Using Artificial Neural Networks to reconstruct cosmology

Kostas Dialektopoulos

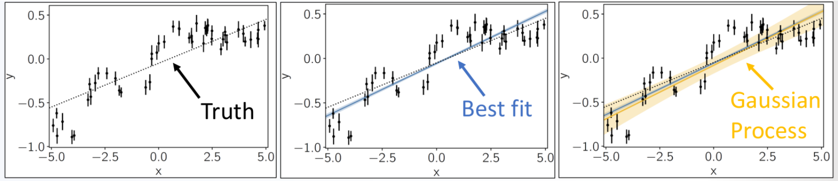
Transilvania University of Brasov

Tensions in cosmology workshop
Kyoto, Japan 2024

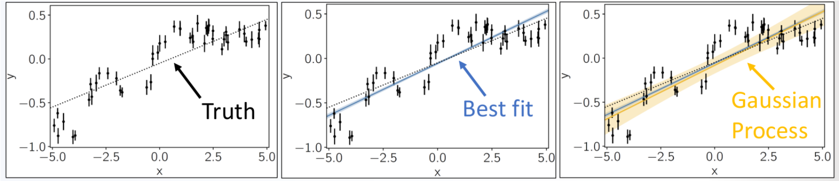
Take home message

- ANNs are a model independent tool that can help us reconstruct cosmological (and not only) parameters.
- We can use them to distinguish between the plethora of theories in the literature, based solely on the data without any physical or statistical assumption.

What are Gaussian processes?



What are Gaussian processes?



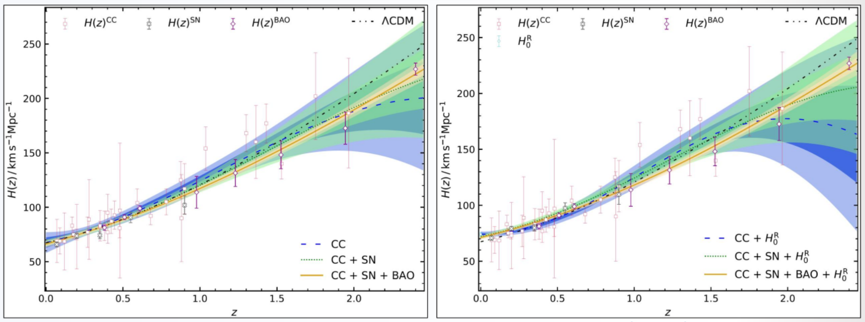
Definition: A GP is a stochastic (random) process where any finite subset is a **multivariant Gaussian distribution** with mean $\mu(x)$ and covariance $k(x, x')$.

Setting each $\mu(x)$ to zero, the **covariance function** can be used to learn the behavior that produced the data points.

Gaussian Process Regression

- The covariance function contains **non-physical hyperparameters** θ which define the distribution $k(\theta, x, x')$.
- Iterating over these values using Bayesian inference (or others) can produce better hyperparameters.
- The result is a **model independent reconstruction** (in physics) of the behabior of some parameter.
- This is superior to regular fitting because it is nonparametric and so **assumes no physical model** whatsoever.

Squared Exponential H_0 GP (GaPP code: Seikel et al. 2012)

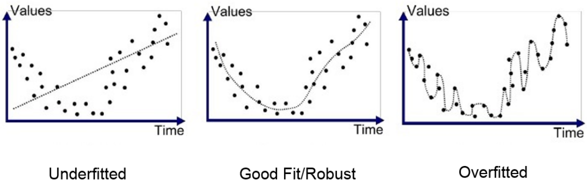


$H_0 = 67.539 \pm 4.772$km/s/Mpc
$H_0 = 67.001 \pm 1.653$km/s/Mpc
$H_0 = 66.197 \pm 1.464$km/s/Mpc

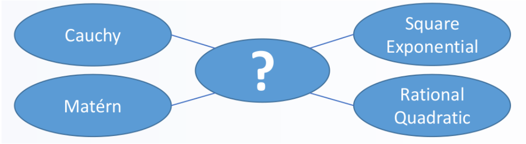
$H_0 = 73.782 \pm 1.374$km/s/Mpc
$H_0 = 72.022 \pm 1.076$km/s/Mpc
$H_0 = 71.180 \pm 1.025$km/s/Mpc

Open problems with GP reconstructions

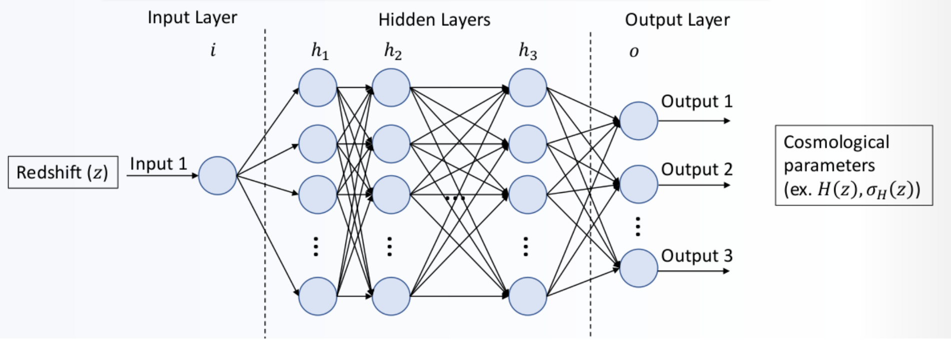
- **Overfitting**: GP is very prone to overfitting for small data sets, which is especially pronounced at the origin, i.e. Hubble constant



- **Kernel Selection Problem**: There is no natural kernel for cosmology

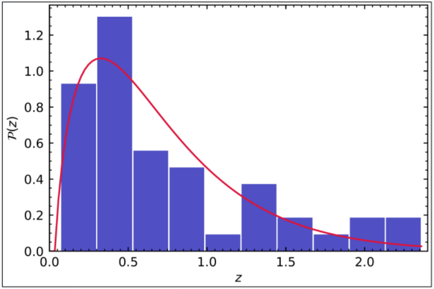


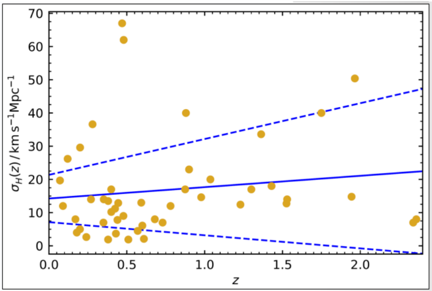
Artificial Neural Networks (ANN)



ReFANN code from Wang et al. (2020)

Training data for the ANN





$$P(z,\alpha,\lambda) = \frac{\lambda^\alpha}{\Gamma(\alpha)} z^{\alpha-1} e^{-\lambda z}$$

Mean: $\sigma_H = 14.25 + 3.42z$
Upper error: $\sigma_H = 21.37 + 10.79z$
Lower error: $\sigma_H = 7.14 - 3.95z$

Designing the ANN

- **Risk**: Optimizes the **number of hidden layers and neurons** in an ANN

$$\mathrm{risk} = \sum_{i=1}^{N} (\mathrm{Bias}_i^2 + \mathrm{Variance}_i) = \sum_{i=1}^{N} \left([H_{\mathrm{obs}}(z_i) - H_{\mathrm{pred}}(z_i)]^2 + \sigma_H^2(z_i) \right)$$

- **Loss**: Balances the **number of iterations** a system needs to predict the observational data

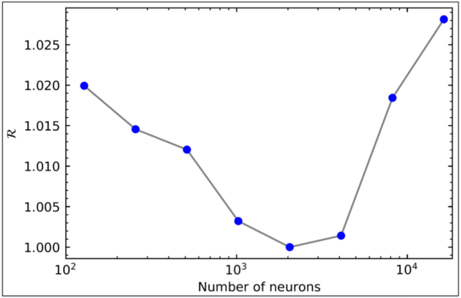
1. Least absolute deviation (**L1**)

$$L1 = \sum_{i=1}^{N} |H_{\mathrm{obs}}(z_i) - H_{\mathrm{pred}}(z_i)|$$

2. Smoothed L1 (**SL1**)
3. Mean Square Error (**MSE**)

$$\mathrm{MSE} = \frac{1}{N} \sum_{i=1}^{N} (H_{\mathrm{obs}}(z_i) - H_{\mathrm{pred}}(z_i))^2$$

Building the ANN



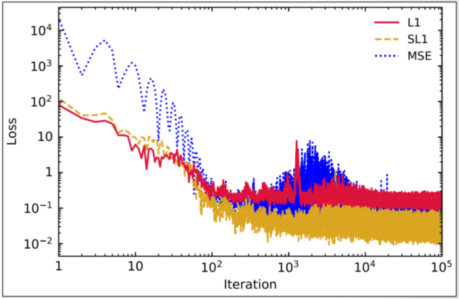


Figure: **Left:** Risk function for **one layer** (number of neurons 2^n, $n \in 7, ..., 14$), **Right:** Evolution of L1, SL1 and MSE loss functions

Using the ANN (KD, Levi Said et al. '21) (KD, Mukherjee et al. '23)

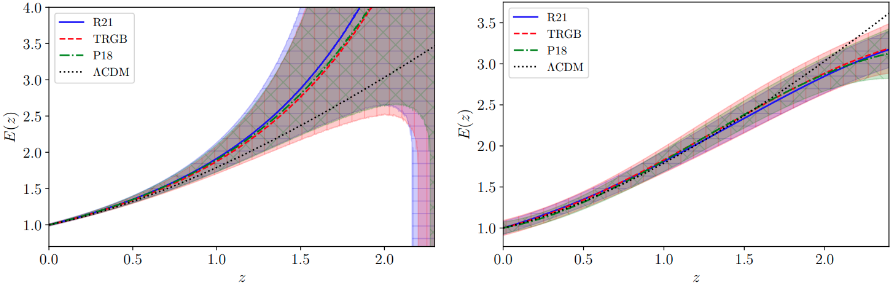


Figure: Reconstructed reduced Hubble parameter from the (i) Pantheon SN compilation (left) and (ii) combined CC+BAO Hubble data set (right), using ANNs.

Om diagnostics (Sahni, Shafieloo, Starobinsky '08) (Shafieloo, Clarkson '10)

Distinguish ΛCDM from alternative dark energy and modified gravity models:

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

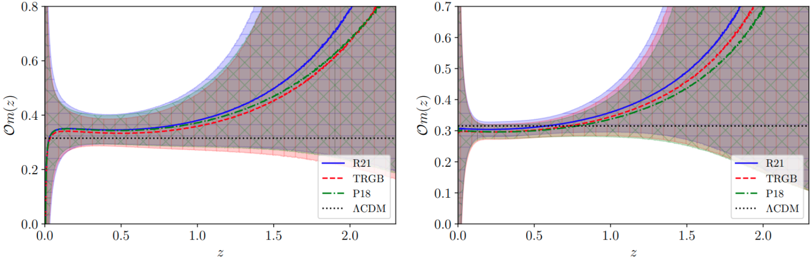


Figure: Reconstructed *Om* diagnostics using (i) ANNs (left) and (ii) GPs (right) from the Pantheon SN data for three different priors.

H0 diagnostics (Krishnan, Colgáin, Sheikh-Jabbari, Yang '20)

It is defined as

$$\mathbf{H0} = \frac{H(z)}{\sqrt{\Omega_{m0}(1+z)^3 + 1 - \Omega_{m0}}},$$

and its non-constancy suggests evidence for new physics beyond ΛCDM.

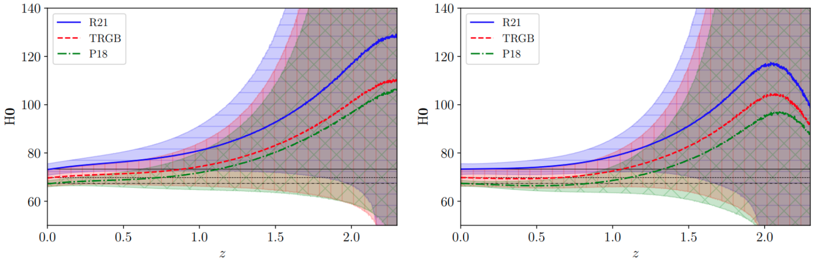


Figure: Reconstructed H_0 diagnostics using (i) ANNs (left) and (ii) GPs (right) from the Pantheon SN data for three different priors.

Constraining theories Arjona, Cardona, Nesseris '19

Example: Horndeski mapping:

$$G_2 = K(X),\ G_3 = G(X),\ G_4 = 1/2, \text{ and } G_5 = 0\,,$$

The action is given by:

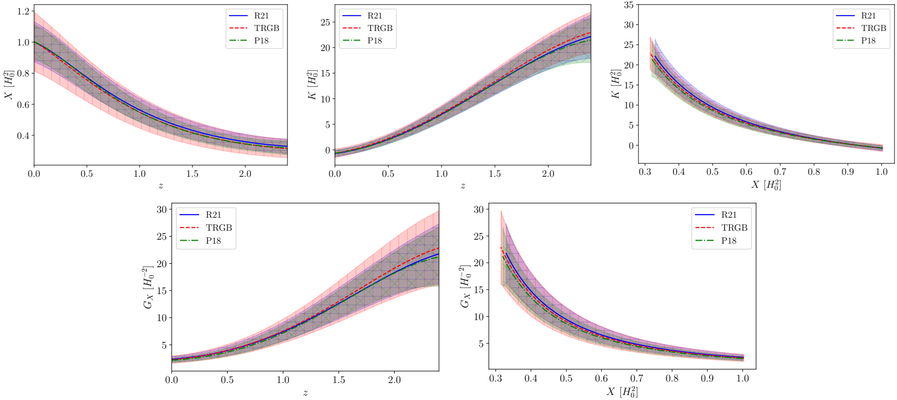
$$S = \int \mathrm{d}^4 x \sqrt{-g}\left(\frac{R}{2} - K(X) - G(X)\Box\phi\right) + S_{\mathrm{mat}}(\psi, g_{\mu\nu})\,.$$

Cosmological equations (flat FLRW):

$$K(X) = -3H_0^2\left(1-\Omega_{m0}\right) + \frac{\mathcal{J}\sqrt{2X}H^2(X)}{H_0^2\Omega_{m0}} - \frac{\mathcal{J}\sqrt{2X}\left(1-\Omega_{m0}\right)}{\Omega_{m0}}\,,$$

and

$$G_X(X) = -\frac{2\mathcal{J}H'(X)}{3H_0^2\Omega_{m0}}\,.$$



(KFD, Mukherjee, Levi Said, Mifsud '23)

We can also compute the DE EoS as

$$w_\phi = \frac{-K + \sqrt{2X}\dot{X}G_X}{K - 2X(K_X + 3\sqrt{2X}HG_X)}$$

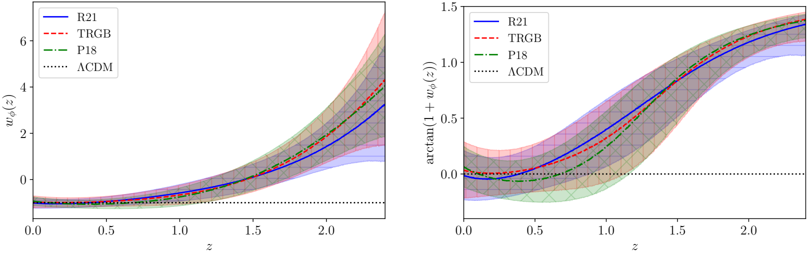


Figure: Plots for dark energy EoS $w_\phi(z)$ (left) and its compactified form $\arctan(1 + w_\phi(z))$ (right) considering R21, TRGB, and P18 H_0 priors. The shaded regions with '$-$', '$|$' and '$\times$' hatches represent the 1σ confidence levels for the above priors respectively.

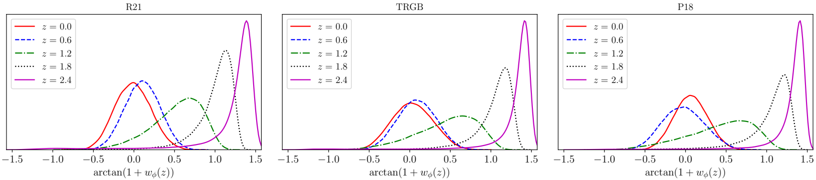


Figure: Plots showing the posteriors of probability distribution of the compactified dark energy EoS for the theory at some sample redshifts for the R21, TRGB, and P18 H_0 priors, respectively.

Conclusion and Prospects

- GP and ANN both have positive features in reconstructing cosmological data sets.
- However, ANN shows greater promise in that they rely on less rigid training data and can model more complex structures of data sets.

From now on, it would be interesting to

- forecast observations for experiments in progress that are about to publish their results,
- use the reconstructed Hubble parameter and its derivative to constrain or even eliminate more alternative cosmological models,
- consider observations related to the perturbative part of the theory, such as LSS or GWs, in the context of ANNs.

Conclusion and Prospects

- GP and ANN both have positive features in reconstructing cosmological data sets.
- However, ANN shows greater promise in that they rely on less rigid training data and can model more complex structures of data sets.

From now on, it would be interesting to

Thank you!

- forecast observations for experiments in progress that are about to publish their results,
- use the reconstructed Hubble parameter and its derivative to constrain or even eliminate more alternative cosmological models,
- consider observations related to the perturbative part of the theory, such as LSS or GWs, in the context of ANNs.