Internal robustness of growth rate data

Bryan Sagredo,[1,*] Savvas Nesseris,[2,†] and Domenico Sapone[1,‡]

[1]*Cosmology and Theoretical Astrophysics group, Departamento de Física, FCFM,*
Universidad de Chile, Blanco Encalada 2008, Santiago, Chile
[2]*Instituto de Física Teórica UAM-CSIC, Universidad Autonóma de Madrid,*
Cantoblanco, 28049 Madrid, Spain

We perform an *internal robustness* analysis (iR) to a compilation of the most recent $f\sigma_8(z)$ data, using the framework of Ref. [1]. The method analyzes combinations of subsets in the data set in a Bayesian model comparison way, potentially finding outliers, subsets of data affected by systematics or new physics. In order to validate our analysis and assess its sensitivity we performed several cross-checks, for example by removing some of the data or by adding artificially contaminated points, while we also generated mock data sets in order to estimate confidence regions of the iR. Applying this methodology, we found no anomalous behavior in the $f\sigma_8(z)$ data set, thus validating its internal robustness.

I. INTRODUCTION

During the last twenty years a plethora of observations suggests that the Universe is undergoing a phase of accelerated expansion at late times. In order to explain this phenomenon, the concept of attractive gravity had to be revised either by introducing a new form of matter dubbed dark energy, see Ref. [2] for a review, or by altering explicitly the laws of gravity [3]. However, the simplest way to account for a phase of accelerated expansion of the Universe within the framework of Friedmann-Lemaître-Robertson-Walker (FLRW) cosmology is to simply introduce a cosmological constant (Λ). While this model gives rise to severe coincidence and fine-tuning problems, current cosmological observations are still compatible with a Universe that is filled by a dark energy component that has the same characteristics of the cosmological constant [4].

Nonetheless, these cosmological observations are not accurate enough at the moment to either constrain any potential time evolution of the cosmological constant, which might lead dark energy to cluster, or any modifications of gravity. Despite the fact that the two aforementioned classes of theories can in principle be arbitrarily similar [5,6], it is still necessary to be able to discriminate between the currently available models.

Future surveys such as Euclid [7], LSST [8], and DESI [9], all of which will gather orders of magnitude more data than current surveys, it would be interesting to constrain the dynamical features of gravity and test our assumptions.

One way to do this is via the growth of matter density perturbations $\delta_m = \delta\rho_m/\rho_m$ and its logarithmic derivative the growth rate $f = d\ln\delta_m/d\ln a$. In practice, most of the growth rate measurements are made via the peculiar velocities obtained from redshift space distortions (RSD) measurements [10] coming from galaxy redshift surveys. In general, these surveys measure the perturbations of the galaxy density δ_g, which can be related to (dark) matter perturbations through the bias b via $\delta_g = b\delta_m$. Thus, initial growth rate measurements measured the growth rate f divided by the bias factor b leading to the parameter $\beta \equiv f/b$. This parameter is very sensitive to the value of the bias which can vary in the range $b \in [1, 3]$, thus making difficult to combine β from different surveys and as a result leading to unreliable data sets of $\beta(z_i)$.

As a result, a more reliable parameter was sought and this was found in the combination $f(z)\sigma_8(z) \equiv f\sigma_8(z)$, which can be shown to be independent of the bias [11], and can be measured either via weak lensing or RSD observations. Still, the current measurements of $f\sigma_8(z)$ (presented in later sections) come from a plethora of different surveys with different assumptions and systematics, thus an approach to study the statistical properties and robustness of the data is imperative.

One such approach would be to perform a tomographic analysis, as was done in Ref. [12], where growth data from different years and different redshifts were split into subsamples. Then, it was found that for the more recent data the agreement with the Planck 15 best-fit ΛCDM cosmology is much improved and is well within 1σ, due to the fact that newer data are at higher redshifts and with higher errors. This approach can clearly highlight inconsistencies in the data, but care must be taken to avoid double-counting of the data.

_____

[*]bryan.sagredo@ing.uchile.cl
[†]savvas.nesseris@csic.es
[‡]dsapone@ing.uchile.cl

Another approach to test the internal consistency of low redshift probes was developed in Ref. [13], where the KiDs data were examined for internal tensions, by performing cuts of the data and examining the cross-correlation measurements of the correlation function that was presented in four tomographic redshift bins by the KiDs collaboration. Then it was found that the KiDs data might have internal tensions of ~ 2.2–3.5σ significance.

In our paper we choose to follow a more direct approach which is based on Bayesian analysis called "internal robustness" method and was pioneered in Ref. [1], see also Refs. [14,15] for applications of the method in supernovae data. This is a fully Bayesian approach which is not only sensitive to the local minimum like a standard χ^2 comparison, but also to the entire likelihood and can in principle detect the presence of systematics in the data set. The main goal of this approach is to identify systematic-contaminated data points, which can then be further analyzed and potentially excluded if they cannot be corrected.

In this paper we present an application of the internal robustness approach to the currently available growth-rate data in the form of $f\sigma_8(z)$ with the aim to examine the data set for systematics and outlier points in a fully automated manner. The reason we specifically use these data is twofold: first, the growth data are dynamic probes that can in principle discriminate between modifications of general relativity [16]; second, currently there is a tension between low redshift probes, such as the growth data, and the Planck 15 best fit cosmology. Therefore, it is imperative to have a data set whose statistical properties have been demonstrated to be robust before inferences about modified gravity models are made or the tension with Planck could be explained.

The layout of our manuscript is as follows: in Sec. II we provide the basic elements of FLRW cosmology related to our models, in Sec. III we briefly review the internal robustness method and its application to the growth data, while in Sec. IV we provide the compilation of growth data used in our analysis and finally, we discuss our results in Sec. V.

II. BASIC EQUATIONS

In this section we present the basic equations required in our analysis. We begin with the Hubble parameter in a flat ΛCDM universe (with a constant equation of state parameter for dark energy $w = -1$), given by

$$H(a)^2 = H_0^2[\Omega_{m,0}a^{-3} + (1 - \Omega_{m,0})], \quad (1)$$

where H_0 is the Hubble constant, and $\Omega_{m,0}$ is the present day value of the matter density parameter and a is the scale factor. The matter density can then also be expressed as a function of the scale factor:

$$\Omega_m(a) = \frac{\Omega_{m,0}a^{-3}}{H(a)^2/H_0^2}. \quad (2)$$

Under the assumption of flat ΛCDM model, the angular diameter distance takes an analytical expression, given by:

$$\frac{H_0}{c}\frac{\sqrt{\Omega_{m,0}}}{2a}d_A(a) = {}_2F_1\left[\frac{1}{2},\frac{1}{6};1+\frac{1}{6};\left(1-\frac{1}{\Omega_{m,0}}\right)\right]$$
$$- \sqrt{a}\cdot{}_2F_1\left[\frac{1}{2},\frac{1}{6};1+\frac{1}{6};\left(1-\frac{1}{\Omega_m(a)}\right)\right], \quad (3)$$

where ${}_2F_1$ is the hypergeometric function. The matter density perturbations in Fourier space $\delta_m(a,k)$ depend on the underlying cosmological model; for the ΛCDM scenario, the linear matter perturbations grow according to

$$\delta_m''(a) + \left(\frac{3}{a} + \frac{H'(a)}{H(a)}\right)\delta_m'(a) - \frac{3}{2}\frac{\Omega_m(a)}{a^2}\delta_m(a) = 0. \quad (4)$$

The equation above has an analytical solution for the growing mode, given by [17–19]

$$\delta_m(a) = a\cdot{}_2F_1\left(\frac{1}{3},1;\frac{11}{6};a^3\left(1-\frac{1}{\Omega_{m,0}}\right)\right). \quad (5)$$

Note that the dependence on the wave number k disappears because of the assumption of small scales approximation.

We define the growth rate f and the root mean square (RMS) normalization of the matter power spectrum σ_8 as:

$$f(a) = \frac{d\log\delta_m}{d\log a}, \quad (6)$$

$$\sigma_8(a) = \sigma_{8,0}\frac{\delta_m(a)}{\delta_m(1)}, \quad (7)$$

$$\sigma_{8,0}^2 = \langle\delta(x)^2\rangle = \frac{1}{2\pi^2}\int_0^\infty P(k)W_R^2(k)k^2 dk, \quad (8)$$

where $W_R(k)$ is the Fourier transform of a top-hat window function. As already mentioned in Sec. I, a more robust and reliable quantity that is measured by redshift surveys is the combination of the growth rate $f(a)$ and the RMS $\sigma_8(a)$:

$$f\sigma_8(a) = a\frac{\delta_m'(a)}{\delta_m(1)}\sigma_{8,0}. \quad (9)$$

Equation (9) will be our key quantity, which will be tested with the most recent data available in the following sections.

Alternatively, $f\sigma_8(a)$ can also be written as [20]

$$f\sigma_8(a) \equiv \frac{\sigma_8^{(vd)}(a)^2}{\sigma_8^{(dd)}(a)}, \quad (10)$$

where $\sigma_8^{(dd)}(a)$ is the usual $\sigma_8(a)$ parameter as defined above and $\sigma_8^{(vd)}(a)$ is the smoothed density-velocity correlation defined in a similar manner, but using instead the correlation power spectrum $P_{vd}(k)$ and $v = -\nabla v_N/H$ where v_N is the peculiar velocity of the baryons and dark matter in the Newtonian-gauge, while d is the total matter density perturbation. Using linear theory for models close to ΛCDM, it is easy to show that $v = -\nabla v_N/H = -ik\frac{v_N}{aH} = f(a)\delta(k,a)$, see Sec. IX.2 in Ref. [21] for a quick derivation. However, the growth can also be written as $\delta(k,a) = \delta_k(k)G(a)$, where δ_k is an initial condition determined from inflation and $G(a) \equiv \frac{\delta(a)}{\delta(1)}$ is the normalized growth, see Eq. (7.8) in Ref. [21]. Then, using the definitions of the σ parameters, we have

$$
\begin{aligned}
\sigma_8^{(vd)}(a)^2 &= \langle v(x,a)\delta(x,a)\rangle \\
&= f(a)G(a)^2\langle\delta(x)^2\rangle \\
&= f(a)G(a)^2\sigma_{8,0}^2,
\end{aligned} \tag{11}
$$

and

$$
\begin{aligned}
\sigma_8^{(dd)}(a)^2 &= \langle\delta(x,a)\delta(x,a)\rangle \\
&= G(a)^2\langle\delta(x)^2\rangle \\
&= G(a)^2\sigma_{8,0}^2.
\end{aligned} \tag{12}
$$

Finally, using the definition of Ref. [20] we have

$$
\begin{aligned}
f\sigma_8(a) &\equiv \frac{\sigma_8^{(vd)}(a)^2}{\sigma_8^{(dd)}(a)} \\
&= \frac{f(a)G(a)^2\sigma_{8,0}^2}{G(a)\sigma_{8,0}} \\
&= f(a)G(a)\sigma_{8,0},
\end{aligned} \tag{13}
$$

which exactly agrees with our original definition of Eq. (9).

III. FORMALISM

Here we report the basic equations that will be used to perform our analysis and we refer to [1] for the details on the derivation of the internal robustness. The statistical definition of the Bayesian evidence is

$$
\mathcal{E}(\mathbf{x}|M) = \int L(\mathbf{x}|\boldsymbol{\theta}^M)\pi(\boldsymbol{\theta}^M)\mathrm{d}\boldsymbol{\theta}^M, \tag{14}
$$

where $\mathbf{x} = (x_1, x_2, ..., x_N)$ are the N random data, $\boldsymbol{\theta}^M = (\theta_1, \theta_2, ..., \theta_n)$ are the n theoretical parameters of the model M. The likelihood function is $L(\mathbf{x}|\boldsymbol{\theta}^M)$, while the prior probability of the parameters of the model is $\pi(\boldsymbol{\theta}^M)$.

By the help of the Bayes' theorem we can find the posterior probability $\mathcal{P}(M|\mathbf{x})$ of having a model M given the data

$$
\mathcal{P}(M|\mathbf{x}) = \mathcal{E}(\mathbf{x}|M)\frac{\pi(M)}{\pi(\mathbf{x})}. \tag{15}
$$

By using the last equation, we can compare two different models by considering the ratio of their probabilities:

$$
\frac{\mathcal{P}(M_1|\mathbf{x})}{\mathcal{P}(M_2|\mathbf{x})} = \mathcal{B}_{12}\frac{\pi(M_1)}{\pi(M_2)}; \tag{16}
$$

with the Bayes factor defined as

$$
\mathcal{B}_{12} = \frac{\mathcal{E}(\mathbf{x}|M_1)}{\mathcal{E}(\mathbf{x}|M_2)}. \tag{17}
$$

If we assume that the prior probabilities of having two different models are the same, then the Bayes factor alone will help us to favor or disfavor a particular model. If $\mathcal{B}_{12} > 1$ then the data favors the model M_1, if it is less than 1, then M_2 is favored.

However, the robustness test needs a further assumption, that is: the data have to come from two different distributions. The reason is two fold: first the total evidence can be factorized as the product of the two evidences and, second, which is the underlying meaning of the robustness test, we would like to prove that data are reliable. If the data is partitioned into two subsets, say $\{\mathbf{x_1}, \mathbf{x_2}\}$ and we assume that they come from two models, say M_1, M_2, then the Bayes factor becomes

$$
\mathcal{B}_{12} = \frac{\mathcal{E}(\mathbf{x}|M_1)}{\mathcal{E}(\mathbf{x_1}|M_1)\mathcal{E}(\mathbf{x_2}|M_2)}. \tag{18}
$$

Finally, we can define the internal robustness as

$$
\mathrm{iR}_{12} = \log\mathcal{B}_{12} = \log\left(\frac{\mathcal{E}(\mathbf{x}|M_1)}{\mathcal{E}(\mathbf{x_1}|M_1)\mathcal{E}(\mathbf{x_2}|M_2)}\right). \tag{19}
$$

This approach allows us to detect if a subset of the data follows another cosmological model or if a specific survey is affected by systematics and hence altering the measurement itself.

However, the assumption of having two different models is not strictly mandatory and we will choose the same cosmological model for both subsets. In this work we will set the cosmological model to be ΛCDM and the parameters for both subsets to $\boldsymbol{\theta} = (\Omega_{m,0}, \sigma_{8,0})$. Hereinafter, we drop the M superscript and the 1,2 subscripts, since we only consider one cosmological model.

Our analysis invokes priors on the parameters, for that we choose a flat prior in the $[0, 1]$ range for $\Omega_{m,0}$ in order to

allow for all physical values possible for the matter density. On the other hand, the choice of a prior for $\sigma_{8,0}$ is less evident; since the prior directly affects the evidence value, so we are going to consider three priors for $\sigma_{8,0}$, to assess the impact of the prior selection on the internal robustness. We choose the following three cases:

  (i) Narrow flat prior: this is a typical flat prior in the range [0.3, 1.5].
  (ii) Broad flat prior: this is a flat prior in the range [0, 10], which allows for high values of $\sigma_{8,0}$.
  (iii) Gaussian prior: the third prior to consider is a Gaussian distribution centered on 0.8150, with a standard deviation of 0.0087, based on the Planck 2015 results (TT, TE, EE + lowP + lensing [4]).

It is clear that we only allow for positive values of $\sigma_{8,0}$ in order to remain physical.

The data considered are $f\sigma_8(z)$ measurements (with $z = -1 + 1/a$ being the redshift of the measurement), and we assume a Gaussian likelihood for the data with a covariance matrix $\mathbf{C}$. We represent the observed data in different redshifts as $\mathbf{m} = (m(z_1), \ldots, m(z_n))$ and its theoretical prediction as $\boldsymbol{\mu}(\boldsymbol{\theta}) = (\mu(z_1), \ldots, \mu(z_n))$, which depends on the cosmological model and parameters. We also take into account the redshift correction of [22], which features a correction factor of

$$\mathrm{fac}(z^i) = \frac{H(z^i)d_A(z^i)}{H^{\mathrm{ref},i}(z^i)d_A^{\mathrm{ref},i}(z^i)}. \tag{20}$$

with the label ref, i stating that the cosmology considered is the reference cosmology used on the corresponding data point on redshift z^i. Hence, the corrected theoretical prediction is [23]

$$\mu_c^i = \frac{\mu^i}{\mathrm{fac}(z^i)}. \tag{21}$$

We are now in the position to define the data vector with the corresponding modification:

$$\mathbf{x} = \mathbf{m} - \boldsymbol{\mu}_c. \tag{22}$$

Then, the chi-squared is

$$\chi^2 = \mathbf{x}^T \mathbf{C}^{-1} \mathbf{x}, \tag{23}$$

which is related to the likelihood via $L = e^{-\chi^2/2}/\sqrt{(2\pi)^n|\mathbf{C}|}$.

To speed up the computations, we note that the $\sigma_{8,0}$ parameter can be marginalized theoretically [24,25]. We rewrite the χ^2:

$$\chi^2 = \mathbf{m}^T \mathbf{C}^{-1} \mathbf{m} - 2\mathbf{m}^T C^{-1} \boldsymbol{\mu}_c + \boldsymbol{\mu}_c^T \mathbf{C}^{-1} \boldsymbol{\mu}_c. \tag{24}$$

The corrected theoretical prediction marginalized over $\sigma_{8,0}$ will be $\boldsymbol{\nu} = \boldsymbol{\mu}_c/\sigma_{8,0}$. Then, the χ^2 can be rewritten as

$$\chi^2 = \xi_{mm} - 2\xi_{m\nu}\sigma_{8,0} + \xi_{\nu\nu}\sigma_{8,0}^2, \tag{25}$$

where the single terms are

$$\xi_{mm} = \mathbf{m}^T \mathbf{C}^{-1} \mathbf{m},$$
$$\xi_{m\nu} = \mathbf{m}^T \mathbf{C}^{-1} \boldsymbol{\nu},$$
$$\xi_{\nu\nu} = \boldsymbol{\nu}^T \mathbf{C}^{-1} \boldsymbol{\nu}. \tag{26}$$

The posterior probability distribution marginalized over $\sigma_{8,0}$ is

$$\mathcal{P}(\Omega_{m,0}) = \int L(\Omega_{m,0}, \sigma_{8,0}) \pi(\Omega_{m,0}, \sigma_{8,0}) \mathrm{d}\sigma_{8,0}. \tag{27}$$

We now consider two cases for the prior probability on $\sigma_{8,0}$: a flat prior between $[a, b]$ and a Gaussian prior with mean s and variance ϵ^2. Let us start by considering the flat prior case. The integration of the posterior is

$$\mathcal{P}_f(\Omega_{m,0}) = \frac{1}{(b-a)\sqrt{(2\pi)^n|\mathbf{C}|}} \int_a^b e^{-\chi^2/2} \mathrm{d}\sigma_{8,0}$$
$$= \frac{1}{(b-a)\sqrt{(2\pi)^n|\mathbf{C}|}} \exp\left(-\frac{1}{2}\left[\xi_{mm} - \frac{\xi_{m\nu}^2}{\xi_{\nu\nu}}\right]\right) I_f, \tag{28}$$

where the quantity I_f is

$$I_f = \int_a^b e^{-\frac{\xi_{\nu\nu}}{2}(\xi_{m\nu}/\xi_{\nu\nu} - \sigma_8)^2} \mathrm{d}\sigma_8$$
$$= \sqrt{\frac{\pi}{2\xi_{\nu\nu}}} \mathrm{erf}\left(\frac{\xi_{m\nu} - x\xi_{\nu\nu}}{\sqrt{2\xi_{\nu\nu}}}\right)\Bigg|_a^b. \tag{29}$$

For the Gaussian prior case we find, by discarding negative values:

$$\pi_g(\sigma_{8,0}) = \frac{e^{-\frac{1}{2}(s-\sigma_{8,0})^2/\epsilon^2}}{\int_0^\infty e^{-\frac{1}{2}(s-\sigma_{8,0})^2/\epsilon^2} \mathrm{d}\sigma_{8,0}}$$
$$= \frac{e^{-\frac{1}{2}(s-\sigma_{8,0})^2/\epsilon^2}}{\sqrt{\pi\epsilon^2/2}[1 + \mathrm{erf}(s/\sqrt{2\epsilon^2})]}$$
$$= A_g e^{-\frac{1}{2}(s-\sigma_{8,0})^2/\epsilon^2}, \tag{30}$$

where we implicitly defined the normalization constant A_g.

The posterior probability distribution function then reads:

$$\mathcal{P}_g(\Omega_{m,0}) = \frac{A_g}{\sqrt{(2\pi)^n|\mathbf{C}|}} \int_0^\infty e^{-[\chi^2 + (s-\sigma_{8,0})^2/\epsilon^2]/2} \mathrm{d}\sigma_{8,0}$$
$$= \frac{A_g}{\sqrt{(2\pi)^n|\mathbf{C}|}} \exp\left(-\frac{1}{2}\left[\xi_{mm} - \frac{\xi_{m\nu}^2}{\xi_{\nu\nu}}\right]\right) I_g, \tag{31}$$

where I_g is equal to:

$$I_g = \int_0^\infty e^{-\frac{\xi_{\nu\nu}}{2}(\xi_{m\nu}/\xi_{\nu\nu}-\sigma_8)^2} e^{-\frac{1}{2}(s-\sigma_{8,0})^2/\epsilon^2} d\sigma_8$$

$$= \sqrt{\frac{\pi\epsilon^2}{2(\epsilon^2\xi_{\nu\nu}+1)}} \exp\left(-\frac{(\xi_{m\nu}-\xi_{\nu\nu}s)^2}{2\xi_{\nu\nu}(\epsilon^2\xi_{\nu\nu}+1)}\right)$$

$$\times \left[1 + \mathrm{erf}\left(\frac{\xi_{m\nu}\epsilon^2+s}{\epsilon\sqrt{2(\epsilon^2\xi_{\nu\nu}+1)}}\right)\right], \quad (32)$$

which is the multiplication of the exponentials of two Gaussians, which is also the exponential part of a Gaussian distribution.

IV. DATA CONSIDERATIONS

A. The data set

The growth rate data set is based on the Gold-2017 compilation from [22], consisting of 18 independent measurements of $f\sigma_8(z)$, obtained from redshift space distortion measurements from a variety of surveys. Among these surveys, it is important to note that the three WiggleZ [26] measurements are correlated, and their covariance matrix is

$$\mathbf{C}_{\mathrm{WiggleZ}} = 10^{-3}\begin{pmatrix} 6.400 & 2.570 & 0.000 \\ 2.570 & 3.969 & 2.540 \\ 0.000 & 2.540 & 5.184 \end{pmatrix}. \quad (33)$$

In addition to the Gold-2017 compilation, we update it with 4 recent measurements from [27]. These points have a covariance matrix given by

$$\mathbf{C}_{\mathrm{SDSS\text{-}IV}} = 10^{-2}\begin{pmatrix} 3.098 & 0.892 & 0.329 & -0.021 \\ 0.892 & 0.980 & 0.436 & 0.076 \\ 0.329 & 0.436 & 0.490 & 0.350 \\ -0.021 & 0.076 & 0.350 & 1.124 \end{pmatrix}.$$
$$\quad (34)$$

Our final data set will be constituted of $N = 22$ data points, shown in Table I, the possible combinations of subsets from them is $2^{N-1} - 1 = 2097151$, and we analyze all of the subsets.[1] The analysis is possible thanks primarily to the marginalization over $\sigma_{8,0}$, as shown in Sec. III.

We do not use the data set of Ref. [16], even though some data points are at high redshifts, as these are measurements of the growth-rate $f(a)$, which is affected by the bias b, and not of the combination $f\sigma_8$ which has been shown to be bias free [11].

---

[1]Note that we do not count the combination of the full data set with the empty set ∅.

TABLE I. Compilation of the $f\sigma_8(z)$ measurements used in this analysis along with the reference matter density parameter Ω_{m_0} (needed for the growth correction) and related references.

z	$f\sigma_8(z)$	$\sigma_{f\sigma_8}(z)$	$\Omega_{m,0}^{\mathrm{ref}}$	Ref.
0.02	0.428	0.0465	0.3	[28]
0.02	0.398	0.065	0.3	[29,30]
0.02	0.314	0.048	0.266	[31,30]
0.10	0.370	0.130	0.3	[32]
0.15	0.490	0.145	0.31	[33]
0.17	0.510	0.060	0.3	[11]
0.18	0.360	0.090	0.27	[34]
0.38	0.440	0.060	0.27	[34]
0.25	0.3512	0.0583	0.25	[35]
0.37	0.4602	0.0378	0.25	[35]
0.32	0.384	0.095	0.274	[36]
0.59	0.488	0.060	0.307115	[37]
0.44	0.413	0.080	0.27	[26]
0.60	0.390	0.063	0.27	[26]
0.73	0.437	0.072	0.27	[26]
0.60	0.550	0.120	0.3	[38]
0.86	0.400	0.110	0.3	[38]
1.40	0.482	0.116	0.27	[39]
0.978	0.379	0.176	0.31	[27]
1.23	0.385	0.099	0.31	[27]
1.526	0.342	0.070	0.31	[27]
1.944	0.364	0.106	0.31	[27]

B. Mock data

An important feature of this work is the comparison of confidence regions for the probability distributions of the internal robustness (iR-PDF). To obtain the confidence regions, we generate mock data sets based on the form

$$f\sigma_8^{\mathrm{mock}}(z_i) = f\sigma_8(z_i|\boldsymbol{\theta}^{\mathrm{bestfit}}) + \mathcal{N}^{\mathrm{random}}, \quad (35)$$

meaning that the mock growth rate data is generated from the best fit parameters $\boldsymbol{\theta}^{\mathrm{bestfit}}$, which are obtained using the complete data set and minimizing the posterior probability (which takes the prior into account). The $\mathcal{N}^{\mathrm{random}}$ term is evaluated by assuming a Gaussian noise with zero mean and standard deviation equal to those given by the data $\sigma_{f\sigma_8}(z_i)$.

The main reason of comparing the results obtained by using the data and the mock catalogues is to compare directly the iR-PDF to the confidence regions. If the iR-PDF from the data falls off the confidence regions, then we can state that either the data set is not internally robust, meaning that the data set could be affected by systematics, or that a better physical model is required in order to better describe the data. In other words, the mock data confidence regions portrait acceptable offset levels from the best fit cosmology obtained from the complete data set.

For each choice of the prior on $\sigma_{8,0}$, we generate 1000 mock data sets. Then, we sample each one of these data sets into 14000 subset combinations, distributed as follows:
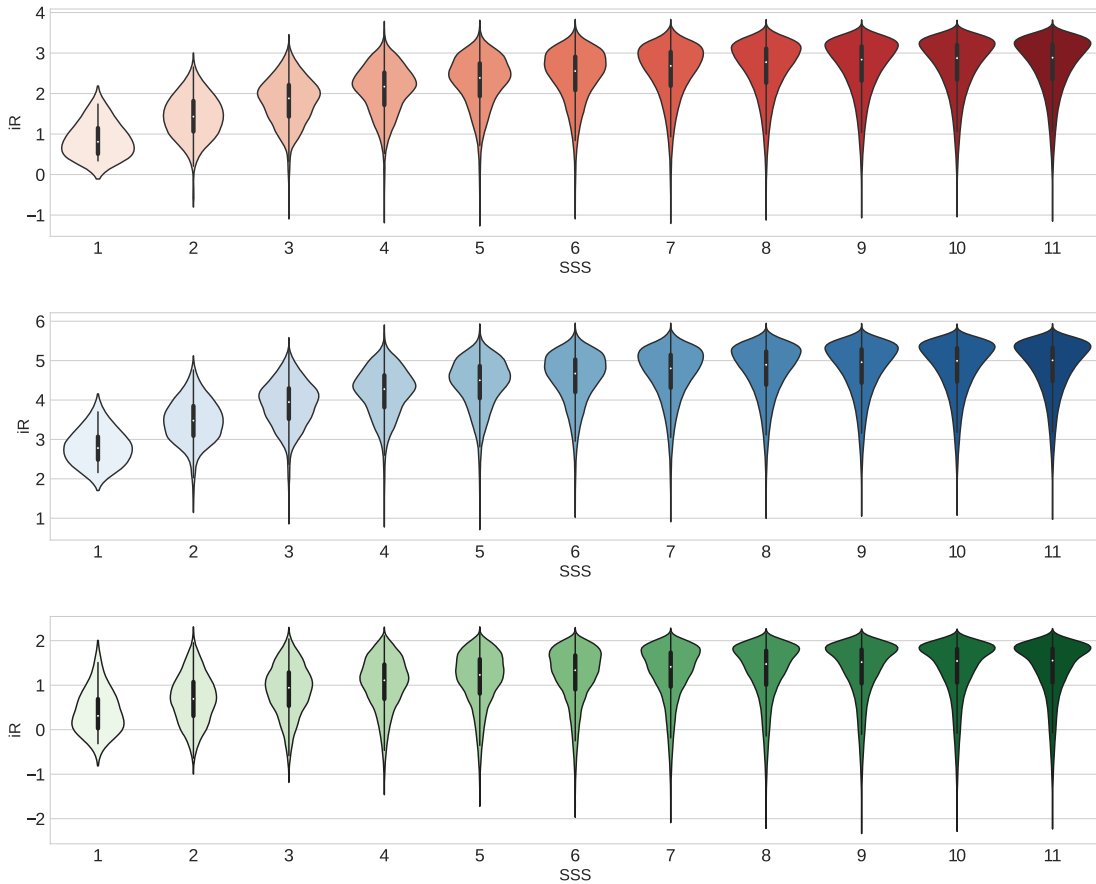
FIG. 1. Violin plots of the internal robustness distributions, for each of the smaller subset sizes of each partition. We employed the narrow flat prior (upper panel), broad flat prior (medium panel), and Gaussian prior (lower panel). The white dots are the mean value of the internal robustness, the bold black line is the 1σ region, and the thinner black line is the 2σ region.

2000 samples for subsets in which the smaller subset size (hereinafter SSS) is 11, another 2000 for subsets with SSS = 10, and so on, until SSS = 4. We stop at SSS = 4 because the number of samples would be larger than the available combinations. As mentioned, the goal is to explore different subset sizes in an equal manner, with the further consideration that, for larger SSS value, we have more possible combinations.

As an ultimate test, we produced mock data sets based on the Planck 2015 best fit parameters [4], for which the parameters are $\Omega_{m,0} = 0.3121$ and $\sigma_{8,0} = 0.815$. The idea behind this is to check whether the tension on measurements of $\sigma_{8,0}$ between cosmic microwave background (CMB) surveys like Planck and galaxy clustering surveys, see Refs. [22,40–47], could be due to inconsistencies in the data themselves.

C. Cross checks

In order to ensure that the method is stable and sensitive to the data set, we decided to opt for two extra cross-checks on our analysis. In brief, the cross-checks have been done using the narrow flat prior only, where we expect the

method to be more sensitive to the final results. The cross-checks are

(i) Data removal: we select one of the combinations with lowest internal robustness and SSS. Then, we remove the data points corresponding to the smaller subset, and evaluate again the complete internal robustness analysis with the new data set. Clearly this procedure forces us to generate a new mock data set with its own best fit. The SSS value will now range from SSS = 4 to the maximum SSS possible. In order to be consistent with the number of points, each SSS will be constituted of 2000 sample subsets.

(ii) Data contamination: we deliberately choose to *contaminate* the first data point of the data set in order to have a worse iR-PDF. This contamination has been implemented by moving the data point by 5σ away from its actual value. In other words, the new first point is constructed as

$$f\sigma_8^{\text{cont}}(z_1) = f\sigma_8(z_1) + 5\sigma_{f\sigma_8}(z_1). \qquad (36)$$

By moving one of the point by 5σ away from its actual position, we expect the iR-PDF to be affected
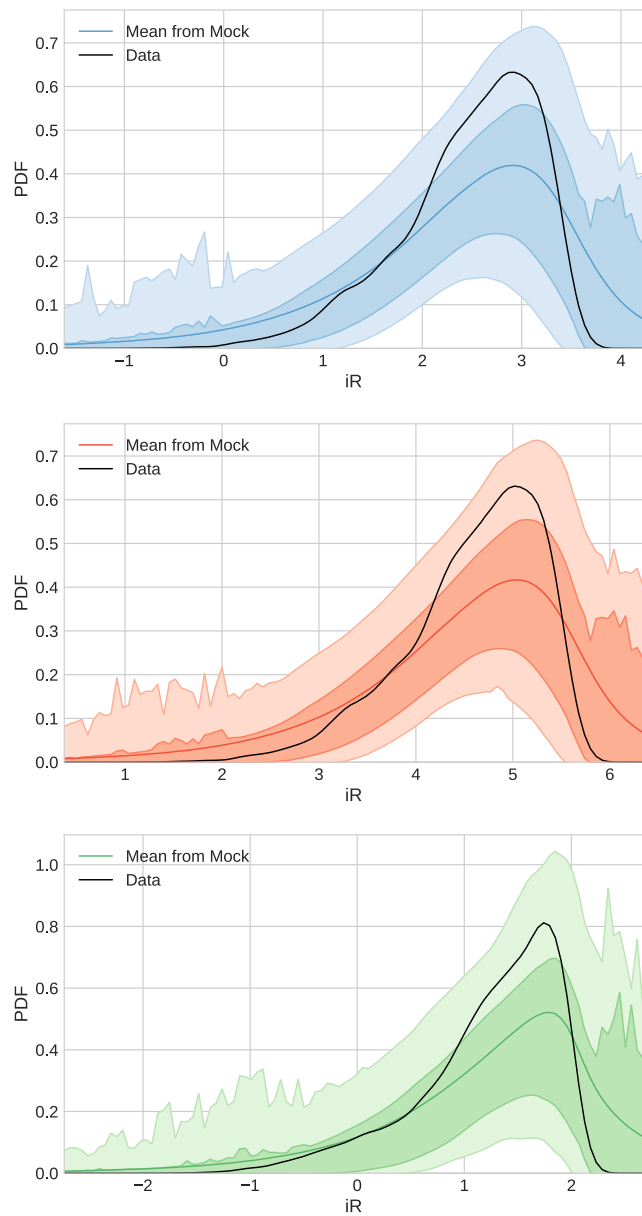
FIG. 2. Internal robustness PDF and confidence regions from mock data based on the best fit cosmology, using each prior. They are: narrow flat prior (upper panel), broad flat prior (medium panel), and Gaussian prior (lower panel).
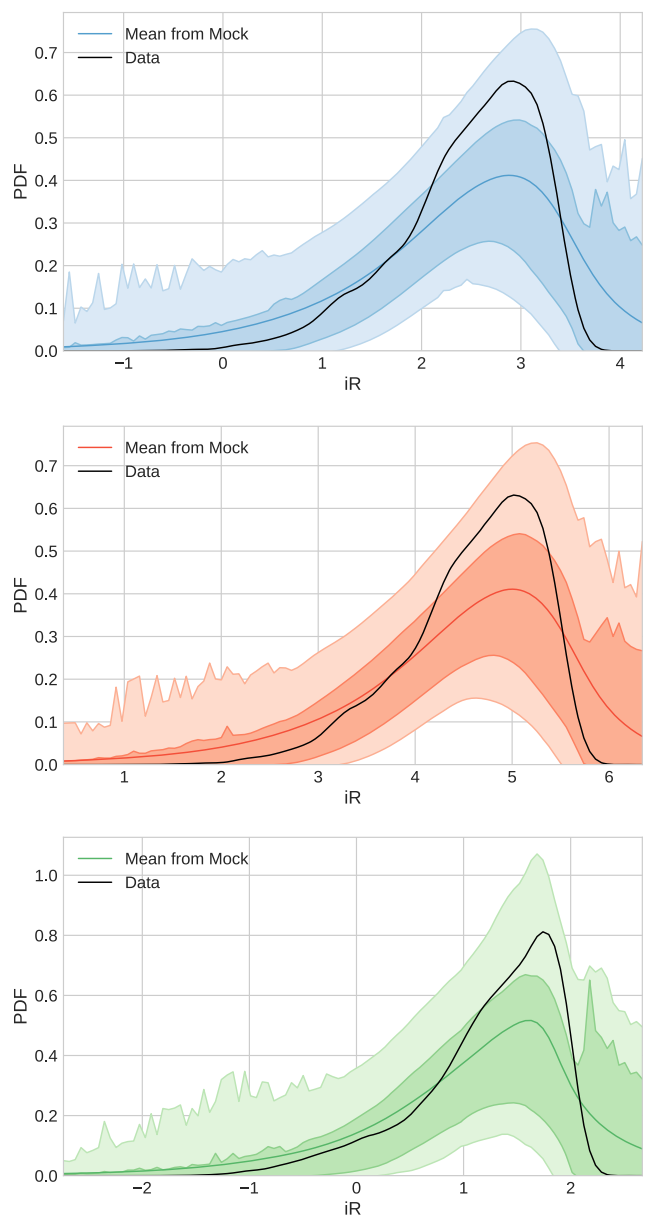
FIG. 3. Same as Fig. 2, but the base parameters for the mock data are from Planck 2015. Priors used: narrow flat prior (upper panel), broad flat prior (medium panel), and Gaussian prior (lower panel).

and fall off the confidence regions, clearly the mean iR has to decrease.

V. RESULTS AND DISCUSSION

The first results are the complete inspection of the data set, comprising all the possible subset combinations. In Fig. 1 we show the iR-PDF in the form of violin plots, arranged by the smaller subset size (SSS) of each subset combination. The three figures differ by the prior used.

From Fig. 1 we can see that the internal robustness increases with the SSS. This results was somehow expected

as a larger data subsets are less prone to manifest outliers, if the data is free of systematic effects. We also see that the broad prior (middle panel in Fig. 1) has much larger iR than the narrow prior (upper panel in Fig. 1). The difference in the iR value is of the order two regardless of the SSS considered. However, the shape of the distributions changes for SSS < 3. We can also see that for the Gaussian prior, the distributions are more stretched for small SSS (SSS < 3) and more clumped up for medium and larger SSS.

In Fig. 2 we show the confidence regions of the mock data sets, as reported in Sec. IV, along with the iR-PDF of the corresponding prior. The data set black lines were
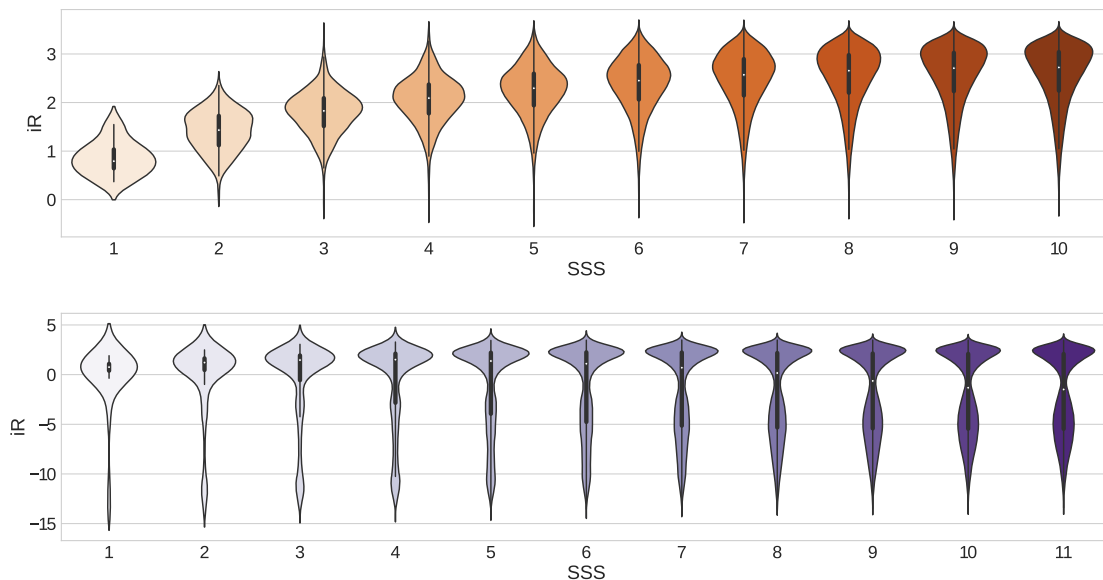
FIG. 4.   Same as Fig. 1, but considering the "r26" data set (upper panel) and the "c1" data set (lower panel), both with a narrow flat prior.

obtained from samples that were equal in size in each SSS as the mock data. We observe that, with the 3 types of prior, the data iR-PDF is within the confidence levels obtained in
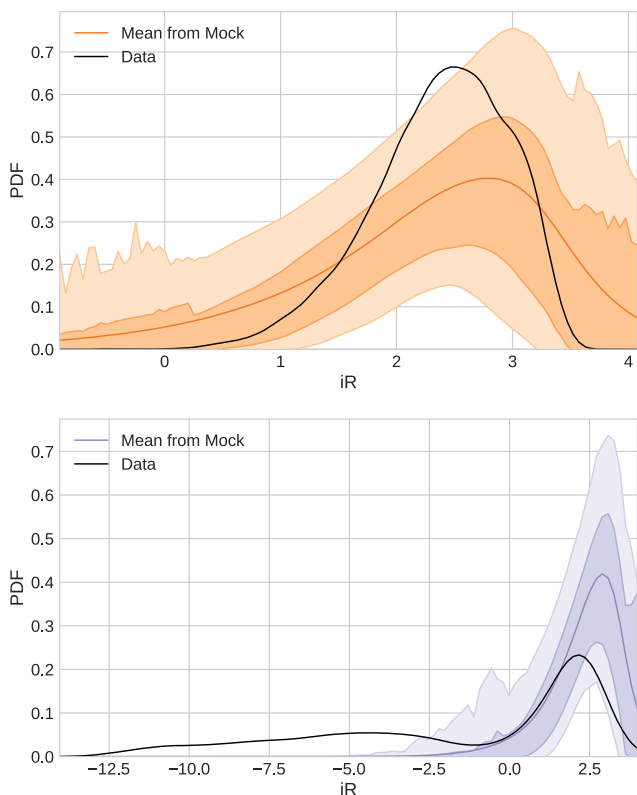

FIG. 5.   Same as Fig. 2, but considering the "r26" data set (upper panel) and the "s1" data set (lower panel), both with a narrow flat prior. The mock data used to generate the confidence regions from the upper panel come from the best fit of the "r26" data set, while the ones for the lower panel are from the original data set.

all the ranges of the internal robustness. This validates the data set, discarding systematic contamination and any other irregularities detectable within the iR formalism.

In Fig. 3 we have plot the confidence regions with the Planck-based cosmology. We see that the confidence regions are nearly identical to those of Fig. 2, with the exception of the Gaussian prior case, where the iR-PDF gets closer to the 1σ region with the Planck-based cosmology mock data. We recall that the Gaussian prior was also chosen based on the Planck 2015 results, so this result is not controversial, although it was not automatically expected, unless we consider the prior to be more constraining than the likelihood alone.

A. Cross-checks results

As mentioned in the previous section, we decided to make a cross-check analysis to ensure that both method and data set gave sensible results. The first check consisted of removing data points from the subset that gave the lowest robustness. In our analysis we found that the lowest SSS that gave a negative lowest robustness was constituted of 2 points (hence SSS = 2) and the data points falling into this subset were the second and sixth data in the Table I. We decided to name this subset "r26."

Our second cross-check was to take the first data point[2] and move $5\sigma s$ away from its actual position. The new dataset is denominated "c1." In Fig. 4 we show the iR-PDFs for the cross-check data sets. We can see from the figure that, for the "r26" data set, the iR for SSS > 6 has a higher minimum but the maximum iR is lower, as well as the mean

---

[2]There is no particular reason of choosing the first point. Since the data set is statistically robust, we are allowed to take randomly any point on the list.

iR values are smaller with respect to the full data set. This is probably due to, when improving a data set by adding robust points, the iR is expected to increase. On the other hand, the improvement in the minimum iR indeed comes from choosing to remove the points with lower iR on the original data set.

For the second cross-check, i.e., the contamination of one data point, we can see immediately that the internal robustness method detects the inconsistency caused by the contaminated data point, by exhibiting a bimodal shape and a decrease of the iR value when we increase the SSS. These are two features that are not proper of a robust data set.

Finally, in Fig. 5 we show the confidence regions for the cross-check data sets. We can see that the confidence regions for the removal cross-check do not fully contain the iR-PDF. The reason is that, by removing some of the data points, the iR for lower SSS is more affected than those with a higher SSS. This is clear if we consider that the effect of dropping data points are more significant for a small data set rather than a large one, assuming they have the similar weights. The anomaly in the iR-PDF for low SSS can be interpreted as the result of an artificial forcing to avoid small iR values.

For the "c1" data set we clearly see that the deviation from the confidence regions from the mock data confirms the efficacy of the methodology presented.

VI. CONCLUSIONS

In this paper we implemented the "internal robustness" of Ref. [1] to the currently available growth-rate data in the form of $f\sigma_8(z)$ shown in Table I, with the aim to examine the data set for systematics and outlier points in a fully automated manner. The internal robustness is a fully Bayesian approach which is not only sensitive to the local minimum like a standard χ^2 comparison, but also to the entire likelihood and can in principle detect the presence of systematics in the data set. The method works by analyzing combinations of subsets in the data set in a Bayesian model comparison way, potentially finding groups of outliers, data affected by systematics or groups that might follow different physics.

The main goal of our approach was to identify systematic-contaminated data points, which can then be further analyzed and potentially excluded if they cannot be corrected. Furthermore, in order to validate our analysis and assess its sensitivity we also performed several cross-checks, for example by removing some of the data points or by artificially contaminating some points, while we also generated mock data sets in order to estimate confidence regions of the iR.

We found that, in the first case, when removing the two points with the least robustness the iR for $SSS > 6$ has a higher minimum but the maximum iR is lower, as well as the mean iR values are smaller with respect to the full data set. In the second case, by adding an artificially contaminated point which was $\sim 5\sigma$ away from its actual value, we found that the internal robustness method indeed detected the inconsistency caused by the contaminated data point.

In conclusion, implementing the iR methodology we found that the $f\sigma_8(z)$ data set, used in our analysis and shown in Table I, is internally robust showing no anomalous behavior, thus ensuring its internal robustness. This is interesting when discussing the tension of the Planck 15 CMB data and the low redshift measurements coming from galaxy surveys, as we can make sure that the discrepancy does not originate from inconsistencies in the data.

[1] L. Amendola, V. Marra, and M. Quartin, Internal robustness: Systematic search for systematic bias in SN Ia data, Mon. Not. R. Astron. Soc. **430,** 1867 (2013).

[2] D. Sapone, Dark energy in practice, Int. J. Mod. Phys. A **25,** 5253 (2010).

[3] S. Tsujikawa, Modified gravity models of dark energy, Lect. Notes Phys. **800,** 99 (2010).

[4] P. A. R. Ade et al., Planck 2015 results. XIII. Cosmological parameters, Astron. Astrophys. **594,** A13 (2016).

[5] M. Kunz and D. Sapone, Crossing the Phantom divide, Phys. Rev. D **74,** 123503 (2006).

[6] S. Nesseris and L. Perivolaropoulos, Crossing the Phantom divide: Theoretical implications and observational status, J. Cosmol. Astropart. Phys. 01 (2007) 018.

[7] L. Amendola et al., Cosmology and fundamental physics with the Euclid satellite, Living Rev. Relativity **16,** 6 (2013).

[8] P. A. Abell et al., LSST science book version 2.0 (unpublished).

[9] A. Aghamousa *et al.*, The DESI experiment part I: Science, targeting, and survey design (unpublished).

[10] N. Kaiser, Clustering in real space and in redshift space, Mon. Not. R. Astron. Soc. **227**, 1 (1987).

[11] Y.-S. Song and W. J. Percival, Reconstructing the history of structure formation using redshift distortions, J. Cosmol. Astropart. Phys. 10 (2009) 004.

[12] L. Kazantzidis and L. Perivolaropoulos, Evolution of the $f\sigma_8$ tension with the Planck15/ΛCDM determination and implications for modified gravity theories, Phys. Rev. D **97**, 103503 (2018).

[13] G. Efstathiou and P. Lemos, Statistical inconsistencies in the KiDS-450 data set, Mon. Not. R. Astron. Soc. **476**, 151 (2018).

[14] C. Heneka, V. Marra, and L. Amendola, Extensive search for systematic bias in supernova Ia data, Mon. Not. R. Astron. Soc. **439**, 1855 (2014).

[15] C. Heneka, A. Posada, V. Marra, and L. Amendola, Searching for bias and correlations in a Bayesian way-Example: SN Ia data, IAU Symp. **10**, 19 (2014).

[16] S. Nesseris and L. Perivolaropoulos, Testing Lambda CDM with the growth function delta(a): Current constraints, Phys. Rev. D **77**, 023504 (2008).

[17] V. Silveira and I. Waga, Decaying Lambda cosmologies and power spectrum, Phys. Rev. D **50**, 4890 (1994).

[18] W. J. Percival, Cosmological structure formation in a homogeneous dark energy background, Astron. Astrophys. **443**, 819 (2005).

[19] A. Bueno Belloso, J. Garcia-Bellido, and D. Sapone, A parametrization of the growth index of matter perturbations in various dark energy models and observational prospects using a Euclid-like survey, J. Cosmol. Astropart. Phys. 10 (2011) 010.

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

[21] S. Dodelson, *Modern Cosmology* (Academic Press, New York, 2003).

[22] S. Nesseris, G. Pantazis, and L. Perivolaropoulos, Tension and constraints on modified gravity parametrizations of $G_{\mathrm{eff}}(z)$ from growth rate and Planck data, Phys. Rev. D **96**, 023542 (2017).

[23] E. Macaulay, I. K. Wehus, and H. K. Eriksen, Lower Growth Rate from Recent Redshift Space Distortion Measurements than Expected from Planck, Phys. Rev. Lett. **111**, 161301 (2013).

[24] S. Basilakos and S. Nesseris, Testing Einstein's gravity and dark energy with growth of matter perturbations: Indications for new physics?, Phys. Rev. D **94**, 123525 (2016).

[25] L. Taddei, M. Martinelli, and L. Amendola, Model-independent constraints on modified gravity from current data and from the Euclid and SKA future surveys, J. Cosmol. Astropart. Phys. 12 (2016) 032.

[26] C. Blake *et al.*, The WiggleZ dark energy survey: Joint measurements of the expansion and growth history at $z < 1$, Mon. Not. R. Astron. Soc. **425**, 405 (2012).

[27] G.-B. Zhao *et al.*, The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic survey DR14 quasar sample: A tomographic measurement of cosmic structure growth and expansion rate based on optimal redshift weights (unpublished).

[28] D. Huterer, D. Shafer, D. Scolnic, and F. Schmidt, Testing LCDM at the lowest redshifts with SN Ia and galaxy velocities, J. Cosmol. Astropart. Phys. 05 (2017) 015.

[29] S. J. Turnbull, M. J. Hudson, H. A. Feldman, M. Hicken, R. P. Kirshner, and R. Watkins, Cosmic flows in the nearby universe from Type Ia Supernovae, Mon. Not. R. Astron. Soc. **420**, 447 (2012).

[30] M. J. Hudson and S. J. Turnbull, The growth rate of cosmic structure from peculiar velocities at low and high redshifts, Astrophys. J. **751**, L30 (2012).

[31] M. Davis, A. Nusser, K. Masters, C. Springob, J. P. Huchra, and G. Lemson, Local gravity versus local velocity: Solutions for β and nonlinear bias, Mon. Not. R. Astron. Soc. **413**, 2906 (2011).

[32] M. Feix, A. Nusser, and E. Branchini, Growth Rate of Cosmological Perturbations at z 0.1 from a New Observational Test, Phys. Rev. Lett. **115**, 011301 (2015).

[33] C. Howlett, A. Ross, L. Samushia, W. Percival, and M. Manera, The clustering of the SDSS main galaxy sample, Mock galaxy catalogues and a measurement of the growth of structure from redshift space distortions at $z = 0.15$, Mon. Not. R. Astron. Soc. **449**, 848 (2015).

[34] C. Blake *et al.*, Galaxy and mass assembly (GAMA): Improved cosmic growth measurements using multiple tracers of large-scale structure, Mon. Not. R. Astron. Soc. **436**, 3089 (2013).

[35] L. Samushia, W. J. Percival, and A. Raccanelli, Interpreting large-scale redshift-space distortion measurements, Mon. Not. R. Astron. Soc. **420**, 2102 (2012).

[36] A. G. Sanchez *et al.*, The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: Cosmological implications of the full shape of the clustering wedges in the data release 10 and 11 galaxy samples, Mon. Not. R. Astron. Soc. **440**, 2692 (2014).

[37] C.-H. Chuang *et al.*, The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: Single-probe measurements from CMASS anisotropic galaxy clustering, Mon. Not. R. Astron. Soc. **461**, 3781 (2016).

[38] A. Pezzotta *et al.*, The VIMOS public extragalactic redshift survey (VIPERS): The growth of structures at $0.5 < z < 1.2$ from redshift-space distortions in the clustering of the PDR-2 final sample, Astron. Astrophys. **604**, A33 (2017).

[39] T. Okumura *et al.*, The Subaru FMOS galaxy redshift survey (FastSound). IV. New constraint on gravity theory from redshift space distortions at $z \sim 1.4$, Publ. Astron. Soc. Jpn. **68**, 38 (2016).

[40] S. Camera, M. Martinelli, and D. Bertacca, Easing tensions with quartessence (unpublished).

[41] F. Kohlinger *et al.* KiDS-450: The tomographic weak lensing power spectrum and constraints on cosmological parameters, Mon. Not. R. Astron. Soc. **471**, 4412 (2017).

[42] T. M. C. Abbott *et al.*, Dark energy survey year 1 results: Cosmological constraints from galaxy clustering and weak lensing, Phys. Rev. D **98**, 043526 (2018).

[43] A. Gomez-Valent and J. Sola, Relaxing the σ_8-tension through running vacuum in the Universe, Europhys. Lett. **120**, 39001 (2017).

[44] A. Gomez-Valent and J. Sola, Density perturbations for running vacuum: A successful approach to structure

formation and to the σ_8-tension, Mon. Not. R. Astron. Soc. **478**, 126 (2018).

[45] B. J. Barros, L. Amendola, T. Barreiro, and N. J. Nunes, Coupled quintessence with a ΛCDM background: Removing the σ_8 tension (unpublished).

[46] G. Lambiase, S. Mohanty, A. Narang, and P. Parashari, Testing dark energy models in the light of σ_8 tension (unpublished).

[47] R. Gannouji and D. Polarski, On the consistency of the expansion with the perturbations (unpublished).