

Alma Mater Studiorum Università di Bologna
Archivio istituzionale della ricerca

An improved model-independent assessment of the late-time cosmic expansion

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

Published Version:

An improved model-independent assessment of the late-time cosmic expansion / Moresco, M.. - In: JOURNAL OF COSMOLOGY AND ASTROPARTICLE PHYSICS. - ISSN 1475-7516. - STAMPA. - 2018:10(2018), pp. 15.1-15.36. [10.1088/1475-7516/2018/10/015]

Availability:

This version is available at: <https://hdl.handle.net/11585/676067> since: 2019-02-27

Published:

DOI: <http://doi.org/10.1088/1475-7516/2018/10/015>

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).
When citing, please refer to the published version.

(Article begins on next page)

This is the final peer-reviewed accepted manuscript of:

Balakrishna S. Haridasu et al., *An improved model-independent assessment of the late-time cosmic expansion*, JCAP 10(2018)015.

The final published version is available at: <https://doi.org/10.1088/1475-7516/2018/10/015>

Rights / License:

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>)

When citing, please refer to the published version.

An improved model-independent assessment of the late-time cosmic expansion

Balakrishna S. Haridasu^{a,b,c} Vladimir V. Luković^{b,c} Michele Moresco^{d,e} Nicola Vittorio^{b,c}

^aGran Sasso Science Institute , Viale Francesco Crispi 7, I-67100 L'Aquila, Italy

^bDipartimento di Fisica, Università di Roma "Tor Vergata", Via della Ricerca Scientifica 1, I-00133, Roma, Italy

^cSezione INFN, Università di Roma "Tor Vergata", Via della Ricerca Scientifica 1, I-00133, Roma, Italy

^dDipartimento di Fisica e Astronomia, Università di Bologna, Via Gobetti 93/2, I-40129, Bologna, Italy

^eINAF - Osservatorio Astronomico di Bologna, Via Gobetti 93/3, I-40129, Bologna, Italy

E-mail: sandeep.haridasu@gssi.it, vladimir.lukovic@roma2.infn.it,
michele.moresco@unibo.it, nicola.vittorio@roma2.infn.it

Abstract. In the current work, we have implemented an extension of the standard Gaussian Process formalism, namely the Multi-Task Gaussian Process with the ability to perform a joint learning of several cosmological data simultaneously. We have utilised the “low-redshift” expansion rate data from Supernovae Type-Ia (SN), Baryon Acoustic Oscillations (BAO) and Cosmic Chronometers (CC) data in a joint analysis. We have tested several possible models of covariance functions and find very consistent estimates for cosmologically relevant parameters. In the current formalism, we also find provisions for heuristic arguments which allow us to select the best-suited kernel for the reconstruction of expansion rate data. We also utilised our method to account for systematics in CC data and find an estimate of $H_0 = 68.52^{+0.94+2.51(\text{sys})}_{-0.94}$ km/s Mpc⁻¹ and a corresponding $r_d = 145.61^{+2.82}_{-2.82-4.3(\text{sys})}$ Mpc as our primary result. Subsequently, we find constraints on the present deceleration parameter $q_0 = -0.52 \pm 0.06$ and the transition redshift $z_T = 0.64^{+0.12}_{-0.09}$. All the estimated cosmological parameters are found to be in good agreement with the standard Λ CDM scenario. Including the local model-independent H_0 estimate to the analysis we find $H_0 = 71.40^{+0.30+1.65(\text{sys})}_{-0.30}$ km/s Mpc⁻¹ and the corresponding $r_d = 141.29^{+1.31}_{-1.31-2.63(\text{sys})}$ Mpc. Also, the constraints on $r_d H_0$ remain consistent throughout our analysis and also with the model-dependent CMB estimate. Using the $Om(z)$ diagnostic, we find that the concordance model is very consistent within the redshift range $z \lesssim 2$ and mildly discrepant for $z \gtrsim 2$.

Contents

1	Introduction	1
2	Method	3
2.1	Single-Task Gaussian Process	3
2.2	Multi-Task Gaussian Process	5
3	Theory and Data	10
3.1	Theoretical framework	10
3.2	Data and Implementation	11
3.3	Kernel selection	14
4	Results and Discussion	15
4.1	Reconstruction of data and cosmological parameters	15
4.2	Reconstruction of Diagnostics	21
4.3	Impact of assumptions in data and method	23
5	Conclusions and Future prospects	28

1 Introduction

The current era of precision cosmology has allowed us to constrain cosmological models to a very high degree of precision [1–4]. However lacking explanations for several physical aspects within these models [5], this creates a need for alternative approaches for analysing the data and making cosmological inferences. Such a need and the ease to conduct phenomenological analysis has led to the use of several model-independent methods, in the last decade.

Amongst several model-independent approaches, Principal Component Analysis [6–11], smoothing methods [12–14], cosmographic and polynomial regression methods [15–19], Gaussian process (GP) formalism [20–22] etc., are a few. See [23] for an overview of some of these methods. GP has been a popular “model-independent” method that has been utilised in the past years to reconstruct the low-redshift cosmic expansion history, henceforth the dynamics of the late-time evolution [21, 24]. GP has been implemented to draw inferences on acceleration [25, 26], estimates of H_0 [18, 27–29], curvature of the universe [30], to perform model falsification by estimating several diagnostics [29, 31–33], and to study the dynamics of the Dark Energy (DE) / scalar field equation of state (EoS) [21, 33–35]. More recently, it has also been implemented to study model selection in [26, 36, 37] and estimation of linear anisotropic stress in [38].

In this work we implement an extended GP formalism, which allows one to perform a Joint-analysis of “low-redshift” observations such as Supernova Type Ia (SN) [4], Cosmic chronometers (CC) [39, 40] and the anisotropic Baryon Acoustic Oscillations (BAO) [41] datasets. To conduct a joint analysis using the cosmological data in an independent/single GP often requires assumptions of the several cosmological parameters such as present expansion rate (H_0), sound horizon at drag epoch (r_d) etc. Such assumptions, in turn, make the overall analysis and thereby any inferences mildly model-dependent (if not significantly). In this respect, we implement a model-independent joint learning method called Multi-Task Gaussian Process (MTGP) [42–44], which provides a provision to model the covariance between different tasks (essentially datasets), thereby allowing various tasks

to complement each other and hence improve the overall reconstruction/prediction for each of the datasets, in comparison to the independent GPs of individual tasks. In particular for a cosmological dataset, one can make better predictions in the redshift range with less significant observations, by complementing the primary dataset with the secondary dataset(s) which provide better measurements in that range of redshifts. Apart from providing improved predictions, MTGP also allows one to infer the parameters (e.g., r_d) from the reconstructions, in contrast to assuming them a priori.

The tension in the local direct estimate of $H_0 = (73.48 \pm 1.66)$ km/s Mpc $^{-1}$ (hereafter R18) in [45] and indirect estimate of $H_0 = (66.93 \pm 0.62)$ km/s Mpc $^{-1}$ (hereafter P16) from the “high-redshift” Planck collaboration [1] (See Table 8 therein) has been a well established problem of precision cosmology. GP formalism has been implemented many a time to estimate H_0 as an intercept ($z = 0$) of the extrapolated predictive reconstruction of CC data in several works [27–29]. [29], have reported $H_0 = 67.42 \pm 4.80$ km/s Mpc $^{-1}$, preferring a lower H_0 value in comparison to R18. More recently, [18] have combined the expansion rate data ($E(z)$) from supernova [4] and CC data to obtain $H_0 = 67.06 \pm 1.68$ km/s Mpc $^{-1}$. These estimates are clearly very consistent with “high-redshift” P16 and retain the well-known tension with R18. As is also mentioned in [18], the GP formalism suffers from the caveat of having to choose the specific covariance function, essentially a prior (*of sorts*) on the reconstruction. In turn, this choice could lead to a bias in the estimate of H_0 . Correcting for this effect would require better ways to model the GP, such as, learning the covariance function instead of assuming a fixed model, or to implement a joint analysis of different datasets. While the earlier option is a statistically rigorous procedure (see e.g., [46]), it is not suitable for the current compilation of limited cosmological data. On the contrary, there lie prospects in the latter option for the same, which we implement here (partly due to the newer $E(z)$ data). It is also important that the joint analysis is complemented with a proper kernel selection criteria to overcome any possible biases/specificities of the assumed covariance functions. However, we anticipate that given the smooth nature of the cosmic expansion rate data, one might not require a kernel selection but on the contrary find very consistent results with each of the kernels implemented.

In this work, we focus on the estimation of constant parameters H_0 , r_d from the MTGP formalism for several suitable combinations of data. As the GP formalism is also able to predict derivatives of the reconstruction regions, evaluation of deceleration parameter ($q(z)$) [20, 25] and estimation of deceleration/acceleration transition redshift (z_T) [47, 48] have been of key interest. The SN distance data was able to provide a reliable reconstruction of $q(z)$ up to $z \sim 0.8$ (see e.g., [20]), however, the CC data is unable to do the same unless otherwise a prior on H_0 is assumed [25]. In [29], it has been reported that no good estimate of z_T is available from the CC data, which we try to revise in this work. We also reconstruct relevant cosmological diagnostics to comment on their estimation and hence the inferences. In particular we reconstruct the $Om(z)$ diagnostic [32, 49, 50] and its derivative. These diagnostics and their variations have been implemented on several occasions to assess the validity of Λ CDM [49]. The CC dataset usually implemented in the literature are derived based on an assumption of a particular model for the stellar evolution [40, 51, 52]. We discuss the differences in the H_0 estimates and hence the systematics, obtained using the alternate data (this issue was also addressed in [18]). We also illustrate several subtleties of the newer method utilised in this work by implementing it with several combinations of the data.

The paper is organised as follows: In Section 2 we describe the newer MTGP method, in Section 3 we then briefly present the datasets, corresponding modelling and also introduce the theoretical framework used to derive inferences. We have also included a brief discussion on the issue of kernel selection in this section. In Section 4 we present results for cosmological inferences from our analysis, along with a discussion on systematic effects and finally summarise our conclusions in Section 5.

2 Method

In this section, we describe the basic formalism of the Multi-Task Gaussian Process (hereafter MTGP), that we have implemented in this work. As the name suggests, MTGP considers different tasks, which are essentially different sets of independent training data and their corresponding cross-covariance to perform a simultaneous learning, which in-turn is utilised to make predictions. We begin by introducing the standard Single-Task Gaussian Process and then proceed to the newer MTGP formalism implemented here.

2.1 Single-Task Gaussian Process

We introduce the individual/Single-Task Gaussian Process (GP), with a brief description. However, we keep the discussion in this subsection to a minimum as it has been introduced numerous times in the cosmological context (see e.g., [20–22])¹. GP is essentially a generalisation of the single Gaussian distribution to the probability distributions of a function over a range of independent variables. In principle, this can be any stochastic process but is way simpler in a gaussian scenario and is also more often the case. While GP can be implemented for both *classification* and *regression* problems, for the cosmological concerns in this paper we focus on *regression*. Given a task to be learnt for a set of n independent variables $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$ and the corresponding dependent outputs $\mathbf{y} = \{y_1, y_2, \dots, y_n\}$, GP provides predictions \mathbf{y}^* at the target inputs \mathbf{x}^* by placing a Gaussian prior on the latent function $f(x)$ that maps \mathbf{x} onto \mathbf{y} . It is in this context that GP is also deemed as the collection of random variables and can be represented as,

$$f(x_1) = \mathcal{GP}(\mu(x_1), \text{cov}[f(x_1), f(x_1)]) \quad (2.1)$$

where $\mu(x_1)$, and $\text{cov}[f(x_1), f(x_1)]$ are the mean and the variance of the random variable at x_1 , respectively. The prior assumption on $\mu(x)$ is usually assumed to be 0, and does not play any role in obtaining the reconstructions, as the constant(deterministic) $\mu(x)$ only remains an additive factor for the final predictions. However, it is not the case if one wants to assume an explicit functional form for $\mu(x)$ with additional parameters [55, 56] and would no longer remain a minimal assumption analysis as the nature of this functional form is now imposed on the reconstructions. Also, note that far away from the regions in which data are available, the reconstructed posterior mean tends to the prior mean. However, we will comment more in detail on the assumption of mean towards the end of this section. Given that one needs to reconstruct the function $f(x)$, we model the covariance between values of this function at different positions as,

$$\text{cov}[f(x), f(x')] = k(x, x') \quad (2.2)$$

where k is an a priori assumed covariance model (*kernel*, in GP terminology) for the particular reconstruction and is appropriately replaced in Eq. (2.1). We use \mathbf{x} to represent a set of inputs values and x for an individual input. Correspondingly, $k(x, x')$ is a single value and $K(\mathbf{x}, \mathbf{x}')$ represents a matrix of dimension $n \times n$, where n is the length of the input vector. In the real-world scenarios often-times the dependent variables are only obtained with uncertainties (*noise*, in the GP terminology) or even possible covariances (Σ) amongst different measurements. In such cases, the learning is performed with an additional noise covariance term Σ along with the modelled $K(\mathbf{x}, \mathbf{x}')$. The standard covariance functions are usually stationary models, where the estimate of the covariance is dependent only on

¹However, for an in-depth discussion, we strongly encourage the reader to refer to [53]. See [54] for a discussion of the Gaussian process with noisy inputs, which is more relevant for cosmological purposes.

the difference between the positions $|x - x'|$ and not on the actual positions. These models for covariance functions are often referred to as stationary kernels. In this work we implement the standard Gaussian/Squared-Exponential (SE) and the Matérn (M_ν) class of kernels. The SE kernel is defined as,

$$k_{SE}(x, x') = \sigma_f^2 \exp\left(-\frac{|x - x'|^2}{2l^2}\right), \quad (2.3)$$

where σ_f is the function/signal variance (amplitude) and l is the length scale that dictates the ability to model features in the predicted region. These two parameters are often called hyperparameters as they are not the parameters of the function but of the covariance function. In the later descriptions to follow we redefine $\tau = |x - x'|$, and explicitly write $|x - x'|$ only when necessary, as this is consistent with all the kernels implemented here, which are stationary in nature. The SE kernel, however, is a very smooth covariance function and so oftentimes not regarded as the best assumption for reconstructing features. To this aid, the Matérn class of kernels are very useful and the general functional form can be written as,

$$k_{M_\nu}(\tau) = \sigma_f^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\sqrt{2\nu}\tau}{l}\right)^\nu K_\nu\left(\frac{\sqrt{2\nu}\tau}{l}\right), \quad (2.4)$$

where K_ν is the modified Bessel function of second kind with ν is strictly a positive parameter and $\Gamma(\nu)$ is the standard Gamma function. $\sigma_f, l > 0$ carry the usual definitions. The modified Bessel function provides an explicit analytic functional form for half-integer values of $\{\nu = 1/2, 3/2, 5/2, \dots\}$ and as $\nu \rightarrow \infty$ the M_ν covariance function converges to the SE kernel. Amongst the several possibilities $\nu = 9/2$ and $\nu = 7/2$ are of particular interest as they are neither too smooth nor do they predict rapid variations¹. We would also like to assert that for the current cosmological data available, the flexibility of the Matérn class kernels is sufficient and there is no immediate need for more complicated kernels, which will become relevant with forthcoming stringent data and for extrapolated predictions. For instance, the Matérn covariance function for $\nu = 9/2$ is given as,

$$k_{M_{9/2}}(\tau) = \sigma_f^2 \exp\left(-\frac{3\tau}{l}\right) \left[1 + \frac{3\tau}{l} + \frac{27\tau^2}{7l^2} + \frac{18\tau^3}{7l^3} + \frac{27\tau^4}{35l^4}\right], \quad (2.5)$$

where $l > 0$. Also, it is worth mentioning that the Matérn kernels are differentiable only to the order $\lfloor \nu \rfloor$ and so the kernels with $\nu < 5/2$ lose their interest for predicting the derivatives¹ and hence relevant cosmological features, however still relevant for predictions at a single target position. Finally, the hyperparameters $\Theta \equiv \{\sigma_f, l\}$ are learned by optimising the log marginal likelihood (see [53] for details) that is defined as,

$$\mathcal{L}(\Theta) = -\frac{1}{2} \mathbf{y}^T K_y^{-1} \mathbf{y} - \frac{1}{2} \ln |K_y| + \frac{n}{2} \ln(2\pi) \quad (2.6)$$

where $K_y = K(\mathbf{x}, \mathbf{x}') + \Sigma_{n \times n}$ for a set of n observations and assuming mean $\mu = 0$. The reconstructions of the regions can be performed by either learning the hyperparameters from a simple optimisation (global) of Eq. (2.6) or by marginalising over them in a Bayesian way. In fact, the latter procedure often puts GP in the light of a non-parametric method². We emphasise that in the context of cosmology, model-independent implies independent of cosmological model, and the assumptions of kernel

¹Please see [21] for a good discussion on the relevancy of these issues for the cosmological data. It has also been argued that reconstructions using Matérn kernels with $\nu \geq 7/2$ cannot be distinguished for finite and noisy training samples (see Sec.4.2 of [53] and references therein).

²In a similar line of reasoning, as we are modelling the covariance but not the latent function ($f(x)$, which is marginalised upon) itself, GP is also deemed as model-independent. In a more recent work [18], discussion on similar issues was presented.

models must however be subjected to proper kernel selection to estimate which of them provides a better description of the data (see Section 3.3). However, it has been argued [53](also in [21]), that optimisation provides equally good predictions, while being computationally lighter. The differences in optimisation *vs.* marginalisation becomes significant when the data are too sparse or if they are too noisy. The latter is, in fact, true with the CC data, as was found in [18]. However, if the data are well constrained the probability distributions of the hyperparameters will be similarly peaked³. In this work we perform global optimisation within a very large prior region ($0 < \sigma_f, l < 10$) for either of the hyperparameters (see Section 3.2). Implementing the optimisation method, once the hyperparameters are learned, one can predict the mean and variance of the latent function at the test/target points through,

$$\begin{aligned}\tilde{f}(\mathbf{x}^*) &= K(\mathbf{x}^*, \mathbf{x})K_y^{-1}\mathbf{y} \\ \text{cov}[f(\mathbf{x}^*)] &= K(\mathbf{x}^*, \mathbf{x}^*) - K(\mathbf{x}^*, \mathbf{x})K_y^{-1}K(\mathbf{x}, \mathbf{x}^*).\end{aligned}\tag{2.7}$$

One can extend the GP predictions also to the derivatives of the latent functions, however limited by the differentiability of an assumed kernel function. As is shown in [58], derivative of a certain GP would also be a Gaussian Process, in fact even an integration or combinations of different orders of derivatives can be approximated as Gaussians within the GP formalism. As described in [31](see also [54]), this formalism can be useful to predict relevant dynamics in cosmology. The covariances involving the derivatives are written as,

$$\begin{aligned}\text{cov}[f(x_1), f'(x_2)] &= \frac{\partial k(x_1, x_2)}{\partial x_2} \\ \text{cov}[f'(x_1), f'(x_2)] &= \frac{\partial^2 k(x_1, x_2)}{\partial x_1 \partial x_2}\end{aligned}\tag{2.8}$$

where f' represent the derivatives with respect to their corresponding independent variables. Once the covariance between the derivative(s) of the latent function are defined we can obtain the mean and variance for the predictions of function derivatives. The mean of the i^{th} derivative and the covariance between i^{th} and j^{th} derivatives can be written as,

$$\begin{aligned}\tilde{f}^{(i)}(\mathbf{x}^*) &= K^{(i)}(\mathbf{x}^*, \mathbf{x})K_y^{-1}\mathbf{y} \\ \text{cov}[f^{(i)}(\mathbf{x}^*), f^{(j)}(\mathbf{x}^*)] &= K^{(i,j)}(\mathbf{x}^*, \mathbf{x}^*) - K^{(i)}(\mathbf{x}^*, \mathbf{x})K_y^{-1}K^{(j)}(\mathbf{x}, \mathbf{x}^*),\end{aligned}\tag{2.9}$$

respectively. For $i = j$ latter of Eq. (2.9) provides the variance of the i^{th} derivative. If the data for derivative functions is available, one can in fact perform a joint analysis, which was by far the only way [59], however requiring to assume physical constants in the cosmological context. Subsequently, one can extend this formalism to obtain predictions for combinations of $f(x)$ and its derivatives ($f'(x), f''(x), \dots$). The prediction at a given target (x^*) for the combination is sampled from a multivariate normal distribution of the necessary components, appropriately considering the covariance amongst them. This is especially crucial for derivatives of higher order with larger uncertainties, where simple propagation of errors will not fare well.

2.2 Multi-Task Gaussian Process

While the Single-Task GP models the covariance between the values of the same task at different positions, the MTGP allows one perform a simultaneous learning of multiple tasks by modelling the

³Although the log marginal likelihood would be peaked at the global maximum, it may also be the case that there exist several local maxima. Appropriate care must be implemented to avoid these local maxima (see Sec 5.4 of [57]). We also provide more discussion on the same in the Section 3.3

covariance between different tasks at their respective positions [43, 44, 60, 61]⁴. For this purpose, similar to Eq. (2.2) one can write,

$$\text{cov}[f_l(\mathbf{x}_l), f_m(\mathbf{x}_m)] = k_{lm}(\mathbf{x}_l, \mathbf{x}_m), \quad (2.10)$$

where $f_l(\mathbf{x}_l)$, $f_m(\mathbf{x}_m)$ represent the latent functions of the l^{th} and the m^{th} task at their corresponding input variables, respectively. The covariance between tasks l and m is modelled by k_{lm} . In this work we implement a more general Multi-Kernel Gaussian process, where the formalism has been developed so one can utilise different kernels for different tasks, unlike in the standard Multi-Task scenario which only allows for single kernel across different tasks. Therefore, one can choose to use $M_{9/2}$ and SE kernels to model two tasks and their covariance to perform a joint learning. Consider two sets of data $(\mathbf{x}_l, \mathbf{y}_l) = \{(x_{l_1}, y_{l_1}), (x_{l_2}, y_{l_2}), \dots, (x_{l_{n_l}}, y_{l_{n_l}})\}$ and $(\mathbf{x}_m, \mathbf{y}_m) = \{(x_{m_1}, y_{m_1}), (x_{m_2}, y_{m_2}), \dots, (x_{m_{n_m}}, y_{m_{n_m}})\}$, with n_l and n_m being the number of data points in each set, respectively. The current formalism is based on the fact that the kernel function can be written as a convolution of the underlying basis function⁵, evaluated at the two positions between which the covariance needs to be modelled. Following [60], one can write the kernel function as,

$$k_{ll}(x, x') \equiv g_l(x_l) \otimes g_l(x'_l) = \int_{-\infty}^{+\infty} g_l(u - x_l) g_l(x'_l - u) du. \quad (2.11)$$

where $g_l(x)$ is the basis function of the kernel (k_{ll}) defining the auto-covariance between two different data points of the same task. Given that one needs to model the cross-covariance which can be written as the convolution of the basis functions of the two different kernels for two individual tasks,

$$k_{lm}(x, x') \equiv g_l(x_l) \otimes g_m(x_m) = \int_{-\infty}^{+\infty} g_l(u - x_l) g_m(x_m - u) du. \quad (2.12)$$

Here k_{lm} denotes the cross-covariance between two different tasks l, m and g_l, g_m represent the basis functions of the two individual kernels, x_l, x_m represent the independent variables of their respective tasks. Given the above formalism, the method to obtain the basis functions has been derived in [60]⁶. The basis function of a given kernel $k(\tau)$ can be written as,

$$g(\tau) = \frac{1}{\sqrt[4]{2\pi}} \mathcal{F}_{s \rightarrow \tau}^{-1} \left[\sqrt{\mathcal{F}_{\tau \rightarrow s} [k(\tau)]} \right], \quad (2.13)$$

where \mathcal{F}^{-1} , \mathcal{F} represent the Inverse Fourier and Fourier transform, respectively. Once the basis functions are evaluated one can write the cross-covariance following Eq. (2.12). Finally, we model the auto-covariance for each of the dataset and the corresponding cross-covariance between different datasets as,

$$\text{cov}[f_l(x_l), f_m(x'_m)] = k_{lm}(x_l, x'_m), \quad (2.14)$$

where $l, m \in 1, 2, \dots, N$ for N number of datasets. This formalism ensures that we obtain a positive semi-definite matrix for the modelled covariance and allows one to perform a joint learning on different

⁴For an historical overview of Multi-Task Gaussian Process, please refer to [42, 53]

⁵The underlying basis function(s) of a kernel models how the strength of covariance/correlation between two inputs decays with increasing separation (τ) between them, see [53] for more details on basis functions.

⁶See also the technical report (ACFR – TR – 2011 – 002) with detailed derivations and explanations, available at <http://www.acfr.usyd.edu.au/techreports/>.

datasets/tasks which can now be incorporated in the joint GP. Given the covariance between the individual tasks, we write the joint covariance matrix as,

$$K_y = \begin{bmatrix} K_{ll} & K_{lm} \\ K_{ml} & K_{mm} \end{bmatrix} + \begin{bmatrix} \Sigma_{n_l \times n_l} & 0 \\ 0 & \Sigma_{n_m \times n_m} \end{bmatrix} \quad (2.15)$$

where the cross-covariance between different datasets is modelled by K_{ml} and $\Sigma_{n_i \times n_i}$ represents the corresponding noise observed on the data, with n_i being the number of data points in task i . If different tasks are correlated as well, the non-diagonal terms in Eq. (2.15) will be non-zero. This formalism can be easily extended to many datasets and the learning can be performed by optimising Eq. (2.6), appropriately replacing K_y and the data $y \equiv \{y_1, y_2, \dots, y_N\}$ for the joint analysis of N tasks. Note that also the reconstruction of the predictive regions is performed taking into account the total covariance in Eq. (2.15). Finally, as an example, we show the cross-covariance terms evaluated for the SE and $M_{9/2}$ kernels. The basis functions of the SE and $M_{9/2}$ kernels can be analytically evaluated and are written as,

$$\begin{aligned} g_{SE}(x) &= \sigma_{SE} \left(\frac{2}{\pi l_{SE}^2} \right)^{1/4} \exp \left(-\frac{x^2}{l_{SE}^2} \right), \\ g_{M_{9/2}}(x) &= \frac{48\sigma_M}{\pi} \sqrt{\frac{3}{35l_M^5}} x^2 K_2 \left(\frac{2|x|}{l_M} \right), \end{aligned} \quad (2.16)$$

where σ_i, l_i have the usual meanings and K_2 represents the modified Bessel function (with $\nu = 2$). Note that the basis function of the $M_{9/2}$ is now given in terms of K_2 , which cannot be explicitly written in an analytical form and hence all further operations should be carried out numerically. With the basis functions defined, the cross-covariance can be evaluated by performing the convolution as in Eq. (2.12). The cross-covariance between two SE kernels with different hyperparameters can be analytically obtained as,

$$k_{SE_1 \times SE_2}(\tau) = \sigma_1 \sigma_2 \sqrt{\frac{2l_1 l_2}{l_1^2 + l_2^2}} \exp \left(-\frac{\tau^2}{l_1^2 + l_2^2} \right). \quad (2.17)$$

where indices 1,2 correspond to the first and the second SE kernels. Note that the $k_{SE_1 \times SE_2}(\tau)$ reduces to the standard k_{SE} (Eq. (2.3)), if hyperparameters of the two kernels are same. Although, we do not have an analytical functional form for the cross-covariance involving Matérn class kernels, they can be evaluated by numerically performing the convolution. In Figure 1 we provide some illustrative comparison of the auto-covariance and cross-covariance functions with several combinations of hyperparameters. In the *left* panel we show the standard kernel auto-covariance as a function of $\tau = |x - x'|$. It can be clearly noticed that the Matérn class kernels provide lesser correlation than the SE kernel for smaller τ and vice-versa. This in turn allows M_ν kernels to provide more features in the predictive region as the value ν is reduced. In the *right* panel we show the cross-covariance between same kernels with different length scale hyperparameter for each. Notice that in this case the covariance between $k_{M_{\nu_1} \times M_{\nu_2}}(\tau)$ is greater for smaller ν_1 and ν_2 , at lower values of τ . This plays a very important role in the determination of posterior at a particular value of x^* for different tasks in the MTGP formalism. Also, the cross-covariance between the derivative of a first kernel and a second kernel is still stationary but anti-symmetric with respect to τ .

While the improved formalism seems very lucrative for cosmological analysis, there lie also a few cautions one must be aware of, before making cosmological inferences. Therefore, we would like to summarise advantages of this formalism and a few cautions to mind, when using MTGP in a

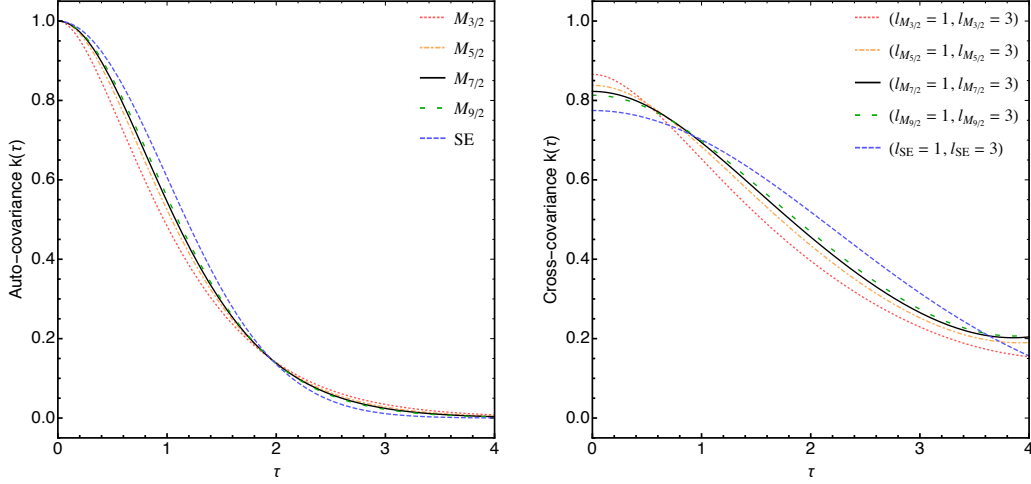


Figure 1. We show the auto-covariance (left) of various kernel functions and cross-covariance (right) between same kernel with different length scale hyperparameter (l) as a function of τ . In all the cases we have assumed the signal variance hyperparameter $\sigma = 1$.

joint analysis.

Advantages:

- Complementary datasets can help improve the reconstructions of data with sparse/fewer or noisy data points. This approach will provide better predictions one knows a priori that the two task are related, however with an unknown functional form, e.g., combining $E(z)$ and $H(z)$ datasets.
- Two closely related datasets that are different due to an unknown multiplicative or additive factor can be modelled together, and in fact the unknown factor(s) can be inferred from MTGP instead of having to assume them a priori. This can also be extended to analyses where the unknown factors are not simply constants but functions, e.g. powers of $(1+z)$, by appropriately adapting the kernels.
- In the current cosmological context, even if a statistical evaluation of the most suited kernel for a particular prediction can be eluding, given the physical nature of data one can infer the better suited kernel through heuristic arguments.
- While MTGP provides the freedom to model different tasks with different underlying latent functions, one can also enforce the equivalence and variedly utilise the shared hyperparameters for an extended analysis.

Disadvantages/cautions:

- One should be careful while combining two unrelated datasets, especially if they have very different constraining strength. As the dataset with stronger significance might force its own features on to the reconstruction of less stronger dataset. Such a phenomenon is called *negative transfer* (please refer to [62–64] for extensive discussion on this issue). We demonstrate this effect in Section 4.3.
- One must be careful in the choice of the covariance functions. As the number of tasks increases, the possible combinations of kernels rise drastically. It would be important to assess

the nature of the data and guess a reasonable covariance function. However, this is only of dire consequence if the data are not functionally related.

- While the data are related, it is still possible to obtain an *over-fitting*, if one of the dataset is much stronger than the other. This can however be easily seen when the stringent posteriors of one dataset over-power the reconstructions of other datasets, in the regions where they are void of data.

Here *over-fitting* and *negative transfer* are very similar phenomenon, which can however be conceptually differentiated. If the datasets are clearly unrelated and the features of one dataset over-power the other to provide incorrect features (nature of the curve) in the reconstruction of second dataset, it is negative transfer. If the datasets are a priori known to be related, and if the stronger dataset over-powers the second dataset to provide stringent reconstruction in the regions where there are no measurements available for the second dataset, this is also negative transfer that is of the over-fitting kind.

On the assumption of the mean $\mu(x) = 0$, often times when no prior knowledge of the latent function is available, the mean is assumed to be 0 by symmetry, to obtain posteriors of $f(x^*)$ i.e., y^* . Although, the reasoning for the same may not be apparent in the function-space viewpoint elaborated here, they become clearer in an equivalent basis function or weight-space⁷ viewpoint [53]. In a weight-space viewpoint the latent function can be written as,

$$f(\mathbf{x}) = \mathbf{w}^T \Phi(\mathbf{x}), \quad (2.18)$$

where $\Phi(x)$ ⁸ is a M -dimensional vector/set of basis functions and ‘ \mathbf{w} ’ is the corresponding vector of weights. This coincides with the formalism presented here, as the covariance kernel can now be written as $k(\mathbf{x}, \mathbf{x}') = \Phi(\mathbf{x})^T \Sigma(\Theta) \Phi(\mathbf{x}')$, which is equivalent to the convolution formulae, Eq. (2.12) and Eq. (2.11). Here $\Sigma(\Theta)$ is the covariance matrix of joint distribution on the weights, equivalently defined as a function of the hyperparameters ‘ Θ ’ and the kernel itself is a kind of prior on the expected nature of the latent functions, which however includes infinite possibilities. A prior assumption of the mean $\mu(x) = 0$ is in fact equivalent to assuming the mean of the prior on probability of weights $p(\mathbf{w}, \Theta) \sim \mathcal{N}(\mathbf{w}|0, \Sigma(\Theta))$ to be 0 in the weight-space view [65]. The probability distribution on ‘ \mathbf{w} ’ therefore characterizes $f(x)$, as $\bar{f}(\mathbf{x}) = \phi(\mathbf{x})^T \bar{\mathbf{w}}(\mathbf{x})$. A non-zero prior on the mean of the weights implies a non-minimal assumption⁸ analysis as a preferred functional form is a priori implied, from amongst all the infinite possibilities. And so, the prior of $\mu(x) = 0$ is seemingly conservative over any other asymmetric prior to make predictions, unless an accurate prior knowledge of the latent function is available or if the motive is ulterior, such as a residual analysis. Also note that the mean of the process itself is not necessarily the mean of the data. However, if the dataset is stringent enough, different assumptions of prior should not effect the reconstructions within the range of data. Therefore, especially if we do not expect that the data can alleviate the effect of different prior mean assumptions, the most unbiased prior that one can assume is $\mu(x) = 0$. Needless to say, such a constant mean prior implies that the posteriors y^* far outside the data range will invariably tend to 0 and varied assumptions of mean could alter the results in the extrapolated regions beyond agreement. An asymmetric assumption of prior mean might also tend to provide stricter, but biased reconstructions of the derivatives. While the symmetric $\mu = 0$ assumption provides fairly less stringent and hence unbiased, conservative reconstructions.

⁷Providing a complete description on these two viewpoints of Gaussian Processes is beyond the scope and necessity of this paper, we urge the reader to refer to Section 2.2 of [53] and Section 6.4 of [65] for more details.

⁸For example, $\Phi(x) = \{1, x, x^2, x^3 \dots, x^{M-1}\}$ is the vector of basis functions for polynomial regression. In fact, this makes

3 Theory and Data

In this section we provide a brief introduction to the theoretical framework of the standard cosmological model and then describe the data, different combinations of the same that we implemented with the MTGP method described in the previous section. In our current work we use the most-recent “low-redshift” BAO, CC, and SN datasets. We utilise the expansion rate measurements available from these data, which are however rescaled according to their corresponding observational methods. Finally, we have also presented a brief discussion on the criteria for kernel selection to infer the best description of the data.

3.1 Theoretical framework

The first Friedmann equation with all necessary degrees of freedom at “low-redshifts” is given by,

$$H(z)^2 \equiv H_0^2 E(z)^2 = H_0^2 \left[\Omega_m (1+z)^3 + \Omega_k (1+z)^2 + \Omega_{DE} f(z) \right], \quad (3.1)$$

where, H_0 is the present expansion rate, while Ω_m , Ω_{DE} and Ω_k are the dimensionless density parameters of matter, dark energy (DE) and curvature, respectively. The dynamics of the DE component are determined by $f(z)$, which is written as,

$$f(z) = \exp \left(3 \int_0^z \frac{1 + w_{DE}(\xi)}{1 + \xi} d\xi \right), \quad (3.2)$$

where $w_{DE}(z)$ is the equation of state (EOS) parameter of dark energy. The dimensionless density parameters are assumed to obey the cosmic sum rule of $\Omega_m + \Omega_{DE} + \Omega_k = 1$. Predominantly, in this work we assumed the concordance Λ CDM, with $\Omega_k = 0$ and a constant DE equation of state parameter ($w_{DE} = -1$) with $f(z) = 1$ for the diagnostics implemented here. Through the second Friedmann equation, $\ddot{a}/a = -4\pi/3G \sum_i \rho_i (1 + 3w_i)$, one can assess the present accelerated state of the universe by estimating the deceleration parameter which is written as,

$$q(z) = -a \frac{\ddot{a}}{\dot{a}^2} \equiv (1+z) \frac{H'(z)}{H(z)} - 1. \quad (3.3)$$

A negative value of the deceleration parameter implies an accelerated expansion and vice-versa. In addition, one can derive $q_0 = q(z=0) = \frac{3}{2}\Omega_m - 1$ in the standard Λ CDM scenario. The transition from deceleration to accelerated expansion is observed as a change in sign of $q(z)$, which is marked as the transition redshift z_T . Having significant constraints on the z_T and q_0 remain crucial for any cosmological inferences in the current GP framework.

The sound horizon at drag epoch is a relevant cosmological scale for the BAO data that is imprinted in the galaxy-clustering BAO data which is necessary to model the data. The sound horizon r_d at the drag epoch is given as,

$$r_d = \int_{z_d}^{\infty} \frac{c_s(z)}{H(z)} dz. \quad (3.4)$$

where c_s is the sound speed and z_d is the redshift at drag epoch. This physical scale is for rescaling the expansion rate data coming from BAO observations.

GP a generalization over the polynomial regression methods. A non-zero constant prior mean of ‘c’ on a 3-dimensional basis function vector of this kind, imposes a prior belief of the functional form to be $f(x) = c + cx + cx^2 + cx^3$, which is clearly a non-minimal supposition.

In the context of the standard framework, we utilise the $O_m(z)$ diagnostic [32, 50],

$$O_m(z) = \frac{(E(z))^2 - 1}{(1+z)^3 - 1}. \quad (3.5)$$

If the underlying expansion history $E(z) = H(z)/H_0$ is given by the Λ CDM, $O_m(z)$ is essentially a constant and is equal to the matter density Ω_m . Therefore, any deviations from this can be used to infer a dynamic nature of dark energy or an intrinsic dynamic nature of matter itself. The derivative of the $O'_m(z)$ w.r.t z provides information regarding the possible variations in $O_m(z)$. Note that $O'_m(z)$ utilises the information from both $H(z)$ and $H'(z)$ reconstructions, which is the additional information inferred from the GP analysis. Therefore, is more often susceptible to the prior assumption of the GP framework (i.e., kernel) itself. Apart from these well-known diagnostics, we also implement a simple modification of the $O_m(z)$ by considering the derivative of the first Friedmann equation (Eq. (3.1)) in the Λ CDM scenario, which provides,

$$O_m^{(2)}(z) = \frac{2E(z)E'(z)}{3(1+z)^2}. \quad (3.6)$$

It is straightforward that the reconstructed $O_m^{(2)}(z)$ should once again be equal to a constant Ω_m if the underlying cosmology is Λ CDM.

3.2 Data and Implementation

In this subsection we summarise the dataset implemented in the current analysis. We have collected datasets which provide information regarding the cosmic expansion history as a function of redshift. The straight-forward relatedness of these datasets allows one to perform the MTGP learning and conduct a joint analysis. The most recent expansion rate data from SN observations plays a very crucial role in our analysis, as it provides significant improvement at lower redshifts.

Baryon Acoustic Oscillations: In this work we utilise $H(z)r_d$ (hereafter abbreviated as BAO) observables provided in [2, 66–70]. In [2], using the Sloan Digital Sky Survey (SDSS) III DR-12, these observables were reported for the galaxy-clustering data set at three effective binned redshifts $z = 0.32, 0.57, 0.61$ (hereafter 3z). From the same data set also a tomographic analysis was performed in [66, 67], providing observables at 9 binned redshifts (9z). Implementing both the 3z and 9z datasets, we find very consistent results within the current methodology and remain to use 3z data set to quote the BAO contribution. More recently, [70] have reported 4 measurements of the same observables at redshifts $z = 0.98, 1.32, 1.52, 1.94$. In fact, these 4 measurements provide data in a much needed redshift range of $1 \lesssim z \lesssim 2$, that could play very important role also in the context of model-independent analysis and reconstructions. We also utilised the “high-redshift” Lyman- α measurements in [69] and [68] at redshifts $z = 2.33$ and $z = 2.4$. For all the datasets mentioned here we appropriately use the covariance matrices that have been provided⁹. While the observations are presented at times for $H(z)(r_{d, fid}/r_d)^{-1}$, we homogenise the dataset to $r_d H(z)$ considering the appropriate $r_{d, fid}$ used in the respective works. r_d clearly being a physical constant set at a much higher redshift ($z \sim 10^3$), the data is interpreted as expansion rate multiplied by a physical constant.

Supernova data: We have implemented the most recent binned expansion rate data provided in [4], where 6 $E(z)^{-1}$ (hereafter SN) data with their covariances have been provided for flat cosmologies ($\Omega_k = 0$). We invert the $E(z)^{-1}$, data to obtain the $E(z)$ points and corresponding covariance. However, the $E(z = 1.5)$ has been reported with an asymmetric uncertainty, while the conversion is unable to

⁹All values of the mean, dispersion and covariances of $H(z)r_d$ observables for the galaxy clustering BAO data are taken from <https://data.sdss.org/sas/dr12/boos/papers/clustering/>.

Table 1. $E(z)$ obtained from the inversion of the $E(z)^{-1}$ data reported in Table 6 of [4]. Note the difference in the estimate of $E(z = 1.5)$, from the actual quoted value. We simply adopt the mean value obtained from our inversion of quoted $E(z)^{-1}$ data.

z	$E(z)$
0.07	0.997 ± 0.023
0.20	1.111 ± 0.021
0.35	1.127 ± 0.037
0.55	1.366 ± 0.062
0.9	1.524 ± 0.121
1.5	2.924 ± 0.675

account for the same. We simply proceed to utilise the symmetric $E(z = 1.5)$ estimate obtained from the inversion, as this data point with its large uncertainty would not be of key importance and also with little significance in the current joint analysis MTGP formalism¹⁰. The $E(z)$ data implemented in this work are summarised in Table 1.

Cosmic Chronometers: The measurements of cosmic expansion rate have been estimated using the differential age method suggested in [39], which considers pairs of massive and passively evolving red galaxies at similar redshifts to obtain dz/dt , which are deemed as the *Standardisable Clocks*. In this work, we utilise two different compilations of CC data, differentiated based on the stellar evolution models assumed for obtaining them. A dataset of 31 (CC) uncorrelated data points is compiled from [40, 51, 52, 71–74] and is the most utilised compilation in literature. Data points in this compilation are obtained utilising the BC03¹¹ models for the stellar evolution and comprise of larger number of data points as it was the only method implemented in earlier works such as [71–73]. Along side this standard dataset we compile a different set of data that are obtained using the MaStro¹¹ stellar evolution models. This compilation consists of 15 (CC_M) points taken from [40, 51, 52]. Given, the fewer number of data points this compilation does not add significant weight in the standard model-dependent residual analysis. However, in a model-independent analysis, the estimate of H_0 as an intercept of the reconstruction of $H(z)$ data can provide significant constraints on the mean with both datasets. Also, complementing with other datasets in our analysis we find suitable arguments to comment on the systematics within the CC dataset. The 15 point compilation of CC_M is compared against the corresponding 15 point compilation obtained using the BC03 (CC_B) method. All the CC data yet available in the literature are summarised in Table 2. Notice that the data at redshifts $z = 1.363$ and $z = 1.965$ are quoted as same for both the methods, as contribution of systematic error for different choice of the stellar evolution method is included in quadrature within the reported error. Repeating our analysis decoupling this systematic error, we however find compatible results to the ones discussed here. Finally, we refer to the recent analysis of [75], that explored the dependence of these data on a possible contamination due to an young underlying component, assessing that for current measurements this effect has been confined by the selection criteria well below the estimated errors.

¹⁰We also verify that the reconstructions show very little to no change, if the $E(z = 1.5)$ data point is omitted. Also, [18] have commented in similar lines, regarding the $E(z = 1.5)$ data point.

¹¹This BC03 compilation was implemented in several works such as [18, 29, 76], also at times including the BAO measurements assuming an r_d . In [40, 51, 52] the estimates from BC03 models were also accompanied with the estimates obtained using MaStro models.

Table 2. Cosmic chronometers data obtained using both the BC03 and MaStro models compiled from all the available literature [40, 51, 52, 71–74] are summarised here.

z	$H(z)$ [km/s Mpc ⁻¹]		Ref.
	BC03 (CC & CC _B)	MaStro (CC _M)	
0.0708	69.0 ± 19.68	-	[73]
0.09	69.0 ± 12.0	-	[71]
0.12	68.6 ± 26.2	-	[73]
0.17	83.0 ± 8.0	-	[71]
0.1791	75.0 ± 4.0	81.0 ± 5.0	[51]
0.1993	75.0 ± 5.0	81.0 ± 6.0	[51]
0.2	72.9 ± 29.6	-	[73]
0.27	77.0 ± 14.0	-	[71]
0.28	88.8 ± 36.6	-	[73]
0.3519	83.0 ± 14.0	88.0 ± 16.0	[51]
0.3802	83.0 ± 13.5	89.3 ± 14.1	[40]
0.4	95.0 ± 17.0	-	[71]
0.4004	77.0 ± 10.2	82.8 ± 10.6	[40]
0.4247	87.1 ± 11.2	93.7 ± 11.7	[40]
0.4497	92.8 ± 12.9	99.7 ± 13.4	[40]
0.47	89.0 ± 34.0	-	[74]
0.4783	80.9 ± 9.0	86.6 ± 8.7	[40]
0.48	97.0 ± 62.0	-	[72]
0.5929	104.0 ± 13.0	110.0 ± 15.0	[51]
0.6797	92.0 ± 8.0	98.0 ± 10.0	[51]
0.7812	105.0 ± 12.0	88.0 ± 11.0	[51]
0.8754	125.0 ± 17.0	124.0 ± 17.0	[51]
0.88	90.0 ± 40.0	-	[72]
0.9	117.0 ± 23.0	-	[72]
1.037	154.0 ± 20.0	113.0 ± 15.0	[51]
1.3	168.0 ± 17.0	-	[72]
1.363	160.0 ± 33.6	160.0 ± 33.6	[52]
1.43	177.0 ± 18.0	-	[72]
1.53	140.0 ± 14.0	-	[72]
1.75	202.0 ± 40.0	-	[72]
1.965	186.5 ± 50.4	186.5 ± 50.4	[52]

All three datasets mentioned above are implemented as different tasks in the MTGP formalism. For this purpose, we start by rescaling/rewriting the data with arbitrary fiducial values such as, $H_0^{fid} = 100$, $r_d^{fid} = 148$ ¹². In principle there is no need for this step and it would not effect the results in any way, but we do so to have all the datasets within similar order(s) of magnitude, which we find also convenient in the numerical computations. The $E(z)$ data is not rescaled but we do include a theoretical data point of $E(z = 0) = 1$. This assumption is justified assuming the results are essentially

¹²Please note the difference from $r_{d,fid}$, which is the fiducial value assumed in the process of acquiring the measurement in various works and r_d^{fid} is simply a rescaling factor used only in our analysis, for convenience.

restricted to a comparison for flat cosmologies within standard framework where $H(z) = H_0 E(z)$ and it also helps improve the features in the reconstructions. To summarise, we have three tasks: $E(z)$ (SN), $r_d H(z) (H_0^{fid} r_d^{fid})^{-1}$ (BAO) and $H(z) (H_0^{fid})^{-1}$ (CC). The CC dataset however is replaced for different compilations that are mentioned above. We perform the MTGP on these three datasets considering several combinations of kernels described in Section 2. Given the known relatedness of the data i.e., all the datasets are essentially providing expansion history, while being rescaled accordingly with physical constants ($H_0, r_d H_0$), we do not immediately perform any inter-kernel analysis. We assume the same covariance function for all the three tasks, however with their own hyperparameters and the appropriate cross-covariance. We will comment in detail about this assumption of the kernels and their implications in the Section 4. Therefore, we have to finally optimise the log Marginal likelihood for six hyperparameters in total, two for each task.

3.3 Kernel selection

The Single-Task GP formalism so far implemented in the cosmological context has been deprived of a proper kernel selection criteria, to evaluate the best possible model of the covariance. Given that the data are not very stringent, there was however no strong motivation to discuss the same. As is also mentioned in [18], this dependence on the prior choice of the covariance functions enforces a subjectivity, that needs to be properly accessed to select the kernel model that provides best description of data and hence for making reliable predictions. As we anticipate the MTGP formalism is capable of providing better reconstructions of the data, the need to discuss kernel selection will become relevant. Note that here kernel selection should not be confused with model selection as is discussed in the context of Gaussian Process literature. The standard idea of model selection in a Gaussian Process regression (GPR) is to select the best possible explanation for the data given the assumed model. In a non-Bayesian approach, as is implemented in this paper, the best explanation for the data is obtained by optimising the log marginal likelihood, $\log[p(\mathcal{D}|\mathcal{M})]$ (equivalent to Eq. (2.6)). Here \mathcal{D} are the observations and $\mathcal{M}(\Theta)$ is the assumed probabilistic model, with parameters Θ . In the Bayesian formalism the posterior distribution of Θ is given according to the Bayes' rule as,

$$p(\Theta|\mathcal{D}, \mathcal{M}) = \frac{p(\mathcal{D}|\Theta, \mathcal{M})p(\Theta|\mathcal{M})}{p(\mathcal{D}|\mathcal{M})}, \quad (3.7)$$

where $p(\Theta|\mathcal{M})$ is the prior on the models. Here $p(\mathcal{D}|\mathcal{M})$ is the marginalised likelihood which is also referred to as *evidence* and is given as,

$$p(\mathcal{D}|\mathcal{M}) = \int p(\mathcal{D}|\Theta, \mathcal{M})p(\Theta|\mathcal{M})d\Theta. \quad (3.8)$$

Therefore a comparison of the *evidence* for two given different models \mathcal{M}_A and \mathcal{M}_B , with equal/uniform priors $p(\mathcal{M}_A) = p(\mathcal{M}_B)$, can be inferred using the Bayes factor as,

$$\frac{p(\mathcal{M}_A|\mathcal{D})}{p(\mathcal{M}_B|\mathcal{D})} = \frac{p(\mathcal{D}|\mathcal{M}_A)}{p(\mathcal{D}|\mathcal{M}_B)}, \quad (3.9)$$

where $p(\mathcal{M}_i|\mathcal{D})$ is the posterior for a respective model (\mathcal{M}_i) given observations \mathcal{D} . The comparison of evidence can also be extended to the non-Bayesian formalism to select the model that performs better for the given data. However, this comparison is not necessarily suitable to perform selection amongst different kernels, with smooth data such as in the current cosmological context. While the comparison of the evidence for different kernels can be performed, it is usually the case with smooth datasets, that one finds the simplest (in this case the SE) covariance function to be probabilistically best suited. To

test the performance of an assumed kernel, one can appropriately split the available data into two separate categories, i.e., the training data ($\sim 75\%$) and test data ($\sim 25\%$). Learning is performed on the training dataset and then the predictions based on the optimised model are compared against the test data by estimating either the *standardised mean squared error* or *negative log probability* [53] (see also [60] for the examples demonstrated therein).

In this work we attempted to evaluate for a better suited kernel combination for available cosmological data, but to no avail. We performed 5 different tests by selecting $\sim 25\%$ or $\sim 33\%$ of the total data points as test data, over different regions, to perform extrapolation (test data chosen towards one end of the dataset) and interpolation (test data chosen in the central regions of the dataset) tests. We find a more flexible kernel (e.g., $M_{3/2}$) to be more suitable in 2 cases, while $M_{9/2}$ and $M_{5/2}$ are the best suited in 3 other cases. We conclude that the current compilation of the cosmological data is unable to provide strict preference for a particular kernel combination. In the current work we implement all the kernels except the $M_{3/2}$, which we a priori deem not very suitable for the reconstructions of the derivatives, given its complexity (higher flexibility) to predict more features, differentiability only to first order, and the fact that it could suffer from possible over-fitting due to higher cross-covariance compared to other kernels. However, we do utilise $M_{3/2}$ to compare estimates of cosmological parameters at particular target points such as H_0 , q_0 etc. While the discussion presented here is minimal, a proper kernel selection criteria is necessary if one intends to draw more robust inferences from future data to arrive, in a cosmology independent way. Kernel selection is a much more intricate problem and needs dedicated attention [46], which we intend to leave for an extended future investigation.

However, given the nature of the cosmological data we utilise an heuristic argument to select from amongst the kernel combinations available. As we expect all the expansion rate data to have the same underlying latent functions, the kernel combination that is capable of predicting the same is clearly a better choice. Therefore, we infer the kernel combination with minimum amount of complexity required to predict equivalent latent functions to provide the best description of data. Increasing the flexibility of the kernels beyond this limit might very well provide same latent functions across datasets, however being over-fitted or predicting additional/unrealistic features. Finally, we choose from amongst the smoother kernels with these criteria to report suitable constraints and comment on the cosmological inferences, *while all the kernel choices remain equally valid in a statistical context of kernel selection.*

4 Results and Discussion

In this section we present the results obtained from the above described formalism Section 2 implemented with the data presented in Section 3. We present the results for estimation of cosmological parameters, reconstructions and then proceed to report the inferred diagnostics, and the corresponding improvement in comparison to the previous works. Finally, we also comment on the issue of systematics within the CC data.

4.1 Reconstruction of data and cosmological parameters

We start by reproducing the existing results through individual GPs and then proceed to discuss our results from the MTGP method. In Figure 2 we show the individual GP reconstructions of all three datasets considered in this work using the $M_{9/2}$ kernel, as an example. The well-known reconstruction of the CC data is very similar to the one obtained using the SE kernel (see also [29]), which we reassert. Also, we find very little to no difference from one kernel to another even for the SN and BAO datasets. Reconstruction of the SN data is clearly at a disadvantage for higher redshifts ($z \gtrsim 1.5$) due to the very large uncertainty of the data point at $E(z = 1.5)$. The inclusion or exclusion of this data

point makes very little difference as was also mentioned in [18] (while they opt to omit this data point we let it remain, with no dire consequences). Similarly, the BAO data is unable to show any features in the reconstructions due to the sparse distribution, and the unavailability of data at redshifts $z \lesssim 0.3$. Note that using Matérn kernels that are necessarily modelled to allow more dynamics/features do not fare immediately well in Single-Task GP. We do not find much difference, except for an increase in the predictive region in the $M_{3/2}$ case compared to $M_{9/2}$. In fact, their utility only becomes significant in our MTGP formalism or when the derivatives are inferred. The estimates for H_0 using several kernels from the reconstruction of CC data alone are summarised in Table 3.

Another important feature worth noticing is the lowering of the posterior mean in the SN and CC reconstructions in the left column of Figure 2. As mentioned earlier in Section 2, this feature could essentially be due to the lack of data in this region and the tendency of reconstructed Gaussian posterior to converge to a 0 (prior) mean with 68% confidence regions given by the optimized σ_f hyperparameter, far away from the data. Clearly, this implies that the extrapolated reconstructions are less reliable in individual GPs for making inferences, e.g., H_0 ¹³. In the MTGP formalism however, these three datasets are modelled to complement each other, thereby providing better reconstructions for all the three datasets (see Figure 3). This essentially implies that the overall range of available data is now an union of all three dataset ranges i.e., $0 < z \lesssim 2.5$ and as expected the effects of extrapolation in all three data planes are now mitigated within this joint redshift range.

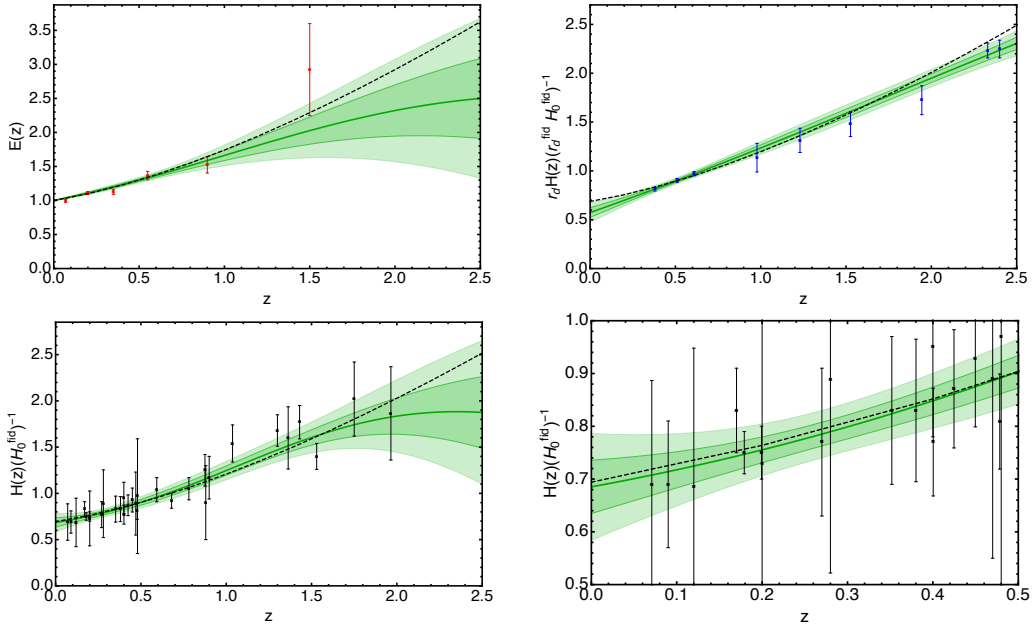


Figure 2. We show the reconstructions of all three datasets, SN (*top-left*), BAO (*top-right*) and CC (*bottom-left*) using the individual GP formalism with an assumed $M_{9/2}$ kernel. In the bottom-right panel we show the CC reconstruction enlarged in the redshift range $z < 0.5$. The data have been rescaled with fiducial values $H_0^{fid} = 100 \text{ km/s Mpc}^{-1}$ and $r_d^{fid} = 148 \text{ Mpc}$. The dashed line corresponds to best-fit ΛCDM to the three datasets, with $\Omega_m = 0.284$ and $H_0 = 69.77 \text{ km/s Mpc}^{-1}$, in all the frames (see Section 4.2). The symmetric prior of $\mu(x) = 0$, with practically unbounded maximum of $\sigma_f \sim 10$ (68% region) reflects the entire region of the plots and more on either side of the y-axis.

¹³However, it can be noticed in Figure 2 that the CC reconstruction does not immediately start tending to the prior $\mu = 0$ as $z \rightarrow 0$.

As mentioned earlier (Section 3.2), we implement the same kernel across all the data planes considered, as one would expect that the dynamic nature of the reconstructions for these three datasets is essentially the same. However, this does not imply that we a priori assume the underlying latent functions of the three tasks to be unique and hence one could find differences in the reconstructions from one plane to another (see Figure 4), as the log marginal likelihood (Eq. (2.6)) is optimised with different length scale hyperparameters for each task. While it might appear inconsistent if the three data planes do not predict unique expansion histories, this freedom provides us with the criteria to evaluate the best kernel combination, that will predict equivalent latent functions, i.e., similar length scale hyperparameters. In fact, the equivalence of the posteriors validates that the observed features in MTGP analysis are no longer due to extrapolated effects but due to the optimized data correlations. After obtaining equivalent predictions in all three planes, the results must be read as the final reconstruction of the CC dataset complemented by the BAO at high-redshift and SN data at low redshifts. Therefore we later proceed to study the dynamics based on the CC reconstruction, obtained from MTGP of all three datasets, although the dynamics inferred from all three reconstructions will be same. We test for a strict equivalence between CC and BAO reconstructions, and expect an agreement amongst CC/BAO and SN to be better than 1σ at higher redshifts (as SN data completely lacks information in this region). We infer H_0 as the intercept of CC reconstruction and use CC and BAO reconstructions to obtain r_d as a rescaling constant at the intercept of BAO reconstruction.

Dataset(s)	Kernel	H_0 [km/s Mpc $^{-1}$]	r_d [Mpc]
CC	SE	67.44 ± 4.75	-
	$M_{9/2}$	68.52 ± 5.06	-
	$M_{7/2}$	68.73 ± 5.18	-
	$M_{5/2}$	68.89 ± 5.43	-
CC+SN+BAO	SE	67.39 ± 1.36	144.24 ± 4.38
	$M_{9/2}$	68.24 ± 1.09	145.35 ± 3.34
	$M_{7/2}$	68.52 ± 0.94	145.61 ± 2.82
	$M_{5/2}$	68.86 ± 0.67	145.78 ± 2.02
CC+SN+BAO+R18	SE	71.58 ± 0.34	140.54 ± 1.28
	$M_{9/2}$	71.35 ± 0.31	141.45 ± 1.35
	$M_{7/2}$	71.40 ± 0.30	141.29 ± 1.31
	$M_{5/2}$	71.61 ± 0.37	141.67 ± 1.39

Table 3. Estimates of H_0 and the corresponding statistical 1σ uncertainty form combinations of the dataset(s) described in Section 3 implemented for several kernels shown in second column. For the joint analysis with the multiple datasets we implement the same kernel for each of them. In the last column we report the r_d estimates and their uncertainties. Highlighted in bold is the best inference for the reasons stated in the text.

In this respect, we would also like to mention that a unique expansion history can be enforced by restricting to the same kernel across all the planes and then imposing that the length scale hyperparameter (l) of the three planes should be the same. This assumption in the MTGP formalism implies that the datasets are in fact rescaled by an absolute constant (i.e., H_0 (CC) and $H_0 r_d$ (BAO), are the rescaling factors with respect to the $E(z)$ (SN) reconstruction), and is clearly not the right assumption to make a priori for the estimation of physical constants. This provision within the formalism becomes extremely useful to estimate the differences in the data once the requirement for unique expansion history is enforced, as will be elaborated later (see Section 4.2). Reconstructions of the three datasets in the MTGP formalism using $M_{7/2}$ kernel are shown in Figure 3. One can immediately

notice a significant improvement in the reconstruction of the individual datasets in comparison to the single task GP shown in Figure 2. The Matérn class kernels with lower $\nu < 7/2$ force the SN reconstruction to be more in agreement with the CC and BAO data reconstructions, which is a consequence of higher cross-covariance between $M_{\nu \leq 5/2}$ kernels for $\tau \rightarrow 0$, as shown in right panel of Figure 1.

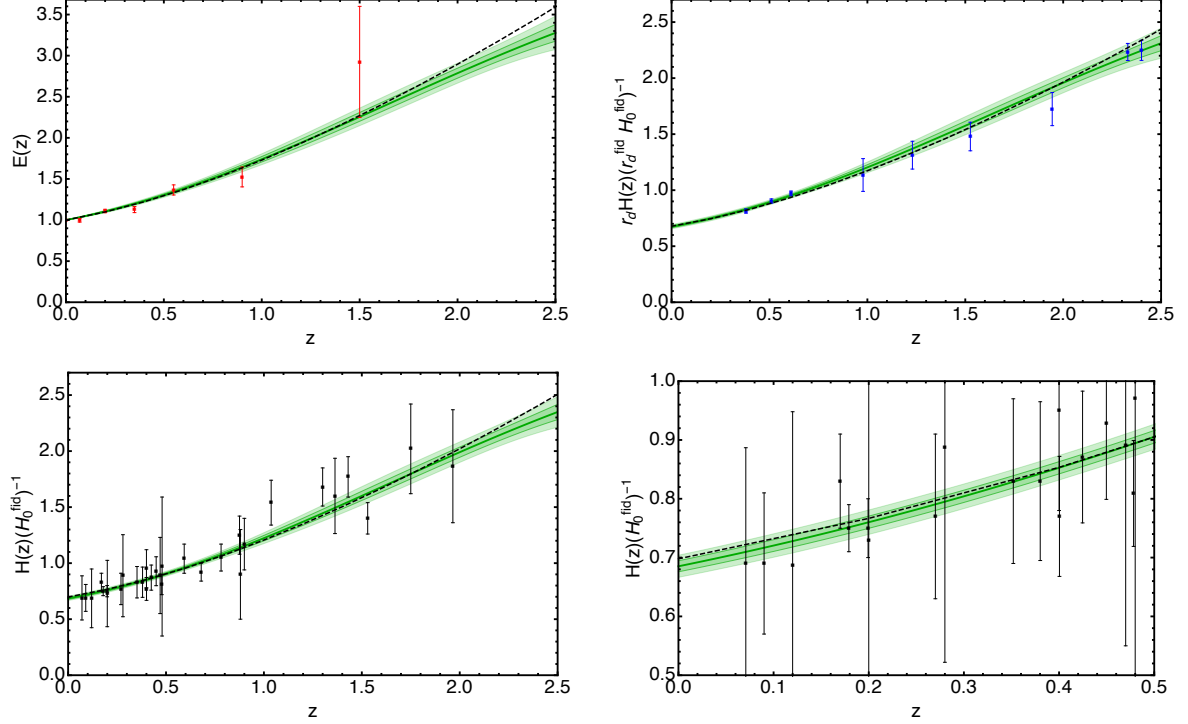


Figure 3. We show the reconstructions of the all the three datasets using the MTGP formalism with an assumed $M_{7/2}$ kernel for each of the dataset. The right-bottom panel shows the enlarged image of the $z < 0.5$ region to emphasise the improvement in the intercept of the CC reconstruction. In all three the panels the data are rescaled with $H_0^{fid} = 100 \text{ km/s Mpc}^{-1}$. The dashed line corresponds to the best-fit Λ CDM with $\Omega_m = 0.284$, $H_0 = 69.77 \text{ km/s Mpc}^{-1}$, to the current data (see Section 4.2). The symmetric prior of $\mu(x) = 0$, with practically unbounded maximum of $\sigma_f \sim 10$ (68% region) reflects the entire region of the plots and more on either side of the y-axis.

Comments on H_0 estimates: The H_0 estimate has been discussed time and again in the context of Gaussian Processes, usually prompting values compatible with the “high-redshift” CMB estimate (P16). We show the improvement in the H_0 estimate and the corresponding uncertainty from standard GP to the MTGP implemented here, with the addition of complementary datasets in Table 3. In the MTGP formalism with the inclusion of BAO data alone to the CC dataset there is only a modest improvement in the H_0 estimate, while the more important contribution is seen at the high redshifts ($2 \leq z \leq 2.5$), where no CC data is available and hence the BAO data aids the CC reconstruction in this redshift range. For example, assuming $M_{9/2}$ kernel we find $H_0 = 64.00 \pm 3.35 \text{ km/s Mpc}^{-1}$ using CC + BAO datasets. In fact, this uncertainty on H_0 is of the order of the single task GP conducted in [29], where BAO (5 points) data were rescaled with an assumed r_d and a Single-Task GP was performed on a total of 36 (31 CC + 5 BAO) data points. We also find a very good agreement in the predicted mean values of the H_0 from this work and [29] (please see Table 2. therein). The SN data helps to improve the CC reconstruction at lower redshifts and hence the H_0 estimate to a much

higher significance. As shown in Figure 2, the SN data is at a severe disadvantage for $z \gtrsim 1$ in the simple GP, while its reconstruction in the MTGP is improved by several orders (see Figure 3). More recently, [18] have utilised the same SN and CC dataset to obtain the H_0 in a modified GP formalism. We do assert the consistency in our findings with the results quoted therein. Our formalism provides slightly stringent constraints, while also allowing for a provision to include the BAO data.

As anticipated, we encounter a subtlety of choosing the best possible reconstruction (i.e., choice of the kernel) to quote the H_0 estimate. While it is appreciable that the constraint on H_0 gets better with more flexible Matérn class kernels ($\nu \leq 5/2$), it is compensated with the overall predictive quality of the reconstruction. For example, using $M_{3/2}$ we find $H_0 = 69.09 \pm 0.38 \text{ km/s Mpc}^{-1}$, which is even more stringent compared to $M_{5/2}$ constraint shown in Table 3. While the predictive region is over-fitted at lower redshifts due to higher cross-correlation with SN data, at the same time, for the higher redshift regions the predictive quality deteriorates in comparison to the kernels with $\nu > 3/2$. Implementing the heuristic argument for equivalent latent functions introduced in Section 3.3 (also elaborated alongside r_d estimates), we argumentatively infer that the assumption of $M_{7/2}$ kernel provides a better description of the data in comparison to other kernel choices for the final constraints on all the cosmologically relevant quantities. Our best estimate of $H_0 = 68.52 \pm 0.94 \text{ km/s Mpc}^{-1}$ from CC + SN + BAO datasets, is in agreement with the “high-redshift” P16 at 1.41σ and in a lesser agreement with the local H_0 measurement of R18 at 2.60σ ¹⁴. This remains to be a very consistent scenario as we have also reported in [80] (see also references therein and [81], for discussion on H_0 estimates from “low-redshift” data), even in a model-dependent analysis, especially if measurements obtained with other stellar population synthesis models are not considered, as we discuss in Section 4.2. Our estimate seems to be consistent with several other low-redshift estimates obtained in [77]. In any case, we find that all the reconstructions assuming each of the SE or Matérn class kernels ($\nu \geq 5/2$) provide very consistent estimates on H_0 .

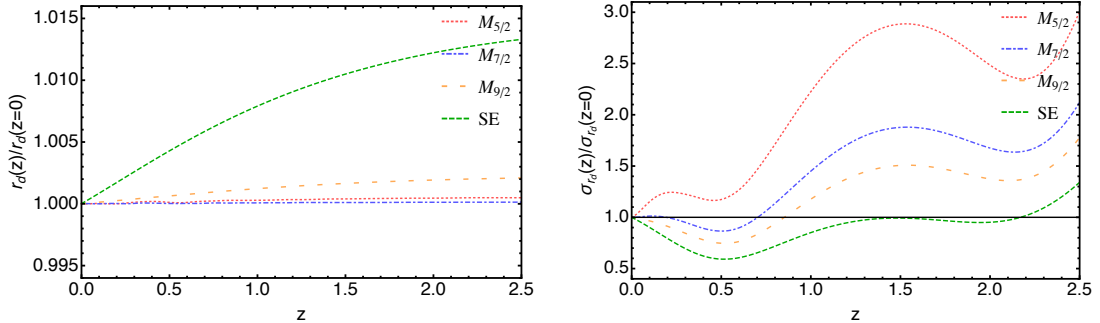


Figure 4. In the left panel we show the variation in the r_d estimate normalised to $r_d(z = 0)$ at the intercept, over redshift using different kernels. Similarly, in the right panel we show the variation of the uncertainty σ_{r_d} in redshift, also normalised to $\sigma_{r_d}(z = 0)$.

Comments on r_d estimates: In Table 3, we also show the constraints on r_d obtained as the rescaling factor between CC and BAO reconstructions. As can be seen in Figure 3, the formalism is able to realise that the CC and BAO data are in fact better described by similar underlying latent functions and hence provide similar length scale hyperparameters. This implies that two reconstructions are well correlated and hence a better description of the data. The reconstructed regions in Figure 3 are obtained after including the covariance between all the datasets. Therefore, the r_d estimates are

¹⁴Our estimate is also consistent with several other previous estimates in [18, 29, 77–80].

simply obtained as the ratios of the intercepts in the BAO and CC reconstruction planes without the need to reconsider the covariance. A cross-check for the same would be if we obtain $\sigma_{r_d} = 0$ when considering the covariance twice while obtaining the uncertainty. In fact, when the SE kernel is considered the MTGP is unable to provide a good equivalence for the latent functions of BAO and CC reconstructions, which leads to a non-zero (0.36) uncertainty when covariance is considered twice.

We would also like to remind that the estimates of r_d quoted in Table 3 are obtained as the intercept of the BAO data at $z = 0$. However, this constraint is supposed to improve at the best-reconstructed redshift of the BAO/CC data. Using the CC + BAO + SN data we find $r_d = 145.61 \pm 2.44$ Mpc ($M_{7/2}$), $r_d = 145.45 \pm 2.49$ Mpc ($M_{9/2}$) and 144.86 ± 2.59 Mpc (SE) estimates at redshifts $z = 0.50$, $z = 0.53$ and $z = 0.51$, respectively. Incidentally, all the kernels predict the best reconstruction at similar redshifts, which is not surprising as the better of BAO data are available in the redshift range of $0.38 < z < 0.61$. This trend is no longer the case when more flexible kernels like $M_{5/2}$ and $M_{3/2}$ are used, as they predict the best BAO reconstruction at $z = 0$ under the influence of strong cross-covariance with SN data. This, in fact, is a clear indication of over-fitting and these kernels are therefore not inferred as to provide the best description of the data. Needless to say, the estimate of mean r_d would also slightly vary when obtained from different redshifts if the optimised underlying latent functions are not exactly the same. For example, the mean of r_d estimated at different redshifts with $M_{9/2}$ kernel varies in the range of $145.35 < r_d < 145.65$, accounting for an additional systematic uncertainty of $\sigma_{r_d}(\text{sys}) = 0.01$ in quadrature. In the left panel of Figure 4 one can notice that using $M_{7/2}$ kernel we find that the latent functions of CC and BAO are completely equivalent, which makes it a better choice compared to SE and $M_{9/2}$. On the other hand, it remains better than $M_{3/2}$ and $M_{5/2}$ as they provide over-fitted predictions (see right panel of Figure 4). All the above-presented arguments show $M_{7/2}$ kernel in the light of being the better choice to make predictions with the current data. Also in the left panel of Figure 4, it can be clearly seen how the uncertainty of the r_d predictions follow the availability of data, by providing better reconstructions around $z \sim 0.5$ and $z \sim 2.3$, compared to their neighbouring regions.

Finally, we also include the R18 estimate of $H_0 = 73.48 \pm 1.66$ km/s Mpc $^{-1}$ [45] to the joint analysis and as expected find improvement in the constraints on H_0 . We find extremely consistent final estimates of H_0 with all the covariance functions utilised in this work. In fact, this high consistency discourages us to make any inference for the best-suited kernel. However, we do utilise $M_{7/2}$ to quote the results throughout the discussion, without loss of reasoning. The estimated H_0 and r_d values are summarised in Table 3. As expected, with the addition of R18 to the CC + SN + BAO data the errors on the parameters decrease. Given the increase in H_0 values we find a consistent decrease in r_d estimates, also retaining very similar reconstructions using all the kernels. Clearly, the value of $r_d H_0$ is conserved while the estimate of r_d gets rescaled to accommodate the increase in the value of H_0 . Using $M_{7/2}$ kernel, with and without the inclusion of R18 we find $r_d H_0 = 10088 \pm 83$ km/s and $r_d H_0 = 9977 \pm 237$ km/s, respectively. While the latter estimate without the inclusion of R18 is extremely consistent with the “high-redshift” CMB estimate of $r_d H_0 = 9952 \pm 97$ km/s [82], the earlier measure is also consistent at $\sim 1\sigma$. Given the agreement in $r_d H_0$, the discrepancy remains within breaking the degeneracy between these two parameters. However, we extend this discussion also to a comparison of Ω_m estimates in Section 4.2.

Illustration of Multi-Kernel Gaussian Process: So far we have performed the analysis assuming same kernel for all the datasets. Now we would like to extend the analysis with the provision within current formalism and utilise different kernels for the datasets and comment on its implications. As anticipated, we find that the initial assumption of one kernel for all the datasets remains best, as we try to implement several combinations. For instance the combination of $M_{9/2}$ (SN), SE (BAO), SE (CC) fares poorly compared to the assumption of SE kernel for all datasets. In this case, we find

CC	SN	BAO	H_0 [km/s Mpc $^{-1}$]
SE	$M_{9/2}$	SE	65.46 ± 2.60
$M_{9/2}$	SE	$M_{9/2}$	65.47 ± 2.60
$M_{9/2}$	SE	SE	68.05 ± 2.89
SE	$M_{9/2}$	$M_{9/2}$	66.69 ± 2.75
$M_{9/2}$	$M_{7/2}$	$M_{7/2}$	67.94 ± 1.94

Table 4. Estimates of H_0 and the corresponding 1σ uncertainty for different combinations of the kernels assumed for different datasets shown in first three columns.

$H_0 = 65.46 \pm 2.6$ km/s Mpc $^{-1}$. In Table 4 we show a few combinations of kernels assumed in the joint analysis and the corresponding H_0 estimates obtained. One can clearly notice that the assumption of different kernels does not improve the results. This can be easily explained based on the strength of cross-covariance that is obtained using a specific combination of kernels and the nature of data. In the Multi-Kernel formalism a combination including $M_{3/2}$ kernel provides stronger covariance for closer data points (i.e., lower $|x - x'|$) compared to $M_{5/2}$, SE and the contrary for far away data points. As we have described earlier the SN data complements the CC data essentially at lower redshifts, and hence giving rise to larger dispersion at the intercept of CC reconstruction for kernel combinations involving $M_{\nu \geq 5/2}$ due to their lower cross-covariance. Although this provision is an added advantage in the current formalism, given the nature of data it does not present any immediate application in the current work. However, it could be of interest in several other implementations such as [83–85], presenting which is beyond the scope of this paper.

4.2 Reconstruction of Diagnostics

In this section we present our results for the reconstructions of $q(z)$ and a few diagnostics inferred from the MTGP analysis described in the Section 3. We find a very significant improvements in the reconstructed regions, moving from single GP performed in [25, 31] to MTGP formalism implemented here. In Table 5 we present the estimates for q_0 and the transition redshift z_T , corresponding to the joint analysis presented in Table 3. We find very consistent results for the estimates of q_0 and z_T using all the kernels implemented here. However, one can notice that the error on estimated q_0 increases with the flexibility of assumed kernel, while z_T is constrained mildly better by the $M_{9/2}$ and $M_{7/2}$ kernels. As expected, using $M_{3/2}$ we find deteriorated constraints of $q_0 = -0.56 \pm 0.12$ and $z_T = 0.51^{+0.43}_{-0.06}$ reassuring that the assumption of $M_{3/2}$ is not suitable for reconstructions of derivatives. We would like to remind that $M_{3/2}$ kernel might retain its importance for reconstructing other cosmological features such as dynamical dark energy equation of state (EoS) $w(z)$. Very similar trend is observed even when R18 is included as data, in the reconstruction. Consistently, the inclusion of R18 predicts higher values of z_T and lower values of q_0 , appropriately suggesting lower Ω_m . It is also the case that inclusion of R18 does not improve the constraints on q_0 and z_T , as can be seen in Table 5.

We find our estimates of z_T obtained using all the kernels shown in Table 5 to be extremely consistent with $z_T = 0.64^{+0.11}_{-0.07}$ quoted in [40], which was however obtained using the CC+ H_0 data alone. Although, the constraint has not improved in comparison, in this analysis we have the advantage of being cosmology-independent (see [40] for more discussion on z_T estimates from various analysis). using our formalism we find as yet the best constraints on z_T in a cosmology-independent formalism. As was also summarised in [49], a stringent constraint on z_T would provide useful insights for accessing the nature of dark energy. As mentioned above, our results are very much consistent with the

Dataset(s)	Kernel	q_0	z_T
CC	SE	$-0.48^{+0.33}_{-0.29}$	$0.56^{+0.22}_{-0.24}$
	$M_{9/2}$	$-0.57^{+0.38}_{-0.33}$	$0.54^{+0.18}_{-0.17}$
	$M_{7/2}$	$-0.57^{+0.39}_{-0.35}$	$0.54^{+0.18}_{-0.16}$
	$M_{5/2}$	$-0.57^{+0.44}_{-0.38}$	$0.54^{+0.18}_{-0.14}$
CC+SN+BAO	SE	-0.44 ± 0.04	$0.63^{+0.12}_{-0.11}$
	$M_{9/2}$	-0.50 ± 0.05	$0.64^{+0.12}_{-0.09}$
	$M_{7/2}$	-0.52 ± 0.06	$0.64^{+0.12}_{-0.09}$
	$M_{5/2}$	-0.54 ± 0.07	$0.63^{+0.15}_{-0.10}$
CC+SN+BAO+R18	SE	-0.57 ± 0.05	$0.74^{+0.09}_{-0.07}$
	$M_{9/2}$	-0.57 ± 0.06	$0.73^{+0.11}_{-0.09}$
	$M_{7/2}$	-0.58 ± 0.06	$0.71^{+0.12}_{-0.09}$
	$M_{5/2}$	-0.59 ± 0.07	$0.69^{+0.17}_{-0.10}$

Table 5. Estimates of q_0 and z_T and the corresponding 1σ uncertainties from combinations of the dataset(s) described in Section 3 implemented for several kernels shown in second column. For the joint analysis with the multiple datasets we implement the same kernel for each of them. The results in this table correspond to the second and third rows of Table 3.

Λ CDM model where we find the constraints to be $H_0 = 69.77 \pm 1.80$ km/s Mpc $^{-1}$, $\Omega_m = 0.284^{+0.018}_{-0.017}$ and $r_d = 143.9^{+4.0}_{-3.9}$ Mpc using CC + SN + BAO. Including the R18 prior in the model-dependent analysis we find $H_0 = 71.80 \pm 1.20$ km/s Mpc $^{-1}$, $\Omega_m = 0.275^{+0.016}_{-0.015}$ and $r_d = 141.1^{+3.4}_{-3.3}$ Mpc. We use these constraints for comparison with the model-independent inferences throughout the paper.

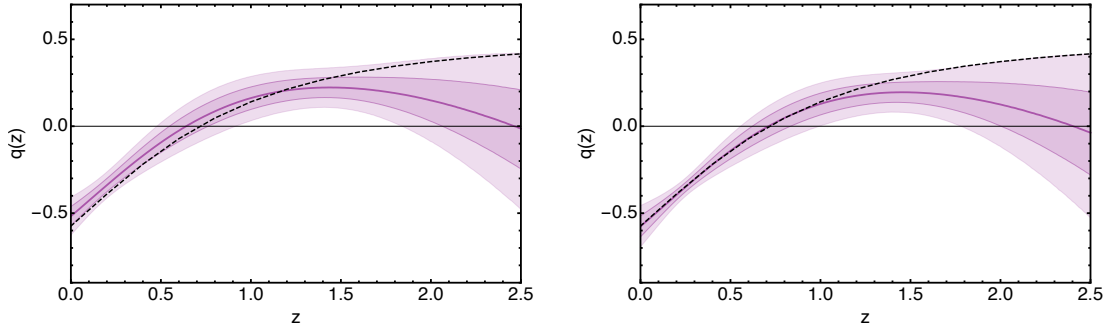


Figure 5. We show the reconstructions of $q(z)$ for the datasets CC + SN + BAO (left) and CC + SN + BAO + R18 (right). All the reconstructions shown here are obtained using $M_{7/2}$ kernel. The dashed line in the left and right panels corresponds to the best-fit Λ CDM model of $(\Omega_m, H_0) = (0.284, 69.77)$ and $(\Omega_m, H_0) = (0.275, 71.80)$, respectively.

In Figure 5 we show the reconstructions of $q(z)$ with (right) and without (left) inclusion of R18. We find that the $q(z)$ reconstructions are consistent with Λ CDM deceleration parameter, however only at upper boundary of 2σ reconstruction. In Figure 6 we show the $Om(z)$ (first row) and its derivative $Om'(z)$ (second row) reconstructions, which are very consistent with a constant Ω_m and flat $\Omega_\Lambda = 1 - \Omega_m$ scenario. However, towards redshifts $z \gtrsim 2$ the diagnostics seem to diverge from a constant $\Omega_m \gtrsim 0.3$, while being consistent with an Ω_m of lower value. In the last row of Figure 5 we show

the reconstructions of $Om^{(2)}(z)$, which shows slightly more deviation from a constant Ω_m scenario in comparison to the $Om(z)$ diagnostic. As can be seen from the $Om(z)$ and $Om^{(2)}(z)$ reconstructions the "high-redshift" CMB constraints on Λ CDM do not agree with these reconstructions even within 2σ at $z \gtrsim 2$. Clearly this mild discrepancy is amplified when R18 is included in the analysis (see right column of Figure 6), although now in terms of Ω_m . The inference of this discrepancy is in agreement with the very recent analysis by [86], and similar observations have also been made earlier in [7, 32]. The negative slope of $Om(z)$ (see also middle row of Figure 6) at higher redshifts tentatively indicates a quintessence like behaviour [49] of freezing kind as it tends to become a constant for $z \lesssim 1.5$, which is contrary to result in [7] and however, requires an in-depth analysis. It is also noteworthy that the general features of $Om(z)$ remain unaltered with the inclusion of R18, please see the right column of Figure 6.

Slowing down of the cosmic acceleration has been an issue of interest in several works [87–92]. In [87] this phenomenon was studied also utilising the $Om(z)$ diagnostic, which suggests that an increase in $Om(z)$ at lower redshift can indicate a slowing of acceleration. In [93, 94] it has been shown that galaxy cluster gas mass fraction samples and strong lensing datasets show such a slowing down of acceleration. In our current analysis and with the datasets implemented here, we do not find any hints for a slowing down of the cosmic acceleration with any of the covariance functions implemented. Clearly, no strong inferences can be drawn from these reconstructions yet, however we do anticipate better inferences with more stringent data to arrive with [95–97]. In particular, obtaining a few additional CC data points in the midrange redshifts $0.5 \leq z \leq 1.5$ could significantly improve the reconstructions.

Our current formalism suggests that the non-accelerating scenario is highly improbable from the estimated $q(z)$ reconstructions, which remains true for all the assumed covariance functions. Given the discussion led in [98–106] and the analysis presented here, the inference for a present accelerated phase, could in fact be claimed as beyond doubt. Also in this respect, we would like to comment on the inference for non-accelerating power-law models ($a(t) \propto t^n$ with $n \leq 1$) [26, 102, 104, 107–109] and linear-coasting-like models [26, 37, 110–114], which have been of particular interest in recent works. We find that with all the model-independent flexibility (even with the most flexible $M_{3/2}$) that can be assumed in the MTGP, these models are disfavoured. However, this inference stands only from the joint analysis of all the three datasets, as the analyses of independent datasets are plagued by their own disadvantages, as discussed in next section, and no strict inferences can be made in a model-independent way. More recently, in [4, 115], similar inference for a non-accelerating power-law model was made, that the high-redshift $E(z = 1.5)$ and CMB data disfavours such models.

To summarise, taking into account the predicted equivalence of different datasets, constraints on q_0 and z_T presented in the Table 5, constraints on H_0 and r_d in Table 3 and, predictive quality of the kernels, we infer $M_{7/2}$ kernel as better choice amongst all kernels utilised here, to quote the final constraints on the cosmologically relevant parameters. Clearly this is not a statistical inference, but in the formalism implemented here, the current dataset(s) is(are) unable to allow for a better inference, which as we mentioned earlier shall be a issue of future investigation.

4.3 Impact of assumptions in data and method

As discussed in the preceding sections, one of the key ingredients in the cosmic chronometers measurements to estimate the Hubble parameter is the stellar population synthesis (SPS) model used to calibrate the method. We recall here that a study of the impact of systematic effects due to the assumption of different SPS models has been already provided in [40, 116]. In [116], the constraints on the dimensionless energy density parameters were explored in an open w_0w_a CDM model (where the equation of state of dark energy is let free to vary as a function of cosmic time according to the

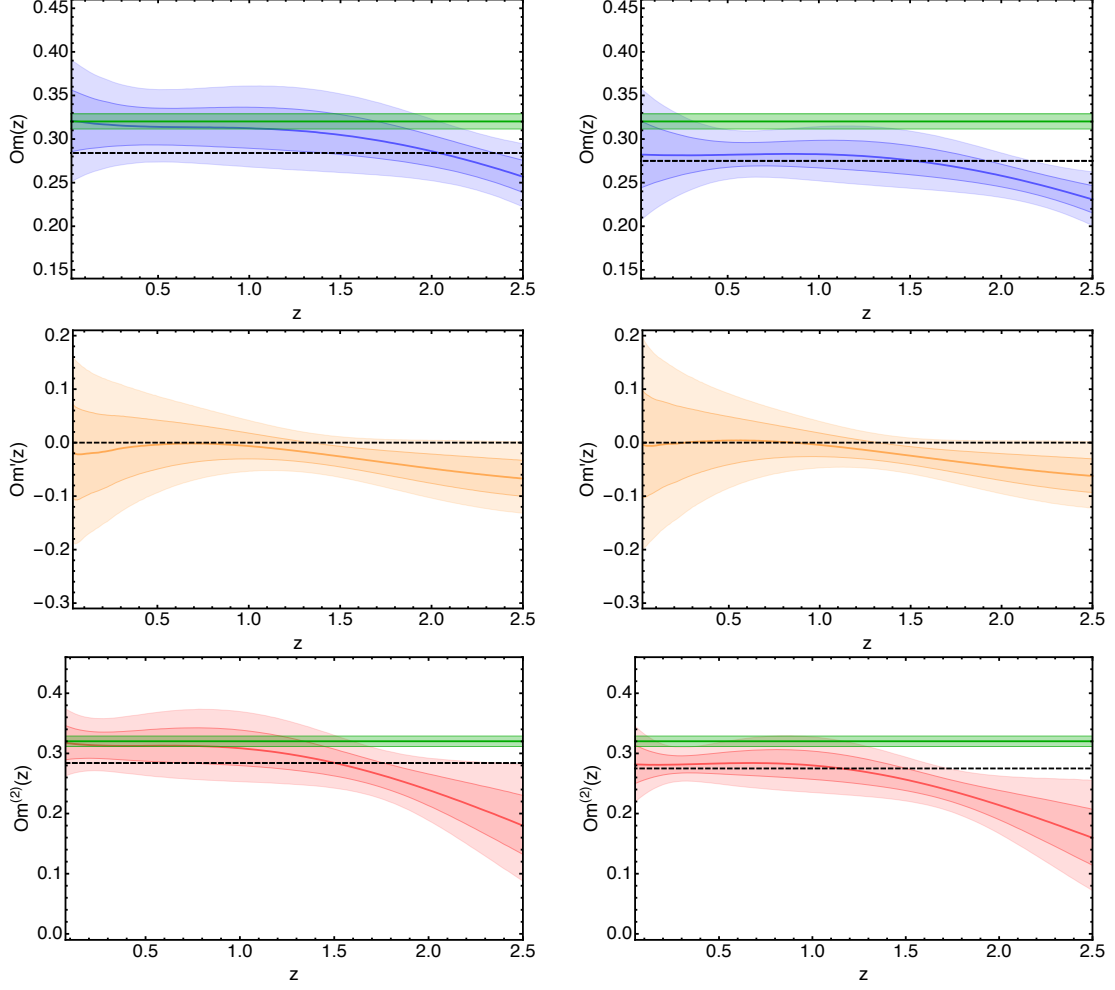


Figure 6. We show the reconstructions of $Om(z)$ (top), $Om'(z)$ (centre) and $Om^{(2)}(z)$ (bottom) for the datasets CC + SN + BAO (left) and CC + SN + BAO + R18 (right). All the reconstructions shown here are obtained using $M_{7/2}$ kernel. The dashed line in the top and bottom rows corresponds to the best-fit Λ CDM model of $\Omega_m = 0.284$ (left) and $\Omega_m = 0.275$ (right) to the current datasets. Similarly, the green region shows the constraint of $\Omega_m = 0.320 \pm 0.0087$ obtained from "high-redshift" CMB fit to Λ CDM model.

parametrisation given by [117, 118]), finding no statistical difference when assuming different SPS models. In [40] studying the flat $w_0 w_a$ CDM model a shift of $\sim 5 \text{ km/s Mpc}^{-1}$ is found in H_0 , with M11 models preferring slightly higher values, but with a difference which is compatible with zero, given the estimated errors. This difference, as shown in [40, 116] and also discussed more recently in [18], is due to the fact that expansion rate measurements from M11 models are slightly larger, thus impacting the estimated normalisation of the $H(z)$ relation, i.e. H_0 .

In parallel to this effect, when different combination of data are explored, they would probe different regions of the parameters space, yielding varied results. Here we present the analysis to evaluate the dependence of the results obtained using both different combinations of data, considering the CC_B and CC_M compilations provided in [40, 116] separately, to estimate the systematics in our approach both due to data and our MTGP method. For the comparison to be homogeneous, we consider in both cases only the datasets drawn from just these analyses, neglecting other measurements that have been obtained using only BC03 models. The results are shown in Figure 7, showing

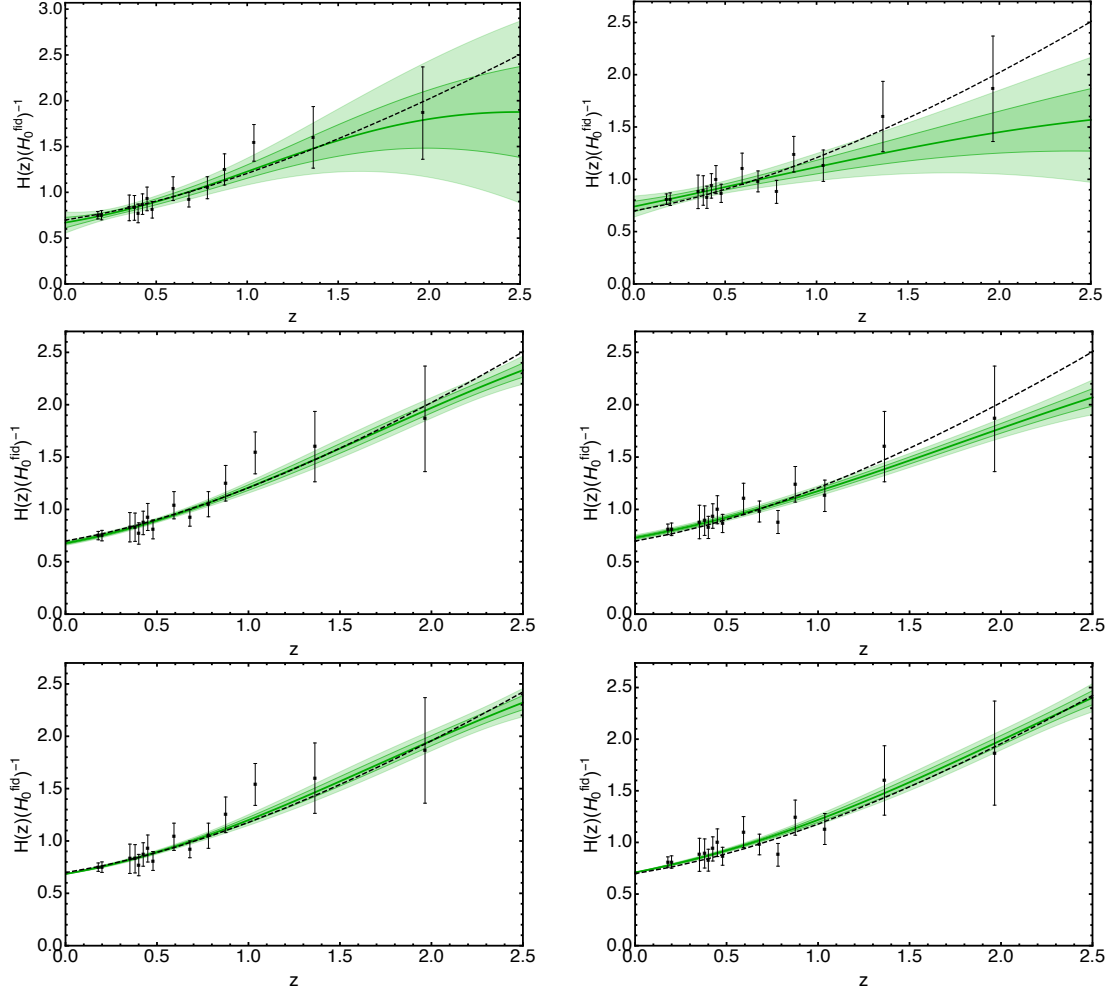


Figure 7. We show the reconstructions of the CC_B (left) and the CC_M (right) data using the Single-Task GP (Top) and MTGP (centre) of $CC_{B/M} + \text{SN} + \text{BAO}$, assuming $M_{7/2}$ kernel. In the bottom row we show the reanalysis of centre row, enforcing unique expansion histories across all the datasets. In all the panels data are rescaled with $H_0^{fid} = 100 \text{ km/s Mpc}^{-1}$ and the dashed line corresponds to ΛCDM with $\Omega_m = 0.284$, $H_0 = 69.77 \text{ km/s Mpc}^{-1}$.

the ST reconstruction of the data in top row with the corresponding H_0 values reported in Table 6. We find that CC_M data predicts a higher value of H_0 , consistent with the estimate from the local model-independent R18 by [4], while CC_B dataset, as discussed in Section 4.1, prefers a value compatible with the estimate from CMB [1]. However, the estimated error bars are large and we confirm that considering the two CC datasets (CC_M and CC_B) alone the measurement from the two SPS model are compatible within errors. This result confirms the findings of [18], who also found, with a different approach, a statistically insignificant difference when considering $CC_{M/B}$ datasets on their own.

The large uncertainty on H_0 is lowered when a joint analysis is performed, and in this case the differences are more significant. The increase in significance of the difference between the two measurements is mostly due, as demonstrated, by the inclusion of the external datasets of SN and BAO. This analysis indicates a tentative systematic error of $\sim 5.7 \text{ km/s Mpc}^{-1}$, on average. This value is estimated as the average of difference between the mean H_0 values in third and fourth column for $CC_{B/M} + \text{SN} + \text{BAO}$ in Table 6, and also includes the variations due to kernel assumptions. However,

Dataset(s)		CC_B	CC_M
	Kernel	H_0 [km/s Mpc ⁻¹]	
	SE	64.08 ± 4.78	73.40 ± 4.90
	$M_{9/2}$	66.19 ± 5.47	73.68 ± 5.00
	$M_{7/2}$	66.92 ± 5.75	73.88 ± 5.04
	$M_{5/2}$	67.56 ± 6.14	74.12 ± 5.18
+SN	SE	67.97 ± 1.01	71.92 ± 0.55
	$M_{9/2}$	68.14 ± 0.90	71.91 ± 0.58
	$M_{7/2}$	68.23 ± 0.82	71.88 ± 0.55
	$M_{5/2}$	68.37 ± 0.68	71.83 ± 0.49
+BAO	SE	62.38 ± 3.67	71.49 ± 3.62
	$M_{9/2}$	62.89 ± 3.79	71.37 ± 3.68
	$M_{7/2}$	62.57 ± 3.73	71.34 ± 3.74
	$M_{5/2}$	62.67 ± 3.75	71.32 ± 3.89
+SN+BAO	SE	66.44 ± 1.51	73.47 ± 1.26
	$M_{9/2}$	67.40 ± 1.20	73.08 ± 1.32
	$M_{7/2}$	67.72 ± 1.03	72.98 ± 1.36
	$M_{5/2}$	68.09 ± 0.73	72.84 ± 1.42
+SN+BAO+R18	SE	71.82 ± 0.63	73.27 ± 0.73
	$M_{9/2}$	71.72 ± 0.56	72.87 ± 0.73
	$M_{7/2}$	71.68 ± 0.53	73.20 ± 0.78
	$M_{5/2}$	71.63 ± 0.45	72.64 ± 0.54

Table 6. Estimates of H_0 and the corresponding 1σ uncertainty form combinations of the dataset(s) using either the CC_B or CC_M , implemented using all the kernels implemented in this work. For the joint analysis with the multiple datasets we implement the same kernel for each of them.

unlike $CC_B + SN + BAO$, we find that the $CC_M + SN + BAO$ combination is unable to provide equivalent expansion histories in all three data planes, even with the utmost flexibility available in the formalism (e.g. see middle row of Figure 7). This in turn devolves the purpose/utility of MTGP, to allow the flexibility/freedom and predict equivalence of data. To explore the reasons for the same we produce a spectrum of H_0 's coming from several data combinations as are shown in Table 6. This also helps us illustrate the subtleties of MTGP formalism listed in Section 2.

As can be seen in the second row of Table 6, the estimates of uncertainties using CC and SN data alone are much tighter than the uncertainties quoted in Table 3, with all three datasets. A more comprehensive way to describe these reconstructions is the over-fitting around $z \sim 0$. This over-fitting in fact is much more pronounced in the CC_M than in the CC_B , due to the difference in the two data points at $z = 0.781$ and $z = 1.037$. While all the data points in CC_M are higher by ~ 6 km/s Mpc⁻¹ w.r.t CC_B data points, at these two redshifts the CC_M measurements are lower than the CC_B and this in turn allows the CC_M data to fit more in concert with the SN data, however, giving rise to over-fitting. A clear indication of this is the constant-like behaviour of the H_0 estimate at intercept, which we have further confirmed by eliminating the first two data points of the CC_M dataset to find an ever more over-fitted uncertainty of 0.2. While CC_B also tends to be over-powered by the SN dataset the effect is not as pronounced as is with CC_M . In the top row of Figure 8, we show the reconstructions using the combinations of CC_M , CC_B with SN data. It is illustrative to see the effect of SN data, which clearly improves the CC reconstruction around $z \sim 0$ than at $z \gtrsim 1.5$. Finally we infer that this effect also takes its toll in the analysis with all three datasets, as this remains to be the clear/only

difference why the data combinations with CC_M are unable to predict equivalent expansion histories.

Therefore to estimate the systematics, we aid to the aforementioned provision within MTGP formalism which enforces the equivalence of latent functions by a priori assuming that the length-scale hyperparameters of all three datasets are the same. The SN and BAO data complement $CC_{M/B}$ data and help predicting equivalent expansion histories, hence ensuring that the difference in the intercept is only due to modification from CC_B to CC_M compilations, while all the other features are primarily driven by SN + BAO datasets (see bottom row of Figure 7). As is already mentioned in the beginning of the results section, this provides a constant like rescaling factor(s) amongst datasets. Having established that these values cannot be inferred as physical constants, we are only interested in the difference in the mean estimates from combinations of CC_M to CC_B . We then infer this difference as the systematic effect arising from the CC_B to CC_M comparison. Using the $M_{7/2}$ kernel we estimate $\sigma_{H_0}(sys) = 2.51 \text{ km/s Mpc}^{-1}$ as the difference in these constant like intercepts which are 70.97 and 68.46 for $CC_M + \text{SN} + \text{BAO}$ and $CC_B + \text{SN} + \text{BAO}$, respectively. An added advantage in this approach is that the difference remains unaltered (with a maximum variation of ~ 0.01) for all assumed kernels, also completely mitigating the variations arising due to different kernel choices. While on the face value the estimates of H_0 from CC_M seem discrepant, through more scrutiny we find consistent results with [18, 29]. However with our slightly larger systematic error, we do not find R18 to be in tension. We refer to a following analysis for a more detailed discussion on the systematic impact on $H(z)$ data due to SPS model assumption in cosmic chronometers. As is known, r_d is estimated as a rescaling factor between BAO and CC and so the systematic in the H_0 estimate also propagates to r_d . We replicate the same formalism to find a systematic error of $\sigma_{r_d}(sys) = 4.3 \text{ Mpc}$ and $\sigma_{r_d H_0}(sys) = 70 \text{ km/s}$, which are also independent of kernel choice. Note that the propagation of the systematics to the $r_d H_0$ estimate are minimal in comparison to the statistical error $\sigma_{r_d H_0}(stat) = 237 \text{ km/s}$ quoted in Section 4.1.

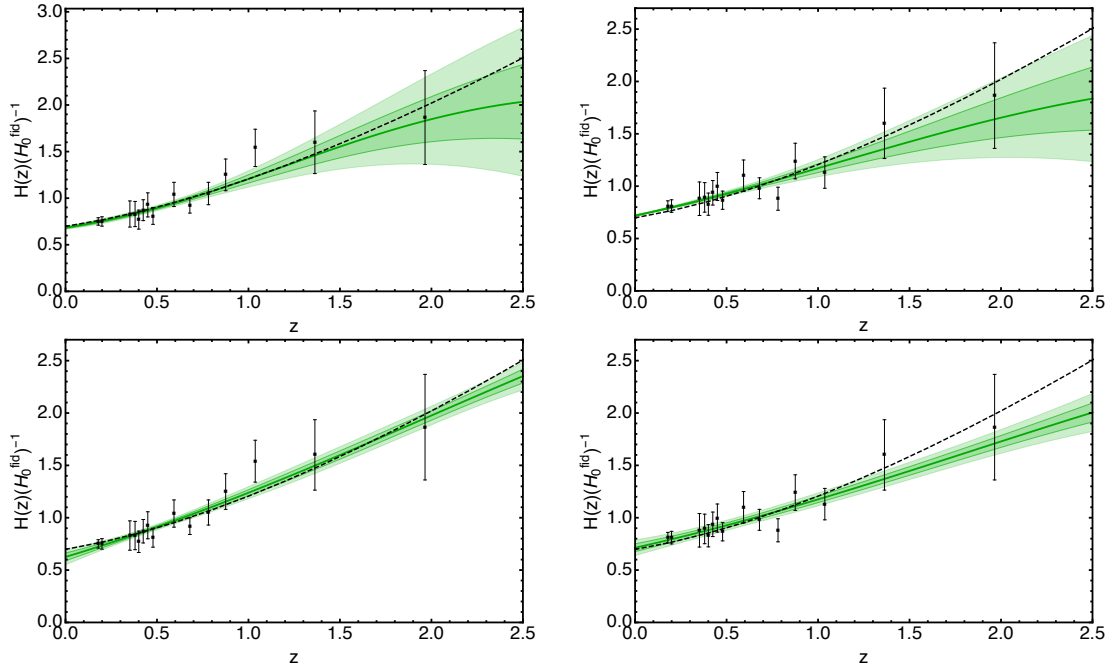


Figure 8. We show the reconstructions of the CC_B (left) and the CC_M (right) data in combination with SN (top) and BAO (bottom) using $M_{7/2}$ kernel. In all the panels data are rescaled with $H_0^{fid} = 100 \text{ km/s Mpc}^{-1}$ and the dashed line corresponds to ΛCDM with $\Omega_m = 0.284$, $H_0 = 69.77 \text{ km/s Mpc}^{-1}$.

In fact when a model-dependent analysis is performed assuming the Λ CDM model, the discrepancy between the H_0 estimates using the CC_M and CC_B compilations separately, reduces in scatter. We find $H_0 = 68.99 \pm 2.04$ km/s Mpc $^{-1}$ and $H_0 = 71.60 \pm 2.30$ km/s Mpc $^{-1}$ for $CC_B + \text{SN} + \text{BAO}$ and $CC_M + \text{SN} + \text{BAO}$, respectively. These values are clearly in no discrepancy among themselves and also very well agree with our best estimate of $H_0 = 68.52^{+0.94+2.51(\text{sys})}_{-0.94}$ km/s Mpc $^{-1}$. We later extend the analysis to estimate the systematic uncertainty even when R18 is included. We find the kernel-independent uncertainties arising from the systematic effect to be $\sigma_{H_0}(\text{sys}) = 1.65$ km/s Mpc $^{-1}$ and $\sigma_{r_d}(\text{sys}) = 2.63$ Mpc. We quote the constraints for this combination of data with the additional systematic error for the corresponding $M_{7/2}$ kernel estimates reported in Table 3.

To illustrate a few more subtleties, in the bottom row of Figure 8 we show the features of BAO data being transferred to both the CC datasets. Clearly the H_0 estimates are lowered due to the linear nature of BAO reconstruction (see third row of Table 6). Notice the agreement in the mean of H_0 estimates from $CC_M + \text{SN}$ and $CC_M + \text{BAO}$ datasets, while the uncertainty is over-fitted in the prior case. The negative transfer from the BAO data is more pronounced in the $CC_B + \text{BAO}$ combination, where the H_0 estimate is further lowered to ~ 62 km/s Mpc $^{-1}$ in comparison to $CC + \text{BAO}$ (~ 64 km/s Mpc $^{-1}$). Given the strength of all the datasets considered in this work, an analysis on the combination of all the three datasets clearly remains necessary for not so trivial reasons, when compared to 2 dataset combinations. One could even suspect that in an analyses utilising only two datasets [18, 29], this could take effect as either of the combinations $CC+\text{SN}$ or $CC+\text{BAO}$ were implemented, although a straightforward comparison is not available.

5 Conclusions and Future prospects

In the present work, we have implemented an extension (MTGP) to the standard “model-independent” GP formalism utilised so far in the cosmological context. The key improvement in this work is to formulate a methodology to combine different datasets to perform joint analysis in a more cosmology-independent way. To this aid, we have utilised the “low-redshift” expansion rate data from SN, BAO and CC observations. Implementing several covariance functions in the MTGP formalism we find an excellent concert amongst them. This, in turn, helps our discussion to override the specificity of choosing a particular covariance function. We have performed several tests, to assess the better kernel assumptions for the cosmological data implemented here, but conclude that it requires a more in-depth investigation and in fact not prospective with the current dataset(s). However, this formalism with its inherent flexibility readily provides us with a heuristic argument to perform a kernel selection, which allows us to infer $M_{7/2}$ as the better kernel choice with the data utilised in this work.

Our model-independent estimates of $H_0 = 68.52^{+0.94+2.51(\text{sys})}_{-0.94}$ km/s Mpc $^{-1}$ is in a good agreement with standard scenario with corresponding $r_d = 145.61^{+2.82}_{-2.82-4.3(\text{sys})}$ Mpc. While our statistical uncertainties using $CC + \text{SN} + \text{BAO}$ are highly competitive, even amongst the existing model-independent analysis, we also included a systematic error specifically utilising the provision in our method and the differences in $CC_{B/M}$ compilations. The statistical estimate obtained here is in agreement with the CMB P16 and local R18 at 1.41σ and 2.60σ , respectively. Very similar inference was also made with a model-dependent analysis in [80]. When we include the R18 measurement in our analysis we find $H_0 = 71.40^{+0.30+1.65(\text{sys})}_{-0.30}$ km/s Mpc $^{-1}$ and $r_d = 141.29^{+1.31}_{-1.31-2.63(\text{sys})}$ Mpc. Amongst several advantages of the current formalism, estimation of the systematic error remains to be one. While our statistical uncertainties remain in excellent concert with the recent analyses in [18, 29], the systematic uncertainty incurred from our analysis does not show local H_0 estimate to be in significant tension at $\sim 1.61\sigma$. Our constraint on $r_d H_0 = 9977 \pm 237(\text{stat})$ km/s is extremely consistent with “high-redshift” CMB and the model-dependent Λ CDM constraints for the datasets implemented here.

The effect of systematics in CC data on this value ($\sigma_{r_d H_0}(\text{sys}) = 70 \text{ km/s}$) is minimal compared to the statistical error.

We have also extended the analysis to reconstruct the derivative regions of the data, through which we find a good agreement with the Λ CDM model up to redshifts $z \lesssim 2.0$. We estimate the deceleration parameter $q(z)$, inferring constraints of $q_0 = -0.52 \pm 0.06$ and $z_T = 0.64^{+0.12}_{-0.09}$, consistent amongst all the kernels implemented here. Utilising the $Om(z)$ diagnostic we find good agreement with Λ CDM within $z \lesssim 2.0$ and a mild discrepancy for $z \gtrsim 2.0$. The inclusion of R18 into the analysis, as expected, magnifies this discrepancy. We find these inferences to be very much in agreement with earlier results in [7, 32, 86]. Work is in progress to extend the current discussion by utilising also the distances and growth rate data, which shall also be implemented to investigate the dark energy EoS in the current model-independent approach. The ability to perform a joint analysis as presented here shall prove extremely useful with more stringent data to arrive [95–97].

Acknowledgments

B.S.H, V.V.L and N.V acknowledge financial support by ASI Grant No. 2016-24-H.0. M.M acknowledges the grants ASI n.I/023/12/0 “Attività relative alla fase B2/C per la missione Euclid” and PRIN MIUR 2015 “Cosmology and Fundamental Physics: illuminating the Dark Universe with Euclid”. B.S.H thanks Federico Tosone for useful discussions and comments.

References

- [1] Planck Collaboration, N. Aghanim, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini et al., *Planck intermediate results. xlv. reduction of large-scale systematic effects in hfi polarization maps and estimation of the reionization optical depth*, *Astron. Astrophys.* **596** (Dec., 2016) A107, [[1605.02985](#)].
- [2] S. Alam, M. Ata, S. Bailey, F. Beutler, D. Bizyaev, J. A. Blazek et al., *The clustering of galaxies in the completed sdss-iii baryon oscillation spectroscopic survey: cosmological analysis of the dr12 galaxy sample*, *arXiv preprint arXiv:1607.03155* (2016) .
- [3] S. Joudaki, A. Mead, C. Blake, A. Choi, J. de Jong, T. Erben et al., *Kids-450: testing extensions to the standard cosmological model*, *Mon. Not. Roy. Astron. Soc.* **471** (Oct., 2017) 1259–1279, [[1610.04606](#)].
- [4] A. G. Riess, S. A. Rodney, D. M. Scolnic, D. L. Shafer, L.-G. Strolger, H. C. Ferguson et al., *Type Ia Supernova Distances at Redshift > 1.5 from the Hubble Space Telescope Multi-cycle Treasury Programs: The Early Expansion Rate*, *Astrophys. J.* **853** (Feb., 2018) 126, [[1710.00844](#)].
- [5] M. J. Mortonson, W. Hu and D. Huterer, *Simultaneous falsification of Λ cdm and quintessence with massive, distant clusters*, *Phys. Rev. D* **83** (Jan., 2011) 023015, [[1011.0004](#)].
- [6] C.-J. Feng and X.-Z. Li, *Probing the Expansion History of the Universe by Model-independent Reconstruction from Supernovae and Gamma-Ray Burst Measurements*, *Astrophys. J.* **821** (Apr., 2016) 30, [[1604.01930](#)].
- [7] G.-B. Zhao, M. Raveri, L. Pogosian, Y. Wang, R. G. Crittenden, W. J. Handley et al., *Dynamical dark energy in light of the latest observations*, *Nature Astronomy* **1** (Sept., 2017) 627–632, [[1701.08165](#)].
- [8] V. Miranda and C. Dvorkin, *Model-independent predictions for smooth cosmic acceleration scenarios*, *ArXiv e-prints* (Dec., 2017) , [[1712.04289](#)].
- [9] M. J. Mortonson and W. Hu, *Observational limits on patchy reionization: Implications for b modes*, *Phys. Rev. D* **81** (Mar., 2010) 067302, [[1001.4803](#)].
- [10] M. J. Mortonson, W. Hu and D. Huterer, *Falsifying paradigms for cosmic acceleration*, *Phys. Rev. D* **79** (Jan., 2009) 023004, [[0810.1744](#)].
- [11] R. A. Vanderveld, M. J. Mortonson, W. Hu and T. Eifler, *Testing dark energy paradigms with weak gravitational lensing*, *Phys. Rev. D* **85** (May, 2012) 103518, [[1203.3195](#)].
- [12] A. Shafieloo, *Crossing statistic: reconstructing the expansion history of the universe*, *J. Cosmology Astropart. Phys.* **8** (Aug., 2012) 002, [[1204.1109](#)].
- [13] A. Shafieloo, U. Alam, V. Sahni and A. A. Starobinsky, *Smoothing supernova data to reconstruct the expansion history of the universe and its age*, *Mon. Not. Roy. Astron. Soc.* **366** (Mar., 2006) 1081–1095, [[astro-ph/0505329](#)].
- [14] B. L’Huillier, A. Shafieloo and H. Kim, *Model-independent cosmological constraints from growth and expansion*, *ArXiv e-prints* (Dec., 2017) , [[1712.04865](#)].
- [15] C. Cattoën and M. Visser, *Cosmographic Hubble fits to the supernova data*, *Phys. Rev. D* **78** (Sept., 2008) 063501, [[0809.0537](#)].
- [16] E. Mörtzell and C. Clarkson, *Model independent constraints on the cosmological expansion rate*, *J. Cosmology Astropart. Phys.* **1** (Jan., 2009) 044, [[0811.0981](#)].
- [17] C. Gruber and O. Luongo, *Cosmographic analysis of the equation of state of the universe through padé approximations*, *Phys. Rev. D* **89** (May, 2014) 103506, [[1309.3215](#)].
- [18] A. Gómez-Valent and L. Amendola, *H_0 from cosmic chronometers and Type Ia supernovae, with Gaussian Processes and the novel Weighted Polynomial Regression method*, *ArXiv e-prints* (Feb., 2018) , [[1802.01505](#)].

- [19] S. Capozziello, R. D’Agostino and O. Luongo, *Cosmographic analysis with chebyshev polynomials*, *Mon. Not. Roy. Astron. Soc.* **476** (May, 2018) 3924–3938, [[1712.04380](#)].
- [20] M. Seikel, C. Clarkson and M. Smith, *Reconstruction of dark energy and expansion dynamics using gaussian processes*, *J. Cosmology Astropart. Phys.* **6** (June, 2012) 036, [[1204.2832](#)].
- [21] M. Seikel and C. Clarkson, *Optimising gaussian processes for reconstructing dark energy dynamics from supernovae*, *ArXiv e-prints* (Nov., 2013) , [[1311.6678](#)].
- [22] A. Shafieloo, A. G. Kim and E. V. Linder, *Gaussian process cosmography*, *Phys. Rev. D* **85** (June, 2012) 123530, [[1204.2272](#)].
- [23] S. D. P. Vitenti and M. Penna-Lima, *A general reconstruction of the recent expansion history of the universe*, *J. Cosmology Astropart. Phys.* **9** (Sept., 2015) 045, [[1505.01883](#)].
- [24] M. Seikel, C. Clarkson and M. Smith, “Gapp: Gaussian processes in python.” Astrophysics Source Code Library, Mar., 2013.
- [25] M.-J. Zhang and J.-Q. Xia, *Test of the cosmic evolution using gaussian processes*, *J. Cosmology Astropart. Phys.* **12** (Dec., 2016) 005, [[1606.04398](#)].
- [26] M. Bilicki and M. Seikel, *We do not live in the $R_h = ct$ universe*, *Mon. Not. Roy. Astron. Soc.* **425** (Sept., 2012) 1664–1668, [[1206.5130](#)].
- [27] V. C. Busti, C. Clarkson and M. Seikel, *The value of h_0 from gaussian processes*, in *Statistical Challenges in 21st Century Cosmology* (A. Heavens, J.-L. Starck and A. Krone-Martins, eds.), vol. 306 of *IAU Symposium*, pp. 25–27, May, 2014, [1407.5227](#), DOI.
- [28] V. C. Busti, C. Clarkson and M. Seikel, *Evidence for a lower value for h_0 from cosmic chronometers data?*, *Mon. Not. Roy. Astron. Soc.* **441** (June, 2014) L11–L15, [[1402.5429](#)].
- [29] H. Yu, B. Ratra and F.-Y. Wang, *Hubble parameter and baryon acoustic oscillation measurement constraints on the hubble constant, the deviation from the spatially-flat λ cdm model, the deceleration-acceleration transition redshift, and spatial curvature*, *ArXiv e-prints* (Nov., 2017) , [[1711.03437](#)].
- [30] J.-J. Wei and X.-F. Wu, *An Improved Method to Measure the Cosmic Curvature*, *Astrophys. J.* **838** (Apr., 2017) 160, [[1611.00904](#)].
- [31] M. Seikel, S. Yahya, R. Maartens and C. Clarkson, *Using $h(z)$ data as a probe of the concordance model*, *Phys. Rev. D* **86** (Oct., 2012) 083001, [[1205.3431](#)].
- [32] V. Sahni, A. Shafieloo and A. A. Starobinsky, *Model-independent Evidence for Dark Energy Evolution from Baryon Acoustic Oscillations*, *Astrophys. J. Lett* **793** (Oct., 2014) L40, [[1406.2209](#)].
- [33] R. Nair, S. Jhingan and D. Jain, *Exploring scalar field dynamics with Gaussian processes*, *J. Cosmology Astropart. Phys.* **1** (Jan., 2014) 005, [[1306.0606](#)].
- [34] T. Holsclaw, U. Alam, B. Sansó, H. Lee, K. Heitmann, S. Habib et al., *Nonparametric dark energy reconstruction from supernova data*, *Physical Review Letters* **105** (Dec., 2010) 241302, [[1011.3079](#)].
- [35] T. Holsclaw, U. Alam, B. Sansó, H. Lee, K. Heitmann, S. Habib et al., *Nonparametric reconstruction of the dark energy equation of state from diverse data sets*, *Phys. Rev. D* **84** (Oct., 2011) 083501, [[1104.2041](#)].
- [36] M. K. Yennapureddy and F. Melia, *Reconstruction of the hii galaxy hubble diagram using gaussian processes*, *J. Cosmology Astropart. Phys.* **11** (Nov., 2017) 029, [[1711.03454](#)].
- [37] F. Melia and M. K. Yennapureddy, *Model selection using cosmic chronometers with gaussian processes*, *ArXiv e-prints* (Feb., 2018) , [[1802.02255](#)].
- [38] A. Marta Pinho, S. Casas and L. Amendola, *Model-independent reconstruction of the linear anisotropic stress η* , *ArXiv e-prints* (Apr., 2018) , [[1805.00027](#)].

- [39] R. Jimenez and A. Loeb, *Constraining Cosmological Parameters Based on Relative Galaxy Ages*, *Astrophys. J.* **573** (July, 2002) 37–42, [[astro-ph/0106145](#)].
- [40] M. Moresco, L. Pozzetti, A. Cimatti, R. Jimenez, C. Maraston, L. Verde et al., *A 6evidence of the epoch of cosmic re-acceleration*, *J. Cosmology Astropart. Phys.* (2016) , [[1601.01701v2](#)].
- [41] D. J. Eisenstein, I. Zehavi, D. W. Hogg, R. Scoccimarro, M. R. Blanton, R. C. Nichol et al., *Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies*, *Astrophys. J.* **633** (Nov., 2005) 560–574, [[astro-ph/0501171](#)].
- [42] R. Caruana, *Multitask learning*, in *Learning to learn*, pp. 95–133. Springer, 1998.
- [43] E. V. Bonilla, F. V. Agakov and C. K. Williams, *Kernel multi-task learning using task-specific features*, in *Artificial Intelligence and Statistics*, pp. 43–50, 2007.
- [44] E. V. Bonilla, K. M. Chai and C. Williams, *Multi-task gaussian process prediction*, in *Advances in neural information processing systems*, pp. 153–160, 2008.
- [45] A. G. Riess, S. Casertano, W. Yuan, L. Macri, J. Anderson, J. W. MacKenty et al., *New parallaxes of galactic cepheids from spatially scanning the hubble space telescope: Implications for the hubble constant*, *Astrophys. J.* **855** (Mar., 2018) 136, [[1801.01120](#)].
- [46] A. B. Abdessalem, N. Dervilis, D. J. Wagg and K. Worden, *Automatic kernel selection for gaussian processes regression with approximate bayesian computation and sequential monte carlo*, *Frontiers in Built Environment* **3** (2017) 52.
- [47] J. P. de Leon, *An analytical model for the transition from decelerated to accelerated cosmic expansion*, *International Journal of Modern Physics D* **15** (2006) 1237–1257, [[gr-qc/0511150](#)].
- [48] J. A. S. Lima, J. F. Jesus, R. C. Santos and M. S. S. Gill, *Is the transition redshift a new cosmological number?*, *ArXiv e-prints* (May, 2012) , [[1205.4688](#)].
- [49] V. Sahni, A. Shafieloo and A. A. Starobinsky, *Two new diagnostics of dark energy*, *Phys. Rev. D* **78** (Nov., 2008) 103502, [[0807.3548](#)].
- [50] C. Zunckel and C. Clarkson, *Consistency tests for the cosmological constant*, *Physical Review Letters* **101** (Oct., 2008) 181301, [[0807.4304](#)].
- [51] M. Moresco, A. Cimatti, R. Jimenez, L. Pozzetti, G. Zamorani, M. Bolzonella et al., *Improved constraints on the expansion rate of the Universe up to $z \sim 1.1$ from the spectroscopic evolution of cosmic chronometers*, *J. Cosmology Astropart. Phys.* **8** (Aug., 2012) 006, [[1201.3609](#)].
- [52] M. Moresco, *Raising the bar: new constraints on the hubble parameter with cosmic chronometers at $z \sim 2$* , *Mon. Not. Roy. Astron. Soc.* **450** (June, 2015) L16–L20, [[1503.01116](#)].
- [53] C. E. Rasmussen and C. K. Williams, *Gaussian processes for machine learning*. 2006, *The MIT Press*, Cambridge, MA, USA **38** (2006) 715–719.
- [54] A. McHutchon and C. E. Rasmussen, *Gaussian process training with input noise*, in *Advances in Neural Information Processing Systems*, pp. 1341–1349, 2011.
- [55] B. Blight and L. Ott, *A bayesian approach to model inadequacy for polynomial regression*, *Biometrika* **62** (1975) 79–88.
- [56] A. O’Hagan and J. Kingman, *Curve fitting and optimal design for prediction*, *Journal of the Royal Statistical Society. Series B (Methodological)* (1978) 1–42.
- [57] C. K. Williams and C. E. Rasmussen, *Gaussian processes for machine learning*, *the MIT Press* **2** (2006) 4.
- [58] E. Solak, R. Murray-Smith, W. E. Leithead, D. J. Leith and C. E. Rasmussen, *Derivative observations in gaussian process models of dynamic systems*, in *Advances in neural information processing systems*, pp. 1057–1064, 2003.

- [59] D. Wang and X.-H. Meng, *Improved constraints on the dark energy equation of state using gaussian processes*, *Phys. Rev. D* **95** (Jan., 2017) 023508, [[1708.07750](#)].
- [60] A. Melkumyan and F. Ramos, *Multi-kernel gaussian processes*, in *IJCAI Proceedings-International Joint Conference on Artificial Intelligence*, vol. 22, p. 1408, 2011.
- [61] S. Vasudevan, *Data fusion with gaussian processes*, *Robotics and Autonomous Systems* **60** (2012) 1528–1544.
- [62] K. M. Chai, *Multi-task learning with gaussian processes*, .
- [63] G. Leen, J. Peltonen and S. Kaski, *Focused multi-task learning in a gaussian process framework*, *Machine learning* **89** (2012) 157–182.
- [64] G. Lee, E. Yang and S. Hwang, *Asymmetric multi-task learning based on task relatedness and loss*, in *International Conference on Machine Learning*, pp. 230–238, 2016.
- [65] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.
- [66] G.-B. Zhao, Y. Wang, S. Saito, D. Wang, A. J. Ross, F. Beutler et al., *The clustering of galaxies in the completed sdss-iii baryon oscillation spectroscopic survey: tomographic bao analysis of dr12 combined sample in fourier space*, *Monthly Notices of the Royal Astronomical Society* **466** (2016) 762–779.
- [67] J. Wang, F. Wang, K. Cheng and Z. Dai, *Measuring dark energy with the eiso–ep correlation of gamma-ray bursts using model-independent methods*, *Astron. Astrophys.* **585** (2016) A68.
- [68] J. E. Bautista, N. G. Busca, J. Guy, J. Rich, M. Blomqvist, H. du Mas des Bourboux et al., *Measurement of baryon acoustic oscillation correlations at $z = 2.3$ with sdss dr12 $\text{Ly}\alpha$ -forests*, *Astron. Astrophys.* **603** (June, 2017) A12, [[1702.00176](#)].
- [69] H. du Mas des Bourboux, J.-M. L. Goff, M. Blomqvist, N. G. Busca, J. Guy, J. Rich et al., *Baryon acoustic oscillations from the complete sdss-iii $\text{Ly}\alpha$ -quasar cross-correlation function at $z = 2.4$* , [1708.02225v1](#).
- [70] G.-B. Zhao, Y. Wang, S. Saito, H. Gil-Marín, W. J. Percival, D. Wang 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*, *ArXiv e-prints* (Jan., 2018) , [[1801.03043](#)].
- [71] J. Simon, L. Verde and R. Jimenez, *Constraints on the redshift dependence of the dark energy potential*, *Phys. Rev. D* **71** (June, 2005) 123001, [[astro-ph/0412269](#)].
- [72] D. Stern, R. Jimenez, L. Verde, S. A. Stanford and M. Kamionkowski, *Cosmic chronometers: Constraining the equation of state of dark energy. ii. a spectroscopic catalog of red galaxies in galaxy clusters*, *Astrophys. J. S* **188** (May, 2010) 280–289, [[0907.3152](#)].
- [73] C. Zhang, H. Zhang, S. Yuan, S. Liu, T.-J. Zhang and Y.-C. Sun, *Four new observational $h(z)$ data from luminous red galaxies in the sloan digital sky survey data release seven*, *Research in Astronomy and Astrophysics* **14** (Oct., 2014) 1221–1233, [[1207.4541](#)].
- [74] A. L. Ratsimbazafy, S. I. Loubser, S. M. Crawford, C. M. Cress, B. A. Bassett, R. C. Nichol et al., *Age-dating luminous red galaxies observed with the southern african large telescope*, *Mon. Not. Roy. Astron. Soc.* **467** (May, 2017) 3239–3254, [[1702.00418](#)].
- [75] M. Moresco, R. Jimenez, L. Verde, L. Pozzetti, A. Cimatti and A. Citro, *Setting the stage for cosmic chronometers i. minimizing frosting with an optimized selection of cosmic chronometers*, *ArXiv e-prints* (Apr., 2018) , [[1804.05864](#)].
- [76] O. Farooq, F. R. Madiyar, S. Crandall and B. Ratra, *Hubble parameter measurement constraints on the redshift of the deceleration-acceleration transition, dynamical dark energy, and space curvature*, *Astrophys. J.* **835** (2017) 26, [[1607.03537](#)].

- [77] G. Chen and B. Ratra, *Median Statistics and the Hubble Constant*, *PASP* **123** (Sept., 2011) 1127–1132, [[1105.5206](#)].
- [78] W. Lin and M. Ishak, *Cosmological discordances. ii. hubble constant, planck and large-scale-structure data sets*, *Phys. Rev. D* **96** (Oct., 2017) 083532, [[1708.09813](#)].
- [79] C. Cheng and Q. Huang, *An accurate determination of the Hubble constant from baryon acoustic oscillation datasets*, *Science China Physics, Mechanics, and Astronomy* **58** (Sept., 2015) 095684, [[1409.6119](#)].
- [80] B. S. Haridasu, V. V. Luković and N. Vittorio, *Isotropic vs. anisotropic components of bao data: a tool for model selection*, *ArXiv e-prints* (Nov., 2017) , [[1711.03929v2](#)].
- [81] V. V. Luković, B. S. Haridasu and N. Vittorio, *Cosmological constraints from low-redshift data*, *ArXiv e-prints* (Jan., 2018) , [[1801.05765](#)].
- [82] P. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi et al., *Planck 2015 results-xiii. cosmological parameters*, *Astron. Astrophys.* **594** (2016) A13.
- [83] A. G. Kim, R. C. Thomas, G. Aldering, P. Antilogus, C. Aragon, S. Bailey et al., *Standardizing Type Ia Supernova Absolute Magnitudes Using Gaussian Process Data Regression*, *Astrophys. J.* **766** (Apr., 2013) 84, [[1302.2925](#)].
- [84] S. Dhawan, S. W. Jha and B. Leibundgut, *Measuring the Hubble constant with Type Ia supernovae as near-infrared standard candles*, *Astron. Astrophys.* **609** (Jan., 2018) A72, [[1707.00715](#)].
- [85] J. Guillochon, M. Nicholl, V. A. Villar, B. Mockler, G. Narayan, K. S. Mandel et al., *MOSFiT: Modular Open-Source Fitter for Transients*, *ArXiv e-prints* (Oct., 2017) , [[1710.02145](#)].
- [86] A. Shafieloo, B. L’Huillier and A. A. Starobinsky, *Falsifying Λ CDM: Model-independent tests of the concordance model with eBOSS DR14Q and Pantheon*, *ArXiv e-prints* (Apr., 2018) , [[1804.04320](#)].
- [87] A. Shafieloo, V. Sahni and A. A. Starobinsky, *Is cosmic acceleration slowing down?*, *Phys. Rev. D* **80** (Nov., 2009) 101301, [[0903.5141](#)].
- [88] A. Shafieloo, V. Sahni and A. A. Starobinsky, *Tentative evidence for slowing down of cosmic acceleration from recent small redshift supernovae and bao data*, in *American Institute of Physics Conference Series* (J.-M. Alimi and A. Fuözfa, eds.), vol. 1241 of *American Institute of Physics Conference Series*, pp. 294–302, June, 2010, DOI.
- [89] M. Shahalam, S. Sami and A. Agarwal, *Om diagnostic applied to scalar field models and slowing down of cosmic acceleration*, *Mon. Not. Roy. Astron. Soc.* **448** (Apr., 2015) 2948–2959, [[1501.04047](#)].
- [90] S. Wang, Y. Hu, M. Li and N. Li, *A Comprehensive Investigation on the Slowing Down of Cosmic Acceleration*, *Astrophys. J.* **821** (Apr., 2016) 60, [[1509.03461](#)].
- [91] M.-J. Zhang and J.-Q. Xia, *Physical condition for the slowing down of cosmic acceleration*, *Nuclear Physics B* **929** (Apr., 2018) 438–451, [[1701.04973](#)].
- [92] A. Bonilla and J. Castillo, *Constraints on dark energy models from galaxy clusters and gravitational lensing data*, *Universe* **4** (Jan., 2018) 21, [[1711.09291](#)].
- [93] J. Magaña, V. H. Cárdenas and V. Motta, *Cosmic slowing down of acceleration for several dark energy parametrizations*, *J. Cosmology Astropart. Phys.* **10** (Oct., 2014) 017, [[1407.1632](#)].
- [94] J. Magaña, V. Motta, V. H. Cárdenas and G. Foëx, *Testing cosmic acceleration for $w(z)$ parametrizations using f_{gas} measurements in galaxy clusters*, *Mon. Not. Roy. Astron. Soc.* **469** (July, 2017) 47–61, [[1703.08521](#)].
- [95] R. Laureijs, J. Amiaux, S. Arduini, J. . Auguères, J. Brinchmann, R. Cole et al., *Euclid definition study report*, *ArXiv e-prints* (Oct., 2011) , [[1110.3193](#)].
- [96] D. Collaboration, A. Aghamousa, J. Aguilar, S. Ahlen, S. Alam, L. E. Allen et al., *The desi experiment part ii: Instrument design*, *ArXiv e-prints* (Oct., 2016) , [[1611.00037](#)].

- [97] D. Collaboration, A. Aghamousa, J. Aguilar, S. Ahlen, S. Alam, L. E. Allen et al., *The desi experiment part i: Science, targeting, and survey design*, *ArXiv e-prints* (Oct., 2016) , [[1611.00036](#)].
- [98] J. T. Nielsen, A. Guffanti and S. Sarkar, *Marginal evidence for cosmic acceleration from type ia supernovae*, *Scientific Reports* **6** (Oct., 2016) 35596, [[1506.01354](#)].
- [99] H. Shariff, X. Jiao, R. Trotta and D. A. van Dyk, *BAHAMAS: New Analysis of Type Ia Supernovae Reveals Inconsistencies with Standard Cosmology*, *Astrophys. J.* **827** (Aug., 2016) 1, [[1510.05954](#)].
- [100] D. Rubin and B. Hayden, *Is the Expansion of the Universe Accelerating? All Signs Point to Yes*, *Astrophys. J. Lett* **833** (Dec., 2016) L30, [[1610.08972](#)].
- [101] H. I. Ringermacher and L. R. Mead, *In Defense of an Accelerating Universe: Model Insensitivity of the Hubble Diagram*, *ArXiv e-prints* (Nov., 2016) , [[1611.00999](#)].
- [102] B. S. Haridasu, V. V. Luković, R. D’Agostino and N. Vittorio, *Strong evidence for an accelerating universe*, *Astron. Astrophys.* **600** (Apr., 2017) L1, [[1702.08244](#)].
- [103] I. Tutusaus, B. Lamine, A. Dupays and A. Blanchard, *Is cosmic acceleration proven by local cosmological probes?*, *Astron. Astrophys.* **602** (June, 2017) A73, [[1706.05036](#)].
- [104] A. I. Lonappan, S. Kumar, Ruchika, B. R. Dinda and A. A. Sen, *Bayesian evidences for dark energy models in light of current observational data*, *ArXiv e-prints* (July, 2017) , [[1707.00603](#)].
- [105] L. H. Dam, A. Heinesen and D. L. Wiltshire, *Apparent cosmic acceleration from type ia supernovae*, *Mon. Not. Roy. Astron. Soc.* **472** (Nov., 2017) 835–851, [[1706.07236](#)].
- [106] H.-N. Lin, X. Li and Y. Sang, *The local probes strongly favor an accelerating universe*, *ArXiv e-prints* (Nov., 2017) , [[1711.05025](#)].
- [107] A. Dolgov, V. Halenka and I. Tkachev, *Power-law cosmology, SN Ia, and BAO*, *J. Cosmology Astropart. Phys.* **10** (Oct., 2014) 047, [[1406.2445](#)].
- [108] D. L. Shafer, *Robust model comparison disfavors power law cosmology*, *Phys. Rev. D* **91** (May, 2015) 103516, [[1502.05416](#)].
- [109] I. Tutusaus, B. Lamine, A. Blanchard, A. Dupays, Y. Zolnierowski, J. Cohen-Tanugi et al., *Power law cosmology model comparison with cmb scale information*, *Phys. Rev. D* **94** (Nov., 2016) 103511, [[1610.03371](#)].
- [110] A. Mitra, *Why the Big Bang Model does not allow inflationary and cyclic cosmologies though mathematically one can obtain any model with favourable assumptions*, *New A* **30** (July, 2014) 46–50.
- [111] M. V. John and K. B. Joseph, *Generalized Chen-Wu type cosmological model*, *Phys. Rev. D* **61** (Apr., 2000) 087304, [[gr-qc/9912069](#)].
- [112] A. Dev, M. Safonova, D. Jain and D. Lohiya, *Cosmological tests for a linear coasting cosmology*, *Physics Letters B* **548** (Nov., 2002) 12–18, [[astro-ph/0204150](#)].
- [113] F. Melia, *On recent claims concerning the $R_h = ct$ Universe*, *Mon. Not. Roy. Astron. Soc.* **446** (Jan., 2015) 1191–1194, [[1406.4918](#)].
- [114] F. Melia, *The linear growth of structure in the $R_h = ct$ universe*, *Mon. Not. Roy. Astron. Soc.* **464** (Jan., 2017) 1966–1976, [[1609.08576](#)].
- [115] I. Tutusaus, B. Lamine and A. Blanchard, *Model-independent cosmic acceleration and type ia supernovae intrinsic luminosity redshift dependence*, *ArXiv e-prints* (Mar., 2018) , [[1803.06197](#)].
- [116] M. Moresco, L. Verde, L. Pozzetti, R. Jimenez and A. Cimatti, *New constraints on cosmological parameters and neutrino properties using the expansion rate of the Universe to $z \sim 1.75$* , *J. Cosmology Astropart. Phys.* **7** (July, 2012) 053, [[1201.6658](#)].
- [117] M. Chevallier and D. Polarski, *Accelerating universes with scaling dark matter*, *International Journal of Modern Physics D* **10** (2001) 213–223, [[gr-qc/0009008](#)].

- [118] E. V. Linder, *Exploring the expansion history of the universe*, *Physical Review Letters* **90** (Mar., 2003) 091301, [[astro-ph/0208512](#)].