \delta^k_i\Delta v^*$, so Eq. (62) gives $\Delta(\mathcal{M}'\Delta v^*) = 0$, which implies $v^* = 0$. In any event, we can always have $h^*_{0i} = 0$ and continue from there as before to show that there is always a gauge where the mixed elements of the metrics vanish.

A. Other choices of the scalar argument of $\mathcal{M}$

Here, I consider the NR limit of theories with other choices of the quadratic scalar argument of $\mathcal{M}$. The main purpose is to see whether these give theories that are different in their NR limits, and so can be distinguished using observations of NR systems such as rotation curves and lensing in galaxies.

Take then the general quadratic argument as given by Eq. (36). All of our procedures in the present section, up to Eq. (52), remain valid. Up to that point, we had not made use of the particular expressions for $\tilde{\mathcal{S}}^\lambda_{\mu\nu}$ and $\hat{\mathcal{S}}^\lambda_{\mu\nu}$, only of the fact that they become equal to first order in the potentials, and this is still the case.[15]

Departure from the above occurs, however, for more general scalars in the employment of Eq. (51). Now, the space components $\bar{\mathcal{S}}^i_{jk}$ do, in general, depend on the gradient of ϕ^*, and it is easy to see that $h^*_{ij} = 0$ is no longer a solution, in general: Substituting $h^*_{ij} = 0$ in Eq. (51) would result in three constraints on the Newtonian potential ϕ^*, which it does not satisfy (ϕ^* is really arbitrary if we allow arbitrary density distributions, including negative ones). So, in general, after ϕ^* is calculated as the Newtonian potential of the system, we have to solve the nine coupled Eqs. (45) (second part) and (51) (only six of which are independent) for the h^*_{ij}. After this is done, we determine ϕ from Eq. (48), and h_{ij} from Eq. (50) with the aid of gauge conditions. Note that in the Newtonian limit, $a_0 \to 0$, ϕ becomes the Newtonian potential, and $h_{ij} \to 0$ as fast as $\mathcal{M}'$ does.

We can derive some scaling properties of the h^*_{ij}, even in the general case: It is easy to see that the second Eq. (45) and (51) are invariant under $m_i \to \lambda m_i$, $\mathbf{r} \to \lambda^{1/2}\mathbf{r}$, where m_i stand for the masses in the system. This is because the only quantities appearing in these equations, beside the variables h^*_{ij}, are $\phi^*_{,i}$ and a_0 both of the same dimensions as m_iG/r^2. This means that the $h^*_{ij,k}$ are invariant under this scaling.

Consider now in more detail a spherically symmetric problem. We are looking for solutions in which the various potential tensors h_{ij}, $\hat{h}_{ij}$, h^*_{ij} are of the form exemplified by

$$h^*_{ij} = h^*_1(r)\delta_{ij} + h^*_2(r)n_in_j, \tag{63}$$

where $\mathbf{n} = \mathbf{r}/r$. It can be shown that h^*_{ij} satisfying the second Eq. (45), namely, annuling H^*_{ij}, is tantamount to $h^*_2 = rh^{*\prime}_1$, which means, in turn, that $h^*_{ij} = q_{,i,j}$ for some $q(r)$. It thus remains to determine q from Eq. (51). Because of the spherical symmetry, the different i components of Eq. (51) give equivalent equations[16]; so we are left with only one equation from which to determine $q(r)$.

Once $q(r)$ is known, we use Eq. (50) to solve for h_{ij}. Write h_{ij} in the form (63). We can use the remaining gauge freedom to eliminate one of the two functions. In the spherically symmetric case, the remaining freedom is to transform $\mathbf{r} = \mathbf{r}'[1 + \epsilon(r')]$ for some $\epsilon(r')$, treated to first order. This transformation takes $h_{ij} \to h_{ij} + 2\epsilon\delta_{ij} + 2r\epsilon' n_in_j$; so, we can use such a transformation to eliminate either of the functions in the expression for h_{ij}. For example, let us choose the gauge in which

$$h_{ij} = \varphi(r)\delta_{ij}. \tag{64}$$

The general NR MOND metric is thus diagonal for this choice of gauge with

$$g_{00} = -1 - 2\phi, \qquad g_{ij} = \delta_{ij}[1 - 2(\phi + \varphi)]. \tag{65}$$

For the form (64) of h_{ij}, we have

$$H_{ij} - \tfrac{1}{2}H\delta_{ij} = -(\varphi'' + r^{-1}\varphi')\delta_{ij} + (\varphi'' - r^{-1}\varphi')n_in_j \equiv a\delta_{ij} + bn_in_j. \tag{66}$$

Note that $a' = -r^{-2}(r^2b)'$, from the fact that the expres-

[15] This follows from the asymmetry of $C^\alpha_{\beta\gamma}$ to interchange of the matrices, and from the fact that to first order we can put everywhere else $\hat{g}_{\mu\nu} \approx g_{\mu\nu} \approx \eta_{\mu\nu}$.

[16] This equation has to read in the spherical case $P[q(r)]\mathbf{r} = 0$, where P is a differential operator acting on $q(r)$, and we get one equation $P[q(r)] = 0$.

sion is divergence free. We now use

$$H_{ij} - \tfrac{1}{2}H\delta_{ij} = -2(\mathcal{M}'\bar{\mathcal{S}}^k_{ij})_{,k}, \tag{67}$$

obtained from Eq. (50), to solve for φ. Since the right-hand side of this equation is already known to be divergence free, from Eq. (51), we get only one independent equation of the form

$$r(r^{-1}\varphi')' = p(r), \tag{68}$$

where $p(r)$ is a known function. This we finally solve for φ, which we permit to behave asymptotically as $\ln(r)$.

Here, we note already an interesting difference from the theories that have Υ as scalar argument, which have Eqs. (2) and (3) as their NR limit (even with general α, β–see the next subsection). In such theories, in the spherical case the MOND acceleration is an algebraic function of the Newtonian one, with the relation being unique for the theory. In the general case this is not so: To get the MOND acceleration in the spherical case, we apply the Gauss theorem to Eq. (48). The expression we then get is some functional of $q(r)$ that cannot be written as a function of the Newtonian acceleration $-d\phi^*/dr$. This can lead to different predictions even for massive-particle motions.

The spherical problem can be solved analytically for the case where $\phi^* = Ar^\theta$–for example, when we are outside the mass where $\theta = -1$, and we are in a region where $M'(z) \propto z^{-\sigma}$, for example, in the deep-MOND regime where we will have to have $\sigma = 1/4$. Then, the solution can be shown to be of the form $q = \lambda r^2\phi^*$, with λ determined from Eq. (51) depending on θ and σ. With this ansatz, $\mathcal{M}' \propto (|A|/a_0)^{-2\sigma}(a + b\lambda + c\lambda^2)^{-\sigma} r^{-2\sigma(\theta-1)}$, and Eq. (51) then gives $(A/a_0)^{-2\sigma}(a+b\lambda+c\lambda^2)^{-\sigma}A(\bar{a}+\bar{b}\lambda)r^\zeta\mathbf{n}=0$, with $\zeta = (1-2\sigma)(\theta-1)-2$ ($=-3$ for the above examples), and $\bar{a}$, $\bar{b}$ depending on θ, σ and the choice of scalar argument of $\mathcal{M}$ for the specific theory ($\bar{a}$ comes from the terms in $\bar{\mathcal{S}}^k_{ij}$ linear in the gradient of ϕ^*, and $\bar{b}$ from those linear in the gradient of h^*_{ij}). So $\lambda = -\bar{a}/\bar{b}$ gives us the solution (for the choice of Υ as argument $\bar{a}=0$). Equation (68) then gives $\varphi = \xi(A/a_0)^{-2\sigma}Ar^{\zeta+3}$ for $\zeta \neq -3$, and $\varphi = \xi(A/a_0)^{-2\sigma}\times A\ln(r)$ for $\zeta = -3$, with the dimensionless ξ determined.

Take, for instance the interesting case where we are asymptotically outside matter and in the deep-MOND regime, where $A = -MG$ and $\zeta = -3$. We then get $\varphi = \xi(MGa_0)^{1/2}\ln(r)$, asymptotically. In this case, we also have $\phi = (MGa_0)^{1/2}\ln(r)$, by definition; so, $\varphi = \xi\phi$. For example, for the choice of argument $-g^{\mu\nu}C^\gamma_{\mu\lambda}C^\lambda_{\nu\gamma}/2a_0^2$, I find, following the above procedure for the deep-MOND ($\sigma = 1/4$), asymptotic ($\theta = -1$) case: $\xi = 4/3$.

In summary, the asymptotic form of the MOND metric is diagonal, with

$$g_{00} = -1 - 2\phi, \qquad g_{ij} = \delta_{ij}[1 - 2\phi(1+\xi)]. \tag{69}$$

Remember that while $\varphi = \xi\phi$ in the MOND regime, in the Newtonian regime ϕ becomes the Newtonian potential, while φ vanishes as fast as dictated by the vanishing of $\mathcal{M}'$ at high values of its argument.

B. General $\alpha\beta$ values

The NR limit of the field equation in a theory governed by the action (13) for general α and β values is

$$\beta(R_{\mu\nu} - \tfrac{1}{2}\eta_{\mu\nu}R) + (\mathcal{M}'\bar{\mathcal{S}}^i_{\mu\nu})_{,i} = -8\pi G\rho\delta_{\mu0}\delta_{\nu0}, \tag{70}$$

$$\alpha(\hat{R}_{\mu\nu} - \tfrac{1}{2}\eta_{\mu\nu}\hat{R}) - (\mathcal{M}'\bar{\mathcal{S}}^i_{\mu\nu})_{,i} = 0, \tag{71}$$

with all quantities taken to first order in the potentials ϕ, $\hat{\phi}$, h_{ij}, $\hat{h}_{ij}$. $\bar{\mathcal{S}}^i_{\mu\nu}$ is linear in the first-order expression for $C^\alpha_{\beta\gamma}$, as given in Eq. (40), and $\mathcal{M}'$ is a function of a quadratic scalar built from these. Multiply the second equation by $-\beta/\alpha$ and add to the first to get

$$\beta\left(R^*_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}R^*\right) + \frac{\alpha+\beta}{\alpha}(\mathcal{M}'\bar{\mathcal{S}}^i_{\mu\nu})_{,i} = -8\pi G\rho\delta_{\mu0}\delta_{\nu0}. \tag{72}$$

The $(0j)$ components of this equation hold identically, as before. Equations (41)–(43) are then used to write the (00) and (ij) components of this equation as

$$\beta\left(\Delta\phi^* - \frac{1}{8}H^*\right) - \frac{\alpha+\beta}{2\alpha}(\mathcal{M}'\bar{\mathcal{S}}^i_{00})_{,i} = 4\pi G\rho,$$
$$H^*_{ij} - \frac{1}{2}H^*\delta_{ij} + \frac{2(\alpha+\beta)}{\alpha\beta}(\mathcal{M}'\bar{\mathcal{S}}^k_{ij})_{,k} = 0. \tag{73}$$

Taking the trace of the second and substituting for H^* in the first, we get

$$\beta\Delta\phi^* - \frac{\alpha+\beta}{2\alpha}[\mathcal{M}'(\bar{\mathcal{S}}^i_{00} + \bar{\mathcal{S}}^i_{mm})]_{,i} = 4\pi G\rho. \tag{74}$$

Equation (73) comprises seven independent equations that can be solved for ϕ^* and the h^*_{ij}. Once these are known, we can get ϕ and h_{ij} from Eq. (70), or equivalently from

$$\beta R_{\mu\nu} + [\mathcal{M}'(\bar{\mathcal{S}}^i_{\mu\nu} - \tfrac{1}{2}\bar{\mathcal{S}}^i\eta_{\mu\nu})] = -4\pi G\rho\delta_{\mu\nu}. \tag{75}$$

Its (00) component gives

$$\Delta\phi = 4\pi G\beta^{-1}\rho + \beta^{-1}[\tfrac{1}{2}\mathcal{M}'(\bar{\mathcal{S}}^i_{00} + \bar{\mathcal{S}}^i_{kk})]_{,i}, \tag{76}$$

from which ϕ can be obtained (since the right-hand side is now known). The (ij) component gives [after making use of Eq. (76)]

$$H_{ij} - \tfrac{1}{2}H\delta_{ij} = -2\beta^{-1}(\mathcal{M}'\bar{\mathcal{S}}^k_{ij})_{,k}. \tag{77}$$

From these, h_{ij} can be obtained. Note that the divergence of Eq. (77) is identically the same as that of the second Eq. (73), so we only have here three independent equations for the six h_{ij}. However, again we have the gauge freedom to eliminate this indeterminacy.

Specializing to our preferred choice of scalar $\Xi = -\Upsilon/2$, the NR limit of the theory again greatly simplifies since from Eq. (53), $\bar{\mathcal{S}}^k_{ij}$ does not contain the derivatives of ϕ^*. This implies that the second term in the second Eq. (73) is linear in the derivatives of h^*_{ij}, and so the solution of this equation is easy to get: $h^*_{ij} = 0$ (again, with the asymptotic boundary conditions $h^*_{ij} \to 0$). This means that the argument of $\mathcal{M}'$ is now a function of $(\vec{\nabla}\phi^*/a_0)^2$, and that, in fact, $\bar{\mathcal{S}}^i_{jk} = 0$. As a result, the first of Eq. (73) becomes identical with the first of Eq. (2), while Eq. (76) becomes identical with the second of Eq. (2). In addition, from Eq. (77), we get $h_{ij} = 0$ as before.

We thus end up with an NR limit in which $g_{\mu\nu} = \eta_{\mu\nu} - 2\phi\delta_{\mu\nu}$ and $\hat{g}_{\mu\nu} = \eta_{\mu\nu} - 2\hat{\phi}\delta_{\mu\nu}$ as in Eq. (57), with the MOND potential ϕ determined now from the NR MOND theory described by Eq. (2). This theory has been discussed at length in [12]. The relation between the first-order MOND metric and the MOND potential is thus, again, the same as that between the first-order GR metric and the Newtonian potential.

Note in this context as well, that using a scalar argument that is a combination of Υ and $\bar{\Upsilon} = g_{\mu\nu}\bar{C}^\mu\bar{C}^\nu$, leads to the same first-order metric and NR limit.

C. Other backgrounds

So far, I assumed that the two metrics have the same cosmological background and so, for systems small on the cosmological scale, both can be taken as nearly Minkowski.

Here, I consider some possible departures from this assumption. One possibility, for example, is that for today's cosmology the two metrics are conformally related $\hat{g}_{\mu\nu} = \lambda g_{\mu\nu}$ with constant λ; this still gives $C^\alpha_{\beta\gamma} = 0$ for the cosmological background. In this case, we can take locally the background metrics to be $g^B_{\mu\nu} = \eta_{\mu\nu}$, $\hat{g}^B_{\mu\nu} = \lambda\eta_{\mu\nu}$ and expand around these:

$$g_{\mu\nu} = \eta_{\mu\nu} - 2\phi\delta_{\mu\nu} + h_{\mu\nu},\qquad \hat{g}_{\mu\nu} = \lambda(\eta_{\mu\nu} - 2\hat{\phi}\delta_{\mu\nu} + \hat{h}_{\mu\nu}), \tag{78}$$

instead of Eq. (38). The expressions of $\hat{\Gamma}^\alpha_{\beta\gamma}$, $C^\alpha_{\beta\gamma}$, and $\hat{R}_{\mu\nu}$ in terms of the potentials remain the same as before. Also, small coordinate transformations of the type shown in Eq. (52) still do not affect the potential differences only the h_{ij}. Work with the general theory for arbitrary α, β. The NR limit of the field equations is now

$$\beta(R_{\mu\nu} - \tfrac{1}{2}\eta_{\mu\nu}R) + \lambda f(\lambda^{-1})(\mathcal{M}'\bar{\mathcal{S}}^i_{\mu\nu})_{,i} = -8\pi G\mathcal{T}_{\mu\nu}, \tag{79}$$

$$\alpha(\hat{R}_{\mu\nu} - \tfrac{1}{2}\eta_{\mu\nu}\hat{R}) - f(\lambda^{-1})(\mathcal{M}'\bar{\mathcal{S}}^i_{\mu\nu})_{,i} = 0, \tag{80}$$

with the λ powers coming from factors such as $\kappa_\pm$, etc. Defining $\tilde{\alpha} = \lambda\alpha$, and $\tilde{\mathcal{M}}' = \lambda f(\lambda^{-1})\mathcal{M}'$, we get back the $\lambda = 1$ case, but with α replaced by $\tilde{\alpha}$, and $\mathcal{M}'$ by $\tilde{\mathcal{M}}'$. As before, with our favorite choice of scalar argument $\Xi = -\Upsilon/2a_0$, we have $h_{ij} = 0$ in the appropriate gauge; so we get the relation between the MOND metric and the MOND potential as before.

For a more general background, we can write the background metrics, locally for a small system,

$$g_{\mu\nu} = \eta_{\mu\nu}, \qquad \hat{g}_{\mu\nu} = \lambda(\eta_{\mu\nu} - u\delta_{\mu 0}\delta_{\nu 0}). \tag{81}$$

This leads to more complex NR limits, which I do not discuss here.

We see then that the NR limit of the theory as applied to system that are small on cosmic scale depends on the background metrics. If the relation between the background metrics vary with cosmological time (I assume that it does not–see Sec. VI) the application of BIMOND to local inhomogeneities also varies with cosmic time. But I will not discuss this possibility further here.

IV. GENERAL RELATIVITY LIMIT

I showed in [12] that the NR theories with $\beta = 1$, but arbitrary α, have a Newtonian limit if $\mathcal{M}'(z) \to 0$ for $z \to \infty$. This carries over to the relativistic theories. The same property of $\mathcal{M}$ causes the relativistic theory to go to GR in the same limit $a_0 \to 0$, because $\mathcal{M}' \to 0$ implies $\tilde{T}_{\mu\nu}$, $\hat{T}_{\mu\nu} \to 0$, and since in this case $\mathcal{M}(z)/z$ must also vanish in the limit, we also have Λ_m, $\hat{\Lambda}_m \to 0$. This gives GR, with $g_{\mu\nu}$ satisfying the Einstein equation with the standard matter EMT as source. In this limit, $\hat{g}_{\mu\nu}$ satisfies its own Einstein equation. It decouples from $g_{\mu\nu}$ anyway, so it affects matter neither directly nor indirectly.

We do not have tight phenomenological constraints on how fast MOND approaches Newtonian dynamics for high accelerations. There are indications from solar system constraints [1,21,22] that $\mathcal{M}'(z)$ decreases at least as fast as z^{-1}, but it might turn out to do so much more precipitously. If so, any departure from GR might be practically wiped out in a very high-acceleration system, such as a laboratory on Earth, the inner solar system, or a close binary pulsar.[17,18]

In the limit $a_0 \to 0$, the vanishing of $\tilde{T}_{\mu\nu}$ and $\hat{T}_{\mu\nu}$ might occur much faster than that of Λ_m and $\hat{\Lambda}_m$. For example, if $\mathcal{M}(\infty)$ is finite, the latter vanishes as a_0^2, while the former might vanish much faster. We can thus, as an intermediate approximation, keep the Λ_m and $\hat{\Lambda}_m$ terms in the theory,

[17]The presence of the galactic field still induces the departures discussed in [21].

[18]A theory like TeVeS also has a surrogate for a_0 that appears in its NR limit, which is constructed out of the constants characterizing the theory. However, TeVeS does not become GR in the limit $a_0 \to 0$, and this leaves possibly detected effects even in an isolated solar system, the binary pulsar, etc., even with the very high accelerations characterizing them [3,8,23–25].

and write the limiting field Eqs. (16) and (17) as

$$R_{\mu\nu} - \tfrac{1}{2}Rg_{\mu\nu} - \Lambda_m(\infty)g_{\mu\nu} = -8\pi G\mathcal{T}_{\mu\nu}, \qquad (82)$$

$$\hat{R}_{\mu\nu} - \tfrac{1}{2}\hat{R}\hat{g}_{\mu\nu} - \hat{\Lambda}_m(\infty)\hat{g}_{\mu\nu} = -8\pi G\hat{\mathcal{T}}_{\mu\nu}. \qquad (83)$$

For general β values, it remains to be checked whether the requirement on $\mathcal{M}'$ from the NR limit, deduced in [12], suffices to guarantee a GR limit for $a_0 \to 0$.

V. GRAVITATIONAL LENSING

One of the two main phenomenological duties we expect from a relativistic MOND theory is to predict gravitational lensing correctly. In particular, we know that lensing analysis of galactic systems indicate mass discrepancies that are not very different to those derived from massive-particle motions (e.g., rotation curves). In other words, lensing in the MOND regime is found observationally to be greatly enhanced over the GR prediction without DM. Reproducing this fact has been a pressing desideratum in constructing relativistic MOND theories, achieved finally in TeVeS through the efforts of Sanders [4] and Bekenstein [3].

In the present BIMOND class such enhanced, MOND-like lensing is predicted naturally by all the theories in the class. For a choice of the scalar argument that is a combinations of Υ, Υ^*, and $\bar{\Upsilon}$ (and with any α, β), relation (57) between the first-order MOND metric and the MOND potential holds. So, to this order the MOND connection $\Gamma^{\lambda}_{\mu\nu}$ is expressed in terms of the MOND potential ϕ in the same way as the GR connection is expressed in terms of the Newtonian potential. This means, in turn, that the MOND potential describes the dynamics of both massive and massless particles in the same way as the Newtonian potential does in GR. In other words, such theories predict that analyzing lensing and massive-particle dynamics by a NR system assuming GR, should give the same effective potential (or the same distribution of "phantom matter"). This is consistent with observations.

We also saw that there are other choices of the scalar argument of $\mathcal{M}$ for which the NR MOND metric is characterized by additional potentials h_{ij}. However, these vanish quickly for accelerations much above a_0, while in the MOND regime they are of the same order as the MOND potential ϕ. We then expect these theories to yield somewhat different lensing to that expected with the GR relation between metric and potential, albeit still with the MOND characteristics.

For example, we saw that far from a central mass M, and in the deep-MOND regime, the form of the MOND metric is given by Eq. (69), and this leads to lensing that is multiplied by a factor $1 + \xi/2$ over that expected from the MOND potential with the GR prescription.

Eventually, by comparing lensing and massive-particle dynamics in the low-acceleration fields of galaxies or other galactic systems, we may be able to differentiate observationally between theories with different scalar arguments.

VI. COSMOLOGY

I cannot at present offer a specific BIMOND cosmology. There are two obstacles to doing this. In the first place, we cannot even pinpoint the exact BIMOND theory out of the various versions possible. NR phenomenology can assist somewhat in pinpointing the NR limit. However, even for a given NR MOND theory, there are different relativistic versions having this limit, which differ greatly in the relativistic regime, and specifically in their application to cosmology.

Second, as we well know from a century of experience with GR, having the underlying theory is one thing; pinpointing the cosmology is another: In dealing with standard GR cosmology, the cosmological evolution and the present state of the Universe does not emerge uniquely from first principles. There are major observational constraints, assumptions about symmetries, initial conditions, and matter content, that are put in by hand into cosmological theory. We do not know, for instance, the initial conditions for our Universe from first principles, so an initial singularity (as opposed say to a steady state universe with continuous matter creation, or to a static universe as Einstein would have it initially) is imposed by hand. Early inflation (the mechanism for which is still moot) is put in by hand to account for various observational facts. Cosmic acceleration, whose cause remains unknown, is imposed by hand by invoking dark energy, modified gravity, or other mechanisms. The material content of the Universe (e.g. the very existence of baryon asymmetry–mechanism still unknown) is an input in cosmology. *A priori*, we could have had a cosmos with a space that is inhomogeneous or anisotropic on cosmological scales; but, the cosmological principle is imposed based on what our eyes tell us about our Universe. All of this is even more acute in light of recent developments in quantum cosmology.

In the case of BIMOND, we are on even shakier ground when coming to construct a cosmology. Here, we are dealing with two space-times, only one of which we can sense directly. We have no direct knowledge of many of the global properties of the other space-time. Did it have a big bang? Did it undergo inflation? Is it spatially flat (in Rosen's theory the auxiliary metric is constrained to be flat)? Is it spatially homogeneous and isotropic? What is the nature of the twin matter? Is it there at all? Is it always homogeneously distributed, or does it clump? Is it characterized by the same baryon asymmetry, etc.?

MOND phenomenology in systems that are small on cosmological scales is particularly simple and clear cut in a double Minkowski background. Such a background applies if in cosmology $\hat{g}_{\mu\nu} = g_{\mu\nu}$. BIMOND cosmology that has disparate metrics, so that the cosmological values of $C^{\alpha}_{\beta\gamma}$

are appreciable, does not seem to make phenomenological sense in small systems.

I thus assume, as an additional cosmological assumption to the many above, that on cosmological scales we have $\hat{g}_{\mu\nu} = g_{\mu\nu}$. This could emerge, as an external constraint on BIMOND, from the world picture that underlies it. Or, more appealingly, it could, at least, correspond to a solution of BIMOND itself in some version.

I thus consider here briefly only cosmologies with $\hat{g}_{\mu\nu} = g_{\mu\nu}$, or with the somewhat relaxed assumption $\hat{g}_{\mu\nu} = \lambda g_{\mu\nu}$, with a constant λ. In either case, we have $C^{\alpha}_{\beta\gamma} = 0$ in cosmology; so, finite values of $C^{\alpha}_{\beta\gamma}$ are produced only due to local inhomogeneities. This greatly simplifies the equations of motion, since the only contributions of the interaction term that survive are the Λ_m terms. With this ansatz, the equations of motion for the more general action (13) are then

$$\begin{aligned} \beta G_{\mu\nu} + q a_0^2 \mathcal{M}(0) g_{\mu\nu} &= -8\pi G \mathcal{T}_{\mu\nu}(g_{\mu\nu}, \psi_i), \\ \alpha G_{\mu\nu} + \hat{q} a_0^2 \mathcal{M}(0) g_{\mu\nu} &= -8\pi G \hat{\mathcal{T}}_{\mu\nu}(\lambda g_{\mu\nu}, \chi_i), \end{aligned} \tag{84}$$

where

$$\begin{aligned} q &= (\lambda/2)[\kappa f(\kappa)]'_{\kappa=\lambda^{-1}}, \\ \hat{q} &= -(\lambda^{-2}/2)[\kappa^{-1} f(\kappa)]'_{\kappa=\lambda^{-1}}, \end{aligned} \tag{85}$$

and I used the fact that with the above ansatz, $\hat{G}_{\mu\nu} = G_{\mu\nu}$. For the two equalities in Eq. (84) to hold simultaneously, we need to start with a BIMOND theory with some symmetry with respect to the two metrics. As an example, consider a special, more symmetric, case of the gravitational Lagrangian in Eq. (13) written as

$$\beta g^{1/2}(R - 2a_0^2 \bar{\mathcal{M}}) + \alpha \hat{g}^{1/2}(\hat{R} - 2a_0^2 \bar{\mathcal{M}}). \tag{86}$$

This choice corresponds to $f(\kappa) = (\beta\kappa + \alpha\kappa^{-1})/(\alpha + \beta)$, and $\mathcal{M} = (\alpha + \beta)\bar{\mathcal{M}}$. It gives $q = \beta/(\alpha + \beta)$, $\hat{q} = \lambda\alpha/(\alpha + \beta)$. So the field equations are now

$$\begin{aligned} G_{\mu\nu} + a_0^2 \bar{\mathcal{M}}(0) g_{\mu\nu} &= -8\pi G \beta^{-1} \mathcal{T}_{\mu\nu}(g_{\mu\nu}, \psi_i), \\ G_{\mu\nu} + \lambda a_0^2 \bar{\mathcal{M}}(0) g_{\mu\nu} &= -8\pi G \alpha^{-1} \hat{\mathcal{T}}_{\mu\nu}(\lambda g_{\mu\nu}, \chi_i). \end{aligned} \tag{87}$$

Clearly, $\hat{g}_{\mu\nu} = g_{\mu\nu}$ ($\lambda = 1$) always corresponds to a vacuum solution of this theory, with both space-times being a de Sitter or anti-de Sitter, with a cosmological constant $\Lambda = -a_0^2 \bar{\mathcal{M}}(0)$. Furthermore, if we also have from symmetry, for two identical configurations in the two sectors, $\hat{\mathcal{T}}_{\mu\nu} = (\alpha/\beta)\mathcal{T}_{\mu\nu}$,[19] the two equations are the same even with matter. The cosmology we then get is, quite interestingly, standard GR cosmology (taking $\beta = 1$) with Λ as cosmological constant. This would be reassuring, since it would automatically ensure that we are not bereft of the successes of standard cosmology regarding inflation, nucleosynthesis, etc.[20]

This picture would also force us to consider more seriously the nature of the twin matter, and its possible visible effects in our space-time. If it is homogeneously distributed, it will be difficult to detect any direct effects of it. If it clumps, it could have various effects; for example, it may produce some effects that are otherwise attributed to cosmological dark matter.

It is also possible that BIMOND can replace cosmological DM by the distribution and fluctuations in $\tilde{T}_{\mu\nu}$, which is constructed from the two metrics alone, not directly from matter. This would be similar in vein to what has been discussed in connection with such a possible role of auxiliary fields in other theories [8,18,23,26].

The above is only one example of a BIMOND theory that has a cosmology with $\hat{g}_{\mu\nu} = g_{\mu\nu}$. This particular version should not be assumed for the special case $\beta + \alpha = 0$, since then $\bar{\mathcal{M}}$ drops from the theory in the NR limit. However, there are other versions of BIMOND that can accommodate our cosmological ansatz for this case. For example, we can take as the gravitational Lagrangian density (say with $\beta = 1$)

$$\begin{aligned} & g^{1/2}(R - 2pa_0^2) - \hat{g}^{1/2}(\hat{R} - 2\hat{p}a_0^2) - 2(g\hat{g})^{1/4} f(\kappa) a_0^2 \mathcal{M} \\ &\quad = g^{1/2}R - \hat{g}^{1/2}\hat{R} - 2a_0^2(g\hat{g})^{1/4} \\ &\qquad \times [p\kappa - \hat{p}\kappa^{-1} + f(\kappa)\mathcal{M}]. \end{aligned} \tag{88}$$

This theory also has the NR limit we discussed above, and has a cosmological solution with $\hat{g}_{\mu\nu} = g_{\mu\nu}$ if a certain relation between p, $\hat{p}$, and $\mathcal{M}(0)$ holds. For example, if $\mathcal{M}(0) = 0$ we have the symmetric theory with $p = \hat{p}$, in which case $-pa_0^2$ is the cosmological constant. More generally, we can take the gravitational Lagrangian density (keeping the symmetry)

$$\begin{aligned} & g^{1/2}(R - 2\bar{\mathcal{M}}a_0^2) - \hat{g}^{1/2}(\hat{R} - 2\bar{\mathcal{M}}a_0^2) - 2(g\hat{g})^{1/4} f(\kappa) a_0^2 \mathcal{M} \\ &\quad = g^{1/2}R - \hat{g}^{1/2}\hat{R} - 2a_0^2(g\hat{g})^{1/4} \\ &\qquad \times [(\kappa - \kappa^{-1})\bar{\mathcal{M}} + f(\kappa)\mathcal{M}] \end{aligned} \tag{89}$$

(with $\bar{\mathcal{M}}$ also a function of the scalar argument Υ/a_0^2). Then, $\bar{\mathcal{M}}$ does not appear in the NR limit on a double Minkowski background, which is governed by $\mathcal{M}$, and a cosmology with our ansatz is a solution, with $-a_0^2\bar{\mathcal{M}}(0)$ as the cosmological constant [for $\mathcal{M}(0) = 0$].

We saw that in the GR limit, $a_0^2\mathcal{M}_\infty$ plays the role of a cosmological constant [$\mathcal{M}_\infty = \mathcal{M}(\infty)$], while in the present context it is $a_0^2\mathcal{M}(0)$. More generally, the $\Lambda_m \propto a_0^2\mathcal{M}$ term, which is variable, and possibly other terms in $\tilde{T}_{\mu\nu}$, may give rise to dark energy effects. $\mathcal{M}(z)$ need not change much for the full range of z, since for small values

[19]This ensures that without the interaction, physics is the same in the two sectors.

[20]This line of thinking seems to indicate that the theories with $\alpha = \beta = 1$are preferable.

it also has to go to some constant, as $\mathcal{M}(z) \approx \mathcal{M}_0 + (4/3)z^{3/4}$ for $z \ll 1$. In fact, changes in $\mathcal{M}(z)$ over the full z range are, generically, of order unity, since $\mathcal{M}_\infty - \mathcal{M}(z) = \int_z^\infty \mathcal{M}'(z)dz$ is of order unity if $\mathcal{M}'$ decreases beyond $z = 1$ faster than z^{-1} (unlike $\mathcal{M}$, which is known only up to an additive constant, $\mathcal{M}'$ is determined by MOND phenomenology). $\mathcal{M}_\infty$ is a dimensionless constant characterizing the theory. If $|\mathcal{M}_\infty| \sim 1$, then $|\mathcal{M}|$ is always of order unity. We then automatically get the well-known, but otherwise mysterious, proximity between a_0^2, as determined from the dynamics of small systems, and Λ, the density of dark energy, as deduced from cosmology.[21]

Note that irrespective of the relevance to MOND, bimetric theories of the type presented here provide a frame for discussing dark energy as modified gravity, which could be an alternative to schemes such as $f(R)$ theories.

Deep-MOND relativistic systems?

In principle, BIMOND theories enable one to study the structure of deep-MOND, relativistic systems such as deep-MOND black holes. As has been stressed many times in the past, such a deep-MOND system would have to have its typical curvature radius much larger than the MOND length $\ell = c^2/a_0$. This length is, however, of the order of the Hubble radius today, and certainly in the past. In practice then, the Universe seems to be the only such low-acceleration (rather, intermediate-acceleration) relativistic system, at present.

VII. DISCUSSION

I have described a class of BIMOND theories. Matter lives in the space-time described by one of the metrics, which, in turn, couples to another through the interaction $\mathcal{M}$ term. If we heuristically view gravity as reflecting an effective elasticity of space-time, we can view the double-metric nature of our theory as representing two coexisting elastic bodies, each with its own elasticity as encapsuled in the respective R, $\hat{R}$ terms in the action. Thus MOND departure from GR is introduced not through a modification of the elasticity properties of space-time, but rather through the interaction of the space-time that is the arena for matter with the auxiliary one. The strength of the interaction between these two space-time "membranes" depends on the gradient difference. The response of our home space-time to matter is affected by its interaction with the other space-time, which modifies its effective elasticity. However, once its shape is determined, this home space-time affects matter in the standard way. With our assumption that on cosmological scales $\hat{g}_{\mu\nu} = g_{\mu\nu}$, the two membranes are, in a sense, stuck together on these scales, and "separate" only locally due to inhomogeneities. Such heuristics may help pinpoint the fundamental concept underpinning the MOND paradigm. For example, it may give meaning to the length $\ell = c^2/a_0$ that appears in the NR limit as a_0.

The BIMOND theories have the (yet unproven) potential to account for all the components of the dark sector (galactic DM, cosmological DM, and dark energy) from one term in the action, all controlled by a_0.

My main objective has been to point out that there exists such a class of relativistic theories that have MOND-like theories as their NR limit, and, which produce enhanced, MOND-like gravitational lensing. We are, however, still far from pinpointing the exact version of the theory that is the most suitable. This is particularly true in the context of cosmology, which depends crucially on the choice of version. Hopefully, theoretical and phenomenological constraints will be brought to bear on this by future studies. It remains to be seen whether a version of BIMOND can be found that pass muster given all such requirements. Recent discussions of matter-of-principle questions, such as the causal structure of bimetric theories of a different type (where the interaction term is a function of the metrics themselves, not their derivatives) can be found, e.g., in [16,17].

Finally, it has to be realized that however useful such theories may turn out to be, they must be only effective, approximate theories, as evinced by the appearance of the *a priori* unspecified function $\mathcal{M}$ and the length ℓ (or the MOND acceleration a_0). These will, hopefully, be calculated from a theory at a deeper stratum.

ACKNOWLEDGMENTS

I am grateful to Jacob Bekenstein and Luc Blanchet for useful comments. This research was supported by a center of excellence grant from the Israel Science Foundation.

[21] The possible proximity of Λ and a_0^2, thus hinges on the dimensionless $\mathcal{M}$ being of order unity. Because of the way the normalization of $\mathcal{M}$ is defined, this means that the scale over which $\mathcal{M}$ varies as a function of Υ, and the scale that determines the magnitude of the $\mathcal{M}$ term in the Lagrangian, which have the same dimensions of length^{-2}, are also of the same magnitude. This need not be the case, just as not all mass parameters that appear in the standard model of particle physics are of similar values. So, the apparent proximity $\Lambda \sim a_0^2$ that we get here is only a plausibility not a corollary.

[1] M. Milgrom, Astrophys. J. **270**, 365 (1983).
[2] J. Bekenstein and M. Milgrom, Astrophys. J. **286**, 7 (1984).
[3] J. D. Bekenstein, Phys. Rev. D **70**, 083509 (2004).
[4] R. H. Sanders, Astrophys. J. **480**, 492 (1997).
[5] J. D. Bekenstein, Contemp. Phys. **47**, 387 (2006).
[6] T. G. Zlosnik, P. G. Ferreira, and G. D. Starkman, Phys. Rev. D **74**, 044037 (2006).
[7] T. G. Zlosnik, P. G. Ferreira, and G. D. Starkman, Phys. Rev. D **75**, 044017 (2007).
[8] C. Skordis, Classical Quantum Gravity **26**, 143001 (2009).
[9] L. Blanchet and A. Le Tiec, Phys. Rev. D **78**, 024031 (2008).
[10] L. Blanchet and A. Le Tiec, Phys. Rev. D **80**, 023524 (2009).
[11] L. Blanchet, Classical Quantum Gravity **24**, 3541 (2007).
[12] M. Milgrom, arXiv:0911.5464 [Mon. Not. R. Astron. Soc. (to be published)].
[13] N. Rosen, Ann. Phys. (N.Y.) **84**, 455 (1974).
[14] N. Boulanger, T. Damour, L. Gualtieri, and M. Henneaux, Nucl. Phys. **B597**, 127 (2001).
[15] T. Damour and I. I. Kogan, Phys. Rev. D **66**, 104024 (2002).
[16] D. Blas, C. Deffayet, and J. Garriga, Classical Quantum Gravity **23**, 1697 (2006).
[17] D. Blas, C. Deffayet, and J. Garriga, Phys. Rev. D **76**, 104036 (2007).
[18] M. Bañados, P. G. Ferreira, and C. Skordis, Phys. Rev. D **79**, 063511 (2009).
[19] R. P. Woodard, Lect. Notes Phys. **720**, 403 (2007).
[20] M. Milgrom, Astrophys. J. **698**, 1630 (2009).
[21] M. Milgrom, Mon. Not. R. Astron. Soc. **399**, 474 (2009).
[22] M. Sereno and Ph. Jetzer, Mon. Not. R. Astron. Soc. **371**, 626 (2006).
[23] R. H. Sanders, Mon. Not. R. Astron. Soc. **363**, 459 (2005).
[24] J. P. Bruneton and G. Esposito-Farese, Phys. Rev. D **76**, 124012 (2007).
[25] E. Sagi, Phys. Rev. D **80**, 044032 (2009).
[26] T. G. Zlosnik, P. G. Ferreira, and G. D. Starkman, Phys. Rev. D **77**, 084010 (2008).