Machine learning unveils the linear matter power spectrum of modified gravity

J. Bayron Orjuela-Quintana[1,*] Savvas Nesseris[2,†] and Domenico Sapone[3,‡]

[1]*Departamento de Física, Universidad del Valle, Ciudad Universitaria Meléndez, Santiago de Cali 760032, Colombia*
[2]*Instituto de Física Teórica UAM-CSIC, Universidad Autonóma de Madrid, Cantoblanco, 28049 Madrid, Spain*
[3]*Cosmology and Theoretical Astrophysics group, Departamento de Física, FCFM, Universidad de Chile, Blanco Encalada 2008, Santiago, Chile*

The matter power spectrum $P(k)$ is one of the main quantities connecting observational and theoretical cosmology. Although for a fixed redshift this can be numerically computed very efficiently by Boltzmann solvers, an analytical description is always desirable. However, accurate fitting functions for $P(k)$ are only available for the concordance model. Taking into account that forthcoming surveys will further constrain the parameter space of cosmological models, it is also of interest to have analytical formulations for $P(k)$ when alternative models are considered. Here, we use the genetic algorithms, a machine learning technique, to find a parametric function for $P(k)$ considering several possible effects imprinted by modifications of gravity. Our expression for the $P(k)$ of modified gravity shows a mean accuracy of around 1–2% when compared with numerical data obtained via modified versions of the Boltzmann solver CLASS, and thus it represents a competitive formulation given the target accuracy of forthcoming surveys.

I. INTRODUCTION

Since its designation as the standard cosmological model around two decades ago [1], the model assuming the existence of the cosmological constant Λ and a cold dark matter (CDM) component, ΛCDM, has passed many stringent observational tests [2–16]. However, intriguing discrepancies have arisen in recent years [17,18]. Arguably, the most representative among these discrepancies is the H_0 tension whose significance is now around 5σ [19–21]. Therefore, the search for a sound theoretical candidate is still plausible, highlighting modified gravity (MG) theories as one the most prominent alternatives [22,23].[1]

The clustering of matter, due to gravitational instabilities, forming the structure of our Universe constitutes a key physical process in the expansion history. This process offers one of the main links between observations and theoretical models through the so-called matter power spectrum $P(k, z)$, which in general is a function of the wave number k and the redshift z [32–34]. For a given redshift, $P(k)$ can be numerically computed using Boltzmann solvers, like the codes CAMB [35] or CLASS [36]. Nonetheless, it is desirable to have an analytical formulation of $P(k)$ in some cases; for instance, when computing costs are demanding and high accuracy is not required. When ΛCDM is the assumed cosmological model, there are two well-known fitting formulas to determine the transfer function involved in the computation of $P(k)$. These are the Bardeen-Bond-Kaiser-Szalay (BBKS) formula [37] and the Eisenstein-Hu (EH) formulas [38,39], which are widely used in the literature [40–47]. However, up to the best of our knowledge, there are no similar formulations considering alternative models.

In Ref. [48] two of the authors here used a machine learning technique known as the genetic algorithms (GAs) to find an analytical expression for the matter power spectrum under the ΛCDM framework and neglecting the representative wiggles of $P(k)$ due to the BAOs (baryon acoustic oscillations). This new formula shows to be as accurate as its predecessors while being considerably simpler; and thus, easier to implement in other numerical computations where the knowledge of the dewiggled $P(k)$ is required. Here, we want to extend that work by considering some possible effects on $P(k)$ introduced by modifications of gravity. Briefly, using the GAs, we find a parametric function for the dewiggled $P_{\text{MG}}(k)$ whose performance is tested by comparing against numerical data

---

[*]john.orjuela@correounivalle.edu.co
[†]savvas.nesseris@csic.es
[‡]domenico.sapone@uchile.cl

[1]We want to stress that modified gravity theories and dark energy models can be treated on equal footing under some approaches [24–31], and thus a clear distinction between them is not necessary.

obtained from the popular Boltzmann solver CLASS. In particular, we use the branches `hi_class` [49,50] and `mgclass` [51,52]. Both allow to numerically compute $P(k)$ for MG theories. The former considers the whole Horndeski theories, while the later uses the slip parameters framework.

The paper is organized as follows. In the next section we give some generalities about the GAs. Then, in Sec. III, we describe the matter power spectrum $P(k)$ in general and the possible effects due to modifications of gravity we take into account. In Sec. IV, we present the parametric function for the $P(k)$ of MG theories, discuss its applicability and its accuracy when compared against numerical data. Finally, in Sec. V, we give our concluding remarks.

II. GENETIC ALGORITHMS

The GAs are an unsupervised machine learning technique in which the extraction of features and patterns from a given dataset is performed by mimicking biological evolution processes [53]. In a nutshell, a random initial population to fit the data is created. Then, GAs improve their description of the data by implementing a natural selection scheme based on reproduction and survival of the fittest individuals. This process, that is stochastic in essence, is carried out for as many generations as desired or needed to achieve the expected accuracy of the fitting.

The working principle of the GAs makes them an excellent tool as a symbolic regression approach. In a few words, symbolic regression consists in discovering a fitting function that accurately matches a given dataset. Now we succinctly describe how to GAs can be used in symbolic regression problems. Firstly, the initial individuals, corresponding to random mathematical expressions, are created. Then, the fitness of each individual in this first generation is measured, which can be done by using a simple mean-squared error, for example. A portion of the fittest individuals are selected from a tournament scheme. The winners of this tournament survive to the next generation to create offspring through the so-called genetic operators, usually corresponding to crossover between individuals and mutation. In the subsequent generations, consisting of the most capable individuals from the previous generation and their offspring, the evolutionary process is repeated until an individual achieves the desired accuracy for data description. See Refs. [53,54] for more details about GAs.

Finally, we point out that GAs have been used in several branches of physics, such as particle physics [55–57], astrophysics [58], and astronomy [59]. Particularly in cosmology, they have been used as a symbolic regression tool [48,60], but also used to perform null tests of cosmological datasets [61,62], constraints on cosmological quantities [63–67], and forecasts for upcoming surveys [68–70]. There are also other approaches to symbolic regression based on different methodologies as described in e.g., Refs. [71–76].

III. THE MATTER POWER SPECTRUM OF MODIFIED GRAVITY

As mentioned in Sec. I, the existence of structure in our Universe follows from the agglomeration of matter due to gravitational instabilities. One of the key statistical correlation functions of this process is the matter power spectrum $P(k, z)$. It can be shown that, when nonlinear effects are neglected, the matter power spectrum for a given redshift can be written as [34]

$$P(k) \equiv P_0(k)T^2(k), \tag{1}$$

where $P_0(k)$ is an initial matter power spectrum and $T(k)$ is the matter transfer function. Now, we firstly comment on the form of $P(k)$ for ΛCDM, and then we discuss some of the possible alterations to this observable when modifications of gravity are considered.

In the standard case, the initial matter power spectrum is approximately given by

$$P_0(k) \equiv A_0 k^{n_s}, \qquad A_0 \equiv \frac{8\pi^2}{25}\frac{\mathcal{A}_s}{\Omega_{m0}^2 H_0^4}k_p^{1-n_s}, \tag{2}$$

where Ω_{m0} is the matter density parameter today, H_0 is the Hubble constant, n_s and $\mathcal{A}_s$ are the spectral tilt and amplitude of primordial scalar perturbations, respectively, and k_p is an arbitrary pivot scale [34].[2] On the other hand, the matter transfer function takes into account the effects of all the interacting species, e.g., acoustic oscillations due to baryonic pressure, amplitude suppression due to the free streaming of massive neutrinos, transition from a radiation-dominated epoch to a matter-dominated epoch, and more. All these physical details are encompassed by the fitting functions presented in Refs. [38,39], which achieve an accuracy of 1–2% while needing around 30 different expressions in their full formulations. However, in Ref. [48], two of the authors here shows that the GAs are able to find much more economical expressions while achieving the same accuracy range.

To the best of our knowledge, there are currently no similar analytical formulations available when considering alternative models. To find derive such a formula, our first step is to characterize the potential effects that MG theories introduce in the matter power spectrum compared to the concordance model. In the following, we undertake this characterization process. Firstly, a background history different to that of ΛCDM can change the growth factor, which in turn modifies the overall amplitude of $P(k)$.

---

[2]Note that $P_0(k)$ is not the primordial power spectrum which is given by $P_{\mathcal{R}}(k) \equiv \mathcal{A}_s(k/k_p)^{1-n_s}$ [14].
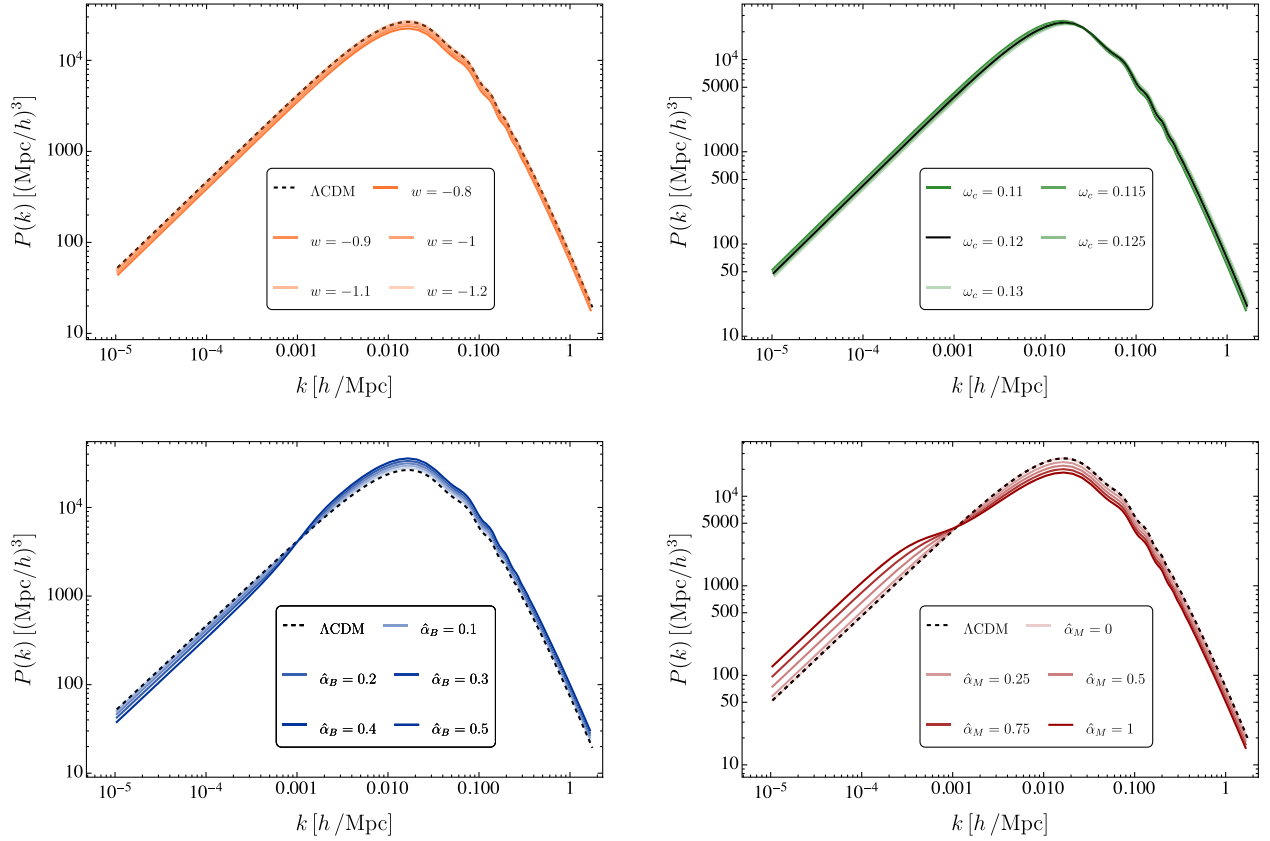
FIG. 1.   Typical effects of modified gravity on the theoretical matter power spectrum. In the upper panels, we consider a wCDM model varying the dark energy equation of state w (left), and the reduced CDM density parameter ω_c (right). In the lower panels, we consider the `proto_scale` models in `hi_class`, varying the braiding $\hat{\alpha}_B$ (left) and the mass running $\hat{\alpha}_M$ (right) parameters. In the former case, the matter power spectrum is suppressed at very large scales while in the later it is enhanced. Then, we see a transition between the large scale spectrum to an enhanced (left) or suppressed (right) spectrum at smaller scales.

Secondly, a variation in the reduced matter density parameter ω_m implies a variation in the redshift and scale at the radiation-matter transition, denoted respectively as $z_{\rm eq}$ and $k_{\rm eq}$. This causes a shift in the position of the maximum of the matter power spectrum. Before continuing, we want to mention that it is possible to estimate $z_{\rm eq}$ and $k_{\rm eq}$ using the approximated expressions given in Ref. [38]. Nevertheless, as shown in the Appendix, the GAs are able to find expressions with an accuracy improved by two orders of magnitude while keeping simplicity. As a third effect, modified gravity can induce a clustering effective dark energy component when the structure formation process is occurring. This of course modifies the amplitude of $P(k)$, [77–79]. Modifications can occur at the largest scales due to the integrated Sachs-Wolfe effect. Finally, there could be a variation in the amplitude and location of the acoustic peaks. However, these effects are principally controlled by the reduced baryon density parameter ω_b which is tightly constrained by observations [2]. Here, we fix it to its default value in the popular Boltzmann solver CLASS; i.e., $\omega_b = 0.0223828$, and neglect the small oscillations in our formulation.

In Fig. 1 we have plotted the matter power spectra for four different models varying some of their parameters in order to exemplify the physical effects introduced by modifications of gravity, as explained above. These plots were obtained using the branch `hi_class` of CLASS. In the upper-left panel, we consider a wCDM model, with a constant speed of sound c_s^2, no anisotropic stress, and the equation of state of dark energy w varying in the interval $[-1.2, -0.8]$. We can see that the overall amplitude of $P(k)$ is modified when w varies. In the upper-right panel, we consider a ΛCDM model, varying the reduced CDM density parameter ω_c in the range $[0.11, 0.13]$. We note small shifts in scale of the maximum of $P(k)$. For the lower panels, we used one of the implemented models by default in `hi_class`; namely, the `propto_scale` model. In this model, the effective field theory (EFT) parameters characterizing the cosmological dynamics of the whole Horndeski theory are taken as proportional to the scale factor. For more details about the EFT of Horndeski theories see Refs. [80,81], and for the code `hi_class` see Refs. [49,50]. In particular, for these plots we fixed the EFT parameters and vary only the proportionality constant

of the braiding parameter $\hat{\alpha}_B$ (lower-left panel), and of the mass running parameter $\hat{\alpha}_M$ (lower-right panel), while the background evolution matches that of ΛCDM. As it can be seen, there is a clear distinction between the behavior of $P(k)$ at large scales and at small scales, with a marked transition occurring before the scale at equality. The matter power spectrum at the largest scales holds the form of the initial matter power spectrum proportional to k^{n_s} while differences can be attributed to variations in the amplitude A. Then, at smaller scales, the spectrum can be enhanced (as in the lower-left panel) or suppressed (as in the lower-right panel) by modified gravity. In all the cases, we have depicted the matter power spectrum of ΛCDM to make notable the differences introduced by MG models.

In the next section, we propose a parametrized function for the matter power spectrum $P(k)$ that incorporates the aforementioned effects.

IV. RESULTS

A. Parametrization of $P(k)$

Summarizing our discussion in the last section, we consider the following modified gravity effects on the $P(k)$:

(1) Variations in the amplitude A of the initial matter power spectrum $P_0(k)$ at large scales.
(2) Overall enhancement or suppression of the amplitude.
(3) Shift of the equality scale $k_{\rm eq}$.

In order to encompass these three effects in one single expression, we assume that the total matter power spectrum considers the contribution of two different spectra. At large scales, the most relevant contribution to $P(k)$ comes from the following expression

$$P_{\rm large}(k; s, n_s) \equiv A(s) \left(\frac{k}{h/\text{Mpc}}\right)^{n_s} \left[(\text{Mpc}/h)^3\right], \quad (3)$$

where

$$A(s) \equiv (1 + s) \frac{P_{\Lambda\text{CDM}}(k_i)}{k_i^{n_s} T_{\Lambda\text{CDM}}^2(k_i)}, \quad (4)$$

is a dimensional constant. The matter power spectrum $P_{\Lambda\text{CDM}}$ and the matter transfer function $T_{\Lambda\text{CDM}}$ for ΛCDM can be retrieved from e.g. CLASS, and thus k_i corresponds to the smallest value of k considered by the Boltzmann solver. Notice that this parameter is basically a modification of A_0 defined in Eq. (2). However, we do not use A_0 directly in the definition of $A(s)$, i.e., we do not define this parameter as $A(s) \equiv A_0(1 + s)$, since the expression for A_0 is an approximation only valid for subhorizon modes [34]. The parameter s takes on negatives values when the matter power spectrum is suppressed (as shown in the lower-left panel of Fig. 1), and positive values when the primordial matter power spectrum is enhanced (as depicted in the lower-right panel of Fig. 1) at large scales. The additional effects, such as an overall modification of the amplitude and the shift of the maximum, are incorporated into the matter power spectrum (referred to as $P_{\rm small}$), which primarily manifest at smaller scales. In the upper-right panel of Fig. 1, it is evident that a change in the parameter w in the wCDM model can result in an overall enhancement or suppression of the amplitude of the matter power spectrum. Hence, we adopt this model as a reference to develop a formula for $P_{\rm small}$ utilizing the GAs. The accuracy of such a formula is measured by the following fitness function:

$$\text{Acc}\% \equiv \left\langle \frac{|P_{\text{CLASS}} - P_{\text{analytical}}|}{P_{\text{CLASS}}} \right\rangle \times 100, \quad (5)$$

where P_{CLASS} is the matter power spectrum numerically computed using CLASS, and $P_{\text{analytical}}$ is the expression obtained from the GAs. In particular, we consider a constant sound speed given by $c_s^2 = 0.001$, fix the spectral tilt n_s to the default value in CLASS, and vary w in the range $[-1.2, -0.8]$. The GAs runs give the following expression[3]:

$$P_{\rm small}(k; \alpha, \omega_m) = \frac{2.52395 \times 10^6 \left(\frac{k}{h/\text{Mpc}}\right)^{0.96452} (1 + 0.710636\alpha^3 + 1.88019\alpha^4 + 0.939217\alpha^5)}{(1 + 59.0998x^{1.49177} + 4658.01x^{4.02755} + 3170.79x^{6.06} + 150.089x^{7.28478})^{1/2}} \left[(\text{Mpc}/h)^3\right], \quad (6)$$

whose accuracy is 1.72%. In the last equation, we have defined the dimensionless variable $x \equiv k/(\omega_m - \omega_b)$ Mpc/h, since k is given in units $[h/\text{Mpc}]$, and we promoted w to a general parameter α in order to avoid possible confusions with the notation. The parameter α, in general, measures the overall modifications in the amplitude of $P(k)$ and it is not restricted to assume values in the range $[-1.2, -0.8]$. However, this parameter has the lower bound $\alpha > -1.68492$ since $P_{\rm small}$ is positive for all k. It is important to emphasize that the wCDM model is used purely as a template in our study, as the variation of w specifically generates the effect we aim to capture, which is

an overall modification of the amplitude of the matter power spectrum.

Finally, we assume that the spectra in Eqs. (3) and (6) are smoothly connected by a transition function given by

$$\sigma(k; k_0, \beta) = [1 + e^{-(\ln k - \ln k_0)/\beta}]^{-1}, \quad (7)$$

---

[3]For details about the specific structure of the GAs, e.g., the genome arrangement and avoidance of overfitting problems, please refer to Ref. [48].

TABLE I.   Best-fit parameters α, k_0, β, and s for several models in the codes `hi_class` and `mgclass`. Just one value of the parameters of each model is shown to keep our presentation simple. We can see that the mean accuracy in each example is around 1–2%.

Model	Parameter	α	$k_0 \times 10^{-3}$	β	s	Accuracy (%)
`propto_scale`	$\hat{\alpha}_B = 0.5$	0.68153	1.16918	0.42243	−0.21975	1.53
`propto_scale`	$\hat{\alpha}_M = 0.5$	0.36380	0.73084	0.45974	0.53958	1.62
`propto_scale`	$\hat{\alpha}_T = -1$	0.53444	0.40205	0.39972	−0.47066	1.52
`propto_omega`	$\tilde{\alpha}_B = 0.625$	0.54866	0.76693	0.45681	−0.07756	1.46
`propto_omega`	$\tilde{\alpha}_M = 3.0$	0.59006	0.39858	0.46187	1.57924	1.48
`propto_omega`	$\tilde{\alpha}_T = -0.25$	0.53215	1.16323	0.42635	0.02292	1.47
`plk_late`	$E_{11} = 0.5, E_{22} = 0.5$	0.54762	0.76681	0.49599	−0.62532	1.43
	$E_{11} = -0.5, E_{22} = -0.5$	0.42644	0.28607	0.51071	2.99343	1.92
HDES	$\tilde{J}_c = 0.005$	0.56331	2.24347	0.443931	0.05417	1.39

where the transition is centered at k_0 and β determines its width.[4] It is noteworthy that the limits of this function are $\sigma \to 0$ when $k \ll k_0$, and $\sigma \to 1$ when $k \gg k_0$, and that we have defined σ using logarithms since usually $P(k)$ is depicted in log space. The transition between these limits occurs around $k \sim k_0$ with a smoothing level determined by β. Consequently, a small β corresponds to a rapid transition, while a large β results in a slower transition. Combining these considerations, the parametrized expression for $P(k)$ in the context of modified gravity is given by

$$P_{\mathrm{MG}}(k; \alpha, \beta, k_0, s, \omega_m, n_s) \equiv [1 - \sigma(k; k_0, \beta)]P_{\mathrm{large}}(k; s, n_s)$$
$$+ \sigma(k; k_0, \beta)P_{\mathrm{small}}(k; \alpha, \omega_m).$$
(8)

The new expression requires six parameters as input; the spectral tilt n_s, the shift s at large scales, the overall enhancement (or suppression) parameter α, the reduced matter density parameter ω_m, the transition center k_0, and the transition width β switching between P_{large} and P_{small}. With the exception of n_s and ω_m, which are determined through observations, the remaining four parameters are free. We want to emphasize that the expression obtained through GAs was P_{small} in Eq. (6), while P_{large} in Eq. (3) and σ in Eq. (7) are assumed and motivated by physical intuitions.

In the following section, we will evaluate the effectiveness of Eq. (8) by estimating the best-fit values of parameters to accurately reproduce the numerical matter power spectra obtained from various MG theories using the `hi_class` and `mgclass` codes.

B. Performance

Here, we numerically compute matter power spectra using the codes `hi_class` and `mgclass` considering several models. Subsequently, we find the best-fit values of the

parameters α, k_0, β, and s by minimizing the fitness function in Eq. (5). This procedure will yield to the best match of $P_{\mathrm{MG}}(k)$ in Eq. (8) with respect to the data from CLASS.

There are several models already implemented in `hi_class`, and we make use of two specific ones; `proto_scale` and `propto_omega`. In the `propto_scale` model, the EFT parameters of the Horndeski theory are proportional to the scale factor a, i.e. to $\alpha_i \equiv \hat{\alpha}_i a$, where $i = K$, B, M, T denotes the kineticity, braiding, mass running, and tensor speed excess parameters, respectively. On the other hand, in the `propto_omega` model, the EFT parameters are proportional to the dark energy density parameter Ω_{DE}, i.e., $\alpha_i \equiv \tilde{\alpha}_i \Omega_{\mathrm{DE}}$. We also add to `hi_class` the designer Horndeski model (HDES) introduced in Ref. [25]. The HDES model has two parameters; namely, $\tilde{J}_c$ and n. We fix $n = 1$, and consider a small value for $\tilde{J}_c$. In all these models, it is assumed that the background evolution exactly matches that of ΛCDM. In `mgclass`, we also use an already implemented model called `plk_late`, in which the slip parameters are parametrized in terms of Ω_{DE} (see Ref. [52]). This model in particular has been used by observational collaborations such as Planck [82] and DES [4].

In Table I, we show the best-fit values of the parameter set $\{\alpha, k_0, \beta, s\}$ along with the associated mean accuracy, achieved by selecting a specific parameter value from each model. It is important to mention that we have explored multiple parameter values for each model. However, for the sake of simplicity in the presentation, we report only one value. The details of the analysis and the parameter exploration for each model can be found in the provided codes (see the Numerical Codes statement). Table I shows that the mean accuracy for the considered models is approximately 1–2%.[5] In Fig. 2, we plot three examples

---

[4]This transition function is not unique. Another common option for a "smooth step function" which would be valid is $\sigma(k; k_0, \beta) \equiv (1 + \tanh[(\ln k - \ln k_0)/\beta])/2$.

[5]We want to mention that since we fix n_s to find P_{small} using the GAs, this could yield to errors when different values of n_s are considered. Although, the effect of n_s on $P(k)$ is small, we verified that the mean accuracy of our formula P_{MG} is still around 1% when n_s is varied in the interval [0.96, 0.97] which is 1σ away from the value measured by Planck [2].
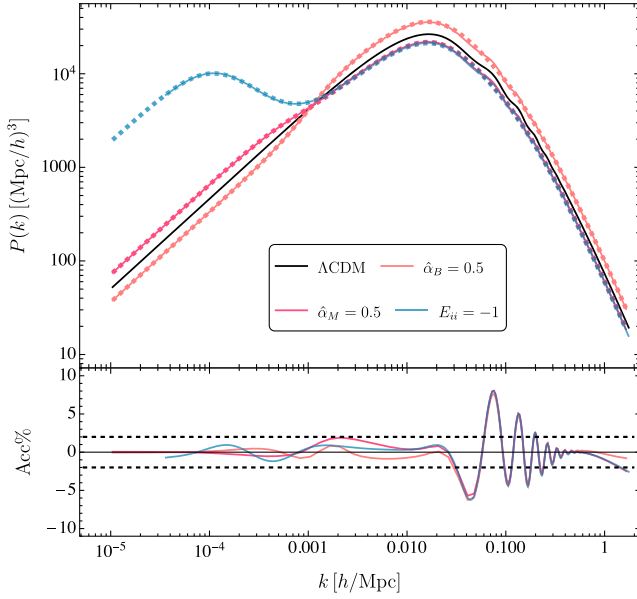
FIG. 2.   Three examples showing the fitting of our formula in Eq. (8) (dots) with respect to numerical results (solid lines). The matter power spectrum of ΛCDM (black line) is plotted to highlight the effects introduced by MG theories. The subplot shows the accuracy of the formulas as a function of k. We can see that most of points are in the 5% region defined by the black dotted lines.

of the matter power spectra from MG theories to visually show the excellent fit provided by our formula with respect to numerical data. In this plot, we use the models `propto_scale` of `hi_class` varying the braiding constant $\hat{\alpha}_B$ and the mass running constant $\hat{\alpha}_M$. From `mgclass`, we used the `plk_late` model assuming the parameters $E_{11} = E_{22} = -1.0$. We also plot the ΛCDM matter power spectrum to highlight the modifications introduced by these alternative theories. In the subplot, we depict the accuracy as a function of k, which can be measure as

$$\text{Acc}\%(k) \equiv \left(\frac{P_{\text{CLASS}} - P_{\text{analytical}}}{P_{\text{CLASS}}}\right) \times 100, \qquad (9)$$

where no mean value is taken, as in Eq. (5). This shows that most of values are in the 2% region defined by the black dotted lines. Remarkably, our formula is able to capture the substantial deviation at large scales observed in the spectrum of the `plk_late` model (depicted in blue) and the significant enhancement around k_{eq} in the `propto_scale` braiding model (depicted in pink). It is also noteworthy that although the expression of P_{small} in Eq. (6) does not consider the wiggles introduced by BAOs in the matter power spectrum, the error of neglecting this oscillatory behavior is around 5% in the most inaccurate cases. The amplitude of the acoustic peaks depends

mainly on the amount of baryonic matter, which translate to the parameter ω_b. However, observations indicate that $\omega_b \approx 0.022$ implying small peaks. Therefore, a line passing through these oscillations seems to be a good approximation.[6]

V. CONCLUSIONS

Accurate fitting formulations for the matter transfer function allow to compute the matter power spectrum within the context of the ΛCDM model [37–39,48]. However, there is no any formulation available when modifications of gravity are considered. In this work, we have put forward a parametric function for $P(k)$ considering several effects introduced by modified gravity theories. The new formula for $P_{\text{MG}}(k)$ was presented in Eq. (8) and it is composed by two main parts. A matter power spectrum P_{large} encompassing modifications at large scales, and the spectrum P_{small} which manifest at smaller scales. The former is proportional to k^{n_s} with amplitude $A(s) \approx A_0(1 + s)$, where s is a parameter measuring the shift of the initial amplitude A_0 which in the case of ΛCDM is approximately given in Eq. (2). The other spectrum; namely P_{small}, was obtained through the machine learning technique of the genetic algorithms using a wCDM model as template. The expression for P_{small} in Eq. (6) involves two parameters; a parameter α measuring overall modifications (enhancement or suppression) in the amplitude of the spectrum, and ω_m which controls the value of the scale at equality k_{eq}. These two spectra; P_{large} and P_{small}, are smoothly connected by the transition function σ defined in Eq. (7). This function marks the transition center at the scale k_0 with a smoothness measured by the parameter β. Therefore, our formulation requires six input parameters; the spectral tilt n_s, the shift s of A_0 at large scales, the enhancement (suppression) factor α measuring modifications in the amplitude of $P_{\text{small}}(k)$, the reduced matter density parameter ω_m, and two parameters, k_0 and β, that determine the center and width of the transition region between the matter power spectra P_{large} and P_{small} [see Eq. (8)].

The performance of our formulation was compared against numerical results obtained from the codes `hi_class` and `mgclass` for several models. The relatively simple expression in Eq. (8) showed a mean accuracy of about 1–2% (see Table I). In Fig. 2, we showed three examples of the fitting of our formula with respect to some MG models. In the subplot, the accuracy of the formulas as a function of k using Eq. (9) was measured. This showed that most of values predicted by our formula lies in the 2% region. Furthermore, although our expression does not consider baryonic acoustic oscillations, neglecting these

---

[6]However, the formula for $P_{\text{MG}}(k)$ should not be used when an accurate description of these wiggles is required. The introduction of these oscillations in $P_{\text{MG}}(k)$ is a further improvement which we leave for a future work.

wiggles introduces an error not greater than 5%. Given that ongoing surveys, like Euclid, expect to reach an observational accuracy of around 1% for some cosmological parameters which are obtained using the matter power spectrum, and that they already use the Eisenstein-Hu formula in their analysis [83], our formula could extend these analysis to model beyond ΛCDM without loosing accuracy. Therefore, the function $P_{MG}(k)$ in Eq. (8) represents a compelling fitting formula for the matter power spectrum of alternative models to ΛCDM.

While this solution does not possess the baryonic wiggles necessary for inferring cosmological parameters from the BAO signal, it can still be used for parameter inference. To illustrate this, let us consider the approach used to forecast cosmological parameters for the Euclid survey [83]. The observed galaxy power spectrum relies on the de-wiggled matter power spectrum, which is the sum of two components; the linear matter power spectrum obtained from a Boltzmann code and the linear matter power spectrum without wiggles but exhibiting the same broad band as the linear matter power spectrum. Both components are modulated by a function that controls the nonlinear damping of the BAO signal in all directions. The no-wiggles power spectrum is often computed using the a fitting formula (for instance, for the ΛCDM models, the EH). It is clear that, when exploring cosmological models beyond the standard framework, it becomes crucial to appropriately parametrize the no-wiggles matter power spectrum in order to accurately capture the amplitude and shape of the observed galaxy power spectrum.

Numerical codes. The genetic algorithm code used here can be found in the GitHub repository [84] of BOQ. This code is based on the GA code by S. N. which can be found at [85].

APPENDIX: GENETIC ALGORITHM EXPRESSIONS FOR k_{eq} AND z_{eq}

Variations in the parameter ω_m implies in general a shift of the radiation-matter transition. In Ref. [38], the scale and redshift at this transition were given by

$$z_{eq}(\omega_m) = 2.5 \times 10^4 \Theta_{2.7}^{-4} \omega_m, \qquad (A1)$$

$$k_{eq}(\omega_m) = 7.46 \times 10^{-2} \Theta_{2.7}^{-2} \omega_m, \qquad (A2)$$

where $\Theta_{2.7}$ defines the cosmic microwave background (CMB) temperature through the relation $T_{CMB} \equiv 2.7\Theta_{2.7}$ K. Taking $T_{CMB} = 2.7255$ K we have the approximated expressions,

$$k_{eq} = 0.0732106\omega_m, \qquad (A3)$$

$$z_{eq} = 24077.4\omega_m. \qquad (A4)$$

In the case of ΛCDM, we used CLASS to get data for k_{eq} and z_{eq} varying ω_m in the range [0.13, 0.15], which is around 10σ from the Planck best-fit. Comparing these datasets with the above formulas, we find that the accuracy of k_{eq} is around 0.37%, while for z_{eq} we get 0.74%. Although these formulas are accurate, the GA was able to find even better analytical formulations. The GAs ended up with the following expressions:

$$k_{eq}(\omega_m) = 0.07222\omega_m + 0.00151\omega_m^{1.3775}, \qquad (A5)$$

$$z_{eq}(\omega_m) = 23896.7\omega_m + 11.6353\omega_m^{1.41595}, \qquad (A6)$$

The accuracy of k_{eq} is 0.023%, while for z_{eq} the GAs achieve an accuracy of 0.0025%, representing an improvement of 1 and 2 orders of magnitude, respectively.

[1] P. J. E. Peebles, *Cosmology's Century: An Inside History of Our Modern Understanding of the Universe* (Princeton University Press, Princeton, NJ, 2020).

[2] N. Aghanim *et al.* (Planck Collaboration), Planck 2018 results. VI. Cosmological parameters, Astron. Astrophys. **641,** A6 (2020); **652,** C4(E) (2021).

[3] M. A. Troxel *et al.* (DES Collaboration), Dark energy survey year 1 results: Cosmological constraints from cosmic shear, Phys. Rev. D **98,** 043528 (2018).

[4] T. M. C. Abbott *et al.* (DES Collaboration), Dark energy survey year 1 results: Constraints on extended cosmological models from galaxy clustering and weak lensing, Phys. Rev. D **99,** 123505 (2019).

[5] T. M. C. Abbott *et al.* (DES Collaboration), Dark energy survey year 3 results: Cosmological constraints from galaxy clustering and weak lensing, Phys. Rev. D **105,** 023520 (2022).

[6] A. G. Riess *et al.* (Supernova Search Team), Observational evidence from supernovae for an accelerating universe and a cosmological constant, Astron. J. **116,** 1009 (1998).

[7] S. Perlmutter *et al.* (Supernova Cosmology Project Collaboration), Measurements of Ω and Λ from 42 high redshift supernovae, Astrophys. J. **517,** 565 (1999).

[8] A. G. Riess *et al.* (Supernova Search Team), Type Ia supernova discoveries at $z > 1$ from the Hubble Space Telescope: Evidence for past deceleration and constraints on dark energy evolution, Astrophys. J. **607,** 665 (2004).

[9] W. percival *et al.*, Baryon acoustic oscillations in the Sloan Digital Sky Survey Data Release 7 galaxy sample, Mon. Not. R. Astron. Soc. **401,** 2148 (2010).

[10] C. Blake *et al.*, The WiggleZ dark energy survey: Mapping the distance-redshift relation with baryon acoustic oscillations, Mon. Not. R. Astron. Soc. **418,** 1707 (2011).

[11] E. Aubourg *et al.*, Cosmological implications of baryon acoustic oscillation measurements, Phys. Rev. D **92,** 123516 (2015).

[12] P. de Bernardis *et al.* (Boomerang Collaboration), A flat universe from high resolution maps of the cosmic microwave background radiation, Nature (London) **404,** 955 (2000).

[13] A. H. Jaffe *et al.*, Recent results from the maxima experiment, New Astron. Rev. **47,** 727 (2003).

[14] Y. Akrami *et al.* (Planck Collaboration), Planck 2018 results. X. Constraints on inflation, Astron. Astrophys. **641,** A10 (2020).

[15] Y. Akrami *et al.* (Planck Collaboration), Planck 2018 results. VII. Isotropy and statistics of the CMB, Astron. Astrophys. **641,** A7 (2020).

[16] T. M. C. Abbott *et al.* (DES Collaboration), Cosmological constraints from multiple probes in the dark energy survey, Phys. Rev. Lett. **122,** 171301 (2019).

[17] L. Perivolaropoulos and F. Skara, Challenges for ΛCDM: An update, New Astron. Rev. **95,** 101659 (2022).

[18] E. Abdalla *et al.*, Cosmology intertwined: A review of the particle physics, astrophysics, and cosmology associated with the cosmological tensions and anomalies, J. High Energy Astrophys. **34,** 49 (2022).

[19] K. C. Wong *et al.*, H0LiCOW—XIII. A 2.4 per cent measurement of H0 from lensed quasars: 5.3σ tension between early- and late-Universe probes, Mon. Not. R. Astron. Soc. **498,** 1420 (2020).

[20] A. G. Riess, S. Casertano, W. Yuan, J. B. Bowers, L. Macri, J. C. Zinn, and D. Scolnic, Cosmic distances calibrated to 1% precision with Gaia EDR3 parallaxes and Hubble Space Telescope Photometry of 75 Milky Way Cepheids confirm tension with ΛCDM, Astrophys. J. Lett. **908,** L6 (2021).

[21] A. G. Riess *et al.*, A comprehensive measurement of the local value of the Hubble constant with $1 \text{ km s}^{-1} \text{ Mpc}^{-1}$ uncertainty from the Hubble Space Telescope and the SH0ES team, Astrophys. J. Lett. **934,** L7 (2022).

[22] E. Di Valentino, O. Mena, S. Pan, L. Visinelli, W. Yang, A. Melchiorri, D. F Mota, A. G. Riess, and J. Silk, In the realm of the Hubble tension—a review of solutions, Classical Quantum Gravity **38,** 153001 (2021).

[23] E. N. Saridakis *et al.* (CANTATA Collaboration), *Modified Gravity and Cosmology: An Update by the CANTATA Network* (2021), .

[24] R. Arjona, W. Cardona, and S. Nesseris, Unraveling the effective fluid approach for $f(R)$ models in the subhorizon approximation, Phys. Rev. D **99,** 043516 (2019).

[25] R. Arjona, W. Cardona, and S. Nesseris, Designing Horndeski and the effective fluid approach, Phys. Rev. D **100,** 063526 (2019).

[26] W. Cardona, J. B. Orjuela-Quintana, and C. A. Valenzuela-Toledo, An effective fluid description of scalar-vector-tensor theories under the sub-horizon and quasi-static approximations, J. Cosmol. Astropart. Phys. 08 (2022) 059.

[27] S. Nesseris, The effective fluid approach for modified gravity and its applications, Universe **9,** 2022 (2023).

[28] J. K. Bloomfield, E. E. Flanagan, M. Park, and S. Watson, Dark energy or modified gravity? An effective field theory approach, J. Cosmol. Astropart. Phys. 08 (2013) 010.

[29] G. Gubitosi, F. Piazza, and F. Vernizzi, The effective field theory of dark energy, J. Cosmol. Astropart. Phys. 02 (2013) 032.

[30] N. Frusciante and L. Perenon, Effective field theory of dark energy: A review, Phys. Rep. **857,** 1 (2020),

[31] M. Kunz and D. Sapone, Dark energy versus modified gravity, Phys. Rev. Lett. **98,** 121301 (2007).

[32] M. Tegmark *et al.* (SDSS Collaboration), The 3-D power spectrum of galaxies from the SDSS, Astrophys. J. **606,** 702 (2004).

[33] S. Alam *et al.* (BOSS Collaboration), The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: Cosmological analysis of the DR12 galaxy sample, Mon. Not. R. Astron. Soc. **470,** 2617 (2017).

[34] S. Dodelson and F. Schmidt, *Modern Cosmology* (Elsevier Science, New York, 2020).

[35] A. Lewis, A. Challinor, and A. Lasenby, Efficient computation of CMB anisotropies in closed FRW models, Astrophys. J. **538,** 473 (2000).

[36] D. Blas, J. Lesgourgues, and T. Tram, The cosmic linear anisotropy solving system (CLASS). Part II: Approximation schemes, J. Cosmol. Astropart. Phys. 07 (2011) 034.

[37] J. M. Bardeen, J. R. Bond, N. Kaiser, and A. S. Szalay, The statistics of peaks of Gaussian random fields, Astrophys. J. **304,** 15 (1986).

[38] D. J. Eisenstein and W. Hu, Baryonic features in the matter transfer function, Astrophys. J. **496,** 605 (1998).

[39] D. J. Eisenstein and W. Hu, Power spectra for cold dark matter and its variants, Astrophys. J. **511,** 5 (1997).

[40] A. Dvornik *et al.*, KiDS-1000: Combined halo-model cosmology constraints from galaxy abundance, galaxy clustering and galaxy-galaxy lensing, Astron. Astrophys. **675,** A189 (2023).

[41] D. Vallés-Pérez, S. Planelles, O. Monllor-Berbegal, and V. Quilis, On the choice of the most suitable indicator for the assembly state of dark matter haloes through cosmic time, Mon. Not. R. Astron. Soc. **519,** 6111 (2023).

[42] J.-E. Campagne et al., JAX-COSMO: An end-to-end differentiable and GPU accelerated cosmology library, Open J. Astrophys. **6,** 1 (2023).

[43] G. C. Petter, R. C. Hickox, D. M. Alexander, A. D. Myers, J. E. Geach, K. E. Whalen, and C. P. Andonie, Host dark matter halos of wide-field infrared survey explorer-selected obscured and unobscured quasars: Evidence for evolution, Astrophys. J. **946,** 27 (2023).

[44] V. Ajani et al. (Euclid Collaboration), Euclid preparation XXIX: Forecasts for 10 different higher-order weak lensing statistics, Astron. Astrophys. **675,** A120 (2023).

[45] L. R. Prole, A. T. P. Schauer, P. C. Clark, S. C. O. Glover, F. D. Priestley, and R. S. Klessen, From dark matter halos to pre-stellar cores: High resolution follow-up of cosmological Lyman–Werner simulations, Mon. Not. R. Astron. Soc. **520,** 2081 (2023).

[46] A. Variu, C. Zhao, D. Forero-Sánchez, C.-H. Chuang, F.-S. Kitaura, C. Tao, A. Tamone, and J.-P. Kneib, Cosmic void exclusion models and their impact on the distance scale measurements from large-scale structure, Mon. Not. R. Astron. Soc. **521,** 4731 (2023).

[47] G. Hütsi, M. Raidal, J. Urrutia, V. Vaskonen, and H. Veermäe, Did JWST observe imprints of axion miniclusters or primordial black holes?, Phys. Rev. D **107,** 043502 (2023).

[48] J. B. Orjuela-Quintana, S. Nesseris, and W. Cardona, Using machine learning to compress the matter transfer function $T(k)$, Phys. Rev. D **107,** 083520 (2023).

[49] M. Zumalacárregui, E. Bellini, I. Sawicki, J. Lesgourgues, and P. G. Ferreira, hi_class: Horndeski in the cosmic linear anisotropy solving system, J. Cosmol. Astropart. Phys. 08 (2017) 019.

[50] E. Bellini, I. Sawicki, and M. Zumalacárregui, hi_class: Background evolution, initial conditions and approximation schemes, J. Cosmol. Astropart. Phys. 02 (2020) 008.

[51] T. Baker and P. Bull, Observational signatures of modified gravity on ultra-large scales, Astrophys. J. **811,** 116 (2015).

[52] Z. Sakr and M. Martinelli, Cosmological constraints on sub-horizon scales modified gravity theories with MGCLASS II, J. Cosmol. Astropart. Phys. 05 (2022) 030.

[53] J. Koza and J. Koza, *Genetic Programming: On the Programming of Computers by Means of Natural Selection* (The MIT Press, Bradford, 1992).

[54] S. Nesseris and J. Garcia-Bellido, A new perspective on dark energy modeling via genetic algorithms, J. Cosmol. Astropart. Phys. 11 (2012) 033.

[55] B. C. Allanach, D. Grellscheid, and F. Quevedo, Genetic algorithms and experimental discrimination of SUSY models, J. High Energy Phys. 07 (2004) 069.

[56] Y. Akrami, P. Scott, J. Edsjo, J. Conrad, and L. Bergstrom, A profile likelihood analysis of the constrained MSSM with genetic algorithms, J. High Energy Phys. 04 (2010) 057.

[57] S. Abel, D. G. Cerdeño, and S. Robles, The power of genetic algorithms: What remains of the pMSSM?, arXiv:1805.03615.

[58] X.-L. Luo, J. Feng, and H.-H. Zhang, A genetic algorithm for astroparticle physics studies, Comput. Phys. Commun. **250,** 106818 (2020).

[59] M. Ho, M. M. Rau, M. Ntampaka, A. Farahi, H. Trac, and B. Póczos, A robust and efficient deep learning method for dynamical mass measurements of galaxy clusters, Astrophys. J. **887,** 25 (2019).

[60] A. Aizpuru, R. Arjona, and S. Nesseris, Machine learning improved fits of the sound horizon at the baryon drag epoch, Phys. Rev. D **104,** 043521 (2021).

[61] S. Nesseris and A. Shafieloo, A model independent null test on the cosmological constant, Mon. Not. R. Astron. Soc. **408,** 1879 (2010).

[62] R. Arjona and S. Nesseris, Novel null tests for the spatial curvature and homogeneity of the Universe and their machine learning reconstructions, Phys. Rev. D **103,** 103539 (2021).

[63] C. Bogdanos and S. Nesseris, Genetic algorithms and supernovae type Ia analysis, J. Cosmol. Astropart. Phys. 05 (2009) 006.

[64] R. Arjona and S. Nesseris, What can machine learning tell us about the background expansion of the Universe?, Phys. Rev. D **101,** 123525 (2020).

[65] R. Arjona and S. Nesseris, Hints of dark energy anisotropic stress using machine learning J. Cosmol. Astropart. Phys. 11 (2020) 042.

[66] R. Arjona, A. Melchiorri, and S. Nesseris, Testing the ΛCDM paradigm with growth rate data and machine learning, J. Cosmol. Astropart. Phys. 05 (2022) 047.

[67] G. Alestas, L. Kazantzidis, and S. Nesseris, Machine learning constraints on deviations from general relativity from the large scale structure of the Universe, Phys. Rev. D **106,** 103519 (2022).

[68] R. Arjona, H.-N. Lin, S. Nesseris, and L. Tang, Machine learning forecasts of the cosmic distance duality relation with strongly lensed gravitational wave events, Phys. Rev. D **103,** 103513 (2021).

[69] M. Martinelli et al. (Euclid Collaboration), Euclid: Forecast constraints on the cosmic distance duality relation with complementary external probes, Astron. Astrophys. **644,** A80 (2020).

[70] S. Nesseris et al. (Euclid Collaboration), Euclid: Forecast constraints on consistency tests of the ΛCDM model, Astron. Astrophys. **660,** A67 (2022).

[71] M. Schmidt and H. Lipson, Distilling free-form natural laws from experimental data, Science **324,** 81 (2009).

[72] S. L. Brunton, J. L. Proctor, and J. N. Kutz, Discovering governing equations from data by sparse identification of nonlinear dynamical systems, Proc. Natl. Acad. Sci. U.S.A. **113,** 3932 (2016).

[73] S.-M. Udrescu and M. Tegmark, AI Feynman: a physics-inspired method for symbolic regression, Sci. Adv. **6,** eaay2631 (2020).

[74] Z. Liu and M. Tegmark, Machine learning hidden symmetries, Phys. Rev. Lett. **128,** 180201 (2022).

[75] M. Cranmer et al., Discovering symbolic models from deep learning with inductive biases, arXiv:2006.11287.

[76] D. J. Bartlett, H. Desmond, and P. G. Ferreira, Exhaustive symbolic regression, IEEE Trans. Evol. Comput. 1 (2024), 10.1109/TEVC.2023.3280250.

[77] D. Sapone and M. Kunz, Fingerprinting dark energy, Phys. Rev. D **80**, 083519 (2009).

[78] D. Sapone and E. Majerotto, Fingerprinting dark energy III: Distinctive marks of viscosity, Phys. Rev. D **85**, 123529 (2012).

[79] D. Sapone, E. Majerotto, M. Kunz, and B. Garilli, Can dark energy viscosity be detected with the Euclid survey?, Phys. Rev. D **88**, 043503 (2013).

[80] E. Bellini and I. Sawicki, Maximal freedom at minimum cost: Linear large-scale structure in general modifications of gravity, J. Cosmol. Astropart. Phys. 07 (2014) 050.

[81] E. Bellini, A. J. Cuesta, R. Jimenez, and L. Verde, Constraints on deviations from ΛCDM within Horndeski gravity, J. Cosmol. Astropart. Phys. 02 (**2016**) 053; 06 (2016) E01.

[82] P. A. R. Ade *et al.* (Planck Collaboration), Planck 2015 results. XIV. Dark energy and modified gravity, Astron. Astrophys. **594**, A14 (2016).

[83] A. Blanchard *et al.* (Euclid Collaboration), Euclid preparation: VII. Forecast validation for Euclid cosmological probes, Astron. Astrophys. **642**, A191 (2020).

[84] https://github.com/BayronO/Pk-in-Modified-Gravity.

[85] https://github.com/snesseris/Genetic-Algorithms.