Resolving the $H_0$ tension with diffusion

Alejandro Perez,¹,∗ Daniel Sudarsky,²,† and Edward Wilson-Ewing³,‡

¹Aix Marseille Univ, Université de Toulon, CNRS, CPT, Marseille, France
²Instituto de Ciencias Nucleares, Universidad Nacional Autónoma de México, México D.F. 04510, México
³Department of Mathematics and Statistics, University of New Brunswick, Fredericton, NB, E3B 5A3, Canada

The tension between the value of the Hubble constant $H_0$ determined from local supernovae data and the one inferred from the cosmic microwave background based on the ΛCDM cosmological model may indicate the need for new physics. Here, we show that this ‘Hubble tension’ can be resolved in models involving an effective energy flux from the matter sector into dark energy resulting naturally from a combination of unimodular gravity and an energy diffusion process. The scheme is one where dark energy has the standard equation of state $w = -1$. This proposal provides an alternative phenomenological paradigm accounting for the observations, while offering a general framework to study diffusion effects coming from novel fundamental physical processes.

I. INTRODUCTION

Over the last 30 years there has been a spectacular development of precision cosmology, with the state of the art including high resolution observations of the cosmic microwave background (CMB) by the Planck collaboration [1, 2], detailed studies of baryon acoustic oscillations by BOSS [3], and extensive supernovae observations [4, 5]. This trend is expected to continue in the near future with information coming from gravitational wave observations by the LIGO collaboration [6], as well as the data from ongoing and future missions like GAIA, EUCLID, and the James Webb Space Telescope. While these observations are, for the most part, consistent with the ΛCDM concordance cosmology based on a still mysterious dark sector including cold dark matter (CDM) and a cosmological constant Λ, the increased precision has brought with it further problematic aspects to our current understanding.

Dark matter was introduced phenomenologically in order to account for local gravitational effects, as observed in the rotation curves of galaxies, and to offer a viable paradigm for structure formation [7]. However, it remains unclear whether dark matter is indeed simply an invisible contribution to the matter content of the universe, or an effect due to modifications of general relativity [8]. Furthermore, while dark energy is perhaps most simply explained by a cosmological constant, from a theoretical perspective it remains unclear why its value is so small (and, in particular, why quantum vacuum fluctuations do not appear to contribute as naively expected [9]).

Dark energy is estimated to contribute about 70% to the mean energy density of the universe today, as first extracted from supernovae measurements [10, 11]. The situation has motivated numerous phenomenological alternatives to a cosmological constant, such as quintessence and modified gravity theories [8]. However, it is safe to say that no fully satisfactory fundamental explanation of these issues seems available at the moment.

More recently, a statistically significant tension has grown between the values of the Hubble expansion rate today, $H_0$, extracted from supernovae data, and the one inferred from the CMB observations based on the ΛCDM model [12, 13], with a higher value for $H_0$ preferred by supernovae data. This discrepancy is currently reported to be at the level of $4.4σ$, for studies based on supernovae observations calibrated using Cepheid stars [4] (although the tension is significantly reduced when using a different distance ladder calibration based on stars lying on the tip of the red giant branch [5]). Could this $H_0$ tension be offering a clue concerning some unknown aspect about the fundamental nature of gravity and matter?

Various proposals have been put forward in order to alleviate this tension. These typically involve modifying the standard ΛCDM cosmology in its dark sector: for example, dark radiation [12], early dark energy [14] or dynamical dark energy [15], and interacting dark matter/energy models [16–18]. The tension can also be alleviated (although only partially) for example either by allowing for the variation of Newton’s constant (and post-Newtonian parameters) [19], or by allowing for a more general equation of state $(w \neq -1)$ for the dark energy sector, with the data preferring the inclusion of regimes where $w < -1$ [2]. From a theoretical viewpoint, the last possibility seems rather problematic, as it would imply the violation of all energy conditions by the dark energy. Particularly worrying is the violation of the dominant energy condition, because, in any field theoretical model, that would imply the existence of a tachyonic degree of freedom and acausal energy fluxes.

Recently, a new mechanism has been proposed that generates an effective cosmological constant out of small departures from the strict conservation of energy momentum in the matter sector [20]. Although within the context of general relativity violations of the conservation of the stress-energy tensor are inconsistent, a slight modification of the theory known as unimodular gravity permits a specific kind of violation generating a time-dependent Λ in the resulting effective Einstein equations [20]. This

∗ Alejandro Perez, alessandro.perez@cpt.univ-mrs.fr
† sudarsky@nucleares.unam.mx
‡ edward.wilson-ewing@unb.ca
framework must, of course, be complemented with detailed input concerning the form and magnitude of the violations of energy-momentum conservation. If such violations are due to a diffusive process resulting from the space-time micro-structure or granularity in the fabric of space-time at the Planck scale, arising in conjunction with curvature (and therefore absent in the Minkowski space-time [21, 22]), the result is a cosmological constant of the order of magnitude of its observed value [23, 24]. It seems, therefore, natural to consider whether a related mechanism could be at play in late cosmological times leading to an effective modification of the cosmological evolution that could resolve the $H_0$ tension.

A simple analysis indicates that the specific model considered in [23, 24] is not able to account for any significant modification of $\Lambda$ in the late universe. However, [25] considers a related diffusion mechanism, connected in this case with the physics of black holes, which naturally places its occurrence after decoupling, and thus the relevant time period for a possible resolution of the $H_0$ discrepancy.

The purpose of the present work is to explore, in a rather general manner, the possibility that a process taking place at relatively late times, involving an effective violation of local energy conservation in the matter sector within the context of unimodular gravity, might resolve the $H_0$ discrepancy; due to the resulting growth of the dark energy at post-CMB times changing the late cosmic evolution. Such models could be regarded, at the phenomenological level, as versions of some models with interacting matter and dark energy such as those considered in [16–18], but the perspective offered by the unimodular framework implies the origin of the interaction comes from a diffusive process \(^1\) (see also [26]).

The outline of the paper is the following: in Sec. II, we review the framework of unimodular gravity that allows for the effective energy transfer from the matter to the dark energy sector. In Sec. III, we review the main observational inputs that lead to the $H_0$ tension and their dependence on the cosmological model. In Sec. IV, we consider some phenomenological models that can resolve the $H_0$ tension. We then offer, in Sec. V, a brief recount of some intermediate red-shift observations that seem to favor some general features of the phenomenological models considered in this work, and we end with a discussion in Sec. VI. To avoid confusion, in the remainder of the paper we will refer to the standard cosmological model as GR-ΛCDM.

II. UNIMODULAR GRAVITY AND VIOLATIONS OF ENERGY MOMENTUM CONSERVATION

General Relativity is incompatible with any violation of the conservation of the stress-energy tensor as a consequence of the (contracted) Bianchi identities of the Einstein tensor, i.e., $\nabla^a G_{ab} = 0$. There is, however, a simple modification of general relativity, known as unimodular gravity (UG) [27], introduced by Einstein himself in 1919, that is more permissive in this regard. The field equations of UG are simply the trace-free part of the Einstein equations, namely

$$R_{ab} - \frac{1}{4} R g_{ab} = 8 \pi G \left(T_{ab} - \frac{1}{4} T g_{ab}\right),$$

which can be written in the more convenient form

$$R_{ab} - \frac{1}{2} R g_{ab} + \frac{1}{4} (R + 8 \pi G T) g_{ab} = 8 \pi G T_{ab}. \quad (2)$$

Unimodular gravity can be derived from the Einstein-Hilbert action by restricting the variational problem to variations that preserve the (four) volume-form \(^2\), i.e., $g_{ab} \delta g^{ab} = 0$ (alternatively, it is possible to add a constraint to the action that requires the volume element of the metric to have a specific value). This is arguably the simplest modification of gravity that completely trivializes the problem of the large contributions to the cosmological constant from vacuum fluctuations in quantum field theory by fully decoupling vacuum energy from the dynamics of the metric [28].

Taking the divergence of (2) and using the Bianchi identities one finds

$$\frac{1}{4} \nabla_a \left(R + 8 \pi G T g_{ab}\right) = 8 \pi G \nabla^b T_{ab} = J_a, \quad (3)$$

where we have introduced the current of energy momentum violation $J_a = 8 \pi G \nabla^b T_{ab}$. The invariance of unimodular gravity under volume-preserving diffeomorphisms implies that $d J = 0$ [20], and therefore (3) can be integrated, leading to

$$R_{ab} - \frac{1}{2} R g_{ab} + \left(\Lambda_0 + \int J\right) g_{ab} = 8 \pi G T_{ab}, \quad (4)$$

\(^1\) The investigation of a possible diffusion mechanism is opened by the perspective adopted here. A possible model is proposed in [25] where the rotational energy of black holes is dissipated via friction produced by fundamental space-time granularity. The details of the energy diffusion during cosmological evolution is determined by the fundamental diffusion equation for each black hole $E = F(M, J)$ together with the expression for the number density of black holes as a function of their mass, angular momentum, and the cosmic time $n(M, J, t)$. In this model the detailed analysis of its effects necessitates the theoretical modeling of the cosmic evolution of black hole abundances. The viability of these models will be constrained in part by astrophysical bounds on the envisioned energy dissipation.

\(^2\) This modification would be natural if at a more fundamental level, as the full quantum gravity regime is approached, there are space-time related structures not fully characterized by the metric description, with the metric corresponding to some sort of ‘mean field’ or effective description of space-time, perhaps due to a fundamental granularity or a four-volume distribution of singular quantum events.
where \( \Lambda_0 \) is a constant of integration and the ‘energy violation current’ \( J \) is integrated along any arbitrary path \( \ell \) from some reference event to the point where the equation is evaluated. Clearly, this generates an effective \( \Lambda(x^a) = \Lambda_0 + \int_\ell J \) which depends on the space-time point. The independence of \( \Lambda(x^a) \) on the choice of path \( \ell \) is guaranteed by the integrability condition \( dJ = 0 \). An analogous system of equations can also be obtained by allowing for certain types of interactions between scalar-field dark energy and the matter fields in Einstein gravity [29, 30].

In the case that the stress-energy tensor is conserved, \( J = 0 \) and (4) reduces to the Einstein equations, with the integration constant \( \Lambda_0 \) becoming the cosmological constant. However, in contrast to general relativity, energy-momentum conservation \( \nabla^b T_{ab} = 0 \) does not follow directly from the equations of motion, and, in fact, it is usually postulated as an additional assumption. This assumption can be replaced by the demand that the matter action be diffeomorphism invariant (related to the assumption that space-time geometry and matter fields are smooth to all scales). For our purposes, the central observation is that, in the context of UG, this extra assumption can be relaxed, leading to a changing cosmological ‘constant’ \( \Lambda(t) \).

There are, indeed, various reasons to consider the possibility that \( \nabla^b T_{ab} \neq 0 \). First of all, the classical setting, where general relativity is typically applied, can be nothing more than a very good approximation. At some point, a quantum description of matter is required and that implies that, at a minimum (in the so-called semiclassical approximation), \( T_{ab} \) must be replaced by \( \langle T_{ab} \rangle \) (the expectation value of the (renormalized) energy-momentum operator in a suitable state). The viability of a semiclassical approximation has been a subject of substantial debate and controversy [31–33]. In fact, such considerations seem to connect, to a large extent, to conceptual difficulties inherent to quantum theory. One of these difficulties is that none of the currently proposed approaches addressing the measurement problem allows for the conservation of energy-momentum at the semiclassical level [34].

Another motivation for \( \nabla^b T_{ab} \neq 0 \) comes from ideas intimately tied to quantum gravity. While there does not yet exist a completely satisfactory quantum theory for gravity, many approaches to the subject predict the existence of some sort of space-time discreteness, in which case the rationale for \( \nabla^b T_{ab} = 0 \) (smoothness at all scales or diffeomorphism invariance) is lost. This opens the door for what could be naively interpreted as an ‘energy diffusion’ resulting from the interaction of the matter sector and the microscopic structure\(^3\) of space-time [23, 24]. For the remainder of the paper, we will avoid further discussion of the possible fundamental source of the violation of \( \nabla^b T_{ab} = 0 \), and we will rather focus on the phenomenology of \( \nabla^b T_{ab} \neq 0 \) in unimodular cosmology, and especially on its possible effects in the context of the \( H_0 \) tension.

As noted in [20], the homogeneous and isotropic Friedman-Lemaître-Robertson-Walker cosmological setting is one where the integrability condition \( dJ = 0 \) is automatically satisfied as a result of the assumed symmetries. We will consider here the spatially flat case corresponding to the metric:

\[
\text{ds}^2 = -dt^2 + a^2(t) \, dx^2,
\]

where the expansion rate of the universe is given by the Hubble rate \( H = a/\dot{a} \), with the dot denoting a derivative with respect to cosmic time \( t \). For simplicity, we choose the normalization of the scale factor \( a(t) \) so that its value today is \( a(t_0) = 1 \). Due to spatial homogeneity, the current \( J_a \) can only have a non-trivial (\( dt \)) component, and can only depend on \( t \), thus we write

\[
J = \dot{\Lambda}(t) \, dt. \tag{6}
\]

The Friedman and Raychaudhuri equations for unimodular gravity follow directly from (4): they are exactly the same as in general relativity, except that \( \Lambda \) is no longer a constant:

\[
H^2 = \frac{8\pi G}{3} \sum_i \rho_i + \frac{\Lambda(t)}{3}, \quad \dot{H} = -4\pi G \sum_i (\rho_i + P_i). \tag{7}
\]

where the index \( i \) denotes the different matter fields in the space-time. It is often convenient to rewrite the Friedman equation as

\[
1 = \Omega_i + \Omega_\Lambda, \quad \Omega_i = 8\pi G \rho_i / 3H^2 \quad \text{and} \quad \Omega_\Lambda = \Lambda / 3H^2.
\]

The growth in \( \Lambda \) resulting from the energy diffusion in the matter sector is governed by \( \dot{\Lambda} = 8\pi G \nabla^a T_{ab} \). Given that we are mostly interested in late-time (post-CMB) physics, and at late times the contribution from radiation to the Friedman equation is negligible, we will focus our attention on the energy density \( \rho_\Lambda \) of dark matter and baryonic matter. That is, we will make the assumption that the growth in \( \Lambda(t) \) is entirely due to diffusion from \( \rho_\Lambda \) (a physical motivation for this exists within the framework of [23, 24]). As is usual in cosmology, we will work in the approximation in which matter is taken to

\(^3\) In a standard quantum mechanical context one should also expect diffusion to be associated with decoherence, which would lead to an apparent violation of unitarity from the perspective of the effective treatment. This has been shown to be potentially useful for the resolution of the information loss paradox in the context of black hole evaporation [35–38]. A contrasting picture is offered by spontaneous collapse modifications of quantum mechanics where unitarity and energy conservation are broken fundamentally. Such an alternative could be helpful in dealing with certain difficulties faced by the current paradigm of structure formation in cosmology [39, 40] as well as in resolving the black hole information loss problem in a different way [41–45]. Such modifications of the quantum theory exist in versions that can both lead to energy diffusion as well as energy increase [40].
have vanishing pressure, and thus the modified continuity equation for \( \rho_m \) is simply

\[
\dot{\rho}_m + 3H \rho_m = -\frac{\dot{\Lambda}(t)}{8\pi G}.
\] (8)

Expressing everything in terms of the redshift variable \( z \equiv (1 - a)/a \), this can be rewritten as

\[
(1 + z)^3 \frac{d}{dz} \left( \frac{\rho_m(z)}{(1 + z)^3} \right) = -\frac{1}{8\pi G} \frac{d\Lambda(z)}{dz}.
\] (9)

Specific solutions can be found, for example, by considering in detail a particular diffusion process as done in [23, 24], or by proposing a phenomenological model for \( \Lambda(z) \) which can be integrated to find \( \rho_m(z) \), or vice versa. In the present work we will adopt the second strategy.

III. THE \( H_0 \) TENSION IN STANDARD AND MODIFIED MODELS

The present value of the Hubble parameter \( H_0 = H(t_0) \) can be measured by different methods. One of the most direct and accurate ways to do it is by studying the magnitude of type Ia supernovae and their redshift \( z = (1 - a)/a \). Inferring the luminosity distance \( d_L \) of each supernova from its apparent and absolute magnitudes gives a relation \( d_L(z) \). In general, we might write \( d_L(z) \) if we have the value of \( H(z') \) for \( z' \) in the intervening cosmological regime between the light emission at \( z = 0 \) and its detection at \( z = 0 \), namely:

\[
d_L = (1 + z) \int_0^z H(z')^{-1} \, dz'.
\] (10)

If we just want to focus on small values of the redshift \( z \), we could instead consider a Taylor expansion of the scale factor \( a(t) \) as a function of cosmic time and write

\[
a(t) = 1 + H_0 \Delta t - \frac{q_0}{2} H_0^2 (\Delta t)^2 + \frac{j_0}{6} H_0^3 (\Delta t)^3 + \ldots
\] (11)

where \( \Delta t = t - t_0 \), \( q_0 \) is the deceleration and \( j_0 \) is the jerk parameter (and recall that \( a_0 = 1 \)). A short calculation shows that, for small \( z \), the relation between the luminosity distance and the redshift is [47]

\[
d_L = \frac{z}{H_0} \left[ 1 + \frac{1 - q_0}{2} z - \frac{1 - q_0 - 3q_0^2 + j_0}{6} z^2 + \ldots \right].
\] (12)

The best fit to this relation for type Ia supernovae data for redshifts \( z \lesssim 0.15 \) gives \( H_0 = 74.03 \pm 1.42 \, \text{km} \, \text{s}^{-1} \, \text{Mpc}^{-1} \) [4], based on the Cepheid distance ladder; the tip of the red giant branch distance ladder gives a lower value, \( H_0 = 69.8 \pm 2.5 \, \text{km} \, \text{s}^{-1} \, \text{Mpc}^{-1} \) [5]. For observations at small \( z \), the observational constraints on \( j_0 \) and higher order terms are weak, and the determination of \( H_0 \) is relatively insensitive to uncertainties in \( q_0, j_0 \), etc.

Concrete cosmological models correspond to specific values of such parameters. For instance, a spatially flat GR-\( \Lambda \)CDM cosmology with \( \Omega_\Lambda^0 = 0.3 \) and \( \Omega_b^0 = 0.7 \), corresponds to \( q_0 = -0.55 \) and \( j_0 = 1 \) (see Sec. V); we put bars on these parameters to denote that they correspond to the ones extracted from the data using the GR-\( \Lambda \)CDM flat model, see below.

Given a particular cosmological model, it is also possible to infer the value of \( H_0 \) via observations of the CMB. For instance, assuming a spatially flat GR-\( \Lambda \)CDM cosmology, comparing the CMB data to the theoretical predictions calculated by a Boltzmann code gives \( H_0 = 67.4 \pm 0.5 \, \text{km} \, \text{s}^{-1} \, \text{Mpc}^{-1} \) [2]. There is a \( \sim 10\% \) difference compared to the supernovae data based on the GR-\( \Lambda \)CDM flat model, see below.

We can gain some insight into how \( H_0 \) can be inferred from the CMB without having to rely on a full Boltzmann code in which many other effects are considered simultaneously, by considering a simpler (and yet nonetheless accurate) calculation. The angular location of the acoustic peaks in the CMB has been measured extremely accurately, and the extracted angular scale \( \theta \) provides an excellent approximation to the ratio of the radius of the sound horizon \( r_s \) (the distance sound waves can travel from the time of reheating to recombination, which is entirely determined by pre-CMB physics) to the radius of the surface of last scattering \( R_{\text{LS}} \) (entirely determined by post-CMB physics),

\[
\theta \approx \frac{r_s}{R_{\text{LS}}},
\] (13)

with \( R_{\text{LS}} \) given by

\[
R_{\text{LS}} = \int_{z_0}^{\infty} \frac{dz}{H(z)} = \int_{z_0}^{\infty} \frac{dz}{H(z)} = \int_{z_0}^{\infty} \frac{dz}{H(z)},
\] (14)

using \( dt = (aH)^{-1} \, da = -a \, dz \). A given cosmological model determines \( H(z) \) via the Friedman equation, and that can be used in the expression above to evaluate \( R_{\text{LS}} \).

For the GR-\( \Lambda \)CDM model,

\[
H(z)^2 = \pi_0 \left[ \pi_0 (1 + z)^4 + \pi_0 (1 + z)^3 + \pi_0 \right],
\] (15)

where \( \pi_0, \pi_0^{\text{m}} \) are the value of the Hubble parameter and the contributions to critical density today in the GR-\( \Lambda \)CDM model. Although both \( \pi_0^{\text{m}} \) and \( \pi_0^{\text{m}} \) are relevant in the previous equation because radiation is an important contribution close to the CMB time, the calculations that follow will evaluate the Friedman equation at \( z = 0 \) where the contribution from radiation can be neglected. For that reason we only need the explicit value of \( \pi_0^{\text{m}} \) [2],

\[
\pi_0^{\text{m}} = 0.315 \pm 0.007,
\] (16)
resolution of the $H_0$ tension will require a clear understanding of the fundamental physics behind the diffusion process, together with a solid modelling of the cosmological and astrophysical conditions that produce it (possible steps aiming at such a scenario are proposed in [25]).

We will consider models where the cosmological constant undergoes a sudden increase approximated by a step function, as well as a model where the matter density decreases—due to energy transfer into the dark energy sector—at a faster rate than in GR-ΛCDM.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{This figure shows the region of the parameter space for $1 < z^\ast < 5$ where the $H_0$ tension is alleviated or resolved for the sudden jump model with $\Lambda_\infty = 2.055$. Here $\alpha$ is plotted using a linear scale. The lines denote the $1\sigma$ and $2\sigma$ contours (assuming Gaussian uncertainty), and the dashed line indicates where the value of $H_0$ according to the sudden jump model exactly agrees with $H_0^\ast$. The predicted value of $\Omega_m^\ast$ according to the sudden jump model is shown by the colour (in the grey region, $\Omega_m^\ast < 0.23$). Near $z^\ast \sim 1.5$, in the best fit region $\Omega_m^\ast \sim 0.25$; while near $z^\ast \sim 5$, in the best fit region $\Omega_m^\ast \sim 0.26$.}
\end{figure}

\section*{A. Sudden transfer model}

A particularly simple model that captures the potential of the general framework to solve the $H_0$ tension is a sudden transfer of energy from the matter density to $\Lambda$ at some instant $z^\ast$. Specifically,

$$\rho_m(z) = \rho_m^0 (1 + z)^3 \left[ \theta_+(z - z^\ast) + \alpha \theta_-(z - z^\ast) \right],$$

with $0 < \alpha < 1$, and $\theta_+(x)$ is the Heaviside function (equal to unity for $x \geq 0$ and zero otherwise), and $\theta_-(x) = 1 - \theta_+(x)$. Here $\rho_m^0$ is the energy density of matter today as inferred from the CMB data for a GR-ΛCDM cosmology. Note that due to the form of $\rho_m(z)$, this sudden jump model predicts the same value of $\rho_m$ at the surface of last scattering as in GR-ΛCDM.

In this section we consider two simple models that, for a suitable choice of parameters, can resolve the $H_0$ tension. These models are simple illustrations of the potential interest of the approach presented here. A convincing
Figure 2. This figure shows the region of the parameter space for $5 < z^* < 10$ where the $H_0$ tension is alleviated or resolved for the sudden jump model with $\Lambda_\infty = 2.055$. Here $\alpha$ is plotted using a linear scale. The lines denote the 1$\sigma$ and 2$\sigma$ contours (assuming Gaussian uncertainty), and the dashed line indicates where the value of $H_0$ according to the sudden jump model exactly agrees with $H_0^{\text{sn}}$. The predicted value of $\Omega_m^0$ according to the sudden jump model is shown by the colour (in the grey region, $\Omega_m^0 < 0.23$). In the best fit region $\Omega_m^0 \sim 0.26$.

Figure 4. This figure shows the region of parameter space where the $H_0$ tension is alleviated or resolved for the sudden jump model with $\alpha = 0.999$. Here $\Lambda_\infty$ is plotted using a linear scale. The lines denote the 1$\sigma$ and 2$\sigma$ contours (assuming Gaussian uncertainty), and the dashed line indicates where the value of $H_0$ according to the sudden jump model exactly agrees with $H_0^{\text{sn}}$. The predicted value of $\Omega_m^0$ according to the sudden jump model is shown by the colour (in the grey region, $\Omega_m^0 < 0.23$ or $\Omega_m^0 > 0.315$). Note that for $\alpha = 0.999$ a negative value of $\Lambda_\infty$ is preferred for $z^* \gtrsim 30$. In the best fit region $\Omega_m^0 \sim 0.26$.

Solving the modified continuity equation (9) gives

$$\dot{\Lambda}(z) \equiv \frac{\Lambda(z)}{H_0^2} = \dot{\Lambda}_\infty + \Delta \theta(z - z^*),$$

(24)
with

$$\Delta = 3(1 - \alpha)(1 + z^*)^3 \frac{\Pi^0_m}{\Pi^0_m}.$$  

(25)

The standard GR-$\Lambda$CDM cosmology is recovered for $\alpha = 1$, and in this limit $\Lambda_\infty = 2.055$ gives $\Omega_\Lambda^0 = 0.685$.

Using equations (23) and (24), $H_0$ can be computed from (21). Due to the sudden drop in the matter energy density at $z^*$, the contribution of matter to the Friedman equation today as predicted by this model will be smaller than the value predicted by GR-$\Lambda$CDM, which implies that $\alpha < 1$ for the ratio

$$\frac{\rho_m^0}{\rho_\Lambda^0} = \frac{\Omega_m^0 H_0^2}{\Pi^0_m H_0^2} < 1.$$  

(26)

This model represents a simple idealization of a rapid diffusion process (instantaneous in terms of the scale of the universe but that could be extended in astrophysical time depending on the value of $z^*$). Note that the value of $\Lambda$ for $z > z^*$ could be positive or negative, depending on the value of $\Lambda_\infty$. In the results plotted in Figs. 1–3, we consider a positive $\Lambda_\infty$, while in Fig. 4 a negative $\Lambda_\infty$ is allowed and is in fact preferred for larger $z^*$. Note that in the case of a negative $\Lambda_\infty$, as the cosmological constant was initially negative, there would occur

To simplify calculations and comparisons with the GR-$\Lambda$CDM model, it is convenient to work in terms of $\rho_m(z)$ defined in (18),

$$\dot{\rho}_m(z) \equiv \frac{G \rho_m(z)}{H_0} = \frac{3\Omega_m^0}{8\pi}(1 + z)^3 \left[ \theta_+ (z - z^*) + \alpha \theta_- (z - z^*) \right].$$  

(23)
a dynamical transition from an early anti-de Sitter era to the current (late-time) nearly de Sitter space-time, as proposed in [48].

To find which parameters for the sudden transfer model resolve the $H_0$ tension without requiring any modifications to pre-CMB physics, we demand that the inferred result for $H_0$—computed from (21)—be compatible with $H_0^\text{SN}$ (the value directly measured from local supernovae data). The results are shown in Figs. 1–4. In Figs. 1–3, the best fit region for $\Lambda_\infty = 2.055$ is shown (for different values of $z^*$), while in Fig. 4, the best fit region for $\alpha = 0.999$ is shown. In Figs. 2–4, the best fit region corresponding to a predicted value of $H_0$ in agreement with supernova observations implies $\Omega_m^0 \sim 0.26$ for the sudden jump model, and a $\sim 1\sigma$ departure occurs for $\Omega_m^0 \sim 0.27$. However, this is no longer the case if smaller values of $z^*$ of order 1 are considered in which case lower values of $\Omega_m^0$ are preferred as can be seen in Fig. 1; for example, for $z^* \sim 1$ the best fit region has $\Omega_m^0 \sim 0.24$ and for $z^* \sim 1.5$ the best fit region has $\Omega_m^0 \sim 0.25$. Also, note that for fixed $\Lambda_\infty$, as $z^*$ becomes larger $\alpha$ becomes closer and closer to 1 in the region where the $H_0$ tension is resolved.

B. Anomalous decay of the matter density

For the second phenomenological model, we assume that

$$\rho_m(z) = \rho_m^0(z) \left[ \theta_+(z-z^*) + \left( \frac{1 + z}{1 + z^*} \right)^\gamma \theta_-(z-z^*) \right],$$

(27)

with $\rho_m^0(z) = \rho_m^0 (1+z)^3$. In this model, the matter density in the past of $z^*$ behaves like normal dust (with an initial value on the surface of last scattering matching the predicted value in the GR-$\Lambda$CDM model), but from $z^*$ on (to smaller values of $z$) diffusion decreases the matter density anomalously as parametrized by $\gamma$. The energy lost is captured by dark energy according to equation (9).

In terms of the dimensionless density (18),

$$\rho_m(z) = \frac{G \rho_m(z)}{H_0^2} = \frac{3 \Omega_m^0 (1+z)^3}{8\pi} \left[ \theta_+(z-z^*) + \left( \frac{1 + z}{1 + z^*} \right)^\gamma \theta_-(z-z^*) \right],$$

(28)

where, because this model is based on a deformation of the standard $\Lambda$CDM model, the $\Lambda$CDM matter contribution to the critical density parameter $\Omega_m^0$ in (16) naturally appears. As matter is diffusing into the dark energy sector,

$$\alpha \equiv \frac{\rho_m^0}{\rho_m} = \frac{\Omega_m^0 H_0^2}{\Omega_m^0 H_0^2} = (1 + z^*)^{-\gamma} < 1,$$

(29)

where the ratio $\alpha$ is expressed as a function of the parameters of the new model in the last equality.

Figure 5. This figure shows the region of the parameter space for small $z^*$ where the $H_0$ tension is alleviated or resolved for the anomalous decay model with $\Lambda_\infty = 2.055$. Here $\gamma$ is plotted using a linear scale. The lines denote the $1\sigma$ and $2\sigma$ contours (assuming Gaussian uncertainty), and the dashed line indicates where the value of $H_0$ according to the sudden jump model exactly agrees with $H_0^\text{SN}$. The predicted value of $\Omega_m^0$ according to the sudden jump model is shown by the colour (in the grey region, $\Omega_m^0 < 0.23$). In the best fit region $\Omega_m^0 \sim 0.26$. The solution to the continuity equation then implies that, in terms of the dimensionless quantity (19),

$$\dot{\lambda}(z) = \frac{\Lambda(z)}{H_0^2} = \dot{\Lambda}_\infty - \frac{3\gamma \Omega_m^0}{(\gamma + 3)} \left[ \left( \frac{z + 1}{z^* + 1} \right)^\gamma (z + 1)^3 - (z^* + 1)^3 \right],$$

(30)

for $z \leq z^*$ and

$$\dot{\Lambda}(z) = \dot{\Lambda}_\infty,$$

(31)

for $z \geq z^*$. As before, $\dot{\Lambda}_\infty$ is a constant of integration and a free parameter in the model. The standard GR-$\Lambda$CDM cosmology is recovered for $\gamma = 0$, and in this limit $\Lambda_\infty = 2.055$ gives $\Omega_m^0 = 0.685$.

As for the first model, to find the parameters for which the anomalous decay model can resolve the $H_0$ tension without requiring any modifications to pre-CMB physics, we calculate the inferred value for $H_0$ using (21) and require that it be compatible with $H_0^\text{SN}$; the results are shown in Figs. 5–7. In Figs. 5 and 6, the best fit region for $\Lambda_\infty = 2.055$ is shown (for small and large values of $z^*$ respectively), while in Fig. 7, the best fit region for $\gamma = 10^{-3}$ is shown. Note that for a fixed choice of $\dot{\Lambda}_\infty$, as $z^*$ increases a smaller value of $\gamma$ preferred. In all three figures, the best fit region corresponding to a predicted value of $H_0$ in agreement with supernova observations implies $\Omega_m^0 \sim 0.26$ and a $\sim 1\sigma$ departure occurs for $\Omega_m^0 \sim 0.27$; this is similar to what was found.
for the sudden jump model. Again, this is no longer the case for smaller \( z^* \sim 1 \) in which case the best fit region has a smaller \( \Omega_m^0 \).

Note that it is not surprising that (for \( z^* \gtrsim 5 \)) the best fit regions for \( H_0 \) in both of the two models considered here give \( \Omega_m^0 \sim 0.26 \). This is because in both cases \( \alpha \approx 1 \) for the best fit region (see Fig. 3 for the sudden transfer model and Fig. 6 together with Eq. (29) for the anomalous decay model) and therefore \( \Omega_m^0 \approx (\Pi^2_0/H_0^2) \Omega_m^0 \approx 0.26 \) for \( H_0 \) close to the value preferred by supernovae observations. Note that for smaller values of \( z^* \), the best fit region will have a smaller value for \( \alpha \) (see for example Fig. 1), and therefore a smaller predicted \( \Omega_m^0 \) as well.

**Figure 6.** This is the same diagram as in Figure 5, but now considering larger values of \( z^* \) in the anomalous decay model and using a logarithmic scale for \( \gamma \). As in Figure 5, \( \Lambda_{\infty} = 2.055 \). As for smaller values of \( z^* \), in the best fit region \( \Omega_m^0 \sim 0.26 \).

**V. OBSERVATIONS AT INTERMEDIATE REDSHIFTs**

It may be possible to have a ‘direct’ detection of a growing \( \Lambda \) from observations at intermediate redshifts \( 1 < z < 10 \).

From the Friedman and Raychaudhuri equations for unimodular gravity with a varying \( \Lambda \) (7), and considering that pressureless matter together with \( \Lambda \) dominate the dynamics (i.e., \( \rho = \rho_m \) and \( p = 0 \)), the deceleration and jerk parameters defined in (11) would be given by \(^4\)

\[
q_0 = -1 + \frac{3}{2} \Omega_m^0, \quad j_0 = 1 + \frac{\Lambda(0)}{2H_0^2}. \tag{32}
\]

\(^4\) For \( z > 1 \), it is better to do a Taylor expansion for observables like \( d_L \) around \( y = z/(1 + z) \) rather than \( z \) [49]; but in any case the prefactors to \( y^n \) in this Taylor series depend only on \( q_0, j_0 \), etc.

It is clear that, in the limit where \( \Lambda \) is constant, the standard GR-CDM cosmological model is recovered with the standard prediction \( j_0 = 1 \). On the other hand, in the context of unimodular gravity, with energy being ‘transferred’ from matter fields to the dark energy sector, we would have \( \dot{\Lambda} > 0 \), and thus \( j_0 > 1 \). We note that the effect of dissipation in unimodular gravity on the jerk parameter has also been studied in [50].

Type Ia supernovae have been observed up to redshifts \( z \approx 2.3 \), and these observations strongly disfavour models with \( q_0 \approx -0.55 \) but all other terms in the expansion (including \( j_0 \)) vanishing [51]. Future observations may serve to solidly constrain the value of \( j_0 \).

Also, the papers [52, 53] report a measurement of the deceleration parameter using local observations (supernovae of type Ia with redshifts \( 0.023 \leq z \leq 0.15 \)) with

\[
q_0 = -1.08 \pm 0.29, \tag{33}
\]

which is 1.9\( \sigma \) away from the \( \Lambda \)CDM value \( q_0 = -0.55 \). Taken at face value, this result would suggest \( \Omega_m^0 < 0.14 \) (see also [54]). Even though these results must be taken with caution, they illustrate the point that there is, at present, a rather important degree of uncertainty concerning the amount of (dark and baryonic) matter present in the universe today, a fact that seems to leave room for a level of energy diffusion of the magnitude required in the models considered here that can resolve the \( H_0 \) tension.

It is clear that observations at larger \( z \) could poten-
VI. DISCUSSION

We have presented a general scenario that can resolve the $H_0$ tension. The new perspective is motivated by the proposal in [23, 24] for a fundamental account of the nature and value of dark energy. The model does not require problematic equations of state for the components of the cosmic pie. However, it does require violations of the local conservation of energy momentum (i.e., deviations from $\nabla^a T_{ab} = 0$), and thus a modified version of GR that can accommodate these: unimodular gravity. We have provided only a brief motivation for this possibility here, because that has been (and will be) the subject of other works. The present analysis has focused on the essential aspects of the phenomenology. The basic effect can be seen as an effective ‘flow of energy’ from the matter sector to dark energy (or ‘cosmological constant’) sector. The viability of the model depends on its ability to successfully overcome several constraints. It must not disrupt the conditions in the CMB in any substantial manner, avoiding any negative repercussion regarding the success of the GR-ΛCDM model. It must, of course, resolve or substantially reduce the $H_0$ tension, thus accommodating the late time supernovae observations, and it must not lead to an excessive reduction in amounts of the energy density of the matter sector. The last point is essential so as not to be in conflict with late time observations of that sector. In other words, the amount of dark matter that, according to the viable scenarios, was converted into dark energy in the period between the surface of last scattering and the present, cannot be more than a small fraction of its initial value.

Concerning this last point, CMB observations, combined with the assumption of GR-ΛCDM, imply that $\Omega_{\text{Baryons}}^0 \sim 0.05$ and $\Omega_{\text{DM}}^0 \sim 0.26$; however, direct present day observations indicate that most of the implied baryonic matter must be invisible (i.e., is not accounted for in the star gas and dust that forms the galaxies and the intergalactic media), namely $\Omega_{\text{Visible-Baryons}}^0 \sim 0.02$ [58]. On the other hand, the dark matter in galaxies and clusters is known to be about 10 times the amount of luminous matter (which itself is a fraction of the baryonic matter as a substantial part of it seems to be in the intergalactic media). Thus the portion of the dark matter component that we can directly infer from galaxy surveys today is of the order of $\sim 0.20$. That is, we have strong confidence that at present there is at least $\Omega_{\text{Matter}}^0 \sim 0.22$. Taking into account the claims of a substantial presence of baryonic matter in the intergalactic media, we could say that a conservative constraint in this regard ought to be $\Omega_{\text{Matter}}^0 \sim 0.25$. The point is that given the uncertainties in the amount of matter directly observed or indirectly inferred to be present in today’s universe there seems to be room for an important missing fraction compared to the GR-ΛCDM value of $\Omega_{\text{Matter}}^0 \sim 0.31$. Thus we consider that our models are viable in this regard if the fraction of $\Omega_{\text{Matter}}^0$ that was lost via our mechanism in the period from the last scattering surface to the present is smaller than, say, 20%. As can be seen in the figures of the previous sections, the models we have presented pass this test for various ranges of the parameters compatible with the other requirements. Notice that the strong constraints on $\Omega_{\text{Baryons}}^0$ coming from nucleosynthesis are not relevant in the present analysis because all the modifications introduced by our models happen after recombination, and diffusion effects can cause the amplitude of $\rho_{\text{Baryons}}$ to decrease faster than what would be expected from cosmological expansion alone.

The simplicity of the two models considered here and their effectiveness in resolving the $H_0$ tension illustrates the potential of the perspective that we put forward here. A more realistic model would require, in addition, a detailed account of the fundamental mechanism behind the diffusion process as well as the astrophysical description of the sources, their abundance and dynamical evolution during the recent cosmological history. Thus, the framework presented here naturally calls for further studies involving the simultaneous analysis of parametrizations characterizing these two combined aspects of the problem, with best fit studies based on the wide range of available data regarding cosmological observations. One such possibility of more complete scenario is the natural generalization of the fundamental mechanism of [23, 24] that would include black holes as diffusive sources [25] (see Footnote 1). We should note that proposals having some resemblance to the present one, but motivated by rather different physical ideas have been put forward in [50, 61].

More generally, unimodular gravity, combined with the relaxation of the assumption of conservation of the
energy-momentum tensor of matter, represents, at a classical level, a relatively mild modification of GR which ought to be subjected to studies and constraining observational tests, just like other modified gravity theories. The present work together with the results of \cite{23, 24} clearly exhibits its potential.

Finally, note that in this paper we have been working under the assumption that the cosmological parameters at the last scattering surface are those determined by the Planck analysis of their data. That analysis in turn has been carried out under the assumption of the validity of the standard GR-ΛCDM model. Our modified model calls for a reevaluation of those parameters, given the Planck observations, but carried out with the modified cosmic evolution, in particular by using a Boltzmann code updated to include the diffusion effects of either of the models considered here (or perhaps another model entirely). This would naturally lead to small changes in the preferred values of some cosmological parameters, but we expect these resulting changes to be of higher order and lead only to small modifications of the results derived here based on the ratio of the sound horizon and the radius of the surface of last scattering. That more complete and complex analysis will be the subject of a forthcoming paper.

**ACKNOWLEDGMENTS**

We thank Marco de Cesare and Violaine Ponsin for help with the figures, and E.W.-E. thanks Aix-Marseille Université for hospitality during the initial stages of this work. E.W.-E. is supported in part by the Natural Science and Engineering Research Council of Canada and by a Harrison McCain Foundation Young Scholars Award. D.S. is supported in part by the grant number FQXi-MGB-1928 from the Foundational Questions Institute and the Fetzer Franklin Fund, a donor advised by the Silicon Valley Community Foundation.

[1] Planck Collaboration, R. Adam et al., “Planck 2015 results. I. Overview of products and scientific results,” Astron. Astrophys. 594 (2016) A1, arXiv:1502.01582.

[2] Planck Collaboration, N. Aghanim et al., “Planck 2018 results. VI. Cosmological parameters,” arXiv:1807.06209.

[3] BOSS Collaboration, S. Alam et al., “The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample,” Mon. Not. Roy. Astron. Soc. 470 (2017) 2617–2652, arXiv:1607.03155.

[4] A. G. Riess, S. Casertano, W. Yuan, L. M. Macri, and D. Scolnic, “Large Magellanic Cloud Cepheid Standards Provide a 1% Foundation for the Determination of the Hubble Constant and Stronger Evidence for Physics beyond ΛCDM,” Astrophys. J. 876 (2019) 85, arXiv:1903.07603.

[5] W. L. Freedman et al., “The Carnegie-Chicago Hubble Program. VIII. An Independent Determination of the Hubble Constant Based on the Tip of the Red Giant Branch,” arXiv:1907.05922.

[6] W. M. Farr, M. Fishbach, J. Ye, and D. Holz, “A Future Percent-Level Measurement of the Hubble Expansion at Redshift 0.8 With Advanced LIGO,” Astrophys. J. 883 (2019) L42, arXiv:1908.09084.

[7] L. Bergström, “Nonbaryonic dark matter: Observational evidence and detection methods,” Rept. Prog. Phys. 63 (2000) 793, arXiv:hep-ph/0002126.

[8] T. Clifton, P. G. Ferreira, A. Padilla, and C. Skordis, “Modified Gravity and Cosmology,” Phys. Rept. 513 (2012) 1–189, arXiv:1106.2476.

[9] J. Martin, “Everything You Always Wanted To Know About The Cosmological Constant Problem (But Were Afraid To Ask),” Comptes Rendus Physique 13 (2012) 566–665, arXiv:1205.3365.

[10] Supernova Search Team, A. G. Riess et al., “Observational evidence from supernovae for an accelerating universe and a cosmological constant,” Astron. J. 116 (1998) 1009–1038, arXiv:astro-ph/9805201.

[11] Supernova Cosmology Project, S. Perlmutter et al., “Measurements of Omega and Lambda from 42 high redshift supernovae,” Astrophys. J. 517 (1999) 565–586, arXiv:astro-ph/9812133.

[12] J. L. Bernal, L. Verde, and A. G. Riess, “The trouble with H0,” JCAP 1610 (2016) 019, arXiv:1607.05617.

[13] W. L. Freedman, “Cosmology at a Crossroads,” Nat. Astron. 1 (2017) 0121, arXiv:1706.02739.

[14] V. Poulin, T. L. Smith, T. Karwal, and M. Kamionkowski, “Early Dark Energy Can Resolve The Hubble Tension,” Phys. Rev. Lett. 122 (2019) 221301, arXiv:1811.04083.

[15] E. Ó Colgáin, M. H. van Putten and H. Yavartanoo, “de Sitter Swampland, H0 tension & observation,” Phys. Lett. B793 (2019) 126-129, arXiv:1807.07451.

[16] S. Kumar, R. C. Nunes and S. K. Yadav, “Dark sector interaction: a remedy of the tensions between CMB and LSS data,” Eur. Phys. J. C79 (2019) 576, arXiv:1903.04865.

[17] E. Di Valentino, A. Melchiorri, O. Mena, and S. Vagnozzi, “Interacting dark energy after the latest Planck, DES, and H0 measurements: an excellent solution to the H0 and cosmic shear tensions,” arXiv:1908.04281.

[18] P. Agrawal, G. Obied, and C. Vafa, “H0 Tension, Swampland Conjectures and the Epoch of Fading Dark Matter,” arXiv:1906.08261.

[19] M. Rossi, M. Ballardini, M. Braglia, F. Finelli, D. Paoletti, A. A. Starobinsky and C. Umiltà, “Cosmological constraints on post-Newtonian parameters in effectively massless scalar-tensor theories of gravity,” Phys. Rev. D100 (2019) 103524, arXiv:1906.10218.

[20] T. Josset, A. Perez, and D. Sudarsky, “Dark Energy from Violation of Energy Conservation,” Phys. Rev. Lett. 118 (2017) 021102, arXiv:1604.04183.
Galaxies," The Astrophysical Journal 862 (Jul, 2018) 3.

[61] M. A. García-Aspeitia, A. Hernández-Almada, J. Magaña, and V. Motta, “On the birth of the cosmological constant and the reionization era,” arXiv:1912.07500.