Numerical Methods in Cosmology

Emmanuel Flores

May 16, 2025

Abstract

Modern cosmology heavily relies on numerical techniques to model the Universe's evolution, from its smooth background expansion to the growth of structure This brief report reviews some of the numerical techniques availables in this vast landscape, covering traditional deterministic and stochastic methods before focusing on the emerging role of Machine Learning. We focus on the the application of Artificial Neural Networks (ANNs), trained via unsupervised learning, to solve the background dynamics for various cosmological models such as ΛCDM, CPL, Quintessence, $f(R)$. All of this to say that numerical techniques, encompassing both established and novel ML approaches, remain crucial for cosmological progress.

1 Motivation and Background

It's well know that one of the main goals of cosmology is to understand the origin, evolution, and large-scale structure of the Universe. While analytical models provide foundational insights, the inherent complexity of cosmological systems-often characterized by significant nonlinearities and a vast number of degrees of freedom (DOF)-frequently renders purely analytical solutions intractable. Which means that phenomena such as structure formation and the dynamics of spacetime itself push the boundaries of what can be solved with pen and paper alone. This is where numerical techniques come to the rescue. As I just mentioned, situations often arise where exact analytical solutions are simply not feasible. In such cases, computational methods offer a powerful alternative, providing approximate solutions that allow us to model and understand these intricate systems. An approximate solution derived numerically is vastly preferable to no solution at all, enabling progress in areas that would otherwise remain inaccessible. Furthermore, the ability of numerical methods to handle highly complex systems makes them essential tools for simulating realistic cosmological scenarios. While employing these techniques, it is crucial to proceed with mathematical rigor. Establishing criteria under which a problem is well-posed and proving that the chosen numerical discretization leads to bounded, convergent solutions ensures the reliability and physical relevance of the results. This brief report will delve into various numerical techniques commonly employed in cosmology, exploring their application, strengths, and the underlying principles that validate their use in unraveling the mysteries of the cosmos.

1.1 GR Essentials

As expected, the standard model of cosmology is built upon the foundation of Einstein's General Relativity (GR). Within this framework, the large-scale Universe is assumed to be homogeneous and isotropic. This assumption leads to the Friedmann-Lematre-Robertson-Walker (FLRW) metric, which describes the geometry of such a spacetime:

$$ds^2 = -c^2dt^2 + a(t)^2 \left[\frac{dr^2}{1 - Kr^2} + r^2(d\theta^2 + \sin^2\theta^d \phi^2) \right],$$

where c is the speed of light and, t is time, $a(t)$ is the scale factor that describes the relative expansion of the universe, and K represents the spatial curvature, which can be positive, negative or zero. Moreover, the dynamics of the scale factor $a(t)$ are governed by the Einstein field equations applied to the FLRW metric. This results in the following Friedmann equations:

$$(\frac{\dot{a}}{a})^2 = H^2 = \frac{8\pi G}{3c^2}\rho - \frac{Kc^2}{a^2},$$

which relates the expansion rate, known as the Hubble parameter to the total energy density and the spatial curvature K, and also

$$(\frac{\ddot{a}}{a})^2 = -\frac{4\pi G}{3c^2}(\rho + 3p).$$

The terms ρ and p represent the total energy density and pressure of all components filling the Universe (radiation, matter, dark energy). These quantities arise from the energy-momentum tensor, $T\mu\nu$, which as we know, acts as the source term in the Einstein field equations. For a perfect fluid in the comoving frame appropriate for the FLRW metric, the energy-momentum tensor takes a simple diagonal form:

$$T^\mu_\nu = \text{diag}(-\rho, p, p, p),$$

Another quite important aspect of GR is the conservation of energy and momentum, which can be expressed mathematically as the vanishing divergence of the energy-momentum tensor $\nabla_\mu T_{\mu\nu} = 0$. And if we apply this

conservation law to the FLRW metric for each non-interacting fluid component (indexed by i) we have the following continuity equation:

$$\dot{\rho} + 3H(\rho + p) = 0.$$

This equation describes how the energy density of each component evolves as the Universe expands. It can be solved if the relationship between pressure and density, known as the equation of state $p_i = w_i \rho_i$, is known. For standard components:

- Non relativistic matter: $\omega_m = 0 \implies \rho_m \propto a^-3$,

- Radiation: $\omega_r = 1/3 \implies \rho_r \propto a^-4$,

- Dark Energy: $\omega_{DE} = -1 \implies \rho_{DE} \propto \text{constant}$

2 A quick Overview of Current Numerical Techniques in Cosmology

Following [1], the use of numerical methods in cosmology can be broadly categorized into two main cathegories based on the nature of the phenomena they describe: deterministic and stochastic. This distinction often aligns with whether we are modeling the smooth background evolution of the Universe or the growth of structures within it. And in the following subsections I briefly elaborate on each one of them.

2.1 Deterministic Approach and Background Observables

This approach deals with the large-scale, homogeneous, and isotropic evolution of the Universe as described by the FLRW metric defined previously. Here, the primary goal is to calculate background observables that depend solely on the smooth expansion history. Some examples include

- Hubble parameter $H(z)$

- Cosmological distances

- Age of the universe

- Some prefictions for observables, such as the distance modulus of Type Ia supernovae.

No, in relationship to the numerical techniques used in this realm, these typically involve solving a relatively small system of Ordinary Differential Equations (ODEs), namely the Friedmann and continuity equations, for the scale factor $a(t)$ and the densities of different cosmic components (ρ_i). Therefore, we can say that given a specific cosmological model and its parameters, the predictions for these background quantities are exact or "deterministic". And now, let's move to the stochastic approach.

2.2 Stochastic Approach and Non Linear Perturbation Theory

While the background model provides the stage, the intricate structure we observe in the Universe arises from small initial deviations from perfect homogeneity. Modeling the evolution of these inhomogeneities constitutes the stochastic approach. Some of the fundamental aspects include:

- Linear perturbation theory, whose purpose is to describe the initial growth of small fluctuations in the early universe.

- Generalizations of linear theory that seek to account non-linear effects that are crucial for the understanding of late-time structure formation of the universe.

- And finally, observables, that are predictions for statistical properties os some structures, such as CMB temperature and polarization, power spectra, matter power spectrum, and galaxy correlation functions, just to name a few.

This approach is considered "stochastic" for two main reasons. First, the initial conditions, are often assumed to originate from quantum fluctuations during an early inflationary phase, are therefore inherently probabilistic. Second, predicting the exact position and properties of every single galaxy or structure is unfeasible and not the primary goal. Instead, numerical techniques aim to predict the statistical distributions and correlations of these structures. The numerical methods involved are significantly more complex, often requiring the solution of large systems of coupled differential equations (Boltzmann codes like CAMB or CLASS) or computationally intensive simulations such as described in [1]. Therefore, the deterministic approach provides the smooth, average evolution, while the stochastic approach models the vast and rich structures that emerge from initial fluctuations on top of this background. Both are essential, and often coupled, in modern numerical cosmology.

3 Background Dynamics and Machine Learning

In recent years, Machine Learning (ML) techniques, particularly artificial neural networks (ANNs), have emerged as powerful tools across various scientific disciplines, including cosmology ([2]).The motivation for applying ML stems from its ability to identify complex patterns and relationships within large, high-dimensional datasets. This is highly relevant in modern cosmology, which deals with vast amounts of observational data and intricate theoretical models describing both the smooth background evolution and the complex formation of structures. As we've just discussed, numerical cosmology often distinguishes between two types of problem: deterministic and

stochastic. Each one presenting it's own challenges as described below.

- Deterministic: Modeling background observables often involves solving systems of Ordinary Differential Equations (ODEs). While deterministic, exploring the parameter space or dealing with complex dark energy models can still pose computational challenges.

- Stochastic: Modeling the inhomogeneous Universe involves describing the evolution from initial, statistically defined fluctuations to the observed structures. Which, as described previously, requires the use of perturbation theory and often non-linear simulations, dealing with inherently statistical predictions and very complex dependencies.

It's clear that the deterministic and stochastic approach present a high degree of complexity, reason for that the use of machine learning techniques can be used to gain some insigth. Probably one of the main results in the theory of ML is the Universal Approximation Theorem, which is described below.

3.1 Universal Approximation Theorem (UAT)

A fundamental concept underpinning the potential of Artificial Neural Networks (ANNs) in both domains is the Universal Approximation Theorem (UAT). In essence, the UAT states that a sufficiently large feedforward neural network with a single hidden layer, containing a finite number of neurons and employing suitable non-linear activation functions, can approximate any continuous function on compact subsets of R^n to an arbitrary degree of accuracy. While different versions of the theorem exist, the core implication remains: neural networks possess the theoretical capacity to represent an extremely broad class of functions, such as the ones used in cosmology. Therefore, this theorem provides a powerful theoretical justification for using neural networks to tackle cosmological challenges. Many problems involve complex mappings or computationally intensive tasks, for instance:

- Mapping parameter-observable relationships

- Analyzing high-dimensional parameter spaces

- Emulating expensive simulations

The UAT suggests that neural networks, given sufficient data and appropriate training, can potentially learn these complex mappings or emulate these processes effectively. This capability opens up new avenues for tackling computationally intensive tasks and extracting insights from complex cosmological data across both the deterministic and stochastic regimes, complementing traditional numerical and analytical methods.

3.2 Models under study

The following discussion relies heavily on the work by Chantada et al. (2023) which investigates the application of cosmology-informed neural networks to solve the background dynamics for several cosmological models, focusing on alternatives to the standard model aimed at explaining the late-time cosmic acceleration. In particular, they work with the following four models:

- Λ CDM: baseline model. It assumes a spatially flat FLRW metric. The Hubble parameter is determined by the standard Friedmann equations, with free parameters typically being the present-day matter density parameter and the Hubble constant.

- Parametric Dark Energy: this model also considers standard matter but replaces the cosmological constant with a dark energy fluid whose equation of follow the previously defined equation of state.

- Quintessence (Exponential Potential): this model introduces a dynamic dark energy component in the form of a minimally coupled scalar field ϕ with a self-interaction potential $V(\phi)$. The specific form of the potential is given by $V(\phi) = V_0 \exp(-\lambda\sqrt{\kappa}\phi)$, where λ is a free parameter.

- f(R) Gravity (Hu-Sawicki Model): this model is probably the most complex of all four, since it represents a modification of GR itself. Instead of the standard Einstein-Hilbert action the gravitational action includes a more general function $f(R)$. The specific form studied is the Hu-Sawicki model in which we have $f(R) = R - 2\Lambda[1 - (R/\Lambda + 1)^n]$.

These four models represent different theoretical approaches, and the main idea is to solve the background evolution using neural networks in the referenced study. In the next subsection I'll describe with more details the simulation and methodology used in the paper.

3.3 Methodology Details

The core task for the ANNs is to find solutions to the systems of differential equations governing the background expansion for each cosmological model. Instead of directly solving for H(z), the paper often employs dimensionless dynamical variables, resulting in the following systems targeted by the ANNs:

- Λ CDM: The evolution is described by the matter density. Using the dimensionless variable $x =$ the single ODE solved by the ANN, as a function of redshift z, is

$$\frac{dx}{dz} = \frac{3x}{1+z},$$

with boundary condition

$$x(z)|_{z=0} = \frac{\kappa\rho_{m,0}}{3H_0^2} = \Omega_{m,0}$$

- Parametric Dark Energy: Here, the focus is on the dark energy density. Using the dimensionless variable $x = \kappa\rho_{DE}/3H_0^2$ and parameters ω_0, ω_1 the ODE solved is:

$$\frac{dx}{dz} = \frac{3x}{1+z}\left(1 + \omega_0 + \frac{\omega_1 z}{1+z}\right)$$

with boundary condition

$$x(z)|_{z=0} = \frac{\kappa\rho_{DE,0}}{3H_0^2} = 1 - \Omega_{m,0}$$

- Quintessence: This model requires solving a system of two coupled ODEs for the dimensionless variables $x = \sqrt{\kappa}\dot{\phi}/\sqrt{6}H$ and $y = \sqrt{\kappa V(\phi)}/\sqrt{3}H$, using $N = -\ln(1+z)$ as the independent variable

$$\frac{dx}{dN} = -3x + \frac{\sqrt{6}}{2}\lambda y^2 + \frac{3}{2}x\left(1 + x^2 - y^2\right),$$

$$\frac{dy}{dN} = -\frac{\sqrt{6}}{2}xy\lambda + \frac{3}{2}y\left(1 + x^2 - y^2\right),$$

and the ANNs solve this system, taking λ and initial conditions (derived from matching ΛCDM at high redshift) as inputs.

- $f(R)$ Gravity: This modified gravity model leads to a more complex system of five coupled ODEs for the dynamical variables $x = \dot{R}f_{RR}/Hf_R$, $y = f/gH^2f_R$, $v = R/6H^2$, $\Omega = \kappa\rho_m/3H^2f_R$, and $r = R/\Lambda$

$$\frac{dx}{dz} = \frac{1}{1+z}\left(-\Omega + 2v + x + 4y + xv + x^2\right),$$

$$\frac{dy}{dz} = -\frac{1}{1+z}\left(vx\Gamma - xy + 4y - 2vy\right),$$

$$\frac{dv}{dz} = -\frac{v}{1+z}\left(x\Gamma + 4 - 2v\right),$$

$$\frac{d\Omega}{dz} = \frac{\Omega}{1+z}\left(-1 + 2v + x\right),$$

$$\frac{dr}{dz} = -\frac{r\Gamma x}{1+z},$$

where $\Gamma = Rf_R/f_{RR}$ is a function of r. The ANNs are trained to solve this coupled system, taking the model parameter b and initial conditions as inputs.

These specific systems of ODEs, expressed in terms of carefully chosen variables, form the mathematical basis that the cosmology-informed neural networks will learn to solve, providing an alternative to traditional numerical integration methods. This method is distinct from many ML applications as it operates in an unsupervised learning framework. This means the training process does not rely on pre-computed analytical or numerical solutions of the ODEs as target outputs.

3.3.1 Core Methodology

The core methodology of the paper is described as follows:

- Optimization Problem: The task is posed as an optimization problem. The internal parameters of the ANN (weights and biases) are adjusted during training to minimize a carefully constructed loss function.

- Loss Function: The loss function quantifies how well the ANN's output satisfies the original differential equations. It is typically based on the sum of squared residuals - the values obtained when plugging the ANN's output (and its derivatives) back into the ODEs. The goal is to drive these residuals towards zero.

- Initial/Boundary Conditions: Initial conditions are enforced by using specific reparametrizations of the ANN output.

- One of the very interesting features of this work is the idea of "bundle solution", which essentially means that the method is extended to create solutions across a continuous range of the model's physical parameters such as $\Omega_{M,0}$, ω_0, ω_1, λ and b. And these parameters become additional inputs to the network together with the independent variable (like z or N).

- And finally, in some models, some specialized Reparametrizations is requiered: in the case of the Dark Matter, Quintessence and $f(R)$ models, they need some specialized reparametrizations, such as using an integrating factor form or a "perturbative" approach where the ANN corrects a known solution, this is the ΛCDM solution at a boundary of the parameter space.

3.3.2 Trainig Details

And now, I'll provide a quick summary of the training details of the network.

- Network Architecture: One ANN is typically assigned to each dependent variable in the system of ODEs. The specific architecture used in the paper consists of two hidden layers with 32 units (neurons) each, employing the tanh as the activation function.

- Optimizer: The ADAM optimizer, a variant of stochastic gradient descent, is used to adjust the network's weights and biases to minimize the loss.

- Training Process: Training proceeds iteratively. In each iteration, the loss is calculated over a "batch" of points randomly sampled from the domain of the independent variable and the parameter space ranges. The optimizer updates the network parameters based on the gradient of this batch loss. This

continues until the loss function converges, indicating that the ANN output closely satisfies the differential equations over the trained domain.

3.3.3 Validation Methodology

The validation of the trained ANNs is primarily achieved by demonstrating their ability to produce scientifically accurate results when used in a standard cosmological analysis pipeline such as:

- Generating Predictions: The trained ANNs are used to predict the necessary cosmological quantities, primarily the Hubble parameter $H(z)$ or related distance measures, as a function of redshift and the model parameters.

- Observational Data: These predictions are compared against real observational data sets, in particular they use Markov Chain Monte Carlo (MCMC) as follows:

- Statistical Analysis (MCMC): Standard likelihood functions are constructed for each data set based on the difference between the ANN predictions and the observed data. A MCMC analysis is then performed to explore the parameter space of each cosmological model and obtain posterior probability distributions and constraints on the model parameters.

- Comparison: And finally, the key validation step involves comparing the parameter constraints obtained using the ANN-based solutions with results published in the literature that used traditional numerical solvers, such as Runge-Kutta methods, for the same cosmological models and similar observational data sets. Agreement between these results demonstrates the accuracy and viability of the ANN method for cosmological inference.

Finally, although not part of the primary validation pipeline, the paper also quantifies the intrinsic accuracy of the ANN solutions by comparing their outputs directly to analytical solutions (for ΛCDM, CPL) or high precision numerical solutions after the training is complete.

3.4 Comparisson with other Approaches

Thw work done by demonstrates the potential of using unsupervised, physics-informed neural networks as a viable alternative for solving the background dynamics in cosmology. Some of the advantages and disadvantages listed in the paper are described below.
Advantages:

- Viability: The method successfully solves the differential equations for various cosmological models.

- Bundle solutions: as described above, a single trained network can efficiently provide solutions across a continuous range of model parameters, ideal for parameter inference.

- Post-Training Efficiency: Evaluating the trained ANN is very fast, which can significantly speed up tasks like MCMC sampling compared to repeatedly running numerical solvers, especially for complex or stiff systems

- Handling Numerical Challenges: Specialized techniques (like perturbative reparametrization) allowed the ANNs to effectively handle issues like singularities that can be quite difficult for traditional solvers.

- Portability: Trained ANNs are lightweight and easy to store and share compared to large grids of numerical solutions.

Disadvantages:

- Training Cost: The initial training phase can be computationally intensive and time-consuming, requiring significant resources.

- Error Quantification: The method, as implemented, lacks a built-in mechanism to rigorously quantify the uncertainty or error of the ANN solution itself; validation relies on comparisons or downstream analysis.

- Hyperparameter Tuning: Finding the optimal network architecture and training parameters often requires empirical tuning and experimentation, something that is quite common in ML. And finally,

- Context-Dependent Speed-up: The overall computational advantage is not guaranteed. For simple problems or scenarios requiring few evaluations, the upfront training cost might make traditional numerical solvers faster.

4 Conclusions

It's clear that numerical techniques are fundamentally indispensable in modern cosmology. The complexity inherent in General Relativity and the vast dynamic range involved in the Universe's evolution quite frequently eludes pure analytical solutions. Traditional numerical methods, such as ODE solvers, Boltzmann codes (like CAMB/CLASS), and N-body simulations, have been crucial for bridging the gap between theoretical models and observational data. However, the recent advent of machine learning, particularly particularly in the field of physics-informed neural networks, represents an exciting new direction. As demonstrated, these methods can effectively solve cosmological ODE systems, offer potential computational advantages in specific scenarios, and yield

results consistent with established techniques. While challenges related to training cost and rigorous error quantification remain, the ongoing development of both traditional and ML-based numerical approaches continues to enhance our ability to model the Universe and interpret the rich of cosmological data.

References

[1] S. Dias Pinto Vitenti and M. Penna-Lima, "Selected topics in numerical methods for cosmology," *Universe*, vol. 5, no. 9, p. 192, 2019.

[2] A. T. Chantada, S. J. Landau, P. Protopapas, C. G. Scóccola, and C. Garraffo, "Cosmology-informed neural networks to solve the background dynamics of the universe," *Physical Review D*, vol. 107, no. 6, p. 063523, 2023.